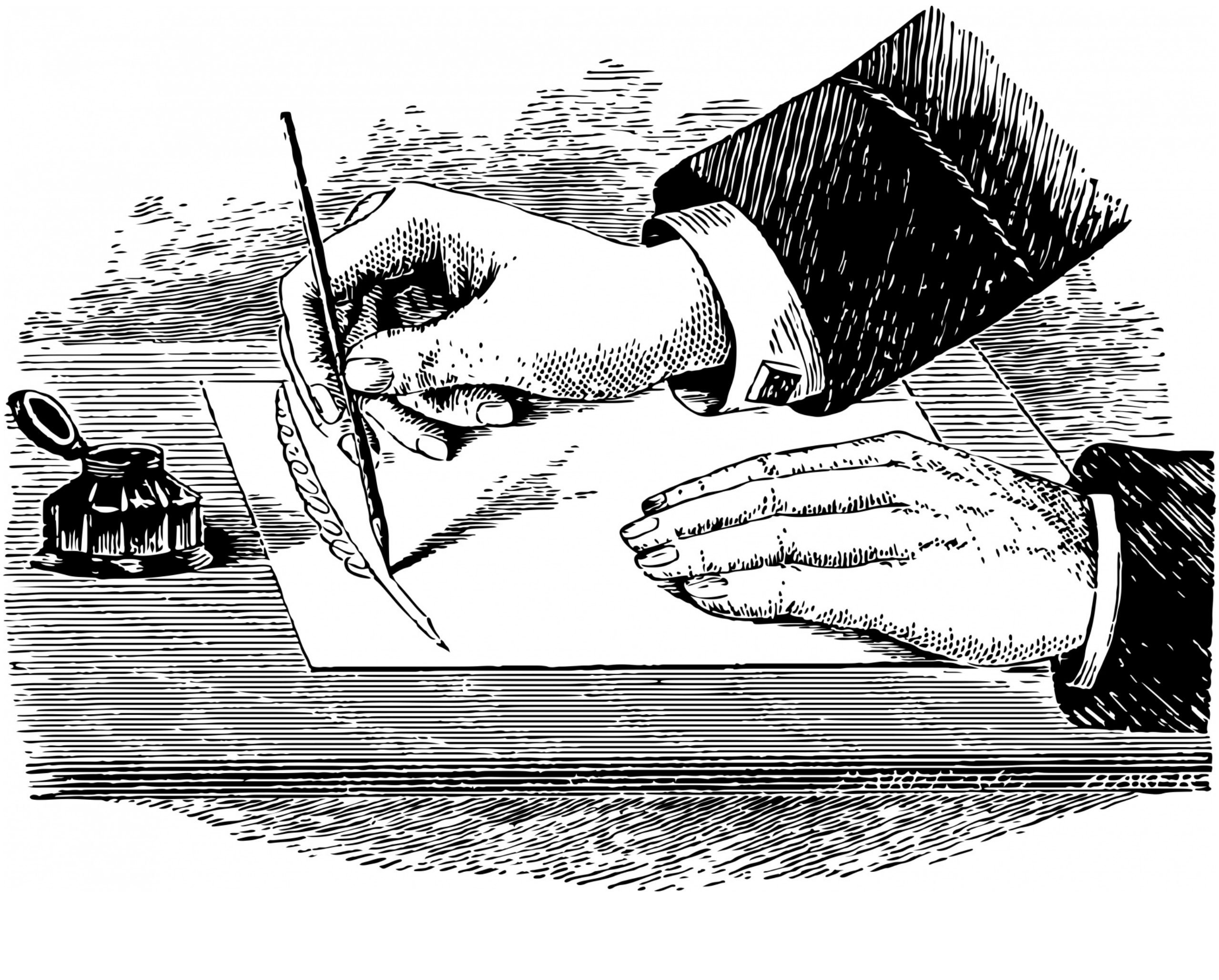
ΟΔΗΓΙΕΣ ΓΙΑ ΣΥΓΓΡΑΦΕΙΣ XƎLATEX

Κεντρική Ομάδα Υποστήριξης

Κάλλιπος Ανοικτές Ακαδημαϊκές Εκδόσεις



ΚΑΛΛΙΠΟΣ
ανοικτές ακαδημαϊκές εκδόσεις

Φωτογραφία εξωφύλλου: "Writing Hand" από τον Dawn Hudson. Άδεια: CC0 Public Domain.

ΜΗ ΠΑΡΑΜΕΤΡΙΚΗ ΣΤΑΤΙΣΤΙΚΗ

Μη Παραμετρική Στατιστική

Υπότιτλος

Σάμης Τρέβεζας
Επίκουρος Καθηγητής Τμήματος Μαθηματικών
Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

ΚΑΛΛΙΠΟΣ
ΑΝΟΙΚΤΕΣ ΑΚΑΔΗΜΑΪΚΕΣ ΕΚ-ΔΟΣΕΙΣ
www.kallipos.gr

Αφιερώνεται σε...

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

ΠΙΝΑΚΑΣ ΣΥΝΤΟΜΕΥΣΕΩΝ - ΑΚΡΩΝΥΜΙΩΝ

APA	American Psychological Association
ΣΕΑΒ	Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών
ΚΟΥ	Κεντρική Ομάδα Υποστήριξης
ΕΥ	Επιστημονικά Υπεύθυνος

ΠΡΟΛΟΓΟΣ

Κεντρική Ομάδα Υποστήριξης
Κάλλιπος, Ανοικτές Ακαδημαϊκές Εκδόσεις
https://helpdesk.kallipos.gr

Πρόλογος θεωρείται το προεισαγωγικό τμήμα του βιβλίου, που τοποθετείται πριν από την εισαγωγή (εφόσον υπάρχει).

Ο πρόλογος είναι ένα πολύ σύντομο κείμενο, συχνά γραμμένο από πρόσωπο διαφορετικό από τον συγγραφέα (από τον εκδότη, από γνωστό ειδικό, από πρόσωπο συνδεόμενο με τον συγγραφέα κ.λπ.). Ο πρόλογος ενίοτε γράφεται και από τον ίδιο τον συγγραφέα, για να περιλάβει σ' αυτόν γενικότερες πληροφορίες, τεχνικά θέματα του βιβλίου, ευχαριστίες κ.λπ. Εφόσον υπάρχει εισαγωγή, ο πρόλογος προηγείται της εισαγωγής και διαφοροποιείται απ' αυτήν πλήρως. Εάν δεν υπάρχει εισαγωγή, ο πρόλογος περιέχει στοιχεία που κανονικά θα δίνονταν σε μια εισαγωγή και είναι εκτενέστερος.

ΕΙΣΑΓΩΓΗ

Κεντρικη Ομαδα Υποστηριξης
Καλλιπος, Ανοικτες Ακαδημαϊκες Εκδοσεις
https://helpdesk.kallipos.gr

Εισαγωγή είναι το προκαταρκτικό τμήμα βιβλίου, επιστημονικού έργου κ.λπ. Η εισαγωγή ενός βιβλίου καταλαμβάνει αρκετές σελίδες και σ' αυτήν ο συγγραφέας δίνει σημαντικές πληροφορίες για το περιεχόμενο του έργου του, τα κύρια προβλήματα που τον απασχόλησαν, τις γενικές θέσεις του κ.λπ.

Μέρος I

ΟΝΟΜΑ ΜΕΡΟΥΣ I

ΚΕΦΑΛΑΙΟ 1

ΕΙΣΑΓΩΓΙΚΑ - ΥΠΕΝΘΥΜΙΣΗ ΒΑΣΙΚΩΝ ΕΝΝΟΙΩΝ

1.1 Γιατί Μη Παραμετρική Στατιστική;

Το κοινό χαρακτηριστικό όλων των παραμετρικών οικογενειών κατανομών είναι ότι ο ακριβής προσδιορισμός της κατανομής μέσα στην αντίστοιχη οικογένεια εξαρτάται από ένα πεπερασμένο πλήθος παραμέτρων ή ισοδύναμα από μία διανυσματική παράμετρο πεπερασμένης διάστασης. Εξάλλου, σε μία πρώτη επαφή με τη Στατιστική, η ενασχόληση κάποιου περιορίζεται σε προβλήματα που εμπλέκουν την εκτίμηση μίας ή το πολύ δύο παραμέτρων, θεωρώντας ότι η αντίστοιχη παραμετρική οικογένεια περιγράφει καλά τη “γκάμα” συμπεριφορών που συνδέεται με το υπό μελέτη στοχαστικό φαινόμενο. Το πολύ ουσιώδες θέμα της καταλληλότητας ενός μοντέλου ελάχιστα θίγεται, αν και ίσως αποτελεί το σημαντικότερο στοιχείο μιας καλής μοντελοποίησης. Διάφορα ερωτήματα ανακύπτουν φυσιολογικά:

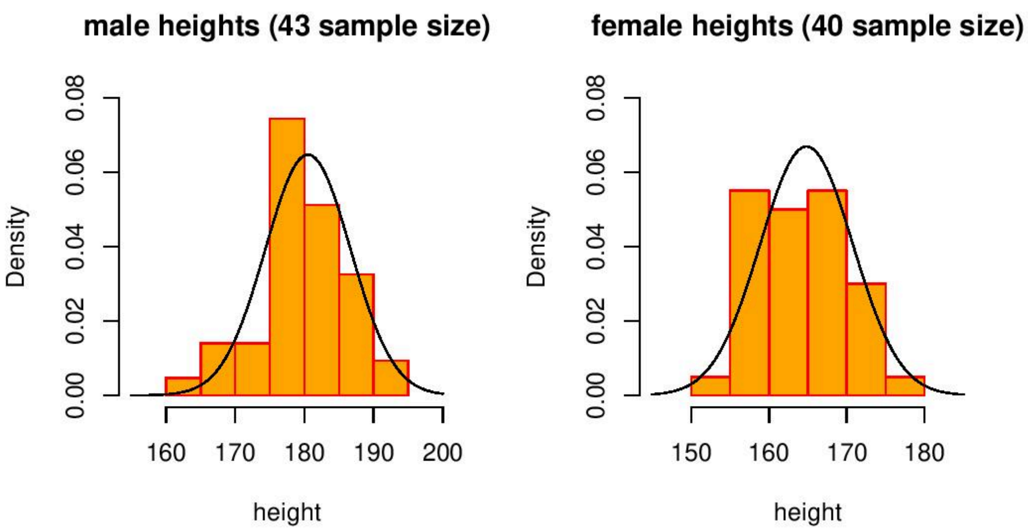
- Πώς δικαιολογείται η χρήση μίας συγκεκριμένης παραμετρικής οικογένειας κατανομών;
- Σε μία πρώτη προσέγγιση σε ένα στατιστικό πρόβλημα θα θέλαμε ίσως να έχουμε πολύ μεγαλύτερη ευελιξία στην κατανομή ενός χαρακτηριστικού. Πώς μπορούμε να κινηθούμε απαλλαγμένοι από στενές παραμετρικές υποθέσεις;
- Σε αρκετές εφαρμογές η κατανομή ενός χαρακτηριστικού ενδιαφέροντος φαίνεται να είναι περίπλοκη, πολυκόρυφη και χωρίς να μπαίνει μέσα σε κάποιο γνωστό παραμετρικό πλαίσιο. Πώς χειριζόμαστε αυτές τις περιπτώσεις;

Σάμης Τρέβεζας (2023). «Μη Παραμετρική Στατιστική».
Αθήνα: Κάλλιπος, Ανοικτές Ακαδημαϊκές Εκδόσεις.
BY-NC-SA Creative Commons Αναφορά Δημιουργού - Μη Εμπορική Χρήση - Παρόμοια Διανομή 4.0

- Η φύση των δεδομένων, π.χ. σειρά προτίμησης (διατάξιμα δεδομένα) ή κατηγορικά δεδομένα, μπορεί να μην επιτρέπει τη χρήση παραμετρικών μοντέλων. Πώς αντιμετωπίζουμε αυτές τις περιπτώσεις;

Ο στόχος λοιπόν αυτού του συγγράμματος είναι να απαντήσει σε αυτά τα ερωτήματα, θέτοντας τις βάσεις ενός πολύ σημαντικού κλάδου της Στατιστικής, που φέρει το όνομα Μη Παραμετρική (ή Απαραμετρική) Στατιστική και συμπληρώνοντας την εικόνα που σχηματίζεται σε ένα εισαγωγικό μάθημα Μαθηματικής Στατιστικής. Τα κεντρικά προβλήματα εκτίμησης και ελέγχων υποθέσεων που εμφανίζονται στην Παραμετρική Στατιστική, εμφανίζονται και στο μη παραμετρικό πλαίσιο και χρειάζεται να αναπτυχθεί η κατάλληλη θεωρία και μεθοδολογία για να μπορέσουν να αντιμετωπιστούν. Ας δούμε δύο σχετικά παραδείγματα. Το πρώτο μας φέρνει κοντά σε μία γνωστή οικογένεια κατανομών, ενώ το δεύτερο μας φέρνει σε μία καινούρια κατάσταση, όπου δε φαίνεται να μας καλύπτει κάποια γνωστή οικογένεια κατανομών.

Παράδειγμα 1.1: Στο Σχήμα 1.1 απεικονίζονται δύο ιστογράμματα σχετικών συχνοτήτων που αφορούν δεδομένα ύψους 43 φοιτητών και 40 φοιτητριών από ερωτηματολόγιο που δόθηκε σε μάθημα Στατιστικής του Τμήματος Μαθηματικών του ΕΚΠΑ[1].



Σχημα 1.1: Ιστογράμματα σχετικών συχνοτήτων και προσαρμοσμένες κανονικές καμπύλες από δεδομένα ύψους 43 φοιτητών (αριστερά) και 40 φοιτητριών (δεξιά).

Θα φαινόταν ίσως λογικό να υποθέσουμε ότι τα δεδομένα αυτά του ύψους αντιστοιχούν σε μετρήσεις από δύο τυχαία δείγματα, ένα του πληθυσμού των φοιτητών και ένα των φοιτητριών που είναι προσεγγιστικά κανονικά κατανεμημένοι, με τη δική τους μέση τιμή και διασπορά, τα οποία θέλουμε να εκτιμήσουμε. Πριν όμως το δεχθούμε, θα θέλαμε να ελέγξουμε στατιστικά αυτόν τον ισχυρισμό στη βάση των τυχαίων δειγμάτων που διαθέτουμε και η Μη

Παραμετρική Στατιστική θα μας δώσει τα κατάλληλα εργαλεία.

Αν δεχθούμε ότι πράγματι βρισκόμαστε σε κανονική οικογένεια κατανομών, τότε μπορούμε να εφαρμόσουμε κλασικές μεθόδους παραμετρικής Στατιστικής για την εκτίμηση των παραμέτρων. Στο Σχήμα 1.1 οι καμπύλες αντιστοιχούν στην προσαρμογή της κανονικής κατανομής στα δεδομένα, όπως προκύπτουν από τα γραφήματα των συναρτήσεων πυκνότητας πιθανότητας της κανονικής με μέσες τιμές και διασπορές τις εκτιμούμενες στα δύο δείγματα. Η μέση τιμή εκτιμήθηκε από το δειγματικό μέσο $\overline{X}$ και η διασπορά από την αμερόληπτη δειγματική διασπορά S^2 (δεύτερη στήλη του Πίνακα 1.2). Υποθέτοντας λοιπόν ότι τα δεδομένα, και στις δύο περιπτώσεις, αποτελούν τυχαίο δείγμα από κανονική κατανομή, μπορούμε: (i) να εκτιμήσουμε μέσες τιμές και διασπορές, (ii) να κατασκευάσουμε διαστήματα εμπιστοσύνης προκαθορισμένου συντελεστή ($(1-\alpha)-$Δ.Ε.) και (iii) να κάνουμε ελέγχους υποθέσεων αναζητώντας κρίσιμες περιοχές προκαθορισμένου μεγέθους ($\alpha-$Κ.Π.). Στον παρακάτω πίνακα συνοψίζουμε τις εκτιμήτριες, τα $(1-\alpha)$ διαστήματα εμπιστοσύνης (Δ.Ε.) για τις δύο παραμέτρους και τις α κρίσιμες περιοχές για τους αμφίπλευρους ελέγχους $H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$ και $H_0 : \sigma^2 = \sigma_0^2$ vs $H_1 : \sigma^2 \neq \sigma_0^2$.

Πινακας 1.1: Στοιχεία στατιστικής συμπερασματολογίας για τ.δ. $\mathcal{N}(\mu, \sigma^2)$

παράμετρος	εκτιμήτρια	$(1-\alpha)-$Δ.Ε.	$\alpha-$Κ.Π.
μ	$\overline{X}$	$\overline{X} \pm t_{n-1,\alpha/2} \dfrac{S}{\sqrt{n}}$	$\dfrac{\left\lvert\overline{X} - \mu_0\right\rvert}{S/\sqrt{n}} \geq t_{n-1,\alpha/2}$
σ^2	S^2	$\left(\dfrac{(n-1)S^2}{\chi^2_{n-1,\alpha/2}}, \dfrac{(n-1)S^2}{\chi^2_{n-1,1-\alpha/2}}\right)$	$\dfrac{(n-1)S^2}{\sigma_0^2} \notin \left(\chi^2_{n-1,1-\alpha/2}, \chi^2_{n-1,\alpha/2}\right)$

Τα παραπάνω θεωρητικά αποτελέσματα μπορούν να εφαρμοστούν στα δεδομένα ύψους των φοιτητών της Στατιστικής, ξεχωριστά σε κάθε ένα από τα δύο δείγματα, των φοιτητών (μεγέθους $n_1 = 43$) και των φοιτητριών (μεγέθους $n_2 = 40$). Υποθέτουμε αρχικά ότι τα δείγματα αυτά προέρχονται από $\mathcal{N}(\mu_1, \sigma_1^2)$ και $\mathcal{N}(\mu_2, \sigma_2^2)$ αντίστοιχα, με άγνωστες μέσες τιμές και διασπορές. Με τη βοήθεια του Πίνακα 1.1, αν $\overline{X_i}$ και S_i^2, $i = 1, 2$ είναι οι δειγματικοί μέσοι και οι αμερόληπτες δειγματικές διασπορές που αντιστοιχούν στα δύο δείγματα, τότε θέτοντας $\alpha = 0.05$ παίρνουμε τα αποτελέσματα που δίνονται στον Πίνακα 1.2. Σημειώνουμε ότι, λόγω της δυϊκότητας διαστημάτων εμπιστοσύνης και ελέγχων υποθέσεων όταν έχουμε αμφίπλευρες εναλλακτικές, μπορούμε άμεσα να πραγματοποιήσουμε τους ελέγχους με μόνη γνώση των διαστημάτων εμπιστοσύνης.

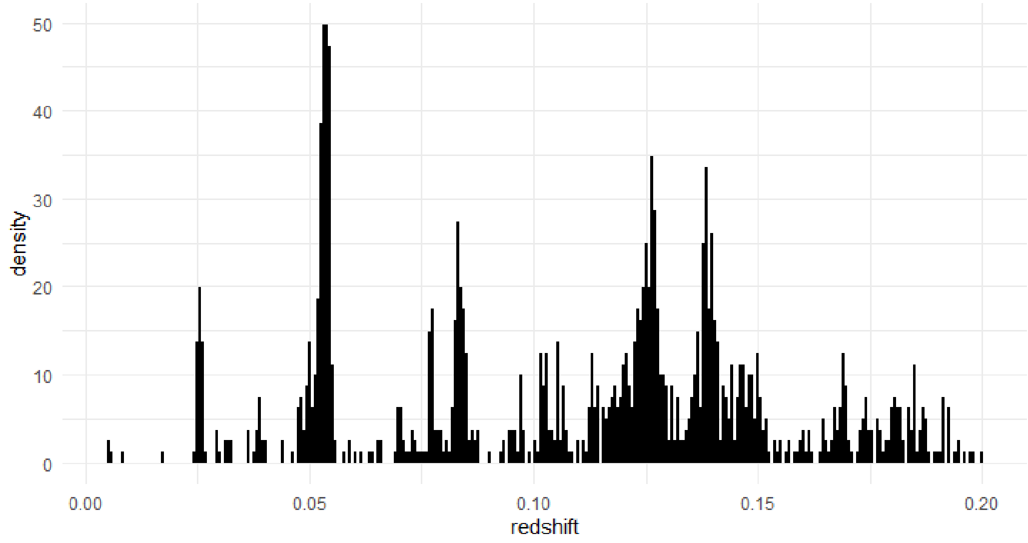
Πινακας 1.2: Εφαρμογή των θεωρητικών αποτελεσμάτων του Πίνακα 1.1 στα δεδομένα του ύψους φοιτητών/φοιτητριών με $\alpha = 0.05$

παράμετρος	εκτιμήτρια	95%−Δ.Ε.	0.05−Κ.Π.
μ_1	$\overline{X_1} = 180.51$	$(178.62\ ,\ 182.41)$	$\mu_{1,0} \notin (178.62\ ,\ 182.41)$
μ_2	$\overline{X_2} = 164.83$	$(162.92\ ,\ 166.73)$	$\mu_{2,0} \notin (162.92\ ,\ 166.73)$
σ_1^2	$S_1^2 = 37.97$	$(25.81\ ,\ 61.34)$	$\sigma_{1,0}^2 \notin (25.81\ ,\ 61.34)$
σ_2^2	$S_2^2 = 35.53$	$(23.84\ ,\ 58.58)$	$\sigma_{2,0}^2 \notin (23.84\ ,\ 58.58)$

Για να έχουν λοιπόν ισχύ όλα τα παραπάνω θα πρέπει να ελέγξουμε συνολικά την κανονικότητα του μοντέλου έναντι της μη κανονικότητας ή τουλάχιστον σε κάποιες συγκεκριμένες περιπτώσεις ελέγχων υποθέσεων να ελέγξουμε μία συγκεκριμένη κανονική κατανομή έναντι οποιασδήποτε άλλης δυνατής κατανομής. Αυτό θα γίνει εφικτό με μεθόδους της Μη Παραμετρικής Στατιστικής που θα συναντήσουμε αργότερα.

Στο Παράδειγμα 1.1 φαίνεται να είμαστε πολύ κοντά σε κάποιο παραμετρικό μοντέλο και θα επικαλεστούμε αργότερα ένα μη παραμετρικό πλαίσιο στην προσπάθειά μας να έχουμε εγγυήσεις για την καταλληλότητα ενός κανονικού μοντέλου. Ας δούμε τώρα και μία άλλη περίπτωση που δεν φαίνεται να εμπίπτει σε κάποιο γνωστό παραμετρικό πλαίσιο.

Παράδειγμα 1.2: Στο Σχήμα 1.2 απεικονίζεται ένα ιστόγραμμα σχετικών συχνοτήτων των αποστάσεων ενός δείγματος $n = 1266$ γαλαξιών από τη Γη. Η απόσταση υπολογίζεται ως συνάρτηση της μετατόπισης που εμφανίζει το ηλεκτρομαγνητικό φάσμα της ακτινοβολίας του κάθε γαλαξία από την ερυθρή περιοχή του φάσματος (redshift). Στην ουσία αυτό που μετριέται είναι η ακτινική ταχύτητα απομάκρυνσης από τη Γη (φαινόμενο Doppler), η οποία είναι άμεσα συνδεδεμένη με την απόσταση, καθώς τα ουράνια αντικείμενα που είναι περισσότερο απομακρυσμένα κινούνται ταχύτερα (διαστολή του σύμπαντος). Για περισσότερες λεπτομέρειες σχετικά με την εφαρμογή αυτή βλ. [1] και τις αναφορές εκεί.



Σχημα 1.2: Ιστόγραμμα σχετικών συχνοτήτων των μετατοπίσεων από το ερυθρό 1266 γαλαξιών. Το πλήθος των κλάσεων που έχει επιλεγεί για την ομαδοποίηση των γαλαξιών είναι 308 και έχει προκύψει από cross-validation.

Η χαρτογράφηση και η ομαδοποίηση των γαλαξιών ανάλογα με την απόστασή τους από τη Γη αποτελεί ένα κεντρικό πρόβλημα της αστρονομίας. Όπως παρατηρούμε η κατανομή των αποστάσεων είναι αρκετά ασύμμετρη και πολυκόρυφη και η ύπαρξη ομαδοποίησης στους γαλαξίες είναι εμφανής. Είναι επίσης φανερό ότι κανένα γνωστό παραμετρικό μοντέλο δε μπορεί να προσαρμοστεί ικανοποιητικά στα δεδομένα αυτά. Η Μη Παραμετρική Στατιστική θα μας δώσει τα κατάλληλα εργαλεία για να προσαρμόσουμε "ικανοποιητικά" μία λεία καμπύλη πάνω στα δεδομένα αυτά. Θα επανέρθουμε λοιπόν σε αυτό το θέμα αργότερα, όταν θα ασχοληθούμε με προβλήματα εκτίμησης με πυκνότητα. Θα πρέπει βέβαια να τονίσουμε ότι ακόμα και το πολύ απλό πρόβλημα κατασκευής ενός αντιπροσωπευτικού ιστογράμματος δεδομένων είναι ένα μη παραμετρικό πρόβλημα, καθώς η επιλογή του πλήθους των κλάσεων, και άρα του επιπέδου λεπτομέρειας ενός ιστογράμματος, όπως στα Παραδείγματα 1.1 και 1.2 δεν πρέπει να γίνεται χειροκίνητα. Η βέλτιστη επιλογή πρέπει να χειρίζεται ισορροπημένα το δίπολο μεροληψία-διασπορά, όπως θα εξηγηθεί αναλυτικά στη συνέχεια μέσα στα πλαίσια της Μη Παραμετρικής Στατιστικής.

Είναι φανερό ότι ένα παραμετρικό μοντέλο θέτει πολύ αυστηρούς περιορισμούς στη φύση της άγνωστης κατανομής. Διαισθητικά, αν το μοντέλο είναι προσεγγιστικά σωστό, τότε αυτό δεν είναι προβληματικό, αντίθετα ισχυροποιεί τα οποιαδήποτε στατιστικά συμπεράσματα. Σε αρκετές περιπτώσεις όμως, υπάρχουν πολλές αμφιβολίες για τη φύση της άγνωστης κατανομής και ιδιαίτερα όταν αντιμετωπίζουμε ένα πρόβλημα για πρώτη φορά. Για το λόγο αυτό, χαλαρώνοντας αρκετά τις υποθέσεις μας, με τη Μη Παραμετρική Στατιστική θα προσεγγίσουμε τα δεδομένα με μεγαλύτερη ασφάλεια, αυξάνοντας κατά πολύ τους βαθμούς ελευθερίας μας και εξηγώντας ένα πολύ μεγαλύτερο εύρος συμπεριφοράς των δεδομένων. Θα πρέπει βέβαια να γίνει κατανοητό από την αρχή ότι αυτό επιφέρει κάποιο κόστος (μικρότερο ή μεγαλύτερο), που συνδέεται γενικά με την απώλεια ισχύος των στατιστικών συμπερασμάτων που εξάγουμε. Αυτό βέβαια είναι προβληματικό μόνο σε περιπτώσεις που μπορούν να βρεθούν καταλληλότερα παραμετρικά μοντέλα.

Στην επόμενη ενότητα υπενθυμίζουμε τις πιο βασικές παραμετρικές οικογένειες κατανομών.

1.2 Παραμετρικές Οικογένειες Βασικών Κατανομών

Σε εισαγωγικά μαθήματα Στατιστικής, μία από τις βασικές υποθέσεις μοντελοποίησης είναι ότι οι παρατηρήσεις $X_1, X_2, \ldots, X_n$, που αντιστοιχούν σε πολλαπλές μετρήσεις κάποιου υπό μελέτη χαρακτηριστικού X, αποτελούν ένα *τυχαίο δείγμα* (ανεξάρτητες και ισόνομες τυχαίες μεταβλητές) από μία παραμετρική οικογένεια κατανομών $\{\mathcal{K}(\theta) : \theta \in \Theta\}$, όπου $\theta = (\theta_1, \ldots, \theta_d) \in \Theta$ συμβολίζει μία εν γένει άγνωστη διανυσματική παράμετρο θ (πεπερασμένης διάστασης d) που παίρνει τιμές σε κάποιον *παραμετρικό χώρο* $\Theta \subset \mathbb{R}^d$ και καθορίζει πλήρως μία κατανομή $\mathcal{K}(\theta)$. Μία σημαντική διάκριση των κατανομών είναι σε *διακριτές* και *συνεχείς*.

Στον Πίνακα 1.3 συνοψίζονται τα βασικότερα παραδείγματα διακριτών κατανομών[2].

Πίνακας Βασικών Κατανομών: Διακριτές					
Όνομα	Σύμβολο	Συν. Πιθανότητας	Στήριγμα	Μέση Τιμή	Διασπορά
Bernoulli	$\mathcal{Be}(p)$	$p^x\,(1-p)^{1-x}$	$[0\,;1]$	p	$p\,(1-p)$
Διωνυμική	$\mathcal{Bin}(N,p)$	$\binom{N}{x}\,p^x\,(1-p)^{N-x}$	$[0\,;N]$	Np	$Np\,(1-p)$
Γεωμετρική	$\mathcal{Ge}(p)$	$p\,(1-p)^x$	$\mathbb{N}$	$(1-p)/p$	$(1-p)/p^2$
Pascal	$\mathcal{NB}(r,p)$	$\binom{r+x-1}{x}p^r\,(1-p)^x$	$\mathbb{N}$	$r(1-p)/p$	$r(1-p)/p^2$
Poisson	$\mathcal{P}(\lambda)$	$e^{-\lambda}\lambda^x/x!$	$\mathbb{N}$	λ	λ

Παρατήρηση 1.1. Στη Στατιστική, όταν αναφερόμαστε στις οικογένειες κατανομών Bernoulli, Διωνυμική, Γεωμετρική και Pascal (ή Αρνητική Διωνυμική), τότε ως βασική άγνωστη παράμετρο θεωρούμε την πιθανότητα $p \in (0,1)$ που ερμηνεύεται ως η πιθανότητα επιτυχίας σε μία δοκιμή Bernoulli. Σε αυτήν την περίπτωση η αντίστοιχη οικογένεια κατανομών είναι μονοπαραμετρική. Οι επιπλέον “παράμετροι” N και r, που εμφανίζονται στη Διωνυμική και στην Pascal, και ερμηνεύονται ως πλήθος ανεξάρτητων δοκιμών Bernoulli και πλήθος επιτυχιών αντίστοιχα συνήθως θεωρούνται γνωστές. Όταν και αυτές είναι άγνωστες, τότε η αντίστοιχη οικογένεια είναι διπαραμετρική και η επιπλέον παράμετρος παίρνει θετικές ακέραιες τιμές. Η παράμετρος λ που εμφανίζεται στην κατανομή Poisson παίρνει θετικές τιμές, όταν δεν υπάρχει κάποιος άλλος περιορισμός.

Στον Πίνακα 1.4 συνοψίζονται τα βασικότερα παραδείγματα συνεχών κατανομών.

Παρατήρηση 1.2 (παραμετρικός χώρος). Το ευρύτερο δυνατό υποσύνολο που

[2] Δεν υπάρχει καθολικά αποδεκτός συμβολισμός για τις κατανομές. Υπάρχουν διάφορες παραλλαγές.

Πίνακας Βασικών Κατανομών: Συνεχείς

Όνομα	Σύμβολο	Συνάρτηση Πυκνότητας	Στήριγμα	Μέση Τιμή	Διασπορά
Ομοιόμορφη	$\mathcal{U}(a,b)$	$1/(b-a)$	(a,b)	$(a+b)/2$	$(b-a)^2/12$
Κανονική	$\mathcal{N}(\mu,\sigma^2)$	$(2\pi\sigma^2)^{-1/2}\,e^{-(x-\mu)^2/2\sigma^2}$	$\mathbb{R}$	μ	σ^2
Εκθετική	$\mathcal{E}(\theta)$	$\theta^{-1}\,e^{-x/\theta}$	$(0,\infty)$	θ	θ^2
Γάμμα	$\mathcal{G}(\alpha,\theta)$	$\Gamma(\alpha)^{-1}\,\theta^{-\alpha}\,x^{\alpha-1}e^{-x/\theta}$	$(0,\infty)$	$\alpha\theta$	$\alpha\theta^2$
Βήτα	$\mathcal{B}(\alpha,\beta)$	$B(\alpha,\beta)^{-1}\,x^{\alpha-1}(1-x)^{\beta-1}$	$(0,1)$	$\alpha/(\alpha+\beta)$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

μπορούμε να πάρουμε ως πεδίο μεταβολής των παραπάνω παραμέτρων καθορίζεται από τους περιορισμούς $\mu\in\mathbb{R}$, $\theta,\alpha,\beta,\sigma^2>0$, και για την ομοιόμορφη κατανομή $a,b\in\mathbb{R}$ με $a<b$. Από τις διπαραμετρικές οικογένειες μπορούν επίσης να προκύψουν κατάλληλες μονοπαραμετρικές υποθέτοντας ότι κάποια από τις δύο παραμέτρους είναι γνωστή.

Παρατήρηση 1.3 (εναλλακτικές μορφές Εκθετικής/Γάμμα). Η Εκθετική κατανομή, και κατ'επέκταση και η Γάμμα, παρουσιάζονται συνήθως σε 2 μορφές. Η μορφή που εμφανίζεται στον Πίνακα 1.4 προτιμάται στη Στατιστική, καθώς η παράμετρος θ συμπίπτει με τη μέση τιμή της Εκθετικής, στοιχείο βολικό για τη στατιστική συμπερασματολογία. Επιπλέον, υπεισέρχεται ως σταθερά αναλογίας και στην έκφραση της μέσης τιμής της Γάμμα. Η δεύτερη μορφή προκύπτει από την πρώτη με αναπαραμέτρηση. Συγκεκριμένα, αν θέσουμε $\beta=1/\theta$, τότε προκύπτει μία ισοδύναμη μορφή της Γάμμα, που θα συμβολίζουμε με $\mathcal{G}(\alpha,\beta)$ με συνάρτηση πυκνότητας

$$f(x;\alpha,\beta)=\frac{\beta^\alpha}{\Gamma(\alpha)}\,x^{\alpha-1}\,e^{-\beta x},\qquad x>0\quad(\alpha,\beta>0).$$

Η μορφή αυτή είναι βολικότερη για υπολογισμούς διάφορων χαρακτηριστικών της κατανομής και προτιμάται σε μαθήματα Πιθανοτήτων. Η διάκριση των δύο θα γίνεται με ρητή αναφορά στο σύμβολο της δεύτερης παραμέτρου. Η παράμετρος θ, καλείται *παράμετρος κλίμακας* και θα αντιστοιχεί στη μορφή $\mathcal{G}(\alpha,\theta)$, ενώ η παράμετρος β, καλείται *παράμετρος ρυθμού* και θα αντιστοιχεί στη μορφή $\mathcal{G}(\alpha,\beta)$. Ανάλογες συμβάσεις θα ισχύουν και για την Εκθετική κατανομή. Καθώς $\mathcal{G}(1,\cdot)\equiv\mathcal{E}(\cdot)$, θα αναφερόμαστε σε 2 μορφές Εκθετικών, την $\mathcal{E}(\theta)$ και την $\mathcal{E}(\beta)$.

Παρατήρηση 1.4 (στήριγμα). Η έννοια *στήριγμα* μπορεί να αναφέρεται είτε σε *στήριγμα κατανομής* (έννοια που δεν θα ασχοληθούμε καθόλου σε αυτό το σύγγραμμα[3]), είτε σε *στήριγμα συνάρτησης* και είναι έννοιες διαφορετικές. Ιδίως για τις συνεχείς κατανομές, αυτό μπορεί να οδηγήσει σε σύγχυση καθώς μία τέτοια κατανομή δεν έχει μονοσήμαντα ορισμένη συνάρτηση πυκνότητας. Αν f είναι μία συνάρτηση πυκνότητας μιας συνεχούς κατανομής $\mathcal{K}$, τότε το στήριγμα της f αντιστοιχεί στο σύνολο

$$S_f=\{x\in\mathbb{R}:f(x)>0\}.$$

Σημαντικό λοιπόν είναι να σημειώσουμε εδώ, ότι ως στήριγμα εννοούμε το στήριγμα

[3] ο ενδιαφερόμενος αναγνώστης μπορεί να ανατρέξει σε ένα σύγγραμμα θεωρίας Πιθανοτήτων

της συνάρτησης πυκνότητας που θα επιλεγεί. Είναι εύκολο τώρα να παρατηρήσουμε ότι μπορούμε να τροποποιήσουμε την f σε μία $\widetilde{f}$, αλλάζοντας τις τιμές της σε κάποια σημεία, έτσι ώστε η $\widetilde{f}$ να είναι πάλι συνάρτηση πυκνότητας της κατανομής $\mathcal{K}$. Αυτό συμβαίνει διότι μία συνεχής κατανομή αποδίδει μηδενική πιθανότητα μεμονωμένα σε κάθε σημείο και άρα αλλάζοντας τις τιμές της f σε ένα σύνολο σημείων που είναι "αμελητέο" (αν αυτό μας εξυπηρετεί) δεν αλλάζουμε την πιθανότητα που αποδίδει η κατανομή $\mathcal{K}$ στα επιτρεπτά υποσύνολα του $\mathbb{R}$ (στην επόμενη ενότητα θα τα πούμε σύνολα Borel). Αυτό για παράδειγμα συμβαίνει αν τροποποιήσουμε την f σε ένα αριθμήσιμο υποσύνολο του $\mathbb{R}$, που είναι πράγματι "αμελητέο" για μία συνεχή κατανομή. Θα αφήσουμε λοιπόν ένα "βαθμό ελευθερίας" στην επιλογή συνάρτησης πυκνότητας μιας κατανομής, και επομένως στηρίγματος. Συνήθως θα επιλέγονται όσο το δυνατόν "ομαλότερες", αλλά και πάλι αυτό είναι σχετικό. Το απλούστερο παράδειγμα είναι η ομοιόμορφη κατανομή $\mathcal{U}(0,1)$. Η συνάρτηση πυκνότητας μπορεί να οριστεί 1 στο $(0,1)$, στο $[0,1]$, στο $(0,1]$ ή στο $[0,1)$ (και 0 διαφορετικά), με συνηθέστερες τις δύο πρώτες περιπτώσεις. Είναι φανερό ότι δεν υπάρχει a priori κάποιος λόγος να ευνοήσουμε τη μία ή την άλλη και εξαρτάται ίσως από το πρόβλημα. Η επιλογή του στηρίγματος επηρεάζει και το στατιστικό μέρος του προβλήματος. Αν για παράδειγμα θεωρήσουμε ένα τ.δ. από την ομοιόμορφη κατανομή $\mathcal{U}(0,\theta)$, τότε η επιλογή του $[0,\theta]$ ή του $(0,\theta)$ ως στήριγμα επηρεάζει την ύπαρξη της εκτιμήτριας μέγιστης πιθανοφάνειας (γιατί;).

Παρατήρηση 1.5 (πυκνότητα με δείκτρια). Σημειώνουμε επίσης ότι αρκετές φορές μία συνάρτηση πυκνότητας εμφανίζεται στη μορφή

$$f(x) = g(x)\,\mathbf{1}_A(x), \qquad x \in \mathbb{R},$$

όπου $\mathbf{1}_A$ είναι η δείκτρια συνάρτηση ενός συνόλου A και g μία συνάρτηση του x που οι τιμές της μας ενδιαφέρουν μόνο στο σύνολο A. Η παραπάνω μορφή της f συνηθίζεται σε πιο μαθηματικά κείμενα και είναι "οικονομικότερη" από την πιο διαισθητική αναπαράσταση

$$f(x) = \begin{cases} g(x), & \text{αν } x \in A, \\ 0, & \text{διαφορετικά.} \end{cases}$$

Έτσι προφανώς η g εκφράζει τον περιορισμό της f στο A και οι τιμές της g εκτός του A (αν ορίζονται) μας είναι αδιάφορες. Είναι αυτός ακριβώς ο περιορισμός πάνω στο στήριγμα S_f της f που αναφέρεται στην τρίτη στήλη του Πίνακα 1.4 ως συνάρτηση πυκνότητας των βασικών συνεχών κατανομών. Σε αυτήν την περίπτωση λοιπόν $A = S_f$ και εκτός του A η συνάρτηση πυκνότητας μηδενίζεται.

1.3 Σύνολα και Συναρτήσεις

Θα χρειαστούμε στη συνέχεια μία καλή εξοικείωση με βασικά στοιχεία της συνολοθεωρίας και μία άνετη χρήση των συνόλων, οικογενειών συνόλων και βασικών ιδιοτήτων τους. Έστω S ένα οποιοδήποτε σύνολο. Θα συμβολίζουμε με $\mathcal{P}(S)$ το *δυναμοσύνολο* του S, δηλ. $\mathcal{P}(S) := \{A : A \subset S\}$. Ως συνήθως, συμβολίζουμε με $A \cup B$ την *ένωση* και με $A \cap B$ ή απλά AB την *τομή* των A και B. Αν δεν υπάρχει κίνδυνος σύγχυσης ως προς το βασικό σύνολο αναφοράς θα γράφουμε A^c για το *συμπλήρωμα* του A ή $S \setminus A$ όταν θέλουμε να δηλώσουμε ξεκάθαρα το βασικό σύνολο αναφοράς. Υπενθυμίζουμε ότι τα διαγράμματα του Venn βοηθούν στην αναπαράσταση άλλων πράξεων μεταξύ συνόλων, όπως της *συνολοθεωρητικής διαφοράς* $A \setminus B$ και της *συμμετρικής διαφοράς* $A \triangle B$.

Η ένωση και η τομή επεκτείνονται για αυθαίρετες οικογένειες υποσυνόλων του S. Αν $\mathcal{A} \subset \mathcal{P}(S)$ είναι μία τέτοια οικογένεια, τότε γράφοντάς την και ως $\mathcal{A} = \{A_i\}_{i \in I}$ προκύπτουν διάφορες ισοδύναμες αναπαραστάσεις της ένωσης των συνόλων αυτών:

$$\bigcup \mathcal{A} \quad \text{ή} \quad \bigcup_{A \in \mathcal{A}} A \quad \text{ή} \quad \bigcup_{i \in I} A_i,$$

ενώ αντίστοιχες είναι και οι εκφράσεις για την τομή τους:

$$\bigcap \mathcal{A} \quad \text{ή} \quad \bigcap_{A \in \mathcal{A}} A \quad \text{ή} \quad \bigcap_{i \in I} A_i.$$

Όταν για συντομία δεν γίνεται αναφορά σε δεικτοσύνολο, π.χ., $\bigcup_i A_i$ ή $\bigcup_n A_n$ (αντίστοιχα για τομές), τότε ως σύνολο δεικτών υπονοείται το $\mathbb{N}$ και η ένωση (ή η τομή) θα θεωρείται αριθμήσιμη. Ερμηνεύουμε λοιπόν την ένωση μίας οικογένειας συνόλων ως το σύνολο εκείνο που ενώνει (κάτω από την ίδια στέγη) όλα τα μέλη των συνόλων που ανήκουν στη συλλογή (χωρίς να επιτρέπει επαναλήψεις στοιχείων). Αντίστοιχα, ερμηνεύουμε την τομή μίας οικογένειας συνόλων ως το σύνολο εκείνο που απομονώνει τα μέλη εκείνα που συνυπάρχουν σε όλα τα σύνολα τα συλλογής. Προφανώς,

$$\emptyset \subset \cap \mathcal{A} \subset \cup \mathcal{A} \subset S.$$

Ιδιαίτερο ενδιαφέρον παρουσιάζεται στις δύο οριακές περιπτώσεις, όταν δηλ. $\cap \mathcal{A} = \emptyset$ ή όταν $\cup \mathcal{A} = S$. Στην πρώτη περίπτωση τα σύνολα είναι "από κοινού ξένα", με την έννοια ότι δεν έχουν κάποιο κοινό σημείο, ενώ στη δεύτερη η συλλογή $\mathcal{A}$ "καλύπτει" το S. Υπενθυμίζουμε ότι αν για δύο σύνολα A, B ισχύει ότι $AB = \emptyset$, τότε τα A και B λέγονται *ξένα μεταξύ τους* και η αντίστοιχη ένωση $A \cup B$ *ξένη ένωση*. Προκειμένου να κατανοηθούν καλύτερα διάφορες ιδιότητες οικογενειών υποσυνόλων που θα μελετήσουμε δίνουμε κάποιους περαιτέρω ορισμούς.

Ορισμός 1.1: Μία οικογένεια υποσυνόλων $\mathcal{A}$ του S καλείται
1. *κατά ζεύγη ξένη*, αν $AB = \emptyset$ για κάθε $A, B \in \mathcal{A}$ με $A \neq B$,
2. *από κοινού ξένη*, αν $\bigcap \mathcal{A} = \emptyset$,
3. *κάλυμμα* του S, αν $S \subset \bigcup \mathcal{A}$,
4. *διαμέριση* του S, αν $\emptyset \notin \mathcal{A}$, είναι κατά ζεύγη ξένη και κάλυμμα του S.

Παρατήρηση 1.6. 1. Η διαμέριση λέγεται πεπερασμένη ή αριθμήσιμη αν η οικογένεια $\mathcal{A}$ είναι πεπερασμένη ή αριθμήσιμη αντίστοιχα. Ανάλογοι χαρακτηρισμοί ισχύουν και για το κάλυμμα.

2. Ο προσδιορισμός *$\mathcal{A}$-κάλυμμα* (*$\mathcal{A}$-διαμέριση*) του S αναφέρεται στο σχηματισμό ενός καλύμματος (διαμέρισης) του S αντίστοιχα με επιλογή συνόλων από μία οικογένεια $\mathcal{A}$.

3. Οι έννοιες κάλυμμα, $\mathcal{A}$-κάλυμμα, διαμέριση και $\mathcal{A}$-διαμέριση του S, επεκτείνονται φυσιολογικά σε οποιοδήποτε μη κενό υποσύνολο B του S, αντικαθιστώντας στον Ορισμό 1.1–(iii) και 1.1–(iv) το S με το B.

Ιδιαίτερο ενδιαφέρον θα έχει και η μελέτη συναρτήσεων μεταξύ τοπολογικών και μετρήσιμων χώρων. Υπενθυμίζουμε πρώτα ότι το σύνολο των συναρτήσεων από ένα σύνολο R σε ένα σύνολο S συμβολίζεται με S^R, δηλ. $S^R := \{f \mid f$ είναι συνάρτηση από το R στο Αν $A \subset R$, $B \subset S$, τότε συμβολίζουμε κατά τα γνωστά με $f(A) := \{f(x) : x \in A\}$ την εικόνα του A μέσω της f και με $f^{-1}(B) := \{x \in R : f(x) \in B\}$ την αντίστροφη εικόνα του B μέσω της f. Υπενθυμίζουμε επίσης ότι ενώ η εναλλαγή των f και f^{-1} με την ένωση είναι πάντα εφικτή:

$$f\left(\bigcup_{i\in I} A_i\right) = \bigcup_{i\in I} f(A_i) \quad \text{και} \quad f^{-1}\left(\bigcup_{i\in I} B_i\right) = \bigcup_{i\in I} f^{-1}(B_i)$$

για οποιεσδήποτε οικογένειες $(A_i)_{i\in I}$ και $(B_i)_{i\in I}$ στην $\mathcal{P}(R)$ και $\mathcal{P}(S)$ αντίστοιχα, δε συμβαίνει το ίδιο με την τομή και το συμπλήρωμα. Την ιδιότητα αυτή απολαμβάνει μόνο η f^{-1}

$$f^{-1}\left(\bigcap_{i\in I} B_i\right) = \bigcap_{i\in I} f^{-1}(B_i) \quad \text{και} \quad f^{-1}(B^{\mathsf{c}}) = \left(f^{-1}(B)\right)^{\mathsf{c}},$$

ενώ για την f δεν ισχύουν οι αντίστοιχες ιδιότητες, παρά μόνο ότι

$$f\left(\bigcap_{i\in I} A_i\right) \subset \bigcap_{i\in I} f(A_i) \quad \text{και} \quad f(R) \setminus f(A) \subset f(R \setminus A).$$

Παρ'όλα αυτά στην ειδική περίπτωση που η f είναι 1-1, οι παραπάνω σχέσεις ισχύουν ως ισότητα και αν υποθέσουμε επιπλέον ότι η f είναι επί, τότε είναι δυνατή η εναλλαγή και με το συμπλήρωμα. Αν $\mathcal{A} \subset \mathcal{P}(R)$ και $\mathcal{B} \subset \mathcal{P}(S)$, τότε συμβολίζουμε με

$$f(\mathcal{A}) := \{f(A) \,:\, A \in \mathcal{A}\} \quad \text{και} \quad f^{-1}(\mathcal{A}) := \{f^{-1}(A) \,:\, A \in \mathcal{A}\}.$$

Η χρήση των παραπάνω συμβολισμών θα είναι ιδιαίτερα βολική στη συνέχεια όταν θα δουλεύουμε με οικογένειες υποσυνόλων. Είναι φανερό ότι $f(\mathcal{A}) \subset \mathcal{P}(S)$ και $f^{-1}(\mathcal{A}) \subset \mathcal{P}(R)$.

Ιδιαίτερα σημαντικές στη θεωρία Πιθανοτήτων και τη Στατιστική είναι οι δείκτριες συναρτήσεις. Οι συναρτήσεις αυτές που παίρνουν τιμές μόνο 0 ή 1 και είναι οι απλούστερες των απλών συναρτήσεων όπως θα δούμε αργότερα και αποτελούν τις βασικές δομικές μονάδες με τη βοήθεια των οποίων προσεγγίζουμε κάθε μετρήσιμη συνάρτηση.

Ορισμός 1.2: Αν $A \subset S$, τότε συμβολίζουμε με $\mathbf{1}_A : S \to \{0,1\}$ τη *δείκτρια συνάρτηση* του A, που ορίζεται ως:

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{αν } x \in A, \\ 0 & \text{αν } x \in A^{\mathsf{c}}. \end{cases}$$

Παρατήρηση 1.7. Η συνάρτηση $f : \mathcal{P}(S) \to \{0,1\}^S \equiv 2^S$, όπου $f(A) = \mathbf{1}_A$ είναι 1-1 και επί, και έτσι μπορούμε να ταυτίσουμε συνολοθεωρητικά κάθε υποσύνολο με τη δείκτρια συνάρτηση $\mathbf{1}_A$. Η ταύτιση αυτή είναι πολύ χρήσιμη για να κατανοήσουμε καλύτερα αρκετές ιδιότητες των συνόλων και όπως θα δούμε αργότερα μας επιτρέπει να ορίσουμε και μία τοπολογία στο $\mathcal{P}(S)$. Προς το παρόν, αφού υπενθυμίσουμε τις στοιχειώδεις ιδιότητες των δεικτριών συναρτήσεων, θα δώσουμε μία ένδειξη προς αυτήν την κατεύθυνση ερμηνεύοντας μία "φυσιολογική" σύγκλιση ακολουθιών υποσυνόλων του S με την κατά σημείο σύγκλιση δεικτριών συναρτήσεων.

Πρόταση 1.1: Έστω $A, B \in \mathcal{P}(S)$, $(A_i)_{i \in I} \subset \mathcal{P}(S)$ αυθαίρετη οικογένεια και $(A_n) \subset \mathcal{P}(S)$ αριθμήσιμη οικογένεια. Ισχύουν τα εξής:

1. $\mathbf{1}_{A^{\mathsf{c}}} = 1 - \mathbf{1}_A$,
2. $\mathbf{1}_{AB} = \mathbf{1}_A \mathbf{1}_B$,
3. $\mathbf{1}_{A \setminus B} = \mathbf{1}_A (1 - \mathbf{1}_B)$,
4. $\mathbf{1}_{A \cup B} = \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_{AB}$,
5. $\mathbf{1}_{\cup_{i \in I} A_i} = \max_{i \in I}\{\mathbf{1}_{A_i}\}$,
6. $\mathbf{1}_{\cap_{i \in I} A_i} = \min_{i \in I}\{\mathbf{1}_{A_i}\}$,
7. $\mathbf{1}_{\cup_n A_n} = \sum_n \mathbf{1}_{A_n}$, αν $A_n A_m = \emptyset$, για $n \neq m$.

Θα ορίσουμε πρώτα ανεξάρτητα μία σύγκλιση ακολουθιών συνόλων και στη συνέχεια θα δείξουμε τη σύνδεση που αυτή έχει με τη σύγκλιση δεικτριών συναρτήσεων.

Ορισμός 1.3: Έστω (A_n) μία ακολουθία υποσυνόλων του S. Το σύνολο

$$\limsup A_n := \bigcap_n \bigcup_{k \geq n} A_k \quad (\text{ή } \ \overline{\lim}\, A_n)$$

λέγεται *ανώτερο όριο* της ακολουθίας (A_n), και το σύνολο

$$\liminf A_n := \bigcup_n \bigcap_{k \geq n} A_k \quad (\text{ή } \ \underline{\lim}\, A_n)$$

λέγεται *κατώτερο όριο* της ακολουθίας (A_n). Θα λέμε ότι η ακολουθία (A_n) *συγκλίνει* αν $\liminf A_n = \limsup A_n$ και η κοινή τιμή θα λέγεται *όριο της ακολουθίας*. Θα γράφουμε $\lim A_n = A$ ή $A_n \to A$ για να δηλώσουμε τη σύγκλιση της (A_n) στο A.

Παρατήρηση 1.8. 1. Το ανώτερο όριο γράφεται αναλυτικά

$$\limsup A_n = (\underbrace{A_1 \cup A_2 \cup \dots}_{B_1}) \cap (\underbrace{A_2 \cup A_3 \cup \dots}_{B_2}) \cap \dots \cap (\underbrace{A_n \cup A_{n+1} \cup \dots}_{B_n}) \cap \dots$$

$$B_1 \supset B_2 \supset \cdots \supset B_n \supset \cdots$$

Από την παραπάνω σχέση είναι φανερό ότι

$$\limsup A_n = \{x \in S : x \in A_n \text{ για άπειρα } n\}, \tag{1.1}$$

και το $\limsup A_n$ προκύπτει ως η τομή της φθίνουσας ακολουθίας (B_n).

2. Το κατώτερο όριο γράφεται αναλυτικά

$$\liminf A_n = \underbrace{(A_1 \cap A_2 \cap \ldots)}_{C_1} \cup \underset{\subset}{} \underbrace{(A_2 \cap A_3 \cap \ldots)}_{C_2} \cup \ldots \underset{\subset \cdots \subset}{} \cup \underbrace{(A_n \cap A_{n+1} \cap \ldots)}_{C_n} \cup \underset{\subset \cdots}{\ldots}$$

Από την παραπάνω σχέση είναι φανερό ότι

$$\liminf A_n = \{x \in S : x \in A_n \text{ τελικά για κάθε } n\}, \tag{1.2}$$

και το $\liminf A_n$ προκύπτει ως η ένωση της αύξουσας ακολουθίας (C_n).

3. Από τους χαρακτηρισμούς (1.1) και (1.2), προκύπτει ότι

$$\emptyset \ \subset \ \cap_n A_n \ \subset \ \liminf A_n \ \subset \ \limsup A_n \ \subset \ \cup_n A_n \ \subset \ S. \tag{1.3}$$

Για να δείξουμε λοιπόν ότι μία ακολουθία (A_n) συγκλίνει, αρκεί $\limsup A_n \subset \liminf A_n$ ή ισοδύναμα ότι κάθε στοιχείο του S που ανήκει σε άπειρα A_n ανήκει τελικά σε κάθε A_n. Από την άλλη, για να δείξουμε ότι μία ακολουθία A_n δε συγκλίνει, αρκεί να βρούμε ένα στοιχείο του S που ανήκει, αλλά και δεν ανήκει σε άπειρα A_n, δηλ. υπάρχει $x \in S$, με $x \in A_n$ για άπειρα n και $x \notin A_n$ για άπειρα n.

4. Υπενθυμίζουμε ότι για μία ακολουθία πραγματικών αριθμών

$$\limsup \alpha_n := \lim_n \sup_{k \geq n} \alpha_k = \inf_n \sup_{k \geq n} \alpha_k \quad \text{και} \quad \liminf \alpha_n := \lim_n \inf_{k \geq n} \alpha_k = \sup_n \inf_{k \geq n} \alpha_k.$$

Αν λοιπόν συμβολίσουμε με $\sup A_n = \cup A_n$ και με $\inf A_n = \cap A_n$, τότε είναι φανερή η αναλογία των ορισμών των $\limsup A_n$ και $\liminf A_n$ για ακολουθίες υποσυνόλων (A_n) με τους αντίστοιχους ορισμούς των $\limsup \alpha_n$ και $\liminf \alpha_n$ για τις ακολουθίες (α_n) πραγματικών αριθμών. Οι έννοιες $\sup A_n$ και $\inf A_n$ έχουν ανάλογες ερμηνείες ως ελάχιστο άνω φράγμα και μέγιστο κάτω φράγμα αντίστοιχα, εδώ βέβαια για ακολουθία υποσυνόλων αυθαίρετου S, αφού με τη σχέση του περιέχεσθαι $\subset$ η ακολουθία (A_n) είναι άνω και κάτω φραγμένη από την $\cup A_n$ και την $\cap A_n$ αντίστοιχα και είναι μάλιστα τα βέλτιστα (ελάχιστο άνω φράγμα και μέγιστο κάτω φράγμα αντίστοιχα) με αυτήν την ιδιότητα.

Υπενθυμίζουμε τώρα την κατά σημείο σύγκλιση ακολουθιών συναρτήσεων σε μία γενική μορφή, καθώς θα τη χρειαστούμε σε αρκετά σημεία στη συνέχεια.

Ορισμός 1.4: Έστω R, S δύο σύνολα και (f_n) μία ακολουθία συναρτήσεων με $f_n \colon R \to S$. Αν στο S έχει οριστεί κάποιου είδους σύγκλιση $\xrightarrow{c}$ ακολουθιών στοιχείων του S (π.χ. μετρική, τοπολογική,...), τότε θα λέμε ότι η ακολουθία (f_n) *συγκλίνει κατά σημείο* σε κάποια συνάρτηση $f\colon R \to S$, και θα γράφουμε

$$f_n \xrightarrow{p.w.} f \quad \Longleftrightarrow \quad f_n(x) \xrightarrow{c} f(x) \quad \text{για κάθε } x \in S,$$

όπου το p.w. είναι συντομογραφία του *pointwise*. Αν θέλουμε να δώσουμε έμφαση στον τρόπο σύγκλισης των στοιχείων του S μπορούμε να γράφουμε $f_n \xrightarrow{p.w.-c} f$.

Παρατήρηση 1.9. Η κατά σημείο σύγκλιση ακολουθιών συναρτήσεων εξαρτάται από τη μεμονωμένη συμπεριφορά των ακολουθιών που προκύπτουν για κάθε σημείο του S και δε συσχετίζονται αυτές μεταξύ τους, όπως θα συνέβαινε για παράδειγμα αν μας ενδιέφερε μία ομοιόμορφη σύγκλιση.

Παρατήρηση 1.10. Το σύμβολο $\xrightarrow{c}$ δε δηλώνει κάποιο συγκεκριμένο τρόπο σύγκλισης, αλλά θα αντικαθίσταται με το σύμβολο που χρησιμοποιείται για να δηλώσει τον ακριβή τρόπο σύγκλισης στο S. Για παράδειγμα, αν αυτή αναφέρεται σε σύγκλιση ακολουθιών ως προς μία μετρική d ή στη σύγκλιση κατά κατανομή ακολουθιών τυχαίων μεταβλητών θα έχουμε $c = d$, αν αναφέρεται στη σύγκλιση κατά πιθανότητα $c = p$, και παρόμοια για άλλου είδους συγκλίσεις.

Στην επόμενη πρόταση αναδεικνύεται η σύνδεση της σύγκλισης ακολουθιών συνόλων με την κατά σημείο σύγκλιση των αντίστοιχων δεικτριών συναρτήσεων.

Πρόταση 1.2: Έστω S σύνολο, $A \subset S$ και (A_n) ακολουθία υποσυνόλων του S. Τότε

$$A_n \to A \quad \text{αν και μόνο αν} \quad \mathbf{1}_{A_n} \xrightarrow{p.w.} \mathbf{1}_A,$$

όπου $\xrightarrow{p.w.}$ (pointwise) συμβολίζει την κατά σημείο σύγκλιση ακολουθιών συναρτήσεων (βλ. Ορισμό 1.4).

Δε συγκλίνουν βέβαια όλες οι ακολουθίες υποσυνόλων, αλλά κάθε ακολουθία είναι φραγμένη ($\emptyset \subset A_n \subset S$) και άρα αν η ακολουθία είναι μονότονη τότε έχουμε πάντα σύγκλιση.

Πρόταση 1.3: Κάθε μονότονη ακολουθία (A_n) υποσυνόλων του S συγκλίνει και ισχύει ότι:

$$\lim A_n = \begin{cases} \bigcup_n A_n & \text{αν } (A_n) \text{ αύξουσα,} \\ \bigcap_n A_n & \text{αν } (A_n) \text{ φθίνουσα.} \end{cases}$$

1.4 Χώρος Πιθανότητας - Τυχαία Μεταβλητή - Κατανομή

Βασική προϋπόθεση μίας βαθύτερης κατανόησης των στατιστικών προβλημάτων που εμφανίζονται σε πραγματικές εφαρμογές είναι μία καλή γνώση του πιθανοθεωρητικού πλαισίου πάνω στο οποίο βασίζεται ένα στατιστικό μοντέλο. Ξεκινάμε με μία υπενθύμιση του χώρου πιθανότητας και της τυχαίας μεταβλητής.

Η μοντελοποίηση ενός στοχαστικού φαινομένου ή ενός πειράματος τύχης γίνεται με την έννοια του χώρου πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$. Το σύνολο Ω περιλαμβάνει τα δυνατά αποτελέσματα (εξαγόμενα) ω ενός πειράματος τύχης, αυτά που λέμε *δειγματικά σημεία* και έτσι καλείται *δειγματικός χώρος*. Η συλλογή $\mathcal{A}$ αποτελείται από όλα τα σύνολα στα οποία μας ενδιαφέρει να αποδώσουμε πιθανότητα, τα λεγόμενα *ενδεχόμενα*, με την απαίτηση να ικανοποιούν κάποιες ιδιότητες κλειστότητας που είναι συμβατές με συνήθεις συνολοθεωρητικές πράξεις. Η προκύπτουσα δομή είναι γνωστή με το όνομα σ-άλγεβρα και μοντελοποιεί μαθηματικά την *πληροφορία* που συνδέεται με ένα πείραμα τύχης.

Ορισμός 1.5 (σ-άλγεβρα): Μία συλλογή (κλάση ή οικογένεια) $\mathcal{A}$ υποσυνόλων του Ω καλείται *σ-άλγεβρα* στο Ω (ή επί του Ω) αν:

(i) $\mathcal{A} \neq \emptyset$ (ή $\emptyset \in \mathcal{A}$ ή $\Omega \in \mathcal{A}$),

(ii) αν $A \in \mathcal{A}$, τότε $A^c \in \mathcal{A}$ [κλειστή στα συμπληρώματα],

(iii) αν (A_n) ακολουθία στην $\mathcal{A}$, τότε $\bigcup\limits_n A_n \in \mathcal{A}$ [κλειστή στις (άπειρες) αριθμήσιμες ενώσεις].

Το ζεύγος $(\Omega, \mathcal{A})$ καλείται *μετρήσιμος χώρος* και τα στοιχεία της $\mathcal{A}$ καλούνται *$\mathcal{A}$-μετρήσιμα σύνολα.*

Ο μετρήσιμος χώρος συμπληρώνεται με μία βασική συνολοσυνάρτηση $\mathbb{P}$, που με τον δικό της τρόπο θα αποδώσει τελικά πιθανότητα, δηλαδή "μέτρο βεβαιότητας", στα διάφορα ενδεχόμενα του πειράματος τύχης. Έτσι, φτάνουμε στον χώρο πιθανότητας.

Ορισμός 1.6 (χώρος πιθανότητας): Έστω $(\Omega, \mathcal{A})$ μετρήσιμος χώρος. Μία συνολοσυνάρτηση $\mathbb{P}\colon \mathcal{A} \to [0,1]$ καλείται *μέτρο πιθανότητας* (ή απλούστερα *πιθανότητα*) στον $(\Omega, \mathcal{A})$ αν

1. $\mathbb{P}(\Omega) = 1$,
2. $\mathbb{P}(\cup_n A_n) = \sum_n \mathbb{P}(A_n)$, για κάθε κατά ζεύγη ξένη ακολουθία[4] (A_n) στην $\mathcal{A}$ (αριθμήσιμη ή σ-προσθετικότητα).

Η τριάδα $(\Omega, \mathcal{A}, \mathbb{P})$ καλείται *χώρος πιθανότητας*.

Παρατήρηση 1.11. Όπως βλέπουμε, οι απαιτήσεις που έχει μία πιθανότητα είναι μικρές, να δίνει μέτρο 1 στο βέβαιο ενδεχόμενο Ω και να είναι αθροιστική σε ξένα ενδεχόμενα, ακόμα και άπειρα το πλήθος, αρκεί να μπορούμε να τα απαριθμήσουμε (σ-προσθετικότητα).

Παρατήρηση 1.12 (η πιθανότητα ως μέτρο). Στη θεωρία μέτρου μία πιθανότητα $\mathbb{P}$ αποτελεί ειδική περίπτωση μέτρου. Μία συνολοσυνάρτηση $\mu\colon \mathcal{A} \to [0,\infty]$ είναι *μέτρο* αν

1. $\mu(\emptyset) = 0$,
2. είναι σ-προσθετική.

---

[4] $A_i A_j = \emptyset$ για $i \neq j$.

Μία συνολοσυνάρτηση πιθανότητας ικανοποιεί τις παραπάνω απαιτήσεις και αυτό αιτιολογεί τον προσδιορισμό μέτρο που δώσαμε στον ορισμό της.

Σε κάθε δειγματικό σημείο ω μας ενδιαφέρει να αποδώσουμε διάφορα ποσοτικά ή ποιοτικά χαρακτηριστικά. Κάθε τέτοιο χαρακτηριστικό X μπορούμε να το δούμε ενιαία ως μία κατάλληλη συνάρτηση με πεδίο ορισμού το Ω. Αν οι τιμές της είναι μέσα στο σύνολο $\mathbb{R}$, τότε οδηγούμαστε στην έννοια της (πραγματικής) τυχαίας μεταβλητής. Για τον τυπικό ορισμό της απαιτείται μία έννοια μετρησιμότητας που θα δούμε παρακάτω. Η ουσία είναι ότι η συνάρτηση X δημιουργεί έναν καινούργιο δειγματικό χώρο. Δεν ενδιαφερόμαστε πλέον για τα δειγματικά σημεία ω, αλλά για τις τιμές του χαρακτηριστικού X, δηλαδή, τα καινούργια σημεία $X(\omega) = x$.

Για λόγους ενοποίησης, αλλά και άλλους πρακτικούς λόγους, αντί να θεωρήσουμε ως δειγματικό χώρο τον $X(\Omega)$, είναι πιο βολικό να δουλεύουμε σε ένα κοινό πλαίσιο για όλα τα χαρακτηριστικά που παίρνουν πραγματικές τιμές. Έτσι μπορούμε να θεωρήσουμε ότι ο κοινός αυτός δειγματικός χώρος είναι το $\mathbb{R}$. Όπως και πριν, έτσι και τώρα, χρειαζόμαστε και μία συλλογή ενδεχομένων που θέλουμε και μπορούμε να τους αποδώσουμε πιθανότητα. Για διάφορους λόγους, που εξηγούνται σε μαθήματα θεωρητικών Πιθανοτήτων, μία συλλογή που μας καλύπτει είναι η οικογένεια των συνόλων *Borel* του $\mathbb{R}$ (δείτε Κεφάλαιο 1.7), την οποία συμβολίζουμε με $\mathcal{B}(\mathbb{R})$. Είναι αρκετά πλούσια για να περιλαμβάνει σχεδόν όλα τα σύνολα που εμφανίζονται σε πρακτικές εφαρμογές, αλλά και για να σέβεται τις ιδιότητες κλειστότητας που αναφέραμε πριν. Είναι ταυτόχρονα όσο μικρή πρέπει, για να είναι συμβατή με τα αξιώματα που θέλουμε να ικανοποιούν οι συνολοσυναρτήσεις πιθανότητας πάνω σε υποσύνολα του $\mathbb{R}$. Υπενθυμίζουμε λοιπόν τώρα τον ορισμό μίας πραγματικής τυχαίας μεταβλητής.

Ορισμός 1.7 (τυχαία μεταβλητή): Έστω $(\Omega, \mathcal{A}, \mathbb{P})$ ένας χώρος πιθανότητας και $\mathcal{B}(\mathbb{R})$ τα σύνολα Borel του $\mathbb{R}$. Μία πραγματική συνάρτηση $X\colon \Omega \to \mathbb{R}$ καλείται *πραγματική τυχαία μεταβλητή* αν ικανοποιεί τη σχέση

$$X^{-1}(B) := \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}, \quad \forall\, B \in \mathcal{B}(\mathbb{R}). \tag{1.4}$$

Το σύνολο $X^{-1}(B)$ το συμβολίζουμε με $\{X \in B\}$.

Παρατήρηση 1.13. Η απαίτηση να επιστρέφει όλα τα *Borel* υποσύνολα του $\mathbb{R}$ μέσα στην σ-άλγεβρα $\mathcal{A}$ είναι αναγκαία από την άποψη ότι μόνο τότε είναι ενδεχόμενα και άρα μπορούμε να τους αποδώσουμε πιθανότητα μέσω της συνολοσυνάρτησης $\mathbb{P}$.

Παρατήρηση 1.14 (μετρήσιμη συνάρτηση). Ο αυστηρός ορισμός μίας τυχαίας μεταβλητής είναι μία εξειδίκευση της έννοιας της μετρήσιμης συνάρτησης όταν αυτή παίρνει τιμές στο $\mathbb{R}$. Συγκεκριμένα, μία συνάρτηση $f\colon \Omega \to S$ μεταξύ 2 μετρήσιμων χώρων $(\Omega, \mathcal{A})$ και $(S, \mathcal{B})$ καλείται $\mathcal{A}/\mathcal{B}-$*μετρήσιμη* αν $f^{-1}(B) \in \mathcal{A}$ για κάθε $B \in \mathcal{B}$. Είναι φανερό λοιπόν ότι ο ορισμός της πραγματικής τυχαίας μεταβλητής προκύπτει για $(S, \mathcal{B}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, όταν ο μετρήσιμος χώρος $(\Omega, \mathcal{A})$ είναι εφοδιασμένος με κάποιο μέτρο πιθανότητας.

Κάθε τυχαία μεταβλητή επάγει με τρόπο φυσιολογικό και ένα μέτρο πιθανότητας στα σύνολα Borel του $\mathbb{R}$ προσδιορίζοντας την κατανομή της τυχαίας μεταβλητής. Σχετικός είναι ο επόμενος ορισμός.

Ορισμός 1.8 (κατανομή τυχαίας μεταβλητής): Έστω $(\Omega, \mathcal{A}, \mathbb{P})$ ένας χώρος πιθανότητας και $X\colon \Omega \to \mathbb{R}$ μία τυχαία μεταβλητή. Το μέτρο πιθανότητας $\mathbb{P}_X\colon \mathcal{B}(\mathbb{R}) \to [0,1]$ που ορίζεται από τη σχέση $\mathbb{P}_X(B) = \mathbb{P}(X \in B)$ στο μετρήσιμο χώρο $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ καλείται *κατανομή της τυχαίας μεταβλητής* X.

Ο όρος κατανομή της τ.μ. X αναφέρεται ακριβώς στη συνολοσυνάρτηση $\mathbb{P}_X$, μπορεί να δειχθεί ότι ικανοποιεί τα αξιώματα της πιθανότητας (ασκ. **??**) και είναι αντικείμενο μελέτης της Θεωρίας Πιθανοτήτων. Έτσι, μπορούμε να αντιληφθούμε τον χώρο $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X)$ ως έναν καινούργιο χώρο πιθανότητας με δειγματικά σημεία τα $x \in \mathbb{R}$, ενδεχόμενα τα σύνολα Borel του $\mathbb{R}$ και μέτρο πιθανότητας την κατανομή $\mathbb{P}_X$ της τ.μ. X. Αυτός ο χώρος συνήθως μας καλύπτει για να πάρουμε ό,τι πληροφορία θέλουμε για τη X, καθώς η φύση του Ω συνήθως είναι αδιάφορη για το πρόβλημα που μελετάμε. Μάλιστα, μας αρκεί πολλές φορές να περιοριστούμε μόνο στη γνώση της κατανομής της X. Έτσι, φυσιολογικά οδηγούμαστε στην έννοια της ισονομίας τυχαίων μεταβλητών.

Ορισμός 1.9 (ισονομία τυχαίων μεταβλητών): Έστω X, Y δύο τυχαίες μεταβλητές με κατανομές $\mathbb{P}_X, \mathbb{P}_Y$ αντίστοιχα. Οι X, Y καλούνται *ισόνομες*, αν $\mathbb{P}_X = \mathbb{P}_Y$, χωρίς να μας ενδιαφέρει καν σε ποιους χώρους ορίζονται.

Παρ'όλο που η έννοια της κατανομής είναι απλή, ως συνολοσυνάρτηση έχει πεδίο ορισμού σύνολα και αυτό δημιουργεί κάποιες δυσκολίες. Η μελέτη των κατανομών διευκολύνεται με τη χρήση απλούστερων συναρτήσεων, πιο βολικών αντικειμένων που χαρακτηρίζουν την κατανομή μιας τυχαίας μεταβλητής. Η συνάρτηση κατανομής είναι ένα τέτοιο παράδειγμα και θα αναφερθούμε εκετενέστερα στις ιδιότητές της στην επόμενη ενότητα. Άλλα παραδείγματα τέτοιων αντικειμένων είναι η χαρακτηριστική συνάρτηση (που ορίζεται μέσω μετασχηματισμού Fourier), που είναι γενικής χρήσης, αλλά και πιο εξειδικευμένα αντικείμενα που χρησιμοποιούνται μόνο για ορισμένες κατηγορίες τ.μ.. Εδώ ανήκουν η συνάρτηση πιθανότητας (για διακριτές τ.μ.), η συνάρτηση πυκνότητας (για απόλυτα συνεχείς τ.μ.) και η ροπογεννήτρια (που ορίζεται μέσω μετασχηματισμού Laplace) που είναι μία πιο "φιλική" έκδοση της χαρακτηριστικής συνάρτησης, αλλά πιο περιορισμένης χρήσης.

1.5 Συνάρτηση Κατανομής

Το πρώτο πράγμα που έρχεται στο νου όταν αναφερόμαστε σε συνάρτηση κατανομής είναι η συνάρτηση κατανομής τυχαίας μεταβλητής. Η σημασία της συνάρτησης κατανομής μιας τ.μ. έγκειται στο ότι χαρακτηρίζει πλήρως την κατανομή της με σχετικά απλό τρόπο. Αν X είναι μία τ.μ., τότε υπενθυμίζουμε ότι με F_X συμβολίζουμε τη συνάρτηση κατανομής της που ορίζεται μέσω της σχέσης $F_X(x) := \mathbb{P}(X \leq x)$. Εύκολα αποδεικνύεται ότι η F_X ικανοποιεί τις ιδιότητες που περιγράφονται στον παρακάτω ορισμό. Είναι σκόπιμο να απομονώσουμε αυτές τις ιδιότητες και να ορίσουμε την έννοια της συνάρτησης κατανομής χωρίς να αντιστοιχεί εξ'αρχής σε κάποια τ.μ. X.

Ορισμός 1.10 (Συνάρτηση Κατανομής): Μία συνάρτηση $F : \mathbb{R} \to [0, 1]$ καλείται *συνάρτηση κατανομής*, αν (i) είναι αύξουσα, (ii) είναι δεξιά συνεχής και (iii) ικανοποιεί $F(-\infty) = 0$ και $F(+\infty) = 1$.

Λόγω της μονοτονίας της, μπορούμε να δείξουμε ότι η F έχει αριστερό όριο σε κάθε σημείο και, μάλιστα, έχει αριθμήσιμο πλήθος σημείων ασυνέχειας. Πρώτα όμως χρειαζόμαστε το εξής γενικότερο αποτέλεσμα.

Λήμμα 1.1: Κάθε μονότονη συνάρτηση $f : \mathbb{R} \to \mathbb{R}$

1. έχει πλευρικά όρια σε κάθε σημείο της και μάλιστα αν η f είναι αύξουσα έχει δεξί όριο
$$f(x+) := \lim_{y \downarrow x} f(y) = \inf \{f(y) : y > x\},$$
και αριστερό όριο
$$f(x-) := \lim_{y \uparrow x} f(y) = \sup \{f(y) : y < x\}.$$
Αν η f είναι φθίνουσα, γίνεται εναλλαγή των *infimum* και *supremum* στις παραπάνω σχέσεις.
2. έχει αριθμήσιμο πλήθος σημείων ασυνέχειας.

Απόδειξη: Δείτε Παράρτημα Α. □

Από το παραπάνω λήμμα, προκύπτουν άμεσα οι ιδιότητες της F.

Πόρισμα 1.1: Κάθε συνάρτηση κατανομής F:

1. έχει αριστερό όριο για κάθε $x \in \mathbb{R}$ και ισχύει ότι
$$F(x-) = \sup \left\{F(y) : y < x\right\},$$
2. είναι συνεχής στα x για τα οποία $F(x-) = F(x)$ [λόγω της δεξιάς συνέχειας εξ'ορισμού] και ασυνεχής όταν $F(x-) < F(x)$ με άλμα ασυνέχειας $F(x) - F(x-)$,
3. έχει αριθμήσιμο πλήθος σημείων ασυνέχειας.

Παρατήρηση 1.15. Είναι απλό να δειχθεί ότι αν μ είναι ένα μέτρο πιθανότητας στο μετρήσιμο χώρο $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ και ορίσουμε $F_\mu \colon \mathbb{R} \to [0, 1]$, όπου
$$F_\mu(x) = \mu\big((-\infty, x]\big), \tag{1.5}$$
τότε η F_μ είναι συνάρτηση κατανομής (ασκ. **??**). Η F_μ καλείται η *συνάρτηση κατα-*

νομής του μέτρου μ.

Παρατήρηση 1.16. Όπως θα δούμε παρακάτω, αν μία συνάρτηση F είναι συνάρτηση κατανομής, τότε είναι και συνάρτηση κατανομής κάποιας τ.μ. X (όχι κατ'ανάγκη μοναδικής), δηλαδή, υπάρχει κατάλληλος χώρος πιθανότητας και τ.μ. X που ορίζεται σε αυτόν έτσι ώστε $F_X = F$, όπου $F_X(x) = \mathbb{P}(X \leq x)$ για κάθε $x \in \mathbb{R}$. Το αποτέλεσμα αυτό θα το αποδείξουμε στην Πρόταση 3.13, αφού ορίσουμε την έννοια της γενικευμένης αντίστροφης.

Με τη βοήθεια της συνάρτησης κατανομής μπορούμε να διακρίνουμε τις τυχαίες μεταβλητές, σε διακριτές και συνεχείς. Είναι τυπικά ορθότερο να κάνουμε πρώτα τη διάκριση γενικότερα για τις κατανομές και στη συνέχεια να το μεταφέρουμε και στις τυχαίες μεταβλητές.

Ορισμός 1.11 (είδη μέτρων πιθανότητας): Έστω μ ένα μέτρο πιθανότητας στο μετρήσιμο χώρο $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ και F_μ η συνάρτηση κατανομής του. Το μ καλείται

1. *διακριτό*, αν η F_μ αυξάνει μόνο με άλματα ασυνέχειας,
2. *συνεχές*, αν η F_μ είναι συνεχής.

Επιπλέον, ένα συνεχές μέτρο μ, καλείται *απόλυτα συνεχές* αν η (συνεχής) F_μ έχει την επιπλέον ιδιότητα οι μεταβολές της μεταξύ 2 σημείων να αναπαρίστανται μέσω ολοκλήρωσης κάποιας μη αρνητικής συνάρτησης στο αντίστοιχο διάστημα, δηλ., αν υπάρχει μία συνάρτηση $f_\mu \geq 0$ τέτοια ώστε για κάθε $x, y \in \mathbb{R}$, με $x < y$,

$$F_\mu(y) - F_\mu(x) = \int_x^y f_\mu(t)\, dt. \tag{1.6}$$

Παρατήρηση 1.17. Αν μ είναι ένα διακριτό μέτρο και A_μ είναι το σύνολο των σημείων ασυνέχειας της F_μ, δηλ.,

$$A_\mu := \big\{x \in \mathbb{R} : F_\mu(x) - F_\mu(x-) > 0\big\},$$

τότε από το Πόρισμα 1.1 αυτό είναι αριθμήσιμο. Από την ιδιότητα της διακριτότητας συμπεραίνουμε λοιπόν ότι το μέτρο πιθανότητας μ συγκεντρώνεται (δίνει πιθανότητα 1) σε ένα αριθμήσιμο σύνολο, αποδίδοντας θετική πιθανότητα σε κάθε ένα από τα σημεία του. Από εδώ προκύπτει και η ιδέα δημιουργίας ενός καινούριου αντικειμένου, της *συνάρτησης πιθανότητας* που περιγράφει τα διακριτά μέτρα με πιο απλό τρόπο. Είναι λοιπόν φανερό ότι αν ορίσουμε τη συνάρτηση

$$f_\mu(x) := \mu\big(\{x\}\big), \qquad \text{για κάθε } x \in \mathbb{R},$$

τότε για κάθε $x \in \mathbb{R}$

$$F_\mu(x) = \sum_{\substack{y \leq x \\ y \in A_\mu}} f_\mu(y),$$

και έτσι η συνάρτηση κατανομής περιγράφεται πλήρως από τη συνάρτηση πιθανότητας. Από την άλλη μεριά, η σχέση (1.6) υποδεικνύει την ύπαρξη μίας *συνάρτησης πυκνότητας* που καθορίζει και αυτή με τρόπο μοναδικό τη συνάρτηση κατανομής ενός απόλυτα συνεχούς μέτρου.

Ο ορισμός των αντίστοιχων εννοιών που αναφέρθηκαν παραπάνω για τυχαίες μεταβλητές είναι πλέον άμεσος.

Ορισμός 1.12 (είδη τυχαίων μεταβλητών): Μία τυχαία μεταβλητή X με κατανομή $\mathbb{P}_X$ καλείται

1. *διακριτή*, αν η κατανομή $\mathbb{P}_X$ είναι διακριτή (δηλ. διακριτό μέτρο πιθανότητας)
2. *συνεχές*, αν η κατανομή $\mathbb{P}_X$ είναι συνεχής (δηλ. συνεχές μέτρο πιθανότητας).

Επιπλέον, μία συνεχής τ.μ., καλείται *απόλυτα συνεχής* αν η συνεχής κατανομή $\mathbb{P}_X$ είναι απόλυτα συνεχής κατανομή (δηλ. απόλυτα συνεχές μέτρο πιθανότητας).

Η ύπαρξη συνάρτησης πιθανότητας και πυκνότητας για διακριτές και απόλυτα συνεχείς τυχαίες μεταβλητές αντίστοιχα προκύπτει τώρα φυσιολογικά από την ύπαρξη συνάρτησης πιθανότητας και πυκνότητας των αντίστοιχων κατανομών τους και συμπίπτουν με αυτές.

1.6 Τοπολογίες και Μετρικοί χώροι

Είδαμε παραπάνω ότι ο ορισμός μιας τ.μ. εμπλέκει την έννοια Borel. Όπως θα δούμε σε λίγο, μετά την εισαγωγή της έννοιας της παραγόμενης σ-άλγεβρας, η ιδιότητα Borel είναι τοπολογική και άρα τα σύνολα Borel του $\mathbb{R}$ καθορίζονται πλήρως από τα ανοικτά σύνολα του $\mathbb{R}$. Υπάρχει λοιπόν και μία ανάγκη καλύτερης κατανόησης τοπολογικών εννοιών που θα ενταθεί όταν θα εργαστούμε σε απειροδιάστατους χώρους. Η επιλογή της μετρικής, αν αυτή υπάρχει, ή γενικότερα της τοπολογίας, αλλά και μιας κατάλληλης σ-άλγεβρας σε τέτοιους χώρους είναι ένα θέμα πιο λεπτό. Γενικά, πολλές διαφορετικές τοπολογίες μπορούν να οριστούν και αυτό επηρεάζει προφανώς και τα σύνολα Borel. Από την άλλη, η κατασκευή σ-αλγεβρών μέσω τοπολογιών δεν θα είναι πάντα η καλύτερη επιλογή. Ξεκινάμε λοιπόν με μία υπενθύμιση της έννοιας της τοπολογίας, που προσδιορίζει τις έννοιες ανοικτό σύνολο και συνεχής συνάρτηση.

Ορισμός 1.13 (τοπολογικός χώρος): Μία συλλογή $\mathcal{T}$ υποσυνόλων του S καλείται *τοπολογία* στο S (ή επί του S) αν:

1. $\emptyset,\ S \in \mathcal{T}$,
2. $A, B \in \mathcal{T} \ \Rightarrow\ AB \in \mathcal{T}$ (η $\mathcal{T}$ είναι κλειστή στις πεπερασμένες τομές),
3. $A_i \in \mathcal{T},\ i \in I \ \Rightarrow\ \cup_{i\in I} A_i \in \mathcal{T}$ για αυθαίρετο σύνολο δεικτών I (η $\mathcal{T}$ είναι κλειστή σε ενώσεις αυθαίρετου πλήθους).

Το ζεύγος $(S, \mathcal{T})$ καλείται *τοπολογικός χώρος* και τα στοιχεία της $\mathcal{T}$ λέγονται *ανοικτά σύνολα ($\mathcal{T}$-ανοικτά)*. Τα υποσύνολα του S που τα συμπληρώματά τους είναι ανοικτά λέγονται *κλειστά σύνολα ($\mathcal{T}$-κλειστά)*. Την οικογένεια των κλειστών συνόλων θα τη συμβολίζουμε με $\mathcal{T}^c$.

Παράδειγμα 1.3: Προκύπτει άμεσα ότι αν $\mathcal{T}$ είναι τοπολογία στο S, τότε

$$\{\emptyset, S\} \subset \mathcal{T} \subset \mathcal{P}(S).$$

- η συλλογή $\{\emptyset, S\}$ είναι τοπολογία και καλείται η *τετριμμένη τοπολογία*.
- το δυναμοσύνολο $\mathcal{P}(S)$ είναι τοπολογία και καλείται η *διακριτή τοπολογία*.

Οι πιο σημαντικές τοπολογίες είναι σίγουρα οι λεγόμενες μετρικές τοπολογίες, αυτές δηλαδή που προέρχονται από μετρικές. Υπενθυμίζουμε παρακάτω την έννοια της μετρικής και του μετρικού χώρου.

Ορισμός 1.14 (μετρικός χώρος): Έστω S σύνολο και $d\colon S \times S \to \mathbb{R}$ συνάρτηση. Η d καλείται *μετρική στο* S και ο (S, d) *μετρικός χώρος*, αν ικανοποιούνται οι εξής ιδιότητες:

1. $d(x, y) = 0 \iff x = y$,
2. $d(x, y) = d(y, x)$, για κάθε $x, y \in S$ (συμμετρική ιδιότητα),
3. $d(x, y) \leq d(x, z) + d(z, y)$, για κάθε $x, y, z \in S$ (τριγωνική ιδιότητα).

Στην περίπτωση που η (i) αντικατασταθεί από την ασθενέστερη συνθήκη

1. $d(x, x) = 0$, για κάθε $x \in S$,

τότε η d καλείται *ψευδομετρική στο* S και ο (S, d) *ψευδομετρικός χώρος*. Σε έναν ψευδομετρικό χώρο επιτρέπονται διακεκριμένα σημεία του χώρου να έχουν μηδενική απόσταση, κάτι που είναι αδύνατον σε ένα μετρικό χώρο. Παρ'όλα αυτά μπορούμε από έναν ψευδομετρικό χώρο να πάμε σε ένα μετρικό χώρο ταυτίζοντας τα σημεία που έχουν μηδενική απόσταση. Αυτό είναι εφικτό μέσω της σχέσης ισοδυναμίας $x \sim y \Leftrightarrow d(x, y) = 0$ (ασκ. **??**).

Παράδειγμα 1.4: Έστω (S, d) ένας μετρικός χώρος. Υπενθυμίζουμε ότι για κάθε $x \in S$ και $\varepsilon > 0$ το σύνολο $B(x, \varepsilon) := \{y \in S : d(x, y) < \varepsilon\}$ καλείται *ανοικτή μπάλα* με κέντρο το x και ακτίνα ε. Ένα σύνολο $A \subset S$ καλείται ανοικτό (ως προς τη μετρική d), αν για κάθε $x \in A$, υπάρχει $\varepsilon_x > 0$ τέτοιο ώστε $B(x, \varepsilon_x) \subset A$. Αν $\mathcal{T}_d = \{A \subset S : A$ ανοικτό ως προς τη μετρική $d\}$, τότε είναι εύκολο να δειχθεί ότι η $\mathcal{T}_d$ ικανοποιεί τις ιδιότητες (i)-(iii) μιας τοπολογίας (ασκ. **??**). Η $\mathcal{T}_d$ καλείται *η μετρική τοπολογία του* X αφού τα μέλη της αντιστοιχούν στα ανοικτά σύνολα ως προς τη μετρική d. Αντίστροφα, αν $\mathcal{T}$ είναι μία τοπολογία επί του S, ιδιαίτερα σημαντικό είναι το ερώτημα κατά πόσον υπάρχει μία μετρική d τέτοια ώστε $\mathcal{T} = \mathcal{T}_d$. Αν υπάρχει, τότε η τοπολογία $\mathcal{T}$ καλείται *μετρικοποιήσιμη*. Για παράδειγμα, η διακριτή τοπολογία $\mathcal{P}(S)$, όπως ορίστηκε παραπάνω, είναι μετρικοποιήσιμη με τη διακριτή μετρική (αυτό αιτιολογεί και το όνομά της), που ορίζεται από τη σχέση $d(x, y) = 1$ για $x \neq y$ και βέβαια $d(x, x) = 0$. Στη διακριτή τοπολογία κάθε σύνολο είναι ανοικτό (και ταυτόχρονα κλειστό).

Υπάρχουν βέβαια και πολλά παραδείγματα μη μετρικοποιήσιμων χώρων. Με την απλή παρατήρηση ότι κάθε μονοσύνολο σε ένα μετρικό χώρο είναι κλειστό σύνολο, συμπεραίνουμε ότι κάθε μετρικοποιήσιμη τοπολογία πρέπει να περιέχει όλα τα σύνολα της μορφής $S \setminus \{x\}$, όπου $x \in S$. Είναι φανερό λοιπόν ότι η τετριμμένη τοπολογία που αντιστοιχεί σε σύνολο S με τουλάχιστον δύο στοιχεία δεν είναι μετρικοποιήσιμη.

Πολλές από τις έννοιες που εμφανίζονται σε μετρικούς χώρους, γενικεύονται για τοπολογικούς χώρους και είναι μάλιστα εκείνες που χαρακτηρίζονται ως τοπολογικές, αφού μπορούν να αναχθούν σε συνθήκες που εμπλέκουν μόνο ανοικτά σύνολα. Παραθέτουμε εδώ 3 σημαντικά παραδείγματα που θα τα συναντήσουμε και παρακάτω.

Ορισμός 1.15 (εσωτερικό/θήκη/σύνορο συνόλου): Έστω $(S, \mathcal{T})$ ένας τοπολογικός χώρος και $A \subset S$. Θα λέμε

- *εσωτερικό* του συνόλου A, το σύνολο

$$A^\circ := \bigcup \Big\{ G \in \mathcal{T} \mid G \subset A \Big\},$$

- *θήκη* του συνόλου A, το σύνολο

$$\overline{A} := \bigcap \Big\{ F \in \mathcal{T}^c \mid F \supset A \Big\},$$

- *σύνορο* του συνόλου A, το σύνολο $\partial A := \overline{A} \setminus A^\circ$.

Παρατήρηση 1.18. Καθώς η ένωση οποιουδήποτε πλήθους ανοικτών συνόλων είναι ανοικτό σύνολο, είναι φανερό ότι το μεγαλύτερο ανοικτό σύνολο που περιέχεται στο A υπάρχει και είναι ακριβώς το A°. Δυϊκά προκύπτει ότι εφόσον η τομή οποιουδήποτε πλήθους κλειστών συνόλων είναι κλειστό σύνολο, το μικρότερο κλειστό σύνολο που περιέχει το A υπάρχει και είναι ακριβώς το $\overline{A}$.

Με τη βοήθεια της θήκης ενός συνόλου καθορίζεται και η έννοια της πυκνότητας, αλλά και της διαχωρισιμότητας ενός τοπολογικού χώρου.

Ορισμός 1.16 (διαχωρίσιμος τοπολογικός χώρος): Έστω $(S, \mathcal{T})$ ένας τοπολογικός χώρος (τ.χ.) και $A \subset S$. Θα λέμε ότι

1. το A είναι ***πυκνό στο*** B, αν $\overline{A} \supset B$ και καλείται απλά ***πυκνό***, αν είναι πυκνό στο S.
2. ο τ.χ. είναι *διαχωρίσιμος*, αν έχει κάποιο αριθμήσιμο πυκνό υποσύνολο.

Παρατήρηση 1.19. Θα ήταν ιδιαιτέρως βολικό να μπορούσαμε να χαρακτηρίσουμε τη θήκη ενός συνόλου A με όρια συγκλινουσών ακολουθιών στοιχείων του A, αφού ορίζαμε βέβαια πρώτα τη σύγκλιση ακολουθιών σε τοπολογικούς χώρους. Τα πράγματα δεν είναι βέβαια τόσο απλά, όπως συμβαίνει στους μετρικούς χώρους που κάτι τέτοιο ισχύει. Ας περιοριστούμε τώρα εδώ σε έναν χαρακτηρισμό που προκύπτει άμεσα από τον ορισμό του:

$$\overline{A} = \{x \in S \mid \text{ για κάθε } G \text{ με } x \in G \in \mathcal{T} \;\Rightarrow\; G \cap A \neq \emptyset\},$$

δηλ. ένα σημείο του χώρου ανήκει στην τοπολογική θήκη $\overline{A}$ ενός συνόλου A, αν και μόνο αν κάθε ανοικτό σύνολο που το περιέχει περιλαμβάνει και σημεία του A.

Η πολύ μεγάλης σημασίας έννοια της συμπάγειας είναι επίσης τοπολογική.

Ορισμός 1.17 (συμπαγές σύνολο): Έστω $(S, \mathcal{T})$ ένας τοπολογικός χώρος και $A \subset S$. Θα λέμε ότι το A είναι *συμπαγές*, αν κάθε *ανοικτό κάλυμμα* του A έχει πεπερασμένο υποκάλυμμα, δηλ. για κάθε υποοικογένεια $\mathcal{G}$ της $\mathcal{T}$ με την ιδιότητα $A \subset \cup\mathcal{G} := \cup_{G\in\mathcal{G}} G$, μπορεί να βρεθεί μία πεπερασμένη υποσυλλογή $\{G_i\}_{i=1}^n$ που να ικανοποιεί την ίδια ιδιότητα, δηλ. $A \subset \cup_{i=1}^n G_i$.

Από τους διαχωρίσιμους τοπολογικούς χώρους που είναι μετρικοποιήσιμοι οι πιο βολικοί είναι εκείνοι για τους οποίους μπορεί να βρεθεί μετρική που κάνει το χώρο πλήρη. Υπενθυμίζουμε εδώ τις σχετικές έννοιες, τονίζοντας ότι η έννοια της πληρότητας δεν είναι τοπολογική και άρα η αναζήτηση αυτή έχει ιδιαίτερο νόημα για μία δεδομένη τοπολογία.

Ορισμός 1.18 (πλήρης/διαχωρίσιμος μετρικός χώρος): Έστω (S, d) ένας μετρικός χώρος.

- Μία ακολουθία (x_n) στοιχείων του S καλείται *βασική*, αν $d(x_n, x_m) \xrightarrow[n,m]{} 0$, δηλ. για κάθε $\varepsilon > 0$, υπάρχει $n_0 \geq 1$ τέτοιο ώστε για κάθε $n, m \geq n_0$, ισχύει $d(x_n, x_m) < \varepsilon$.
- Μία ακολουθία (x_n) στοιχείων του S καλείται *συγκλίνουσα*, αν υπάρχει $x \in S$ με $d(x_n, x) \to 0$ και γράφουμε $x_n \xrightarrow{d} x$, δηλ. για κάθε $\varepsilon > 0$, υπάρχει $n_0 \geq 1$ τέτοιο ώστε για κάθε $n \geq n_0$, ισχύει $d(x_n, x) < \varepsilon$.
- Ο μετρικός χώρος καλείται *πλήρης*, αν κάθε βασική ακολουθία στοιχείων του S είναι συγκλίνουσα.
- Ο μετρικός χώρος καλείται *διαχωρίσιμος*, αν έχει ένα αριθμήσιμο πυκνό υποσύνολο, δηλ. αν υπάρχει ένα αριθμήσιμο $A \subset S$ με $\overline{A} = S$.

Παρατήρηση 1.20. Η ιδιότητα μιας ακολουθίας να είναι βασική, θα μπορούσε να αναδιατυπωθεί με τη χρήση της έννοιας της ουράς μιας ακολουθίας και της διαμέτρου ενός υποσυνόλου ενός μετρικού χώρου. Λέμε *ουρά* μιας ακολουθίας, οποιοδήποτε τμήμα μιας ακολουθίας της μορφής $(x_n)_{n\geq k}$, δηλ. όλους τους όρους της από κάποιον όρο και πέρα. Από την άλλη, για κάθε υποσύνολο A ενός μετρικού χώρου, μπορούμε να ορίσουμε τη διάμετρό του, ας πούμε,

$$\operatorname{diam}(A) := \sup_{x,y\in A} d(x, y),$$

ένα μέτρο της έκτασης ενός συνόλου με όρους μετρικής, που καθορίζει τη μεγαλύτερη δυνατή απόσταση που μπορούν να βρεθούν δύο οποιαδήποτε στοιχεία του συνόλου (με την έννοια του supremum). Αν με $\mathrm{diam}(x_n)$ συμβολίσουμε λοιπόν τη διάμετρο του συνόλου που καθορίζεται από την ουρά της ακολουθίας (x_n), τότε είναι φανερό ότι

$$(x_n) \text{ είναι βασική} \quad \Longleftrightarrow \quad \mathrm{diam}(x_n) \to 0,$$

και άρα οι βασικές ακολουθίες είναι εκείνες των οποίων η ουρά τους "συρρικνώνεται" απεριόριστα.

Παρατήρηση 1.21. Για να αποδείξουμε ότι μία μετρική είναι πλήρης, αρκεί να περιοριστούμε σε βασικές ακολουθίες των οποίων η απόσταση των διαδοχικών τους όρων συγκλίνει γεωμετρικά γρήγορα στο 0, π.χ. αρκετά σύνηθες είναι το 2^{-n}. Πράγματι, κάθε βασική ακολουθία έχει υπακολουθία που ικανοποιεί τη σχέση $d(x_n, x_{n+1}) < 2^{-n}$, για κάθε $n \geq 1$ (προκύπτει άμεσα με επαγωγή από τον ορισμό της βασικής ακολουθίας). Αν έχουμε δείξει ότι αυτή συγκλίνει, τότε η αρχική βασική ακολουθία έχει συγκλίνουσα υπακολουθία και άρα συγκλίνει (γιατί;). Έτσι συμπεραίνουμε ότι ο μετρικός χώρος είναι πλήρης.

Παρατήρηση 1.22. Μία πολύ χρήσιμη αναδιατύπωση της σύγκλισης μιας ακολουθίας (x_n) σε κάποιο σημείο x ως προς μία μετρική d χρησιμοποιεί ανοικτές μπάλες. Μία ακολουθία (x_n) συγκλίνει σε κάποιο $x \in S$ (υποχρεωτικά μοναδικό, όπως αποδεικνύεται άμεσα), αν και μόνο αν οποιαδήποτε ανοικτή μπάλα και αν θεωρήσουμε με κέντρο το x, η ουρά της ακολουθίας θα βρεθεί τελικά μέσα σε αυτήν. Αν $\mathcal{B}_x := \{B(x, \varepsilon) : \varepsilon > 0\}$ συμβολίσουμε την οικογένεια των ανοικτών μπαλών με κέντρο το x, τότε μπορούμε να γράψουμε

$$x_n \xrightarrow{d} x \quad \Longleftrightarrow \quad \text{για κάθε } B \in \mathcal{B}_x, \text{ υπάρχει } n_0 \geq 1 : \text{ για κάθε } n \geq n_0, \; x_n \in B.$$

Καθώς κάθε ανοικτό σύνολο που περιέχει το x περιέχει και μία ανοικτή μπάλα της παραπάνω κατηγορίας, θα δούμε αργότερα ότι αυτός είναι και ο τρόπος που θα ορίσουμε τη σύγκλιση ακολουθίας σε τοπολογικό χώρο, αντικαθιστώντας τις ανοικτές μπάλες από κατάλληλα βασικά ανοικτά σύνολα που θα σχηματίζουν μία βάση ανοικτών περιοχών γύρω από το σημείο x.

Οι παραπάνω ιδιότητες συνδυαστικά προσδιορίζουν μία πολύ σημαντική κατηγορία τοπολογικών χώρων που φέρουν το όνομα Πολωνικοί, προς τιμήν μιας σπουδαίας ομάδας Πολωνών μαθηματικών που τους μελέτησαν, κυρίως τους Casimir Kuratowski, Alfred Tarski και Wacław Sierpiński[5].

> **Ορισμός 1.19:** Ένας τοπολογικός χώρος $(S, \mathcal{T})$ καλείται *Πολωνικός* αν είναι διαχωρίσιμος και πλήρως μετρικοποιήσιμος, δηλ. μετρικοποιείται από μία πλήρη μετρική.

[5] η ορολογία αυτή εισήχθει από τον Roger Godement το 1949 στον τόμο της Γενικής Τοπολογίας της ομάδας Bourbaki

Παράδειγμα 1.5: Τα σημαντικότερα παραδείγματα Πολωνικών χώρων είναι οι χώροι $\mathbb{R}$ και $\mathbb{R}^d$ εφοδιασμένοι με τη συνήθη τοπολογία. Οι χώροι αυτοί είναι διαχωρίσιμοι, π.χ. ένα αριθμήσιμο πυκνό υποσύνολο είναι το σύνολο των ρητών αριθμών $\mathbb{Q}$ και το $\mathbb{Q}^d$ αντίστοιχα, και η συνήθης τοπολογία μετρικοποιείται από την ευκλείδια μετρική, που είναι πλήρης, και άρα είναι πλήρως μετρικοποιήσιμη τοπολογία. Το ανοικτό διάστημα $(0, 1)$, εφοδιασμένο με τη σχετική τοπολογία (της απόλυτης τιμής) είναι και αυτός Πολωνικός χώρος, αν και η σχετική μετρική $d_1(x, y) = |x - y|$ δεν είναι πλήρης. Πράγματι, η ακολουθία $(1/n)_{n\geq 2}$ είναι βασική ως προς d_1, αλλά δε συγκλίνει στο $(0, 1)$. Όμως, μπορούμε να βρούμε ισοδύναμη μετρική, π.χ. θέτουμε $d_2 = d_1/(1 + d_1)$, που είναι πλήρης. Κάθε πεπερασμένος ή αριθμησίμως άπειρος διακριτός τοπολογικός χώρος, π.χ. το $\{1, 2, \ldots, n\}$ ή το $\mathbb{N}$ εφοδιασμένα με τη διακριτή τοπολογία είναι Πολωνικοί χώροι, καθώς είναι αριθμήσιμοι και μετρικοποιήσιμοι από τη διακριτή μετρική που είναι πλήρης. Η πληρότητα της διακριτής μετρικής έπεται από το γεγονός ότι οι μόνες βασικές ακολουθίες είναι οι τελικά σταθερές και άρα προφανώς συγκλίνουν. Αντίθετα, ένας υπεραριθμήσιμος χώρος εφοδιασμένος με τη διακριτή τοπολογία δεν είναι Πολωνικός, αφού δεν είναι διαχωρίσιμος. Η μη διαχωρισιμότητα έπεται από το γεγονός ότι κάθε υποσύνολο ενός διακριτού χώρου είναι κλειστό και έτσι το μόνο πυκνό υποσύνολο του χώρου είναι το ίδιο το βασικό σύνολο που είναι υπεραριθμήσιμο.

1.7 Παραγόμενες Τοπολογίες και σ-άλγεβρες

Στο Παράδειγμα 1.4 είναι φανερό ότι οι ανοικτές μπάλες, που εξαρτώνται άμεσα από τη μετρική του χώρου, καθορίζουν τα ανοικτά σύνολα. Κάθε ανοικτό σύνολο μπορεί να περιγραφεί ως μία ένωση από ανοικτές μπάλες. Σε έναν τοπολογικό χώρο δεν υπάρχει κατ'ανάγκη μετρική, άρα θα ήταν χρήσιμο να απομονώσουμε τη σχέση που συνδέει τα ανοικτά σύνολα με τις μπάλες, αλλά και τις ιδιότητες εκείνες που οι μπάλες ικανοποιούν, έτσι ώστε να μπορέσουμε να γενικεύσουμε την έννοια αυτή στους τοπολογικούς χώρους. Αυτό θα μας επιτρέψει να έχουμε μία απλή περιγραφή των ανοικτών συνόλων, αλλά και ιδιότητες με τη βοήθεια των οποίων θα μπορούμε να παράγουμε τοπολογίες από μικρότερες οικογένειες υποσυνόλων.

Ορισμός 1.20: Έστω $(S, \mathcal{T})$ ένας τοπολογικός χώρος. Μία οικογένεια $\mathcal{B} = \{B_i : i \in I\}$ στοιχείων της $\mathcal{T}$ καλείται *βάση* της τοπολογίας $\mathcal{T}$, αν για κάθε $A \in \mathcal{T}$ υπάρχει $J \subset I$ με

$$A = \bigcup_{j\in J} B_j, \tag{1.7}$$

δηλαδή αν κάθε ανοικτό σύνολο είναι ένωση στοιχείων της $\mathcal{B}$.

Προφανώς μία τοπολογία έχει πολλές διαφορετικές βάσεις, αλλά θα θέλαμε επίσης να γνωρίζουμε πότε μία οικογένεια υποσυνόλων του S μπορεί να αποτελέσει βάση τοπολογίας. Θα θέλαμε δηλαδή, σύμφωνα με τα παραπάνω, οι ενώσεις των στοιχείων της να ικανοποιούν τις ιδιότητες του ορισμού μιας τοπολογίας. Στο επόμενο θεώρημα δίνουμε ικανές και αναγκαίες συνθήκες.

Θεώρημα 1.1: Έστω $\mathcal{B} = \{B_i : i \in I\}$ μία οικογένεια υποσυνόλων του S. Η $\mathcal{B}$ είναι βάση κάποιας τοπολογίας του S, αν και μόνο αν ικανοποιούνται οι εξής ιδιότητες:

-

(ii) αν $B_1, B_2 \in \mathcal{B}$ και $x \in B_1 \cap B_2$, τότε υπάρχει $B_3 \in \mathcal{B}$ ώστε $x \in B_3 \subset B_1 \cap B_2$.

Παρατήρηση 1.23. Μία ικανή (και χρήσιμη) συνθήκη για να ικανοποιείται το (ii) είναι η εξής:

(iii) αν $B_1, B_2 \in \mathcal{B}$ και $B_1 \cap B_2 \neq \emptyset$, τότε $B_1 \cap B_2 \in \mathcal{B}$.

Η συνθήκη αυτή εκφράζει την κλειστότητα στις μη κενές πεπερασμένες τομές. Προφανώς, αν το $\emptyset$ περιέχεται στη $\mathcal{B}$, τότε η συνθήκη αυτή ισοδυναμεί με την κλειστότητα στις πεπερασμένες τομές. Για παράδειγμα, οι ανοικτές μπάλες του $\mathbb{R}$ ταυτίζονται με τα ανοικτά φραγμένα διαστήματα του $\mathbb{R}$ και άρα ικανοποιούν τη συνθήκη (iii) αφού κάθε μη κενή τομή δύο ανοικτών και φραγμένων διαστημάτων του $\mathbb{R}$ είναι ανοικτό και φραγμένο διάστημα. Αντίθετα, οι ανοικτές μπάλες του $\mathbb{R}^2$, δεν ικανοποιούν την (iii), αν και ικανοποιούν τις (i) και (ii).

Αν και μία αυθαίρετη (μη κενή) οικογένεια υποσυνόλων $\mathcal{C}$ του S δεν είναι υποχρεωτικά βάση κάποιας τοπολογίας, εύκολα μπορούμε να δούμε ότι αν φτιάξουμε μία μεγαλύτερη οικογένεια $\mathcal{B}$ που αποτελείται από όλες τις πεπερασμένες τομές στοιχείων της $\mathcal{C}$, τότε η $\mathcal{B}$ είναι βάση μιας μοναδικής τοπολογίας $\mathcal{T}$ που μάλιστα είναι η μικρότερη τοπολογία που περιέχει την οικογένεια $\mathcal{C}$. Έτσι οδηγούμαστε φυσιολογικά στον επόμενο ορισμό.

Ορισμός 1.21 (παραγόμενη τοπολογία): Έστω $\mathcal{C}$ μία μη κενή οικογένεια υποσυνόλων του S. Λέμε *παραγόμενη τοπολογία* της $\mathcal{C}$ ή *τοπολογία που παράγεται* από τη $\mathcal{C}$, και τη συμβολίζουμε με $\tau(\mathcal{C})$, την οικογένεια

$$\tau(\mathcal{C}) := \bigcap \{\mathcal{T} \mid \mathcal{C} \subset \mathcal{T} \text{ και } \mathcal{T} \text{ τοπολογία στο } S\}. \tag{1.8}$$

Παρατήρηση 1.24. Αποδεικνύεται άμεσα (ασκ. **??**) ότι η τομή οποιουδήποτε πλήθους τοπολογιών είναι τοπολογία. Έτσι η $\tau(\mathcal{C})$ που ορίζεται από την (1.8) είναι πράγματι τοπολογία. Ο τρόπος κατασκευής της υποδηλώνει ότι είναι η ελάχιστη τοπολογία που περιέχει τη $\mathcal{C}$, καθώς προφανώς περιέχει τη $\mathcal{C}$ και επιπλέον είναι η τομή όλων των τοπολογιών με αυτήν την ιδιότητα, άρα και η μικρότερη.

Παρατήρηση 1.25. Να σημειώσουμε εδώ ότι κάθε οικογένεια υποσυνόλων $\mathcal{C}$ έχει μία μοναδική παραγόμενη τοπολογία, ενώ αντίθετα, αν $\mathcal{T}$ είναι μία τοπολογία, τότε αυτή δεν παράγεται από μοναδική οικογένεια υποσυνόλων. Κάθε οικογένεια $\mathcal{C}$ που ικανοποιεί τη σχέση $\mathcal{T} = \tau(\mathcal{C})$ καλείται *γεννήτορας* ή *υποβάση* της $\mathcal{T}$. Οι έννοιες αυτές είναι ισοδύναμες, αλλά έχουν διαφορετικές ερμηνείες. Το όνομα γεννήτορας, υποδηλώνει ότι η $\mathcal{T}$ είναι η ελάχιστη τοπολογία που περιέχει τη $\mathcal{C}$, ενώ το όνομα υποβάση συνδέεται άμεσα με τον τρόπο αναπαράστασης των στοιχείων της $\mathcal{T}$ από τα στοιχεία της $\mathcal{C}$. Συγκεκριμένα, είναι σαφές από τα παραπάνω ότι η οικογένεια όλων των πεπερασμένων τομών της $\mathcal{C}$ συνιστά μία βάση $\mathcal{B}$ της τοπολογίας $\mathcal{T}$ και άρα κάθε $\mathcal{T}$-ανοικτό σύνολο αναπαρίσταται στη μορφή

$$A = \bigcup_{i\in I} \bigcap_{j\in J_i} C_{ij}, \tag{1.9}$$

με $C_{ij} \in \mathcal{C}$, I αυθαίρετο σύνολο δεικτών και J_i πεπερασμένο για κάθε $i \in I$.

Η αντίστοιχη έννοια για μία σ-άλγεβρα είναι η παρακάτω.

Ορισμός 1.22 (παραγόμενη σ-άλγεβρα): Έστω $\mathcal{C}$ μία μη κενή οικογένεια υποσυνόλων του S. Λέμε *παραγόμενη σ-άλγεβρα* της $\mathcal{C}$ ή *σ-άλγεβρα που παράγεται* από τη $\mathcal{C}$, και τη συμβολίζουμε με $\sigma(\mathcal{C})$, την οικογένεια

$$\sigma(\mathcal{C}) := \bigcap \big\{\mathcal{B} \mid \mathcal{C} \subset \mathcal{B} \text{ και } \mathcal{B} \text{ σ-άλγεβρα στο } S\big\}. \tag{1.10}$$

Παρατήρηση 1.26. Αποδεικνύεται άμεσα (ασκ. **??**) ότι η τομή οποιουδήποτε πλήθους σ-αλγεβρών είναι σ-άλγεβρα. Έτσι η $\sigma(\mathcal{C})$ που ορίζεται από την (1.10) είναι πράγματι σ-άλγεβρα. Ο τρόπος κατασκευής της υποδηλώνει ότι είναι η ελάχιστη σ-άλγεβρα που περιέχει τη $\mathcal{C}$, καθώς προφανώς περιέχει τη $\mathcal{C}$ και επιπλέον είναι η τομή όλων των σ-αλγεβρών με αυτήν την ιδιότητα, άρα και η μικρότερη. Σε αντίθεση με την παραγόμενη τοπολογία, εκτός από ειδικές περιπτώσεις, κάθε στοιχείο της $\sigma(\mathcal{C})$ δεν έχει μία απλή αναπαράσταση ως συνάρτηση των στοιχείων της $\mathcal{C}$. Εξαίρεση αποτελούν οι σ-άλγεβρες που παράγονται από διαμερίσεις.

Παρατήρηση 1.27. Να σημειώσουμε εδώ ότι κάθε οικογένεια υποσυνόλων $\mathcal{C}$ έχει μία μοναδική παραγόμενη σ-άλγεβρα, ενώ αντίθετα, αν $\mathcal{A}$ είναι μία σ-άλγεβρα, τότε αυτή δεν παράγεται από μοναδική οικογένεια υποσυνόλων. Κάθε οικογένεια $\mathcal{C}$ που ικανοποιεί τη σχέση $\mathcal{A} = \sigma(\mathcal{C})$ καλείται *γεννήτορας* της $\mathcal{A}$.

Είμαστε τώρα σε θέση να ορίσουμε τα σύνολα Borel, δημιουργώντας μία αλληλεπίδραση μεταξύ σ-αλγεβρών και τοπολογιών.

Ορισμός 1.23 (Borel σ-άλγεβρα): Έστω $(S, \mathcal{T})$ ένας τοπολογικός χώρος και $\mathcal{B}(S) := \sigma(\mathcal{T})$.

- Τα στοιχεία της οικογένειας $\mathcal{B}(S)$ λέγονται *σύνολα Borel* στο S και η $\mathcal{B}(S)$ καλείται η *σ-άλγεβρα Borel* του S. Αντιπροσωπεύει την ελάχιστη σ-άλγεβρα που περιέχει τα ανοικτά σύνολα του S. Ο χώρος $(S, \mathcal{B}(S))$ καλείται *χώρος Borel*.
- Στην ειδική περίπτωση που ο τοπολογικός χώρος είναι Πολωνικός, τότε ο $(S, \mathcal{B}(S))$ καλείται *τυπικός χώρος Borel* (standard Borel space).

Παρατήρηση 1.28. Στην ειδική περίπτωση που ο S είναι μετρικός χώρος με μετρική τη d, τότε ως σύνολα Borel του S θα εννοούνται αυτά που παράγονται από τα ανοικτά σύνολα ως προς τη μετρική d.

1.8 Υπόχωροι τοπολογικών και μετρήσιμων χώρων

Πολλές φορές εργαζόμαστε σε χώρους οι οποίοι προκύπτουν φυσιολογικά από ήδη υπάρχοντες και καλά μελετημένους χώρους με έναν κατάλληλο περιορισμό. Άλλες φορές ο περιορισμός αυτός μπορεί να επιβάλλεται για λόγους πρακτικούς. Σε αυτήν την ενότητα ορίζουμε και μελετάμε τις στοιχειώδεις ιδιότητες αυτών των υπόχωρων έτσι όπως προκύπτουν φυσιολογικά από τις ιδιότητες του χώρου από τον οποίο έχουν προκύψει. Αυτό θα γίνει πρώτα για τους τοπολογικούς και στη συνέχεια για τους μετρήσιμους.

Ορισμός 1.24 (τοπολογία ίχνος): Έστω $(S, \mathcal{T})$ ένας τοπολογικός χώρος και $R \subset S$. Η οικογένεια

$$\mathcal{T}_R := \{A \cap R \mid A \in \mathcal{T}\} \tag{1.11}$$

καλείται *τοπολογία ίχνος* (ή *σχετική τοπολογία*, ή *τοπολογία υπόχωρος*, ή *επαγόμενη τοπολογία* ή *τοπολογία περιορισμός*) της $\mathcal{T}$ στο R. Το ζεύγος $(R, \mathcal{T}_R)$ καλείται (τοπολογικός) *υπόχωρος* του $(S, \mathcal{T})$.

Παρατήρηση 1.29. Επαληθεύεται άμεσα ότι η οικογένεια $\mathcal{T}_R$ είναι πράγματι τοπολογία στο R (βλ. ασκ. **??**). Σημειώνουμε επίσης ότι δεν απαιτείται από τον ορισμό το σύνολο R να είναι μέλος της τοπολογίας. Στην περίπτωση που $R \in \mathcal{T}$, τότε είναι φανερό ότι $\mathcal{T}_R \subset \mathcal{T}$ και μάλιστα

$$\mathcal{T}_R = \{A \in \mathcal{T} \mid A \subset R\}.$$

Παρατήρηση 1.30. Έστω S κάποιο σύνολο. Ανεξάρτητα από το αν μία οικογένεια $\mathcal{S}$ υποσυνόλων του S έχει κάποια δομή ή όχι θα ήταν χρήσιμο να ορίσουμε την έννοια του ίχνους πάνω σε κάποιο υποσύνολο R του S. Άρα θα συμβολίζουμε με

$$\mathcal{S}_R := \{A \cap R \mid A \in \mathcal{S}\},$$

και θα λέμε ότι η οικογένεια $\mathcal{S}_R$ είναι το *ίχνος της $\mathcal{S}$ στο R*. Θα μας απασχολήσουν βέβαια οι περιπτώσεις κατά τις οποίες η $\mathcal{S}$ έχει κάποια δομή και θα εξετάζουμε κάτα πόσον η δομή αυτή διατηρείται όταν οδηγηθούμε στο $\mathcal{S}_R$.

Παρατήρηση 1.31. Η τοπολογία ίχνος κληρονομεί φυσιολογικά ιδιότητες από τη "μητέρα" τοπολογία. Κάθε μέλος της $\mathcal{T}_R$ καλείται *ανοικτό στο* R και έχει προκύψει από τον περιορισμό (μέσω της τομής) τουλάχιστον ενός $\mathcal{T}$-ανοικτού συνόλου στο R. Η ανάλογη ιδιότητα ισχύει βέβαια και για τα κλειστά στο R παίρνοντας συμπληρώματα.

Η επόμενη πρόταση κινείται στην ίδια κατεύθυνση και καθορίζει με φυσιολογικό τρόπο υποβασικά και βασικά σύνολα της τοπολογίας ίχνος από αντίστοιχα σύνολα της "μητέρας" τοπολογίας.

Πρόταση 1.4: Έστω $(S, \mathcal{T})$ ένας τοπολογικός χώρος και $\mathcal{C}, \mathcal{B}$ ένας γεννήτορας (υποβάση) και μία βάση της $\mathcal{T}$ αντίστοιχα. Αν $R \subset S$, τότε η

1. $\mathcal{C}_R$ είναι γεννήτορας (υποβάση) της $\mathcal{T}_R$,
2. $\mathcal{B}_R$ είναι βάση της $\mathcal{T}_R$.

Παρατήρηση 1.32. Μπορούμε να αναδιατυπώσουμε συμβολικά το συμπέρασμα (i) της παραπάνω πρότασης ως εξής:

$$\mathcal{T} = \tau(\mathcal{C}) \quad \Longrightarrow \quad \mathcal{T}_R = \tau_R(\mathcal{C}_R),$$

όπου $\tau_R(\mathcal{C}_R)$ είναι η τοπολογία στο R που παράγεται από την οικογένεια $\mathcal{C}_R$. Από αυτή τη σχέση μπορούμε λοιπόν να γράφουμε αλγεβρικά

$$\tau(\mathcal{C})_R = \tau_R(\mathcal{C}_R), \tag{1.12}$$

όπου $\tau(\mathcal{C})_R$ είναι το ίχνος της οικογένειας $\tau(\mathcal{C})$ στο R.

Δίνουμε τώρα αντίστοιχους ορισμούς και αποτελέσματα για τους μετρήσιμους χώρους.

Ορισμός 1.25: Έστω $(S, \mathcal{A})$ ένας μετρήσιμος χώρος και $R \subset S$. Η οικογένεια

$$\mathcal{A}_R = \{A \cap R \mid A \in \mathcal{A}\} \tag{1.13}$$

καλείται *σ-άλγεβρα ίχνος* (ή *σχετική σ-άλγεβρα*, ή *σ-άλγεβρα υπόχωρος*, ή *επαγόμενη σ-άλγεβρα* ή *σ-άλγεβρα περιορισμός*) της $\mathcal{A}$ στο R. Το ζεύγος $(R, \mathcal{A}_R)$ καλείται (μετρήσιμος) *υπόχωρος* του $(S, \mathcal{A})$.

Παρατήρηση 1.33. Επαληθεύεται άμεσα ότι η οικογένεια $\mathcal{A}_R$ είναι πράγματι σ-άλγεβρα στο R (βλ. ασκ. **??**). Σημειώνουμε επίσης ότι δεν απαιτείται από τον ορισμό το σύνολο R να είναι μέλος της $\mathcal{A}$. Στην περίπτωση που $R \in \mathcal{A}$, τότε είναι φανερό ότι $\mathcal{A}_R \subset \mathcal{A}$ και μάλιστα

$$\mathcal{A}_R = \{A \in \mathcal{A} \mid A \subset R\}.$$

Παρατήρηση 1.34. Η σ-άλγεβρα ίχνος κληρονομεί και αυτή φυσιολογικά ιδιότητες από τη "μητέρα" σ-άλγεβρα. Κάθε μέλος της $\mathcal{A}_R$ προκύπτει από τον περιορισμό (μέσω της τομής) τουλάχιστον ενός $\mathcal{A}$-μετρήσιμου συνόλου στο R.

Η επόμενη πρόταση καθορίζει με φυσιολογικό τρόπο γεννήτορες της σ-άλγεβρας ίχνους από γεννήτορες της "μητέρας" σ-άλγεβρας.

Πρόταση 1.5: Έστω $(S, \mathcal{A})$ ένας μετρήσιμος χώρος και $\mathcal{C}$ ένας γεννήτορας της $\mathcal{A}$. Αν $R \subset S$, τότε η $\mathcal{C}_R$ είναι γεννήτορας της $\mathcal{A}_R$.

Παρατήρηση 1.35. Μπορούμε να αναδιατυπώσουμε συμβολικά την παραπάνω πρόταση ως εξής:

$$\mathcal{A} = \sigma(\mathcal{C}) \quad \Longrightarrow \quad \mathcal{A}_R = \sigma_R(\mathcal{C}_R),$$

όπου $\sigma_R(\mathcal{C}_R)$ είναι η σ-άλγεβρα στο R που παράγεται από την οικογένεια $\mathcal{C}_R$. Από αυτή τη σχέση μπορούμε λοιπόν να γράφουμε αλγεβρικά

$$\sigma(\mathcal{C})_R = \sigma_R(\mathcal{C}_R), \tag{1.14}$$

όπου $\sigma(\mathcal{C})_R$ είναι το ίχνος της οικογένειας $\sigma(\mathcal{C})$ στο R.

Είμαστε τώρα σε θέση να δούμε και τη σχέση που συνδέει τη σ-άλγεβρα των συνόλων Borel της σ-άλγεβρας ίχνους με την αρχική σ-άλγεβρα Borel ενός τοπολογικού χώρου.

Πόρισμα 1.2: Έστω $(S, \mathcal{T})$ ένας τοπολογικός χώρος και $(S, \mathcal{B}(S))$ ο αντίστοιχος χώρος Borel. Αν $R \subset S$ και θέσουμε $\mathcal{B}(R)$ τη σ-άλγεβρα Borel του υπόχωρου $(R, \mathcal{T}_R)$, τότε

$$\mathcal{B}(R) = \mathcal{B}(S)_R. \tag{1.15}$$

Γενικότερα, αν $\mathcal{C}$ είναι ένας γεννήτορας της σ-άλγεβρας $\mathcal{B}(S)$, τότε η $\mathcal{C}_R$ είναι γεννήτορας της $\mathcal{B}(R)$.

Απόδειξη: Το πρώτο συμπέρασμα έπεται άμεσα από την Πρόταση 1.5 για $\mathcal{A} = \mathcal{B}(S)$ και $\mathcal{C} = \mathcal{T}$. Το δεύτερο είναι ένα γενικότερο συμπέρασμα που έπεται από την ίδια Πρόταση για οποιοδήποτε γεννήτορα. □

1.9 Χώροι Μέτρου

Στην Παρατήρηση 1.12 αναφέραμε ότι ένα μέτρο πιθανότητας είναι ειδική περίπτωση μέτρου που αποδίδει σε όλο το χώρο μέτρο 1. Στην ενότητα αυτή θα αναφέρουμε εν συντομία κάποιες βασικές έννοιες και στοιχειώδη αποτελέσματα που θα μας χρειαστούν στη συνέχεια.

Ορισμός 1.26 (χώρος μέτρου): Έστω $(S, \mathcal{A})$ μετρήσιμος χώρος. Μία συνολοσυνάρτηση $\mu\colon \mathcal{A} \to [0, +\infty]$ καλείται *μέτρο* στο $(S, \mathcal{A})$ αν

1. $\mu(\emptyset) = 0$,
2. $\mu(\cup_n A_n) = \sum_n \mu(A_n)$, για κάθε κατά ζεύγη ξένη ακολουθία (A_n) στην $\mathcal{A}$ (αριθμήσιμη ή σ-προσθετικότητα).

Η τριάδα $(S, \mathcal{A}, \mu)$ καλείται *χώρος μέτρου*. Αν επιπλέον,

1. $\mu(S) < +\infty$, τότε το μ λέγεται *πεπερασμένο μέτρο* και ο $(S, \mathcal{A}, \mu)$ *χώρος πεπερασμένου μέτρου*,
2. $\mu(S) = 1$, τότε το μ λέγεται *μέτρο πιθανότητας* και ο $(S, \mathcal{A}, \mu)$ *χώρος πιθανότητας*,
3. υπάρχει μία ακολουθία (A_n) στην $\mathcal{A}$ τέτοια ώστε $S = \cup_n A_n$ με $\mu(A_n) < +\infty$ για κάθε n, τότε το μ λέγεται *σ–πεπερασμένο μέτρο* και ο χώρος $(S, \mathcal{A}, \mu)$ *χώρος σ-πεπερασμένου μέτρου*.

Παρατήρηση 1.36. Ο παραπάνω ορισμός αντιστοιχεί στα λεγόμενα *θετικά μέτρα*, καθώς στον ορισμό της συνολοσυνάρτησης υπάρχει η απαίτηση $\mu(A) \geq 0$, για κάθε $A \in \mathcal{A}$ (ενδεχομένως $+\infty$). Η θετικότητα του μέτρου δεν έπεται από τις ιδιότητες (i) και (ii), άρα πρέπει να επαληθεύεται ως ξεχωριστή ιδιότητα. Μάλιστα, μπορούν να οριστούν συνολοσυναρτήσεις που ικανοποιούν μόνο τις (i) και (ii) με τιμές στο $\mathbb{R}$, οι οποίες αναφέρονται ως *προσημασμένα μέτρα*.

Δίνουμε παρακάτω κάποια σημαντικά παραδείγματα μέτρων.

Παράδειγμα 1.6 (μέτρο Dirac): Έστω $(S, \mathcal{A})$ ένας μετρήσιμος χώρος. Η συνολοσυνάρτηση $\delta_x : \mathcal{A} \to [0, +\infty]$ με

$$\delta_x(A) = \mathbf{1}_A(x) = \begin{cases} 1 & \text{αν } x \in A, \\ 0 & \text{αν } x \in A^{\mathsf{c}}, \end{cases}$$

καλείται *μέτρο Dirac* στο $x \in S$. Το μέτρο αυτό μας "πληροφορεί" ποιά σύνολα της $\mathcal{A}$ περιέχουν το x και ποιά όχι. Η διαφορά της με μία δείκτρια συνάρτηση (που δεν είναι βέβαια μέτρο) είναι ότι η τελευταία είναι συνάρτηση του x, με σταθερό το σύνολο A, και μας υποδεικνύει ποιά σημεία είναι στο A και ποιά όχι (αλλαγή οπτικής γωνίας). Αποδεικνύουμε τώρα ότι είναι πράγματι μέτρο. Κατ'αρχάς $\delta_x(A) \geq 0$ για κάθε $A \in \mathcal{A}$. Επιπλέον,

-

(ii) Έστω (A_n) κατά ζεύγη ξένη ακολουθία στην $\mathcal{A}$ (για συντομία $(A_n)+$). Ισχύει ότι

$$\delta_x(\cup_n A_n) = \mathbf{1}_{\cup_n A_n}(x) \overset{(A_n)+}{=} \sum_n \mathbf{1}_{A_n}(x) = \sum_n \delta_x(A_n).$$

Παρατηρούμε επίσης βέβαια ότι $\delta_x(S) = 1$ και άρα κάθε μέτρο Dirac (για κάθε $x \in S$) είναι μέτρο πιθανότητας. Το μέτρο Dirac μπορεί να οριστεί σε όλο το δυναμοσύνολο του S χωρίς πρόβλημα. Στην περίπτωση τώρα που $(S, \mathcal{A}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, τότε το μέτρο Dirac στο $c \in \mathbb{R}$ αντιστοιχεί στην κατανομή των εκφυλισμένων στο c τυχαίων μεταβλητών, δηλ. των τυχαίων μεταβλητών που παίρνουν την τιμή c με πιθανότητα 1. Με τη βοήθεια των μέτρων Dirac μπορούμε να αναπαραστήσουμε και κάθε διακριτό μέτρο πιθανότητας στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ που όπως προαναφέραμε αντιστοιχούν στις κατανομές των διακριτών τυχαίων μεταβλητών. Πράγματι, αν S είναι το στήριγμα της συνάρτησης πιθανότητας f μιας διακριτής τυχαίας μεταβλητής X (το σύνολο των σημείων που αποδίδει θετική πιθανότητα), τότε η κατανομή $\mathbb{P}_X$ της X αναπαρίσταται ως:

$$\mathbb{P}_X = \sum_{x \in S} f(x)\delta_x,$$

δηλ. ως μία αριθμήσιμη μίξη (πεπερασμένη ή αριθμησίμως άπειρη) εκφυλισμένων κατανομών στα σημεία του στηρίγματος με μικτικά βάρη αυτά που αποδίδει η συνάρτηση πιθανότητας. Έτσι αν $A \in \mathcal{B}(\mathbb{R})$, έχουμε

$$\mathbb{P}_X(A) = \sum_{x \in S} f(x)\delta_x(A) = \sum_{x \in AS} f(x),$$

δηλ. η πιθανότητα που αποδίδεται στο A από την κατανομή της X, υπολογίζεται μέσω της σειράς (πεπερασμένης ή άπειρης) των πιθανοτήτων των σημείων του στηρίγματος που βρίσκονται εντός του A.

Παράδειγμα 1.7 (αριθμητικό μέτρο): Έστω $(S, \mathcal{A})$ μετρήσιμος χώρος. Η συνολοσυνάρτηση : $\nu\colon \mathcal{A} \to [0, +\infty]$ με

$$\nu(A) = |A| = \begin{cases} 0 & \text{αν } A = \emptyset, \\ n & \text{αν } A = \{x_1, x_2, \ldots, x_n\}, \\ +\infty & \text{αν } A \text{ απειροσύνολο}, \end{cases}$$

λέγεται *αριθμητικό μέτρο* στο S. Το μέτρο αυτό καταγράφει το πλήθος των στοιχείων (τον πληθάριθμο) κάθε μέλους της $\mathcal{A}$, όταν αυτό είναι πεπερασμένο. Σε κάθε απειροσύνολο αποδίδει την τιμή $+\infty$, χωρίς να ξεχωρίζει διαφορετικές τάξεις απείρου. Αποδεικνύουμε τώρα ότι είναι πράγματι μέτρο. Κατ'αρχάς $\nu(A) \geq 0$ για κάθε $A \in \mathcal{A}$. Επιπλέον,

-

(ii) Έστω (A_n) κατά ζεύγη ξένη ακολουθία στην $\mathcal{A}$. Μπορούμε να διακρίνουμε 3 περιπτώσεις:
1. A_n απειροσύνολο για τουλάχιστον ένα n,
2. A_n όλα πεπερασμένα και μη κενά για άπειρα n,
3. A_n όλα πεπερασμένα και μη κενά μόνο για πεπερασμένο πλήθος n.

Στις περιπτώσεις (1) και (2), η $\cup_n A_n$ είναι απειροσύνολο και η ιδιότητα (ii) ενός μέτρου ικανοποιείται. Για την περίπτωση (3), αν $I = \{n \in \mathbb{N} : A_n \neq \emptyset\}$, τότε το I είναι πεπερασμένο και άρα

$$\nu(\cup_n A_n) = \nu(\cup_{n\in I} A_n) = |\cup_{n\in I} A_n| \overset{(A_n)+}{=} \sum_{n\in I} |A_n| = \sum_{n\in I} \nu(A_n) \overset{\nu(\emptyset)=0}{=} \sum_n \nu(A_n).$$

Το αριθμητικό μέτρο μπορεί να είναι είτε πεπερασμένο, είτε σ–πεπερασμένο, είτε να μην είναι τίποτα από τα δύο. Παίρνουμε για ευκολία την περίπτωση που $\mathcal{A} = \mathcal{P}(S)$. Αν το S είναι πεπερασμένο, τότε $\nu(S) < +\infty$ και άρα το ν είναι πεπερασμένο μέτρο. Αν το S είναι αριθμησίμως άπειρο, τότε $S = \cup_n \{x_n\}$ και $\nu(\{x_n\}) = 1 < +\infty$ για κάθε n. Σε αυτήν την περίπτωση συμπεραίνουμε ότι το ν είναι σ–πεπερασμένο. Τέλος, θα δείξουμε ότι το ν δε μπορεί να είναι σ-πεπερασμένο, αν το S είναι υπεραριθμήσιμο. Ας υποθέσουμε λοιπόν ότι είναι σ-πεπερασμένο. Τότε, υπάρχει μία ακολουθία (A_n) υποσυνόλων του S με $S = \cup_n A_n$ και $\nu(A_n) < +\infty$ για κάθε n. Αυτό όμως σημαίνει ότι το S είναι αριθμήσιμη ένωση πεπερασμένων συνόλων και άρα αριθμήσιμο σύνολο, το οποίο είναι άτοπο, καθώς υποθέσαμε ότι είναι υπεραριθμήσιμο.

Το πιο σημαντικό συνεχές[6] μέτρο είναι το μέτρο Lebesgue και αναφέρουμε εδώ κάποια απαραίτητα στοιχεία, χωρίς να μπούμε σε τεχνικές λεπτομέρειες.

[6]η έννοια της συνέχειας για μέτρα επεκτείνεται και σε μέτρα που δεν είναι κατ'ανάγκη πιθανότητας.

Παράδειγμα 1.8 (μέτρο Lebesgue): Στο μετρήσιμο χώρο $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ μπορεί να κατασκευαστεί ένα μέτρο που επεκτείνει την έννοια του μήκους από τα διαστήματα σε όλα τα σύνολα Borel. Μία τέτοια επέκταση θα πρέπει να διατηρεί τα βασικά γεωμετρικά χαρακτηριστικά της έννοιας του μήκους. Το μήκος ενός συνόλου δεν θα πρέπει να μεταβάλλεται όταν μεταθέσουμε τα σημεία του κατά μία σταθερά ή όταν του εφαρμόσουμε μία ανάκλαση (σε μεγαλύτερες διαστάσεις εμπλέκονται και οι περιστροφές). Το αναλλοίωτο του μήκους ως προς αυτές τις ισομετρίες πρέπει να χαρακτηρίζει οποιαδήποτε επέκταση της έννοιας του μήκους από διαστήματα σε πιο πολύπλοκα υποσύνολα των πραγματικών αριθμών. Οι σκέψεις αυτές οδήγησαν στην κατασκευή του μέτρου Lebesgue, το οποίο περιορισμένο στα διαστήματα του $\mathbb{R}$ αποδίδει ως τιμή το μήκος τους και έχει ως χαρακτηριστική ιδιότητα το *αναλλοίωτο στις ισομετρίες* του $\mathbb{R}$, δηλαδή στους μετασχηματισμούς που διατηρούν τις αποστάσεις (βλ. κεφ. 2). Το πεδίο ορισμού του μέτρου Lebesgue μπορεί να είναι τα σύνολα Borel του $\mathbb{R}$, αλλά μπορεί να επεκταθεί μέχρι και τα λεγόμενα Lebesgue-μετρήσιμα υποσύνολα $\mathcal{M}_{\lambda^*}$ του $\mathbb{R}$. Η απόδειξη ύπαρξης μη Lebesgue-μετρησίμων συνόλων που γίνεται σε μαθήματα θεωρίας μέτρου οδηγεί στο συμπέρασμα ότι το μέτρο Lebesgue δεν μπορεί να επεκταθεί σε όλα τα υποσύνολα του $\mathbb{R}$. Στην ουσία βέβαια, η μη επεκτασιμότητά του είναι αποτέλεσμα της αδυναμίας που υπάρχει να διατηρηθεί η σ–-προσθετικότητα του μέτρου Lebesgue σε όλα τα υποσύνολα του $\mathbb{R}$. Όπως συνήθως συμβαίνει σε τέτοιες περιπτώσεις, αυτό οδήγησε τους μαθηματικούς σε πολύ ενδιαφέρουσες εναλλακτικές προσεγγίσεις. Σε μία από αυτές ο *Robert M. Solovay* κατασκεύασε ένα μοντέλο, γνωστό πλέον με το όνομα *μοντέλο του Solovay* [2], στο οποίο χωρίς το αξίωμα της επιλογής, αλλά με την ισχύ όλων των άλλων αξιωμάτων της *συνολοθεωρίας ZF (Zermelo-Fraenkel)* και μία ασθενέστερη έκδοση του αξιώματος της επιλογής έδειξε ότι όλα τα υποσύνολα του $\mathbb{R}$ μπορούν να θεωρηθούν Lebesgue–μετρήσιμα. Για μία σύντομη ιστορική αναδρομή και ισοδύναμες μορφές του αξιώματος της επιλογής, βλ. The Axiom of Choice.

Θα μπορούσε βέβαια κάποιος να αναρωτηθεί αν η σ-προσθετικότητα του μέτρου Lebesgue είναι αναγκαία. Συγκεκριμένα, ένα σ-προσθετικό μέτρο είναι και πεπερασμένα προσθετικό (δηλ. προσθετικό στις πεπερασμένες ξένες ενώσεις, βλ. Πρόταση 1.6), χωρίς βέβαια να ισχύει το αντίστροφο. Αν κρατήσουμε λοιπόν μόνο την πεπερασμένη προσθετικότητα, είναι δυνατόν να κατασκευαστεί ένα πεπερασμένα προσθετικό μέτρο Lebesgue ορισμένο σε όλα τα υποσύνολα του $\mathbb{R}$; Σε αυτήν την περίπτωση βέβαια θα πρέπει να εγκαταλείψουμε την απαίτηση να είναι σ–προσθετικό όπως αναφέρθηκε παραπάνω. Η απάντηση είναι καταφατική για το $\mathbb{R}$ και το $\mathbb{R}^2$, αλλά όχι γενικότερα για τον $\mathbb{R}^n$, για $n \geq 3$. Είναι διάσημο το *παράδοξο των Banach-Tarski* που έδειξαν ότι στον $\mathbb{R}^3$ μία σφαίρα μπορεί να αποδομηθεί σε ένα πεπερασμένο πλήθος κομματιών (τουλάχιστον 5 χρειάζονται για την κατασκευή) και στη συνέχεια με μεταθέσεις και περιστροφές να κατασκευαστούν 2 ακριβή αντίγραφα της αρχικής σφαίρας! Εφόσον οι μετασχηματισμοί που χρησιμοποιούνται είναι ισομετρίες, είναι βέβαια φανερό ότι το παράδοξο είναι συνέπεια του γεγονότος ότι σε κάποια από αυτά τα κομμάτια δεν μπορεί να αποδωθεί η έννοια του όγκου, και έτσι στον $\mathbb{R}^3$ δεν μπορεί να υπάρξει ούτε ένα πεπερασμένα προσθετικό μέτρο Lebesgue ορισμένο σε όλα τα υποσύνολα του $\mathbb{R}^3$.

Η πιο δημοφιλής μέθοδος κατασκευής του μέτρου Lebesgue είναι μέσω της έννοιας του εξωτερικού μέτρου και η γενικότερη ιδέα κατασκευής μέτρων με τη βοήθεια εξωτερικών μέτρων οφείλεται στον Κωνσταντίνο Καραθεοδωρή (1873-1950), ενός από τους θεμελιωτές της θεωρίας Μέτρου. Η ιδέα κατασκευής στηρίζεται σε θετικές συνολοσυναρτήσεις με πεδίο ορισμού όλα τα υποσύνολα του $\mathbb{R}$ που απαιτούν μόνο μονοτονία και σ-υποπροσθετικότητα, ιδιότητες δηλ. που ικανοποιούνται από ένα μέτρο (βλ. Πρόταση 1.6), χωρίς να απαιτείται η σ-προσθετικότητα. Αφού οριστούν κατάλληλα κατά περίπτωση, στη συνέχεια η συνολοσυνάρτηση περιορίζεται σε μία σ-άλγεβρα με τρόπο ώστε εκεί να ικανοποιούνται οι ιδιότητες του μέτρου. Με τον τρόπο αυτό, από ένα εξωτερικό μέτρο μπορούμε να κατασκευάσουμε ένα μέτρο στην ευρύτερη δυνατή οικογένεια υποσυνόλων του $\mathbb{R}$ (όπως βέβαια και σε άλλα αυθαίρετα σύνολα) όπου πλέον σε αυτήν την οικογένεια ικανοποιείται η ιδιότητα της σ–-προσθετικότητας.

Στην επόμενη πρόταση διατυπώνονται βασικές ιδιότητες ενός μέτρου, οι οποίες ικανοποιούνται βέβαια και από τα μέτρα πιθανότητας, και στη συνέχεια αποδεικνύονται για πληρότητα.

Πρόταση 1.6 (ιδιότητες μέτρου): Έστω $(S, \mathcal{A}, \mu)$ ένας χώρος μέτρου. Τότε

1. $\mu(A \cup B) = \mu(A) + \mu(B)$, για κάθε $A, B \in \mathcal{A}$ με $AB = \emptyset$ (πεπερασμένη προσθετικότητα).
2. $\mu(A) \leq \mu(B)$, για κάθε $A, B \in \mathcal{A}$ με $A \subset B$ (μονοτονία).
3. $\mu(\cup_n A_n) \leq \sum_n \mu(A_n)$, για κάθε ακολουθία (A_n) στην $\mathcal{A}$ (σ-υποπροσθετικότητα).
4. $\mu(\cup_n A_n) = \lim_n \mu(A_n)$, για κάθε αύξουσα ακολουθία (A_n) στην $\mathcal{A}$.
5. $\mu(\cap_n A_n) = \lim_n \mu(A_n)$, για κάθε φθίνουσα ακολουθία (A_n) στην $\mathcal{A}$ με $\mu(A_n) < \infty$ τελικά για κάθε n.

Απόδειξη:

1.
$$\mu(A \cup B) = \mu(A \cup B \cup \emptyset \cup \emptyset \ldots) = \mu(A) + \mu(B) + \mu(\emptyset) + \mu(\emptyset) + \ldots = \mu(A) + \mu(B).$$

2.
$$\mu(B) = \mu\big(A \cup (B \setminus A)\big) \overset{\text{ξένα}}{=} \mu(A) + \mu(B \setminus A) \geq \mu(A).$$

3. Έστω (A_n) τυχούσα ακολουθία στην $\mathcal{A}$. Μπορούμε να την αντικαταστήσουμε με μία κατά ζεύγη ξένη ακολουθία στην $\mathcal{A}$. Θέτουμε $B_1 = A_1$ και για κάθε

$n \geq 2$

$$B_n := A_n \setminus (\cup_{k=1}^{n-1} A_k). \tag{1.16}$$

Κάθε σύνολο B_n περιλαμβάνει τα στοιχεία του A_n που δεν συμπεριλαμβάνονται στην ένωση των προηγούμενων όρων A_k για $1 \leq k \leq n-1$. Ἀμεσα προκύπτουν οι εξής ιδιότητες για την καινούρια ακολουθία (B_n):

(α) Η (B_n) είναι κατά ζεύγη ξένη ακολουθία στην $\mathcal{A}$.

(β) $\cup_{k=1}^{n} B_k = \cup_{k=1}^{n} A_k$, για κάθε $n \geq 1$.

(γ) $\cup_{n=1}^{\infty} B_n = \cup_{n=1}^{\infty} A_n$.

Από τις παραπάνω ιδιότητες έχουμε

$$\mu(\cup_n A_n) \overset{(3)}{=} \mu(\cup_{n=1}^{\infty} B_n) \overset{(1)}{=} \sum_n \mu(B_n) \overset{B_n \subset A_n}{\leq} \sum_n \mu(A_n),$$

και έτσι καταλήγουμε στην σ–υποπροσθετικότητα.

4. Ἐστω (A_n) αύξουσα ακολουθία στην $\mathcal{A}$. Για (B_n) όπως ορίστηκε στην (1.16) έχουμε

$$\begin{aligned} \mu(\cup_n A_n) &\overset{(3)}{=} \mu(\cup_n B_n) \overset{(1)}{=} \sum_n \mu(B_n) = \lim_n \sum_{k=1}^{n} \mu(B_k) \\ &\overset{(1)}{=} \lim_n \mu(\cup_{k=1}^{n} B_k) \overset{(2)}{=} \lim_n \mu(\cup_{k=1}^{n} A_k) = \lim_n \mu(A_n), \end{aligned}$$

όπου η τελευταία ισότητα προκύπτει από την υπόθεση ότι η (A_n) είναι αύξουσα και άρα $\cup_{k=1}^{n} A_k = A_n$.

5. Έστω (A_n) φθίνουσα ακολουθία στην $\mathcal{A}$. Από υπόθεση υπάρχει $n_0 \in \mathbb{N}$ τέτοιο ώστε $\mu(A_{n_0}) < \infty$. Η ακολουθία $(A_{n_0} \setminus A_n)$ είναι αύξουσα ακολουθία στην $\mathcal{A}$. Από το (iv) έχουμε ότι

$$\mu\big(A_{n_0} \setminus (\cap_n A_n)\big) = \mu\big(\cup_n (A_{n_0} \setminus A_n)\big) = \lim_n \mu(A_{n_0} \setminus A_n) = \mu(A_{n_0}) - \lim_n \mu(A_n),$$

όπου η τελευταία ισότητα προκύπτει από το γεγονός ότι $A_n \subset A_{n_0}$ για κάθε $n \geq n_0$, καθώς η (A_n) έχει υποτεθεί φθίνουσα και $\mu(A_{n_0}) < \infty$. Για τον ίδιο λόγο έχουμε ότι το αριστερό μέλος της πρώτης ισότητας γράφεται $\mu(A_{n_0}) - \mu(\cap_n A_n)$ και έτσι καταλήγουμε στη ζητούμενη σχέση.

□

1.10 Πλήρωση σ-άλγεβρας ως προς μέτρο

Όταν δίνεται μία σ-άλγεβρα $\mathcal{A}$ και ένα μέτρο μ ορισμένο στην $\mathcal{A}$, τότε φυσιολογικά μπορούμε να προσδιορίσουμε τα στοιχεία της εκείνα που έχουν μηδενικό μέτρο. Όταν τα σύνολα αυτά εμπλέκονται σε πράξεις συνόλων, π.χ. ενώσεις, τομές, συμμετρική διαφορά, κ.τ.λ., δεν έχουν καμία επίδραση στον υπολογισμό του μέτρου των συνόλων που προκύπτουν. Είναι λοιπόν αμελητέα. Εκτός όμως από αυτό το χαρακτηριστικό τους, μπορούμε με τη βοήθειά τους να εμπλουτίσουμε την αρχική σ-άλγεβρα με πολλά ακόμα σύνολα, κάνοντας μία επέκταση του αρχικού μέτρου μ σε ένα μέτρο $\overline{\mu}$ που θα είναι πλέον ορισμένο σε μία μεγαλύτερη σ-άλγεβρα $\mathcal{A}_\mu$. Ο σημαντικότερος λόγος για τον οποίο θέλουμε να κάνουμε μία τέτοια επέκταση είναι γιατί μας ενδιαφέρει να ικανοποιούνται συγκεκριμένες ιδιότητες που χαρακτηρίζουν το μέτρο αυτό στην ευρύτερη δυνατή οικογένεια υποσυνόλων του χώρου που είναι ορισμένο. Για παράδειγμα, στην περίπτωση του μέτρου Lebesgue θέλουμε ένα μέτρο που να αποδίδει ως τιμή το μήκος στα υποσύνολα του $\mathbb{R}$ που είναι διαστήματα και να είναι αναλλοίωτο στις ισομετρίες του $\mathbb{R}$. Η πρώτη σκέψη λοιπόν είναι να φτιάξουμε ένα μέτρο στην ελάχιστη σ-άλγεβρα που περιέχει τα διαστήματα του $\mathbb{R}$ και αυτή είναι βέβαια η σ-άλγεβρα των συνόλων Borel. Αποδεικνύεται λοιπόν ότι το μέτρο Lebesgue λ έχει την επιθυμητή αυτή ιδιότητα στη συλλογή $\mathcal{B}(\mathbb{R})$ και έτσι όταν γράφουμε $\lambda(A)$ για $A \in \mathcal{B}(\mathbb{R})$ μπορούμε να το ερμηνεύουμε ως *μήκος* του A. Μπορούμε να πάμε παραπέρα; Μπορούμε να μεγαλώσουμε και άλλο αυτή τη συλλογή, ώστε να αποδώσουμε "μήκος" και σε άλλα υποσύνολα του $\mathbb{R}$; Σε αυτήν την ενότητα λοιπόν, θα δούμε πώς μπορούμε γενικά, για οποιοδήποτε μέτρο και όχι μόνο για το μέτρο Lebesgue, να το επεκτείνουμε φυσιολογικά, φτιάχνοντας πολλά περισσότερα σύνολα μηδενικού μέτρου, από αυτά που αρχικά έχει στη σ-άλγεβρα $\mathcal{A}$. Αυτό θα το λέμε *πλήρωση* της σ-άλγεβρας ως προς το μέτρο, και το μέτρο που θα προκύπτει καλείται *πλήρες*.

Ορισμός 1.27: Έστω $(S, \mathcal{A}, \mu)$ ένας χώρος μέτρου. Θέτουμε

$$\mathcal{A}_\mu^0 := \{N \subset S \mid \text{υπάρχει } A \in \mathcal{A} \text{ με } N \subset A \text{ και } \mu(A) = 0\}.$$

Αν $N \in \mathcal{A}_\mu^0$, τότε το N καλείται *μ-μηδενικό*. Θέτουμε επίσης

$$\mathcal{A}_\mu := \{B \subset S \mid B = A \cup N \text{ με } A \in \mathcal{A} \text{ και } N \in \mathcal{A}_\mu^0\}.$$

Η $\mathcal{A}_\mu$ καλείται *πλήρωση της σ-άλγεβρας $\mathcal{A}$ ως προς μ*. Επιπλέον, το μέτρο μ καλείται *πλήρες μέτρο*, αν $\mathcal{A}_\mu^0 \subset \mathcal{A}$ και ο αντίστοιχος χώρος μέτρου $(S, \mathcal{A}, \mu)$ καλείται *πλήρης χώρος μέτρου*.

Στην επόμενη πρόταση δίνονται οι βασικότερες ιδιότητες που συνδέονται με την πλήρωση ενός χώρου μέτρου.

Πρόταση 1.7: Έστω $(S, \mathcal{A}, \mu)$ ένας χώρος μέτρου. Για κάθε $B \in \mathcal{A}_\mu$, θέτουμε

$$\overline{\mu}(B) = \mu(A), \quad \text{αν } B = A \cup N, \text{ με } A \in \mathcal{A} \text{ και } N \in \mathcal{A}_\mu^0.$$

Τότε,

1. η $\mathcal{A}_\mu$ είναι σ-άλγεβρα με $\mathcal{A} \subset \mathcal{A}_\mu$,
2. η συνολοσυνάρτηση $\overline{\mu}$ είναι πλήρες μέτρο στο μετρήσιμο χώρο $(S, \mathcal{A}_\mu)$.

3. το μέτρο μ έχει μοναδική επέκταση από την $\mathcal{A}$ στην $\mathcal{A}_\mu$ το μέτρο $\overline{\mu}$,
Ο χώρος μέτρου $(S, \mathcal{A}_\mu, \overline{\mu})$ καλείται η *πλήρωση* του χώρου μέτρου $(S, \mathcal{A}, \mu)$.

Πόρισμα 1.3: Σε ένα μετρήσιμο χώρο $(S, \mathcal{A})$ ένα μέτρο μ είναι πλήρες, αν και μόνο αν $\mathcal{A}_\mu = \mathcal{A}$.

1.11 Χώροι Γινόμενο

Σε αυτήν την ενότητα θα αναφερθούμε σε χώρους, τοπολογικούς ή μετρήσιμους, που παράγονται "φυσιολογικά" από ένα πεπερασμένο πλήθος τοπολογικών ή μετρήσιμων χώρων μέσα από το καρτεσιανό τους γινόμενο. Οι έννοιες και τα αποτελέσματα θα επεκταθούν αργότερα και σε απειροδιάστατους χώρους γινόμενο, είτε αυτό αφορά ένα αριθμήσιμο, είτε ένα υπεραριθμήσιμο πλήθος. Σε ότι αφορά το συμβολισμό, για δύο σύνολα A, B το καρτεσιανό τους γινόμενο θα συμβολίζεται με $A \times B$, ενώ για δύο συλλογές συνόλων $\mathcal{A}$ και $\mathcal{B}$, για να μην προκληθεί σύγχυση με το καρτεσιανό γινόμενο $\mathcal{A} \times \mathcal{B}$ θα συμβολίζουμε με το συλλογή που αποτελείται από όλα τα καρτεσιανά γινόμενα στοιχείων της $\mathcal{A}$ με στοιχεία της $\mathcal{B}$. Όταν $\mathcal{A} = \mathcal{P}(R)$ και $\mathcal{B} = \mathcal{P}(S)$, δηλαδή τα $\mathcal{A}$ και $\mathcal{B}$ αντιστοιχούν στα δυναμοσύνολα δύο συνόλων R και S αντίστοιχα, τότε το $\mathcal{A} \boxtimes \mathcal{B}$ αντιστοιχεί στα λεγόμενα *ορθογώνια υποσύνολα* του $R \times S$. Για n σύνολα $\{A_i\}_{i=1}^n$ θα γράφουμε $\prod_{i=1}^n A_i$ για το καρτεσιανό τους γινόμενο και A^n, αν όλα είναι ίσα με A. Για n συλλογές $\{\mathcal{A}_i\}_{i=1}^n$ θα γράφουμε $\boxtimes_{i=1}^n \mathcal{A}_i$ για τη συλλογή που έχει ως στοιχεία όλα τα καρτεσιανά γινόμενα $\prod_{i=1}^n A_i$ με $A_i \in \mathcal{A}_i$ και θα γράφουμε $\mathcal{A}^{\boxtimes n}$, αν $\mathcal{A}_i = \mathcal{A}$ για κάθε $1 \leq i \leq n$.

Ορισμός 1.28: Έστω $(S_i, \mathcal{T}_i)_{1 \leq i \leq n}$ ένα πεπερασμένο πλήθος τοπολογικών χώρων και $S := \prod_{i=1}^n S_i$ το καρτεσιανό τους γινόμενο. Αν $\boxtimes_{i=1}^n \mathcal{T}_i$ είναι η οικογένεια των ανοικτών ορθογωνίων (open rectangles) του S, τότε η τοπολογία $\mathcal{T}^{pr} := \tau\left(\boxtimes_{i=1}^n \mathcal{T}_i\right)$ που παράγεται από τα ανοικτά ορθογώνια του S καλείται *τοπολογία γινόμενο* ή *καρτεσιανή τοπολογία* του S, και ο $(S, \mathcal{T}^{pr})$ καλείται *(τοπολογικός) χώρος γινόμενο.*

Παρατήρηση 1.37. Η οικογένεια των ανοικτών ορθογωνίων $\mathcal{T}^{\boxtimes n}$ είναι βάση κάποιας τοπολογίας αφού είναι κλειστή στις πεπερασμένες τομές και καλύπτει τον S^n (βλ. θεώρημα 1.1 και παρατήρηση 1.23; ασκ. **??**). Αυτό δίνει άμεσα μία περιγραφή των μελών της τοπολογίας γινόμενο $\mathcal{T}^{pr}$ ως ενώσεων (οποιουδήποτε πλήθους) ανοικτών ορθογωνίων.

Έναν αντίστοιχο ορισμό μπορούμε να δώσουμε για την σ-άλγεβρα γινόμενο.

Ορισμός 1.29: Έστω $(S_i, \mathcal{A}_i)_{1 \leq i \leq n}$ ένα πεπερασμένο πλήθος μετρήσιμων χώρων και $S := \prod_{i=1}^n S_i$ το καρτεσιανό τους γινόμενο. Αν $\boxtimes_{i=1}^n \mathcal{A}_i$ είναι η οικογένεια των μετρήσιμων ορθογωνίων (measurable rectangles) του S, τότε η σ-άλγεβρα

$\bigotimes_{i=1}^{n}\mathcal{A}_i := \sigma\left(\boxtimes_{i=1}^{n}\mathcal{A}_i\right)$ που παράγεται από τα μετρήσιμα ορθογώνια του S καλείται *σ-άλγεβρα γινόμενο* και ο $(S, \bigotimes_{i=1}^{n}\mathcal{A}_i)$ καλείται *(μετρήσιμος) χώρος γινόμενο*. Η σ-άλγεβρα γινόμενο των $\{\mathcal{A}_i\}_{i=1}^{n}$ θα συμβολίζεται και με $\mathcal{A}^{pr}$. Αν $S_i = S$ και $\mathcal{A}_i = \mathcal{A}$ για κάθε $1 \le i \le n$, τότε θα γράφουμε και $(S^n, \mathcal{A}^{\bigotimes n})$ για το χώρο γινόμενο.

Ιδιαίτερο ενδιαφέρον έχει η σύνδεση της σ-άλγεβρας των συνόλων Borel ενός χώρου γινόμενο με τη σ-άλγεβρα γινόμενο των συνιστώντων χώρων Borel. Η περίπτωση που θα μας απασχολήσει είναι όταν έχουμε δυνάμεις ενός τοπολογικού χώρου.

Πρόταση 1.8: Έστω $(S, \mathcal{T})$ ένας τοπολογικός χώρος και $(S^n, \mathcal{T}^{pr})$ ο αντίστοιχος χώρος γινόμενο. Αν θέσουμε $\mathcal{B}(S^n) = \sigma(\mathcal{T}^{pr})$ και $\mathcal{B}^n(S) := \mathcal{B}(S)^{\bigotimes n}$ την αντίστοιχη σ-άλγεβρα γινόμενο, τότε $\mathcal{B}^n(S) \subset \mathcal{B}(S^n)$. Αν ο τοπολογικός χώρος $\mathcal{T}$ είναι *δεύτερος αριθμήσιμος*, δηλ. έχει αριθμήσιμη βάση, τότε $\mathcal{B}^n(S) = \mathcal{B}(S^n)$.

Πόρισμα 1.4: $\mathcal{B}(\mathbb{R}^n) = \mathcal{B}^n(\mathbb{R})$.

Απόδειξη: Κατ'αρχάς η τοπολογία γινόμενο του $\mathbb{R}^n$ συμπίπτει με την τοπολογία της ευκλείδιας μετρικής, αφού οι μετρικές που τους αντιστοιχούν είναι ισοδύναμες (γιατί;). Επιπλέον, ο $\mathbb{R}^n$ είναι διαχωρίσιμος, π.χ. από το $\mathbb{Q}^n$, και άρα έχει και αριθμήσιμη βάση (π.χ. τις ανοικτές μπάλες με κέντρα ρητών συντεταγμένων και με ακτίνες ρητού μήκους). Τελικά, από την Πρόταση 1.8 συμπεραίνουμε το ζητούμενο. □

1.12 Συνέχεια και Μετρησιμότητα συναρτήσεων

Σε αυτήν την ενότητα θα αναφέρουμε πολύ σύντομα τις πιο σημαντικές συναρτήσεις μεταξύ τοπολογικών και μετρήσιμων χώρων, δηλ. τις συνεχείς και τις μετρήσιμες αντίστοιχα. Ιδιαίτερη βάση δίνουμε στις Borel μετρήσιμες συναρτήσεις καθώς αποτελούν την απαραίτητη εξειδίκευση που χρειαζόμαστε όταν αναφερόμαστε σε τυχαίες μεταβλητές.

Ορισμός 1.30: Έστω $(R, \mathcal{R})$ και $(S, \mathcal{S})$ δύο τοπολογικοί χώροι. Μία συνάρτηση $f\colon R \to S$ καλείται *συνεχής*, αν $f^{-1}(V) \in \mathcal{R}$ για κάθε $V \in \mathcal{S}$ ή ισοδύναμα, αν $f^{-1}(\mathcal{S}) \subset \mathcal{R}$, όπου $f^{-1}(\mathcal{S}) := \{f^{-1}(S) \mid S \in \mathcal{S}\}$, δηλ. η συλλογή των αντίστροφων εικόνων των στοιχείων της $\mathcal{S}$ μέσω της f.

Ο ορισμός είναι αντίστοιχος για τις μετρήσιμες συναρτήσεις αντικαθιστώντας τις τοπολογίες με τις σ-άλγεβρες.

Ορισμός 1.31: Έστω $(R, \mathcal{A})$ και $(S, \mathcal{B})$ δύο μετρήσιμοι χώροι. Μία συνάρτηση $f\colon R \to S$ καλείται

1. *μετρήσιμη*, αν $f^{-1}(B) \in \mathcal{A}$ για κάθε $B \in \mathcal{B}$ ή ισοδύναμα, αν $f^{-1}(\mathcal{B}) \subset \mathcal{A}$,
2. *Borel–μετρήσιμη* ή απλά *Borel*, αν ο S είναι τοπολογικός χώρος και $\mathcal{B} =$

$\mathcal{B}(S)$, δηλαδή η σ-άλγεβρα αποτελείται από τα σύνολα Borel.

Παρατήρηση 1.38. Οι παραπάνω ορισμοί δείχνουν την απλότητα με την οποία ορίζονται οι έννοιες της συνέχειας και της μετρησιμότητας. Μία συνάρτηση είναι συνεχής αν "επιστρέφει" ανοικτά σε ανοικτά και είναι μετρήσιμη αν "επιστρέφει" μετρήσιμα σε μετρήσιμα σύνολα.

Όπως θα δούμε αμέσως τώρα, αν περιορίσουμε μία συνεχή ή μία μετρήσιμη συνάρτηση σε κάποιο υποσύνολο του αρχικού χώρου, τότε η συνάρτηση παραμένει συνεχής ή μετρήσιμη αντίστοιχα στον υπόχωρο που προκύπτει. Θυμίζουμε ότι ως υπόχωρο, θεωρούμε το υποσύνολο που περιορίζεται η συνάρτηση εφοδιασμένο με την τοπολογία ίχνος ή τη σ-άλγεβρα ίχνος αντίστοιχα.

Πρόταση 1.9: Έστω $f: R \to S$ μία συνάρτηση.

- Αν $(R, \mathcal{T})$ και $(S, \mathcal{T}^*)$ είναι τοπολογικοί χώροι, η f είναι συνεχής και $A \subset R$, τότε ο περιορισμός $f_A := f|A$ της f στο A είναι συνεχής συνάρτηση.
- Αν $(R, \mathcal{A})$ και $(S, \mathcal{A}^*)$ είναι μετρήσιμοι χώροι, η f είναι μετρήσιμη και $A \subset R$, τότε ο περιορισμός $f_A := f|A$ της f στο A είναι μετρήσιμη συνάρτηση.

Απόδειξη: Ας αποδείξουμε πρώτα την πρόταση για τις συνεχείς. Αρκεί να αποδείξουμε ότι για κάθε $B \in \mathcal{T}^*$, η αντίστροφη εικόνα $f_A^{-1}(B) \in \mathcal{T}_A$. Έστω λοιπόν $B \in \mathcal{T}^*$. Έχουμε

$$f_A^{-1}(B) = \{x \in A \mid f(x) \in B\} = A \cap f^{-1}(B) \in \mathcal{T}_A,$$

αφού από τη συνέχεια της f προκύπτει ότι $f^{-1}(B) \in \mathcal{T}$. Αν αντικαταστήσουμε τους παραπάνω χώρους με μετρήσιμους και τη συνέχεια με μετρησιμότητα, τότε προκύπτει το ζητούμενο και για τις μετρήσιμες συναρτήσεις. □

Η απλούστερη πράξη που μπορούμε να ορίσουμε μεταξύ συναρτήσεων ορισμένων σε τοπολογικούς ή μετρήσιμους χώρους είναι η σύνθεση και είναι άμεσο ότι η προκύπτουσα συνάρτηση διατηρεί τη συνέχεια ή τη μετρησιμότητα αντίστοιχα.

Ορισμός 1.32: Έστω R, S, T σύνολα και $f: R \to S$, $g: S \to T$ συναρτήσεις. Υπενθυμίζουμε ότι ως σύνθεση της f με τη g, ορίζουμε τη συνάρτηση $g \circ f: R \to T$ με $g \circ f(x) = g(f(x))$.

1. Αν $(R, \mathcal{R})$, $(S, \mathcal{S})$ και $(T, \mathcal{T})$ είναι τοπολογικοί χώροι και f, g συνεχείς, τότε η $g \circ f$ είναι συνεχής.
2. Αν $(R, \mathcal{A})$, $(S, \mathcal{B})$ και $(T, \mathcal{C})$ είναι μετρήσιμοι χώροι και f, g είναι μετρήσιμες, τότε η $g \circ f$ είναι μετρήσιμη.

Θα ήταν επιθυμητό βέβαια να εξασφαλίζουμε τη συνέχεια ή τη μετρησιμότητα ελέγχοντας τη συνθήκη συνέχειας ή μετρησιμότητας αντίστοιχα σε μία μικρότερη οικογένεια υποσυνόλων. Έτσι φυσιολογικά οδηγούμαστε στον έλεγχο πάνω σε γεννήτορες.

Πρόταση 1.10:
1. Έστω $\mathcal{C}$ ένας γεννήτορας (υποβάση) μιας τοπολογίας $\mathcal{S}$ σε κάποιο σύνολο S και $(R, \mathcal{R})$ ένας άλλος τοπολογικός χώρος. Μία συνάρτηση $f\colon R \to S$ είναι συνεχής, αν και μόνο αν $f^{-1}(\mathcal{C}) \subset \mathcal{R}$.
2. Έστω $\mathcal{C}$ ένας γεννήτορας μιας σ-άλγεβρας $\mathcal{B}$ σε κάποιο σύνολο S και $(R, \mathcal{A})$ ένας άλλος μετρήσιμος χώρος. Μία συνάρτηση $f\colon R \to S$ είναι μετρήσιμη, αν και μόνο αν $f^{-1}(\mathcal{C}) \subset \mathcal{A}$.

Η παραπάνω πρόταση μας οδηγεί άμεσα στο εξής αποτέλεσμα.

Πόρισμα 1.5: Κάθε συνεχής συνάρτηση είναι Borel.

Απόδειξη: Έστω $f\colon R \to S$ συνεχής. Τα ανοικτά σύνολα του S είναι γεννήτορας των συνόλων Borel του S. Η f ως συνεχής επιστρέφει τα ανοικτά του S σε ανοικτά του R και άρα σε σύνολα Borel του R. Σύμφωνα με την προηγούμενη πρόταση συμπεραίνουμε ότι η f είναι Borel. □

Καθώς τα σύνολα Borel παράγονται από τα ανοικτά σύνολα θα περίμενε ίσως κάποιος οι Borel μετρήσιμες συναρτήσεις, ας τις συμβολίσουμε με $\mathcal{B}(R, S)$, να "παράγονται" και αυτές από τις συνεχείς $\mathcal{C}(R, S)$. Η παραγωγή μιας μεγαλύτερης οικογένειας συναρτήσεων από μία υπάρχουσα οικογένεια συναρτήσεων γίνεται φυσιολογικά παίρνοντας το κατά σημείο όριο ακολουθιών συναρτήσεων της αρχικής οικογένειας και η διαδικασία αυτή συνεχίζεται "επαγωγικά" (με υπερπεπερασμένη επαγωγή) μέχρι να επέλθει κλειστότητα στο κατά σημείο όριο ακολουθιών συναρτήσεων. Το σχήμα αυτό "δουλεύει" για την παραγωγή των Borel συναρτήσεων από τις συνεχείς σε πολλές σημαντικές περιπτώσεις, αλλά όχι πάντα και αφήνουμε το σημείο αυτό για αργότερα.

Ιδιαίτερο ενδιαφέρον έχουν οι 1-1 αντιστοιχίες (συναρτήσεις 1-1 και επί) μεταξύ δύο συνόλων R και S, τα οποία είναι εφοδιασμένα με κάποια από τις δύο παραπάνω δομές. Τέτοιες είδους μετασχηματισμοί μπορούν να βρεθούν μόνο σε ισοπληθικά σύνολα και καθώς είναι εφοδιασμένα με δομή αναζητούμε ιδιαιτέρως εκείνους που κάνουν τις δομές αυτές συμβατές, τους λεγόμενους ισομορφισμούς. Οι ισομορφισμοί μας δίνουν τη δυνατότητα να ταυτίζουμε χώρους που φαινομενικά είναι διαφορετικοί, αλλά συμπεριφέρονται με τον ίδιο τρόπο ως προς τη δομή που μελετάμε.

Ορισμός 1.33: Δύο σύνολα R, S λέγονται *ισοπληθικά* αν υπάρχει μία συνάρτηση $f : R \to S$ που να είναι 1-1 και επί. Σε αυτήν την περίπτωση ορίζεται μία μοναδική αντίστροφη $f^{-1} : S \to R$ με $f^{-1}(y) = x$ αν και μόνο αν $f(x) = y$. Ισοδύναμα, έχουμε $f^{-1} \circ f = \mathrm{id}_R$ και $f \circ f^{-1} = \mathrm{id}_S$.
1. Έστω $(R, \mathcal{R})$ και $(S, \mathcal{S})$ δύο τοπολογικοί χώροι. Αν υπάρχει μία 1-1 και επί συνάρτηση $f\colon R \to S$, τέτοια ώστε f και f^{-1} να είναι συνεχείς, τότε η f καλείται *ομοιομορφισμός* ή *τοπολογικός ισομορφισμός* και οι χώροι $(R, \mathcal{R})$ και $(S, \mathcal{S})$ *ομοιομορφικοί* ή *τοπολογικά ισόμορφοι*.
2. Έστω $(R, \mathcal{A})$ και $(S, \mathcal{B})$ δύο μετρήσιμοι χώροι. Αν υπάρχει μία 1-1 και επί

συνάρτηση $f\colon R \to S$, τέτοια ώστε f και f^{-1} να είναι μετρήσιμες, τότε η f καλείται *μετρήσιμος ισομορφισμός* και οι χώροι $(R, \mathcal{R})$ και $(S, \mathcal{S})$ *μετρήσιμα ισόμορφοι (ισομορφικοί)*. Αν οι χώροι είναι επιπλέον τοπολογικοί και οι σ-άλγεβρες $\mathcal{A}$ και $\mathcal{B}$ είναι οι σ-άλγεβρες Borel που τους αντιστοιχούν, τότε λέγονται *Borel ισομορφικοί*.

Παρόλο που οι ιδιότητες 1-1 και επί μεταφέρονται στην αντίστροφη, δε συμβαίνει το ίδιο "αυτόματα" στη συνέχεια, αλλά ούτε και στη μετρησιμότητα. Αυτός είναι και ο λόγος που απαιτούμε στον παραπάνω ορισμό να έχει και η f^{-1} την αντίστοιχη ιδιότητα. Εξαίρεση στο παραπάνω αποτελούν οι τυπικοί χώροι Borel, όπως προκύπτει άμεσα ως συνέπεια ενός θεωρήματος του Souslin. Εδώ παραθέτουμε τη συνέπεια αυτή ως θεώρημα.

Θεώρημα 1.2: Αν $f\colon (R, \mathcal{B}(R)) \to (S, \mathcal{B}(S))$ είναι μία μετρήσιμη 1-1 και επί συνάρτηση μεταξύ δύο τυπικών χώρων Borel, τότε και η αντίστροφη f^{-1} είναι μετρήσιμη.

Είναι μάλιστα εφικτό μέσω των Borel ισομορφισμών να γίνει μία ιδιαίτερα απλή ταξινόμηση των τυπικών χώρων Borel.

Θεώρημα 1.3 (Kuratowski): Αν $(R, \mathcal{B}(R))$ είναι ένας τυπικός χώρος Borel (με Πολωνική τοπολογία), τότε είναι Borel ισομορφικός με τον $(S, \mathcal{B}(S))$, όπου

$$S = \begin{cases} \{1, 2, \ldots, n\}, & \text{αν } |S| = n, \\ \mathbb{N}, & \text{αν το } S \text{ είναι άπειρο αριθμήσιμο,} \\ \mathbb{R}, & \text{αν το } S \text{ είναι υπεραριθμήσιμο} \end{cases}$$

και τα σύνολα Borel $\mathcal{B}(S)$ αντιστοιχούν σε αυτά που παράγονται από τη διακριτή τοπολογία (όλα τα υποσύνολα του S) αν το S είναι πεπερασμένο ή αριθμησίμως άπειρο, και από τη συνήθη τοπολογία του $\mathbb{R}$ (της απόλυτης τιμής) αν το S είναι υπεραριθμήσιμο.

1.13 Ιστορικά Στοιχεία - Περαιτέρω Μελέτη

Το κεφάλαιο αυτό είναι εισαγωγικό. Στην ελληνική βιβλιογραφία, για βασικές Πιθανότητες ο αναγνώστης παραπέμπεται στα συγγράμματα των [3] και [4]. Στοιχεία θεωρητικών πιθανοτήτων καλύπτονται στις ηλεκτρονικές σημειώσεις του [5] και στις Σημειώσεις Παραδόσεων του [6]. Ένα ολοκληρωμένο σύγγραμμα Θεωρίας Μέτρου είναι το βιβλίο των [7]. Στη ξενόγλωσση βιβλιογραφία υπάρχουν πολλά καλά συγγράμματα. Εξαιρετικό και πιο σύγχρονο είναι του [8], ενώ ένα πιο προχωρημένο βιβλίο που αποτελεί σημείο αναφοράς είναι του [9].

Ασκήσεις

1.13.1 Να υπολογιστούν η μέση τιμή και η διασπορά των βασικών διακριτών και συνεχών κατανομών που εμφανίζονται στους Πίνακες 1.3 και 1.4.

1.13.2 Έστω $(X_k)_{k\geq 1}$ μία ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με $X_k \sim \mathcal{B}e(p)$, όπου $0 < p < 1$. Η ακολουθία αυτή καλείται *διαδικασία Bernoulli* (*Bernoulli process*) και τυποποιεί μαθηματικά αυτό που στις Πιθανότητες αναφέρεται ως ακολουθία ανεξάρτητων δοκιμών Bernoulli με σταθερή πιθανότητα επιτυχίας p. Με τη βοήθεια της διαδικασίας Bernoulli μπορούμε να ορίσουμε αυστηρά τις βασικές διακριτές κατανομές που αναφέρονται στον Πίνακα 1.3 (για τη σύνδεση με την Poisson θα αναφερθούμε στο Κεφάλαιο...). Για $n, i, r \geq 1$, ορίζουμε τ.μ. Y_n, Z_i, W_r, W_r^* ως εξής:

$$\begin{aligned}
Y_n &= \#\ \text{επιτυχιών σε } n \text{ δοκιμές},\\
Z_1 &= \#\ \text{αποτυχιών μέχρι την 1η επιτυχία},\\
Z_i &= \#\ \text{αποτυχιών μεταξύ της } i-1 \text{ και της } i \text{ επιτυχίας},\ \ i \geq 2,\\
W_r &= \#\ \text{αποτυχιών μέχρι την } r \text{ επιτυχία},\\
W_r^* &= \#\ \text{δοκιμών μέχρι την } r \text{ επιτυχία}.
\end{aligned}$$

Είναι φανερό ότι μπορούμε να ορίσουμε την W_1^* στη μορφή

$$W_1^* = \inf\{k \geq 1 : X_k = 1\}$$

και επίσης ότι $Z_1 = W_1 = W_1^* - 1$. Ορίζουμε επαγωγικά την ακολουθία (W_r^*), αν θέσουμε

$$W_r^* = \inf\{k > W_{r-1}^* : X_k = 1\}, \qquad r \geq 2.$$

Ως κατανομές $\mathcal{B}in(n,p)$, $\mathcal{G}e(p)$ και $\mathcal{NB}(r,p)$ μπορούμε να ορίσουμε τις κατανομές των τ.μ. Y_n, Z_1 και W_r αντίστοιχα.

1. Να επαληθεύσετε ότι οι τ.μ. Y_n, Z_1 και W_r, έτσι όπως ορίστηκαν, έχουν συνάρτηση πιθανότητας αυτήν που αντιστοιχεί στον Πίνακα 1.3.

2. Να βρεθεί σχέση που συνδέει την τ.μ. Y_n με τις τ.μ. $(X_k)_{1\leq k\leq n}$ και να υπολογιστεί η μέση τιμή και η διασπορά της Διωνυμικής $\mathcal{B}in(n,p)$ με τη βοήθεια αυτής της σχέσης.

3. Να δειχθεί ότι για κάθε $(z_1, z_2, \ldots, z_r) \in \mathbb{N}^r$ η πιθανότητα

$$\mathbb{P}(Z_1 = z_1, Z_2 = z_2, \ldots, Z_r = z_r) = p^r(1-p)^{z_1+z_2+\ldots+z_r}$$

   και να αιτιολογηθεί από το παραπάνω αποτέλεσμα ότι οι τ.μ. $\{Z_i\}_{i=1}^n$ είναι ανεξάρτητες και ακολουθούν $\mathcal{G}e(p)$.

4. Να βρεθεί σχέση που συνδέει την τ.μ. W_r με τις τ.μ. $(Z_i)_{1\leq i\leq r}$ και να υπολογιστεί η μέση τιμή και η διασπορά της κατανομής Pascal $\mathcal{NB}(r,p)$ με τη βοήθεια αυτής της σχέσης, θεωρώντας γνωστές τη μέση τιμή και τη διασπορά της $\mathcal{G}e(p)$.

5. Να βρεθεί η σχέση που συνδέει τις W_r^* και W_r και με τη βοήθεια αυτής και της συνάρτησης πιθανότητας της $\mathcal{NB}(r,p)$ να βρεθεί η συνάρτηση πιθανότητας της W_r^*. Η κατανομή της W_r^* αναφέρεται επίσης ως κατανομή *Pascal* (ή *Αρνητική Διωνυμική*) με στήριγμα το $\{r, r+1, \ldots\}$.

1.13.3 Έστω $(X_k)_{k\geq 1}$ μία διαδικασία Bernoulli με $X_k \sim \mathcal{B}\boldsymbol{e}(p)$, όπου $0 < p < 1$. Σύμφωνα με τα θεωρήματα ολικής μέσης τιμής και διασποράς, για δύο τ.μ. Y, Z ισχύει ότι

$$\begin{aligned}\mathbb{E}(Z) &= \mathbb{E}\big[\,\mathbb{E}(Z|Y)\big]\\ \mathbb{V}(Z) &= \mathbb{V}\big[\,\mathbb{E}(Z|Y)\big] + \mathbb{E}\big[\,\mathbb{V}(Z|Y)\big].\end{aligned}$$

Θέτουμε Z την τ.μ. που εκφράζει το πλήθος των αποτυχιών μέχρι την 1η επιτυχία στην παραπάνω διαδικασία $\mathcal{B}\boldsymbol{e}(p)$ και όπως γνωρίζουμε $Z \sim \mathcal{G}\boldsymbol{e}(p)$.

1. Να υπολογιστεί η $\mathbb{E}(Z)$ και η $\mathbb{V}(Z)$ χρησιμοποιώντας τα παραπάνω θεωρήματα [υπόδειξη: να τεθεί η $Y = X_1$].
2. Θέτουμε $(Y_k)_{k\geq 1}$, όπου $Y_k = 1 - X_k,\ k \geq 1$. Να δείξετε ότι η $(Y_k)_{k\geq 1}$ είναι διαδικασία $\mathcal{B}\boldsymbol{e}(q)$, όπου $q = 1 - p$ και ότι

$$Z = \sum_{n=1}^{\infty} \prod_{k=1}^{n} Y_k.$$

   Να υπολογιστεί η $\mathbb{E}(Z)$ και η $\mathbb{V}(Z)$ χρησιμοποιώντας την παραπάνω σχέση.
3. Να επαληθευτούν οι ιδιότητες της συνάρτησης κατανομής για την F_X.

1.13.4 Να βρεθεί διακριτή συνάρτηση κατανομής με την ιδιότητα:

$$0 < F(x) < 1, \quad \forall x \in \mathbb{R}.$$

1.13.5 Έστω S σύνολο και

$$\begin{aligned}\mathcal{A}_1 &= \{A \subset S : A \ \text{ ή } \ A^c \ \text{ πεπερασμένο}\} \cup \{\emptyset\},\\ \mathcal{A}_2 &= \{A \subset S : A \ \text{ ή } \ A^c \ \text{ αριθμήσιμο}\} \cup \{\emptyset\}.\end{aligned}$$

Να εξεταστεί αν οι $\mathcal{A}_1$, $\mathcal{A}_2$ είναι σ-άλγεβρες στο S.

1.13.6 Για μία τ.μ. $X\colon \Omega \to \mathbb{R}$ να δειχθεί ότι η συνολοσυνάρτηση $\mathbb{P}_X\colon \mathcal{B}(\mathbb{R}) \to [0,1]$ που ορίζεται από τη σχέση $\mathbb{P}_X(B) = \mathbb{P}(X \in B)$ είναι μέτρο πιθανότητας.

1.13.7 Στην Παρατήρηση **1.13.7** να δειχθεί ότι η F_μ είναι συνάρτηση κατανομής.

1.13.8 Έστω S σύνολο και

$$\begin{aligned}\mathcal{T}_1 &= \{A \subset S : A^c \ \text{ πεπερασμένο}\} \cup \{\emptyset\},\\ \mathcal{T}_2 &= \{A \subset S : A^c \ \text{ αριθμήσιμο}\} \cup \{\emptyset\}.\end{aligned}$$

Να δειχθεί ότι οι $\mathcal{T}_1$ και $\mathcal{T}_2$ είναι τοπολογίες στο S.

Σχόλιο: Η $\mathcal{T}_1$ καλείται συμπεπερασμένη τοπολογία και η $\mathcal{T}_2$ συναριθμήσιμη τοπολογία.

1.13.9 Στο Παράδειγμα 1.4 να δειχθεί ότι η $\mathcal{T}_d$ είναι τοπολογία.

1.13.10 Έστω S σύνολο. Στο Παράδειγμα 1.4 εισάγεται η έννοια της ψευδομετρικής. Θα δούμε μία εφαρμογή της στην έννοια της ισονομίας τυχαίων μεταβλητών. Έστω d μία ψευδομετρική στο S.

1. Να δειχθεί ότι η σχέση $x \sim y$, αν και μόνο αν $d(x, y) = 0$, καθορίζει μία σχέση ισοδυναμίας (βλ. παράρτημα **??**) στο S.

2. Θέτουμε $\widetilde{x} = \{y \in S \,:\, y \sim x\}$ την κλάση ισοδυναμίας του x. Σχηματίζουμε το χώρο πηλίκο $S/\!\sim\; = \{\,\widetilde{x} : x \in S\}$ και θέτουμε $\widetilde{d}\,(\widetilde{x}, \widetilde{y}) = d(x, y)$. Να δείξετε ότι η συνάρτηση $\widetilde{d}$ είναι μετρική στο $S/\!\sim$. Υπό αυτήν την έννοια λοιπόν, όταν ταυτίζουμε ισοδύναμα σημεία (εδώ σημεία μηδενικής απόστασης), ένας ψευδομετρικός χώρος γίνεται μετρικός.

3. Έστω T ένα ακόμα σύνολο και $f\colon S \to T$ συνάρτηση. Να δειχθεί ότι η σχέση $x \sim y$, αν και μόνο αν $f(x) = f(y)$, καθορίζει μία σχέση ισοδυναμίας στο S. Επιπλέον, αν p είναι μετρική στο T, τότε η συνάρτηση $p_f\colon S^2 \to \mathbb{R}_+$ με $p_f(x) = p(f(x), f(y))$ είναι ψευδομετρική στο S.

4. Έστω $(\Omega, \mathcal{A}, \mathbb{P})$ ένας χώρος πιθανότητας, $\mathcal{X} = \{X\colon \Omega \to \mathbb{R} \mid X \text{ τυχαία μεταβλητή}\}$, δηλ., ο χώρος όλων των πραγματικών τυχαίων μεταβλητών και $\mathcal{P} = \{\mu\colon \mathcal{B}(\mathbb{R}) \to [0, 1] \mid \mu \text{ μέτρο πιθανότητας}\}$, δηλ., ο χώρος όλων των μέτρων πιθανότητας στο μετρήσιμο χώρο $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Αν $\overset{d}{=}$ είναι μία σχέση στο $\mathcal{X}$ με $X \overset{d}{=} Y$, αν και μόνο αν $\mathbb{P}_X = \mathbb{P}_Y$ (δηλ. αν X και Y ισόνομες), τότε να δειχθεί ότι η σχέση ισονομίας είναι σχέση ισοδυναμίας στο χώρο $\mathcal{X}$.

5. Έστω $A \subset \mathbb{R}$. Θέτουμε

$$A^{\varepsilon} := \big\{y \in \mathbb{R} \mid \text{υπάρχει } x \in A \text{ με } |y - x| < \varepsilon\big\} = \bigcup_{x \in A} (x - \varepsilon, x + \varepsilon),$$

και $\pi : \mathcal{P}^2 \to \mathbb{R}_+$ με

$$\pi(\mu, \nu)^{\ddot{}} := \inf \big\{\varepsilon > 0 \mid \mu(A) \le \nu(A^{\varepsilon}) + \varepsilon, \text{ για κάθε } A \in \mathcal{B}(\mathbb{R})\big\}.$$

Να δειχθεί ότι η π είναι μετρική στο $\mathcal{P}$. Η μετρική αυτή είναι γνωστή ως *μετρική Lèvy-Prokhorov* και μετρικοποιεί την ασθενή σύγκλιση των μέτρων πιθανότητας στον $\big(\mathbb{R}, \mathcal{B}(\mathbb{R})\big)$ και επομένως και τη σύγκλιση κατά κατανομή τ.μ., στοιχείο που θα αναδειχθεί στα επόμενα τρία ερωτήματα.

1.13.11 Για 2 τ.μ. $X, Y \in \mathcal{X}$, θέτουμε $p(X, Y) = \pi(\mathbb{P}_X, \mathbb{P}_Y)$, όπου π η μετρική Lèvy-Prokhorov. Να συμπεράνετε ότι η p είναι ψευδομετρική στο $\mathcal{X}$ και ότι ο χώρος $\widetilde{X} := \mathcal{X}/_{\overset{d}{=}}$ γίνεται “φυσιολογικά” μετρικός με τη $\widetilde{p}$.

1.13.12 Έστω $\partial A := \overline{A} \setminus A^\circ$ το σύνορο του $A \subset \mathbb{R}$ και $\mathcal{C}_\mu := \{A \in \mathcal{B}(\mathbb{R}) \mid \mu(\partial A) = 0\}$ για $\mu \in \mathcal{M}$. Η συλλογή $\mathcal{C}_\mu \subset \mathcal{B}(\mathbb{R})$ περιλαμβάνει τα λεγόμενα σύνολα συνεχείας του μέτρου μ, που δεν επιτρέπουν "μάζα" στο σύνορό τους. Θα λέμε ότι μία ακολουθία μέτρων πιθανότητας (μ_n) *συγκλίνει ασθενώς* σε ένα μέτρο μ, και θα γράφουμε $\mu_n \Rightarrow \mu$, αν $\mu_n(A) \to \mu(A)$, για κάθε $A \in \mathcal{C}_\mu$. Να δειχθεί ότι

$$\mu_n \ \Rightarrow \ \mu \iff \pi(\mu_n, \mu) \to 0.$$

Η παραπάνω σχέση μας δείχνει ότι η ασθενής σύγκλιση μέτρων πιθανότητας είναι μετρικοποιήσιμη από τη μετρική Lèvy-Prokhorov.

1.13.13 Η σύγκλιση κατά κατανομή πραγματικών τ.μ. μπορεί τώρα να οριστεί άμεσα από τα παραπάνω. Συγκεκριμένα, λέμε ότι μία ακολουθία τ.μ. (X_n) *συγκλίνει κατά κατανομή* σε μία τ.μ. X, και γράφουμε

$$X_n \xrightarrow{d} X \iff \mathbb{P}_{X_n} \Rightarrow \mathbb{P}_X\,. \tag{1.17}$$

Εξηγείστε από τα παραπάνω με ποιά έννοια η κατά κατανομή σύγκλιση ακολουθιών τ.μ. είναι μετρικοποιήσιμη.

1.13.14 Σε ένα πρώτο μάθημα Πιθανοτήτων δίνεται συνήθως ο εξής ορισμός σύγκλισης κατά κατανομής:

$$X_n \xrightarrow{d} X \iff F_n(x) \to F(x), \quad \text{για κάθε } x \in C_F, \tag{1.18}$$

όπου F_n, F οι σ.κ. των X_n, X και C_F το σύνολο των σημείων συνεχείας της F. Να δειχθεί ότι οι ορισμοί 1.17 και 1.18 είναι ισοδύναμοι.

1.13.15 Έστω $(\mathcal{A}_i)_{i\in I}$ μια οικογένεια σ-αλγεβρών επί του S. Τότε η $\mathcal{A} = \bigcap_{i\in I} \mathcal{A}_i$ είναι σ-άλγεβρα στο S. Να αποδειχθεί ότι το ίδιο ισχύει και για τοπολογίες.

1.13.16 Έστω S σύνολο και $\mathcal{A}_1$, $\mathcal{A}_2$ δύο σ-άλγεβρες στο S. Να δειχθεί ότι η ένωσή τους $\mathcal{A}_1 \cup \mathcal{A}_2 = \{A \subset S \mid A \in \mathcal{A}_1 \text{ ή } A \in \mathcal{A}_2\}$ δεν είναι υποχρεωτικά σ-άλγεβρα. Να δειχθεί ότι το ίδιο ισχύει και για τοπολογίες.

1.13.17 Έστω $S_2 = \{1, 2\}$ και $S_3 = \{1, 2, 3\}$.

1. Να βρεθούν όλες οι σ-άλγεβρες και όλες οι τοπολογίες στα S_2 και S_3.
2. Να βρεθούν όλες οι σ-άλγεβρες στο $S_4 = \{1, 2, 3, 4\}$.
3. Με τη βοήθεια των παραπάνω να γίνει μία κατάλληλη περιγραφή όλων των σ-αλγεβρών στο $S_n = \{1, 2, \dots, n\}$ και να βρεθεί το πλήθος τους.

Το πρόβλημα εύρεσης κλειστού τύπου για το πλήθος των τοπολογιών n-στοιχείων είναι ανοιχτό (βλ. εδώ για μία βατή εισαγωγή στο θέμα & αναφορές).

1.13.18 Να δειχθεί ότι η οικογένεια $\mathcal{C}^o$ των ανοικτών ορθογωνίων του $\mathbb{R}^n$

1. είναι βάση της τοπολογίας γινόμενο,
2. έχει ως υποβάση την $\mathcal{C}^{o;1} := \{\pi_i^{-1}(A_i) \mid 1 \leq i \leq n,\ A_i \text{ ανοικτό στο } \mathbb{R}\}$,

3. παράγει τη σ-άλγεβρα γινόμενο $\mathcal{B}^n(\mathbb{R})$.

1.13.19 Αν Δ είναι διάστημα του $\mathbb{R}$, τότε έχει κάποιον από τους επόμενους 8 τύπους: $\{(a,b),\ [a,b)\,,(a,b],\ [a,b],\ (-\infty,b),\ (-\infty,b],\ (a,+\infty),\ [a,+\infty)\}$. Αν συμβολίσουμε με $\mathcal{D}_i, 1 \leq i \leq 8$, τις οικογένειες των 8 διαφορετικών τύπων διαστημάτων του $\mathbb{R}$ αντίστοιχα, τότε να δειχθεί ότι $\sigma(\mathcal{D}_i) = \mathcal{B}(\mathbb{R})$.

1.13.20 Έστω S ένα (μη κενό) σύνολο και $\mathcal{A}$ μία σ-άλγεβρα στο S. Η $\mathcal{A}$ καλείται *τοπολογικά παραγόμενη* ή απλούστερα *Borel*, αν υπάρχει μία τοπολογία $\mathcal{T}$ στο S με $\mathcal{A} = \sigma(\mathcal{T})$, δηλ. η $\mathcal{A}$ γίνεται Borel του S με τοπολογία την $\mathcal{T}$. Να δειχθεί ότι:

1. κάθε πεπερασμένη σ-άλγεβρα είναι *Borel*,
2. μία σ-άλγεβρα *Borel* μπορεί να παράγεται από διαφορετικές τοπολογίες
3. υπάρχει σ-άλγεβρα που δεν είναι *Borel*.

1.13.21 Να δειχθεί ότι η τοπολογία ίχνος (βλ. ορισμό 1.24) είναι πράγματι τοπολογία. Να δειχθεί το ίδιο και για τη σ-άλγεβρα ίχνος (βλ. ορισμό 1.25).

ΚΕΦΑΛΑΙΟ 2

ΤΥΧΑΙΑ ΔΙΑΝΥΣΜΑΤΑ ΚΑΙ ΠΟΛΥΔΙΑΣΤΑΤΕΣ ΚΑΤΑΝΟΜΕΣ

Η χρήση τυχαίων διανυσμάτων (τ.δ.) ή ισοδύναμα διανυσματικών τυχαίων μεταβλητών (δ.τ.μ.) είναι απαραίτητη για την από κοινού μελέτη περισσότερων του ενός χαρακτηριστικού ενδιαφέροντος σε ένα πείραμα τύχης και είναι το πλαίσιο μέσα στο οποίο εντάσσονται οι εκτιμήτριες σε πολυπαραμετρικά προβλήματα. Ο τυπικός ορισμός ενός τυχαίου διανύσματος είναι μία εξειδίκευση της έννοιας της μετρήσιμης συνάρτησης (βλ. παρατήρηση 1.14), όταν αυτή παίρνει τιμές στον $\mathbb{R}^d$. Όπως και στη μονοδιάστατη περίπτωση, έτσι και εδώ εμπλέκεται η έννοια των συνόλων *Borel*, συμβολικά $\mathcal{B}(\mathbb{R}^d)$, που αφορούν τα "καλά" υποσύνολα του $\mathbb{R}^d$ στα οποία κάποιος θα ήθελε να αποδώσει πιθανότητα για να μπορέσει να απαντήσει σχεδόν σε όλα τα προβλήματα που τον απασχολούν στην πράξη. Καθώς ο βασικός μας χώρος είναι ο $\mathbb{R}^n$ και χρειαζόμαστε κάποιες στοιχειώδεις έννοιες από διανυσματικούς χώρους υπενθυμίζουμε εδώ βασικές αλγεβρικές δομές, δίνοντας κυρίως παραδείγματα από χώρους που μας ενδιαφέρουν στο δικό μας πλαίσιο.

2.1 Ομάδες

Από τις πιο στοιχειώδεις και σημαντικές αλγεβρικές δομές είναι οι ομάδες.

Ορισμός 2.1: Ένα ζεύγος $(G, \circ)$ που αποτελείται από ένα σύνολο G και μία πράξη $\circ \colon G \times G \to G$ καλείται *ομάδα*, αν ικανοποιεί τις ιδιότητες:

Σάμης Τρέβεζας (2023). «Μη Παραμετρική Στατιστική».
Αθήνα: Κάλλιπος, Ανοικτές Ακαδημαϊκές Εκδόσεις.
BY-NC-SA Creative Commons Αναφορά Δημιουργού - Μη Εμπορική Χρήση - Παρόμοια Διανομή 4.0

1. $a \circ (b \circ c) = (a \circ b) \circ c$ [προσεταιριστική ιδιότητα],
2. υπάρχει $e \in G :\ a \circ e = e \circ a = e$, για κάθε $a \in G$ [ύπαρξη ουδετέρου στοιχείου],
3. για κάθε $a \in G$, υπάρχει $b \in G :\ a \circ b = b \circ a = e$, το οποίο το συμβολίζουμε με a^{-1} [ύπαρξη αντιστρόφου στοιχείου].

Αν επιπλέον ικανοποιείται και η ιδιότητα

(iv) $a \circ b = b \circ a$, για κάθε $a, b \in G$ [αντιμεταθετική ιδιότητα],

τότε η ομάδα καλείται *αντιμεταθετική* ή αβελιανή. Κάθε υποσύνολο H της G, που ικανοποιεί τις ιδιότητες της ομάδας ως προς την ίδια πράξη, αλλά μέσα στο H, καλείται *υποομάδα* της G και γράφουμε $H \leq G$.

Η επαλήθευση της ιδιότητας της υποομάδας μέσα σε μία ομάδα γίνεται απλά, ελέγχοντας την επόμενη συνθήκη.

Πρόταση 2.1: Έστω $(G, \circ)$ μία ομάδα και $H \subset G$.

$$H \leq G \quad \Longleftrightarrow \quad a \circ b^{-1} \in H \quad \text{για κάθε } a, b \in H.$$

Τα σημαντικότερα παραδείγματα ομάδων όταν δουλεύουμε με τους πραγματικούς αριθμούς είναι η προσθετική ομάδα $(\mathbb{R}, +)$ και η πολλαπλασιαστική ομάδα $(\mathbb{R}^*, \cdot)$ με πράξεις την πρόσθεση και τον πολλαπλασιασμό πραγματικών αριθμών αντίστοιχα, που είναι μάλιστα και αβελιανές ομάδες. Όταν δουλεύουμε βέβαια με συναρτήσεις η πιο "φυσιολογική" πράξη είναι η σύνθεση συναρτήσεων. Το επόμενο παράδειγμα αφορά σύνολα με πεπερασμένο σύνολο στοιχείων.

Παράδειγμα 2.1: Έστω $S = \{1, 2, \ldots, n\}$. Ορίζουμε ως *μετάθεση* κάθε συνάρτηση $\sigma\colon S \to S$ που είναι 1-1 και επί. Αν σχηματίσουμε το σύνολο όλων των μεταθέσεων $S_n = \{\sigma\colon S \to S \mid \sigma \text{ είναι 1-1 και επί}\}$ και το εφοδιάσουμε με την πράξη της σύνθεσης, τότε το ζεύγος $(S_n, \circ)$ αποτελεί ομάδα, γνωστή με το όνομα *συμμετρική ομάδα* n στοιχείων. Η ομάδα αυτή δεν είναι αβελιανή για $n \geq 3$. Για παράδειγμα, αν κινηθούμε από τα αριστερά προς τα δεξιά:

$$(1\ 2) \circ (1\ 3) = (1\ 2\ 3), \quad \text{ενώ} \quad (1\ 3) \circ (1\ 2) = (1\ 3\ 2),$$

και άρα η σειρά με την οποία εφαρμόζουμε τις αντιμεταθέσεις $(1\ 2)$ και $(1\ 3)$ μας οδηγεί σε διαφορετική μετάθεση. Το παράδειγμα αυτό μη αντιμεταθετικότητας "δουλεύει" για κάθε $n \geq 3$, αφού ακόμα και αν υπάρχουν και άλλα στοιχεία, οι μεταθέσεις που θα προκύψουν θα διαφέρουν τουλάχιστον σε 3 σημεία. Η σπουδαιότητα της συμμετρικής ομάδας έγκειται στο γεγονός ότι κάθε πεπερασμένη ομάδα n στοιχείων είναι ισομορφική με μία υποομάδα της S_n, όπως μπορεί να δειχθεί από το θεώρημα του Cayley. Αυτό το γεγονός αναδεικνύει βέβαια και τη σημασία της S_n. Στο παράδειγμα της S_3, η οποία έχει 3!=6 μεταθέσεις, τα στοιχεία της γράφονται ως εξής:

$$S_3 = \{(1), (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}.$$

Οι υποομάδες που μπορούμε να φτιάξουμε από την S_3 θα είναι της μορφής:

$$H_1 = \{(1)\},\ H_2 = \{(1), (\alpha\ \beta)\},\ H_3 = \{(1), (\alpha\ \beta\ \gamma), (\alpha\ \gamma\ \beta)\},\ H_6 = S_3,$$

όπου ως δείκτη δηλώνουμε το πλήθος των στοιχείων της υποομάδας και ως σύμβαση παίρνουμε $\alpha < \beta < \gamma$. Η σημασία των ισομορφισμών έγκειται λοιπόν στο γεγονός ότι με την παραπάνω αναπαράσταση έχουμε κάνει μία περίληψη όλων των δυνατών τύπων υποομάδων που μπορούμε να έχουμε, χωρίς να τις καταγράψουμε όλες. Για παράδειγμα υπάρχουν 3 διαφορετικές υποομάδες τύπου H_2 που προκύπτουν με $\binom{3}{2} = 3$ διαφορετικούς τρόπους που μπορούμε να επιλέξουμε το ζεύγος $(\alpha\ \beta)$. Αυτού του τύπου οι υποομάδες είναι ισόμορφες, διότι έχουν ταυτόσημη αλγεβρική δομή, με άλλα λόγια μεταφερόμαστε από τη μία στην άλλη μόνο με αλλαγές στις "ετικέτες" των στοιχείων.

Η συμμετρική ομάδα ορίζεται και για απειροσύνολα S.

Ορισμός 2.2: Έστω S σύνολο και

$$\text{Sym}(S) := \{f\colon S \to S \mid\ f \text{ είναι 1-1 και επί}\}.$$

Κάθε στοιχείο του $\text{Sym}(S)$ καλείται *μετάθεση* του S, και εκφράζει μία αναδιάταξη των στοιχείων του S. Το σύνολο $\text{Sym}(S)$ θα θεωρείται εφοδιασμένο με την πράξη της σύνθεσης και καλείται *συμμετρική ομάδα* του S. Διάφορες υποομάδες της $\text{Sym}(S)$ έχουν ιδιαίτερο ενδιαφέρον. Για παράδειγμα, αν ο S είναι και τοπολογικός χώρος, τότε το σύνολο των ομοιομορφισμών (τοπολογικών ισομορφισμών), δηλ.

$$\text{Iso}^t(S) := \{f : S \to S \mid f \text{ συνεχής, 1-1 και επί και } f^{-1} \text{ συνεχής}\}$$

είναι υποομάδα της $\text{Sym}(S)$. Το ίδιο ισχύει και για τους Borel ισομορφισμούς,

$$\text{Iso}^b(S) := \{f : S \to S \mid f \text{ Borel, 1-1 και επί και } f^{-1} \text{ Borel}\},$$

αν θεωρήσουμε την αντίστοιχη Borel σ–άλγεβρα. Αυτό προκύπτει άμεσα από το γεγονός ότι $g^{-1} \circ f$ είναι συνεχής ή Borel μετρήσιμη αντίστοιχα, αν οι f, g έχουν την ίδια ιδιότητα. Επιπλέον, αφού κάθε συνεχής συνάρτηση είναι και Borel μετρήσιμη, έχουμε τελικά

$$\text{Iso}^t(S) \leq \text{Iso}^b(S) \leq \text{Sym}(S).$$

Η μελέτη μιας οποιασδήποτε δομής πάνω σε ένα σύνολο διευκολύνεται με τους ισομορφισμούς αυτής της δομής που μας δίνει τη δυνατότητα να μεταφερθούμε σε ένα απλούστερο σύνολο, που το καταλαβαίνουμε καλύτερα και αντί να μελετάμε τις ιδιότητες της δομής που μας ενδιαφέρει στο αρχικό σύνολο, μπορούμε να τις μελετάμε στο απλούστερο σύνολο, χωρίς να χάσουμε καμία πληροφορία.

Ορισμός 2.3: Έστω $(G, \circ)$ και $(H, *)$ δύο ομάδες. Μία συνάρτηση $f\colon G \to H$ καλείται *ομομορφισμός ομάδων*, αν ισχύει:

$$f(a \circ b) = f(a) * f(b), \quad \text{για κάθε } a, b \in G.$$

Αν η f είναι επιπλέον 1-1 και επί, τότε καλείται *ισομορφισμός ομάδων*. Στην περίπτωση που υπάρχει ένας τέτοιος ισομορφισμός οι ομάδες G και H λέγονται *ισόμορφες*. Αν $(G, \circ) = (H, *)$, τότε ο ισομορφισμός καλείται *αυτομορφισμός* και το σύνολο των αυτομορφισμών συμβολίζεται με $\mathrm{Aut}(G)$.

Παρατήρηση 2.1. Η συνθήκη του ομομορφισμού ομάδων καθορίζει τη συμβατότητα στις δομές των ομάδων τους. Αυτό γίνεται ακόμα καλύτερα κατανοητό αν το H δεν ήταν εφοδιασμένο με κάποια πράξη. Πώς θα έπρεπε να οριστεί τότε μία πράξη στο H ώστε να είναι αντίστοιχη της G; Ο τρόπος σύνδεσης των στοιχείων της G επάγει μέσω της f έναν τρόπο σύνδεσης στο H:

$$(a, b) \longmapsto a \circ b \quad \Longrightarrow \quad (f(a), f(b)) \longmapsto f(a \circ b).$$

Η τελευταία σχέση λοιπόν μας δείχνει ότι για να υπάρχει συμβατότητα με οποιαδήποτε πράξη $*$ στο H, πρέπει $f(a) * f(b) = f(a \circ b)$. Οι επιπλέον συνθήκες ότι η f είναι 1-1 και επί εξασφαλίζει επιπλέον ότι τα σύνολα G και H είναι ισοπληθικά και μαζί με τις συμβατές τους πράξεις αποτελούν στην ουσία ταυτόσημες ομάδες.

Παρατήρηση 2.2. Θα περίμενε ίσως κανείς ότι για να εξασφαλίσουμε τον ισομορφισμό θα έπρεπε να απαιτήσουμε επιπλέον ότι η f^{-1} είναι και αυτή ομομορφισμός ομάδων. Γενικά, όταν δείχνουμε ισομορφισμούς δομών δεν πρέπει να ξεχνάμε ότι ως η f^{-1} πρέπει να έχει τις ίδιες ιδιότητες με την f και όταν αυτό δεν έπεται από τον τύπο της δομής που εξετάζουμε, τότε πρέπει να προστεθεί ως συνθήκη στον ορισμό. Εδώ, βέβαια παρατηρούμε ότι αν h_1, $h_2 \in H$, τότε από την ιδιότητα επί, έχουμε ότι υπάρχουν g_1, $g_2 \in G$ με $h_1 = f(g_1)$ και $h_2 = f(g_2)$, από το οποίο έπεται ότι

$$f^{-1}(h_1 * h_2) = f^{-1}(f(g_1) * f(g_2)) = f^{-1}(f(g_1 \circ g_2)) = g_1 \circ g_2 = f^{-1}(h_1) \circ f^{-1}(h_2),$$

όπου στην τελευταία σχέση χρησιμοποιήσαμε και την ιδιότητα 1-1. Τελικά, η παραπάνω σχέση μας δείχνει ότι η f^{-1} είναι ομομορφισμός ομάδων και άρα δεν απαιτείται ως επιπλέον συνθήκη στον ορισμό του ισομορφισμού.

Παράδειγμα 2.2: Το σύνολο των αυτομορφισμών $\mathrm{Aut}(G)$ αποτελεί ομάδα με πράξη τη σύνθεση που μάλιστα είναι υποομάδα της $\mathrm{Sym}(G)$. Πράγματι, εφόσον ισχύει $\mathrm{Aut}(G) \subset \mathrm{Sym}(G)$ και $\mathrm{Sym}(G)$ είναι ομάδα, αρκεί να δειχθεί ότι ικανοποιεί τη συνθήκη $f \circ g^{-1} \in \mathrm{Aut}(G)$, για κάθε $f, g \in \mathrm{Aut}(G)$. Η συνθήκη 1-1 και επί ικανοποιείται, άρα αρκεί να δειχθεί η συνθήκη του ομομορφισμού ομάδων. Έχουμε $f, g^{-1} \in \mathrm{Aut}(G)$ και άρα πράγματι

$$f \circ g^{-1}(a * b) = f\left(g^{-1}(a) * g^{-1}(b)\right) = \left[f \circ g^{-1}(a)\right] * \left[f \circ g^{-1}(b)\right].$$

Στο σχηματισμό των ομάδων, μεγαλύτερο ενδιαφέρον παρουσιάζεται όταν ένα σύνολο G έχει αρχικά κάποια δομή. Στο σύγγραμμα αυτό μας ενδιαφέρει η περίπτωση που το G είναι είτε ένας τοπολογικός, είτε ένας μετρήσιμος χώρος. Σε αυτήν την περίπτωση η διαδικασία της αντιστροφής και η πράξη που καθορίζει το G ως ομάδα είναι συναρτήσεις μεταξύ χώρων με δομή και θα ήταν επιθυμητό να είναι συμβατές με τη δομή που έχει το σύνολο G. Αν το G είναι τοπολογικός χώρος, τότε η συμβατότητα επιτυγχάνεται με το να απαιτήσουμε οι συναρτήσεις αυτές να είναι συνεχείς και αν είναι μετρήσιμος χώρος, με το να είναι μετρήσιμες συναρτήσεις. Στην περίπτωση της δυαδικής πράξης της ομάδας, ο χώρος γινόμενο $S \times S$ νοείται εφοδιασμένος με την τοπολογία ή τη σ–άλγεβρα γινόμενο στο S αντίστοιχα. Έτσι οδηγούμαστε φυσιολογικά στον ορισμό των τοπολογικών και των μετρήσιμων ομάδων.

Ορισμός 2.4: Έστω S σύνολο, $\circ\colon S \times S \to S$ μία πράξη, $\mathcal{T}$ μια τοπολογία και $\mathcal{A}$ μία σ–άλγεβρα στο S.

- Η τριάδα $(S, \circ, \mathcal{T})$ καλείται *τοπολογική ομάδα* στο S, αν το ζεύγος $(S, \mathcal{T})$ είναι τοπολογικός χώρος, το ζεύγος $(S, \circ)$ είναι ομάδα και η αντιστροφή $g \mapsto g^{-1}$, αλλά και η πράξη της ομάδας $(f, g) \mapsto f \circ g$ είναι συνεχείς συναρτήσεις.
- Η τριάδα $(S, \circ, \mathcal{A})$ καλείται *μετρήσιμη ομάδα* στο S, αν το ζεύγος $(S, \mathcal{A})$ είναι μετρήσιμος χώρος, το ζεύγος $(S, \circ)$ είναι ομάδα και η αντιστροφή $g \mapsto g^{-1}$, αλλά και η πράξη της ομάδας $(f, g) \mapsto f \circ g$ είναι μετρήσιμες συναρτήσεις.
- Μία μετρήσιμη ομάδα $(S, \circ, \mathcal{A})$ καλείται *ομάδα Borel* αν $\mathcal{A} = \sigma(\mathcal{T})$ για κάποια τοπολογία $\mathcal{T}$.

Κάθε υποσύνολο R του S που είναι τοπολογική ομάδα, μετρήσιμη ομάδα ή ομάδα Borel, καλείται και *τοπολογική υποομάδα, μετρήσιμη υποομάδα ή υποομάδα Borel* του S αντίστοιχα.

Ο σχηματισμός υποομάδων από τις παραπάνω ομάδες πρέπει λοιπόν φυσιολογικά να "σέβεται" και τις υπάρχουσες δομές. Αν όμως επαληθευτεί η κλειστότητα στην πράξη της ομάδας, τότε τα υπόλοιπα εξασφαλίζονται άμεσα διότι ο περιορισμός μιας συνεχούς ή μετρήσιμης συνάρτησης διατηρεί την ιδιότητά της σε οποιοδήποτε υποσύνολο του αρχικού χώρου με την τοπολογία ή σ–άλγεβρα ίχνος αντίστοιχα (βλ. πρόταση 1.9). Άρα αυτό συμβαίνει και με τον περιορισμό της πράξης της ομάδας.

Πόρισμα 2.1: Αν ένα υποσύνολο R είναι αλγεβρική υποομάδα μιας τοπολογικής, μετρήσιμης ή Borel ομάδας S, τότε είναι και τοπολογική, μετρήσιμη ή Borel υποομάδα της S.

Ας δούμε τώρα μία ενδιαφέρουσα εφαρμογή των παραπάνω εννοιών στο σχηματισμό των λεγόμενων ομαδικών οικογενειών κατανομών. Όπως και στις εκθετικές οικογένειες κατανομών, ο όρος αυτός δεν αναφέρεται σε μία συγκεκριμένη οικογένεια κατανομών, αλλά σε έναν τρόπο σύνδεσης των μελών της οικογένειας.

Ορισμός 2.5: Έστω $\mathcal{K}$ μία οικογένεια κατανομών στο $\mathbb{R}$. Θα λέμε ότι η $\mathcal{K}$ αποτελεί *ομαδική οικογένεια κατανομών*, αν υπάρχει μία κατανομή P και μία ομάδα (Borel-)μετρήσιμων μεταθέσεων $G \leq \mathrm{Sym}(\mathbb{R})$, έτσι ώστε για κάθε $K \in \mathcal{K}$, να υπάρχει $g \in G$:

$$K = P \circ g^{-1} \equiv P_g,$$

όπου υπενθυμίζουμε ότι για κάθε $B \in \mathcal{B}(\mathbb{R})$, έχουμε

$$P_g(B) = P \circ g^{-1}(B) = P(g \in B).$$

Ειδικότερα, αν επιλέξουμε

$$G = \{g \colon \mathbb{R} \to \mathbb{R} \mid g(x) = x + \alpha,\ \ \alpha \in \mathbb{R}\},$$

τότε η $\mathcal{K}$ καλείται *οικογένεια θέσης*.
Αν επιλέξουμε

$$G = \{g \colon \mathbb{R} \to \mathbb{R} \mid g(x) = \beta x,\ \ \beta > 0\},$$

τότε η $\mathcal{K}$ καλείται *οικογένεια κλίμακας*.
Αν επιλέξουμε

$$G = \{g \colon \mathbb{R} \to \mathbb{R} \mid g(x) = \alpha + \beta x,\ \ \alpha \in \mathbb{R},\ \ \beta > 0\},$$

τότε η $\mathcal{K}$ καλείται *οικογένεια θέσης-κλίμακας*.

Παρατήρηση 2.3. Οι οικογένειες που προκύπτουν μέσα από τους μετασχηματισμούς θέσης, κλίμακας και θέσης-κλίμακας είναι πράγματι ομάδες όπως επαληθεύεται εύκολα. Πράγματι, επειδή τα μέλη της G είναι συναρτήσεις 1-1 και επί, άρα μεταθέσεις του $\mathbb{R}$, αρκεί να δειχθεί η σχέση ότι $g^{-1} \circ f \in G$ για $f, g \in G$, το οποίο έπεται εύκολα. Επιπλέον, η μετρησιμότητα έπεται από το ότι αντιστοιχούν σε συνεχείς μετασχηματισμούς.

Πρόταση 2.2: Μία οικογένεια κατανομών στο $\mathbb{R}$ είναι ομαδική, αν και μόνο αν συμπίπτει με την οικογένεια των κατανομών των τυχαίων μεταβλητών $X = g(Z)$, όπου Z είναι πραγματική τυχαία μεταβλητή και $g \in G$ για κάποια ομάδα μετρήσιμων μεταθέσεων του $\mathbb{R}$. Η ομαδική οικογένεια είναι

- θέσης, αν $X = \alpha + Z$ για $\alpha \in \mathbb{R}$,
- κλίμακας, αν $X = \beta Z$ για $\beta > 0$,
- θέσης-κλίμακας, αν $X = \alpha + \beta Z$ για $\alpha \in \mathbb{R}$ και $\beta > 0$.

Απόδειξη: Αν Z είναι πραγματική τυχαία μεταβλητή ορισμένη σε χώρο πιθανότητας

$(\Omega, \mathcal{A}, \mathbb{P})$ και θεωρήσουμε την οικογένεια των τ.μ. $X = g(Z)$, τότε

$$\mathbb{P}_X = \mathbb{P}_{g\circ Z} = \mathbb{P} \circ (g \circ Z)^{-1} = \mathbb{P} \circ Z^{-1} \circ g^{-1} = \mathbb{P}_Z \circ g^{-1},$$

και έτσι προκύπτει ομαδική οικογένεια κατανομών. Αντίστροφα, αν δοθεί ομαδική οικογένεια κατανομών με κάποιο P και ομάδα G, τότε υπάρχει χώρος πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$ και τ.μ. Z με $\mathbb{P}_Z = P$. Είναι τώρα άμεσο πάλι να δειχθεί ότι $P_g = \mathbb{P}_{g(Z)}$, αφού

$$P_g = P \circ g^{-1} = \mathbb{P} \circ Z^{-1} \circ g^{-1} = \mathbb{P} \circ (g \circ Z)^{-1} = \mathbb{P}_{g(Z)}.$$

Οι διάφορες εξειδικεύσεις της ομάδας μετασχηματισμών δίνουν άμεσα τις υπόλοιπες οικογένειες. □

Χρησιμοποιώντας συναρτήσεις κατανομής, που όπως γνωρίζουμε χαρακτηρίζουν τις κατανομές μπορούμε άμεσα να έχουμε ισοδύναμες αναπαραστάσεις.

Πόρισμα 2.2: Αν $\mathcal{F}$ είναι μία οικογένεια συναρτήσεων κατανομής (είτε τυχαίων μεταβλητών, είτε κατανομών), τότε αντιστοιχούν σε ομαδική οικογένεια κατανομών, αν και μόνο αν υπάρχει $F_e \in \mathcal{F}$ και μία ομάδα μετρήσιμων μεταθέσεων G, έτσι ώστε αν Z τ.μ. με $F_Z = F_e$, τότε $F \in \mathcal{F}$ αν και μόνο αν $F = F_g$, όπου F_g είναι η σ.κ. της $g(Z)$.

- Στην ειδική περίπτωση που η ομάδα αποτελείται από γνησίως αύξουσες συναρτήσεις, τότε

  $$F_g(x) = F_e(g^{-1}(x)), \qquad \text{για κάθε } x \in \mathbb{R}.$$

  Η προκύπτουσα οικογένεια είναι:
  1. θέσης, αν $F_g(x) \equiv F_\alpha(x) = F_0(x - \alpha)$ για $\alpha \in \mathbb{R}$,
  2. κλίμακας, αν $F_g(x) \equiv F_\beta(x) = F_1(x/\beta)$ για $\beta > 0$,
  3. θέσης-κλίμακας, αν $F_g(x) \equiv F_{\alpha,\beta}(x) = F_{0,1}((x-\alpha)/\beta)$ για $\alpha \in \mathbb{R}$ και $\beta > 0$.
- Αν επιπλέον είναι απόλυτα συνεχείς και άρα έχουν συνάρτηση πυκνότητας, τότε

  $$f_g(x) = f_e\big(g^{-1}(x)\big)\frac{dg^{-1}(x)}{dx}, \qquad \text{για κάθε } x \in \mathbb{R}.$$

  και για μετασχηματισμούς θέσης-κλίμακας είναι:
  1. θέσης, αν $f_g(x) \equiv f_\alpha(x) = f_0(x - \alpha)$ για $\alpha \in \mathbb{R}$,
  2. κλίμακας, αν $f_g(x) \equiv f_\beta(x) = \beta^{-1} f_1\big(x/\beta\big)$ για $\beta > 0$,
  3. θέσης-κλίμακας, αν $f_g(x) \equiv f_{\alpha,\beta}(x) = \beta^{-1} f_{0,1}\big((x-\alpha)/\beta\big)$ για $\alpha \in \mathbb{R}$ και $\beta > 0$.

Παράδειγμα 2.3: Αρκετές συνεχείς παραμετρικές οικογένειες κατανομών αποτελούν ομαδικές οικογένειες κατανομών. Δίνουμε εδώ κάποια παραδείγματα οικογενειών κατανομών:

- Κανονική: $X \sim \mathcal{N}(\mu, \sigma^2) \iff X = \mu + \sigma Z$ με $Z \sim \mathcal{N}(0,1)$ (θέσης-κλίμακας),

- Εκθετική: $X \sim \mathcal{E}(\theta) \iff X = \theta\, Z$, με $Z \sim \mathcal{E}(1)$ (κλίμακας),
- Ομοιόμορφη: $X \sim \mathcal{U}(\theta_1, \theta_2) \iff X = \theta_1 + (\theta_2 - \theta_1)\, U$, με $U \sim \mathcal{U}(0,1)$ (θέσης-κλίμακας),
- Cauchy: $X \sim \mathcal{C}(m, \sigma) \iff X = m + \sigma Z$ με $Z \sim \mathcal{C}(0,1)$ (θέσης-κλίμακας).

Σε κάθε οικογένεια θέσης η παράμετρος θέσης κινείται στο $\mathbb{R}$, ενώ σε οικογένεια κλίμακας η παράμετρος κλίμακας στο $(0, +\infty)$. Σε κάθε οικογένεια θέσης-κλίμακας συνδυάζονται και οι δύο παράμετροι. Είναι σημαντικό να παρατηρήσουμε ότι από μία οποιαδήποτε κατανομή μπορούμε να φτιάξουμε την αντίστοιχη οικογένεια θέσης-κλίμακας εφαρμόζοντας τους μετασχηματισμούς θέσης-κλίμακας. Για παράδειγμα, από μία Student $\mathcal{T}_n$ με n-βαθμούς ελευθερίας προκύπτει η αντίστοιχη οικογένεια θέσης-κλίμακας $\mathcal{T}_n(\mu, \sigma)$ που είναι αρκετά χρήσιμη σε εφαρμογές. Τέλος, η αρχική παραμέτρηση μιας οικογένειας θέσης-κλίμακας (ή μόνο θέσης, ή μόνο κλίμακας) μπορεί να μην αντιστοιχεί στις παραμέτρους θέσης και κλίμακας που μετασχηματίζουν την "τυπική" κατανομή της οικογένειας, αλλά να είναι κάποιας άλλης βολικής μορφής. Για παράδειγμα, η ομοιόμορφη $\mathcal{U}(\theta_1, \theta_2)$ που είδαμε παραπάνω έχει ως παράμετρο κλίμακας το εύρος $\theta_2 - \theta_1$ και όχι κάποια από τις δύο παραμέτρους.

2.2 Διανυσματικοί Χώροι

Η μελέτη των πεπερασμένης διάστασης διανυσματικών (γραμμικών) χώρων και των γραμμικών απεικονίσεων μεταξύ τους αποτελεί το αντικείμενο μελέτης της Γραμμικής Άλγεβρας. Πριν υπενθυμίσουμε τον ορισμό του διανυσματικού χώρου, ας θυμηθούμε επίσης την έννοια του σώματος.

Ορισμός 2.6: Μία τριάδα $(\mathbb{F}, +, *)$ που αποτελείται από ένα σύνολο $\mathbb{F}$ και δύο πράξεις, μία πρόσθεση $+ \colon \mathbb{F} \times \mathbb{F} \to \mathbb{F}$ και έναν πολλαπλασιασμό $* \colon \mathbb{F} \times \mathbb{F} \to \mathbb{F}$ καλείται *σώμα*, αν

1. το ζεύγος $(\mathbb{F}, +)$ είναι αντιμεταθετική ομάδα με ουδέτερο στοιχείο που το συμβολίζουμε με 0,
2. το ζεύγος $(\mathbb{F}^*, *)$ είναι αντιμεταθετική ομάδα με ουδέτερο στοιχείο που το συμβολίζουμε με 1,
3. ισχύει η επιμεριστική ιδιότητα του πολλαπλασιασμού ως προς την πρόσθεση, δηλ.

$$x * (y + z) = x * y + x * z, \qquad \text{για κάθε } x, y, z \in \mathbb{F}.$$

Κάθε υποσύνολο $\mathbb{K}$ του $\mathbb{F}$, που ικανοποιεί τις ιδιότητες του σώματος ως προς τις ίδιες πράξεις, αλλά μέσα στο $\mathbb{K}$, καλείται *υπόσωμα* του $\mathbb{F}$ και γράφουμε $\mathbb{K} \leq \mathbb{F}$.

Η επαλήθευση της ιδιότητας του υποσώματος μέσα σε ένα σώμα γίνεται απλά, ελέγχοντας τις επόμενες συνθήκες.

Πρόταση 2.3: Έστω $(\mathbb{F}, +, *)$ ένα σώμα και $\mathbb{K} \subset \mathbb{F}$ με $\mathbb{K}^* \neq \emptyset$. Τότε, έχουμε

$$\mathbb{K} \leq \mathbb{F} \quad \Longleftrightarrow \quad x - y \in \mathbb{K} \quad \text{και} \quad x * y^{-1} \in \mathbb{K} \quad \text{για κάθε } x, y \in \mathbb{K}.$$

Τα σημαντικότερα παραδείγματα σωμάτων είναι το σώμα $\mathbb{Q}$ των ρητών αριθμών και το σώμα των πραγματικών αριθμών $\mathbb{R}$ με πράξεις τη συνήθη πρόσθεση και τον πολλαπλασιασμό πραγματικών αριθμών, αλλά και το σώμα των μιγαδικών αριθμών $\mathbb{C}$, όπου ο πολλαπλασιασμός επεκτείνεται σε όλους τους μιγαδικούς αριθμούς. Το σώμα θα το χρειαστούμε εδώ για τον ορισμό του διανυσματικού χώρου, αφού θα χρειαστεί να δηλώσουμε την ύπαρξη βαθμωτών (scalars) που επιδρούν πάνω στα αντικείμενα (στοιχεία) του χώρου, που τα λέμε διανύσματα. Η γεωμετρική ερμηνεία της επίδρασης των βαθμωτών πάνω στα αντικείμενα (στοιχεία) στα οποία εφαρμόζονται είναι σημαντική. Η βασική τους επίδραση είναι να μεγεθύνουν ή να συρρικνώνουν το αντικείμενο (στοιχείο) πάνω στο οποίο επιδρούν. Επιπλέον, ανάλογα με το ποιο σώμα επικαλείται θα μπορούσαν να αλλάζουν τη φορά ενός διανύσματος (π.χ. με το $\mathbb{R}$) ή ακόμα και να τα περιστρέφουν (με το $\mathbb{C}$).

Ορισμός 2.7: Ένα σύνολο V εφοδιασμένο με δύο πράξεις, μία εσωτερική $+ \colon V \times V \to V$ που τη λέμε *πρόσθεση*, και μία (αριστερά) εξωτερική $\cdot : \mathbb{F} \times V \to V$ που τη λέμε *βαθμωτό πολλαπλασιασμό* ή *βαθμωτό γινόμενο* με βαθμωτά που ανήκουν σε ένα σώμα $\mathbb{F}$, καλείται $\mathbb{F}-$*διανυσματικός χώρος*, ή *διανυσματικός χώρος επί του* $\mathbb{F}$, αν:

- το ζεύγος $(V, +)$ είναι αντιμεταθετική ομάδα,
- το βαθμωτό γινόμενο $\cdot : \mathbb{F} \times V \to V$ ικανοποιεί για κάθε $\lambda, \mu \in \mathbb{F}$, $x, y \in V$,
  1. $\lambda \cdot (\mu \cdot x) = (\lambda\, \mu) \cdot x$ (συμβατότητα πολλαπλασιασμών σώματος και βαθμωτού),
  2. $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$ (το βαθμωτό επιμεριστικό ως προς "+" στο σώμα),
  3. $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$ (το βαθμωτό επιμεριστικό ως προς "+" διανυσμάτων),
  4. $1 \cdot x = x$ (ταυτοτικό στοιχείο βαθμωτού πολ/σμού)

Κάθε υποσύνολο W του V, που είναι και το ίδιο διανυσματικός χώρος με τις ίδιες πράξεις, αλλά μέσα στο W, καλείται *διανυσματικός υπόχωρος* ή *γραμμικός υπόχωρος* του V και γράφουμε $W \leq V$.

Η επαλήθευση της ιδιότητας του διανυσματικού υπόχωρου ενός διανυσματικού χώρου γίνεται απλά, ελέγχοντας την επόμενη συνθήκη.

Πρόταση 2.4: Έστω V ένας $\mathbb{F}$-διανυσματικός χώρος και $W \subset V$ με $W \neq \emptyset$. Τότε,

$$W \leq V \quad \Longleftrightarrow \quad \lambda \cdot x + y \in W \quad \text{για κάθε } \lambda \in \mathbb{F},\ \ x, y \in W.$$

Παρατήρηση 2.4. Το βαθμωτό γινόμενο μπορεί να ερμηνευθεί και διαφορετικά, δίνοντας μία διαφορετική αναπαράσταση του διανυσματικού χώρου με τη βοήθεια της αρχικής ομάδας και μιας ομάδας αυτομορφισμών της. Αν λοιπόν $\cdot : \mathbb{F} \times V \to V$

είναι το βαθμωτό γινόμενο, τότε θέτουμε

$$\mathcal{M} := \{M_\lambda \colon V \to V \text{ με } M_\lambda(x) = \lambda \cdot x \mid \lambda \in \mathbb{F}\},$$

δηλαδή όλους τους μετασχηματισμούς του V που φτιάχνονται με τη βοήθεια του $\mathbb{F}$. Παρατηρούμε βέβαια ότι είναι καλά ορισμένες συναρτήσεις με τιμές στο V, αφού το βαθμωτό γινόμενο είναι εξωτερική πράξη στο V. Επιπλέον αποτελούν μεταθέσεις του V (να γίνει επαλήθευση του 1-1 και επί). Μπορούμε λοιπόν να αναδιατυπώσουμε τις 4 συνθήκες που καθορίζουν το βαθμωτό γινόμενο ως εξής:

$$\begin{aligned} M_\lambda \circ M_\mu(x) &= M_{\lambda\mu}(x) \\ M_{\lambda+\mu}(x) &= M_\lambda(x) + M_\mu(x) \\ M_\lambda(x+y) &= M_\lambda(x) + M_\lambda(y) \\ M_1(x) &= \mathrm{id}_V(x). \end{aligned}$$

Οι παραπάνω συνθήκες μας λένε ότι ένα βαθμωτό γινόμενο καθορίζει στην ομάδα V μία υποομάδα $\mathcal{M}$ της $\mathrm{Aut}(V)$, δηλ. των αυτομορφισμών της V, η οποία επιπλέον είναι κλειστή στην πρόσθεση συναρτήσεων (που ορίζεται κατά σημείο) και άρα είναι και προσθετική ομάδα. Έτσι από μία προσθετική ομάδα $(V, +)$, ο βαθμωτός πολλαπλασιασμός καθορίζει ένα σώμα $(\mathcal{M}, +, \circ)$ που είναι υπόσωμα του $(\mathrm{Aut}(G), +, \circ)$.

Παράδειγμα 2.4: Έστω $(\mathbb{F}, +, *)$ ένα σώμα. Θέτουμε $\mathbb{F}^n$ το καρτεσιανό γινόμενο n αντιγράφων του $\mathbb{F}$ και ορίζουμε 2 πράξεις:

$$\begin{aligned} + &: \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}^n \quad \text{με} \quad (x, y) \longmapsto x + y := (x_1 + y_1, \ldots x_n + y_n), \\ \cdot &: \mathbb{F} \times \mathbb{F}^n \to \mathbb{F}^n \quad \text{με} \quad (\lambda, x) \longmapsto \lambda \cdot x := (\lambda * x_1, \ldots, \lambda * x_n), \end{aligned}$$

που προκύπτουν από τις πράξεις του $\mathbb{F}$ κατά συντεταγμένη. Είναι άμεσα επαληθεύσιμο ότι το σύνολο $\mathbb{F}^n$ με αυτές τις πράξεις είναι $\mathbb{F}-$διανυσματικός χώρος. Ο $\mathbb{R}-$διανυσματικός χώρος $\mathbb{R}^n$ είναι ειδική περίπτωση του παραπάνω για $\mathbb{F} = \mathbb{R}$.

Παράδειγμα 2.5: Είναι φανερό ότι αν διαθέτουμε ένα σώμα $\mathbb{F}$ και θέσουμε $\mathbb{F}^T$ το σύνολο όλων των συναρτήσεων $f \colon T \to \mathbb{F}$, όπου T ένα οποιοδήποτε μη κενό σύνολο, τότε ο χώρος $\mathbb{F}^T$ είναι $\mathbb{F}$-διανυσματικός χώρος με πράξη την κατά σημείο πρόσθεση συναρτήσεων και βαθμωτό γινόμενο τον κατά σημείο πολλαπλασιασμό του βαθμωτού με τη συνάρτηση. Τα σημαντικότερα παραδείγματα τέτοιων χώρων είναι οι χώροι της μορφής $\mathbb{R}^{\{1,2,\ldots,n\}}$, $n \geq 1$, δηλ. οι χώροι όλων των πεπερασμένων ακολουθιών πραγματικών αριθμών μήκους n, ο χώρος $\mathbb{R}^{\mathbb{N}}$ ή $\mathbb{R}^{\mathbb{N}^*}$ που αντιστοιχεί στο χώρο όλων των (άπειρων) ακολουθιών πραγματικών αριθμών και γενικότερα ο $\mathbb{R}^T$, όπου T κάποιο υποσύνολο των πραγματικών αριθμών.

Ας θυμηθούμε τώρα εδώ στοιχειώδεις έννοιες και αποτελέσματα σε διανυσματικούς χώρους. Σε ότι ακολουθεί, αν δε δημιουργείται σύγχυση, θα παραλείπουμε το σύμβολο του βαθμωτού γινομένου και θα γράφουμε λx αντί για $\lambda \cdot x$.

Ορισμός 2.8: Έστω V ένας $\mathbb{F}$-διανυσματικός χώρος. Καλούμε

- *γραμμικό συνδυασμό* ενός πεπερασμένου πλήθους στοιχείων του V, ας πούμε των $x_1, x_2, \ldots, x_n$, κάθε διάνυσμα της μορφής $\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n$,
- *παραγόμενο διανυσματικό χώρο* του υποσυνόλου $S \subset V$, ή *διανυσματικό χώρο που παράγεται* από το S, το σύνολο
$$\langle S \rangle := \left\{ \sum_{i=1}^{n} \lambda_i x_i \;\middle|\; \lambda_i \in \mathbb{F},\; x_i \in S,\; i = 1, \ldots, n, \quad n \geq 1 \right\},$$
δηλ. το σύνολο όλων των γραμμικών συνδυασμών όλων των πεπερασμένων υποσυνόλων του S,
- *γεννήτορα* του V, κάθε υποσύνολο $S \subset V$ που παράγει τον V, δηλ. που ικανοποιεί τη σχέση $\langle S \rangle = V$,
- το διανυσματικό χώρο (δ.χ.) V *πεπερασμένης διάστασης*, αν έχει πεπερασμένο γεννήτορα, δηλ. αν παράγεται από ένα πεπερασμένο σύνολο και το ελάχιστο πλήθος στοιχείων ενός γεννήτορα καλείται *διάσταση* του δ.χ. και συμβολίζεται με $\dim V$ ή $\dim_{\mathbb{F}} V$ αν θέλουμε να εμφανίζεται το σώμα υπεράνω του οποίου έχει αυτή τη διάσταση,
- *βάση* ενός διανυσματικού χώρου πεπερασμένης διάστασης, κάθε πεπερασμένο υποσύνολο $S \subset V$, που παράγει τον V και έχει πληθάριθμο $|S| = \dim V$.

Παρατήρηση 2.5. Η μελέτη των διανυσματικών χώρων διευκολύνεται με την έννοια της γραμμικής ανεξαρτησίας. Υπενθυμίζουμε ότι n το πλήθος διανύσματα $x_1, x_2, \ldots x_n$ λέγονται *γραμμικά ανεξάρτητα*, ή το σύνολο των διανυσμάτων αυτών καλείται γραμμικά ανεξάρτητο, αν $\dim\langle\{x_1, x_2, \ldots x_n\}\rangle = n$. Έτσι καταλήγουμε και στη συνηθισμένη διατύπωση της έννοιας της βάσης, ως ενός συνόλου διανυσμάτων που είναι γραμμικά ανεξάρτητο και γεννήτορας του διανυσματικού χώρου. Το "δημοφιλές" τέχνασμα του ελέγχου της γραμμικής ανεξαρτησίας προκύπτει από την ισχύ της παρακάτω συνεπαγωγής:
$$\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n = 0 \quad \Longrightarrow \quad \lambda_1 = \lambda_2 = \ldots = \lambda_n = 0,$$
αφού σε αντίθετη περίπτωση μπορούμε κάποιο από αυτά να το εκφράσουμε ως γραμμικό συνδυασμό των υπολοίπων, και άρα η διάσταση του διανυσματικού υπόχωρου που αυτά παράγουν είναι μικρότερη από το n. Όταν δεν είναι γραμμικά ανεξάρτητα, τα λέμε *γραμμικά εξαρτημένα*.

Παρατήρηση 2.6. Ο λόγος που ο παραγόμενος διανυσματικός χώρος ενός υποσυνόλου $S \subset V$ ορίστηκε απευθείας μέσω γραμμικών συνδυασμών είναι καθαρά χρηστικός, έχει μία απλή αναπαράσταση. Μπορεί να δειχθεί ότι αντιστοιχεί στον ελάχιστο διανυσματικό χώρο που περιέχει το υποσύνολο S. Θα μπορούσαμε λοιπόν να τον ορίσουμε, όπως και για τις παραγόμενες τοπολογίες ή τις παραγόμενες σ–άλγεβρες μέσα από τη σχέση:

$$\langle S\rangle := \bigcap\{W \mid S \subset W \text{ και } W \text{ διανυσματικός υπόχωρος του } V\}.$$

Είναι άμεσο πάλι να δειχθεί ότι η τομή οποιουδήποτε πλήθους διανυσματικών υπόχωρων του V είναι διανυσματικός χώρος και άρα ο $\langle S\rangle$ που θα περιέχει προφανώς τον S, θα είναι ο ελάχιστος διανυσματικός χώρος που θα περιέχει το S. Ο έλεγχος ότι οι ορισμοί αυτοί είναι ισοδύναμοι αφήνεται ως άσκηση (βλ. ασκ. **??**).

Είδαμε ότι ένας ομομορφισμός μεταξύ 2 ομάδων είναι μία συνάρτηση μεταξύ τους που κάνει συμβατές τις πράξεις των δύο ομάδων. Έτσι και στους διανυσματικούς χώρους, που έχουμε 2 πράξεις, αν θέλουμε μία συνάρτηση μεταξύ των 2 διανυσματικών χώρων να κάνει συμβατές τις πράξεις τους, πρέπει να είναι διπλός ομομορφισμός, των προσθετικών τους ομάδων, αλλά και των συμμετρικών υποομάδων που αυτές καθορίζουν. Φτάνουμε λοιπόν στην έννοια της γραμμικής απεικόνισης που αποτελεί το σημαντικότερο είδος συνάρτησης μεταξύ διανυσματικών χώρων.

Ορισμός 2.9: Έστω $f\colon V \to W$ μία συνάρτηση μεταξύ δύο διανυσματικών χώρων επί ενός σώματος $\mathbb{F}$. Λέμε ότι η f είναι *γραμμική*, αν για κάθε $\lambda \in \mathbb{F}$ και $x, y \in V$

$$\begin{aligned} f(x+y) &= f(x)+f(y)\\ f(\lambda\cdot x) &= \lambda\cdot f(x). \end{aligned}$$

Το σύνολο όλων των γραμμικών απεικονίσεων από τον V στον W το συμβολίζουμε με $\mathcal{L}(V,W)$ και είναι $\mathbb{F}-$διανυσματικός χώρος με πρόσθεση την κατά σημείο πρόσθεση συναρτήσεων και βαθμωτό γινόμενο τον κατά σημείο πολ/σμό του βαθμωτού με τη συνάρτηση.

- Αν η γραμμική συνάρτηση f είναι επιπλέον 1-1, τότε η f καλείται *γραμμική εμφύτευση* και λέμε ότι ο V *εμφυτεύεται γραμμικά* στον W.
- Αν η f είναι 1-1 και επί, τότε η f καλείται *γραμμικός ισομορφισμός* και οι V, W λέγονται *γραμμικά ισομορφικοί* ή *γραμμικά ισόμορφοι* και το συμβολίζουμε με $V \overset{l}{\cong} W$.
- Στην περίπτωση που $V = W$, τότε η f καλείται και *γραμμικός αυτομορφισμός*. Το σύνολο των γραμμικών αυτομορφισμών σχηματίζουν τη *γενική γραμμική ομάδα*, που συμβολίζεται με $\mathrm{GL}(V)$.
- Αν $W = \mathbb{F}$, τότε η γραμμική συνάρτηση $f\colon V \to \mathbb{F}$ καλείται *γραμμική μορφή*.

Παρατήρηση 2.7. Καθώς $\mathcal{L}(V,W) \subset W^V$ και ο W^V είναι διανυσματικός χώρος, για να δείξουμε ότι $\mathcal{L}(V,W)$ είναι διανυσματικός χώρος με τις ίδιες πράξεις του W^V, αρκεί να δείξουμε ότι $\lambda f + g \in \mathcal{L}(V,W)$, για κάθε $\lambda \in \mathbb{F}$ και $f, g \in \mathcal{L}(V,W)$. Πράγματι, εφόσον οι πράξεις ορίζονται κατά σημείο, η γραμμικότητα της συνάρτησης $\lambda f + g$ έπεται από τη γραμμικότητα του χώρου W.

Παρατήρηση 2.8. Για να ελεγχθεί η γραμμικότητα μιας απεικόνισης μεταξύ δύο διανυσματικών χώρων μπορούμε να απλοποιήσουμε τις 2 συνθήκες που απαιτούνται στον ορισμό με μία συνθήκη της μορφής $f(\lambda \cdot x + y) = \lambda \cdot f(x) + f(y)$, για κάθε $\lambda \in \mathbb{F}$ και $x, y \in V$.

Παρατήρηση 2.9. Για τους γραμμικούς ισομορφισμούς διανυσματικών χώρων ισχύει το ανάλογο της παρατήρησης 2.2 για ισομορφισμούς ομάδων. Για να κατοχυρώσουμε έναν ισομορφισμό διανυσματικών χώρων, δε χρειάζεται ο έλεγχος ότι η f^{-1} είναι γραμμική, διότι αυτό έπεται από τις ιδιότητες 1-1, επί και γραμμική της f. Πράγματι, αν $y_1, y_2 \in W$, τότε υπάρχουν μοναδικά $x_1, x_2 \in V$ με $y_1 = f(x_1)$ και $y_2 = f(x_2)$, και για τυχόν $\lambda \in \mathbb{F}$, έχουμε

$$f^{-1}(\lambda y_1 + y_2) = f^{-1}(\lambda f(x_1) + f(x_2)) = f^{-1}(f(\lambda x_1 + x_2)) = \lambda x_1 + x_2 = \lambda f^{-1}(y_1) + f^{-1}(y_2).$$

Τελικά, η παραπάνω σχέση μας δείχνει ότι η f^{-1} είναι γραμμική απεικόνιση και άρα δεν απαιτείται ως επιπλέον συνθήκη στον ορισμό του ισομορφισμού διανυσματικών χώρων.

Πρόταση 2.5: Κάθε $\mathbb{F}-$διανυσματικός χώρος V διάστασης n, είναι γραμμικά ισομορφικός με τον $\mathbb{F}^n$. Γράφουμε $V_{\mathbb{F}} \overset{l}{\cong} \mathbb{F}^n$ και ειδικότερα για $\mathbb{F} = \mathbb{R}$, έχουμε $V_{\mathbb{R}} \overset{l}{\cong} \mathbb{R}^n$.

Απόδειξη: Έστω V ένας $\mathbb{F}-$διανυσματικός χώρος διάστασης n και $\{x_1, x_2, \ldots, x_n\}$ μία βάση του, την οποία και σταθεροποιούμε. Κάθε στοιχείο του V γράφεται κατά μοναδικό τρόπο στη μορφή $x = \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n$. Θεωρούμε λοιπόν τη συνάρτηση

$$f \colon V \to \mathbb{F}^n \text{ με } x \longmapsto (\lambda_1, \lambda_2, \ldots, \lambda_n).$$

Η f είναι καλά ορισμένη, 1-1, επί και γραμμική όπως επαληθεύεται άμεσα και άρα οι V και $\mathbb{F}^n$ είναι γραμμικά ισόμορφοι. □

Το παρακάτω έπεται άμεσα από την Πρόταση 2.5.

Πόρισμα 2.3: Ο διανυσματικός χώρος $\mathbb{R}^{\{1,2,\ldots,n\}}$ είναι γραμμικά ισόμορφος με τον $\mathbb{R}^n$.

Είναι λοιπόν φανερό από τα παραπάνω ότι η σημασία του $\mathbb{R}^n$ έγκειται στο γεγονός ότι αποτελεί τον απλούστερο αντιπρόσωπο όλων των $\mathbb{R}$-διανυσματικών χώρων διάστασης n, με τους οποίους είναι γραμμικά ισόμορφος. Για να απλοποιήσουμε και την περιγραφή των στοιχείων του ορίζουμε και τη συνήθη βάση του $\mathbb{R}^n$.

Ορισμός 2.10: Λέμε *συνήθη βάση* του $\mathbb{R}^n$ το σύνολο των διανυσμάτων $\{e_1, e_2, \ldots, e_n\}$, όπου

$$e_i = \underbrace{(0, \ldots, 0, 1, 0, \ldots, 0)}_{1 \text{ στην } i \text{ θέση του διανύσματος}}, \quad i = 1, 2, \ldots, n.$$

Αν $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, τότε γράφεται στη μορφή $x = \sum_{i=1}^n x_i\, e_i$. Οι πραγματικοί αριθμοί $\{x_i\}_{i=1}^n$, που είναι οι $i-$συνιστώσες του x, λέγονται και *συντεταγμένες* του διανύσματος x ως προς τη συνήθη βάση του $\mathbb{R}^n$.

Με τη βοήθεια των γραμμικών απεικονίσεων μπορούμε να επεκταθούμε σε διγραμμικές απεικονίσεις και τετραγωνικές μορφές που έχουν πολλές εφαρμογές στη Στατιστική.

Ορισμός 2.11: Έστω V, W, Z διανυσματικοί χώροι επί ενός κοινού σώματος $\mathbb{F}$. Μία απεικόνιση $B\colon V \times W \to Z$ καλείται *διγραμμική*, αν

(i) η $B_v\colon W \to Z$ με $B_v(w) = B(v, w)$ είναι γραμμική, για κάθε $v \in V$,
(ii) η $B_w\colon V \to Z$ με $B_w(v) = B(v, w)$ είναι γραμμική, για κάθε $w \in W$.

Η διγραμμική απεικόνιση B καλείται

- *συμμετρική*, αν $W = V$ και $B(v, w) = B(w, v)$ για κάθε $v, w \in V$,
- *διγραμμικό γινόμενο*, αν $Z = W = V$ και τότε η $B\colon V \times V \to V$ καθορίζει μία επιπρόσθετη εσωτερική πράξη στο δ.χ. V,
- *διγραμμική μορφή*, αν $W = V$ και $Z = \mathbb{F}$. Αν η B είναι επιπλέον συμμετρική, τότε καλείται *συμμετρική διγραμμική μορφή*.

Επιπλέον, μία απεικόνιση $Q\colon V \to \mathbb{F}$ καλείται *τετραγωνική μορφή*, αν

$$Q(\lambda v) = \lambda^2\, Q(v), \quad \text{για κάθε } \lambda \in \mathbb{F}, v \in V.$$

Παρατήρηση 2.10. Κάθε διγραμμική απεικόνιση $B\colon V \times W \to Z$ ικανοποιεί:

$$\begin{aligned} B(\lambda v, w) &= B(v, \lambda w) = \lambda B(v, w), \\ B(v_1 + v_2, w) &= B(v_1, w) + B(v_2, w), \\ B(v, w_1 + w_2) &= B(v, w_1) + B(v, w_2), \end{aligned}$$

για κάθε $\lambda \in \mathbb{F}$, $v_1, v_2 \in V$, $w_1, w_2 \in W$.

Παράδειγμα 2.6: Αν $\mathcal{L}(V, W)$ και $\mathcal{L}(W, Z)$ είναι οι δ.χ. των γραμμικών απεικονίσεων μεταξύ των παραπάνω δ.χ., τότε η σύνθεση απεικονίσεων:

$$\circ : \mathcal{L}(V, W) \times \mathcal{L}(W, Z) \to \mathcal{L}(V, Z) \quad \text{με} \quad (f, g) \mapsto g \circ f$$

είναι διγραμμική, ενώ αν $Z = W = V$, τότε είναι και διγραμμικό γινόμενο.

Οι δ.χ. μπορούν να εφοδιαστούν με επιπρόσθετες εσωτερικές πράξεις. Στην περίπτωση που η πράξη είναι διγραμμική, οδηγούμαστε στην κατασκευή αλγεβρών.

Ορισμός 2.12: Έστω V ένας $\mathbb{F}-$διανυσματικός χώρος και $\star : V \times V \to V$ μία επιπρόσθετη εσωτερική πράξη στο V. Αν η πράξη είναι διγραμμικό γινόμενο, τότε ο V καλείται *$\mathbb{F}-$άλγεβρα* ή *άλγεβρα επί του* $\mathbb{F}$.

Αν θεωρήσουμε τους δ.χ. εφοδιασμένους με τη συνήθη βάση τους, τότε μπορούμε να δώσουμε μία απλή αναπαράσταση των γραμμικών απεικονίσεων μεταξύ δ.χ. με τη βοήθεια πινάκων, οι οποίοι αποτελούν ένα πολύ δυναμικό εργαλείο.

Ορισμός 2.13: Έστω $(\mathbb{F},+,*)$ ένα σώμα. Κάθε ορθογώνια διευθέτηση $m \cdot n$ στοιχείων του $\mathbb{F}$ σε m γραμμές και n στήλες καλείται *πίνακας* (διάστασης) $m \times n$. Το σύνολο όλων των $m \times n$ πινάκων με στοιχεία του $\mathbb{F}$ συμβολίζεται με $\mathbb{F}^{m\times n}$. Αν $A \in \mathbb{F}^{m\times n}$, τότε γράφουμε $A = (a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}}$. Αν ορίσουμε ως *πρόσθεση πινάκων*, την πρόσθεση κατά συντεταγμένη για κάθε ζεύγος (i,j), δηλ. $A+B := (a_{ij}+b_{ij})$ και ως βαθμωτό γινόμενο $\lambda \cdot A = (\lambda * a_{ij})$, τότε το σύνολο όλων των πινάκων διάστασης $m \times n$ γίνεται $\mathbb{F}-$διανυσματικός χώρος και αν δεν υπάρχει σύγχυση ως προς το σώμα το συμβολίζουμε και με $\mathcal{M}_{m,n}$. Αν $m = n$, τότε το σύνολο όλων των $n \times n$ τετραγωνικών πινάκων το συμβολίζουμε επίσης και $\mathcal{M}_n$.

Η παρουσίαση κάποιων αποτελεσμάτων που θα δώσουμε παρακάτω και σχετίζονται με ακριβείς ή ασυμπτωτικές κατανομές τυχαίων διανυσμάτων, διευκολύνεται με την εισαγωγή δύο ειδών γραμμικών απεικονίσεων, που αφορούν τα διαδοχικά μερικά αθροίσματα και τις διαδοχικές μερικές διαφορές των στοιχείων ενός διανύσματος.

Ορισμός 2.14: Συμβολίζουμε με $s, \delta \colon \mathbb{R}^m \to \mathbb{R}^m$ τις συναρτήσεις με

$$s(x) = (x_1, x_1 + x_2, \ldots, x_1 + x_2 + \cdots + x_m)^\top, \tag{2.1}$$

και

$$\delta(x) = (x_1, x_2 - x_1, \ldots, x_m - x_{m-1})^\top, \tag{2.2}$$

που αντιστοιχίζουν σε κάθε διάνυσμα $x \in \mathbb{R}^m$ το διάνυσμα των διαδοχικών μερικών αθροισμάτων και διαφορών αντίστοιχα.

Παρατήρηση 2.11. Είναι φανερό ότι οι s, δ είναι γραμμικές απεικονίσεις και ως αντίστροφες διαδικασίες ικανοποιούν:

$$s \circ \delta = \delta \circ s = \mathrm{id}_{\mathbb{R}^m},$$

ενώ αν S, Δ είναι οι πίνακες που τους αντιστοιχούν, τότε

$$S = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix} \qquad \text{και} \qquad \Delta = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{pmatrix},$$

και προφανώς ικανοποιούν ότι $\Delta = S^{-1}$ και $S = \Delta^{-1}$.

Το επόμενο αποτέλεσμα θα φανεί επίσης χρήσιμο.

Λήμμα 2.1: Αν $x = (x_1, \ldots, x_m)^\top$, τότε

(i)
$$S\,\mathrm{dg}(x)\,S^\top = \big(s_{i\wedge j}(x)\big)_{i,j},$$

(ii)
$$S\,x\,x^\top\,S^\top = \big(s_i(x)\,s_j(x)\big)_{i,j}.$$

Απόδειξη: (i) Προκύπτει άμεσα ως εξής:

$$S\,\mathrm{dg}(x)\,S^\top = \begin{pmatrix} x_1 & x_1 & x_1 & \cdots & x_1 \\ x_1 & x_1+x_2 & x_1+x_2 & \cdots & x_1+x_2 \\ x_1 & x_1+x_2 & x_1+x_2+x_3 & \cdots & x_1+x_2+x_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1 & x_1+x_2 & x_1+x_2+x_3 & \cdots & x_1+x_2+\cdots+x_m \end{pmatrix} = \big(s_{i\wedge j}(x)\big)_{i,j}.$$

(ii) Παρατηρούμε ότι

$$S\,x\,x^\top\,S^\top = s(x)\,s^\top(x) = \big(s_i(x)\,s_j(x)\big)_{i,j}.$$

□

2.3 Πίνακες Ορθογώνιας Προβολής

Στην ενότητα αυτή κάνουμε μία υπενθύμιση εννοιών από τη Γραμμική Άλγεβρα που συνδέονται με το ευθύ άθροισμα διανυσματικών υπόχωρων και τους πίνακες ορθογώνιας προβολής.

Ορισμός 2.15 (ευθύ άθροισμα): Έστω $\{V_i\}_{i=1}^p$, p διανυσματικοί υπόχωροι του $\mathbb{R}^n$.

- Οι V_i και V_j, $i \neq j$, καλούνται *κάθετοι*, και γράφουμε $V_i \perp V_j$, αν για κάθε $v_i \in V_i$ και $v_j \in V_j$, έχουμε $v_i \perp v_j$, ή ισοδύναμα, $\langle v_i, v_j\rangle = v_i^\top v_j = 0$.
- Οι υπόχωροι $\{V_i\}_{i=1}^p$ καλούνται *κάθετοι μεταξύ τους*, αν είναι κάθετοι ανά δύο, δηλ., $V_i \perp V_j$ για κάθε $i \neq j$.
- Το σύνολο
$$V_1 + \ldots + V_p = \big\{v_1 + \ldots + v_p \;:\; v_i \in V_i,\ \text{για κάθε } 1 \leq i \leq p\big\},$$
καλείται *άθροισμα* των διανυσματικών υπόχωρων $\{V_i\}_{i=1}^p$ και δείχνεται άμεσα ότι αποτελεί διανυσματικό υπόχωρο του $\mathbb{R}^n$.
- Το άθροισμα των $\{V_i\}_{i=1}^p$ καλείται *ευθύ*, και το συμβολίζουμε $V_1 \oplus \cdots \oplus V_p$, αν οι υπόχωροι είναι κάθετοι μεταξύ τους.

- Το σύνολο

$$V^{\perp} := \left\{u \in \mathbb{R}^n \ : \ \langle u, v\rangle = 0, \text{ για κάθε } v \in V\right\},$$

καλείται *ορθογώνιος υπόχωρος* ή *ορθογώνιο συμπλήρωμα* του V και αποδεικνύεται άμεσα ότι είναι πράγματι διανυσματικός χώρος και ότι $\mathbb{R}^n = V \oplus V^{\perp}$.

Στην επόμενη πρόταση δίνουμε μία στοιχειώδη ιδιότητα που συνδέεται με την ανάλυση ενός διανύσματος v πάνω σε ένα ευθύ άθροισμα διανυσματικών υπόχωρων.

Πρόταση 2.6: Έστω V και $\{V_i\}_{i=1}^{p}$ διανυσματικοί υπόχωροι του $\mathbb{R}^n$. Αν $V = V_1 \oplus \ldots \oplus V_p$, τότε για κάθε $v \in V$ υπάρχουν μοναδικά $v_i \in V_i$, $1 \leq i \leq p$, τέτοια ώστε $v = v_1 + \ldots + v_p$.

Στον επόμενο ορισμό υπενθυμίζουμε έννοιες σχετικές με την προβολή διανυσμάτων σε διανυσματικούς υπόχωρους.

Ορισμός 2.16 (ορθογώνια προβολή):
- Έστω $u, v \in \mathbb{R}^n$. Το μοναδικό διάνυσμα w που είναι παράλληλο (συγγραμμικό) με το v, έτσι ώστε το w και το $u - w$ να είναι κάθετα καλείται *ορθογώνια προβολή του u στο διάνυσμα v* και συμβολίζεται $\mathbf{pr}_v u$,
- Έστω $v \in \mathbb{R}^n$ και $V \leq \mathbb{R}^n$. Το μοναδικό διάνυσμα w που ανήκει στον V με την ιδιότητα το w και το $u - w$ να είναι κάθετα καλείται *ορθογώνια προβολή του u στο διανυσματικό χώρο V* και συμβολίζεται $\mathbf{pr}_V u$.
- Έστω $V \leq \mathbb{R}^n$. Η γραμμική απεικόνιση $\pi_V : \mathbb{R}^n \to \mathbb{R}^n$, όπου

$$\pi_V(u) = \mathbf{pr}_V u$$

καλείται *συνάρτηση ορθογώνιας προβολής* στο διανυσματικό υπόχωρο V. Ο αντίστοιχος πίνακας Π_V της γραμμικής απεικόνισης π_V καλείται *πίνακας ορθογώνιας προβολής* στον V.

Οι επόμενες ιδιότητες των ορθογώνιων προβολών προκύπτουν εύκολα.

Πρόταση 2.7: Έστω $u, v \in \mathbb{R}^n$ και $V \leq \mathbb{R}^n$.

(i)

$$\mathbf{pr}_v u = \frac{\langle u, v\rangle}{\langle v, v\rangle} v.$$

(ii) Αν $\{v_i\}_{i=1}^d$ είναι μία ορθογώνια βάση του V, τότε

$$\mathbf{pr}_V u = \sum_{i=1}^d \mathbf{pr}_{v_i} u = \sum_{i=1}^d \frac{\langle u, v_i\rangle}{\langle v_i, v_i\rangle} v_i.$$

Αν η βάση είναι επιπλέον ορθοκανονική, τότε

$$\mathbf{pr}_V u = \sum_{i=1}^d \langle u, v_i\rangle\, v_i.$$

Δίνουμε τώρα έναν χαρακτηριστικό κατάλογο ιδιοτήτων των πινάκων ορθογώνιας προβολής που θα χρειαστούμε στη συνέχεια.

Πρόταση 2.8: Έστω $\mathbb{R}^n = V_1 \oplus \ldots \oplus V_p$, όπου $V_i \leq \mathbb{R}^n$ διάστασης d_i και Π_i ο πίνακας της ορθογώνιας προβολής στον V_i, $1 \leq i \leq p$. Ισχύουν τα παρακάτω:

(i) $\Pi_i^2 = \Pi_i$, $i = 1, \ldots, p$,

(ii) $\Pi_i \Pi_j = 0_{n\times n}$, για κάθε $i \neq j$,

(iii) $\Pi_1 + \ldots + \Pi_p = I_n$,

(iv) αν $\{v_{ij}\}_{j=1}^{d_i}$ είναι ορθοκανονική βάση του V_i και B_i ο πίνακας με στήλες τα διανύσματα της παραπάνω ορθοκανονικής βάσης, τότε $\Pi_i = B_i B_i^\top$,

(v) ο πίνακας Π_i είναι συμμετρικός και άρα ορθογώνια διαγωνοποιήσιμος στη μορφή $\Pi_i = \Lambda_i D_i \Lambda_i^\top$, όπου Λ_i ορθογώνιος πίνακας, δηλ. $\Lambda_i^\top = \Lambda_i^{-1}$ και D_i διαγώνιος πίνακας με d_i το πλήθος 1 και $n - d_i$ το πλήθος 0.

Απόδειξη: Τα (i)–(iii) προκύπτουν άμεσα. Αποδεικνύουμε τώρα το (iv). Αρκεί να δείξουμε ότι για κάθε $u \in \mathbb{R}^n$, $\Pi_i u = B_i B_i^\top u$. Έστω $u \in \mathbb{R}^n$. Από την υπόθεση και την Πρόταση 2.8–(ii)

$$\begin{aligned}
\Pi_i u &= \langle u, v_{i1}\rangle v_{i1} + \ldots + \langle u, v_{id_i}\rangle v_{id_i} \\
&= \underbrace{\begin{pmatrix} v_{i1} & \ldots & v_{id_i}\end{pmatrix}}_{B_i} \cdot \begin{pmatrix} \langle u, v_{i1}\rangle \\ \vdots \\ \langle u, v_{id_i}\rangle \end{pmatrix} = B_i \cdot \begin{pmatrix} v_{i1}^\top \\ \vdots \\ v_{id_i}^\top \end{pmatrix} \cdot u = B_i B_i^\top u,
\end{aligned}$$

και έτσι καταλήγουμε στο ζητούμενο. Για το (v), κατ'αρχάς είναι φανερό ότι ο Π είναι συμμετρικός από τη σχέση $\Pi = B_i B_i^\top$, άρα και ορθογώνια διαγωνοποιήσιμος στη μορφή $\Pi_i = \Lambda_i D_i \Lambda_i^\top$, όπου Λ_i πίνακας με $\Lambda_i^\top = \Lambda_i^{-1}$ και D_i διαγώνιος πίνακας με στοιχεία τις πραγματικές ιδιοτιμές του Π_i. Επίσης, από τη σχέση (i) είναι φανερό ότι ο A ως ταυτοδύναμος έχει δυνατές ιδιοτιμές μόνο 0 ή 1 και καθώς η διάστασή του είναι d_i έχει ακριβώς d_i ίσες με 1 και τις υπόλοιπες μηδενικές.

□

Στη συνέχεια τοπολογικοί και μετρήσιμοι διανυσματικοί χώροι.

2.4 Τυχαία Διανύσματα

Ξεκινάμε, λοιπόν, με τον τυπικό ορισμό ενός τυχαίου διανύσματος, την έννοια της κατανομής και την ισονομία τυχαίων διανυσμάτων.

Ορισμός 2.17 (τυχαίο διάνυσμα-κατανομή-ισονομία):
Έστω $(\Omega, \mathcal{A}, \mathbb{P})$ ένας χώρος πιθανότητας και $d \geq 2$.

1. Μία συνάρτηση $X \colon \Omega \to \mathbb{R}^d$ καλείται *τυχαίο διάνυσμα* (τ.δ.) ή *διανυσματική τ.μ.* (δ.τ.μ.), αν $X^{-1}(B) \in \mathcal{A}$ για κάθε $B \in \mathcal{B}(\mathbb{R}^d)$.
2. Το μέτρο πιθανότητας $\mathbb{P}_X \colon \mathcal{B}(\mathbb{R}^d) \to [0,1]$ με $\mathbb{P}_X(B) = \mathbb{P}(X \in B) = \mathbb{P}\left(X^{-1}(B)\right)$ καλείται *κατανομή* της X, όπως στη μονοδιάστατη περίπτωση.
3. Δύο τυχαία διανύσματα X, Y (όχι κατ'ανάγκη ορισμένα στον ίδιο χώρο πιθανότητας) καλούνται *ισόνομα* αν $\mathbb{P}_X = \mathbb{P}_Y$ και τότε γράφουμε $X \stackrel{d}{=} Y$.

Η έννοια της συνάρτησης κατανομής επεκτείνεται φυσιολογικά και για τυχαία διανύσματα. Σε ό,τι ακολουθεί, θεωρούμε το $\mathbb{R}^d$ εφοδιασμένο με τη *διάταξη κατά συνιστώσες*, που είναι η σχέση μερικής διάταξης (βλ. παράρτημα B, **??**) που προκύπτει μέσω της σχέσης $x \leq y$ αν, και μόνο αν, $x_i \leq y_i$, για κάθε $1 \leq i \leq d$.

Ορισμός 2.18: Έστω X ένα τυχαίο διάνυσμα με τιμές στον $\mathbb{R}^d$. Η συνάρτηση $F_X \colon \mathbb{R}^d \to [0,1]$, όπου

$$F_X(x) := \mathbb{P}(X \leq x) = \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_d \leq x_d) = \mathbb{P}\left(\bigcap_{i=1}^{d} \{X_i \leq x_i\}\right),$$

καλείται *συνάρτηση κατανομής* του τ.δ. X ή *από κοινού συνάρτηση κατανομής* των $X_1, X_2, \ldots, X_d$ όταν θέλουμε να δώσουμε έμφαση στις συνιστώσες τ.μ.. Σε αυτήν την περίπτωση γράφεται και ως $F_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$.

Δίνουμε παρακάτω χωρίς απόδειξη τρεις χρήσιμους χαρακτηρισμούς της ισονομίας τυχαίων διανυσμάτων. Η πλήρης απόδειξη γίνεται σε μαθήματα θεωρητικών Πιθανοτήτων.

Θεώρημα 2.1: Δύο τυχαία διανύσματα X,Y είναι ισόνομα αν, και μόνο αν,

1. $F_X = F_Y$, δηλαδή έχουν κοινή συνάρτηση κατανομής.
2. $u^\top X \stackrel{d}{=} u^\top Y$, για κάθε $u \in \mathbb{R}^d$, δηλαδή κάθε γραμμικός συνδιασμός των συνιστωσών της X είναι ισόνομος με τον αντίστοιχο γραμμικό συνδυασμό των συνιστωσών της Y.
3. $\mathbb{E}[f(X)] = \mathbb{E}[f(Y)]$, για κάθε $f : \mathbb{R}^d \to \mathbb{R}$ συνεχή και φραγμένη συνάρτηση.

Απόδειξη: (σκιαγράφηση)

1. Αν $X \overset{d}{=} Y$, τότε $\mathbb{P}_X = \mathbb{P}_Y$ και άρα για κάθε $a = (a_1, \ldots, a_d) \in \mathbb{R}^d$ έχουμε
$$F_X(a) = \mathbb{P}(X \leq a) = \mathbb{P}_X(\Delta_a) = \mathbb{P}_Y(\Delta_a) = \mathbb{P}(Y \leq a) = F_Y(a),$$
όπου
$$\Delta_a = \prod_{i=1}^{d}(-\infty, a_i]. \tag{2.3}$$
Συμπεραίνουμε λοιπόν ότι $F_X = F_Y$. Αντίστροφα, θέτουμε
$$\mathcal{D} := \{\Delta_a \mid a \in \mathbb{R}^d\} \tag{2.4}$$
την οικογένεια όλων των διαστημάτων της μορφής (2.3). Αν $F_X = F_Y$, τότε σύμφωνα με τα παραπάνω, τα μέτρα πιθανότητας $\mathbb{P}_X$ και $\mathbb{P}_Y$ συμφωνούν στην οικογένεια υποσυνόλων $\mathcal{D}$. Πρέπει τώρα να δειχθεί ότι η ισότητα αυτή επεκτείνεται σε όλα τα $\mathcal{B}(\mathbb{R}^d)$ και άρα $\mathbb{P}_X = \mathbb{P}_Y$. Ικανές συνθήκες είναι οι εξής:
   (α) $\sigma(\mathcal{D}) = \mathcal{B}(\mathbb{R}^d)$, δηλ. η $\mathcal{D}$ είναι γεννήτορας της $\mathcal{B}(\mathbb{R}^d)$,
   (β) η οικογένεια $\mathcal{D}$ είναι κλειστή στις πεπερασμένες τομές,
   (γ) υπάρχει αύξουσα ακολουθία $(\Delta_n)_{n\geq 1}$ στην $\mathcal{D}$ με $\mathbb{R}^d = \cup_{n\geq 1}\Delta_n$.

   Η επαλήθευση αυτών των ιδιοτήτων οδηγεί τελικά στο ζητούμενο.
2. Αν $X \overset{d}{=} Y$, τότε $f(X) \overset{d}{=} f(Y)$ για οποιαδήποτε Borel μετρήσιμη συνάρτηση $f\colon \mathbb{R}^d \to \mathbb{R}$. Ένας γραμμικός συνδυασμός αντιστοιχεί σε μία γραμμική απεικόνιση και άρα Borel μετρήσιμη ως συνεχής. Το αντίστροφο είναι πιο απαιτητικό και το δεχόμαστε (βγαίνει άμεσα αν χρησιμοποιηθούν χαρακτηριστικές συναρτήσεις).
3. Αν $X \overset{d}{=} Y$, τότε $f(X) \overset{d}{=} f(Y)$ για οποιαδήποτε Borel μετρήσιμη συνάρτηση $f\colon \mathbb{R}^d \to \mathbb{R}$, άρα και για τις συνεχείς και φραγμένες. Η μέση τιμή είναι ιδιότητα της κατανομής και έτσι συμπεραίνουμε ότι $\mathbb{E}[f(X)] = \mathbb{E}[f(Y)]$. Το αντίστροφο μπορεί να δειχθεί αποδεικνύοντας ότι η απόδοση μέσης τιμής στην οικογένεια των συνεχών και φραγμένων συναρτήσεων χαρακτηρίζει πλήρως τις μέσες τιμές των δεικτριών συναρτήσεων $\mathbf{1}_\Delta$, όπου $\Delta \in \mathcal{D}$, την οικογένεια των διαστημάτων που ορίζεται στη σχέση (2.4). Αυτά προσδιορίζουν πλήρως τη συνάρτηση κατανομής και έτσι προκύπτει το ζητούμενο.

□

Παρατήρηση 2.12. Ο πρώτος χαρακτηρισμός είναι φυσιολογική γενίκευση του αντίστοιχου χαρακτηρισμού για μονοδιάστατες τ.μ.. Ο δεύτερος είναι αρκετά χρήσιμος σε αποδείξεις και δίνει τη δυνατότητα χαρακτηρισμού της κατανομής ενός τ.δ. μέσω της γνώσης μιας (άπειρης) οικογένειας κατανομών μονοδιάστατων τ.μ. που προκύπτουν μάλιστα μόνο μέσω γραμμικών μετασχηματισμών. Είναι μάλιστα θεωρητικά πιο γόνιμος, αφού επεκτείνεται και σε απειροδιάστατους διανυσματικούς χώρους με νόρμα μέσω της έννοιας του γραμμικού φραγμένου τελεστή (αντικείμενο μελέτης της Συναρτησιακής Ανάλυσης). Ο τελευταίος χαρακτηρισμός είναι ο πιο γενικός και μπορεί να χρησιμοποιηθεί γενικότερα σε μετρικούς χώρους.

Οι έννοιες της μέσης τιμής και της διασποράς, όπως και της ροπής οποιασδήποτε τάξης, επεκτείνονται σε τυχαία διανύσματα. Θα περιοριστούμε εδώ μόνο στη μέση τιμή και τη διασπορά, που είναι και οι πιο χρήσιμες. Είναι φανερό από τον χαρακτηρισμό της κατανομής που δίνεται στο Θεώρημα 2.1–(ii) ότι μία επέκταση κάποιας έννοιας στην πολυδιάστατη περίπτωση θα μπορούσε να εμπνέεται απο την εφαρμογή της αντίστοιχης έννοιας στη μονοδιάστατη περίπτωση σε όλους τους γραμμικούς συνδυασμούς $u^\top X$. Λόγω της γραμμικότητας της μέσης τιμής, η μέση τιμή $\mathbb{E}(u^\top X)$ καθορίζεται πλήρως από το διάνυσμα των μέσων τιμών $\big(\mathbb{E}(X_1), \ldots, \mathbb{E}(X_d)\big)$. Στην περίπτωση της διασποράς, δεν αρκεί η γνώση των συνιστωσών διασπορών, αλλά χρειάζεται η γνώση και των συνδιακυμάνσεων $Cov(X_i, X_j)$ που θα συμβολίζουμε για απλότητα από εδώ και στο εξής ως $\mathbb{C}(X_i, X_j)$. Έτσι, με τον πίνακα διασπορώνσυνδιασπορών καθορίζεται πλήρως η διασπορά $\mathbb{V}(u^\top X)$, δηλαδή η διασπορά κάθε γραμμικού συνδυασμού της X. Οδηγούμαστε, λοιπόν, στους επόμενους ορισμούς.

Ορισμός 2.19: Έστω X ένα τυχαίο διάνυσμα με τιμές στον $\mathbb{R}^d$.

- τότε λέμε ότι το X *έχει μέση τιμή* $\mathbb{E}(X)$ που ορίζεται από τη σχέση:

$$\mu_X \equiv \mathbb{E}(X) := \big(\mathbb{E}(X_1), \mathbb{E}(X_2), \ldots, \mathbb{E}(X_d)\big) .$$

Η συνθήκη ύπαρξης μέσης τιμής είναι ισοδύναμη με τη συνθήκη $\mathbb{E}\|X\| < +\infty$, όπου $\|\cdot\|$ είναι η ευκλείδεια νόρμα στον $\mathbb{R}^d$, την οποία και θα χρησιμοποιούμε για απλότητα.

(ii) Αν $\mathbb{E}(X_i^2) < +\infty$ για κάθε $1 \le i \le d$, τότε ορίζουμε ως *διασπορά* του X τον πίνακα:

$$\Sigma_X \equiv \mathbb{V}(X) := \big(\mathbb{C}(X_i, X_j)\big)_{1\le i,j\le d} .$$

Ο πίνακας $\mathbb{V}(X)$ καλείται και *πίνακας διασπορών-συνδιασπορών* (για εμφανείς λόγους) ή *πίνακας συνδιακύμανσης* και συμβολίζεται και με $\mathrm{Cov}(X)$. Η συνθήκη ύπαρξης της διασποράς είναι ισοδύναμη με τη συνθήκη $\mathbb{E}\|X\|^2 < +\infty$, την οποία και θα χρησιμοποιούμε για απλότητα.

(iii) Αν Y είναι ένα ακόμα τυχαίο διάνυσμα με τιμές στον $\mathbb{R}^s$ και $\mathbb{E}\|Y\|^2 < +\infty$, τότε ορίζουμε ως (σταυρωτή) *συνδιακύμανση* των τ.δ. X και Y, τον πίνακα

$$\Sigma_{X,Y} \equiv \mathbb{C}(X, Y) := \begin{pmatrix} \mathbb{C}(X_1, Y_1) & \mathbb{C}(X_1, Y_2) & \ldots & \mathbb{C}(X_1, Y_s) \\ \mathbb{C}(X_2, Y_1) & \mathbb{C}(X_2, Y_2) & \ldots & \mathbb{C}(X_2, Y_s) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{C}(X_d, Y_1) & \mathbb{C}(X_d, Y_2) & \ldots & \mathbb{C}(X_d, Y_s). \end{pmatrix}$$

Τα τ.δ. X και Y λέγονται *ασυσχέτιστα*, αν ο $\mathbb{C}(X, Y)$ είναι ο μηδενικός πίνακας.

Παρατήρηση 2.13. Από τους παραπάνω ορισμούς, είναι φανερό ότι $\mathbb{C}(X, X) = \mathbb{V}(X)$ και $\mathbb{C}(Y, X) = \mathbb{C}(X, Y)^\top$. Παρατηρούμε λοιπόν εδώ ότι ο πίνακας συνδιακύμανσης μεταξύ δύο τ.δ. (έτσι όπως τον ορίσαμε) δεν είναι γενικά συμμετρικός, εκτός αν $X = Y$, οπότε και ταυτίζεται με τον συμμετρικό πίνακα $\mathbb{V}(X)$.

Στην επόμενη πρόταση δίνουμε κάποιες στοιχειώδεις ιδιότητες της μέσης τιμής τυχαίων διανυσμάτων.

Πρόταση 2.9: Έστω X, Y τυχαία διανύσματα με τιμές στον $\mathbb{R}^d$ με $\mathbb{E}\|X\|, \mathbb{E}\|Y\| < +\infty$ και A, b πίνακας και διάνυσμα κατάλληλης διάστασης. Τότε, ισχύουν τα εξής:

$$\begin{aligned} \mathbb{E}(AX+b) &= A\,\mathbb{E}(X)+b, \\ \mathbb{E}(X+Y) &= \mathbb{E}(X)+\mathbb{E}(Y). \end{aligned}$$

Στην επόμενη πρόταση δίνουμε στοιχειώδεις ιδιότητες της διασποράς τυχαίων διανυσμάτων.

Πρόταση 2.10: Έστω $X, Y \in \mathbb{R}^d$ τυχαία διανύσματα με $\mathbb{E}\|X\|^2, \mathbb{E}\|Y\|^2 < +\infty$ και A, b πίνακας και διάνυσμα κατάλληλης διάστασης. Τότε, ισχύουν τα εξής:

$$\begin{aligned} \mathbb{V}(X) &= \mathbb{E}\left((X-\mu_X)(X-\mu_X)^\top\right) = \mathbb{E}(XX^\top) - \mathbb{E}(X)\,\mathbb{E}(X^\top), \\ \mathbb{V}(AX+b) &= A\,\mathbb{V}(X)\,A^\top, \\ \mathbb{V}(X+Y) &= \mathbb{V}(X)+\mathbb{V}(Y)+\mathbb{C}(X,Y)+\mathbb{C}(Y,X). \end{aligned}$$

Ακολουθούν ιδιότητες που συνδέονται με τη συνδιακύμανση δύο τυχαίων διανυσμάτων.

Πρόταση 2.11: Έστω X, Y τυχαία διανύσματα με τιμές στον $\mathbb{R}^d$ και $\mathbb{R}^s$ αντίστοιχα. Αν υποθέσουμε ότι $\mathbb{E}\|X\|^2, \mathbb{E}\|Y\|^2 < +\infty$, τότε ισχύουν τα εξής:

$$\begin{aligned} \mathbb{C}(X,Y) &= \mathbb{E}\left((X-\mu_X)(Y-\mu_Y)^\top\right) = \mathbb{E}(XY^\top) - \mathbb{E}(X)\,\mathbb{E}(Y^\top), \\ \mathbb{C}(AX+b, BY+c) &= A\,\mathbb{C}(X,Y)\,B^\top. \end{aligned}$$

Αν επιπλέον, Z, W τυχαία διανύσματα με τιμές στον $\mathbb{R}^d$ και $\mathbb{R}^s$ αντίστοιχα με $\mathbb{E}\|Z\|^2, \mathbb{E}\|W\|^2 < +\infty$, τότε

$$\mathbb{C}(X+Z, Y+W) = \mathbb{C}(X,Y)+\mathbb{C}(X,W)+\mathbb{C}(Z,Y)+\mathbb{C}(Z,W).$$

Ορισμός 2.20 (ανεξαρτησία τυχαίων διανυσμάτων):
Δύο τ.δ. $X \in \mathbb{R}^d$, $Y \in \mathbb{R}^s$ καλούνται *ανεξάρτητα*, αν για κάθε $A \in \mathcal{B}(\mathbb{R}^d)$ και $B \in \mathcal{B}(\mathbb{R}^s)$

$$\mathbb{P}(X \in A,\ Y \in B) = \mathbb{P}(X \in A)\,\mathbb{P}(Y \in B). \tag{2.5}$$

Χρησιμοποιείται αρκετά ο συμβολισμός $X \perp\!\!\!\perp Y$ για να δηλώσουμε την ανεξαρτησία των X και Y.

Μπορούν να αποδειχθούν και οι επόμενοι χρήσιμοι χαρακτηρισμοί της ανεξαρτησίας τ.δ..

Πρόταση 2.12: Δύο τ.δ. $X \in \mathbb{R}^d$, $Y \in \mathbb{R}^s$ είναι ανεξάρτητα αν, και μόνο αν,

- $$F_{X,Y}(x,y) = F_X(x)F_Y(y),$$

  όπου ως $F_{X,Y}$ συμβολίζουμε την από κοινού συνάρτηση κατανομής των τ.δ. X, Y,

(ii) $u^\top X \perp\!\!\!\perp v^\top Y$, για κάθε $u \in \mathbb{R}^d$, $v \in \mathbb{R}^s$, δηλαδή κάθε γραμμικός συνδυασμός των συνιστωσών του X είναι ανεξάρτητος από κάθε γραμμικό συνδυασμό των συνιστωσών του Y,

(iii) για κάθε $f : \mathbb{R}^d \to \mathbb{R}$, $g : \mathbb{R}^s \to \mathbb{R}$ συνεχείς και φραγμένες,

$$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\,\mathbb{E}[g(Y)]. \tag{2.6}$$

Η παραπάνω σχέση ισχύει και για οποιεσδήποτε f, g (μετρήσιμες), αρκεί να υπάρχουν οι μέσες τιμές.

Από τα παραπάνω, συμπεραίνουμε τα εξής:

Πόρισμα 2.4: Αν τα τ.δ. $X \in \mathbb{R}^d, Y \in \mathbb{R}^s$ είναι ανεξάρτητα, τότε, όταν ορίζονται καλά οι ακόλουθες ποσότητες, ισχύουν τα εξής:

$$\begin{aligned}
\mathbb{E}(XY^\top) &= \mathbb{E}(X)\,\mathbb{E}(Y^\top), \\
\mathbb{C}(X,Y) &= 0_{d\times s} \qquad (\Rightarrow\ X,Y \text{ ασυσχέτιστες}), \\
\mathbb{V}(X+Y) &= \mathbb{V}(X)+\mathbb{V}(Y), \qquad \text{για } s=d,
\end{aligned}$$

όπου $0_{d\times s}$ είναι ο μηδενικός πίνακας διάστασης $d \times s$.

Στη συνέχεια, θα αναφέρουμε και θα μελετήσουμε τις στοιχειώδεις ιδιότητες της πιο σημαντικής διακριτής και της πιο σημαντικής συνεχούς δ.τ.μ., που είναι η πολυωνυμική και η πολυδιάστατη κανονική τ.μ., αντίστοιχα.

2.5 Πολυωνυμική Κατανομή

Η πολυωνυμική κατανομή γενικεύει τη διωνυμική κατανομή. Είναι γνωστό ότι μία τ.μ. που ακολουθεί τη διωνυμική κατανομή ερμηνεύεται ως το πλήθος των επιτυχιών σε ένα δεδομένο αριθμό ανεξάρτητων δοκιμών *Bernoulli* με σταθερή πιθανότητα επιτυχίας σε κάθε δοκιμή. Αν η επιτυχία δεν είναι μονοσήμαντα ορισμένη, αλλά θεωρήσουμε για παράδειγμα ότι υπάρχουν d είδη επιτυχιών, τότε είναι φυσιολογικό να αναζητήσουμε την κατανομή της δ.τ.μ. που εκφράζει την από κοινού καταγραφή του πλήθους των επιτυχιών i−είδους σε ένα δεδομένο αριθμό ανεξάρτητων δοκιμών, όπου σε κάθε δοκιμή έχουμε το πολύ μία επιτυχία με σταθερές πιθανότητες επιτυχίας για το κάθε είδος. Για παράδειγμα, σε 20 ρίψεις ενός ζαριού μπορεί κάποιος να ενδιαφέρεται για την απο κοινού κατανομή του πλήθους των ρίψεων που ήρθαν 1 μαζί με αυτές που ήρθαν 6. Εδώ, σε κάθε ρίψη μπορούμε να θεωρήσουμε την ένδειξη 1 ως επιτυχία πρώτου είδους και την ένδειξη 6 ως επιτυχία δεύτερου είδους. Αν το ζάρι είναι δίκαιο, τότε και τα δύο είδη επιτυχιών έχουν σταθερή πιθανότητα $1/6$ σε κάθε δοκιμή. Αν στο ίδιο τυχαίο πείραμα σε κάθε δοκιμή αντιστοιχίσουμε την αποτυχία, δηλαδή την εμφάνιση κάποιας από τις έδρες 2, 3, 4 ή 5, σε επιτυχία τρίτου είδους, και συμπεριλάβουμε την τ.μ. που καταγράφει το πλήθος των επιτυχιών τρίτου είδους στις δύο προηγούμενες, τότε είναι φανερό ότι δεν κερδίζουμε παραπάνω πληροφορία, αφού η τελευταία καθορίζεται πλήρως ως η διαφορά του πλήθους των συνολικών δοκιμών από το άθροισμα των επιτυχιών πρώτου και δεύτερου είδους. Παρ'όλα αυτά, συναντάμε στη βιβλιογραφία και τις δύο αυτές ισοδύναμες μορφές της πολυωνυμικής κατανομής. Πριν δώσουμε τον τυπικό ορισμό τους, ας δούμε μία χρήσιμη επέκταση του διωνυμικού συντελεστή.

Υπενθυμίζουμε ότι ο διωνυμικός συντελεστής $\binom{n}{k}$ φέρει το όνομα αυτό λόγω του ότι προκύπτει ως συντελεστής στο διωνυμικό ανάπτυγμα. Συγκεκριμένα,

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}, \quad \text{όπου } \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Εύκολα δείχνεται (με επαγωγή στο n) η ισχύς του πολυωνυμικού θεωρήματος, το οποίο μας επιτρέπει να γενικεύσουμε το διωνυμικό ανάπτυγμα από πολυώνυμα 2 μεταβλητών σε πολυώνυμα d μεταβλητών:

$$(x_1+x_2+\ldots+x_d)^n = \sum_{\substack{n_1,n_2,\ldots,n_d:\\ n_1+n_2+\cdots+n_d=n}} \frac{n!}{n_1!n_2!\ldots n_d!}\, x_1^{n_1}x_2^{n_2}\ldots x_d^{n_d}. \tag{2.7}$$

Αν κάνουμε μία αρίθμηση (διάκριση) των n παραγόντων που εμφανίζονται στο αριστερό μέλος:

$$\underbrace{(x_1+x_2+\ldots+x_d)}_{1} * \underbrace{(x_1+x_2+\ldots+x_d)}_{2} * \ldots * \underbrace{(x_1+x_2+\ldots+x_d)}_{n},$$

τότε είναι φανερό ότι το ανάπτυγμα συνίσταται σε ένα άθροισμα διακεκριμένων μονωνύμων βαθμού n (επιλέγοντας μία μεταβλητή από κάθε παράγοντα) της μορφής $x_1^{n_1}x_2^{n_2}\ldots x_d^{n_d}$, όπου κάθε εκθέτης αντιστοιχεί στο πλήθος των φορών που επιλέχθηκε η αντίστοιχη μεταβλητή, και σε κάθε μονώνυμο αντιστοιχεί ένας συντελεστής που εκφράζει το πλήθος των διαφορετικών τρόπων με τους οποίους μπορούμε να επιλέξουμε μέσα από το σύνολο των n διακεκριμένων παραγόντων (θέσεων) τις n_1 θέσεις για την επιλογή του x_1, τις n_2 θέσεις για την επιλογή του x_2 και τις υπολοιπόμενες n_d θέσεις για την επιλογή του x_d. Η σκέψη αυτή μας οδηγεί σε μία δεύτερη συνδυαστική απόδειξη (χωρίς επαγωγή) του πολυωνυμικού αναπτύγματος (2.7) κάνοντας χρήση της πολλαπλασιαστικής αρχής (ολοκληρώστε την απόδειξη).

Ορισμός 2.21: Έστω $n, n_i \in \mathbb{N}$, $i = 1, 2, \ldots, d$ με $n_1 + n_2 + \ldots n_d = n$. Λέμε *πολυωνυμικό συντελεστή* τον αριθμό

$$\binom{n}{n_1, n_2, \ldots, n_d} := \frac{n!}{n_1! n_2! \ldots n_d!}, \tag{2.8}$$

που εμφανίζεται ως συντελεστής στο πολυωνυμικό ανάπτυγμα (2.7).

Παρατήρηση 2.14.

1. Στην ειδική περίπτωση που $d = 2$, ο πολυωνυμικός συντελεστής αντιστοιχεί στο διωνυμικό συντελεστή, έτσι έχουμε $\binom{n}{k, n-k} = \binom{n}{k}$. Παραδοσιακά, προτιμάται ο τελευταίος συμβολισμός, που είναι και ο πιο σύντομος.

2. Κάποιες φορές, ο πολυωνυμικός συντελεστής (2.8) εμφανίζεται στη μορφή

$$\binom{n}{n_1, n_2, \ldots, n_{d-1}} := \frac{n!}{n_1! \ldots n_{d-1}!(n - n_1 - \ldots - n_{d-1})!}. \tag{2.9}$$

   Αυτό δεν θα πρέπει να προκαλεί σύγχυση, αφού οι δύο εκφράσεις διακρίνονται από το αν οι δείκτες που εκφράζουν το πλήθος των επιλογών αθροίζουν ή όχι στο συνολικό πλήθος των στοιχείων, n.

3. Είναι φανερό από τα σχόλια που ακολουθούν το πολυωνυμικό ανάπτυγμα ότι ο πολυωνυμικός συντελεστής εκφράζει το πλήθος των διαφορετικών τρόπων με τους οποίους μπορούμε να διαμερίσουμε ένα σύνολο n στοιχείων σε d (διακεκριμένα) υποσύνολα A_1, $A_2, \ldots, A_d$, με n_1, $n_2, \ldots, n_d$ στοιχεία αντίστοιχα.

Είμαστε τώρα σε θέση να ορίσουμε την πολυωνυμική κατανομή, η οποία αντιπροσωπεύει την κατανομή ενός τυχαίου διανύσματος που καταγράφει το πλήθος των επιτυχιών όλων των ειδών σε n ανεξάρτητες δοκιμές με σταθερές πιθανότητες επιτυχίας κάθε είδους σε κάθε δοκιμή, όταν σε κάθε μία από αυτές έχουμε επιτυχία μόνο ενός είδους ή το πολύ ενός. Είναι φανερό ότι αν ονομάσουμε την αποτυχία ως επιτυχία ενός επιπρόσθετου είδους με ταυτόχρονη καταγραφή του πλήθους τους, τότε έχουμε μία ισοδύναμη αναπαράσταση. Οι δύο αυτές ισοδύναμες μορφές θα διακρίνονται ανάλογα με το αν οι πιθανότητες επιτυχίας έχουν άθροισμα μικρότερο του 1 ή όχι. Στη στατιστική είναι επίσης βολικό να έχουμε έναν ορισμό που μας καλύπτει ακόμα και για ενδεχόμενα επιτυχίας μηδενικής πιθανότητας. Στην κατεύθυνση αυτή δίνουμε τον παρακάτω ορισμό.

Ορισμός 2.22: Έστω $X = (X_1, \ldots, X_d)$ ένα τ.δ. με $d \geq 1$, $n \geq 1$ και $p_1, \ldots, p_d \geq 0$ με $\sum_{i=1}^d p_i \leq 1$. Θέτουμε $p_{d+1} = 1 - \sum_{i=1}^d p_i$ και

$$\widetilde{S}_X = \left\{ x = (x_1, \ldots, x_d) \in \mathbb{N}^d : \sum_{i=1}^d x_i \leq n \right\}.$$

Θα λέμε ότι το X ακολουθεί την *πολυωνυμική κατανομή* $\mathcal{M}(n, p_1, \ldots, p_d)$ με παραμέτρους n και $p_1, \ldots, p_d$, αν για κάθε $x \in \widetilde{S}_X$

$$\mathbb{P}(X_1 = x_1, \ldots, X_d = x_d) = \binom{n}{x_1, \ldots, x_d} p_1^{x_1} \ldots p_d^{x_d} p_{d+1}^{x_{d+1}}, \tag{2.10}$$

όπου $x_{d+1} = n - \sum_{i=1}^d x_i$. Για $n = 1$, η πολυωνυμική κατανομή $\mathcal{M}(1, p_1, \ldots, p_d)$ αναφέρεται ως ***κατηγορική*** και θα λέμε ότι το τ.δ. $C \sim \boldsymbol{\mathcal{C}at}(p_1, \ldots, p_d)$.

Παρατήρηση 2.15. 1. Στην ειδική περίπτωση που $p_1, \ldots, p_d > 0$ και $\sum_{i=1}^d p_i = 1$ (άρα $p_{d+1} = 0$), αν συμβολίσουμε με

$$S_X = \left\{ x = (x_1, \ldots, x_d) \in \mathbb{N}^d : \sum_{i=1}^d x_i = n \right\},$$

τότε το S_X αντιστοιχεί στο στήριγμα της κατανομής του X και ισχύει πάλι η (2.10) με $x_{d+1} = 0$ και $p_{d+1}^{x_{d+1}} = 0^0 = 1$ (οπότε απλά παραλείπεται ο τελευταίος όρος). Είναι φανερό επίσης ότι αν $p_1, \ldots, p_d, p_{d+1} > 0$, τότε το στήριγμα της κατανομής του X είναι το $\widetilde{S}_X$.

2. Με τον παραπάνω ορισμό, η διωνυμική κατανομή $\boldsymbol{\mathcal{B}in}(n, p)$ ταυτίζεται με την $\mathcal{M}(n, p)$.

3. Η κατηγορική κατανομή γενικεύει την κατανομή *Bernoulli*, όπου αντί για τον καθορισμό μόνο μιας πιθανότητας επιτυχίας, προσδιορίζει τις πιθανότητες επιτυχίας όλων των ειδών. Αν $p_1, \ldots, p_d > 0$ και $p_{d+1} = 0$, τότε ένα τ.δ. $C \sim \boldsymbol{\mathcal{C}at}(p_1, \ldots, p_d)$ έχει στήριγμα το σύνολο $\{e_1, \ldots, e_d\}$ (δηλ. τα διανύσματα της ορθοκανονικής βάσης του $\mathbb{R}^d$). Αν όμως $p_{d+1} > 0$, τότε στα διανύσματα αυτά πρέπει να προστεθεί και το μηδενικό διάνυσμα 0_d.

4. Είναι φανερό από τον ορισμό και την ερμηνεία της πολυωνυμικής κατανομής ότι αν
$X \sim \mathcal{M}(n, p_1, \ldots, p_d)$, τότε μπορούμε να έχουμε την εξής χρήσιμη αναπαράσταση

$$X = \sum_{i=1}^n C_i, \tag{2.11}$$

όπου $C_i \sim \boldsymbol{\mathcal{C}at}(p_1, \ldots, p_d)$, $i = 1, 2, \ldots, n$, και είναι μεταξύ τους ανεξάρτητα τ.δ.. Η έκφραση αυτή του X ως άθροισμα ανεξάρτητων και ισόνομων κατηγορικών τ.δ. είναι η γενίκευση στα πολυδιάστατα της αντίστοιχης αναπαράστασης μιας τυχαίας μεταβλητής που ακολουθεί τη διωνυμική κατανομή ως άθροισμα ανεξάρτητων και ισόνομων τ.μ. που ακολουθούν κατανομή *Bernoulli*. Οι αναπαραστάσεις αυτές είναι χρήσιμες, όπως γνωρίζουμε και από τις στοιχειώδεις Πιθανότητες, για τον υπολογισμό μέσων τιμών και διασπορών.

Πρόταση 2.13: Αν ένα τ.δ. $X \sim \mathcal{M}(n, p)$ με $p = (p_1, \ldots, p_d)$, τότε η μέση τιμή

του είναι
$$\mathbb{E}(X) = np = n\,(p_1, \ldots, p_d)$$
και η διασπορά του (ο πίνακας διασποράς) είναι
$$\mathbb{V}(X) = n\,[\,\boldsymbol{dg}(p) - pp^\top] = n \begin{pmatrix} p_1(1-p_1) & -p_1p_2 & \ldots & -p_1p_d \\ -p_1p_2 & p_2(1-p_2) & \ldots & -p_2p_d \\ \vdots & \vdots & \ddots & \vdots \\ -p_1p_d & -p_2p_d & \ldots & p_d(1-p_d) \end{pmatrix}.$$

Απόδειξη: Χρησιμοποιώντας την αναπαράσταση (2.11), έχουμε
$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{i=1}^n C_i\right) = \sum_{i=1}^n \mathbb{E}(C_i) = n\,\mathbb{E}(C),$$
όπου $C \sim \boldsymbol{\mathcal{C}at}(p)$. Όμως
$$\mathbb{E}(C) = \big(\mathbb{E}(C_1), \ldots, \mathbb{E}(C_d)\big) = \big(\mathbb{P}(C_1 = 1), \ldots, \mathbb{P}(C_d = 1)\big) = (p_1, \ldots, p_d) = p,$$
καθώς $\{C_i = 1\} = \{C = e_i\}$ και $\mathbb{P}(C = e_i) = p_i$, $1 \leq i \leq d$. Επίσης, λόγω της ανεξαρτησίας, για τον υπολογισμό της διασποράς έχουμε
$$\mathbb{V}(X) = \mathbb{V}\left(\sum_{i=1}^n C_i\right) = \sum_{i=1}^n \mathbb{V}(C_i) = n\,\mathbb{V}(C).$$
Για τον τελευταίο πίνακα, έχουμε
$$\mathbb{V}(C) = \mathbb{E}(CC^\top) - \mathbb{E}(C)\,\mathbb{E}(C^\top) = \boldsymbol{dg}(p) - pp^\top,$$
χρησιμοποιώντας ότι $C_i \sim \boldsymbol{Be}(p_i)$, ότι $C_iC_j = 0$ για $i \neq j$, και άρα
$$\mathbb{E}(CC^\top) = \begin{pmatrix} \mathbb{E}(C_1^2) & \mathbb{E}(C_1C_2) & \ldots & \mathbb{E}(C_1C_d) \\ \mathbb{E}(C_2C_1) & \mathbb{E}(C_2^2) & \ldots & \mathbb{E}(C_2C_d) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}(C_dC_1) & \mathbb{E}(C_dC_2) & \ldots & \mathbb{E}(C_d^2) \end{pmatrix} = \begin{pmatrix} p_1 & 0 & \ldots & 0 \\ 0 & p_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & p_d \end{pmatrix}.$$
□

2.6 Πολυδιάστατη Κανονική Κατανομή

Η πιο σημαντική συνεχής πολυδιάστατη κατανομή είναι η κανονική κατανομή, για τους ίδιους λόγους που είναι και στη μονοδιάστη περίπτωση. Θα στηρίξουμε τον επόμενο ορισμό της κανονικής κατανομής στο Θεώρημα 2.1-(ii), χρησιμοποιώντας τις κατανομές όλων των γραμμικών συνδυασμών των συνιστωσών ενός τυχαίου διανύσματος.

Ορισμός 2.23: Έστω $X = (X_1, \ldots, X_m)$ μία δ.τ.μ. με $m \geq 1$, $\mu \in \mathbb{R}^m$ ένα διάνυσμα και Σ ένας συμμετρικός και θετικά ημιορισμένος πίνακας. Θα λέμε ότι η X ακολουθεί m-διάστατη *κανονική κατανομή* $\mathcal{N}_m(\mu, \Sigma)$ με παραμέτρους μ και Σ, αν για κάθε $u \in \mathbb{R}^m$ ισχύει

$$u^\top X \sim \mathcal{N}_1(u^\top \mu,\, u^\top \Sigma\, u). \tag{2.12}$$

Στην επόμενη πρόταση δείχνουμε ότι οι παράμετροι μ και Σ της κανονικής αντιστοιχούν σε μέση τιμή και διασπορά.

Πρόταση 2.14: Αν μία δ.τ.μ. $X \sim \mathcal{N}_m(\mu,\, \Sigma)$, τότε $\mathbb{E}(X) = \mu$ και $\mathbb{V}(X) = \Sigma$.

Απόδειξη: Από την (2.12) για $u = e_i$, $i = 1, \ldots, m$, έχουμε

$$X_i = e_i^\top X \sim \mathcal{N}_1\left(e_i^\top \mu,\, e_i^\top \Sigma\, e_i\right) = \mathcal{N}_1(\mu_i,\, \Sigma_{ii}),$$

από το οποίο συμπεραίνουμε ότι $\mathbb{E}(X_i) = \mu_i$ και $\mathbb{V}(X_i) = \Sigma_{ii}$. Από τις συνιστώσες μέσες τιμές συμπεραίνουμε ότι $\mathbb{E}(X) = \mu$ και από τις διασπορές ότι τα διαγώνια στοιχεία του Σ συμπίπτουν με αυτές. Θέτοντας τώρα $u = e_i + e_j$, $i, j = 1, \ldots, m$, $i \neq j$, έχουμε

$$\mathbb{V}(X_i + X_j) = \mathbb{V}[(e_i + e_j)^\top X] = (e_i^\top + e_j^\top)\,\Sigma\,(e_i + e_j) = \Sigma_{ii} + \Sigma_{jj} + 2\Sigma_{ij}.$$

Από τη σχέση

$$\mathbb{V}(X_i + X_j) = \mathbb{V}(X_i) + \mathbb{V}(X_j) + 2\,\mathbb{C}(X_i, X_j)$$

και τα παραπάνω, καταλήγουμε στο ότι $\mathbb{C}(X_i, X_j) = \Sigma_{ij}$ και έτσι φτάνουμε στο ζητούμενο και για τον πίνακα διασποράς.

$\square$

Παρατήρηση 2.16. Γίνεται σαφές από την παραπάνω απόδειξη ότι αν για μία οποιαδήποτε δ.τ.μ. X ισχύει ότι για κάθε $u \in \mathbb{R}^m$, η μέση τιμή $\mathbb{E}(u^\top X) = u^\top \mu$ και η διασπορά $\mathbb{V}(u^\top X) = u^\top \Sigma\, u$, τότε $\mathbb{E}(X) = \mu$ και $\mathbb{V}(X) = \Sigma$. Σε αλγεβρική γλώσσα, θα λέγαμε ότι η μέση τιμή μ καθορίζει μία γραμμική μορφή $f : \mathbb{R}^m \to \mathbb{R}$, όπου $f(u) = u^\top \mu$. Αντίστοιχα, ο πίνακας διασποράς Σ καθορίζει μία θετικά ημιορισμένη τετραγωνική μορφή $Q : \mathbb{R}^m \to \mathbb{R}$, όπου $Q(u) = u^\top \Sigma\, u$.

Η πιο απλή περίπτωση πολυδιάστατης κανονικής είναι βέβαια μία δ.τ.μ. Z που αποτελείται από ανεξάρτητες τ.μ. που ακολουθούν την τυπική κανονική.

Λήμμα 2.2: Αν $Z \sim \mathcal{N}_m(0_m, I_m)$, όπου 0_m είναι το μηδενικό διάνυσμα και I_m είναι ο ταυτοτικός πίνακας m-τάξης, τότε οι συνιστώσες της είναι ανεξάρτητες και ισόνομες τ.μ. που ακολουθούν την τυπική κανονική. Λέμε ότι η Z ακολουθεί την *τυπική κανονική διάστασης* m.

Απόδειξη: Ας υποθέσουμε ότι $Z \sim \mathcal{N}_m(0_m, I_m)$. Τότε, από τον ορισμό της πολυδιάστατης κανονικής, έχουμε $u^\top Z \sim \mathcal{N}_1(0, u^\top u)$. Ας υποθέσουμε τώρα ότι παίρνουμε μία δ.τ.μ. $W = (W_1, \ldots, W_m)^\top$ που αποτελείται από τ.μ. $\{W_i\}_{i=1}^m$ που είναι ανεξάρτητες τυπικές κανονικές. Τότε $u^\top W \sim \mathcal{N}_1(0, u^\top u)$, από γνωστό αποτέλεσμα για γραμμικούς συνδιασμούς ανεξάρτητων τ.μ. που ακολουθούν κανονική κατανομή. Καταλήγουμε στο συμπέρασμα ότι $u^\top Z \stackrel{d}{=} u^\top W$ για κάθε $u \in \mathbb{R}^m$ και, άρα, $Z \stackrel{d}{=} W$. □

Η έννοια της συνάρτησης πυκνότητας πιθανότητας (σ.π.π.) για απόλυτα συνεχείς τ.μ. είναι ήδη γνωστή από τις στοιχειώδεις Πιθανότητες, όπως επίσης και η από κοινού σ.π.π. για δ.τ.μ.. Αν γνωρίζουμε τη σ.π.π. για μία απόλυτα συνεχή δ.τ.μ. X, τότε μπορούμε να χαρακτηρίσουμε την κατανομή της X. Η εύρεση της σ.π.π. για μετασχηματισμούς της μορφής $AX + b$ (αφφινικούς μετασχηματισμούς) είναι εντελώς ανάλογη της περίπτωσης για πραγματικές τ.μ. και το μόνο που αλλάζει είναι η υπολογιστική πολυπλοκότητα. Υπενθυμίζουμε ότι στη μονοδιάστατη περίπτωση, αν $Y = aX + b$ και $f_X(x)$ είναι η σ.π.π. της X, τότε για $a \neq 0$,

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right).$$

Δίνουμε παρακάτω το πολυδιάστατο ανάλογο της παραπάνω σχέσης.

Πρόταση 2.15: Αν $X \in \mathbb{R}^m$ είναι μία δ.τ.μ. που έχει σ.π.π. $f_X(x)$, A είναι ένας $m \times m$ αντιστρέψιμος πίνακας και $b \in \mathbb{R}^m$ ένα σταθερό διάνυσμα, τότε η $Y = AX + b$ έχει σ.π.π. στον $\mathbb{R}^m$ που δίνεται από τη σχέση

$$f_Y(y) = \frac{1}{|\det(A)|} f_X\left(A^{-1}(y-b)\right).$$

Στην επόμενη πρόταση δίνονται κάποιες σημαντικές ιδιότητες της πολυδιάστατης κανονικής κατανομής.

Πρόταση 2.16: Έστω $X \sim \mathcal{N}_m(\mu, \Sigma)$, $\mu \in \mathbb{R}^m$ και Σ πίνακας συνδιακύμανσης.

(i) Αν $A \in \mathbb{R}^{n \times m}$ και $b \in \mathbb{R}^n$, τότε

$$AX + b \sim \mathcal{N}_n\left(A\mu + b,\ A\,\Sigma\,A^\top\right), \tag{2.13}$$

(ii) αν ο Σ είναι θετικά ορισμένος (αντιστρέψιμος), τότε η X έχει συνάρτηση πυκνότητας πιθανότητας στον $\mathbb{R}^m$ της μορφής

$$f_X(x) = (2\pi)^{-m/2} \left(\det(\Sigma)\right)^{-1/2} \exp\left\{-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)\right\}, \quad \forall x \in \mathbb{R}^m. \tag{2.14}$$

Απόδειξη:

(i) Έστω $u \in \mathbb{R}^n$. Τότε,

$$u^\top(AX+b) = (u^\top A)X+u^\top b = (A^\top u)^\top X+u^\top b \sim \mathcal{N}_1\Big((A^\top u)^\top \mu,\, (A^\top u)^\top \Sigma\, A^\top u\Big)+u^\top b,$$

όπου η τελευταία σχέση έπεται από την υπόθεση ότι $X \sim \mathcal{N}_m(\mu, \Sigma)$ και άρα $v^\top X \sim \mathcal{N}_1\left(v^\top \mu,\, v^\top \Sigma\, v\right)$ για $v = A^\top u \in \mathbb{R}^m$. Τελικά, η παραπάνω σχέση ξαναγράφεται στη μορφή

$$u^\top(AX+b) \sim \mathcal{N}_1\left(u^\top(A\mu+b),\, u^\top A\Sigma\, A^\top u\right),$$

από την οποία συμπεραίνουμε την (2.13).

(ii) Αν ο Σ είναι θετικά ορισμένος, τότε γράφεται στη μορφή $\Sigma = AA^\top$, με A αντιστρέψιμο. Ο πίνακας A παίζει τον ρόλο μιας τετραγωνικής ρίζας του Σ. Συμπεραίνουμε λοιπόν από το ερώτημα (i) ότι η δ.τ.μ.

$$Z = A^{-1}(X-\mu) \sim \mathcal{N}_m\left(0_m,\, A^{-1}\,\Sigma\,(A^{-1})^\top\right)$$

και αφού $\Sigma = AA^\top$ συμπεραίνουμε ότι $Z \sim \mathcal{N}_m(0_m, I_m)$. Η σ.π.π. του Z είναι

$$f_Z(z) = \prod_{i=1}^m f_{Z_i}(z_i) = \prod_{i=1}^m (2\pi)^{-1/2} e^{-z_i^2/2} = (2\pi)^{-m/2} \exp\left\{-\frac{1}{2} z^\top z\right\}, \quad \forall z \in \mathbb{R}^m.$$

Από την Πρόταση 2.15, συμπεραίνουμε ότι η δ.τ.μ. $X = AZ+\mu$ έχει σ.π.π. στον $\mathbb{R}^m$ της μορφής

$$f_X(x) = (2\pi)^{-m/2} |\det(A)|^{-1} \exp\left\{-\frac{1}{2}\left(A^{-1}(x-\mu)\right)^\top \left(A^{-1}(x-\mu)\right)\right\}.$$

Η παραπάνω έκφραση συμπίπτει με την (2.14). □

Ένα πολύ χρήσιμο θεώρημα που συνδέεται με την πολυδιάστατη κανονική κατανομή και έχει αρκετές εφαρμογές στη Στατιστική είναι το θεώρημα του Cochran.

Θεώρημα 2.2 (Cochran): Έστω X δ.τ.μ. με $X \sim \mathcal{N}_n(\mu, \sigma^2 I_n)$, όπου $\mu \in \mathbb{R}^n$ και $\sigma^2 > 0$. Έστω επίσης $\mathbb{R}^n = V_1 \oplus \cdots \oplus V_p$ μία ανάλυση του $\mathbb{R}^n$ σε ένα ευθύ άθροισμα $p \geq 2$ διανυσματικών υπόχωρων (κάθετων ανά δύο) με $\dim V_i = d_i$. Αν Π_i είναι ο πίνακας ορθογώνιας προβολής στον V_i και $Y_i = \Pi_i X$, $1 \leq i \leq p$, δηλαδή Y_i είναι η ορθογώνια προβολή του X στον V_i, τότε

(i) $\{Y_1, \ldots, Y_p\}$ είναι ανεξάρτητες δ.τ.μ. του $\mathbb{R}^n$,

(ii) $Y_i \sim \mathcal{N}_n\left(\Pi_i\, \mu,\, \sigma^2\, \Pi_i\right)$, $1 \leq i \leq p$,

(iii) $\left\{\|Y_1 - \Pi_1\mu\|^2, \ldots, \|Y_p - \Pi_p\mu\|^2\right\}$ είναι ανεξάρτητες τυχαίες μεταβλητές,

(iv) $\frac{1}{\sigma^2}\left\|Y_i - \Pi_i\mu\right\|^2 \sim \chi^2_{d_i}$, $1 \le i \le p$.

Απόδειξη: Κατ'αρχάς από υπόθεση έχουμε $X \sim \mathcal{N}_n\left(\mu,\, \sigma^2 I_n\right)$ και άρα

$$Y := \begin{pmatrix} Y_1 \\ \vdots \\ Y_p \end{pmatrix} = \begin{pmatrix} \Pi_1 X \\ \vdots \\ \Pi_p X \end{pmatrix} = \underbrace{\begin{pmatrix} \Pi_1 \\ \vdots \\ \Pi_p \end{pmatrix}}_{=:\Pi} X = \Pi X \sim \mathcal{N}_{np}\big(\Pi\mu,\, \sigma^2 \Pi\,\Pi^\top\big) = \mathcal{N}_{np}\big(\Pi\mu,\, \sigma^2\, \mathrm{dg}(\Pi)\big), \tag{2.15}$$

ως γραμμικός μετασχηματισμός της X και το γεγονός ότι

$$\Pi\Pi^\top = \big(\Pi_i\Pi_j\big)_{i,j} = \mathrm{dg}(\Pi),$$

όπου η τελευταία ισότητα προκύπτει από την Πρόταση 2.8, καθώς $\Pi_i^2 = \Pi_i$ και $\Pi_i\Pi_j = 0_{n\times n}$ για $i \ne j$. Από την (2.15) έπονται άμεσα τα (i) και (ii). Πράγματι, καθώς τα $\{Y_1, \ldots, Y_p\}$ είναι από κοινού κανονικά κατανεμημένα, θα είναι ανεξάρτητα αν, και μόνο αν είναι ασυσχέτιστα, ή ισοδύναμα $\mathbb{C}(Y_i, Y_j) = 0_{n\times n}$, για $i \ne j$. Η ανεξαρτησία λοιπόν έπεται από το γεγονός ότι $\mathbb{V}(Y) = \sigma^2\, \mathrm{dg}(\Pi)$ και άρα αντιστοιχεί σε μπλοκ διαγώνιο πίνακα. Το (ii) προκύπτει παίρνοντας τις περιθώριες κατανομές της Y από την (2.15). Για το (iii), παρατηρούμε ότι

$$\left\{\left\|Y_1 - \Pi_1\mu\right\|^2, \ldots, \left\|Y_p - \Pi_p\mu\right\|^2\right\} = \left\{g_1(Y_1), \ldots, g_p(Y_p)\right\},$$

όπου $g_i(y_i) = \left\|y_i - \Pi_i\mu\right\|^2$, $1 \le i \le p$. Καθώς οι δ.τ.μ. $\{Y_i\}_{i=1}^p$ είναι ανεξάρτητες από το (i), το ίδιο θα ισχύει και για τις $\{g_i(Y_i)\}_{i=1}^p$. Τέλος, θα δείξουμε τώρα και την (iv). Από τη (ii) και την Πρόταση 2.8–(v), όπου ο Π_i διαγωνοποιείται στη μορφή $\Lambda_i D_i \Lambda_i^\top$, με $\Lambda_i^\top = \Lambda_i^{-1}$, έχουμε

$$Y_i - \Pi_i\mu \sim \mathcal{N}_n(0_n,\, \sigma^2\Pi_i) \quad\Longrightarrow\quad \Lambda_i^\top\left(Y_i - \Pi_i\mu\right) \sim \mathcal{N}_n\left(0,\, \sigma^2 D_i\right), \quad 1 \le i \le p.$$

Χωρίς βλάβη της γενικότητας υποθέτουμε ότι $D_i = \left(\begin{array}{c|c} I_{d_i} & O \\ \hline O & O \end{array}\right)$, και άρα

$$\frac{1}{\sigma}\Lambda_i^\top\left(Y_i - \Pi_i\mu\right) \overset{d}{=} \big(Z_1, \ldots, Z_{d_i}, 0, \ldots, 0\big) \sim \mathcal{N}_n(0_n,\, D_i). \tag{2.16}$$

Όμως

$$\begin{aligned}
\frac{1}{\sigma^2}\left\|\Lambda_i^\top(Y_i - \Pi_i\mu)\right\|^2 &= \frac{1}{\sigma^2}\Big(\Lambda_i^\top(Y_i - \Pi_i\mu)\Big)^\top\Big(\Lambda_i^\top(Y_i - \Pi_i\mu)\Big) \\
&= \frac{1}{\sigma^2}(Y_i - \Pi_i\mu)^\top \underbrace{\Lambda_i\Lambda_i^\top}_{I_n}(Y_i - \Pi_i\mu) = \frac{1}{\sigma^2}\left\|Y_i - \Pi_i\mu\right\|^2,
\end{aligned}$$

και άρα από την (2.16) και την παραπάνω ισότητα προκύπτει ότι

$$\frac{1}{\sigma^2}\|Y_i - \Pi_i\mu\|^2 \overset{d}{=} \sum_{k=1}^{d_i} Z_k^2 \sim \chi^2_{d_i}, \quad 1 \leq i \leq p.$$

□

Μία ενδιαφέρουσα εφαρμογή του Θεωρήματος Cochran δίνεται στην παρακάτω πρόταση, από την οποία προκύπτει άμεσα ότι σε ένα τυχαίο δείγμα από κανονική κατανομή με άγνωστη μέση τιμή και διασπορά, ο δειγματικός μέσος και η αμερόληπτη δειγματική διασπορά είναι ανεξάρτητες τ.μ., καθώς επίσης προσδιορίζονται και οι ακριβείς κατανομές τους.

Πρόταση 2.17: Έστω $\{X_i\}_{i=1}^n$ α.ι.τ.μ. από την κανονική $\mathcal{N}(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$. Θέτουμε $\overline{X}_n$ το δειγματικό μέσο και $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n \left(X_i - \overline{X}_n\right)^2$ την αμερόληπτη δειγματική διασπορά. Τότε,

1. οι $\overline{X}_n$ και S_n^2 είναι ανεξάρτητες τυχαίες μεταβλητές,
2. $\overline{X}_n \sim \mathcal{N}\left(\mu, \dfrac{\sigma^2}{n}\right)$,
3. $\dfrac{(n-1)S_n^2}{\sigma^2} \sim \chi^2_{n-1}$.

Απόδειξη: Έχουμε ότι $\mathbb{R}^n = V_1 \oplus V_2$, όπου $V_1 = \langle \underbrace{(1, 1, \ldots, 1)^\top}_{1_n} \rangle = \langle \underbrace{\frac{1}{\sqrt{n}} 1_n}_{\text{νόρμα } 1} \rangle$ και $V_2 = V_1^\perp$. Από την Πρόταση 2.8 προκύπτει ότι $\Pi_2 = I_n - \Pi_1$, όπου

$$\Pi_1 = \left(\frac{1}{\sqrt{n}} 1_n\right)\left(\frac{1}{\sqrt{n}} 1_n^\top\right) = \frac{1}{n} 1_n 1_n^\top = \frac{1}{n}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}.$$

Αν $X = (X_1, \ldots, X_n)^\top$, τότε από υπόθεση $X \sim \mathcal{N}_n(\mu\, 1_n, \sigma^2 I_n)$ και από το (i) του θεωρήματος Cochran, οι $Y_1 = \Pi_1 X$ και $Y_2 = \Pi_2 X$ είναι ανεξάρτητες δ.τ.μ.. Όμως

$$Y_1 = \frac{1}{n}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} \overline{X}_n \\ \vdots \\ \overline{X}_n \end{pmatrix},$$

$$Y_2 = \Pi_2 X = (I_n - \Pi_1)X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} - \begin{pmatrix} \overline{X}_n \\ \vdots \\ \overline{X}_n \end{pmatrix} = \begin{pmatrix} X_1 - \overline{X}_n \\ \vdots \\ X_n - \overline{X}_n \end{pmatrix},$$

άρα $\overline{X}_n = g_1(Y_1)$ και $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2 = g_2(Y_2)$. Καθώς τα Y_1 και Y_2 είναι ανεξάρτητες δ.τ.μ., συμπεραίνουμε ότι οι $\overline{X}_n$ και S_n^2 είναι ανεξάρτητες τ.μ. και έτσι ισχύει το (i). Το (ii) δείχνεται άμεσα με εφαρμογή του (ii) του θεωρήματος Cochran. Συγκεκριμένα, $Y_1 = (\overline{X}_n, \ldots, \overline{X}_n)^\top \sim \mathcal{N}_n\left(\mu\,\Pi_1 1_n,\, \sigma^2\,\Pi_1\right) = \mathcal{N}_n\left(\mu\, 1_n,\, \sigma^2\,\Pi_1\right)$ και άρα

$$\overline{X}_n = e_1^\top Y_1 \sim \mathcal{N}_1\left(\mu\,, \frac{\sigma^2}{n}\right).$$

Για την απόδειξη του (iv) παρατηρούμε ότι $\dim V_2 = n - \dim V_1 = n-1$ και ότι $\mathbb{E}(Y_2) = 0_n$. Από από το Cochran – (iv) συμπεραίνουμε λοιπόν ότι

$$\frac{n-1}{\sigma^2}S_n^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}Y_{2,i}^2 = \frac{1}{\sigma^2}\left\|Y_2 - \mathbb{E}(Y_2)\right\|^2 \sim \chi^2_{n-1}.$$

□

2.7 Κατανομή Dirichlet

Μία αρκετά ενδιαφέρουσα πολυδιάστατη κατανομή με πολλές εφαρμογές στη Μπεϋζιανή Στατιστική είναι η κατανομή Dirichlet που αποτελεί την πολυδιάστατη γενίκευση της κατανομής Βήτα. Όπως στη σ.π.π. της Βήτα κατανομής εμπλέκεται η συνάρτηση Βήτα ως σταθερά κανονικοποίησης, έτσι και για την κατανομή Dirichlet χρειαζόμαστε την πολυμεταβλητή συνάρτηση Βήτα.

Ορισμός 2.24: Έστω $\alpha = (\alpha_1, \ldots, \alpha_m)$, $m \geq 2$, ένα διάνυσμα θετικών πραγματικών αριθμών και α_0 το άθροισμά τους. Η συνάρτηση

$$B(\alpha) = \int_{x \in \Delta^{m-1}} \prod_{i=1}^{m} x_i^{\alpha_i - 1}\, dx \tag{2.17}$$

όπου Δ^{m-1} είναι το $(m-1)$-simplex πιθανοτήτων του $\mathbb{R}^m$ καλείται *πολυμεταβλητή συνάρτηση Βήτα*. Πολλές φορές ορίζεται απ'ευθείας από τη σχέση που τη συνδέει με τη συνάρτηση Γάμμα:

$$B(\alpha) = \frac{\prod_{i=1}^{s}\Gamma(\alpha_i)}{\Gamma(\alpha_0)}.$$

Παρατήρηση 2.17. Το ολοκλήρωμα (2.17) μπορεί να δειχθεί εύκολα ότι δίνεται από την (2.7) με επαγωγή. Για $m = 2$ είναι η γνωστή σχέση που ικανοποιεί η συνάρτηση Βήτα. Ας υποθέσουμε τώρα ότι ισχύει για $m-1$ και θα το δείξουμε για m.

Πράγματι, παρατηρούμε ότι

$$B(\alpha) = \int_0^1 \int_0^{1-x_1} \cdots \int_0^{1-\sum_{i=1}^{m-2} x_i} \left(\prod_{i=1}^{m-2} x_i^{\alpha_i - 1}\right) x_{m-1}^{\alpha_{m-1}-1}(1-\sum_{i=1}^{m-2} x_i - x_{m-1})^{\alpha_m - 1} dx_{m-1} \ldots dx_2 dx_1. \tag{2.18}$$

Κάνοντας την αλλαγή μεταβλητής $y = x_{m-1}/(1-\sum_{i=1}^{m-2} x_i)$ στο τελευταίο ολοκλήρωμα, καταλήγουμε στη σχέση

$$B(\alpha_1, \ldots, \alpha_{m-2}, \alpha_{m-1}, \alpha_m) = B(\alpha_1, \ldots, \alpha_{m-2}, \alpha_{m-1} + \alpha_m) B(\alpha_{m-1}, \alpha_m). \tag{2.19}$$

Από την επαγωγική υπόθεση και την ισχύ για $m = 2$ καταλήγουμε άμεσα στο ζητούμενο.

Ορισμός 2.25: Έστω $X = (X_1, X_2, \ldots, X_m)$ ένα τ.δ. με $m \geq 2$, $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m)$ με $\alpha_i > 0$, $1 \leq i \leq m$. Θα λέμε ότι το X ακολουθεί την m-διάστατη *κατανομή Dirichlet* $\mathcal{D}_m(\alpha)$ με παράμετρο το διάνυσμα α, αν έχει σ.π.π. στο $(m-1)$-simplex πιθανοτήτων του $\mathbb{R}^m$ της μορφής

$$f_X(x) = \frac{1}{B(\alpha)} \prod_{i=1}^{m} x_i^{\alpha_i - 1},$$

όπου $B(\alpha)$ είναι η πολυμεταβλητή συνάρτηση Βήτα που δίνεται από τη σχέση (2.7).

Είναι φανερό ότι αν $X \sim \mathcal{D}_m(\alpha)$, τότε οι συνιστώσεις $\{X_i\}_{i=1}^m$ είναι εξαρτημένες και ικανοποιούν $\sum_{i=1}^m X_i = 1$. Θα ήταν βολικό να ορίζαμε και μία κατανομή που είναι της μορφής Dirichlet, αλλά έχει πυκνότητα απ'ευθείας στον $\mathbb{R}^m$. Η κατανομή αυτή θα αντιστοιχεί στην περιθώρια κατανομή οποιουδήποτε τυχαίου υποδιανύσματος διάστασης m φτιάξουμε από ένα τ.δ. $X \sim \mathcal{D}_{m+1}(\alpha)$. Ο παρακάτω ορισμός εξυπηρετεί αυτό το σκοπό.

Ορισμός 2.26: Έστω $X = (X_1, \ldots, X_m)$ ένα τ.δ. με $m \geq 1$, $\alpha = (\alpha_1, \ldots, \alpha_m, \alpha_{m+1})$ με $\alpha_i > 0$, $1 \leq i \leq m+1$. Θα λέμε ότι το X ακολουθεί την m-διάστατη *κατανομή Dirichlet* $\mathcal{D}_m^*(\alpha)$ με παράμετρο το διάνυσμα α, αν έχει σ.π.π. $f_X(x)$ της μορφής (2.7), αλλά $x \in (0,1)^m$ με $\sum_{i=1}^m x_i < 1$.

Παρατήρηση 2.18. Είναι φανερό ότι $\mathcal{B}(\alpha, \beta) \equiv \mathcal{D}^*(\alpha, \beta)$.

Η κατανομή Dirichlet έχει την ιδιότητα ότι οι περιθώριες κατανομές παραμένουν Dirichlet.

Πρόταση 2.18: Αν $X \sim \mathcal{D}_m(\alpha)$ και Y υποδιάνυσμα του X διάστασης l, με $1 \leq l \leq m-1$, τότε $Y \sim \mathcal{D}_l^*(\beta, \gamma_0)$, όπου β το υποδιάνυσμα του α που αντιστοιχεί στις συνιστώσεις του Y και γ αυτό που απομένει μετά την αφαίρεση των συνιστωσών.

Απόδειξη: Έστω $l \in \{1, 2, \ldots, m-1\}$. Χωρίς βλάβη της γενικότητας θα το δείξουμε για τις l πρώτες συνιστώσες. Θέτουμε $X = (Y, Z)$ και $\alpha = (\beta, \gamma)$. Αν $l = m-1$, τότε το αποτέλεσμα ισχύει αφού τότε $Z = X_m = 1 - \sum_{i=1}^{m-1} X_i$ και η $Y \sim \mathcal{D}^*_{m-1}(\beta)$. Αν $l < m-1$, $x = (y, z)$ και $y_0 = \sum_{i=1}^{l} y_i$ τότε

$$f_Y(y) = \frac{1}{B(\alpha)} \left(\prod_{i=1}^{l} y_i^{\beta_i - 1} \right) \int_0^{1-y_0} \cdots \int_0^{1-y_0-\sum_{i=1}^{m-l-2} z_i} \left(\prod_{i=1}^{m-l-1} z_i^{\gamma_i - 1} \right) (1-y_0-\sum_{i=1}^{m-l-2} z_i)^{\gamma_{m-l}-1} dz^*,$$

όπου $z^* = (z_1, \ldots, z_{m-l})$. Με την αλλαγή μεταβλητών $w_i = z_i/(1-y_0)$ προκύπτει ότι

$$f_Y(y) = \frac{B(\gamma)}{B(\alpha)} \left(\prod_{i=1}^{l} y_i^{\beta_i - 1} \right) (1-y_0)^{\gamma_0 - 1}. \tag{2.20}$$

Όμως $B(\alpha) = B(\beta, \gamma)$ και από τη σχέση (2.19) με μία απλή γενίκευση έχουμε $B(\beta, \gamma)/B(\gamma) = B(\beta, \gamma_0)$. Αντικαθιστώντας το παραπάνω στην (2.20) παίρνουμε το ζητούμενο.

□

Μία χρήσιμη ιδιότητα της κατανομής Dirichlet που προκύπτει άμεσα από τη μορφή της σ.π.π. δίνεται στο παρακάτω λήμμα.

Λήμμα 2.3: Αν $X \sim \mathcal{D}_m(\alpha)$ με $\alpha = (\alpha_1, \ldots, \alpha_m)$ και $\beta = (\beta_1, \ldots, \beta_m)$ ένα ακόμα διάνυσμα θετικών πραγματικών αριθμών, τότε

$$\mathbb{E}\left(\prod_{i=1}^{m} X_i^{\beta_i} \right) = \frac{B(\alpha+\beta)}{B(\alpha)}.$$

Στην παρακάτω πρόταση δίνεται μέση τιμή και διασπορά.

Πρόταση 2.19: Αν $X \sim \mathcal{D}_m(\alpha)$ με $\alpha = (\alpha_1, \ldots, \alpha_m)$, τότε η μέση τιμή του είναι

$$\mathbb{E}(X) = \frac{1}{\alpha_0}\alpha = \frac{1}{\alpha_0}(\alpha_1, \ldots, \alpha_m)$$

και η διασπορά του (ο πίνακας διασποράς) είναι

$$\mathbb{V}(X) = \frac{1}{\alpha_0^2(\alpha_0+1)} \left[\alpha_0 \, \mathrm{dg}(\alpha) - \alpha\, \alpha^\top \right],$$

που αντιστοιχεί στον πίνακα

$$\frac{1}{\alpha_0^2(\alpha_0+1)}\begin{pmatrix} \alpha_1(\alpha_0-\alpha_1) & -\alpha_1\alpha_2 & \dots & -\alpha_1\alpha_s \\ -\alpha_1\alpha_2 & \alpha_2(\alpha_0-\alpha_2) & \dots & -\alpha_2\alpha_s \\ \vdots & \vdots & \ddots & \vdots \\ -\alpha_1\alpha_s & -\alpha_2\alpha_s & \dots & \alpha_s(\alpha_0-\alpha_s) \end{pmatrix}.$$

Απόδειξη: Από το Λήμμα 2.3 για $\beta = e_i$ και τη σχέση $\Gamma(k+1) = k\Gamma(k)$ η έκφραση για τις μέσες τιμές προκύπτει ως εξής:

$$\mathbb{E}(X_i) = \frac{B(\alpha+e_i)}{B(\alpha)} = \frac{\Gamma(\alpha_i+1)}{\Gamma(\alpha_i)}\frac{\Gamma(\alpha_0)}{\Gamma(\alpha_0+1)} = \frac{\alpha_i}{\alpha_0}.$$

Επιπλέον, για $i \neq j$ και $\beta = e_i + e_j$ έχουμε

$$\mathbb{E}(X_iX_j) = \frac{\Gamma(\alpha_i+1)\Gamma(\alpha_j+1)}{\Gamma(\alpha_i)\Gamma(\alpha_j)}\frac{\Gamma(\alpha_0)}{\Gamma(\alpha_0+2)} = \frac{\alpha_i\alpha_j}{\alpha_0(\alpha_0+1)},$$

ενώ για $i = j$ και $\beta = 2e_i$ έχουμε

$$\mathbb{E}(X_i^2) = \frac{\Gamma(\alpha_i+2)}{\Gamma(\alpha_i)}\frac{\Gamma(\alpha_0)}{\Gamma(\alpha_0+2)} = \frac{\alpha_i(\alpha_i+1)}{\alpha_0(\alpha_0+1)}.$$

Από τις παραπάνω σχέσεις έχουμε για $i \neq j$

$$\mathbb{C}(X_i, X_j) = -\frac{\alpha_i\alpha_j}{\alpha_0^2(\alpha_0+1)},$$

και

$$\mathbb{V}(X_i) = \frac{\alpha_i(\alpha_0-\alpha_i)}{\alpha_0^2(\alpha_0+1)}.$$

Έτσι ολοκληρώνεται η απόδειξη. □

2.8 Σύγκλιση ακολουθίας τυχαίων διανυσμάτων και οριακά θεωρήματα

Σε αυτήν την ενότητα υπενθυμίζουμε τους σημαντικότερους τρόπους σύγκλισης ακολουθίας τ.μ., δίνουμε τις πολυδιάστατες επεκτάσεις τους και καταλήγουμε σε νόμους μεγάλων αριθμών και ένα πολυδιάστατο κεντρικό οριακό θεώρημα.

Ορισμός 2.27: Μία ακολουθία τ.μ. (X_n) λέμε ότι συγκλίνει *με πιθανότητα 1* ή *σχεδόν βεβαίως*, ή *ισχυρά*, σε μία τ.μ. X, και γράφουμε $X_n \xrightarrow{a.s.} X$, αν

$$\mathbb{P}\left(\left\{\omega \in \Omega : \exists \lim X_n(\omega) \text{ και } \lim X_n(\omega) = X(\omega)\right\}\right) = 1.$$

Το παραπάνω ενδεχόμενο γράφεται και συνοπτικά ως $\{\lim X_n = X\}$ και έτσι σύντομα γράφουμε $\mathbb{P}(\lim X_n = X) = 1$ για να δηλώσουμε την παραπάνω σύγκλιση. Από τον ορισμό αυτόν, είναι φανερό ότι το μόνο που χρειάζεται να ορίσει κανείς είναι η σύγκλιση στο 0, αφού

$$X_n \xrightarrow{a.s.} X \iff |X_n - X| \xrightarrow{a.s.} 0.$$

Ο τρόπος αυτός σύγκλισης επεκτείνεται εντελώς φυσιολογικά για ακολουθίες τ.δ. στον $\mathbb{R}^d$. Ο τυπικός ορισμός είναι όπως και στη μονοδιάστατη περίπτωση, απλώς τώρα η έννοια του ορίου νοείται ως προς την ευκλείδεια νόρμα.

Ορισμός 2.28: Μία ακολουθία τ.δ. (X_n) λέμε ότι συγκλίνει *με πιθανότητα 1* ή *σχεδόν βεβαίως*, ή *ισχυρά*, σε ένα τ.δ. X, και γράφουμε $X_n \xrightarrow{a.s.} X$, αν

$$\|X_n - X\| \xrightarrow{a.s.} 0 \quad \text{ή ισοδύναμα} \quad \mathbb{P}\left(\lim X_n = X\right) = 1.$$

Ασθενέστερη της παραπάνω σύγκλισης είναι η στοχαστική σύγκλιση.

Ορισμός 2.29: Μία ακολουθία τ.μ. (X_n) λέμε ότι συγκλίνει *κατά πιθανότητα* ή *στοχαστικά* σε μία τ.μ. X, και γράφουμε $X_n \xrightarrow{p} X$, αν για κάθε $\varepsilon > 0$

$$\mathbb{P}\left(|X_n - X| > \varepsilon\right) \xrightarrow[n\to\infty]{} 0 \quad \text{ή ισοδύναμα} \quad \mathbb{P}\left(|X_n - X| \le \varepsilon\right) \xrightarrow[n\to\infty]{} 1.$$

Όπως για την ισχυρή σύγκλιση, έτσι και για τη στοχαστική, έχουμε

$$X_n \xrightarrow{p} X \iff |X_n - X| \xrightarrow{p} 0.$$

Έτσι, οδηγούμαστε στην ακόλουθη επέκταση.

Ορισμός 2.30: Μία ακολουθία τ.δ. (X_n) λέμε ότι συγκλίνει *κατά πιθανότητα* ή *στοχαστικά* σε ένα τ.δ. X, και γράφουμε $X_n \xrightarrow{p} X$, αν

$$\|X_n - X\| \xrightarrow{p} 0.$$

Παρόλο που η ισχυρή και η στοχαστική σύγκλιση εμπλέκουν τη συμπεριφορά της νόρμας στον $\mathbb{R}^d$, η μελέτη τους απλοποιείται και είναι ισοδύναμη με την εξέταση των αντίστοιχων ορίων κατά συνιστώσα. Συγκεκριμένα, είναι εύκολο να δειχθεί το επόμενο αποτέλεσμα.

Πρόταση 2.20: Έστω $(X_n) = (X_{n,1}, X_{n,2}, \ldots, X_{n,d})$ ακολουθία τ.δ. και $X = (X_1, X_2, \ldots, X_d)$ ένα ακόμα τ.δ.. Τότε,

$$X_n \xrightarrow{a.s./p} X \iff X_{n,i} \xrightarrow{a.s./p} X_i, \quad \text{για κάθε } i = 1, 2, \ldots, d,$$

δηλαδή, και οι δύο συγκλίσεις είναι ισοδύναμες με τις αντίστοιχες συγκλίσεις κατά συνιστώσα.

Με βάση την παραπάνω πρόταση, καταλήγουμε άμεσα σε πολυδιάστατες επεκτάσεις του Ισχυρού Νόμου των Μεγάλων Αριθμών (Ι.Ν.Μ.Α.), αλλά και του Ασθενούς Νόμου των Μεγάλων Αριθμών (Α.Ν.Μ.Α.).

Πόρισμα 2.5: Έστω (X_n) μία ακολουθία ανεξάρτητων και ισόνομων τ.δ. με $\mathbb{E}\,\|X_1\| < +\infty$. Αν $\overline{X}_n$ είναι ο δειγματικός μέσος, τότε

$$\overline{X}_n \xrightarrow{a.s.} \mathbb{E}(X_1) \qquad \text{(Ι.Ν.Μ.Α)}$$

και

$$\overline{X}_n \xrightarrow{p} \mathbb{E}(X_1) \qquad \text{(Α.Ν.Μ.Α).}$$

Όπως βέβαια και στη μονοδιάστατη περίπτωση, ο Α.Ν.Μ.Α. είναι συνέπεια του Ι.Ν.Μ.Α., καθώς η στοχαστική σύγκλιση προκύπτει άμεσα από την ισχυρή (βλ. Πρόταση 2.21).

Πιο σύνθετη είναι η κατάσταση στη σύγκλιση κατά κατανομή. Η σημασία βέβαια της σύγκλισης αυτής είναι τεράστια και έχει πολλές εφαρμογές, ιδιαίτερα δε στη Στατιστική. Έτσι, έχουν δοθεί αρκετοί ισοδύναμοι χαρακτηρισμοί, κάποιους από τους οποίους θα δούμε σε λίγο. Υπενθυμίζουμε πρώτα τον ορισμό που συνήθως δίνεται για ακολουθίες πραγματικών τ.μ. και είναι και αυτός που παρουσιάζεται ως απλούστερος και στις στοιχειώδεις Πιθανότητες.

Ορισμός 2.31: Μία ακολουθία τ.μ. (X_n) λέμε ότι συγκλίνει *κατά κατανομή* ή *ασθενώς* σε ένα τ.δ. X, και γράφουμε $X_n \xrightarrow{d} X$, αν για κάθε $x \in \mathbb{R}$ με $F(x-) = F(x)$ (σημεία συνεχείας της F),

$$F_n(x) \xrightarrow[n\to\infty]{} F(x),$$

όπου F_n είναι η συνάρτηση κατανομής της X_n και F είναι η σ.κ. της X.

Όπως ήδη βέβαια προαναφέραμε, η συνάρτηση κατανομής είναι ένα βολικό αντικείμενο περιγραφής της κατανομής $\mathbb{P}_X$ μίας πραγματικής, αλλά και μίας διανυσματικής τ.μ.. Δίνουμε τώρα έναν γενικότερο ορισμό, που ισχύει και σε απειροδιάστατους χώρους, αλλάζοντας βέβαια κατάλληλα τους χώρους στους οποίους οι τ.μ. παίρνουν τιμές.

Ορισμός 2.32: Μία ακολουθία τ.δ. (X_n) λέμε ότι συγκλίνει *κατά κατανομή* ή *ασθενώς* σε ένα τ.δ. X, και γράφουμε $X_n \xrightarrow{d} X$, αν για κάθε $B \in \mathcal{B}(R^d)$ με $\mathbb{P}(X \in \partial B) = 0$,

$$\mathbb{P}(X_n \in B) \xrightarrow[n\to\infty]{} \mathbb{P}(X \in B),$$

όπου $\partial B = \overline{B} \setminus B^\circ$ είναι το τοπολογικό σύνορο του B.

Στην ισχυρή και στη στοχαστική σύγκλιση, όλες οι τ.μ. πρέπει να είναι ορισμένες στον ίδιο χώρο πιθανότητας. Στην ασθενή σύγκλιση αυτό δεν είναι απαραίτητο και στον παραπάνω ορισμό μπορούμε να αντικαταστήσουμε την $\mathbb{P}(X_n \in B)$ με $\mathbb{P}_n(X_n \in B)$.

Τονίζουμε, επίσης, το γεγονός ότι η ασθενής σύγκλιση μιας ακολουθίας τ.δ. κατά συνιστώσες, σε αντίθεση με τους άλλους δύο τρόπους σύγκλισης, δεν αρκεί για να συμπεράνουμε ολικά τη σύγκλιση. Υπάρχουν αρκετοί ισοδύναμοι χαρακτηρισμοί, βολικοί κατά περίπτωση, και δίνουμε τώρα μερικούς από αυτούς σε αντιστοιχία με το Θεώρημα 2.1.

Θεώρημα 2.3: Μία ακολουθία τ.δ. (X_n) συγκλίνει κατά κατανομή σε ένα τ.δ. $X \in \mathbb{R}^d$ αν, και μόνο αν,

1. για κάθε σημείο συνεχείας $x \in \mathbb{R}^d$ της F,
$$F_n(x) \xrightarrow[n\to\infty]{} F(x),$$
όπου F_n και F είναι οι σ.κ. των X_n και X αντίστοιχα,
2. κάθε γραμμικός συνδυασμός των συνιστωσών της X_n συγκλίνει κατά κατανομή στον αντίστοιχο γραμμικό συνδυασμό των συνιστωσών της X, δηλαδή, για κάθε $u \in \mathbb{R}^d$
$$u^\top X_n \xrightarrow{d} u^\top X,$$
3. για κάθε $f : \mathbb{R}^d \to \mathbb{R}$ συνεχή και φραγμένη συνάρτηση,
$$\mathbb{E}[f(X_n)] \xrightarrow[n\to\infty]{} \mathbb{E}[f(X)].$$

Στην παρακάτω πρόταση υπενθυμίζουμε τη σχέση που έχουν οι διάφοροι τρόποι σύγκλισης μεταξύ τους.

Πρόταση 2.21: Έστω (X_n) μία ακολουθία τ.δ. και $X \in \mathbb{R}^d$ ένα ακόμα τ.δ. ορισμένα σε κοινό χώρο πιθανότητας. Τότε,
$$X_n \xrightarrow{a.s.} X \quad \Rightarrow \quad X_n \xrightarrow{p} X \quad \Rightarrow \quad X_n \xrightarrow{d} X.$$
Στην ειδική περίπτωση που $X \overset{a.s.}{=} c$, όπου c σταθερά, τότε
$$X_n \xrightarrow{p} c \iff X_n \xrightarrow{d} c.$$

Είμαστε τώρα σε θέση να παρουσιάσουμε το πολυδιάστατο Κεντρικό Οριακό Θεώρημα σε δύο ισοδύναμες μορφές.

Πρόταση 2.22: Έστω (X_n) μία ακολουθία ανεξάρτητων και ισόνομων τ.δ. με $\mathbb{E}\,\|X_1\|^2 < +\infty$. Θέτουμε $\mu_X = \mathbb{E}(X_1)$ και $\Sigma_X = \mathbb{V}(X_1)$. Αν $\left(\overline{X}_n\right)$ είναι η ακολουθία των δειγματικών μέσων και (S_n) η ακολουθία των μερικών αθροισμάτων της (X_n), τότε
$$\sqrt{n}\left(\overline{X}_n - \mu_X\right) \xrightarrow{d} \mathcal{N}_d\left(0_d, \Sigma_X\right)$$
και
$$\frac{S_n - n\,\mu_X}{\sqrt{n}} \xrightarrow{d} \mathcal{N}_d\left(0_d, \Sigma_X\right).$$

Απόδειξη: Θα αποδείξουμε την πρώτη από τις δύο ισοδύναμες μορφές. Η δεύτερη έπεται από την παρατήρηση ότι $S_n = n\overline{X}_n$. Έστω, λοιπόν, U ένα τ.δ. με $U \sim \mathcal{N}_d(0_d, \Sigma_X)$ και $u \in \mathbb{R}^d$. Από το Θεώρημα 2.3, αρκεί να δείξουμε ότι

$$u^\top \left\{ \sqrt{n} \left(\overline{X}_n - \mu_X \right) \right\} \xrightarrow{d} u^\top U \sim \mathcal{N}_1(0, u^\top \Sigma_X u). \tag{2.21}$$

Παρατηρούμε όμως ότι

$$u^\top \left\{ \sqrt{n} \left(\overline{X}_n - \mu_X \right) \right\} = \sqrt{n} \left(\overline{u^\top X_n} - u^\top \mu_X \right) = \sqrt{n} \left(\overline{Y}_n - \mu_Y \right),$$

όπου (Y_n) είναι ακολουθία ανεξάρτητων και ισόνομων τ.μ. με $Y_n = u^\top X_n$ και $\mu_Y = \mathbb{E}(Y_1) = u^\top \mu_X$. Από το κλασικό Κεντρικό Οριακό Θεώρημα, έχουμε

$$\sqrt{n} \left(\overline{Y}_n - \mu_Y \right) \xrightarrow{d} \mathcal{N}_1(0, \mathbb{V}(Y_1)) = \mathcal{N}_1(0, u^\top \Sigma_X u),$$

που είναι η οριακή κατανομή (2.21) στην οποία θέλαμε να φτάσουμε. □

Παρατήρηση 2.19. Αν ο πίνακας Σ_X είναι αντιστρέψιμος, τότε μπορούμε να γράψουμε τη σύγκλιση στην πολυμεταβλητή κανονική στη μορφή

$$\sqrt{n}\, \Sigma_X^{-1/2} \left(\overline{X}_n - \mu_X \right) \xrightarrow{d} \mathcal{N}_d(0_d, I_d),$$

όπου $\Sigma_X^{1/2}$ είναι μία οποιαδήποτε τετραγωνική ρίζα του Σ_X, δηλαδή, οποιαδήποτε επιλογή αντιστρέψιμου πίνακα A που ικανοποιεί τη σχέση $\Sigma_X = AA^\top$.

2.9 Ασυμπτωτικές ιδιότητες συναρτήσεων τυχαίων διανυσμάτων

Θα αναφέρουμε σε αυτήν την ενότητα κάποια οριακά θεωρήματα που μας επιτρέπουν να βρίσκουμε εύκολα τις ασυμπτωτικές ιδιότητες συναρτήσεων τυχαίων μεταβλητών ή τυχαίων διανυσμάτων, όταν γνωρίζουμε τις ασυμπτωτικές ιδιότητες των τελευταίων. Στη συνέχεια, θα κάνουμε κάποιες επεκτάσεις σε ένα ευρύτερο πλαίσιο. Ξεκινάμε με ένα πολύ χρήσιμο θεώρημα, το οποίο απλοποιεί σε σημαντικό βαθμό αποδείξεις σύγκλισης κατά κατανομή.

Θεώρημα 2.4 (αναπαράστασης του Skorokhod): Έστω (X_n), X τυχαία διανύσματα με τιμές στον $\mathbb{R}^d$, τέτοια ώστε $X_n \xrightarrow{d} X$. Τότε, υπάρχουν (X_n^*) και X^* τυχαία διανύσματα με τιμές στον $\mathbb{R}^d$, ορισμένα σε κοινό χώρο πιθανότητας, τέτοια ώστε $X_n^* \overset{d}{=} X_n$, $\forall n \geq 1$, $X^* \overset{d}{=} X$ και $X_n^* \xrightarrow{a.s.} X^*$.

Παρατήρηση 2.20. Το θεώρημα αυτό μας επιτρέπει να αντικαταστήσουμε την $\xrightarrow{d}$ με $\xrightarrow{a.s.}$ σύγκλιση, χρησιμοποιώντας ισόνομα αντίγραφα των αρχικών. Μία χρήση του θεωρήματος θα γίνει παρακάτω.

Θεώρημα Συνεχούς Απεικόνισης (τυχαία διανύσματα) (Θ.Σ.Α.): Έστω (X_n), X τυχαία διανύσματα με τιμές στον $\mathbb{R}^d$. Αν $g : \mathbb{R}^d \to \mathbb{R}^\ell$ και η g είναι συνεχής σε κάποιο $C \in \mathcal{B}(\mathbb{R}^d)$, τέτοιο ώστε $\mathbb{P}(X \in C) = 1$, τότε

$$X_n \xrightarrow{a.s./p/d} X \quad \Longrightarrow \quad g(X_n) \xrightarrow{a.s./p/d} g(X). \tag{2.22}$$

Η απόδειξη του παραπάνω θεωρήματος θα γίνει σε ένα γενικότερο πλαίσιο όταν γίνει αναφορά σε τυχαία στοιχεία.

Θυμόμαστε ότι η χαρακτηριστική συνάρτηση ϕ_X μιας τυχαίας μεταβλητής X είναι η συνάρτηση $\phi_X \colon \mathbb{R} \to \mathbb{C}$, με

$$\phi_X(t) = \mathbb{E}\left[e^{itX}\right] = \mathbb{E}\Big(\cos(tX) + i\sin(tX)\Big) = \mathbb{E}\Big(\cos(tX)\Big) + i\,\mathbb{E}\Big(\sin(tX)\Big), \quad \forall t \in \mathbb{R}.$$

Η έννοια της χαρακτηριστικής συνάρτησης επεκτείνεται φυσιολογικά και στα τυχαία διανύσματα.

> **Ορισμός 2.33:** Αν $X = (X_1, \dots, X_d)$, τότε η συνάρτηση $\phi_X \colon \mathbb{R}^d \to \mathbb{C}$ με $\phi_X(u) = \mathbb{E}\left[e^{i\,u^\top X}\right]$, είναι καλά ορισμένη και καλείται χαρακτηριστική συνάρτηση του τυχαίου διανύσματος X.

Όπως και στη μονοδιάστατη, έτσι και στην πολυδιάστατη περίπτωση η χαρακτηριστική συνάρτηση χαρακτηρίζει πλήρως την κατανομή μίας διανυσματικής τυχαίας μεταβλητής. Μπορούμε, μάλιστα, τώρα να δείξουμε άμεσα ότι

$$X \overset{d}{=} Y \quad \Longleftrightarrow \quad u^\top X \overset{d}{=} u^\top Y, \quad \forall\, u \in \mathbb{R}^d.$$

Θεώρημα Συνέχειας Lévy: Για τις χαρακτηριστικές συναρτήσεις, ισχύει η ισοδυναμία

$$X_n \xrightarrow{d} X \quad \Longleftrightarrow \quad \phi_{X_n}(u) \to \phi_X(u), \quad \forall\, u \in \mathbb{R}^d. \tag{2.23}$$

2.10 Λήμμα Slutsky για τυχαία διανύσματα

Για πραγματικές τυχαίες μεταβλητές χρησιμοποιείται ευρύτατα το εξής πολύ χρήσιμο αποτέλεσμα.

Λήμμα Slutsky (τυχαίες μεταβλητές): Αν X_n, Y_n, X είναι πραγματικές τυχαίες μεταβλητές, και c σταθερά με $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{p} c$, τότε $X_n + Y_n \xrightarrow{d} X + c$, $Y_n X_n \xrightarrow{d} cX$ και, αν $c \neq 0$, τότε $Y_n^{-1} X_n \xrightarrow{d} c^{-1} X$.

Μπορείτε να κάνετε μία απόδειξή του με το θεώρημα του Skorokhod; Θα γενικεύσουμε τώρα αυτό το αποτέλεσμα με τη βοήθεια κάποιων άλλων αποτελεσμάτων που έχουν ανεξάρτητο ενδιαφέρον.

Λήμμα 2.4: Αν $X_n \xrightarrow{d} X$ και $X_n - Y_n \xrightarrow{p} 0$, τότε $Y_n \xrightarrow{d} X$.

Παράδειγμα 2.7: Έστω (X_n) ακολουθία α.ι.τ.μ. με $\mathbb{E}(X_1^4) < +\infty$, όπου τα $\mu = \mathbb{E}(X_1)$ και $\sigma^2 = \mathbb{V}(X_1)$ είναι άγνωστα. Θέτουμε $\widehat{\theta}_n = \left(\widehat{\mu}_n, \widehat{\sigma^2_n}\right)^\top = \left(\overline{X}_n, M_n\right)^\top$, όπου $\overline{X}_n$ είναι ο δειγματικός μέσος και $M_n = \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X}_n)^2$ η (μεροληπτική) δειγματική διασπορά. Σε τυχαίο δείγμα από κανονική κατανομή $N(\mu, \sigma^2)$, οι $\widehat{\mu}_n$ και $\widehat{\sigma^2_n}$ είναι ανεξάρτητες, γενικά όμως όχι. Αναζητούμε την ασυμπτωτική κατανομή του $\sqrt{n}\left(\widehat{\theta}_n - \theta\right)$. Θέτουμε $N_n = \frac{1}{n}\sum_{i=1}^n (X_i - \mu)^2$ και, τότε,

$$M_n - N_n = \frac{1}{n}\sum_{i=1}^n \left[(X_i - \overline{X}_n)^2 - (X_i - \mu)^2\right] = -\left(\overline{X}_n - \mu\right)^2$$

και

$$\begin{aligned} \sqrt{n}\left(\hat{\theta}_n - \theta\right) &= \sqrt{n}\left[\begin{pmatrix} \overline{X}_n \\ M_n \end{pmatrix} - \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}\right] \\ &= \sqrt{n}\left[\begin{pmatrix} \overline{X}_n \\ N_n \end{pmatrix} - \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}\right] - \sqrt{n}\begin{pmatrix} 0 \\ (\overline{X}_n - \mu)^2 \end{pmatrix}. \end{aligned}$$

Έχουμε ότι $\sqrt{n}(\overline{X}_n - \mu) \xrightarrow{d} N(0, \sigma^2)$ και $\overline{X}_n - \mu \xrightarrow{p} 0$. Από το μονοδιάστατο Slutsky, έχουμε

$$\sqrt{n}(\overline{X}_n - \mu)^2 = (\overline{X}_n - \mu) \cdot \sqrt{n}\,(\overline{X}_n - \mu) \xrightarrow{d} 0 \cdot N(0, \sigma^2) = 0,$$

και άρα συμπεραίνουμε ότι

$$\begin{pmatrix} 0 \\ -\sqrt{n}(\overline{X}_n - \mu)^2 \end{pmatrix} \xrightarrow{p} \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Από τα παραπάνω, το αρχικό πρόβλημα ανάγεται στην εύρεση της οριακής κατανομής της $\sqrt{n}\left[(\overline{X}_n \quad N_n)^\top - (\mu \quad \sigma^2)^\top\right]$. Παρατηρούμε τώρα ότι

$$N_n = \frac{1}{n}\sum_{i=1}^n (X_i - \mu)^2 =: \overline{(X - \mu)^2_{[n]}},$$

άρα

$$\sqrt{n}\left[\begin{pmatrix} \overline{X}_n \\ \overline{(X - \mu)^2_{[n]}} \end{pmatrix} - \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}\right] = \sqrt{n}\left(\overline{Y}_n - \mathbb{E}(Y_1)\right),$$

όπου $Y_n = \begin{pmatrix} X_n \\ (X_n-\mu)^2 \end{pmatrix} = g(X_n)$ για $g(x) = (x,\ (x-\mu)^2)^\top$ και οι $g(X_n)$ είναι α.ι.τ.δ. (διότι οι X_n είναι α.ι.τ.μ.), με

$$\mathbb{E}(Y_n) = \mathbb{E}(Y_1) = \begin{pmatrix} \mathbb{E}(X_1) \\ \mathbb{E}(X_1-\mu)^2 \end{pmatrix} = \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}.$$

Προφανώς, ο πίνακας διασπορών-συνδιασπορών Σ υπάρχει και

$$\Sigma := \mathbb{V}(Y_1) = \begin{pmatrix} \mathbb{V}(X_1) & \mathbb{C}\left(X_1, (X_1-\mu)^2\right) \\ \mathbb{C}\left(X_1, (X_1-\mu)^2\right) & \mathbb{V}\left((X_1-\mu)^2\right) \end{pmatrix}.$$

Υπολογίζουμε τα στοιχεία του πίνακα:

- $\mathbb{C}\left(X_1, (X_1-\mu)^2\right) = \mathbb{C}\left(X_1-\mu, (X_1-\mu)^2\right) = \underbrace{\mathbb{E}\left((X_1-\mu)^3\right)}_{\mu_3} - \underbrace{\mathbb{E}(X_1-\mu)}_{0}\mathbb{E}(X_1-\mu)^2 = \mu_3$
- $\mathbb{V}\left((X_1-\mu)^2\right) = \mathbb{E}(X_1-\mu)^4 - \mathbb{E}^2(X_1-\mu)^2 = \mu_4 - \sigma^4$.

Από τα παραπάνω, συμπεραίνουμε ότι

$$\Sigma = \begin{pmatrix} \sigma^2 & \mu_3 \\ \mu_3 & \mu_4-\sigma^4 \end{pmatrix}.$$

Με εφαρμογή του πολυδιάστατου Κεντρικού Οριακού Θεωρήματος για την ακολουθία των α.ι.τ.δ. (Y_n), και λόγω των παραπάνω σχέσεων, έχουμε τελικά ότι

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow{d} \mathcal{N}_2(0, \Sigma).$$

Λήμμα 2.5: Αν $X_n \xrightarrow{d} X$ και $Y_n \xrightarrow{p} c$, τότε $(X_n, Y_n) \xrightarrow{d} (X, c)$.

Απόδειξη: Έχουμε ότι $Y_n - c \xrightarrow{p} 0$, άρα

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

Από το προηγούμενο λήμμα αρκεί λοιπόν να δειχθεί ότι $(X_n, c) \xrightarrow{d} (X, c)$. Όμως, από υπόθεση, $X_n \xrightarrow{d} X$ και έτσι το αποτέλεσμα έπεται άμεσα από οποιονδήποτε από τους χαρακτηρισμούς της σύγκλισης κατά κατανομή (π.χ., με γραμμικούς συνδυασμούς). □

Αποδείξαμε το Λήμμα 2.5 από το Λήμμα 2.4. Αποδείξτε τώρα το Λήμμα 2.4 από το Λήμμα 2.5 και το Θ.Σ.Α.. Μπορούμε να αποδείξουμε το Λήμμα 2.5 και από το θεώρημα του Skorokhod (να γίνει).

Θεώρημα 2.5 (Γενίκευση Λήμματος Slutsky): Αν X_n, Y_n, X τυχαία διανύσματα (ή πίνακες), c σταθερό διάνυσμα (ή πίνακας) και $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{p} c$, τότε όταν οι πράξεις είναι καλά ορισμένες $X_n + Y_n \xrightarrow{d} X + c$, $Y_n X_n \xrightarrow{d} cX$ και, αν c αντιστρέψιμος, τότε $Y_n^{-1} X_n \xrightarrow{d} c^{-1} X$.

Απόδειξη: Αφού $X_n \xrightarrow{d} X$ και $Y_n \xrightarrow{p} c$, από το Λήμμα 2.5 παίρνουμε ότι $(X_n, Y_n) \xrightarrow{d} (X, c)$. Οι συναρτήσεις $g(x, y) = x + y, yx, y^{-1}x$ είναι συνεχείς (η τελευταία περιορίζοντάς την στα αντιστρέψιμα y). Από το Θεώρημα Συνεχούς Απεικόνισης, παίρνουμε ότι $g(X_n) \xrightarrow{d} g(X)$, όπως θέλαμε. □

Παρατήρηση 2.21. Αν $X_n \xrightarrow{as/p} X$, $Y_n \xrightarrow{as/p} Y$, τότε $X_n + Y_n \xrightarrow{as/p} X + Y$, $X_n Y_n \xrightarrow{as/p} XY$ και $Y_n^{-1} X_n \xrightarrow{as/p} Y^{-1} X$, αν υπάρχουν τα αντίστροφα. Γενικά όμως δεν ισχύει για τη σύγκλιση κατά κατανομή. Σκεφτείτε γιατί εδώ δεν μπορεί να εφαρμοστεί η τεχνική του Skorokhod.

Πόρισμα 2.6: Αν (r_n) ακολουθία με $r_n \to +\infty$ και $r_n(X_n - \theta) \xrightarrow{d} X$, τότε $X_n \xrightarrow{p} \theta$. Ειδικά για εκτιμήτριες $\widehat{\theta}_n$, ισχύει ότι αν η ασυμπτωτική κατανομή τους προκύπτει όπως παραπάνω, τότε η $\hat{\theta}_n$ είναι συνεπής.

Απόδειξη: Έχουμε ότι

$$X_n = X_n - \theta + \theta = \frac{1}{r_n} r_n(X_n - \theta) + \theta.$$

Όμως $\dfrac{1}{r_n} \to 0$, $\quad r_n(X_n - \theta) \xrightarrow{d} X$ και θ σταθερό, άρα με εφαρμογή του Λήμματος Slutsky παίρνουμε ότι $X_n \xrightarrow{d} 0 \cdot X + \theta = \theta$. □

2.11 Μέθοδος Δέλτα

Η μέθοδος δέλτα είναι μία από τις πιο χρήσιμες μεθόδους για να βρίσκουμε την ασυμπτωτική κατανομή εκτιμητριών που είναι συναρτήσεις κάποιας άλλης εκτιμήτριας (συνήθως μέσω plug-in) της οποίας γνωρίζουμε την ασυμπτωτική της κατανομή. Παρουσιάζουμε τη μέθοδο κατ'ευθείαν στην πολυδιάστατη περίπτωση.

Θεώρημα 2.6: Έστω $\phi : D_\phi \subset \mathbb{R}^d \to \mathbb{R}^m$ μία διαφορίσιμη απεικόνιση στο $\theta \in D_\phi$ και $X_n \in D_\phi$ για κάθε $n \geq 1$. Αν (r_n) ακολουθία με $r_n \to +\infty$ και $r_n(X_n - \theta) \xrightarrow{d} X$, τότε

$$r_n\Big(\phi(X_n) - \phi(\theta)\Big) \xrightarrow{d} \Phi'_\theta \cdot X, \tag{2.24}$$

όπου $\Phi'_\theta := \left(\dfrac{\partial \phi_i}{\partial x_j}\right)\Bigg|_{x=\theta}$ είναι ο Ιακωβιανός πίνακας της ϕ υπολογισμένος στο θ.

Ισοδύναμα, το παραπάνω αποτέλεσμα αναδιατυπώνεται και ως εξής:

$$r_n\Big(\phi(X_n) - \phi(\theta)\Big) \xrightarrow{d} \phi_\theta'(X), \tag{2.25}$$

όπου η $\phi_\theta' : \mathbb{R}^d \to \mathbb{R}^m$ είναι η παράγωγος απεικόνιση ή το διαφορικό της ϕ στο θ.

Απόδειξη:

Θυμίζουμε πρώτα την εξής σημαντική παρατήρηση.

Παρατήρηση του Καραθεοδωρή (για διανυσματικές συναρτήσεις): Έστω $\phi : D_\Phi \subset \mathbb{R}^d \to \mathbb{R}^m$. Η ϕ είναι διαφορίσιμη στο $\theta \in D_\phi^\circ$, με Ιακωβιανό πίνακα Φ_θ' αν, και μόνο αν, υπάρχει συνάρτηση $g : \mathbb{R}^d \to \mathbb{R}^{m\times d}$ με $\lim\limits_{x\to\theta} g(x) = g(\theta) = \Phi_\theta'$ και τέτοια ώστε $\phi(x) = \phi(\theta) + g(x)(x-\theta)$.

Σύμφωνα με τις υποθέσεις του θεωρήματος και την Παρατήρηση του Καραθεοδωρή, έχουμε

$$r_n\Big(\phi(X_n) - \phi(\theta)\Big) = r_n\, g(X_n)(X_n - \theta) = g(X_n)\cdot r_n(X_n-\theta). \tag{2.26}$$

Όμως, $r_n(X_n - \theta) \xrightarrow{d} X \Rightarrow X_n \xrightarrow{p} \theta$ (από το Πόρισμα 2.6). Επιπλέον η g είναι συνεχής στο θ, και άρα από το Θεώρημα Συνεχούς Απεικόνισης έπεται ότι $g(X_n) \xrightarrow{p} g(\theta) = \Phi_\theta'$. Με μία εφαρμογή του Slutsky στο δεξί μέλος της (2.26) έχουμε

$$r_n\Big(\phi(X_n) - \phi(\theta)\Big) \xrightarrow{d} \Phi_\theta' \cdot X.$$

□

Πόρισμα 2.7: Με τις υποθέσεις του Θεωρήματος 2.6, αν $\sqrt{n}(X_n - \theta) \xrightarrow{d} \mathcal{N}_d(0,\Sigma)$, τότε

$$\sqrt{n}\,\Big(\phi(X_n) - \phi(\theta)\Big) \xrightarrow{d} \mathcal{N}_d\Big(0, \Phi_\theta'\,\Sigma\,(\Phi_\theta')^\top\Big). \tag{2.27}$$

Ειδικά για $d = 1$, έχουμε

$$\sqrt{n}\,\Big(\phi(X_n) - \phi(\theta)\Big) \xrightarrow{d} \mathcal{N}\Big(0, (\phi'(\theta))^2\,\sigma^2\Big). \tag{2.28}$$

Παράδειγμα 2.8: Έστω (X_n) ακολουθία α.ι.τ.μ. με $\mathbb{E}(X_1^4) < +\infty$.

1. Να βρεθεί η οριακή κατανομή της ακολουθίας $\sqrt{n}\left[\begin{pmatrix}\overline{X}_n\\ \overline{X_n^2}\end{pmatrix} - \begin{pmatrix}m_1\\ m_2\end{pmatrix}\right]$, όπου $m_1 = \mathbb{E}(X_1) = \mu$ και $m_2 = \mathbb{E}(X_1^2)$ (γενικά, $m_k = \mathbb{E}(X_1^k)$).
2. Με τη βοήθεια της μεθόδου Δέλτα, να βρεθεί η οριακή κατανομή της ακολουθίας $\sqrt{n}\left[\begin{pmatrix}\overline{X}_n\\ M_n\end{pmatrix} - \begin{pmatrix}\mu\\ \sigma^2\end{pmatrix}\right]$, όπου $\sigma^2 = \mathbb{V}(X_1)$ και $M_n = \frac{1}{n}\sum\limits_{i=1}^{n}(X_i -$

$\overline{X}_n)^2$.

Λύση: (i) Θέτουμε $g(x) = \begin{pmatrix} x \\ x^2 \end{pmatrix}$ και $Y_n = g(X_n) = \begin{pmatrix} X_n \\ X_n^2 \end{pmatrix}$. Τότε,

$$\overline{Y}_n = \overline{g(X_n)} = \begin{pmatrix} \overline{X}_n \\ \overline{X_n^2} \end{pmatrix} \quad \text{και} \quad \mathbb{E}(Y_1) = \begin{pmatrix} \mathbb{E}(X_1) \\ \mathbb{E}(X_1^2) \end{pmatrix} = \begin{pmatrix} m_1 \\ m_2 \end{pmatrix}.$$

Από υπόθεση, η (X_n) είναι ακολουθία α.ι.τ.μ., άρα και η (Y_n) είναι ακολουθία α.ι.τ.δ., με $\mathbb{E}\|Y_1\|^2 = \mathbb{E}\left(X_1^2 + X_1^4\right) = \underbrace{\mathbb{E}(X_1^2)}_{<+\infty} + \underbrace{\mathbb{E}(X_1^4)}_{<+\infty} < +\infty$. Άρα εφαρμόζεται το πολυδιάστατο Κεντρικό Οριακό Θεώρημα και παίρνουμε ότι

$$\sqrt{n}\left[\begin{pmatrix} \overline{X}_n \\ \overline{X_n^2} \end{pmatrix} - \begin{pmatrix} m_1 \\ m_2 \end{pmatrix}\right] = \sqrt{n}\Big(\overline{Y}_n - \mathbb{E}(Y_1)\Big) \xrightarrow{d} \mathcal{N}_2(0, \Sigma),$$

όπου

$$\Sigma = \begin{pmatrix} \mathbb{V}(X_1) & \mathbb{C}(X_1, X_1^2) \\ \mathbb{C}(X_1, X_1^2) & \mathbb{V}(X_1^2) \end{pmatrix}.$$

Όμως

$$\begin{aligned} \mathbb{V}(X_1) &= \mathbb{E}(X_1^2) - \mathbb{E}^2(X_1) = m_2 - m_1^2, \\ \mathbb{V}(X_1^2) &= \mathbb{E}(X_1^4) - \mathbb{E}^2(X_1^2) = m_4 - m_2^2, \\ \mathbb{C}(X_1, X_1^2) &= \mathbb{E}(X_1^3) - \mathbb{E}(X_1)\,\mathbb{E}(X_1^2) = m_3 - m_1 m_2. \end{aligned}$$

Τελικά,

$$\sqrt{n}\left[\begin{pmatrix} \overline{X}_n \\ \overline{X_n^2} \end{pmatrix} - \begin{pmatrix} m_1 \\ m_2 \end{pmatrix}\right] \xrightarrow{d} \mathcal{N}_2(0, \Sigma), \quad \text{με} \quad \Sigma = \begin{pmatrix} m_2 - m_1^2 & m_3 - m_1 m_2 \\ m_3 - m_1 m_2 & m_4 - m_2^2 \end{pmatrix}.$$

(ii) Έχουμε $M_n = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2 = \overline{X_n^2} - \overline{X}_n^2$, άρα θέτουμε $\phi(x, y) = (x, y - x^2)^\top$, η οποία είναι διαφορίσιμη σε όλο το $\mathbb{R}^2$. Έχουμε τώρα ότι $\phi\begin{pmatrix} \overline{X}_n \\ \overline{X_n^2} \end{pmatrix} = \begin{pmatrix} \overline{X}_n \\ M_n \end{pmatrix}$ και, επιπλέον, $\phi\begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \begin{pmatrix} m_1 \\ m_2 - m_1^2 \end{pmatrix} = \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}$, επομένως ισχύει ότι

$$\sqrt{n}\left[\begin{pmatrix} \overline{X}_n \\ M_n \end{pmatrix} - \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}\right] = \sqrt{n}\left[\phi\begin{pmatrix} \overline{X}_n \\ \overline{X_n^2} \end{pmatrix} - \phi\begin{pmatrix} m_1 \\ m_2 \end{pmatrix}\right].$$

Από τη μέθοδο Δέλτα και το (i), έχουμε ότι

$$\sqrt{n}\left[\begin{pmatrix}\overline{X}_n\\ M_n\end{pmatrix}-\begin{pmatrix}\mu\\ \sigma^2\end{pmatrix}\right]\xrightarrow{d}\mathcal{N}_2(0,\Sigma')=\mathcal{N}_2\Big(0,\Phi'_\theta\,\Sigma\,(\Phi'_\theta)^\top\Big),$$

όπου $\theta=(m_1,m_2)$ και $\Phi'_\theta=\begin{pmatrix}\dfrac{\partial\phi_1}{\partial x}(\theta) & \dfrac{\partial\phi_1}{\partial y}(\theta)\\ \dfrac{\partial\phi_2}{\partial x}(\theta) & \dfrac{\partial\phi_2}{\partial y}(\theta)\end{pmatrix}=\begin{pmatrix}1 & 0\\ -2m_1 & 1\end{pmatrix}.$

Τελικά,

$$\begin{aligned}\Sigma'&=\begin{pmatrix}1 & 0\\ -2m_1 & 1\end{pmatrix}\begin{pmatrix}m_2-m_1^2 & m_3-m_1m_2\\ m_3-m_1m_2 & m_4-m_2^2\end{pmatrix}\begin{pmatrix}1 & -2m_1\\ 0 & 1\end{pmatrix}=\cdots=\\ &=\begin{pmatrix}m_2-m_1^2 & m_3-3m_2m_1+2m_1^3\\ m_3-3m_2m_1+2m_1^3 & m_4-4m_3m_1-m_2^2+8m_2m_1^2-4m_1^4\end{pmatrix}.\end{aligned}$$

Με άλλη μέθοδο είχαμε βρει (βλ. Παράδειγμα 2.7) ότι ο ασυμπτωτικός πίνακας συνδιακύμανσης

$$\Sigma'=\begin{pmatrix}\sigma^2 & \mu_3\\ \mu_3 & \mu_4-\sigma^4\end{pmatrix}.$$

Είναι εύκολο να δειχθεί ότι οι παραπάνω δύο εκφράσεις συμπίπτουν. Ενδεικτικά, π.χ.,

$$\begin{aligned}\mu_3&=\mathbb{E}(X_1-m_1)^3=\mathbb{E}(X_1^3-3X_1^2m_1+3m_1^2X_1-m_1^3)\\ &=m_3-3m_2m_1+3m_1^2m_1-m_1^3=m_3-3m_2m_1+2m_1^3\end{aligned}$$

και $\sigma^2=m_2-m_1^2$. Συμπληρώστε τη λύση για το $\mu_4-\sigma^4$. □

Ασκήσεις

2.11.1 Έστω V ένας διανυσματικός χώρος. Να δειχθεί ότι ο τρόπος με τον οποίο ορίζεται ο παραγόμενος διανυσματικός χώρος $\langle S\rangle$ ενός υποσυνόλου $S \subset V$ σύμφωνα με τον Ορισμό 2.8 είναι ισοδύναμος με αυτόν που δίνεται στην Παρατήρηση 2.6.

2.11.2 Έστω $\{S_i\}_{i=1}^n$ n ολικά διατεταγμένα σύνολα και $S = \prod_{i=1}^n S_i$ το καρτεσιανό τους γινόμενο. Αν $x, y \in S$, τότε θέτουμε $i_{x,y} = \min\{1 \leq i \leq n \mid x_i < y_i\}$, και $\min \emptyset = n + 1$.

1. Να δειχθεί ότι η *λεξικογραφική διάταξη* έτσι όπως υποδεικνύεται από τον τρόπο που ταξινομούνται οι λέξεις σε ένα λεξικό αντιστοιχεί στον ορισμό $x \leq y$, αν και μόνο αν $i_{x,y} \leq i_{y,x}$.

2. Να δειχθεί ότι η λεξικογραφική διάταξη είναι σχέση ολικής διάταξης στο S.

3. Η *διάταξη κατά συνιστώσες* ή *διάταξη κατά συντεταγμένες* ορίζεται από τη σχέση $x \leq y$, αν και μόνο αν $x_i \leq y_i$ για κάθε $1 \leq i \leq n$. Να δειχθεί ότι η διάταξη αυτή είναι σχέση μερικής διάταξης στο S.

2.11.3 Στην άσκηση αυτή εισάγονται έννοιες στοχαστικής διάταξης μεταξύ τυχαίων μεταβλητών και διανυσμάτων.

1. Έστω F, G οι συναρτήσεις κατανομής 2 τ.μ. X, Y και $\overline{F}, \overline{G}$ οι συμπληρωματικές τους συναρτήσεις με $\overline{F}(x) := \mathbb{P}(X > x)$ για κάθε $x \in \mathbb{R}$ και ανάλογα για τη $\overline{G}$. Ορίζουμε τη σχέση $X \leq_{st} Y$, αν και μόνο αν $\overline{F} \leq \overline{G}$. Να δειχθεί ότι η σχέση $\leq_{st}$ είναι σχέση μερικής διάταξης και χρησιμοποιώντας απ'ευθείας τις συναρτήσεις κατανομής γράφεται ισοδύναμα $F \geq G$. Η διάταξη αυτή καλείται *στοχαστική διάταξη* (stochastic order) και είναι χρήσιμο εργαλείο στη σύγκριση των τ.μ..

2. Συμπεράνετε ότι αν $X \leq_{st} Y$ και $Y \leq_{st} X$, τότε $X \stackrel{d}{=} Y$.

3. Έστω F, G οι συναρτήσεις κατανομής 2 τυχαίων διανυσμάτων X, Y και $\overline{F}(x) := \mathbb{P}(X > x)$ για κάθε $x \in \mathbb{R}^d$ και ανάλογα για τη $\overline{G}$. Ορίζουμε τις σχέσεις $X \leq_{uo} Y$ (upper orthant order), αν και μόνο αν $\overline{F} \leq \overline{G}$ (διάταξη κατά συνιστώσα) και $X \leq_{lo} Y$ (lower orthant order), αν και μόνο αν $F \geq G$ (κατά συνιστώσα). Να δειχθεί ότι οι σχέσεις $\leq_{uo}$ και $\leq_{lo}$ είναι σχέσεις μερικής διάταξης.

4. Να δειχθεί ότι $X \leq_{uo} Y$ και $Y \leq_{uo} X$, τότε $X \stackrel{d}{=} Y$. Να δειχθεί το ίδιο και για τη σχέση $\leq_{lo}$.

5. Να εξεταστεί αν οι παραπάνω σχέσεις διάταξης είναι ισοδύναμες.

2.11.4 Έστω $X = (X_1, X_2),\ Y = (Y_1, Y_2)$ δύο τυχαία διανύσματα. Να δειχθεί ότι

$$X_1 \stackrel{d}{=} Y_1, \quad X_2 \stackrel{d}{=} Y_2 \quad \not\Rightarrow \quad X \stackrel{d}{=} Y.$$

Να δώσετε παραδείγματα για τη μη ισχύ της συνεπαγωγής για

1. X_1, X_2 διακριτές τ.μ.
2. X_1, X_2 απόλυτα συνεχείς τ.μ.

2.11.5 Έστω S σύνολο και $\mathcal{D}$ μία οικογένεια υποσυνόλων του S. Η $\mathcal{D}$ καλείται *λ–κλάση* (ή *λ-σύστημα*) στο S, αν ικανοποιεί τις ιδιότητες: (i) $\mathcal{D} \neq \emptyset$, (ii) είναι κλειστή στα συμπληρώματα και (iii) είναι κλειστή στις αριθμήσιμες ξένες ενώσεις. Αποδεικνύεται ότι η λ–κλάση είναι ισοδύναμη με τη λεγόμενη *κλάση Dynkin* που ορίζεται από μία αναδιατύπωση των παραπάνω συνθηκών (παραλείπεται). Όταν λοιπόν λέμε κλάση Dynkin θα αναφερόμαστε στις παραπάνω ιδιοτητες. Παρατηρείστε ότι κάθε σ–άλγεβρα είναι κλάση Dynkin. Στην άσκηση αυτή θα δειχθεί ότι ένα μέτρο πιθανότητας δεν καθορίζεται μονοσήμαντα από τις τιμές του σε έναν γεννήτορα και άρα χρειάζονται επιπλέον συνθήκες, όπως αυτές στην απόδειξη του Θεωρήματος 2.1–(i).

1. Έστω μ, ν δύο μέτρα πιθανότητας στον $(S, \mathcal{A})$ και $\mathcal{D} = \{A \in \mathcal{A} \mid \mu(A) = \nu(A)\}$, δηλ. η οικογένεια των υποσυνόλων που τα μέτρα συμφωνούν. Να δειχθεί ότι η $\mathcal{D}$ είναι κλάση Dynkin.
2. Έστω $S = \{1, 2, 3, 4\}$ και $\mathcal{D} = \{\emptyset, \{1, 2\}, \{2, 3\}, \{3, 4\}, \{1, 4\}, S\}$. Να δειχθεί ότι η $\mathcal{D}$ είναι κλάση Dynkin, αλλά όχι σ–άλγεβρα.
3. Να δειχθεί ότι η $\mathcal{D}$ είναι γεννήτορας της $\mathcal{A}$.
4. Βρείτε 2 μέτρα πιθανότητας μ, ν που να συμφωνούν στην οικογένεια $\mathcal{D}$ με $\mu(A) = \nu(A) = 0.5$ για κάθε δισύνολο της $\mathcal{D}$, αλλά είναι διαφορετικά μεταξύ τους.

2.11.6 Έστω $X = (X_1, X_2, ..., X_d)$ ένα τυχαίο διάνυσμα. Ν.δ.ο.

$$E(X_i^2) < +\infty, \quad \forall 1 \leq i \leq d \quad \Leftrightarrow \quad E||X||^2 < \infty.$$

2.11.7 Έστω $\mathbf{X} \in \mathbb{R}^d, \mathbf{Y} \in \mathbb{R}^s$ τ.δ. με $E||X||^2, E||Y||^2 < +\infty$ και A, B, b, c πίνακες και διανύσματα κατάλληλης διάστασης για να ορίζονται οι παρακάτω πράξεις. Ν.δ.ο.

1. $C(\mathbf{X}, \mathbf{Y}) = E\left[(\mathbf{X} - \mu_{\mathbf{X}})(\mathbf{Y} - \mu_{\mathbf{Y}})^\top\right] = E(\mathbf{X}\mathbf{Y}^T) - \mu_{\mathbf{X}}\mu_{\mathbf{Y}}^\top$.
2. $C(A\mathbf{X} + b, B\mathbf{Y} + c) = AC(\mathbf{X}, \mathbf{Y})B^\top$.
3. $C(X_1 + X_2, Y_1 + Y_2) = \sum_{1 \leq i,j \leq 2} C(X_i, Y_j)$.

Συμπεράνετε τις ιδιότητες της διασποράς.

2.11.8 Έστω $\mathbf{X} \in \mathbb{R}^d, \mathbf{Y} \in \mathbb{R}^s$ δύο ανεξάρτητα τυχαία διανύσματα. Ν.δ.ο.

1. $E(\mathbf{X}\mathbf{Y}^\top) = E(\mathbf{X})E(\mathbf{Y}^\top)$.
2. $C(\mathbf{X}, \mathbf{Y}) = \mathbf{O}_{d \times s}$.
3. $V(\mathbf{X} + \mathbf{Y}) = V(\mathbf{X}) + V(\mathbf{Y})$, όπου εδώ $d = s$.

2.11.9 Έστω $\mathbf{X} = (X_1, X_2, ..., X_d) \sim \mathcal{M}_d(n, p), \quad p^T = (p_1, p_2, ..., p_d)$.

1. Να βρεθεί η κατανομή του τ.δ.

$$Y = \mathbf{X}_{i_{1:k}} \coloneqq (X_{i_1}, X_{i_2}, ..., X_{i_k}), \qquad i_1, i_2, ..., i_k \in \{1, 2, ..., d\},$$

   όπου $k < d$ και $i_1, i_2, ..., i_k$ διακεκριμένοι δείκτες.

2. Προσδιορίστε τα $E(Y)$ και $V(Y)$.
3. Να βρεθεί η συνάρτηση πιθανότητας του Y.
4. Σε μια ακολουθία ανεξάρτητων ρίψεων ενός δίκαιου ζαριού να βρεθεί η από κοινού κατανομή των N_1 και N_6, όπου N_i = # εμφανίσεων της ένδειξης $i \in \{1, 6\}$ σε 10 ρίψεις.

2.11.10
1. Να εξεταστεί αν ο πίνακας $\Sigma_1 = V(X)$ με $\mathbf{X} \sim \mathcal{M}_d(n, p)$, $p^\top = (p_1, p_2, ..., p_d)$ είναι αντιστρέψιμος.
2. Να γίνει το ίδιο για τον $\Sigma_2 = V(X)$ με $\mathbf{X} \sim \mathcal{M}_d^*(n, p)$ και να βρεθεί ο αντίστροφος αν υπάρχει.

2.11.11 Έστω $\mathbf{X}, \mathbf{S}$ τα τυχαία διανύσματα όπου:

$$\mathbf{X} = (X_1, X_2, ..., X_d) \sim \mathcal{M}_d^\star(n, p), \quad p = (p_1, p_2, ..., p_d).$$
$$\mathbf{S} = (S_1, S_2, ..., S_d), \quad \text{όπου} \quad S_k = X_1 + X_2 + ... + X_k, \quad \forall 1 \leq k \leq d.$$

1. Να βρεθεί το στήριγμα και η συνάρτηση πιθανότητας του $\mathbf{S}$.
2. Προσδιορίστε τα $E(\mathbf{S})$ και $V(\mathbf{S})$.

ΚΕΦΑΛΑΙΟ 3

Η ΕΜΠΕΙΡΙΚΗ ΠΡΟΣΕΓΓΙΣΗ

Η πιο ακραία περίπτωση, την οποία θα εξετάσουμε τώρα είναι όταν υποθέτουμε ότι η κατανομή του υπό μελέτη χαρακτηριστικού είναι εντελώς άγνωστη. Εφόσον η κατανομή μίας τ.μ. χαρακτηρίζεται πλήρως από τη συνάρτηση κατανομής της (που ορίζεται και για διακριτές, αλλά και για συνεχείς τ.μ.) μπορούμε λοιπόν να δουλεύουμε με συναρτήσεις κατανομής (σ.κ.). Θα υποθέσουμε λοιπόν τώρα ότι η άγνωστη παράμετρος του στατιστικού μοντέλου είναι η ίδια η σ.κ. F. Θα συμβολίζουμε το χώρο όλων των σ.κ. με $\mathcal{F}$ και θα παίζει το ρόλο του παραμετρικού χώρου, όπως έπαιζε το Θ στην παραμετρική Στατιστική. Έτσι η έκφραση $\theta \in \Theta$ μετατρέπεται εδώ σε $F \in \mathcal{F}$. Υπάρχει βέβαια μία μικρή δυσκολία. Ενώ στην παραμετρική Στατιστική $\Theta \subset \mathbb{R}^d$ για κατάλληλο πεπερασμένο d, εδώ η κλάση $\mathcal{F}$ είναι απειροδιάστατη και έτσι το πρόβλημα επιλογής του κατάλληλου συναρτησιακού χώρου μέσα στον οποίο θα λυθούν τα διάφορα στατιστικά προβλήματα που θα μας απασχολήσουν είναι πιο λεπτό. Ας ξεκινήσουμε τώρα με το πρώτο πρόβλημα που αφορά την εύρεση κατάλληλης εκτιμήτριας για την άγνωστη συνάρτηση κατανομής.

3.1 Εμπειρική Συνάρτηση Κατανομής

Έστω λοιπόν ότι διαθέτουμε n παρατηρήσεις $x_1, x_2, \ldots, x_n$, οι οποίες υποθέτουμε ότι αποτελούν n ανεξάρτητες πραγματοποιήσεις κάποιου υπό μελέτη χαρακτηριστικού $X \sim F$. Με τη βοήθεια των δεδομένων αυτών, η μόνη πληροφορία που έχουμε είναι ότι η X μπορεί να πάρει αυτές τις τιμές. Αν κάποιος τώρα μας ζητήσει μία τιμή

Σάμης Τρέβεζας (2023). «Μη Παραμετρική Στατιστική».
Αθήνα: Κάλλιπος, Ανοικτές Ακαδημαϊκές Εκδόσεις.
Creative Commons Αναφορά Δημιουργού - Μη Εμπορική Χρήση - Παρόμοια Διανομή 4.0

αυτής της τ.μ. δεν θα είχαμε κάποιο λόγο να ευνοήσουμε κάποια από αυτές. Θα μπορούσαμε να τις βάλουμε σε μία κληρωτίδα και να τραβήξουμε στην τύχη μία από αυτές. Με άλλα λόγια θα επιλέγαμε κάποια από αυτές ισοπίθανα. Αυτή η απλή σκέψη μας οδηγεί στην κατασκευή μίας καινούριας τ.μ. X_n^* που ακολουθεί τη διακριτή ομοιόμορφη κατανομή στο σύνολο $\{x_1, x_2, \dots x_n\}$. Θα ήταν βέβαια ιδιαίτερα επιθυμητό η συνάρτηση κατανομής F_n^* της X_n^* να προσεγγίζει (με κάποια έννοια) την άγνωστη συνάρτηση κατανομής F για μεγάλο n. Εφόσον βέβαια η X_n^* είναι διακριτή τ.μ. έχουμε ότι $\forall x \in \mathbb{R}$,

$$F_n^*(x) = \mathbb{P}(X_n^* \le x) = \sum_{k: x_k \le x} \mathbb{P}(X_n^* = x_k) = \sum_{k=1}^{n} \frac{1}{n} \mathbf{1}_{x_k \le x} = \frac{\sum_{k=1}^{n} \mathbf{1}_{x_k \le x}}{n}. \tag{3.1}$$

Τη διαδικασία αυτή μπορούμε να την επαναλαμβάνουμε ανεξάρτητα από το ποιες είναι οι συγκεκριμένες τιμές των δεδομένων $\{x_k\}_{k=1}^n$. Έτσι, όπως και στην παραμετρική Στατιστική, οδηγούμαστε στην κατασκευή μίας εκτιμήτριας $\mathbb{F}_n(x)$ του $F(x)$, μία για κάθε $x \in \mathbb{R}$. Συγκεκριμένα, αντικαθιστώντας με τ.μ. έχουμε

$$\mathbb{F}_n(x) = \frac{\sum_{k=1}^{n} \mathbf{1}_{X_k \le x}}{n}, \qquad \forall x \in \mathbb{R}. \tag{3.2}$$

Ορισμός 3.1: Έστω $\{X_k\}_{k=1}^n$ ένα τυχαίο δείγμα από κάποια κατανομή με σ.κ. $F \in \mathcal{F}$. Η τυχαία συνάρτηση $\mathbb{F}_n := \mathbb{F}_{X_1,\dots,X_n} = (\mathbb{F}_n(x))_{x\in\mathbb{R}}$, όπου $\mathbb{F}_n(x)$ δίνεται από την (3.2) καλείται *εμπειρική συνάρτηση κατανομής*.

Παρατήρηση 3.1. Κάθε πραγματοποίηση F_n^* της $\mathbb{F}_n$ αντιστοιχεί σε μία συνάρτηση κατανομής όπως προαναφέραμε και άρα παίρνει τιμές στο χώρο συναρτήσεων $\mathcal{F}$. Μπορούμε λοιπόν να σκεφτόμαστε την $\mathbb{F}_n$ ως μία τυχαία συνάρτηση του x ή ισοδύναμα ως μία τυχαία μεταβλητή με τιμές συναρτήσεις στον χώρο $\mathcal{F}$. Ο ρόλος της εμπειρικής συνάρτησης κατανομής ως στοχαστικής διαδικασίας θα τονιστεί στη συνέχεια.

Παρατήρηση 3.2. Η σχέση (3.2) μας δείχνει ότι η εμπειρική συνάρτηση κατανομής $\mathbb{F}_n$ έχει μία πολύ απλή ερμηνεία. Για κάθε $x \in \mathbb{R}$, η τιμή $\mathbb{F}_n(x)$ εκφράζει το ποσοστό των παρατηρήσεων του δείγματος που είναι μικρότερες ή ίσες του x.

Καθώς η ε.σ.κ. είναι εκτιμήτρια της άγνωστης συνάρτησης κατανομής μας ενδιαφέρει βέβαια να μελετήσουμε διάφορες ιδιότητές της.

3.2 Σημειακές ιδιότητες της εμπειρικής συνάρτησης κατανομής

Εξετάζουμε πρώτα τις σημειακές ιδιότητες της $\mathbb{F}_n$.

Πρόταση 3.1: Για κάθε $x \in \mathbb{R}$ ισχύει ότι

$$\mathbb{F}_n(x) \sim \frac{1}{n} \mathcal{B}\boldsymbol{in}(n, F(x)).$$

Απόδειξη: Έχουμε ότι

$$\mathbb{F}_n(x) = \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{(-\infty,x]}(X_k) = \frac{1}{n} \sum_{k=1}^{n} Y_k, \tag{3.3}$$

όπου $Y_k = \mathbf{1}_{(-\infty,x]}(X_k)$ για κάθε $1 \leq k \leq n$. Είναι φανερό ότι $Y_k \sim \mathcal{B}\boldsymbol{e}(F(x))$ για κάθε $1 \leq k \leq n$ ως δείκτριες τυχαίες μεταβλητές και επιπλέον οι $\{Y_k\}_{k=1}^n$ είναι ανεξάρτητες, καθώς είναι της μορφής $g(X_k)$ και οι $\{X_k\}_{k=1}^n$ είναι ανεξάρτητες τυχαίες μεταβλητές. Από τα παραπάνω και τη σχέση (3.3) συμπεραίνουμε το ζητούμενο καθώς το $\sum_{k=1}^{n} Y_k$ ακολουθεί $\mathcal{B}\boldsymbol{in}(n, F(x))$ ως άθροισμα n ανεξάρτητων $\mathcal{B}\boldsymbol{e}(F(x))$.

□

Πόρισμα 3.1: Έστω $\{X_k\}_{k=1}^n$ ένα τυχαίο δείγμα από την F. Για κάθε $x \in \mathbb{R}$:

$$\begin{aligned} \mathbb{E}\left(\mathbb{F}_n(x)\right) &= F(x), \\ \mathbb{V}\left(\mathbb{F}_n(x)\right) &= \frac{1}{n} F(x)(1 - F(x)). \end{aligned}$$

Παρατήρηση 3.3. Η παραπάνω σχέση που αφορά τη μέση τιμή εκφράζει την **αμεροληψία** της $\mathbb{F}_n(x)$ ως εκτιμήτριας του $F(x)$, για κάθε $x \in \mathbb{R}$. Από τη σχέση που αφορά τη διασπορά συμπεραίνουμε ότι είναι και **συνεπής εκτιμήτρια** του $F(x)$, για κάθε $x \in \mathbb{R}$. Πράγματι, ισχύει ότι $\mathbb{V}(\mathbb{F}_n(x)) \to 0$ και η $\mathbb{F}_n(x)$ είναι αμερόληπτη εκτιμήτρια του $\mathbb{F}(x)$. Άρα σε κάθε σημείο το Μέσο Τετραγωνικό Σφάλμα μηδενίζεται ασυμπτωτικά, συμπεραίνοντας έτσι τη σημειακή (ασθενή) συνέπεια. Μάλιστα, θα δείξουμε στην επόμενη Πρόταση ότι έχουμε σημειακή **ισχυρή συνέπεια**.

Πρόταση 3.2: Η εμπειρική συνάρτηση κατανομής $\mathbb{F}_n$ είναι σημειακά ισχυρά συνεπής, δηλ., για κάθε $x \in \mathbb{R}$:

$$\mathbb{F}_n(x) \xrightarrow{a.s.} F(x).$$

Απόδειξη: Από τη σχέση (3.3) συμπεραίνουμε ότι η $\mathbb{F}_n(x)$ είναι δειγματικός μέσος της ακολουθίας των α.ι.τ.μ. $(Y_n)_{n\geq 1}$, με $Y_n \sim \mathcal{B}\boldsymbol{e}(F(x))$. Εφόσον $\mathbb{E}(Y_1) = F(x)$, ικανοποιούνται οι προϋποθέσεις του I.N.M.A. και άρα $\mathbb{F}_n(x) = \overline{Y}_n \xrightarrow{a.s.} F(x)$.

□

Εύκολα προκύπτει και η ασυμπτωτική κανονικότητα της εκτιμήτριας $\mathbb{F}_n(x)$.

Πρόταση 3.3: Η εμπειρική συνάρτηση κατανομής $\mathbb{F}_n$ είναι σημειακά ασυμπτωτικά κανονική, δηλ., για κάθε $x \in \mathbb{R}$:

$$\sqrt{n}\,\left(\mathbb{F}_n(x) - F(x)\right) \xrightarrow{d} \mathcal{N}\left(0,\, F(x)\left(1 - F(x)\right)\right).$$

Απόδειξη: Παρόμοια με την απόδειξη της Πρότασης 3.2, έχουμε $\mathbb{F}_n(x) = \overline{Y}_n$, όπου $(Y_n)_{n\geq 1}$ ακολουθία α.ι.τ.μ. με $Y_n \sim \mathcal{B}e(F(x))$. Ισχύουν οι προϋποθέσεις του Κεντρικού Οριακού Θεωρήματος, με $\mathbb{V}(Y_1) = F(x)(1 - F(x))$. Επομένως,

$$\sqrt{n}\,\left(\mathbb{F}_n(x) - F(x)\right) = \sqrt{n}\,\left(\overline{Y}_n - \mathbb{E}(Y_1)\right) \xrightarrow{d} \mathcal{N}\left(0,\, \mathbb{V}(Y_1)\right) = \mathcal{N}\left(0,\, F(x)\left(1 - F(x)\right)\right).$$

□

Παρατήρηση 3.4. Το παραπάνω αποτέλεσμα προκύπτει και από την Πρόταση 3.1 και την προσέγγιση της Διωνυμικής κατανομής $\mathcal{B}in(n, p)$ από την Κανονική κατανομή $\mathcal{N}(np,\, np(1-p))$. Επομένως, για $p = F(x)$ και μεγάλο n:

$$n\,\mathbb{F}_n(x) \approx \mathcal{N}\left(n\,F(x),\, n\,F(x)\left(1 - F(x)\right)\right) \implies \mathbb{F}_n(x) \approx \mathcal{N}\left(F(x),\, \frac{F(x)(1 - F(x))}{n}\right),$$

όπου με $\approx$ εννοούμε ότι ακολουθεί προσεγγιστικά αυτήν την κατανομή.

Τα παραπάνω αποτελέσματα μας δείχνουν τη σημειακή συμπεριφορά της εμπειρικής συνάρτησης κατανομής, δηλαδή το πώς συμπεριφέρονται οι ακολουθίες των τυχαίων μεταβλητών $\left(\mathbb{F}_n(x)\right)_n$ για κάθε $x \in \mathbb{R}$. Η σημειακή ασυμπτωτική κανονικότητα μας δίνει τη δυνατότητα να κατασκευάσουμε ασυμπτωτικά διαστήματα εμπιστοσύνης (Δ.Ε.) για το $F(x)$ σε κάθε σημείο $x \in \mathbb{R}$, επιτρέποντας έτσι την κατασκευή σημειακών λωρίδων εμπιστοσύνης για την F. Είναι φανερό ότι η κατασκευή Δ.Ε. σε συγκεκριμένα σημεία ανάγεται στην κατασκευή Δ.Ε. για την παράμετρο p ενός τ.δ. από $\mathcal{B}e(p)$. Για παράδειγμα, ας υποθέσουμε ότι μας ενδιαφέρει να εκτιμήσουμε την πιθανότητα $p := F(1)$ και να κατασκευάσουμε ένα διάστημα εμπιστοσύνης (Δ.Ε.), π.χ. 95%–Δ.Ε.. Η εμπειρική εκτιμήτρια του $F(1)$ είναι η

$$\mathbb{F}_n(1) = \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{(-\infty,1]}(X_k) = \overline{Y}_n,$$

όπου $Y_n = \mathbf{1}_{(-\infty,1]}(X_n) \sim \mathcal{B}e(p)$. Από την Πρόταση 3.1 έχουμε $n\mathbb{F}_n(1) \sim \mathrm{Bin}(n, F(1))$. Υπάρχουν πολλές διαφορετικές μέθοδοι κατασκευής Δ.Ε. για παράμετρο που αντιστοιχεί σε πιθανότητα και ο ενδιαφερόμενος αναγνώστης παραπέμπεται στο άρθρο των [10] για μία σύγκριση πολλών από αυτών. Στην ενότητα αυτή θα αντιπαραβάλλουμε 3 διαφορετικές μεθόδους κατασκευής Δ.Ε., μία επονομαζόμενη “ακριβής” και 2 ασυμπτωτικές. Ξεκινάμε με τις τελευταίες, καθώς η κατασκευή τους είναι ευκολότερη. Στις ασκήσεις..

Πρόταση 3.4: Έστω $Y_1, \ldots, Y_n$ ένα τ.δ. από $\mathcal{B}e(p)$, $0 < p < 1$ και θέτουμε $\widehat{p}_n = \overline{Y}_n$. Δύο ασυμπτωτικά $(1-\alpha)-$Δ.Ε. για το p, δίνονται ως εξής:

$$I_{1-\alpha}^{Wald}(\widehat{p}_n) = \widehat{p}_n \pm z_{a/2}\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}} \cap (0,1),$$

το οποίο καλείται *Wald-Δ.Ε.* ή *ασυμπτωτικό* και

$$I_{1-\alpha}^{Wilson}(\widehat{p}_n) = \left(1 + \frac{z_{a/2}^2}{n}\right)^{-1} \left(\widehat{p}_n + \frac{z_{a/2}^2}{2n} \pm z_{a/2}\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n) + \frac{z_{a/2}^2}{4n}}{n}}\right) \cap (0,1),$$

το οποίο καλείται *Wilson-Δ.Ε.*.

Παράδειγμα 3.1: Η επονομαζόμενη "ακριβής" μέθοδος καλείται και Clopper-Pearson [[11]] και υπολογίζεται ως εξής:

$$I_{1-a}^{CP}(s) = \left[\left(1 + \frac{n-s+1}{s\, f_{2s,2(n-s+1);1-a/2}}\right)^{-1}, \left(1 + \frac{n-s}{(s+1)\, f_{2(s+1),2(n-s);a/2}}\right)^{-1}\right],$$

όπου $s = y_1 + \ldots + y_n$ είναι το πλήθος των μονάδων στο δείγμα μεγέθους n και $f_{n_1,n_2;a}$ το a-άνω ποσοστημόριο της κατανομής Fisher $\mathcal{F}_{n_1,n_2}$ με n_1 και n_2 βαθμούς ελευθερίας. Να σημειώσουμε βέβαια εδώ ότι ο όρος "ακριβές" είναι παραπλανητικός καθώς η μέθοδος αυτή εξασφαλίζει ότι η διαστηματική εκτιμήτρια που κατασκευάζεται με τον παραπάνω τρόπο ($s \to S_n$) περιλαμβάνει την άγνωστη παράμετρο με πιθανότητα τουλάχιστον $1-a$ για όλα τα $p \in (0,1)$, αλλά όχι ακριβώς, και η ισότητα επιτυγχάνεται μόνο ως μέγιστο κάτω φράγμα για όλες τις τιμές του p. Η μέθοδος CP δεν είναι ασυμπτωτική. Παρουσιάζουμε τώρα 2 ασυμπτωτικές μεθόδους που χρησιμοποιούνται αρκετά για την κατασκευή ασυμπτωτικών Δ.Ε., τη μέθοδο Wald και τη μέθοδο Wilson.
Για την κατασκευή ενός ασυμπτωτικού Δ.Ε. τύπου Wald, ξεκινάμε με εφαρμογή του Κ.Ο.Θ.

$$\sqrt{n} \cdot \frac{\overline{Y}_n - p}{\sqrt{p(1-p))}} \xrightarrow{d} Z \sim N(0,1)$$

και με αντικατάσταση, όπου $\widehat{p}_n = \mathbb{F}_n(1) = \overline{Y}_n$ παίρνουμε

$$\sqrt{n} \cdot \frac{\widehat{p}_n - p}{\sqrt{p(1-p)}} \xrightarrow{d} Z \sim N(0,1). \tag{3.4}$$

Καθώς $\widehat{p}_n \xrightarrow{p} p$ (προκύπτει άμεσα από το παραπάνω ή απ'ευθείας με χρήση του Α.Ν.Μ.Α.), από το Θ.Σ.Α. προκύπτει ότι $\sqrt{\widehat{p}_n(1-\widehat{p}_n)} \xrightarrow{p} \sqrt{p(1-p)}$. Από το Λήμμα του Slutsky συμπεραίνουμε ότι

$$\sqrt{n} \cdot \frac{\widehat{p}_n - p}{\sqrt{\widehat{p}_n(1-\widehat{p}_n)}} \xrightarrow{d} Z \sim N(0,1).$$

Λύνοντας ως προς $p(= F(1))$ τη σχέση

$$-z_{a/2} \leq \sqrt{n} \cdot \frac{\widehat{p}_n - p}{\sqrt{\widehat{p}_n(1-\widehat{p}_n)}} \leq z_{a/2},$$

όπου $z_{a/2}$ είναι το $a/2$-άνω ποσοστημόριο της τυπικής κανονικής παίρνουμε

$$\widehat{p}_n - z_{a/2}\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}} \leq p \leq \widehat{p}_n + z_{a/2}\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}}.$$

Τελικά, λαμβάνοντας υπ'όψη και τον περιορισμό ότι $0 < p < 1$, συρρικνώνουμε το Δ.Ε. στη μορφή $I_n(1) = [L_n(1), U_n(1)]$, όπου

$$L_n(1) = \max\left\{0,\, \widehat{p}_n - z_{a/2}\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}}\right\} \tag{3.5}$$

και

$$U_n(1) = \min\left\{1,\, \widehat{p}_n + z_{a/2}\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}}\right\}. \tag{3.6}$$

Αυτού του τύπου το ασυμπτωτικό Δ.Ε. είναι γνωστό και ως τύπου Wald.
Για την κατασκευή ενός ασυμπτωτικού Δ.Ε. τύπου Wilson, χρησιμοποιούμε απ'ευθείας ως ασυμπτωτικό οδηγό την ποσότητα που εμφανίζεται στην (3.4) και λύνουμε ως προς $p = F(1)$ τη διπλή ανισότητα που προκύπτει όταν χρησιμοποιήσουμε τα $a/2$-άνω και $a/2$-κάτω ποσοστημόρια της τυπικής κανονικής. Έτσι, λύνοντας μία ανισότητα δευτέρου βαθμού προκύπτει ένα ακόμα $(1-a)$–Δ.Ε. για το $F(1)$ με άκρα:

$$\left(1 + \frac{z_{a/2}^2}{n}\right)^{-1} \left(\widehat{p}_n + \frac{z_{a/2}^2}{2n} \pm z_{a/2}\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n) + \frac{z_{a/2}^2}{4n}}{n}}\right)$$

βάζοντας κάτω και άνω φράγμα το 0 και το 1 αντίστοιχα, όπως στην περίπτωση των (3.5) και (3.6) για το διάστημα τύπου Wald.

Αντιπαραβάλλουμε τώρα τα αποτελέσματα των 3 μεθόδων για μεγέθη δείγματος $n = 10$ και $n = 100$ στα παρακάτω δεδομένα που παράγονται από $\mathcal{N}(0,1)$. Η υλοποίηση στην R φαίνεται παρακάτω:

```r
library(binom)
set.seed(1293)
```

```
n1 <- 10; n2 <- 100
x <- rnorm(100)
s1 <- sum(x[1:n1] < 1); s2<-sum(x < 1)

>binom.confint(s1, n1, conf.level = 0.95, methods = c("exact","
   asymptotic","wilson"))

method        x  n   mean   lower      upper
1 wald        8 10   0.8 0.5520820 1.0479180
2 CP          8 10   0.8 0.4439045 0.9747893
3 wilson      8 10   0.8 0.4901625 0.9433178

> binom.confint(s2, n2, conf.level = 0.95, methods = c("exact","
   asymptotic","wilson"))
      method  x   n  mean    lower     upper
1 wald        85 100 0.85 0.7800153 0.9199847
2 CP          85 100 0.85 0.7646925 0.9135456
3 wilson      85 100 0.85 0.7671644 0.9069401
> pnorm(1)
[1] 0.8413447
```

Σημειώνουμε ότι η πραγματική τιμή της ζητούμενης πιθανότητας είναι $p = \Phi(1) \approx 0.841$ (όπως φαίνεται παραπάνω) και επομένως και με τις τρεις μεθόδους, αλλά και για τα δύο μεγέθη δείγματος τα 95%–Δ.Ε. καλύπτουν την πραγματική τιμή. Παρατηρούμε ότι τα Wald-Δ.Ε. είναι στενότερα και δεν γίνεται διόρθωση άνω φράγματος (δείτε για $n = 10$). Από την άλλη, τα Wilson-Δ.Ε. είναι πιο κοντά στα CP. Αναφορικά με την αβεβαιότητα στην εκτίμηση, το πλάτος των Δ.Ε. μικραίνει αρκετά καθώς το μέγεθος του δείγματος αυξάνει από $n = 10$ σε $n = 100$ (όπως αναμένεται από τη θεωρία). Τέλος, στην περίπτωση αυτή που έχουμε προσομοιωμένα δεδομένα μπορούμε να υπολογίσουμε και τις πιθανότητες κάλυψης των παραπάνω Δ.Ε., ώστε να αντιπαραβάλλουμε τις τιμές τους με την ονομαστική τους αξία, που αντιστοιχεί βέβαια στον κοινό τους συντελεστή εμπιστοσύνης $1 - a = 0.95$. Γενικότερα, τα Wilson και τα CP ανταποκρίνονται καλύτερα στο ζητούμενο συντελεστή, όμως τα CP είναι συνήθως ελαφρώς συντηρητικά.

Τα παραπάνω Δ.Ε. έχουν προκύψει από μη παραμετρικές υποθέσεις για την F. Ενώ προσομοιώσαμε από την $\mathcal{N}(0,1)$, προσποιηθήκαμε ότι δε γνωρίζουμε τίποτα για την άγνωστη κατανομή. Ας υποθέσουμε τώρα ότι γνωρίζουμε ότι έχει προσομοιωθεί από μία οικογένεια κανονικών κατανομών της μορφής $N(\mu, 1)$ και μόνο η παράμετρος μ είναι άγνωστη. Μπορούμε έτσι να εκτιμήσουμε την $\mathbb{P}(X \leq 1) = F(1)$ παραμετρικά, βρίσκοντας κατάλληλο 95%–Δ.Ε., όταν $n = 10$ και όταν $n = 100$. Τώρα, η περίπτωση αυτή αντιστοιχεί στην εκτίμηση και στην κατασκευή ενός $(1-a)$–Δ.Ε. για την "παράμετρο"

$$F(1) = \mathbb{P}_\mu(X \le 1) = \mathbb{P}_\mu(X - \mu \le 1 - \mu) = \Phi(1-\mu) =: g(\mu) \tag{3.7}$$

σε ένα τ.δ. $\{X_i\}_{i=1}^n$ από $\mathcal{N}(\mu, 1)$. Με τη μέθοδο της μέγιστης πιθανοφάνειας, τη μέθοδο της αντικατάστασης (plug-in) και την ιδιότητα του αναλλοίωτου μπορούμε εύκολα να βρούμε την ε.μ.π. του $F(1)$ και ένα Δ.Ε.. Πράγματι, καθώς η ε.μ.π. $\widehat{\mu}_n = \overline{X}_n$ έχουμε

$$\widehat{F(1)}_n = g(\widehat{\mu}_n) = \Phi(1 - \widehat{\mu}_n) = \Phi(1 - \overline{X}_n).$$

Όμως $\overline{X}_n \sim N\left(\mu, \frac{1}{n}\right)$ και άρα $\sqrt{n}\left(\overline{X}_n - \mu\right) \sim N(0,1)$. Από το $(1-a)$–Δ.Ε. για το μ:

$$\left[\overline{X}_n - \frac{z_{a/2}}{\sqrt{n}}\,,\, \overline{X}_n + \frac{z_{a/2}}{\sqrt{n}}\right],$$

και το ότι η g είναι γνησίως φθίνουσα ως προς μ έχουμε $(1-a)$–Δ.Ε. για το $F(1)$:

$$\left[g\left(\overline{X}_n + \frac{Z_{a/2}}{\sqrt{n}}\right)\,,\, g\left(\overline{X}_n - \frac{Z_{a/2}}{\sqrt{n}}\right)\right],$$

όπου η g θυμίζουμε ότι δίνεται από τη σχέση (3.7). Σε αυτήν την περίπτωση, η πιθανότητα κάλυψης συμπίπτει με το συντελεστή εμπιστοσύνης για κάθε τιμή του $\mu \in \mathbb{R}$. Τα 95%–Δ.Ε. που προκύπτουν για $n = 10$ και $n = 100$ δίνονται παρακάτω:

```
[1] CI10  = [0.6355148, 0.9436400]
[1] CI100 = [0.7922591, 0.8861448]
```

Η αντιπαραβολή των παραμετρικών Δ.Ε. με τα αντίστοιχα μη παραμετρικά αναδεικνύει το γεγονός ότι τα παραμετρικά είναι πιο στενά. Αυτό συμβαίνει διότι στα παραμετρικά υποθέτουμε ότι η άγνωστη κατανομή ανήκει σε μια συγκεκριμένη οικογένεια κατανομών. Έτσι, η πληροφορία αυτή μειώνει την αβεβαιότητα που συνδέεται με την υπό εκτίμηση παράμετρο.

Η ανάγκη μελέτης της καθολικής συμπεριφοράς της εμπειρικής συνάρτησης κατανομής μας οδηγεί στο πλαίσιο των στοχαστικών διαδικασιών (ανελίξεων) και μία μικρή εισαγωγή τους γίνεται στην Ενότητα 3.4. Καθώς μία στοχαστική διαδικασία εμπλέκει γενικά απειροδιάστατους χώρους, στην επόμενη ενότητα δίνονται κάποια θεωρητικά προαπαιτούμενα για την καλύτερη κατανόηση του θεωρητικού υποβάθρου.

3.3 Κύλινδροι και κυλινδρική σ-άλγεβρα*

Ιδιαίτερο ενδιαφέρον για μας θα έχουν συναρτήσεις της μορφής $f\colon T \to \mathbb{R}$, όπου το T είναι κάποιο υποσύνολο του $\mathbb{R}$. Υπενθυμίζουμε ότι με $\mathbb{R}^T$ συμβολίζουμε το χώρο όλων αυτών των συναρτήσεων και κάθε μία από αυτές αναπαρίσταται ως ένα σημείο του. Συνήθως το T θα είναι απειροσύνολο και έτσι ο $\mathbb{R}^T$ εκφράζει ένα απειροδιάστατο καρτεσιανό γινόμενο αντιγράφων του $\mathbb{R}$.

Ορισμός 3.2: Έστω T ένα αυθαίρετο δεικτοσύνολο. Θέτουμε

$$\mathbb{R}^T := \{x = (x_t)_{t\in T} \mid x_t \in \mathbb{R},\ \text{για κάθε } t \in T\}.$$

- Αν $t \in T$, τότε η συνάρτηση

$$\pi_t \colon \mathbb{R}^T \to \mathbb{R} \quad \text{με} \quad \pi_t(x) = x_t,$$

καλείται *t-προβολή* και αντιστοιχίζει σε κάθε σημείο του $\mathbb{R}^T$ την t-συνιστώσα του. Αν $A \subset \mathbb{R}$ και $t \in T$ τότε

$$\pi_t^{-1}(A) = \{x \in \mathbb{R}^T \mid x_t \in A\}.$$

- Αν $I \subset T$, τότε γενικεύοντας το παραπάνω, ορίζουμε τη συνάρτηση

$$\pi_I \colon \mathbb{R}^T \to \mathbb{R}^I \quad \text{με} \quad \pi_I(x) = (x_t)_{t\in I} =: x_I,$$

που καλείται *I-προβολή* και αντιστοιχίζει σε κάθε σημείο του $\mathbb{R}^T$ τις συνιστώσες που αντιστοιχούν στο σύνολο I. Αντίστοιχα, αν $A \subset \mathbb{R}^I$, τότε το σύνολο

$$\pi_I^{-1}(A) = \{x \in \mathbb{R}^T \mid x_I \in A\}.$$

- Αν $I \subset T$ πεπερασμένο και $A = \prod_{t\in I} A_t$, δηλαδή το A είναι ένα ορθογώνιο σύνολο του $\mathbb{R}^I$, τότε το σύνολο

$$\pi_I^{-1}(A) = \bigcap_{t\in I} \pi_t^{-1}(A_t) = \{x \in \mathbb{R}^T \mid x_t \in A_t,\ \text{για κάθε } t \in I\},$$

καλείται *κύλινδρος* του $\mathbb{R}^T$. Ο κύλινδρος καλείται *ανοικτός*, αν A_t είναι ανοικτό υποσύνολο του $\mathbb{R}$ για κάθε $t \in I$, και καλείται *μετρήσιμος*, αν $A_t \in \mathcal{B}(\mathbb{R})$ για κάθε $t \in I$.

Από τους παραπάνω ορισμούς είναι φανερό ότι κάθε κυλινδρικό σύνολο αφήνει ελεύθερες όλες τις συνιστώσες του $\mathbb{R}^T$, εκτός από ένα πεπερασμένο πλήθος από αυτές, τις οποίες και υποχρεώνει να βρίσκονται μέσα σε ένα πεπερασμένης διάστασης ορθογώνιο σύνολο. Είναι λοιπόν φανερό, ότι ένα κυλινδρικό σύνολο σε έναν απειροδιάστατο χώρο αντιστοιχεί σε μία μεγάλη συλλογή συναρτήσεων (κάθε σημείο είναι συνάρτηση), που ο μόνος περιορισμός είναι οι τιμές τους σε ένα πεπερασμένο σύνολο σημείων να βρίσκονται μέσα σε κάποια προκαθορισμένα σύνολα. Με τη βοήθεια των κυλινδρικών συνόλων μπορούμε να ορίσουμε χρήσιμες σ–άλγεβρες στο χώρο $\mathbb{R}^T$.

Ορισμός 3.3: Έστω $\mathcal{C}^{op}$ η συλλογή όλων των ανοικτών κυλίνδρων, δηλαδή,

$$\mathcal{C}^{op} := \{\pi_I^{-1}(A) : I \subset T \text{ πεπερασμένο και } A \text{ ανοικτό ορθογώνιο του } \mathbb{R}^{|I|}\}.$$

και $\mathcal{T}^{pr} := \tau(\mathcal{C}^{op})$ την τοπολογία που παράγεται από τους ανοικτούς κυλίνδρους. Η $\mathcal{T}^{pr}$ καλείται *τοπολογία γινόμενο* ή *καρτεσιανή τοπολογία* του $\mathbb{R}^T$. Μπορούμε λοιπόν να ορίσουμε την Borel σ–άλγεβρα του $\mathbb{R}^T$,

$$\mathcal{B}(\mathbb{R}^T) = \sigma(\mathcal{T}^{pr}) = \sigma\big(\tau(\mathcal{C}^{op})\big), \tag{3.8}$$

ως τη σ–άλγεβρα που παράγεται από την τοπολογία γινόμενο.
Αν επιπλέον $\mathcal{C}^m$ είναι η συλλογή όλων των μετρήσιμων κυλίνδρων του $\mathbb{R}^T$, δηλαδή

$$\mathcal{C}^m := \{\pi_I^{-1}(A) : I \subset T \text{ πεπερασμένο και } A \text{ μετρήσιμο ορθογώνιο του } \mathbb{R}^{|I|}\},$$

τότε η σ–άλγεβρα που παράγεται από τους μετρήσιμους κυλίνδρους, καλείται *κυλινδρική σ–άλγεβρα* ή *σ–άλγεβρα γινόμενο*. Θα τη συμβολίζουμε με $\mathcal{B}^T(\mathbb{R})$ και άρα

$$\mathcal{B}^T(\mathbb{R}) = \sigma(\mathcal{C}^m). \tag{3.9}$$

Παρατήρηση 3.5. Η $\mathcal{C}^{op}$ είναι βάση κάποιας τοπολογίας αφού είναι κλειστή στις πεπερασμένες τομές και καλύπτει τον $\mathbb{R}^T$ (βλ. Θεώρημα 1.1 και Παρατήρηση 1.23; ασκ. **??**). Αυτό δίνει άμεσα μία περιγραφή των μελών της τοπολογίας γινόμενο $\mathcal{T}^{pr}$ ως ενώσεων ανοικτών κυλίνδρων (οποιουδήποτε πλήθους).

Στην επόμενη πρόταση δίνεται μία περιγραφή της κυλινδρικής σ–άλγεβρας και της σχέσης της με τη σ–άλγεβρα των συνόλων Borel του $\mathbb{R}^T$.

Πρόταση 3.5: Έστω $\mathcal{B}^T(\mathbb{R})$ η κυλινδρική σ–άλγεβρα και $\mathcal{B}(\mathbb{R}^T)$ η σ–άλγεβρα Borel που παράγεται από την τοπολογία γινόμενο. Έστω επίσης S αριθμήσιμο υποσύνολο του T, $\pi_S\colon \mathbb{R}^T \to \mathbb{R}^S$ η S-προβολή και

$$\mathcal{C}^{m;S} := \left\{ \pi_S^{-1}(A) \;\middle|\; A = \prod_{t\in S} A_t,\; A_t \in \mathcal{B}(\mathbb{R}),\; \text{για κάθε } t \in S \right\}. \tag{3.10}$$

Τότε,

- αν το T είναι αριθμήσιμο,

$$\mathcal{B}^T(\mathbb{R}) = \sigma(\mathcal{C}^{m;T}) = \mathcal{B}(\mathbb{R}^T), \tag{3.11}$$

- αν το T είναι υπεραριθμήσιμο,

$$\mathcal{B}^T(\mathbb{R}) = \bigcup \Big\{ \sigma(\mathcal{C}^{m;S}) \mid S \subset T \text{ αριθμήσιμο} \Big\} \subsetneqq \mathcal{B}(\mathbb{R}^T). \tag{3.12}$$

Παρατήρηση 3.6. Είναι σημαντικό να παρατηρήσουμε ότι, σύμφωνα με την παραπάνω αναπαράσταση για T υπεραριθμήσιμο, για να είναι ένα υποσύνολο του $\mathbb{R}^T$ μετρήσιμο ως προς την κυλινδρική σ–άλγεβρα, ικανή και αναγκαία συνθήκη είναι να καθορίζεται πλήρως μόνο από ένα αριθμήσιμο πλήθος συνιστωσών. Με όρους συναρτήσεων, θα λέγαμε ισοδύναμα ότι μετρήσιμες μπορεί να είναι μόνο συλλογές συναρτήσεων $f : T \to \mathbb{R}$ που οι τιμές τους δεσμεύονται ή καθορίζονται μόνο σε ένα αριθμήσιμο υποσύνολο του T. Έτσι, π.χ., αν $T = [0, 1]$ και

$$\begin{aligned} F_1 &= \{f \colon [0,1] \to \mathbb{R} \mid f(0) = f(1) = 0\}, \\ F_2 &= \{f \colon [0,1] \to \mathbb{R} \mid f(q) \in \mathbb{Q},\ \text{για κάθε } q \in \mathbb{Q}\}, \end{aligned}$$

τότε οι F_1, F_2 είναι κυλινδρικά μετρήσιμες. Διαισθητικά, όπως έχουμε αναφέρει, μία σ–άλγεβρα $\mathcal{A}$ εκφράζει πληροφορία, και έτσι η $\mathcal{A}$-μετρησιμότητα μας δίνει ακριβώς τον τύπο των ερωτήσεων που έχει νόημα να κάνουμε. Οι F_1, F_2 πράγματι καθορίζονται από ένα αριθμήσιμο πλήθος συνιστωσών της f. Για να ξέρουμε αν μία $f \in F_1$ ή αν $f \in F_2$, αυτό εξαρτάται από τη γνώση των τιμών της f σε ένα αριθμήσιμο πλήθος σημείων. Δε συμβαίνει όμως το ίδιο, αν θεωρήσουμε π.χ. το σύνολο των συνεχών συναρτήσεων $\mathcal{C}[0, 1]$. Για να μπορούμε να καθορίσουμε αν μία συνάρτηση f είναι συνεχής ή όχι, δεν αρκεί η γνώση της συμπεριφοράς της μόνο σε κάποιο αριθμήσιμο υποσύνολο S του T, όπως απαιτεί η κυλινδρική μετρησιμότητα. Είναι φανερό ότι ακόμα και αν το σύνολο αυτό S είναι πυκνό στο T δεν θα μπορούσαμε να προσδιορίσουμε αν μία f είναι συνεχής ούτε καν σε ένα σημείο του S, αφού δεν θα είχαμε πρόσβαση σε πληροφορία που αφορά τις τιμές της f στα σημεία του $T \setminus S$ που είναι επίσης πυκνό στο T. Το παραπάνω παράδειγμα μας δείχνει ότι οι συνεχείς συναρτήσεις δεν είναι κυλινδρικά μετρήσιμες.

3.4 Στοχαστικές Διαδικασίες

Η εμπειρική συνάρτηση κατανομής $\mathbb{F}_n$ είναι μία συλλογή τυχαίων μεταβλητών $\{\mathbb{F}_n(x)\}_{x\in\mathbb{R}}$, οι οποίες είναι συσχετισμένες. Η μελέτη της συμπεριφοράς της $\mathbb{F}_n$ γίνεται καλύτερα κατανοητή αναγνωρίζοντας ότι αποτελεί στοχαστική διαδικασία. Άρα, ως ακολουθία η $(\mathbb{F}_n)$ είναι μία ακολουθία στοχαστικών διαδικασιών.

Ορισμός 3.4: Έστω $(\Omega, \mathcal{A}, \mathbb{P})$ ένας χώρος πιθανότητας. Μία οικογένεια τυχαίων μεταβλητών $X := (X_t)_{t\in T}$ με $X_t : \Omega \to \mathbb{R}$, για κάθε $t \in T$, και $T \subset \mathbb{R}$ ή $\overline{\mathbb{R}}$, καλείται *στοχαστική διαδικασία* (σ.δ.) ή ανέλιξη, με *δεικτοσύνολο* T. Αν το T είναι διακριτό (συνήθως $T = \mathbb{N}$ ή $\mathbb{Z}$), τότε η σ.δ. καλείται *διακριτού χρόνου*, ενώ αν το T είναι κάποιο διάστημα του $\mathbb{R}$ ή του $\overline{\mathbb{R}}$, τότε καλείται *συνεχούς χρόνου*. Συνήθως $T = [0, +\infty)$ ή $T = \mathbb{R}$. Το ευρύτερο σύνολο μέσα στο οποίο οι X_t παίρνουν τις τιμές τους από κοινού καλείται *χώρος καταστάσεων* και συνήθως θα το συμβολίζουμε με S.

Παρατήρηση 3.7. Σύμφωνα με τον παραπάνω ορισμό, μία στοχαστική διαδικασία X έχει δύο ορίσματα, το $t \in T$ και το $\omega \in \Omega$. Αν σταθεροποιήσουμε το $t \in T$, τότε προκύπτει μία τυχαία μεταβλητή $X_t\colon \Omega \to \mathbb{R}$. Αν σταθεροποιήσουμε το $\omega \in \Omega$, τότε γράφουμε

$$X(\omega) := \big(X_t(\omega)\big)_{t\in T} \in S^T,$$

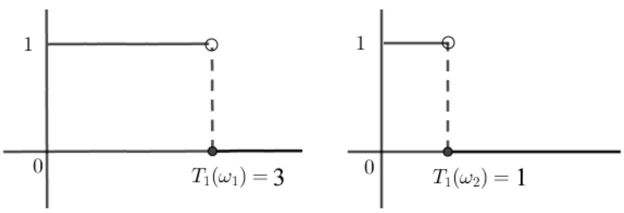
και άρα μία πραγματοποίηση της X αντιστοιχεί σε μία συνάρτηση με πεδίο ορισμού το T και τιμές στο S. Για το λόγο αυτό, μία πραγματοποίηση της X αναφέρεται και ως *δειγματοσυνάρτηση*. Χρησιμοποιείται επίσης η ορολογία *δειγματικό μονοπάτι*, ή απλούστερα *μονοπάτι*. Συχνά, γράφουμε $X = X(t, \omega)$ και άρα στην πρώτη περίπτωση έχουμε $X_t = X(t, \cdot)$, ενώ στη δεύτερη $X(\omega) = X(\cdot, \omega)$.

Παρατήρηση 3.8. Ο χώρος καταστάσεων S συνδέεται με το πεδίο τιμών των X_t. Εδώ μπορεί να αναφέρεται σε ένα πεπερασμένο σύνολο καταστάσεων, σε ένα αριθμήσιμο σύνολο ή σε κάποιο υπεραριθμήσιμο υποσύνολο του $\mathbb{R}$ (συνήθως διάστημα) για πραγματικές τυχαίες μεταβλητές. Η περίπτωση που $S = \mathbb{R}^d$ εμφανίζεται επίσης συχνά και αφορά τυχαία διανύσματα. Στην επόμενη ενότητα γίνεται μία σύντομη αναφορά στην έννοια των τυχαίων στοιχείων που καλύπτουν γενικότερες περιπτώσεις, όπως χώρους καταστάσεων που είναι συναρτησιακοί ή μετρικοί χώροι.

Η μοντελοποίηση ενός προβλήματος μέσω μίας στοχαστικής διαδικασίας, συνήθως συνδέεται με το ενδιαφέρον που έχουμε να μελετήσουμε την εξέλιξη της κατάστασης ενός συστήματος στον χρόνο, είτε αυτός νοείται διακριτός είτε συνεχής. Δίνουμε τώρα ορισμένα παραδείγματα στοχαστικών διαδικασιών.

Παράδειγμα 3.2: Μία ακολουθία πραγματικών τυχαίων μεταβλητών (X_n) είναι μία στοχαστική διαδικασία διακριτού χρόνου με δεικτοσύνολο $T = \mathbb{N}$ και χώρο καταστάσεων το σύνολο $\mathbb{R}$ ή κάποιο υποσύνολό του που υποδεικνύεται συνήθως από τη φύση του προβλήματος που μας απασχολεί. Το είδος εξάρτησης και οι κατανομές των τυχαίων μεταβλητών καθορίζουν και διαφορετικά είδη στοχαστικών διαδικασιών. Για παράδειγμα, η απλούστερη περίπτωση αφορά μία ακολουθία $(X_n)_{n\geq 1}$ ανεξάρτητων και ισόνομων τ.μ.. Παρόλο που η παραπάνω ακολουθία είναι τετριμμένη ως στοχαστική διαδικασία καθώς δεν εμπλέκονται εξαρτήσεις μεταξύ των μεταβλητών, με τη βοήθειά της μπορούμε να χτίσουμε πολύ πιο ενδιαφέρουσες ακολουθίες. Αν $S_0 = 0$ και για $n \geq 1$ θέσουμε $S_n = \sum_{k=1}^{n} X_k$, τότε η ακολουθία $(S_n)_{n\geq 0}$ καλείται *τυχαίος περίπατος* (random walk). Οι τ.μ. $\{X_n\}_{n\geq 1}$ αναφέρονται ως βήματα του τυχαίου περιπάτου και η κοινή κατανομή τους ως βηματική κατανομή. Ο *απλός τυχαίος περίπατος* αφορά την περίπτωση που $X_k \in \{-1, 1\}$, ενώ αν επιπλέον είναι *συμμετρικός*, τότε τα παραπάνω σημεία είναι ισοπίθανα.

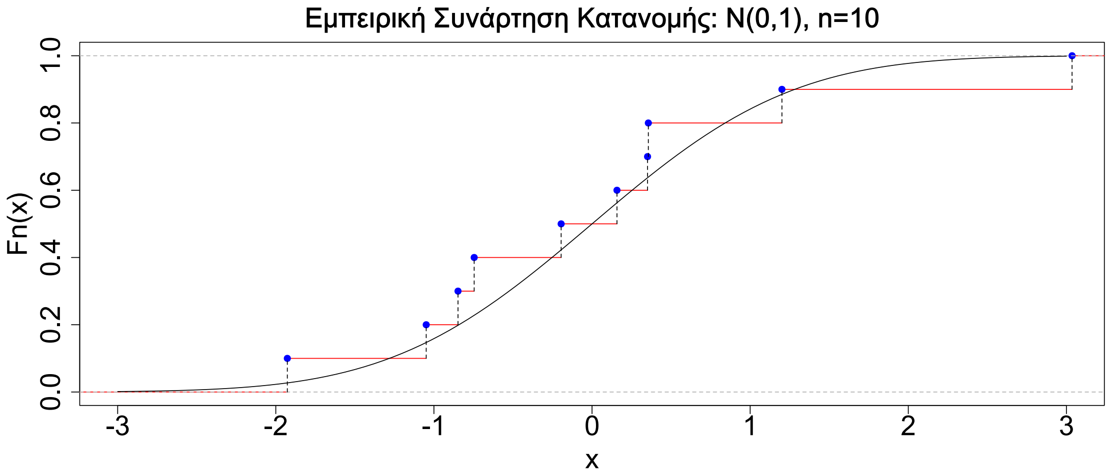
Παράδειγμα 3.3: Έστω T_1 μία τυχαία μεταβλητή που εκφράζει το χρόνο ζωής ενός λαμπτήρα και $X_t = \mathbf{1}_{\{t<T_1\}}$, για κάθε $t \geq 0$. Η προκύπτουσα οικογένεια των τ.μ. $(X_t)_{t\geq 0}$ αποτελεί μία στοχαστική διαδικασία συνεχούς χρόνου, που εκφράζει την κατάσταση



Σχημα 3.1: Δύο πραγματοποιήσεις της σ.δ. $(X_t)_{t\geq 0}$

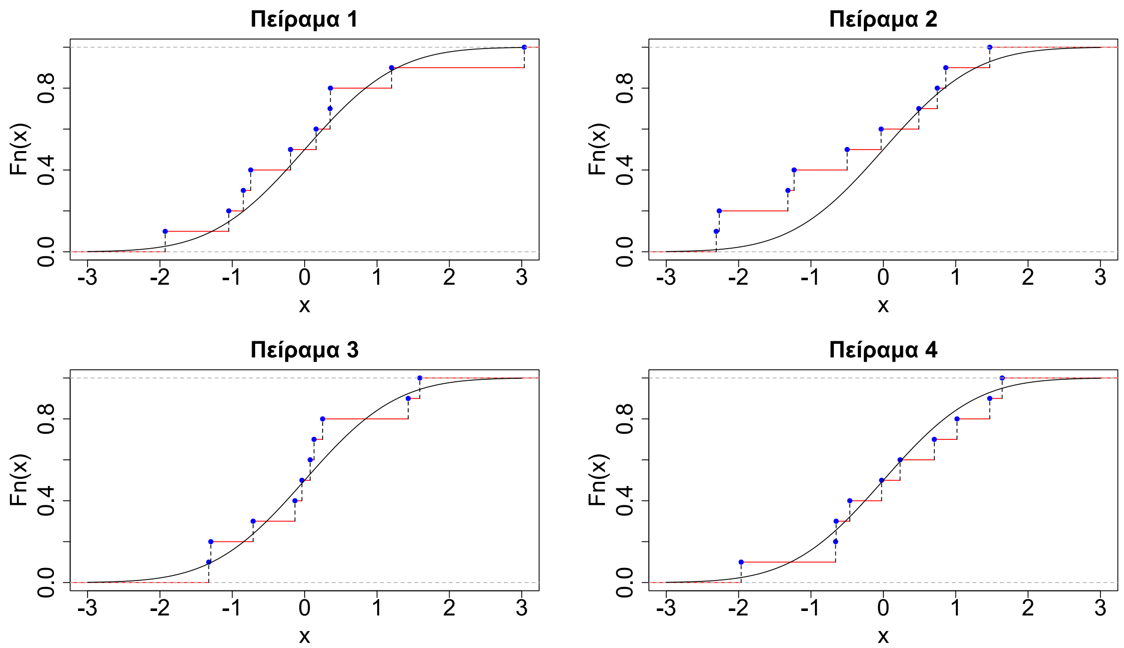
λειτουργίας του λαμπτήρα κάθε χρονική στιγμή. Εδώ, ο χώρος καταστάσεων είναι το σύνολο $\{0, 1\}$ (1: λειτουργεί, 0: δε λειτουργεί), ενώ το δεικτοσύνολο είναι το $T = [0, +\infty)$. Δύο δυνατές πραγματοποιήσεις φαίνονται στο Σχήμα 3.1.

Παράδειγμα 3.4: Η εμπειρική συνάρτηση κατανομής $\mathbb{F}_n = \big(\mathbb{F}_n(x)\big)_{x\in\mathbb{R}}$ αποτελεί μία στοχαστική διαδικασία συνεχούς χρόνου με δεικτοσύνολο $T = \mathbb{R}$ και χώρο καταστάσεων $S = [0, 1]$. Η ε.σ.κ. γράφεται και στη μορφή $\mathbb{F}_n = \big(\mathbb{F}_n(t)\big)_{t\in\mathbb{R}}$, αν ως μεταβλητή χρησιμοποιήσουμε το t. Κάθε πραγματοποίηση $\mathbb{F}_n(\omega)$ αντιστοιχεί σε μία συνάρτηση κατανομής. Στο Σχήμα 3.2 μπορεί εύκολα να παρατηρήσει κανείς ότι όταν το μέγεθος του δείγματος είναι αρκετά μικρό, η προσέγγιση της θεωρητικής από την εμπειρική σ.κ. δεν είναι καλή.



Σχημα 3.2: Μία πραγματοποίηση (προσομοίωση) της εμπειρικής σ.κ. με δείγμα μεγέθους 10 από $\mathcal{N}(0, 1)$. Η θεωρητική συνάρτηση κατανομής Φ της $\mathcal{N}(0, 1)$ αντιστοιχεί στη μαύρη καμπύλη.

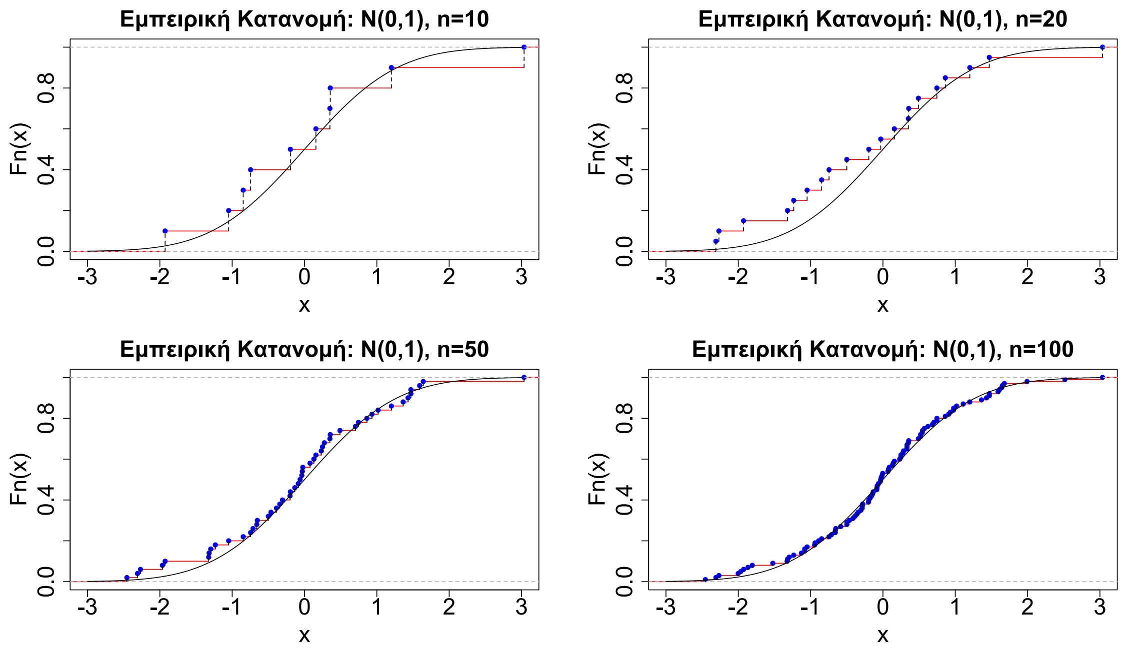
Στο Σχήμα 3.3 είναι σαφές ότι το παραπάνω ποιοτικό αποτέλεσμα δεν εξαρτάται από τη συγκεκριμένη πραγματοποίηση, αλλά είναι συνέπεια του μικρού μεγέθους δείγματος και της μορφής της εμπειρικής συνάρτησης κατανομής.



Σχημα 3.3: 4 διαφορετικές πραγματοποιήσεις της εμπειρικής σ.κ. που προκύπτουν από 4 προσομοιωμένα δείγματα μεγέθους 10 της $\mathcal{N}(0,1)$. Η θεωρητική συνάρτηση κατανομής της $\mathcal{N}(0,1)$ αντιστοιχεί στη μαύρη καμπύλη.

Σε κάθε παρατήρηση αντιστοιχεί ένα άλμα ασυνέχειας μεγέθους $1/n$. Έτσι, για μικρό n προκύπτουν αισθητές μεταβολές στο γράφημα της συνάρτησης, αλλά και μεταξύ διαφορετικών πραγματοποιήσεων είναι φανερή η μεταβλητότητα που παρουσιάζεται.

Παράδειγμα 3.5: Θέλουμε να κάνουμε προσομοίωση της εξέλιξης της εμπειρικής συνάρτησης κατανομής για $n = 10, 20, 50, 100$ όταν $X_i \sim \mathcal{N}(0,1), \quad 1 \leq i \leq n$. Μία πιθανή εξέλιξη της εμπειρικής συνάρτησης κατανομής απεικονίζεται στο Σχήμα 3.4 σε 4 διαδοχικά στιγμιότυπα για $n = 10^k$, $1 \leq k \leq 4$. Καθώς το μέγεθος του δείγματος αυξάνεται η προσέγγιση της θεωρητικής σ.κ. από την εμπειρική βελτιώνεται αισθητά. Ενώ για $n = 10$ (πάνω αριστερά) η προσέγγιση δεν είναι καλή, βελτιώνεται σημαντικά για $n = 100$ (κάτω αριστερά) και δύσκολα ξεχωρίζουμε τις καμπύλες για $n = 1000$ (πάνω δεξιά). Για $n = 10000$ (κάτω δεξιά) επιτυγχάνεται πλήρης ταύτιση, τουλάχιστον γραφικά.



Σχημα 3.4: Μία πραγματοποίηση της εξέλιξης της εμπειρικής σ.κ. σε 4 διαδοχικά γραφήματα καθώς το μέγεθος του δείγματος της $\mathcal{N}(0,1)$ αυξάνει (n=10, 20,50 και 100). Η θεωρητική συνάρτηση κατανομής της $\mathcal{N}(0,1)$ αντιστοιχεί στη μαύρη καμπύλη.

3.5 Κατανομή Στοχαστικής Διαδικασίας

Η έννοια της κατανομής μίας στοχαστικής διαδικασίας μπορεί να οριστεί ανάλογα με την περίπτωση της κατανομής μίας πραγματικής τυχαίας μεταβλητής ή ενός τυχαίου διανύσματος, όπως στους Ορισμούς 1.8 και 2.17 αντίστοιχα. Αν $(S, \mathcal{S})$ είναι ο μετρήσιμος χώρος μέσα στον οποίο οι X_t παίρνουν τιμές, τότε συνήθως η κατανομή ορίζεται πάνω στην κυλινδρική σ-άλγεβρα (σ-άλγεβρα γινόμενο) $\mathcal{S}^T$, που ορίζεται με ανάλογο τρόπο όπως στον Ορισμό 3.3, αντικαθιστώντας τη σ-άλγεβρα $\mathcal{B}(\mathbb{R})$ με την $\mathcal{S}$.

Ορισμός 3.5: Έστω $X = (X_t)_{t\in T}$ μία στοχαστική διαδικασία ορισμένη στο χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$ με τιμές στον S^T. Το μέτρο πιθανότητας $\mathbb{P}_X : \mathcal{S}^T \to [0,1]$ με $\mathbb{P}_X(B) = \mathbb{P}(X \in B) = \mathbb{P}\left(X^{-1}(B)\right)$ καλείται *κατανομή* της σ.δ. X.

Το επόμενο αποτέλεσμα απλοποιεί την εύρεση της κατανομής μίας σ.δ., καθώς περιορίζει το πρόβλημα αυτό στην εύρεση της οικογένειας όλων των πεπερασμένης διάστασης κατανομών, δηλ. αυτών που προκύπτουν από τις κατανομές όλων των τυχαίων διανυσμάτων που μπορούμε να σχηματίσουμε από τη στοχαστική διαδικασία.

Πρόταση 3.6: *Η κατανομή μίας στοχαστικής διαδικασίας $X = (X_t)_{t\in T}$ καθορίζεται πλήρως από τις κατανομές όλων των τυχαίων διανυσμάτων που μπο-*

ρούμε να σχηματίσουμε από την οικογένεια $\{X_t\}_{t\in T}$, δηλ. τις κατανομές των $X_{t_{1:m}} := (X_{t_1}, X_{t_2}, \ldots, X_{t_m})$, για κάθε $m \geq 1$ και επιλογή δεικτών $t_1, t_2, \ldots, t_m \in T$.

Σύμφωνα με το παραπάνω αποτέλεσμα μπορούμε να ορίσουμε την ισονομία στοχαστικών διαδικασιών μόνο με όρους κατανομών πεπερασμένης διάστασης.

Ορισμός 3.6: Δύο στοχαστικές διαδικασίες X και Y θα λέγονται *ισόνομες*, και θα γράφουμε $X \stackrel{d}{=} Y$, αν για κάθε $m \geq 1$ και επιλογή δεικτών $t_{1:m} \in T^m$ έχουμε $X_{t_{1:m}} \stackrel{d}{=} Y_{t_{1:m}}$.

Παρατήρηση 3.9. Ο όρος ισόνομες σ.δ. δεν είναι κοινά αποδεκτός στη βιβλιογραφία. Χρησιμοποιούνται και άλλοι, όπως *στοχαστικά ισοδύναμες*. Όμως ο τελευταίος χρησιμοποιείται και για εκείνες για τις οποίες $X_t \stackrel{a.s.}{=} Y_t$ για κάθε $t \in T$. Για να μην προκληθεί σύγχυση θα υιοθετήσουμε τον όρο ισόνομες.

Παρατήρηση 3.10. Όπως και για τα τυχαία διανύσματα, όταν αναφερόμαστε σε ισονομία στοχαστικών διαδικασιών, δεν χρειάζεται αυτές να είναι ορισμένες στον ίδιο χώρο πιθανότητας.

Είμαστε τώρα σε θέση να προσδιορίσουμε την κατανομή της εμπειρικής συνάρτησης κατανομής. Για το σκοπό αυτό θα ήταν χρήσιμο να ορίσουμε την κατανομή του τυχαίου διανύσματος που έχει ως συνιστώσες τα διαδοχικά μερικά αθροίσματα των συνιστωσών ενός τυχαίου διανύσματος που ακολουθεί πολυωνυμική κατανομή.

Ορισμός 3.7: Έστω $X \sim \mathcal{M}_m(n, p)$. Θα λέμε ότι το τυχαίο διάνυσμα $S = (S_1, \ldots, S_m)$ ακολουθεί την m-διάστατη *αθροιστική πολυωνυμική κατανομή* με παραμέτρους n και p, και θα γράφουμε $S \sim \mathcal{SM}_m(n, p)$, αν $S_i = X_1 + \cdots + X_i$ για κάθε $1 \leq i \leq m$.

Η επόμενη πρόταση γενικεύει την Πρόταση 3.1.

Πρόταση 3.7: Έστω $n, m \geq 1$ και $t_1, t_2, \ldots, t_m \in \mathbb{R}$ με $t_1 < t_2 < \cdots < t_m$. Τότε,

$$\mathbb{F}_n(t_{1:m}) = \big(\mathbb{F}_n(t_1), \mathbb{F}_n(t_2), \ldots, \mathbb{F}_n(t_m)\big) \sim \frac{1}{n}\,\mathcal{SM}_m(n, p),$$

όπου $p = (p_1, p_2, \ldots, p_m) = \big(F(t_1), F(t_2) - F(t_1), \ldots, F(t_m) - F(t_{m-1})\big)$.

Απόδειξη: Τα σημεία $t_1, t_2, \ldots, t_m$ διαμερίζουν το $\mathbb{R}$ σε $m+1$ το πλήθος διαστήματα, τα $(-\infty, t_1]$, $(t_1, t_2]$, $\ldots$, $(t_{m-1}, t_m]$ και $(t_m, +\infty)$. Θέτουμε $t_0 = -\infty$ και έστω

$$C_k = (C_{k,1}, C_{k,2}, \ldots, C_{k,m}), \text{ όπου } C_{k,i} = \mathbf{1}_{(t_{i-1}, t_i]}(X_k), \ k \geq 1 \text{ και } 1 \leq i \leq m.$$

Η X_k θα βρεθεί σε ακριβώς ένα από αυτά τα διαστήματα αν $X_k \le t_m$. Άρα

$$C_k \sim \mathcal{C}\boldsymbol{at}_m(p), \quad \text{όπου } p = (p_1, p_2, \dots, p_m),$$

και για κάθε $1 \le i \le m$,

$$p_i = \mathbb{P}(C_{k,i} = 1) = \mathbb{P}(t_{i-1} < X_k \le t_i) = F(t_i) - F(t_{i-1}).$$

Καθώς τα $\{C_k\}_{k=1}^n$ είναι ανεξάρτητα κατηγορικά τ.δ. με κοινή παράμετρο p, συμπεραίνουμε ότι

$$N = \sum_{k=1}^n C_k \sim \mathcal{M}_m(n, p),$$

όπου η ανεξαρτησία των $\{C_k\}_{k=1}^n$ έπεται από την ανεξαρτησία των $\{X_k\}_{k=1}^n$. Άρα αν

$$S = (S_1, S_2, \dots, S_m), \text{ με } S_i = N_1 + N_2 + \dots + N_i, \ 1 \le i \le m, \text{ τότε } S \sim \mathcal{SM}_m(n, p).$$

Παρατηρούμε τώρα ότι για κάθε $1 \le i \le m$,

$$S_i = \sum_{\ell=1}^i N_\ell = \sum_{\ell=1}^i \sum_{k=1}^n \mathbf{1}_{(t_{\ell-1}, t_\ell]}(X_k) = \sum_{k=1}^n \mathbf{1}_{(-\infty, t_i]}(X_k) = n\,\mathbb{F}_n(t_i),$$

από το οποίο συμπεραίνουμε ότι $\mathbb{F}_n(t_{1:m}) = S/n$ και η απόδειξη είναι πλήρης.

$\square$

Πόρισμα 3.2: Για $n, m \ge 1$ και $t_1 < t_2 < \dots < t_m$, έχουμε

•

(ii) $$\mathbb{V}\left[\mathbb{F}_n(t_{1:m})\right] = \frac{1}{n}\begin{pmatrix} F(t_1)\big(1-F(t_1)\big) & F(t_1)\big(1-F(t_2)\big) & \cdots & F(t_1)\big(1-F(t_m)\big) \\ F(t_1)\big(1-F(t_2)\big) & F(t_2)\big(1-F(t_2)\big) & \cdots & F(t_2)\big(1-F(t_m)\big) \\ \vdots & \vdots & \ddots & \vdots \\ F(t_1)\big(1-F(t_m)\big) & F(t_2)\big(1-F(t_m)\big) & \cdots & F(t_m)\big(1-F(t_m)\big) \end{pmatrix}.$$

Απόδειξη: Από την Πρόταση 3.7 συμπεραίνουμε ότι το τ.δ. $X = n\,\mathbb{F}_n(t_{1:m}) \sim \mathcal{SM}_m(n, p)$. Αρκεί λοιπόν να δειχθεί ότι $\mathbb{E}(X_i) = nF(t_i)$ για $1 \le i \le m$ και $\mathbb{C}(X_i, X_j) = nF(t_i)(1 - F(t_j))$ για $1 \le i \le j \le m$. Αν $N \sim \mathcal{M}_m(n, p)$, τότε για κάθε $1 \le i \le m$,

$$X_i = N_1 + \ldots + N_i \sim \mathcal{B}\boldsymbol{in}(n, p_1 + \ldots + p_i) = \mathcal{B}\boldsymbol{in}(n, F(t_i)).$$

Το αποτέλεσμα (i) έπεται άμεσα παίρνοντας μέσες τιμές. Για το (ii) αρκεί να παρατηρήσουμε ότι για κάθε $1 \le i \le j \le n$ το ζεύγος $(X_i, n - X_j) \sim \mathcal{M}_2\big(n, F(t_i), 1 - F(t_j)\big)$ και άρα

$$\mathbb{C}(X_i, X_j) = -\,\mathbb{C}(X_i, n - X_j) = nF(t_i)(1 - F(t_j)).$$

Εναλλακτικά, τα αποτελέσματα προκύπτουν από το γεγονός ότι $X = SN$, όπου S είναι ο πίνακας με στοιχεία $s_{ij} = 1$ για $1 \le j \le i \le n$ και 0 διαφορετικά. Άρα $\mathbb{E}(X) = S\,\mathbb{E}(N)$ και $\mathbb{V}(S) = S\,\mathbb{V}(N)S^\top$ και τα αποτελέσματα προκύπτουν μετά από πράξεις.

□

3.6 Στοχαστικές Διαδικασίες ως συλλογές τυχαίων στοιχείων*

Ορισμός 3.8: Έστω $(\Omega, \mathcal{A}, \mathbb{P})$ ένας χώρος πιθανότητας και $(S, \mathcal{S})$ ένας μετρήσιμος χώρος. Ένα *τυχαίο στοιχείο* (*random element*) του S, ή του $(S, \mathcal{S})$ (αν θέλουμε να τονίσουμε τη σ–άλγεβρα $\mathcal{S}$ που έχουμε επιλέξει), είναι μία συνάρτηση $X : \Omega \to S$, με την απαίτηση $X^{-1}(\mathcal{S}) \subset \mathcal{A}$, δηλαδή,

$$\{X \in B\} := X^{-1}(B) \in \mathcal{A}, \text{ για κάθε } B \in \mathcal{S}.$$

Αν η X ικανοποιεί την παραπάνω σχέση, τότε καλείται και $\mathcal{S}$-μετρήσιμη.

Παράδειγμα 3.6: Αν $(S, \mathcal{S}) = \big(\mathbb{R}, \mathcal{B}(\mathbb{R})\big)$ ή $\big(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\big)$, τότε ο ορισμός αυτός αντιστοιχεί στις τυχαίες μεταβλητές ή στα τυχαία διανύσματα αντίστοιχα. Έτσι κάθε πραγματική τυχαία μεταβλητή είναι ένα τυχαίο στοιχείο του $\mathbb{R}$ και κάθε τυχαίο διάνυσμα είναι ένα τυχαίο στοιχείο του $\mathbb{R}^d$.

Παράδειγμα 3.7: Μία ακολουθία πραγματικών τυχαίων μεταβλητών $(X_n)_{n\geq 0}$, ορισμένων σε ένα κοινό χώρο πιθανότητας, μπορούμε να την ερμηνεύσουμε ως ένα τυχαίο στοιχείο του $\mathbb{R}^{\mathbb{N}}$ με μία κατάλληλη επιλογή σ–άλγεβρας.

Παράδειγμα 3.8: Έστω $\Theta \subset \mathbb{R}^d$ ο παραμετρικός χώρος που αντιστοιχεί σε ένα πρόβλημα παραμετρικής Στατιστικής. Υποθέτουμε ότι η παράμετρος $\theta \in \Theta$ προσδιορίζει πλήρως μία συνάρτηση πιθανότητας ή πυκνότητας πιθανότητας $f(x;\theta)$ μιας άγνωστης κατανομής που θέλουμε να εκτιμήσουμε. Αν διαθέτουμε ένα τυχαίο δείγμα $X_1, X_2, \ldots, X_n$ από αυτήν την κατανομή, τότε η συνάρτηση πιθανοφάνειας $L_n(\theta)$ και η συνάρτηση λογαριθμοπιθανοφάνειας $\ell_n(\theta) := \log L_n(\theta)$ που αντιστοιχούν σε αυτό το τυχαίο δείγμα είναι τυχαίες συναρτήσεις του θ, άρα είναι και αυτές τυχαία στοιχεία ενός κατάλληλου συναρτησιακού χώρου. Για παράδειγμα, αν η $L_n(\theta)$ είναι συνεχής και φραγμένη στο Θ και με $C_b(\Theta)$ συμβολίσουμε όλες τις συνεχείς και φραγμένες συναρτήσεις με πεδίο ορισμού το Θ, τότε η $L_n(\theta)$ είναι ένα τυχαίο στοιχείο του $C_b(\Theta)$. Ο χώρος αυτός γίνεται μετρικός με επιλογή της $d_\infty(f,g) = \sup_\theta |f(\theta) - g(\theta)|$ και έτσι φυσιολογικά μπορούμε να επιλέξουμε ως σ–άλγεβρα τα σύνολα *Borel* αυτού του μετρικού χώρου.

Στο Παράδειγμα 3.6 είναι φανερό ότι τα σύνολα *Borel* προσδιορίζονται από τη συνήθη τοπολογία του $\mathbb{R}$ ή του $\mathbb{R}^d$ και δεν υπάρχει σύγχυση. Για να πάμε σε πιο ενδιαφέροντα παραδείγματα πρέπει να τοποθετηθούμε σε απειροδιάστατους χώρους. Στο Παράδειγμα 3.7 θα θέλαμε, και πράγματι είναι εφικτό, οι τυχαίες μεταβλητές X_n να είναι ορισμένες σε έναν κοινό χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, έτσι ώστε να μπορούμε να κατοχυρώσουμε την ακολουθία αυτή ως ένα τυχαίο στοιχείο. Χρειαζόμαστε βέβαια και το πεδίο τιμών της $X = (X_n)_{n\geq 0}$. Παρατηρούμε ότι κάθε πραγματοποίηση $X(\omega)$ της X αντιστοιχεί σε μία ακολουθία πραγματικών αριθμών. Έτσι, το ευρύτερο δυνατό υποσύνολο μέσα στο οποίο η X παίρνει τιμές είναι ο χώρος $\mathbb{R}^{\mathbb{N}}$ που αντιστοιχεί ακριβώς σε όλες τις δυνατές ακολουθίες πραγματικών αριθμών. Μπορούμε να αντιληφθούμε τον χώρο αυτό ως ένα αριθμησίμως άπειρο καρτεσιανό γινόμενο αντιγράφων του $\mathbb{R}$. Για να γίνει το X τυχαίο στοιχείο χρειάζεται επιπλέον και τον καθορισμό μιας σ–άλγεβρας $\mathcal{S}$ στον $\mathbb{R}^{\mathbb{N}}$ ή σε κάποιο υποσύνολο του $\mathbb{R}^{\mathbb{N}}$, ας το πούμε S, με τρόπο τέτοιο ώστε να ικανοποιείται η απαίτηση της μετρησιμότητας. Η πολυπλοκότητα προφανώς αυξάνει όταν εξετάζουμε οικογένειες τυχαίων μεταβλητών, όπως για παράδειγμα στοχαστικές διαδικασίες συνεχούς χρόνου, αφού οι πραγματοποιήσεις τους είναι μέσα σε υποσύνολα S του χώρου $\mathbb{R}^T$ (T κάποιο διάστημα του $\mathbb{R}$), που αντιστοιχούν σε ένα υπεραριθμήσιμο καρτεσιανό γινόμενο αντιγράφων του $\mathbb{R}$. Η επιλογή του S, αλλά και της σ–άλγεβρας $\mathcal{S}$ εξαρτώνται από τη φύση του προβλήματος και πολλές φορές δεν είναι προφανής. Στην καλύτερη περίπτωση, θέλουμε το S να έχει τη δομή μετρικού χώρου και η $\mathcal{S}$ να αντιστοιχεί φυσιολογικά στα *Borel* υποσύνολα που καθορίζονται από τη μετρική του τοπολογία.

Ορισμός 3.9: Έστω $(\Omega, \mathcal{A}, \mathbb{P})$ ένας χώρος πιθανότητας, T ένα δεικτοσύνολο και $(S, \mathcal{S})$ ένας μετρήσιμος χώρος. Μία συλλογή $\{X_t : t \in T\}$ τυχαίων στοιχείων $X_t : \Omega \to S$ καλείται ***στοχαστική διαδικασία*** (*stochastic process*) με δεικτοσύνολο T.

Παρατήρηση 3.11. Παραδοσιακά, σε μία στοχαστική διαδικασία, που στην ελληνική βιβλιογραφία αναφέρεται και ως στοχαστική ανέλιξη, το δεικτοσύνολο T είναι είτε διακριτό, είτε κάποιο διάστημα του $\mathbb{R}$ και ερμηνεύεται ως χρόνος. Έτσι μιλάμε για στοχαστικές διαδικασίες διακριτού ή συνεχούς χρόνου. Αυτές είναι και οι στοχαστικές διαδικασίες που έχουν μελετηθεί περισσότερο και μοντελοποιούν την εξέλιξη ενός στοχαστικού φαινομένου στο χρόνο.

Παρατήρηση 3.12. Αν $X = (X_t)_{t\in T}$, τότε είναι φανερό ότι μία στοχαστική διαδικασία μπορούμε να τη θεωρήσουμε ως ένα τυχαίο στοιχείο του S^T με μία κατάλληλη επιλογή σ–άλγεβρας γενικεύοντας τις έννοιες της προηγούμενης παραγράφου από το χώρο γινόμενο $\mathbb{R}^T$ στο χώρο S^T. Παρατηρούμε ότι κάθε στοχαστική διαδικασία $X = X(t, \omega)$ εξαρτάται από δύο μεταβλητές. Για t σταθερό, η $X(t, \cdot)$ ως συνάρτηση του ω είναι τυχαίο στοιχείο του S, ενώ για ω σταθερό, η $X(\cdot, \omega)$ είναι μία συνάρτηση του t. Καθώς η τελευταία αντιστοιχεί σε μία πραγματοποίηση της στοχαστικής διαδικασίας, αναφερόμαστε σε αυτήν ως δειγματοσυνάρτηση (*sample function*) ή δειγματικό μονοπάτι (*sample path*). Το σύνολο S αναφέρεται ως χώρος καταστάσεων.

Παρατήρηση 3.13. Πολλές φορές η στοχαστική διαδικασία περιορίζεται έτσι ώστε να έχει μονοπάτια σε ένα $M \subset S^T$. Αυτό γίνεται για να εξασφαλίσουμε καλύτερες ιδιότητες στις δειγματοσυναρτήσεις της. Οι ιδιότητες αυτές εκφράζονται ακριβώς με την επιλογή του M και μιας κατάλληλης σ–άλγεβρας, π.χ., αν $S = \mathbb{R}$ και $T = [0, 1]$, θα ήταν προτιμότερο η X να έχει μονοπάτια στο $M = C[0, 1]$, δηλαδή στις συνεχείς συναρτήσεις του $[0, 1]$, παρά στον άτακτο χώρο $\mathbb{R}^{[0,1]}$. Σε αυτήν την περίπτωση η X

αντιστοιχεί σε ένα τυχαίο στοιχείο του M.

Η ερμηνεία μιας στοχαστικής διαδικασίας ως ενός τυχαίου στοιχείου του S^T μας επιτρέπει να ορίσουμε την κατανομή της με ανάλογο τρόπο όπως στον Ορισμό 3.5 και να την χαρακτηρίσουμε όπως στην Πρόταση 3.6. Όπως στην περίπτωση τυχαίων διανυσμάτων, έτσι και για τυχαία στοιχεία (π.χ., στοχαστικές διαδικασίες) μπορούμε να χρησιμοποιήσουμε κάποια οριακά θεωρήματα που απλοποιούν την εύρεση ασυμπτωτικών ιδιοτήτων ακολουθιών τυχαίων στοιχείων. Σε αυτήν την κατεύθυνση, ιδιαίτερο ενδιαφέρον έχει το θεώρημα αναπαράστασης του Skorokhod που απλοποιεί σε σημαντικό βαθμό αποδείξεις σύγκλισης κατά κατανομή.

Θεώρημα 3.1 (αναπαράστασης του Skorokhod): Έστω (X_n), X τυχαία στοιχεία με τιμές σε μετρικό χώρο S, όπου υποθέτουμε ότι $X_n \xrightarrow{d} X$ και επιπλέον ότι το $\mathbb{P}_X$ έχει *διαχωρίσιμο στήριγμα*. Τότε υπάρχουν (X_n^*) και X^* τυχαία στοιχεία με τιμές στο S, ορισμένα σε κοινό χώρο πιθανότητας, τέτοια ώστε $X_n^* \overset{d}{=} X_n$, $\forall n \geq 1$, $X^* \overset{d}{=} X$ και $X_n^* \xrightarrow{a.s.} X^*$.

Το θεώρημα συνεχούς απεικόνισης γενικεύεται και αυτό για τυχαία στοιχεία.

Θεώρημα 3.2 (Θεώρημα Συνεχούς Απεικόνισης (τυχαία στοιχεία)): Έστω (X_n), X τυχαία στοιχεία με τιμές σε έναν μετρικό χώρο (S_1, d_1). Αν $g : S_1 \to S_2$, όπου (S_2, d_2) είναι ένας άλλος μετρικός χώρος, και η g είναι συνεχής σε κάποιο $C \in \mathcal{B}(S_1)$, με $\mathbb{P}(X \in C) = 1$, τότε

$$X_n \xrightarrow{a.s./p/d} X \quad \Longrightarrow \quad g(X_n) \xrightarrow{a.s./p/d} g(X). \tag{3.13}$$

Απόδειξη: Μελετάμε χωριστά καθε μία από τις τρεις συγκλίσεις.

- $g(X_n) \xrightarrow{a.s.} g(X)$.
  Το αποτέλεσμα αυτό έπεται άμεσα (ελέγξτε το).

- $g(X_n) \xrightarrow{p} g(X)$.
  Για τη σύγκλιση κατά πιθανότητα υπάρχει μία εύκολη απόδειξη, που προκύπτει χρησιμοποιώντας υπακολουθίες. Συγκεκριμένα, μπορεί να δειχθεί ότι

  $$X_n \xrightarrow{p} X \quad \Longrightarrow \quad Y_n \xrightarrow{a.s.} X,$$

  όπου (Y_n) είναι κάποια υπακολουθία της (X_n). Αυτό διευκολύνει αρκετά τα πράγματα. Μπορείτε τώρα να δείξετε ότι $g(X_n) \xrightarrow{p} g(X)$ με εις άτοπον απαγωγή;

  Μία διαφορετική απόδειξη προκύπτει ως εξής: Θέλουμε να δείξουμε ότι για κάθε $\varepsilon > 0$,

  $$\mathbb{P}\Big(d_2\big(g(X_n)\,,\, g(X)\big) > \varepsilon\Big) \xrightarrow[n\to\infty]{} 0. \tag{3.14}$$

  Θα το δείξουμε για g που είναι συνεχής παντού και μετά έπεται εύκολα για g συνεχή σε κάποιο C με $\mathbb{P}(X \in C) = 1$ (γιατί;). Έστω $\varepsilon > 0$. Για κάθε $m \geq 1$, θέτουμε

$$A_m = \Big\{x \in S_1 : \exists\, y \in S_1 \text{ με } d(y,x) < \frac{1}{m} \text{ και } d_2(g(y), g(x)) > \varepsilon\Big\}.$$

Αφού η g είναι συνεχής, έχουμε $A_m \downarrow \varnothing$, καθώς $m \to \infty$. Έχουμε λοιπόν ότι $\forall m \geq 1$,

$$\begin{aligned}\Big\{d_2\big(g(X_n)\,,\, g(X)\big) > \varepsilon\Big\} &\subset \Big\{d_1(X_n, X) \geq \frac{1}{m}\Big\} \cup \Big\{d_1(X_n, X) < \frac{1}{m}\,,\, d_2\big(g(X_n)\,,\, g(X)\big) > \varepsilon\Big\} \\ &\subset \Big\{d_1(X_n, X) \geq \frac{1}{m}\Big\} \cup \Big\{X \in A_m\Big\}.\end{aligned}$$

Συμπεραίνουμε ότι για κάθε $m \geq 1$,

$$\mathbb{P}\left(d_2\big(g(X_n)\,,\, g(X)\big) > \varepsilon\right) \leq \mathbb{P}\left(d_1(X_n, X) \geq \frac{1}{m}\right) + \mathbb{P}\left(X \in A_m\right).$$

Τότε, παίρνοντας $\limsup_n$ στην τελευταία σχέση, έχουμε ότι για κάθε $m \geq 1$

$$\limsup_n \mathbb{P}\left(d_2\big(g(X_n)\,,\, g(X)\big) > \varepsilon\right) \leq 0 + \mathbb{P}(X \in A_m).$$

Παίρνοντας όριο καθώς $m \to \infty$ στο δεξί μέλος της παραπάνω ανισότητας, έχουμε

$$\limsup_n \mathbb{P}\left(d_2\big(g(X_n)\,,\, g(X)\big) > \varepsilon\right) = 0,$$

αφού $A_m \downarrow \varnothing \Rightarrow \lim_m \mathbb{P}(X \in A_m) = \mathbb{P}(X \in \varnothing) = 0$. Τελικά, δείξαμε τη σχέση (3.14) και η απόδειξη είναι πλήρης.

- $g(X_n) \xrightarrow{d} g(X)$.
  Η απόδειξη για τη σύγκλιση κατά κατανομή μπορεί να γίνει με διάφορους τρόπους. Η απλούστερη απόδειξη προκύπτει από τον χαρακτηρισμό της σύγκλισης κατά κατανομή μέσω της σύγκλισης της ακολουθίας των μέσων τιμών κάθε συνεχούς και φραγμένης συνάρτησης (η οποία επεκτείνεται και για συνεχείς μόνο σε σύνολο που το οριακό τυχαίο στοιχείο παίρνει τιμές με πιθανότητα 1). Ας δούμε μία εφαρμογή του θεωρήματος του Skorokhod στην ειδική περίπτωση που το $\mathbb{P}_X$ έχει διαχωρίσιμο στήριγμα. Από υπόθεση, $X_n \xrightarrow{d} X$ και άρα υπάρχει ακολουθία τυχαίων στοιχείων (X_n^*) και X με $X_n^* \stackrel{d}{=} X_n$, $X^* \stackrel{d}{=} X$ και $X_n^* \xrightarrow{a.s.} X^*$. Από τα παραπάνω, $g(X_n^*) \xrightarrow{a.s.} g(X^*)$, αφού $\mathbb{P}(X^* \in C) = \mathbb{P}(X \in C) = 1$ και η g είναι συνεχής στο C. Άρα, $g(X_n^*) \xrightarrow{d} g(X^*)$. Τελικά, $g(X_n) \xrightarrow{d} g(X)$, αφού $g(X_n^*) \stackrel{d}{=} g(X_n)$ και $g(X^*) \stackrel{d}{=} g(X)$.

$\square$

3.7 Ισχυροί τρόποι σύγκλισης ακολουθίας σ.δ.

Πριν εμβαθύνουμε στο θέμα της σύγκλισης ακολουθιών στοχαστικών διαδικασιών είναι απαραίτητο να κατανοήσουμε διαφορετικούς τρόπους σύνδεσης στοχαστικών διαδικασιών.

Ορισμός 3.10: Δύο στοχαστικές διαδικασίες X και Y λέγονται

- *μη διακρινόμενες* (indistinguishable) και γράφουμε $X \overset{a.s.}{=} Y$ ή $X \overset{a.s.-p.w.}{=\!=} Y$, αν

$$\mathbb{P}\left(\left\{\omega \in \Omega : \; X(\omega) = Y(\omega)\right\}\right) = \mathbb{P}\left(\left\{\omega \in \Omega : \; X(t,\omega) = Y(t,\omega), \; \forall t \in T\right\}\right) = 1,$$

  δηλαδή αν με πιθανότητα 1 τα μονοπάτια τους ταυτίζονται.
- *τροποποιήσεις* (modifications) και γράφουμε $X \overset{\mathrm{mdf}}{=\!=} Y$ ή $X \overset{p.w.-a.s.}{=\!=} Y$, αν

$$X(t) \overset{a.s.}{=} Y(t), \quad \text{για κάθε } t \in T,$$

  δηλαδή αν οι αντίστοιχες τυχαίες μεταβλητές τους είναι ίσες με πιθανότητα 1.

Η σύνδεση δύο στοχαστικών διαδικασιών μέσω της σχέσης της τροποποίησης είναι στη γενική περίπτωση ασθενέστερη από την ισότητα με πιθανότητα 1.

Παρατήρηση 3.14. Είναι προφανές ότι

$$X \overset{a.s.}{=} Y \implies X \overset{\mathrm{mdf}}{=\!=} Y, \quad \text{καθώς} \quad \left\{X \overset{a.s.}{=} Y \implies X(t) \overset{a.s.}{=} Y(t), \quad \forall t \in T\right\}.$$

Η αντίστροφη συνεπαγωγή ισχύει στην ειδική περίπτωση που το δεικτοσύνολο T είναι αριθμήσιμο.

Πρόταση 3.8: Αν $X = (X_t)_{t\in T}$ και $Y = (Y_t)_{t\in T}$ είναι στοχαστικές διαδικασίες, τότε αν το T είναι αριθμήσιμο

$$X \overset{a.s.}{=} Y \iff X \overset{\mathrm{mdf}}{=\!=} Y,$$

ή ισοδύναμα

$$X \overset{a.s.-p.w.}{=\!=} Y \iff X \overset{p.w.-a.s.}{=\!=} Y.$$

Η αντίστροφη συνεπαγωγή δεν ισχύει όμως στη γενική περίπτωση, όπως φαίνεται στο επόμενο παράδειγμα.

Παράδειγμα 3.9: Έστω T_1 μία συνεχής τυχαία μεταβλητή και ορίζουμε $X = (X_t)_{t\in\mathbb{R}}$, όπου $X_t = \mathbf{1}_{\{t=T_1\}}$, για κάθε $t \in \mathbb{R}$. Είναι φανερό ότι κάθε πραγματοποίηση της X δίνει ένα μονοπάτι που παίρνει την τιμή 1 στο σημείο που αντιστοιχεί στην πραγματοποίηση της T_1 και είναι 0 σε κάθε άλλο σημείο. Εφόσον η T_1 είναι συνεχής, έχουμε ότι

$$\mathbb{P}(X_t = 1) = \mathbb{P}(T_1 = t) = 0, \quad \forall t \in \mathbb{R}.$$

Συμπεραίνουμε ότι $X_t \overset{a.s.}{=} 0$ (μηδενική τ.μ.) για κάθε $t \in \mathbb{R}$, και άρα η $Y = 0$ (μηδενική σ.δ.) είναι μία τροποποίηση της X. Όμως οι X και Y είναι διακρινόμενες, αφού $X(\omega) \neq 0 = Y(\omega)$ (μηδενική συνάρτηση) για κάθε $\omega \in \Omega$ και άρα $\mathbb{P}(X \neq Y) = 1$. Τελικά,

$$X \overset{\text{mdf}}{=} Y \;\not\Longrightarrow\; X \overset{a.s.}{=} Y, \quad \text{ή ισοδύναμα,} \quad X \overset{p.w.-a.s.}{=} Y \;\not\Longrightarrow\; X \overset{a.s.-p.w.}{=} Y.$$

Από το παραπάνω παράδειγμα είναι φανερό ότι μεταξύ τροποποιήσεων η συμπεριφορά των μονοπατιών τους μπορεί να είναι αρκετά διαφορετική.

Στην Πρόταση 3.2 αποδείχθηκε ότι η εμπειρική συνάρτηση κατανομής είναι σημειακά ισχυρά συνεπής. Η ε.σ.κ. είναι μία ακολουθία στοχαστικών διαδικασιών και για την καλύτερη κατανόηση της οριακής της συμπεριφοράς χρειάζεται να οριστούν κάποια είδη σύγκλισης.

Ορισμός 3.11: Έστω $X_n = \big(X_n(t)\big)_{t\in T}$ μία ακολουθία στοχαστικών διαδικασιών και $X = \big(X(t)\big)_{t\in T}$ μία ακόμα στοχαστική διαδικασία. Θα λέμε ότι

- η X_n *συγκλίνει με πιθανότητα 1 σημειακά* στη σ.δ. X και θα γράφουμε

  $$X_n \xrightarrow{a.s.-p.w.} X \quad \Longleftrightarrow \quad \mathbb{P}\left(\left\{\omega \in \Omega : X_n(\cdot,\omega) \xrightarrow[n\to\infty]{p.w.} X(\cdot,\omega)\right\}\right) = 1,$$

  ή ισοδύναμα (και παραλείποντας το πρώτο όρισμα), σχεδόν για κάθε $\omega \in \Omega$,

  $$X_n(\omega) \xrightarrow{p.w.} X(\omega) \qquad (\text{δηλαδή } X_n(t,\omega) \longrightarrow X(t,\omega), \quad \forall t \in T).$$

- η X_n *συγκλίνει σημειακά με πιθανότητα 1* στη σ.δ. X και θα γράφουμε

  $$X_n \xrightarrow{p.w.-a.s.} X \quad \Longleftrightarrow \quad \forall t \in T, \quad X_n(t) \xrightarrow[n\to\infty]{a.s.} X(t).$$

Με τη βοήθεια των παραπάνω ορισμών έχουμε την παρακάτω πρόταση.

Πρόταση 3.9: Έστω $X_n = (X_n(t))_{t\in T}$ και $X = (X(t))_{t\in T}$, με T αριθμήσιμο. Τότε,

$$X_n \xrightarrow{a.s.-p.w.} X \quad \Longleftrightarrow \quad X_n \xrightarrow{p.w.-a.s.} X.$$

Απόδειξη: Αφήνεται ως άσκηση. □

Παρατήρηση 3.15. Από το Παράδειγμα 3.9 είναι σαφές ότι

$$X_n \xrightarrow{p.w.-a.s.} X \;\not\Longrightarrow\; X_n \xrightarrow{a.s.-p.w.} X, \tag{3.15}$$

καθώς δεν ισχύει ούτε στη σχέση της ισότητας για T υπεραριθμήσιμο. Συμπεραίνουμε λοιπόν ότι η εναλλαγή επιτρέπεται μόνο για T αριθμήσιμο, όπως περιγράφεται στην παραπάνω πρόταση.

Από τις Προτάσεις 3.2 και 3.9 προκύπτει άμεσα το παρακάτω αποτέλεσμα.

Πόρισμα 3.3: Η εμπειρική συνάρτηση κατανομής $\mathbb{F}_n$ ικανοποιεί το εξής: για κάθε αριθμήσιμο σύνολο δεικτών $I \subset \mathbb{R}$ ισχύει

$$\big(\mathbb{F}_n(t)\big)_{t\in I} \xrightarrow{a.s.-p.w.} \big(F(t)\big)_{t\in I}.$$

Σύμφωνα με τα παραπάνω δε μπορούμε να εξασφαλίσουμε ακόμη ότι η εμπειρική συνάρτηση κατανομής συγκλίνει με πιθανότητα 1 σημειακά (κατά σημείο) στη θεωρητική συνάρτηση κατανομής. Χρειαζόμαστε λοιπόν ένα ισχυρότερο αποτέλεσμα. Το αποτέλεσμα αυτό συνοψίζεται στο Θεώρημα Glivenko-Cantelli που θα αποδείξουμε στην Ενότητα 3.9 και είναι μάλιστα ισχυρότερο καθώς εμπλέκει την ομοιόμορφη σύγκλιση. Για μία ακολουθία συναρτήσεων γνωρίζουμε ότι η ομοιόμορφη σύγκλισή της σε κάποια συνάρτηση είναι ισχυρότερη από την κατά σημείο σύγκλιση. Έτσι οδηγούμαστε φυσιολογικά σε μία ισχυρότερη σύγκλιση ακολουθίας στοχαστικών διαδικασιών.

Ορισμός 3.12: Έστω $X_n = \big(X_n(t)\big)_{t\in T}$ μία ακολουθία στοχαστικών διαδικασιών. Θα λέμε ότι η X_n συγκλίνει *με πιθανότητα 1 ομοιόμορφα* στη στοχαστική διαδικασία $X = \big(X(t)\big)_{t\in T}$ και θα γράφουμε

$$X_n \xrightarrow{a.s.-\|\cdot\|_\infty} X \iff \mathbb{P}\Big(\Big\{\omega\in\Omega : \big\|X_n(\cdot,\omega) - X(\cdot,\omega)\big\|_\infty \to 0\Big\}\Big) = 1,$$

ή, απλούστερα,

$$\|X_n - X\|_\infty = \sup_{t\in T} |X_n(t) - X(t)| \xrightarrow{a.s.} 0.$$

3.8 Η ε.σ.κ. ως cadlag διαδικασία

Τα μονοπάτια της ε.σ.κ. είναι συναρτήσεις κατανομής και επομένως αντιστοιχούν σε συναρτήσεις που είναι δεξιά συνεχείς και έχουν αριστερό όριο σε κάθε σημείο. Πολλές σημαντικές στοχαστικές διαδικασίες έχουν τροποποιήσεις που τα μονοπάτια τους αντιστοιχούν σε τέτοιες συναρτήσεις και για το λόγο αυτό η μελέτη τους έχει ιδιαίτερο ενδιαφέρον.

Ορισμός 3.13: Έστω $T = [a,b] \subset \overline{\mathbb{R}}$. Ορίζουμε

$$D_{[a,b]} = \Big\{ f : [a,b] \to \mathbb{R} \;\Big|\; f \text{ δεξιά συνεχής } \forall t \in [a,b) \;\&\; \exists \text{αριστερό όριο } \forall t \in (a,b] \Big\}.$$

Αν $f \in D_{[a,b]}$, τότε η f καλείται *cadlag* (continue à droite, limite à gauche). Μία στοχαστική διαδικασία που έχει cadlag μονοπάτια καλείται *διαδικασία cadlag* (cadlag process).

Παρατήρηση 3.16. (α) Ο χώρος $\left(D_{[a,b]}, +, \cdot\right)$ εφοδιασμένος με τη συνήθη πρόσθεση συναρτήσεων (κατά σημείο) και βαθμωτό πολλαπλασιασμό τον κατά σημείο πολλαπλασιασμό μιας συνάρτησης με βαθμωτά πραγματικούς αριθμούς είναι διανυσματικός χώρος. Επιπλέον, ο χώρος $\left(D_{[a,b]}, \|\cdot\|_\infty\right)$ είναι χώρος Banach (πλήρης διανυσματικός χώρος με νόρμα), αλλά μη διαχωρίσιμος. Η μη διαχωρισιμότητα του παραπάνω χώρου έδωσε ώθηση στην ανάπτυξη διαφορετικών μετρικών και αντίστοιχων τοπολογιών με καλύτερες ιδιότητες. Μία μικρή εισαγωγή και επιπλέον αναφορές μπορούν να βρεθούν στο άρθρο[1] του Jakubowski[2]. Ο A. Skorokhod (βλ. Σχήμα 3.16) κατασκεύασε 4 τοπολογίες που φέρουν τα ονόματα J_1, J_2, M_1, M_2 (με πιο γνωστή την J_1) που μπορούν να περιγραφούν από πλήρεις μετρικές, κάνοντας τον χώρο διαχωρίσιμο.

(β) Η μη διαχωρισιμότητα του χώρου $\left(D_{[a,b]}, \|\cdot\|_\infty\right)$ συνδέεται με κάποια μειονεκτήματα, περισσότερο θεωρητικής φύσης, μέσα στο πλαίσιο των στοχαστικών διαδικασιών, αλλά δεν θα επεκταθούμε παραπέρα σε αυτό το σημείο. Θα δεχτούμε ότι ο χώρος $\mathcal{F}$ των συναρτήσεων κατανομής ικανοποιεί $\mathcal{F} \subset D_{[-\infty,+\infty]}$, όπου ο τελευταίος είναι εφοδιασμένος με την ομοιόμορφη μετρική. Αν

$$\ell^\infty_{[a,b]} := \left\{ f : [a,b] \to \mathbb{R} \;\middle|\; \sup_{t\in[a,b]} |f(t)| < \infty \right\},$$

τότε, αν $f \in \ell^\infty_{[a,b]}$, ορίζουμε $\|f\|_\infty := \sup\limits_{t\in[a,b]} |f(t)|$ και ισχύει η επόμενη σχέση:

$$C_{[a,b]} \subset D_{[a,b]} \subset \ell^\infty_{[a,b]},$$

όπου $C_{[a,b]}$ αποτελούν τις συνεχείς συναρτήσεις στο $[a,b]$ (άρα και φραγμένες) και όλοι οι παραπάνω χώροι θεωρούνται εφοδιασμένοι με την ομοιόμορφη νόρμα.

3.9 Το Θεώρημα Glivenko-Cantelli

Το παρακάτω λήμμα είναι βοηθητικό για την απόδειξη του Θεωρήματος Glivenko-Cantelli, αλλά έχει και ανεξάρτητο ενδιαφέρον.

Λήμμα 3.1: Αν για μία θετική τ.μ. X ισχύει ότι $X \overset{a.s.}{\leq} \varepsilon$ για κάθε $\varepsilon > 0$, τότε $X \overset{a.s.}{=} 0$.

[1] The Skorokhod topologies in functional convergence - A short introduction

[2] Adam Jakubowski, Καθηγητής στο Τμήμα Μαθηματικών και Πληροφορικής του Πανεπιστημίου Nicolaus Copernicus, στο Torun της Πολωνίας.

Απόδειξη: Αρκεί να δείξουμε ότι $\mathbb{P}(X = 0) = 1$ ή ισοδύναμα ότι $F(0) = 1$, αφού $\mathbb{P}(X < 0) = 0$ από υπόθεση (X θετική τ.μ.) και έτσι $F(0) = \mathbb{P}(X = 0)$. Από την υπόθεση έχουμε επίσης ότι για κάθε $\varepsilon > 0$, $F(\varepsilon) = \mathbb{P}(X \le \varepsilon) = 1$. Όμως, η σ.κ. F είναι δεξιά συνεχής (πάντα) και άρα $F(0) = F(0+) = \lim_{\varepsilon \to 0^+} F(\varepsilon) = 1$.

□

Μία εφαρμογή του παραπάνω λήμματος δίνεται στο παρακάτω αποτέλεσμα.

Πόρισμα 3.4: $X_n \xrightarrow{a.s.} X \iff \forall \varepsilon > 0,\ \limsup |X_n - X| \overset{a.s.}{\le} \varepsilon.$

Απόδειξη: Αν $X_n \xrightarrow{a.s.} X$, τότε $\lim |X_n - X| \overset{a.s.}{=} 0$ και άρα $\limsup |X_n - X| \overset{a.s.}{=} 0 < \varepsilon$ ($\forall \varepsilon > 0$).
Αντίστροφα, από την υπόθεση και το Λήμμα 3.1 έχουμε $\limsup |X_n - X| \overset{a.s.}{=} 0$. Όμως

$$0 \le \liminf |X_n - X| \le \limsup |X_n - X|,$$

και άρα $\liminf |X_n - X| \overset{a.s.}{=} \limsup |X_n - X| \overset{a.s.}{=} 0$. Συμπεραίνουμε λοιπόν ότι $\lim |X_n - X| \overset{a.s.}{=} 0$ και άρα $X_n \xrightarrow{a.s.} X$.

□

Είμαστε τώρα σε θέση να αποδείξουμε ένα σχετικά απλό, αλλά σημαντικό αποτέλεσμα, το οποίο αποδείχθηκε από τους Glivenko και Cantelli το 1933. Μας εξασφαλίζει ότι η ε.σ.κ. είναι ομοιόμορφα ισχυρά συνεπής εκτιμήτρια της θεωρητικής σ.κ.. Πρακτικά αυτό σημαίνει ότι καθώς το μέγεθος του τυχαίου δείγματος αυξάνει, η μέγιστη κάθετη απόκλιση μεταξύ των γραφημάτων της εμπειρικής και της θεωρητικής σ.κ. θα ελαττώνεται απεριόριστα. Έτσι αυτό εξασφαλίζει ότι, τυπικά, από κάποιο πλήθος παρατηρήσεων και πέρα, το γράφημα της ε.σ.κ. δεν θα διακρίνεται πλέον από το γράφημα της θεωρητικής σ.κ.. Από στατιστικής πλευράς, εφόσον η θεωρητική συνάρτηση σ.κ. είναι γενικά άγνωστη, έχουμε λοιπόν τη δυνατότητα να χρησιμοποιήσουμε την ε.σ.κ. ως μία λογική εκτιμήτρια της.

Θεώρημα 3.3 (Glivenko-Cantelli): Η εμπειρική συνάρτηση κατανομής $\mathbb{F}_n$ συγκλίνει με πιθανότητα 1 ομοιόμορφα στην F, δηλαδή,

$$\|\mathbb{F}_n - F\|_\infty \xrightarrow{a.s.} 0.$$

Απόδειξη: Θα δείξουμε ότι $\limsup_n \|\mathbb{F}_n - F\|_\infty \overset{a.s.}{\le} \varepsilon$ για κάθε $\varepsilon > 0$ και άρα το ζητούμενο προκύπτει από το Πόρισμα 3.4. Έστω λοιπόν $\varepsilon > 0$. Επειδή η F είναι αύξουσα και παίρνει τιμές στο $[0, 1]$, είναι φανερό ότι θα έχει το πολύ ένα πεπερασμένο πλήθος σημείων ασυνέχειας με άλμα ασυνέχειας $\ge \varepsilon$. Συμπεραίνουμε ότι υπάρχει ένα πεπερασμένο d και σημεία $-\infty = t_0 < t_1 < \ldots < t_d < t_{d+1} = +\infty$ (εδώ περιλαμβάνονται τα παραπάνω σημεία ασυνέχειας), έτσι ώστε $F(t_{i+1}^-) - F(t_i) < \varepsilon$ για κάθε $0 \le i \le d$. Παρατηρούμε τώρα ότι αν $x \in [t_i, t_{i+1})$,

$$\begin{aligned}
\mathbb{F}_n(x) - F(x) &\le \mathbb{F}_n(t_{i+1}^-) - F(t_i) \le \mathbb{F}_n(t_{i+1}^-) - F(t_{i+1}^-) + \varepsilon, \\
F(x) - \mathbb{F}_n(x) &\le F(t_{i+1}^-) - \mathbb{F}_n(t_i) \le F(t_i) - \mathbb{F}_n(t_i) + \varepsilon.
\end{aligned}$$

Επειδή τα φράγματα στις παραπάνω σχέσεις είναι ανεξάρτητα του x, για $x \in [t_i, t_{i+1})$, έχουμε

$$\sup_{x\in[t_i,t_{i+1})} |\mathbb{F}_n(x) - F(x)| \leq \max\{\mathbb{F}_n(t_{i+1}^-) - F(t_{i+1}^-), F(t_i) - \mathbb{F}_n(t_i)\} + \varepsilon.$$

Άρα

$$\|\mathbb{F}_n - F\|_\infty = \max_{0\leq i\leq d} \sup_{x\in[t_i,t_{i+1})} |\mathbb{F}_n(x) - F(x)| \leq \max_{0\leq i\leq d} \max\{\mathbb{F}_n(t_{i+1}^-) - F(t_{i+1}^-), F(t_i) - \mathbb{F}_n(t_i)\} + \varepsilon. \tag{3.16}$$

Από τον Ισχυρό Νόμο των Μεγάλων Αριθμών, έχουμε

$$\frac{\sum_{k=1}^n \mathbf{1}_{X_k\in A}}{n} \xrightarrow[n\to\infty]{a.s.} \mathbb{E}(\mathbf{1}_{X\in A}) = \mathbb{P}(X \in A). \tag{3.17}$$

Επιλέγοντας $A = (-\infty, t_i)$ και $(-\infty, t_i]$ έχουμε ότι $\mathbb{F}_n(t_i^-) \stackrel{a.s}{\to} F(t_i^-)$ και $\mathbb{F}_n(t_i) \stackrel{a.s}{\to} F(t_i)$. Από τα παραπάνω συμπεραίνουμε ότι όλες οι ακολουθίες τ.μ. που υπεισέρχονται στο max της σχέσης (3.16) συγκλίνουν με πιθανότητα 1 στο 0 και άρα $\limsup_n \|\mathbb{F}_n - F\|_\infty \stackrel{a.s.}{\leq} \varepsilon$.

□

3.10 Ρυθμός σύγκλισης στο Θεώρημα Glivenko-Cantelli

Το Θεώρημα Glivenko-Cantelli δεν μας δίνει καμία πληροφορία για το ρυθμό σύγκλισης της ε.σ.κ. στη θεωρητική. Για να μπορέσουμε να αποφανθούμε πόσο δείγμα θα πρέπει να έχουμε στη διάθεσή μας, έτσι ώστε με μεγάλη πιθανότητα να προσεγγίσουμε με δεδομένη ακρίβεια τη θεωρητική σ.κ. χρειαζόμαστε ισχυρότερα αποτελέσματα. Μία απάντηση σε αυτό το ερώτημα δίνει η ανισότητα Dvoretzky-Kiefer-Wolfowitz (DKW).

Θεώρημα 3.4 (ανισότητα DKW: Dvoretzky-Kiefer-Wolfowitz):
Για κάθε $\varepsilon > 0$, $n \geq 1$ και F σ.κ.

$$\mathbb{P}\left(\|\mathbb{F}_n - F\|_\infty > \varepsilon\right) \leq 2\, e^{-2\,n\,\varepsilon^2}. \tag{3.18}$$

Η παραπάνω ανισότητα έχει μία σχετικά μακρά ιστορία. Ξεκίνησε το 1956 με την απόδειξη της ανισότητας για ακαθόριστη πολλαπλασιαστική σταθερά C. Μετά από κάποιες προσπάθειες, η βέλτιστη σταθερά $C = 2$ προσδιορίστηκε από τον Pascal Massart. Περισσότερες πληροφορίες μπορεί να βρει ο αναγνώστης στο άρθρο “*The Tight Constant in the Dvoretzky-Kiefer-Wolfowitz Inequality*”, του Pascal Massart, που δημοσιεύτηκε στο επιστημονικό περιοδικό *The Annals of Probability*, το 1990. Η περίληψη φαίνεται στην ακόλουθη εικόνα.

The Annals of Probability
1990, Vol. 18, No. 3, 1269–1283

THE TIGHT CONSTANT IN THE DVORETZKY–KIEFER–WOLFOWITZ INEQUALITY

BY P. MASSART

Université de Paris-Sud

Let $\hat{F}_n$ denote the empirical distribution function for a sample of n i.i.d. random variables with distribution function F. In 1956 Dvoretzky, Kiefer and Wolfowitz proved that

$$P\left(\sqrt{n}\sup_x\left(\hat{F}_n(x) - F(x)\right) > \lambda\right) \le C\exp(-2\lambda^2),$$

where C is some unspecified constant. We show that C can be taken as 1 (as conjectured by Birnbaum and McCarty in 1958), provided that $\exp(-2\lambda^2) \le \frac{1}{2}$. In particular, the two-sided inequality

$$P\left(\sqrt{n}\sup_x|\hat{F}_n(x) - F(x)| > \lambda\right) \le 2\exp(-2\lambda^2)$$

holds without any restriction on λ. In the one-sided as well as in the two-sided case, the constants cannot be further improved.

Σχημα 3.5: Η περίληψη από το σημαντικό άρθρο του P. Massart

Η ανισότητα DKW μας δίνει το αποτέλεσμα του Θεωρήματος Glivenko-Cantelli, αλλά επιπλέον εξασφαλίζει και τον ρυθμό σύγκλισης. Για να αποδείξουμε ότι η ανισότητα DKW είναι ισχυρότερη, θα χρειαστούμε το 1ο Λήμμα Borel-Cantelli.

Λήμμα 3.2 (1ο Λήμμα Borel-Cantelli):
Αν (A_n) είναι μία ακολουθία ενδεχόμενων, τότε

$$\sum_{n=1}^{\infty}\mathbb{P}(A_n) < +\infty \quad \Longrightarrow \quad \mathbb{P}(\limsup A_n) = 0.$$

Έτσι μπορούμε να αποδείξουμε με απλό τρόπο ότι η ε.σ.κ. συγκλίνει πλήρως ομοιόμορφα στη σ.κ. F, άρα και με πιθανότητα 1 ομοιόμορφα.

Ορισμός 3.14 (πλήρης σύγκλιση): Μία ακολουθία τυχαίων μεταβλητών (X_n) λέμε ότι *συγκλίνει πλήρως* σε μία τ.μ. X και γράφουμε

$$X_n \xrightarrow[n\to\infty]{c} X \quad \Longleftrightarrow \quad \text{για κάθε } \varepsilon > 0, \ \sum_{n=1}^{\infty}\mathbb{P}\left(|X_n - X| > \varepsilon\right) < +\infty.$$

Είναι προφανές ότι η πλήρης σύγκλιση είναι ισχυρότερη από τη στοχαστική. Όμως, είναι ισχυρότερη και από τη σύγκλιση με πιθανότητα 1, δίνοντας σε πολλές περιπτώσεις μία εύκολα επαληθεύσιμη ικανή συνθήκη της σύγκλισης με πιθανότητα

1.

Πρόταση 3.10: Αν (X_n) είναι μία ακολουθία τ.μ. και X μία ακόμη, ορισμένες σε κοινό χώρο πιθάνοτητας, τότε

$$X_n \xrightarrow{c} X \quad \Longrightarrow \quad X_n \xrightarrow{a.s.} X.$$

Απόδειξη: Έστω λοιπόν ότι η (X_n) συγκλίνει πλήρως στη X. Σταθεροποιούμε $\varepsilon > 0$ και θέτουμε $A_n = \{|X_n - X| > \varepsilon\}$, $n \geq 1$. Από την υπόθεση και το 1ο Λήμμα Borel-Cantelli, συμπεραίνουμε ότι $\mathbb{P}(\limsup A_n) = 0$. Όμως

$$\{\limsup |X_n - X| > \varepsilon\} \subset \limsup \{|X_n - X| > \varepsilon\} = \limsup A_n.$$

Παίρνοντας πιθανότητες στην παραπάνω σχέση συμπεραίνουμε ότι $\limsup |X_n - X| \overset{a.s.}{\leq} \varepsilon$, για κάθε $\varepsilon > 0$ και άρα από το Λήμμα 3.1 έχουμε $\limsup |X_n - X| \overset{a.s.}{=} 0$.

□

Πόρισμα 3.5: Αν δεχθούμε την ισχύ της ανισότητας DKW, τότε για κάθε $F \in \mathcal{F}$ ισχύει

$$\|\mathbb{F}_n - F\|_\infty \xrightarrow[n\to\infty]{a.s.} 0. \tag{3.19}$$

Απόδειξη: Με τη βοήθεια της ανισότητας DKW, παρατηρούμε ότι για κάθε $\varepsilon > 0$ ισχύει

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\|\mathbb{F}_n - F\|_\infty > \varepsilon\right) \leq 2 \sum_{n=1}^{\infty} \Big(\underbrace{e^{-2\varepsilon^2}}_{<1}\Big)^n < +\infty.$$

Συμπεραίνουμε ότι η $\|\mathbb{F}_n - F\|_\infty$ συγκλίνει πλήρως στο 0, και άρα με πιθανότητα 1.

□

Ορισμός 3.15: Αν υποθέσουμε ότι η F είναι γνωστή, τότε η στατιστική συνάρτηση

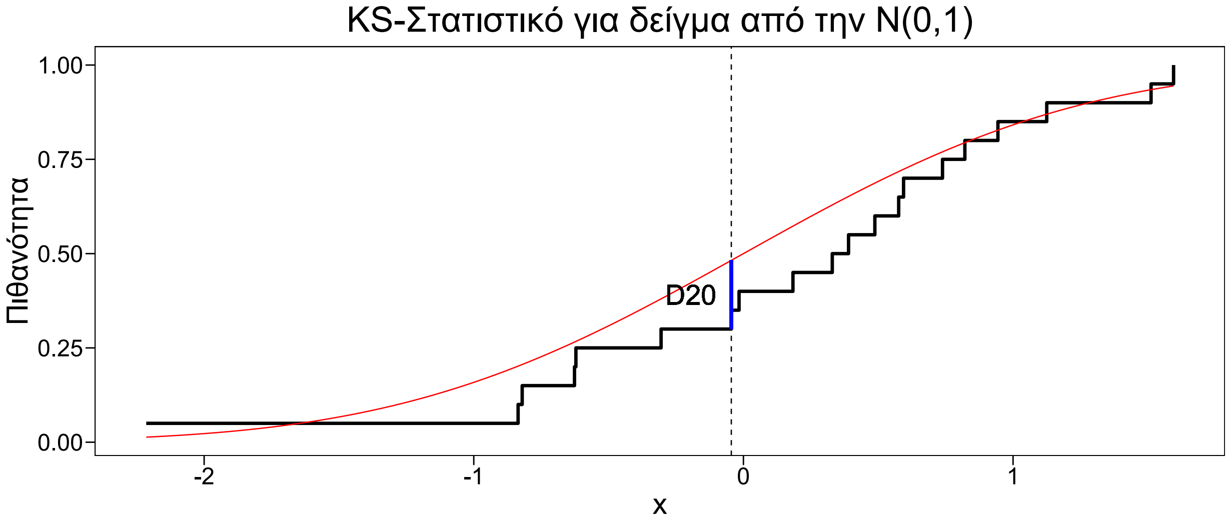
$$D_n = \|\mathbb{F}_n - F\|_\infty \tag{3.20}$$

καλείται *στατιστικό Kolmogorov-Smirnov* ή απλούστερα *KS-στατιστικό*. Εκφράζει την απόσταση της εμπειρικής συνάρτησης κατανομής από τη συνάρτηση κατανομής F. Γενικότερα, η D_n καλείται και *απόσταση Kolmogorov-Smirnov*, καθώς εκφράζει την απόσταση της $\mathbb{F}_n$ από την F με την ομοιόμορφη μετρική (βλ. Σχήμα 3.6).

Παρατήρηση 3.17. Αν στη σχέση (3.18) θέσουμε $\varepsilon = \dfrac{x}{\sqrt{n}}$, τότε έχουμε

$$\mathbb{P}(\sqrt{n} D_n > x) \leq 2e^{-2x^2}, \quad \forall x > 0.$$

Η παραπάνω ανισότητα μας δίνει ένα άνω φράγμα για την ουρά της κατανομής της



Σχημα 3.6: Γραφική αναπαράσταση της απόστασης KS μεταξύ της εμπειρικής σ.κ. (μαύρη καμπύλη) και της σ.κ. της τυπικής κανονικής (κόκκινη καμπύλη). Αντιστοιχεί στο μήκος του μπλε κάθετου ευθύγραμμου τμήματος, που καθορίζει τη μέγιστη απόκλιση μεταξύ των δύο γραφημάτων. Το γράφημα αυτό αντιστοιχεί σε ένα προσομοιωμένο δείγμα μεγέθους 20 από την τυπική κανονική.

τυχαίας μεταβλητής $\sqrt{n}D_n$, για κάθε $n \geq 1$.

3.11 Λωρίδες Εμπιστοσύνης

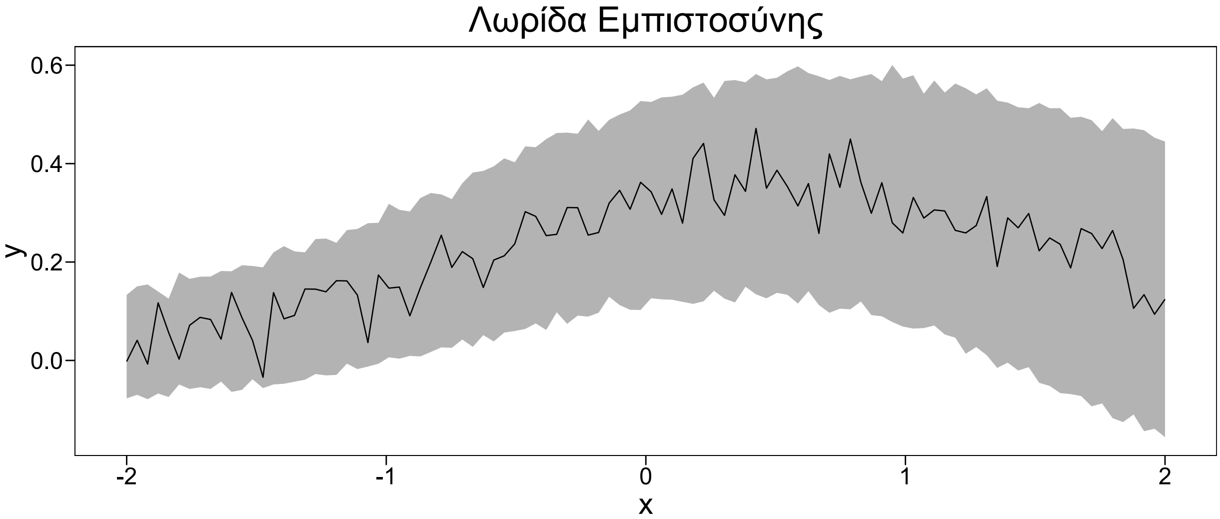
Ας δούμε τώρα πώς μπορούμε να κατασκευάσουμε λωρίδες εμπιστοσύνης για την άγνωστη συνάρτηση κατανομής F. Οι λωρίδες ή ζώνες εμπιστοσύνης στη μη παραμετρική Στατιστική παίζουν το ρόλο που έπαιζαν τα διαστήματα ή οι περιοχές εμπιστοσύνης στην παραμετρική Στατιστική. Με τον προκαθορισμό ενός συντελεστή εμπιστοσύνης, μας δίνουν μία ακρίβεια εκτίμησης, ώστε να μπορούμε μαζί με την εκτιμώμενη συνάρτηση να καθορίζουμε μία ζώνη μέσα στην οποία θα βρίσκεται η πραγματική σ.κ., όποια και αν είναι αυτή, με προκαθορισμένη πιθανότητα. Για το σκοπό αυτό θα χρειαστεί να ορίσουμε και κατάλληλα “διαστήματα” συναρτήσεων, γνωστές με το όνομα λωρίδες.

Ορισμός 3.16: Έστω $f, g\colon \mathbb{R} \to \mathbb{R}$ με $f \leq g$. Καλούμε *λωρίδα συναρτήσεων* μεταξύ της f και της g, το διάστημα

$$[f, g] := \left\{ h\colon \mathbb{R} \to \mathbb{R} \;\middle|\; f(x) \leq h(x) \leq g(x), \text{ για κάθε } x \in \mathbb{R} \right\}.$$

Στο παρακάτω σχήμα αναπαρίσταται μία τέτοια λωρίδα.

Ορισμός 3.17: Έστω F μία άγνωστη σ.κ. και L_n, U_n δύο τυχαίες συναρτήσεις του x, που έχουν προκύψει από ένα τυχαίο δείγμα μεγέθους n, με $L_n \leq U_n$. Η τυχαία λωρίδα $[L_n, U_n]$ καλείται



Σχημα 3.7: Λωρίδα Εμπιστοσύνης

- *σημειακή λωρίδα εμπιστοσύνης* με συντελεστή εμπιστοσύνης $1-a$ $(0 < a < 1)$, ή απλούστερα $(1-a)$–σ.λ.ε., αν ικανοποιεί τη σχέση:

$$\inf_{F\in\mathcal{F}} \mathbb{P}_F\big(F(x) \in [L_n(x), U_n(x)]\big) \geq 1-a, \quad \text{για κάθε } x \in \mathbb{R},$$

- *ομοιόμορφη (ή ταυτόχρονη) λωρίδα εμπιστοσύνης* με συντελεστή εμπιστοσύνης $1-a$ $(0 < a < 1)$, ή απλούστερα $(1-a)$–ο.λ.ε., αν ικανοποιεί τη σχέση:

$$\inf_{F\in\mathcal{F}} \mathbb{P}_F\big(F \in [L_n, U_n]\big) \geq 1-a.$$

Στις περιπτώσεις που το παραπάνω infimum είναι ακριβώς ίσο με $1-a$, τότε οι λωρίδες καλούνται *ακριβείς*. Αν όμως είναι $> 1-a$, τότε καλούνται *συντηρητικές*.

Παρατήρηση 3.18. (α) Σε μία ομοιόμορφη λωρίδα εμπιστοσύνης το ενδεχόμενο $\big\{F \in [L_n, U_n]\big\}$ ικανοποιείται αν η συνάρτηση F βρεθεί εξ'ολοκλήρου μέσα στη λωρίδα (ή ζώνη) που καθορίζεται με κάτω όριο την L_n και άνω όριο τη U_n. Σημειώνουμε δε, πως πολλές φορές, όπως και στην παραμετρική Στατιστική όταν γράφουμε $\big\{\theta \in I_n\big\}$, με I_n τυχαίο διάστημα, η έκφραση $\big\{F \in [L_n, U_n]\big\}$ δημιουργεί σύγχυση ως προς το τί είναι τυχαίο. Η έκφραση $\big\{[L_n, U_n] \ni F\big\}$, που αντιπροσωπεύει το ίδιο ενδεχόμενο και αντανακλά περισσότερο το γεγονός ότι η τυχαία λωρίδα καλύπτει ή περιλαμβάνει εξ'ολοκλήρου την F, θα μπορούσε ίσως να δημιουργεί μικρότερη σύγχυση, αλλά δε συνηθίζεται.

(β) Γενικά οι ομοιόμορφες λωρίδες εμπιστοσύνης είναι πλατύτερες από τις αντίστοιχες σημειακές για τον ίδιο συντελεστή εμπιστοσύνης. Αυτό οφείλεται στο γεγονός ότι με τις πρώτες θέλουμε να εξασφαλίσουμε να βρίσκεται όλη η συνάρτηση εντός της λωρίδας και, έτσι, η απαίτηση αυτή "αναγκάζει" τη λωρίδα να μεγαλώσει ώστε να εξασφαλίσει τον απαιτούμενο συντελεστή εμπιστοσύνης.

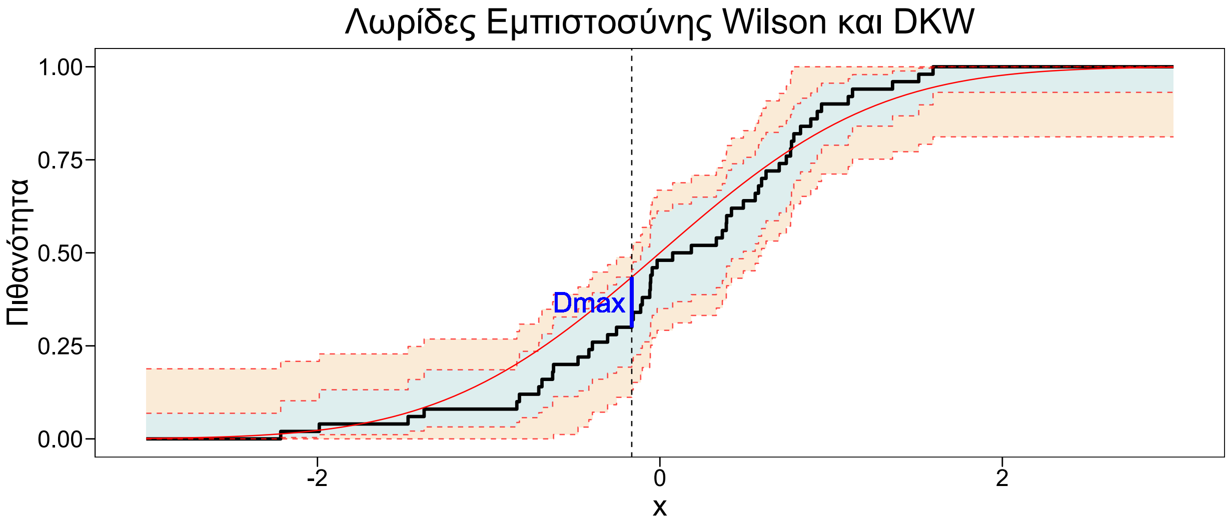
Πράγματι, παρατηρούμε ότι

$$\{F \in [L_n, U_n]\} = \bigcap_{x\in\mathbb{R}} \{F(x) \in [L_n(x), U_n(x)]\} \subset \Big\{F(x) \in [L_n(x), U_n(x)]\Big\}, \quad \text{για κάθε } x \in \mathbb{R}.$$

Αν λοιπόν η $[L_n, U_n]$ είναι μία ακριβής $(1-a)$–ο.λ.ε., τότε παίρνοντας infimum ως προς $F \in \mathcal{F}$ στις πιθανότητες των παραπάνω ενδεχομένων, συμπεραίνουμε ότι

$$1-a \;\leq \inf_{F\in\mathcal{F}} \mathbb{P}_F \Big(F(x) \in [L_n(x), U_n(x)]\Big), \quad \text{για κάθε } x \in \mathbb{R}.$$

Η παραπάνω σχέση δείχνει ότι η τυχαία λωρίδα $[L_n, U_n]$ είναι και $(1-a)$–σ.λ.ε.. Προκειμένου λοιπόν να κατασκευάσουμε μία ακριβή $(1-a)$–σ.λ.ε., αν αυτό είναι εφικτό, θα έπρεπε να θεωρήσουμε διαστήματα που είναι μικρότερα των $[L_n(x), U_n(x)]$. Έτσι, η αντίστοιχη σημειακή λωρίδα θα είναι στενότερη της ομοιόμορφης. Στο παρακάτω σχήμα δίνεται μία γραφική αναπαράσταση μίας σημειακής και μίας ομοιόμορφης 0.95-λ.ε. για τη θεωρητική συνάρτηση κατανομής.



Σχημα 3.8: Σύγκριση της σημειακής (γαλάζια λωρίδα καθορισμένη με 0.95–δ.ε. Wilson) με την ομοιόμορφη 0.95 λωρίδα εμπιστοσύνης όπου η τελευταία προεκτείνεται με κόκκινο μέχρι τα όρια που καθορίζουν η κατώτερη και η ανώτερη διακεκομμένη γραμμή.

Στην Παραμετρική Στατιστική, πολλές φορές δεν είναι εύκολο ή εφικτό να εξασφαλίσουμε ακριβή $(1-a)$–δ.ε. ή περιοχές εμπιστοσύνης. Το πρόβλημα αυτό εντείνεται ακόμα περισσότερο στη Μη Παραμετρική Στατιστική, όταν θέλουμε να κατασκευάσουμε ακριβείς $(1-a)$–λ.ε.. Για ευκολία λοιπόν, πολλές φορές καταφεύγουμε σε προσεγγιστικές μεθόδους. Αυτό είναι εφικτό είτε με μη ασυμπτωτικές προσεγγίσεις, που επιτυγχάνουμε με κατάλληλες ανισότητες συγκέντρωσης (όπως η DKW), είτε με ασυμπτωτικές προσεγγίσεις, όπως θα δείξουμε στο επόμενο κεφάλαιο με τη βοήθεια της ασυμπτωτικής κατανομής του Kolmogorov.

Στη συνέχεια αυτής της ενότητας θα κινηθούμε προσεγγιστικά και θα κατασκευάσουμε λωρίδα εμπιστοσύνης για τη συνάρτηση F με τη βοήθεια της ανισότητας

DKW.

Πρόταση 3.11: Έστω συνάρτηση κατανομής F και $X_1, \ldots, X_n$ ένα τυχαίο δείγμα από την F. Μία ομοιόμορφη λωρίδα εμπιστοσύνης $[L_n, U_n]$ για την F με συντελεστή εμπιστοσύνης $1-a$ δίνεται από την $[L_n, U_n] = \mathbb{F}_n \pm \sqrt{\frac{1}{2n}\log\frac{2}{a}}$ περιορίζοντάς τη στο $[0,1]$, δηλαδή

$$[L_n, U_n] = \left[\mathbb{F}_n - \sqrt{\frac{1}{2n}\log\frac{2}{a}}\ ,\ \mathbb{F}_n + \sqrt{\frac{1}{2n}\log\frac{2}{a}}\right] \cap [0,1].$$

Απόδειξη: Αν εξισώσουμε το άνω φράγμα της ανισότητας DKW με a, τότε η λύση ως προς ε δίνεται από το $\varepsilon_n = \sqrt{\log(2/a)}\Big/\sqrt{2n}$. Τότε, ισχύει ότι για κάθε $n \geq 1$

$$\mathbb{P}_F\left(\|\mathbb{F}_n - F\|_\infty > \varepsilon_n\right) \leq a, \quad \forall F \in \mathcal{F}.$$

Άρα, για κάθε $n \geq 1$

$$\mathbb{P}\left(\|\mathbb{F}_n - F\|_\infty \leq \varepsilon_n\right) = \mathbb{P}\left(F \in [\mathbb{F}_n - \varepsilon_n\, ,\, \mathbb{F}_n + \varepsilon_n]\right) \geq 1-a, \quad \forall F \in \mathcal{F}.$$

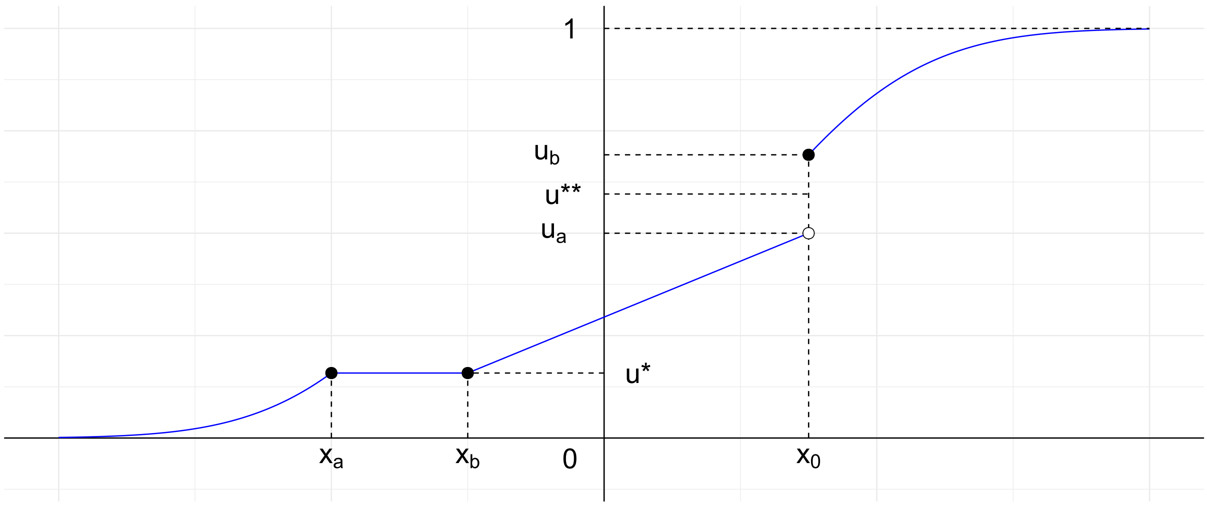
Παίρνοντας το infimum ως προς $F \in \mathcal{F}$ στην παραπάνω σχέση προκύπτει άμεσα το ζητούμενο, καθώς ο περιορισμός στο $[0,1]$ αφήνει αμετάβλητο το αποτέλεσμα (οι τιμές των συναρτήσεων είναι πιθανότητες).

□

Θα επανέλθουμε στο θέμα κατασκευής λωρίδων εμπιστοσύνης, αφού εξετάσουμε πιο διεξοδικά την κατανομή του στατιστικού Kolmogorov-Smirnov, που είναι απαραίτητη και για τη διεξαγωγή ελέγχων υποθέσεων, όταν θέλουμε να ελέγξουμε αν μία σ.κ. F_0 βρίσκεται πίσω από την παραγωγή κάποιων δεδομένων. Στην κατεύθυνση αυτή θα φανούν χρήσιμα και τα αποτελέσματα της επόμενης ενότητας που αφορούν τη γενικευμένη αντίστροφη μίας συνάρτησης κατανομής. Τα αποτελέσματα αυτά έχουν και ανεξάρτητο ενδιαφέρον, καθώς μπορούν να χρησιμοποιηθούν για την προσομοίωση τυχαίων μεταβλητών.

3.12 Γενικευμένη Αντίστροφη Συνάρτησης Κατανομής

Η έννοια της γενικευμένης αντίστροφης παίζει σημαντικό ρόλο στη θεωρία Πιθανοτήτων και στη Στατιστική. Μας επιτρέπει να αντιστρέψουμε μία συνάρτηση κατανομής F, παρόλο που σε πολλές περιπτώσεις δεν υπάρχει η συνήθης αντίστροφη. Τα προβλήματα που εμφανίζονται στην ύπαρξη μίας αντίστροφης $F^{-1} : (0,1) \to \mathbb{R}$ παρουσιάζονται στο Σχήμα 3.9.



Σχημα 3.9: "Προβληματικές" περιπτώσεις για την ύπαρξη αντίστροφης $F^{-1} : (0,1) \to \mathbb{R}$ μίας συνάρτησης κατανομής F. Η εξίσωση $F(x) = u$ (i) έχει άπειρες λύσεις στο $u = u^*$ (περιοχή σταθεροποίησης της F) και (ii) δεν έχει καμία λύση για $u \in (u_a, u_b)$ (άλμα ασυνέχειας της F)

Τα προβλήματα αυτά αντιμετωπίζονται με την κατασκευή μίας γενικευμένης αντίστροφης που διατηρεί επιθυμητές ιδιότητες της συνήθους αντίστροφης, όπως θα δούμε παρακάτω.

Ορισμός 3.18 (γενικευμένη αντίστροφη συνάρτησης κατανομής): Έστω $F : \mathbb{R} \to [0,1]$ μία συνάρτηση κατανομής. Η συνάρτηση $F^{-} : (0,1) \to \mathbb{R}$, όπου

$$F^{-}(u) = \inf\{x : F(x) \geq u\}$$

καλείται *γενικευμένη αντίστροφη* της F.

Παρατήρηση 3.19. Η συνάρτηση F^{-} είναι καλά ορισμένη. Πράγματι, είναι απλό να δειχθεί ότι $F^{-}(u) \in \mathbb{R}$ για κάθε $u \in (0,1)$. Για το σκοπό αυτό ορίζουμε το σύνολο

$$I_u := \{x : F(x) \geq u\} \quad \text{και άρα} \quad I_u^{\mathsf{c}} = \{x : F(x) < u\}. \tag{3.21}$$

Θα δείξουμε ότι για το τυχόν $u \in (0,1)$ το σύνολο I_u είναι (i) μη κενό και (ii) κάτω φραγμένο. Σε αυτήν την περίπτωση, από την πληρότητα του $\mathbb{R}$, ο πραγματικός αριθμός $F^{-}(u)$ ορίζεται καλά ως το μέγιστο κάτω φράγμα του συνόλου I_u. Το (i) έπεται άμεσα από την ιδιότητα $F(+\infty) = 1$, καθώς σε αντίθετη περίπτωση η σ.κ. F θα ήταν άνω φραγμένη από το $u < 1$. Για το (ii) παρατηρούμε ότι το σύνολο I_u^{c} είναι επίσης μη κενό, καθώς $F(-\infty) = 0$ και κάθε στοιχείο του I_u^{c} είναι κάτω φράγμα του I_u από την ιδιότητα που έχει η F να είναι αύξουσα. Από την τελευταία ιδιότητα είναι επίσης σαφές ότι το $\mathbb{R} = I_u^{\mathsf{c}} \cup I_u$ και αν $x \in I_u^{\mathsf{c}}$ και $y \in I_u$, τότε $x < y$. Συμπεραίνουμε ότι για κάθε $u \in (0,1)$ θα έχουμε $I_u = (F^{-}(u), +\infty)$ ή $[F^{-}(u), +\infty)$. Όμως, από τη δεξιά συνέχεια της F και το γεγονός ότι $F(x) \geq u$ για $x > F^{-}(u)$, καταλήγουμε στη σχέση

$$F\left(F^{-}(u)\right) = \lim_{x \to \left(F^{-}(u)\right)^{+}} F(x) \geq u \quad \Longrightarrow \quad \left\{x : F(x) \geq u\right\} = \left[F^{-}(u), +\infty\right). \tag{3.22}$$

Στην παρακάτω πρόταση συνοψίζουμε κάποιες στοιχειώδεις ιδιότητες της F^{-}.

Πρόταση 3.12: Η γενικευμένη αντίστροφη $F^{-} : (0,1) \to \mathbb{R}$ ικανοποιεί τις ιδιότητες:

1. $F^{-}(u) = \min\left\{x : F(x) \geq u\right\}$, για κάθε $u \in (0,1)$,
2. $F\left(F^{-}(u)\right) \geq u$, για κάθε $u \in (0,1)$,
3. $F^{-}(u) \leq x \iff u \leq F(x)$, για κάθε $x \in \mathbb{R}$ και $u \in (0,1)$,
4. $F^{-}\left(F(x)\right) \leq x$, για κάθε $x \in \mathbb{R}$,
5. είναι αύξουσα,
6. είναι αριστερά συνεχής.

Απόδειξη: Τα (i), (ii), (iii) προκύπτουν άμεσα από την (3.22).
(iv) Προκύπτει από το ($\Leftarrow$) της (iii) αν θέσουμε $u = F(x)$.
(v) Αρκεί να παρατηρήσουμε ότι για $u < w$ έχουμε $I_u \supset I_w$ και άρα $F^{-}(u) \leq F^{-}(w)$.
(vi) Έστω $u \in (0,1)$. Από το (iv) και το Λήμμα 1.1 υπάρχει το αριστερό όριο της F^{-} στο u και $F^{-}(u-) \leq F^{-}(u)$. Αρκεί λοιπόν να δειχθεί και η αντίστροφη ανισότητα. Καθώς η F^{-} είναι αύξουσα, έχουμε ότι για κάθε $w < u$,

$$F^{-}(u-) \geq F^{-}(w) \quad \Longrightarrow \quad F\left(F^{-}(u-)\right) \geq F(F^{-}(w)) \overset{\text{(ii)}}{\geq} w \quad \overset{w \uparrow u}{\Longrightarrow} \quad F\left(F^{-}(u-)\right) \geq u.$$

Από την τελευταία ανισότητα και την (iii) συμπεραίνουμε ότι $F^{-}(u-) \geq F^{-}(u)$.

□

Στην επόμενη πρόταση παρουσιάζεται μία πολύ ενδιαφέρουσα εφαρμογή της γενικευμένης αντίστροφης. Αποδεικνύεται ότι για κάθε συνάρτηση κατανομής F υπάρχει τυχαία μεταβλητή X με $F_X = F$.

Πρόταση 3.13: Αν $U \sim \mathcal{U}(0,1)$, τότε η τ.μ. $X = F^{-}(U) \sim F$.

Απόδειξη: Κατ'αρχάς χ.β.γ. θεωρούμε ότι $U \in (0,1)$ και άρα η $X = F^{-}(U)$ είναι καλά ορισμένη. Αν $x \in \mathbb{R}$, τότε

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}\left(F^{-}(U) \leq x\right) = \mathbb{P}\left(U \leq F(x)\right) = F(x), \tag{3.23}$$

όπου η τρίτη ισότητα έπεται από την Πρόταση 3.12–(iii) και η τελευταία από τη σχέση $F_U(u) = u$ που ισχύει για τη σ.κ. της ομοιόμορφης για κάθε $u \in [0,1]$.

□

Αν η X είναι συνεχής τ.μ., δηλαδή έχει συνεχή συνάρτηση κατανομής F, τότε ισχύει το εξής ενδιαφέρον αποτέλεσμα.

Πρόταση 3.14: Αν η τ.μ. $X \sim F$ και η X είναι συνεχής τ.μ., τότε
(i) $F\left(F^{-}(u)\right) = u$, για κάθε $u \in (0,1)$,
(ii) η τ.μ. $U = F(X) \sim \mathcal{U}(0,1)$.

Απόδειξη: (i) Έστω $u \in (0,1)$ και $x_u := F^{-}(u)$. Αν $x < x_u$, τότε $F(x) < u$ και άρα

$$F(x_u-) = \lim_{x \uparrow x_u} F(x) \leq u \quad \overset{F \text{ συνεχής}}{\Longrightarrow} \quad F(x_u) \leq u.$$

Από την Πρόταση 3.12–(ii) ισχύει και η αντίστροφη ανισότητα, άρα τελικά $F(x_u) = u$.
(ii) Έστω $u \in (0,1)$. Τότε,

$$\mathbb{P}(U \leq u) = \mathbb{P}\left(F(X) \leq u\right) = \mathbb{P}\left(F(X) < u\right) + \mathbb{P}\left(F(X) = u\right). \tag{3.24}$$

Από υπόθεση η F είναι συνεχής και άρα η εξίσωση $F(x) = u$ έχει πάντα λύση. Επιπλέον, λόγω μονοτονίας της F έχουμε $F^{-1}\left(\{u\}\right) = [a,b]$ για κάποια $a, b \in \mathbb{R}$ με $a \leq b$ (μάλιστα $a = F^{-}(u)$). Καθώς $F(a) = F(b) = u$, συμπεραίνουμε από τα παραπάνω ότι

$$\mathbb{P}\left(F(X) = u\right) = \mathbb{P}\left(X \in F^{-1}\left(\{u\}\right)\right) = \mathbb{P}\left(X \in [a,b]\right) = F(b) - F(a-) = F(b) - F(a) = 0. \tag{3.25}$$

Από την Πρόταση 3.12–(iii) έχουμε $F(x) < u \Leftrightarrow x < F^{-}(u)$ και άρα

$$\mathbb{P}\left(F(X) < u\right) = \mathbb{P}\left(X < F^{-}(u)\right) = \mathbb{P}\left(X \leq F^{-}(u)\right) = F(F^{-}(u)) = u, \tag{3.26}$$

όπου η δεύτερη ισότητα προκύπτει από τη συνέχεια της X και η τρίτη από το (i). Αντικαθιστώντας τις (3.25) και (3.26) στην (3.24) φτάνουμε στο ζητούμενο. Εναλλακτικά, θα μπορούσαμε να δείξουμε πρώτα ότι $\mathbb{P}(U < u) = u$ για κάθε $u \in (0,1)$ και στη συνέχεια να συμπεράνουμε ότι $\mathbb{P}(U \leq u) = \lim_{w \downarrow u} \mathbb{P}(U < w) = u$.

□

Παρατήρηση 3.20. Αν η $X \sim F$ δεν είναι συνεχής τ.μ., δηλαδή η F είναι ασυνεχής σε κάποιο $x \in \mathbb{R}$, τότε δεν ισχύει το παραπάνω αποτέλεσμα. Πράγματι, σε αυτήν την περίπτωση $F(x) - F(x-) > 0$. Συμπεραίνουμε ότι $F(X) \notin \left(F(x-), F(x)\right)$, άρα $\mathbb{P}\left(F(X) \in \left(F(x-), F(x)\right)\right) = \mathbb{P}(\emptyset) = 0$. Έτσι η $F(X)$ δεν ακολουθεί την ομοιόμορφη κατανομή.

Παράδειγμα 3.10: Έστω ότι θέλουμε να παράγουμε παρατηρήσεις από την $Exp(1)$. Υπενθυμίζουμε ότι:

$$F(x) = 1 - e^{-x}, \quad x > 0,$$

και εύκολα βλέπουμε ότι η F είναι συνεχής και γνησίως αύξουσα στο $(0, +\infty)$. Έτσι μπορούμε να ορίσουμε τη συνήθη αντίστροφη της F στο $(0,1)$. Συγκεκριμένα, έχουμε

$$u = 1 - e^{-x} \Leftrightarrow e^{-x} = 1 - u \Leftrightarrow -x = \log(1-u) \Leftrightarrow x = -\log(1-u).$$

Συμπεραίνουμε ότι $F^{-1}(u) = -\log(1-u)$ και από τη μέθοδο της αντιστροφής έχουμε ότι αν $U \sim \mathcal{U}(0,1)$, τότε η $X = -\log(1-U) \sim Exp(1)$. Στην περίπτωση της ομοιόμορφης κατανομής παρατηρούμε ότι $U \stackrel{d}{=} 1-U$ και έτσι μπορούμε να απλοποιήσουμε ακόμα πιο πολύ στη μορφή $X = -\log(U) \sim Exp(1)$.

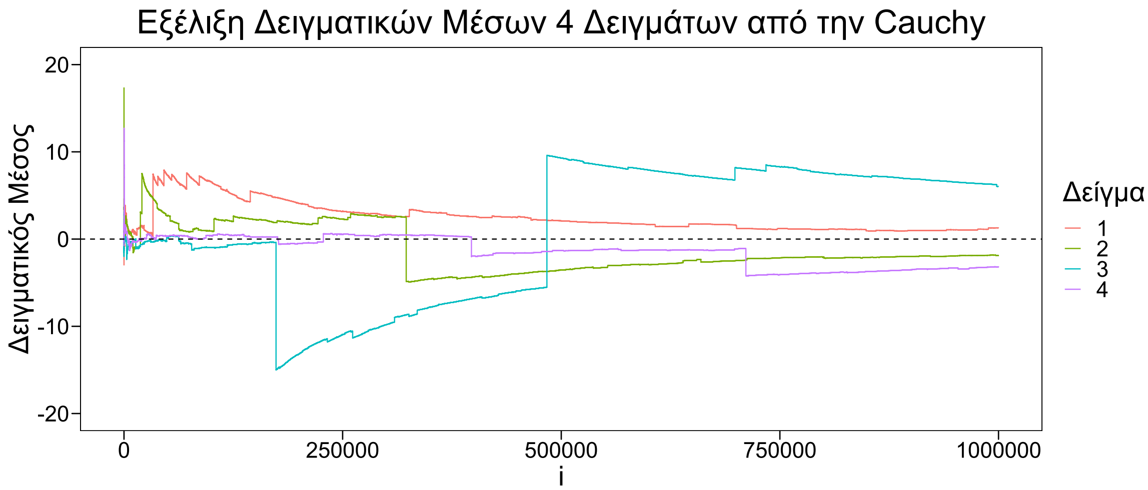
Παράδειγμα 3.11: Έστω ότι θέλουμε να παράγουμε παρατηρήσεις από την $Bin(2, 0.3)$. Εύκολα υπολογίζουμε ότι:

$$\mathbb{P}(X=x) = \begin{cases} 0.49, & \text{x=0} \\ 0.42, & \text{x=1} \\ 0.09, & \text{x=2} \end{cases}, \quad F(x) = \begin{cases} 0.49, & \text{x=0} \\ 0.91, & \text{x=1} \\ 1.00, & \text{x=2} \end{cases}.$$

Η συνάρτηση κατανομής F έχει άλματα ασυνέχειας, επομένως η συνήθης αντίστροφη της F στο $(0,1)$ δεν ορίζεται. Από τον ορισμό της γενικευμένης αντίστροφης έχουμε ότι:

$$F^{-}(u) = \inf\{x \in \mathbb{R} : F(x) \geq u\} = \begin{cases} 0, & 0.00 \leq \text{u} < 0.49 \\ 1, & 0.49 \leq \text{u} < 0.91 \\ 2, & 0.91 \leq \text{u} < 1.00 \end{cases}.$$

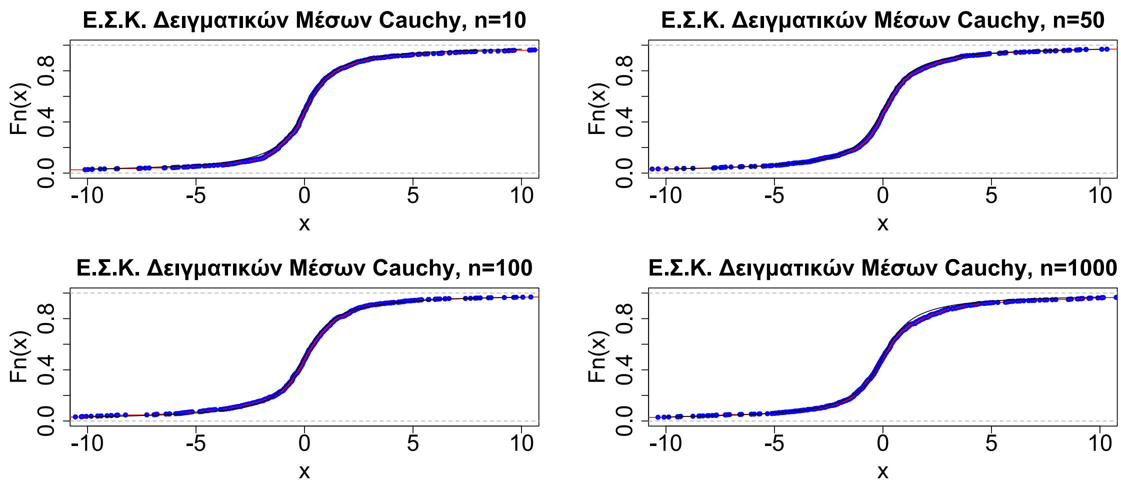
Παράδειγμα 3.12: Σε μαθήματα Πιθανοτήτων και Στατιστικής εισάγεται η κατανομή Cauchy ως μία “περίεργη” κατανομή που η μέση τιμή της δεν ορίζεται. Θα ήταν ενδιαφέρον να μελετήσουμε τη συμπεριφορά του δειγματικού μέσου μίας ακολουθίας α.ι.τ.μ. που ακολουθούν κατανομή Cauchy. Στο Σχ. 3.10 απεικονίζεται η συμπεριφορά του δειγματικού μέσου, καθώς το μέγεθος του δείγματος αυξάνει, σε 4 διαφορετικές περιπτώσεις που αντιστοιχούν σε ανεξάρτητα δείγματα που έχουν παραχθεί από την τυπική Cauchy. Είναι φανερό ότι δεν παρατηρείται σύγκλιση σε κάποια τιμή και μάλιστα υπάρχει μεγάλη μεταβλητότητα. Μπορούμε να αμφισβητήσουμε τον Ισχυρό Νόμο των Μεγάλων Αριθμών;



Σχημα 3.10: Η εξέλιξη των δειγματικών μέσων τεσσάρων ανεξάρτητων τυχαίων δειγμάτων μεγέθους 10^6 από την τυπική κατανομή Cauchy.

Αναδιατάσσουμε τις παρατηρήσεις ώστε να σχηματίσουμε 1.000 ανεξάρτητα δείγματα μεγέθους 1.000 από την τυπική Cauchy. Έστω $\overline{X}_n^{(i)}$ ο δειγματικός μέσος των n πρώτων παρατηρήσεων του i-δείγματος, όταν $1 \leq n \leq 1000$. Οι τυχαίες μεταβλητές $\overline{X}_n^{(1)}, \overline{X}_n^{(2)}, \ldots, \overline{X}_n^{(1000)}$ αποτελούν ένα τυχαίο δείγμα από την κατανομή του δειγματικού μέσου $\overline{X}_n$. Στο Σχ. 3.11 απεικονίζονται οι εμπειρικές συναρτήσεις κατανομής $\mathbb{F}_{n\,;\,1000}$ που αντιστοιχούν στο τυχαίο δείγμα $\left\{\overline{X}_n^{(1)}, \overline{X}_n^{(2)}, \ldots, \overline{X}_n^{(1000)}\right\}$, για $n = 10, 50, 100, 1000$, όπου για κάθε $x \in \mathbb{R}$ και $n \geq 1$

$$\mathbb{F}_{n\,;\,1000}(x) = \frac{1}{1000} \sum_{i=1}^{1000} \mathbf{1}_{(-\infty,x]}(\,\overline{X}_n^{(i)}\,).$$



Σχημα 3.11: Εμπειρικές συναρτήσεις κατανομής που αντιστοιχούν σε τυχαία δείγματα μεγέθους 1000 της κατανομής των δειγματικών μέσων $\overline{X}_n$ για $n = 10, 50, 100, 1000$. Η συνάρτηση κατανομής της τυπικής Cauchy απεικονίζεται με μαύρο.

Από τα παραπάνω γραφήματα προκύπτει ότι η εμπειρική συνάρτηση κατανομής του τυχαίου δείγματος των δειγματικών μέσων προσεγγίζει πολύ καλά τη σ.κ. της τυπικής Cauchy, ενδεικτικό του ότι ο δειγματικός μέσος $\overline{X}_n$ ακολουθεί την τυπική Cauchy για κάθε $n \geq 1$. Για μία εύκολη απόδειξη θα γίνει χρήση χαρακτηριστικών συναρτήσεων. Η χαρακτηριστική συνάρτηση της τυπικής Cauchy είναι η $\phi(t) = e^{-|t|}$. Ισχύει λοιπόν ότι

$$\phi_{\overline{X}_n}(t) = \mathbb{E}\left[e^{it\overline{X}_n}\right] = \left(\mathbb{E}\left[e^{\frac{itX_1}{n}}\right]\right)^n = \left(e^{-\frac{|t|}{n}}\right)^n = e^{-|t|}$$

από όπου έπεται το ζητούμενο. Εξηγεί το αποτέλεσμα αυτό τη μη ισχύ του I.N.M.A. στο συγκεκριμένο παράδειγμα; Πράγματι, αν η ακολουθία των δειγματικών μέσων συγκλίνει με πιθ. 1 σε μια σταθερά μ, τότε θα έπρεπε να συγκλίνει και κατά κατανομή στο μ, το οποίο είναι άτοπο καθώς η κατανομή των δειγματικών μέσων είναι σταθερά η τυπική Cauchy. Αυτό εξηγεί και γιατί ο I.N.M.A. δε μπορεί να εφαρμοστεί στη συγκεκριμένη περίπτωση και γιατί οι υποθέσεις του δεν μπορούν να παραλειφθούν. Μάλιστα, θα μπορούσαμε, ακόμη και αν δεν το γνωρίζαμε, να συμπεράνουμε ότι η μέση τιμή της κατανομής Cauchy δεν υπάρχει. Πράγματι, αν υπήρχε, τότε από τον I.N.M.A. θα έπρεπε οι δειγματικοί μέσοι να συγκλίνουν με πιθ. 1 στη μέση τιμή, το οποίο όπως είδαμε παραπάνω δε συμβαίνει.

3.13 Υπολογισμός και ακριβής κατανομή του KS-στατιστικού

Βασικό συστατικό της απόδειξης της ανισότητας DKW, του υπολογισμού του KS στατιστικού αλλά και της εύρεσης της κατανομής του (ακριβούς, προσεγγιστικής ή ασυμπτωτικής), είναι η παρατήρηση ότι αν η F είναι συνεχής συνάρτηση κατανομής, τότε η μελέτη των ιδιοτήτων της εμπειρικής συνάρτησης κατανομής ανάγεται στη μελέτη των ιδιοτήτων της ομοιόμορφης εμπειρικής συνάρτησης κατανομής, δηλαδή αυτής που προκύπτει όταν πάρουμε ένα τυχαίο δείγμα από την ομοιόμορφη κατανομή $\mathcal{U}(0,1)$. Μάλιστα, όπως θα δούμε, η κατανομή του KS-στατιστικού δεν εξαρτάται από τη συγκεκριμένη F που βρίσκεται πίσω από την παραγωγή των δεδομένων και, άρα, είναι ελεύθερο κατανομής (distribution free).

Πρόταση 3.15: Έστω $\mathbb{F}_n$ η εμπειρική συνάρτηση κατανομής που αντιστοιχεί σε ένα τυχαίο δείγμα $\{X_i\}_{i=1}^n$, από μία κατανομή με σ.κ. F. Αν η κατανομή είναι η ομοιόμορφη $\mathcal{U}(0,1)$, τότε η ε.σ.κ. καλείται *ομοιόμορφη ε.σ.κ.*, συμβολίζεται με $\mathbb{U}_n$, και ισχύει ότι:

$$\mathbb{F}_n = \mathbb{U}_n \circ F, \tag{3.27}$$

ή ισοδύναμα: για κάθε $x \in \mathbb{R}$

$$\mathbb{F}_n(x) = \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{(-\infty,\, F(x)]}(U_k). \tag{3.28}$$

Απόδειξη: Έστω $x \in \mathbb{R}$. Έχουμε

$$\mathbb{F}_n(x) = \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\{X_k \le x\}} = \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\{F^-(U_k) \le x\}}, \tag{3.29}$$

όπου χρησιμοποιήσαμε ότι για κάθε $k \ge 1$ υπάρχει τυχαία μεταβλητή $U_k \sim \mathcal{U}(0,1)$ τέτοια ώστε $X_k = F^-(U_k)$. Οι τυχαίες μεταβλητές $\{U_k\}_{k=1}^n$ μπορούν βέβαια να επιλεγούν ανεξάρτητες μεταξύ τους ώστε να διατηρείται η ανεξαρτησία των $\{X_k\}_{k=1}^n$. Από την ανισότητα της γενικευμένης αντίστροφης $F^-(u) \le x \Leftrightarrow u \le F(x)$ και την (3.29) έχουμε

$$\mathbb{F}_n(x) = \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\{U_k \le F(x)\}} = \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{(-\infty, F(x)]}(U_k),$$

και έτσι καταλήγουμε στη σχέση (3.28) ή ισοδύναμα στην (3.27).

□

Ορισμός 3.19: Έστω $F = F_0$ μία γνωστή συνάρτηση κατανομής. Θέτουμε

$$D_n^+ = \sup_{x \in \mathbb{R}} \Big(\mathbb{F}_n(x) - F_0(x) \Big), \tag{3.30}$$

$$D_n^- = \sup_{x \in \mathbb{R}} \Big(F_0(x) - \mathbb{F}_n(x) \Big), \tag{3.31}$$

$$D_n = \max\{D_n^+, D_n^-\}. \tag{3.32}$$

Παρατήρηση 3.21. Η D_n^+ αναπαριστά τη μέγιστη απόκλιση μεταξύ των γραφημάτων της $\mathbb{F}_n$ και της F_0, για το τμήμα που η $\mathbb{F}_n$ ξεπερνά την F_0, ενώ η D_n^- αναπαριστά την αντίστοιχη μέγιστη απόκλιση για το τμήμα που η $\mathbb{F}_n$ υπολείπεται της F_0. Η μέγιστη λοιπόν κάθετη απόκλιση μεταξύ των δύο γραφημάτων που ορίστηκε μέσω του στατιστικού Kolmogorov-Smirnov D_n στη σχέση (3.20) θα είναι η μέγιστη των παραπάνω δύο αποκλίσεων, όπως εκφράζεται στη σχέση (3.32).

Στην απόδειξη του Θεωρήματος Glivenko-Cantelli διαμερίσαμε το $\mathbb{R}$ σε ένα κατάλληλο πεπερασμένο πλήθος διαστημάτων και δείξαμε τη σύγκλιση με πιθανότητα 1 ενός supremum που εμπλέκει μία άπειρη οικογένεια τυχαίων μεταβλητών, φράσσοντάς το κατάλληλα από ένα μέγιστο ενός πεπερασμένου πλήθους τυχαίων μεταβλητών. Λόγω της μονοτονίας της F, μεταφερόμασταν κάθε φορά στα άκρα του διαστήματος και έτσι, τελικά, σε ένα πεπερασμένο πλήθος σημείων. Εδώ, παρόμοια, ο υπολογισμός των παραπάνω στατιστικών απλοποιείται σημαντικά και, για τον σκοπό αυτό, θα χρησιμοποιήσουμε τη διαμέριση του $\mathbb{R}$ που υποδεικνύεται από τις ασυνέχειες της εμπειρικής συνάρτησης κατανομής.

Πρόταση 3.16: Έστω F_0 μία συνεχής συνάρτηση κατανομής, $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ οι διατεταγμένες παρατηρήσεις που προκύπτουν από το τυχαίο δείγμα $X_1, X_2, \ldots, X_n$, και $U_i = F_0(X_i)$ για κάθε $i = 1, 2, \ldots, n$. Τότε,

$$D_n^+ = \max_{1\leq i\leq n}\left\{\frac{i}{n} - U_{(i)}\right\} \quad \text{και} \quad D_n^- = \max_{1\leq i\leq n}\left\{U_{(i)} - \frac{i-1}{n}\right\}, \tag{3.33}$$

όπου $U_{(1)}, U_{(2)}, \ldots, U_{(n)}$ οι αντίστοιχες διατεταγμένες παρατηρήσεις των $U_1, U_2, \ldots, U_n$.

Απόδειξη: Λόγω της υπόθεσης ότι η F_0 είναι συνεχής, η $\mathbb{F}_n$ μεταβάλλεται από 0 σε 1 με άλματα ασυνέχειας μεγέθους $\frac{1}{n}$ σε κάθε σημείο ασυνέχειάς της. Τα σημεία αυτά συμπίπτουν με τις διατεταγμένες παρατηρήσεις $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$. Είναι τώρα φανερό ότι

$$D_n^+ = \max_{1\leq i\leq n} D_{n,i}^+ \quad \text{και} \quad D_n^- = \max_{1\leq i\leq n} D_{n,i}^-,$$

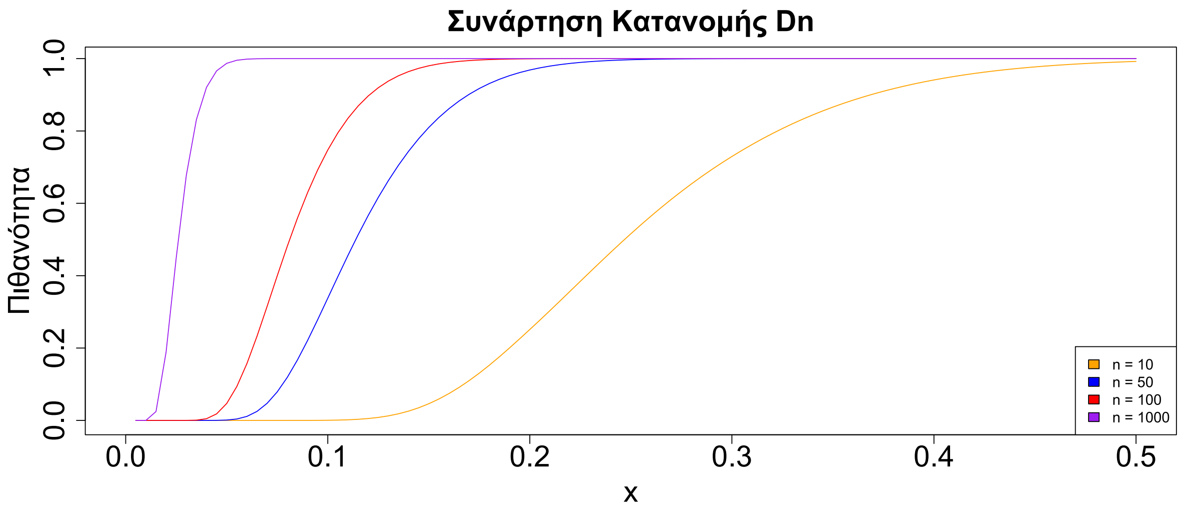
όπου τα $D_{n,i}^+$ και $D_{n,i}^-$ είναι τα αντίστοιχα supremum των συναρτήσεων που εμφανίζονται στα D_n^+ και D_n^-, αλλά περιορισμένα στο διάστημα $\left[X_{(i)}, X_{(i+1)}\right)$ (υποθέσαμε ότι $X_{(0)} = -\infty$ και $X_{(n+1)} = +\infty$). Από τη μονοτονία και τη συνέχεια της F_0, έχουμε ότι το $D_{n,i}^+$ επιτυγχάνεται στο αριστερό άκρο του διαστήματος, ενώ το $D_{n,i}^-$ στο δεξί. Χρησιμοποιώντας ότι $U_{(i)} = F_0\big(X_{(i)}\big)$ και ότι $D_{n,0}^+ = D_{n,0}^- = 0$, καταλήγουμε στο ζητούμενο.

□

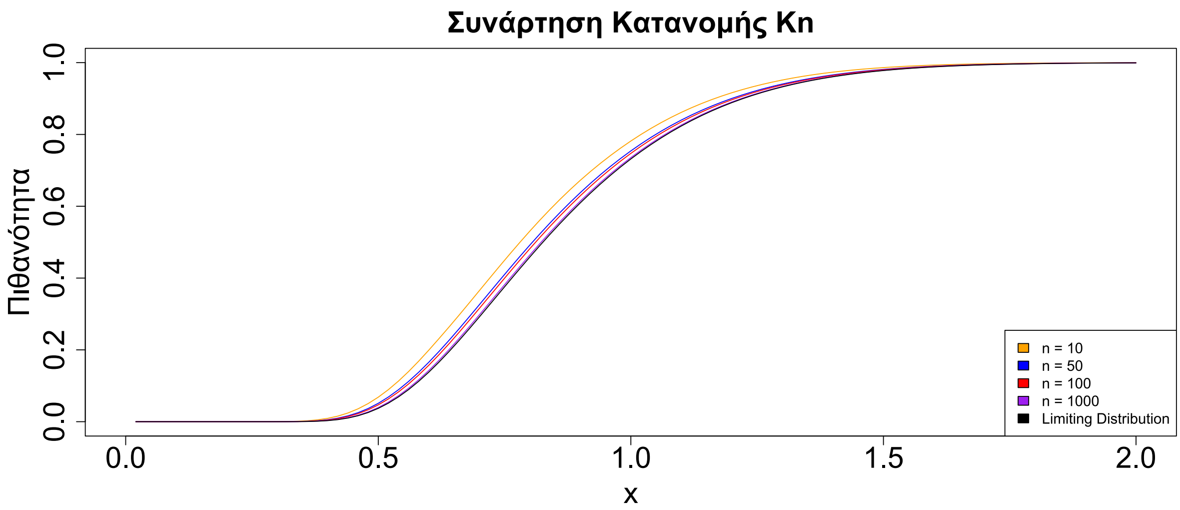
Παρατήρηση 3.22. (α) Στην περίπτωση που η F_0 είναι συνεχής, τότε είναι φανερό ότι οι παρατηρήσεις $U_{(1)}, U_{(2)}, \ldots, U_{(n)}$ αντιστοιχούν σε ένα διατεταγμένο δείγμα από την ομοιόμορφη $\mathcal{U}(0,1)$ και, άρα, τα D_n^+, D_n^- και D_n είναι ελεύθερα κατανομής, καθώς η κατανομή τους δεν εξαρτάται από την F_0.

(β) Τα κριτήρια αυτά μπορούν να εφαρμοστούν και για F_0 διακριτή, αλλά οδηγούν σε συντηρητικούς ελέγχους.

Το στατιστικό Kolmogorov-Smirnov D_n στο οποίο θα αναφερόμαστε σύντομα ως KS-στατιστικό, ή το αντίστοιχο κανονικοποιημένο $K_n = \sqrt{n}D_n$, χρησιμοποιείται ευρύτατα και πάνω από όλα ως ένα μέτρο καλής προσαρμογής μιας υποτιθέμενης συνεχούς κατανομής σε ένα σύνολο δεδομένων. Στα Σχ. 3.12 και 3.13 δίνεται η μορφή της συνάρτησης κατανομής των D_n και K_n αντίστοιχα για διάφορα μεγέθη δείγματος. Η σύγκλιση κατά κατανομή του K_n είναι εμφανής στο παραπάνω σχήμα, ένα αποτέλεσμα που είναι γνωστό ως Θεώρημα του Kolmogorov και θα παρουσιαστεί στο Κεφάλαιο 6 (βλ. **??**)



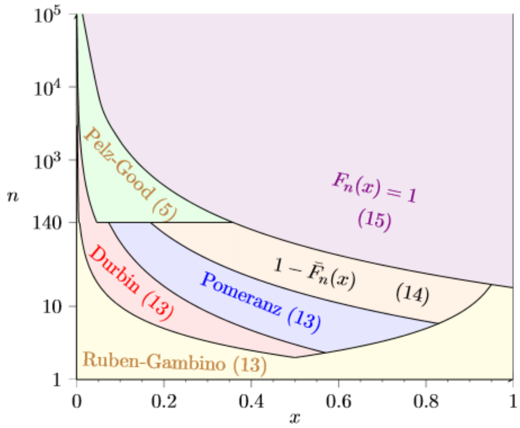
Σχημα 3.12: Εξέλιξη της σ.κ. του KS-στατιστικού D_n για διάφορα μεγέθη δείγματος.



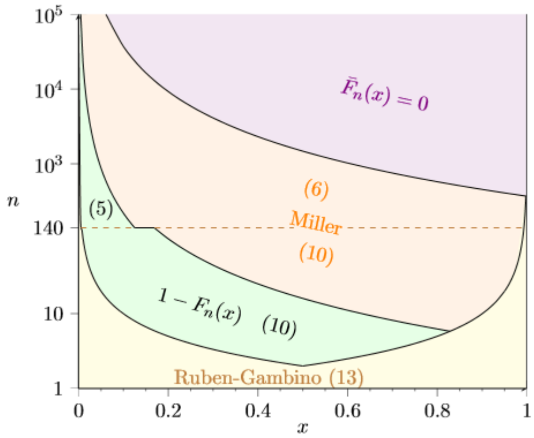
Σχημα 3.13: Εξέλιξη της σ.κ. του κανονικοποιημένου KS-στατιστικού K_n για διάφορα μεγέθη δείγματος.

3.14 Ιστορικά Στοιχεία - Περαιτέρω Μελέτη

Θα περίμενε ίσως κανείς ότι η εύρεση της κατανομής του KS-στατιστικού, που είναι απαραίτητη για να υπολογιστεί η p-value ή οι κρίσιμες τιμές του αντίστοιχου ελέγχου σε διάφορα επίπεδα στατιστικής σημαντικότητας (ε.σ.σ.), δεν θα παρουσίαζε ιδιαίτερες δυσκολίες. Παρ'όλα αυτά, ήταν ακετά απαιτητικό. Ξεκίνησε με την εργασία του A. Kolmogorov (*1933, "Sulla determinazione empirica di una legge di distribuzione", Giornale dell'Instituto Italiano degli Attuari*), τη δημοσίευση πίνακα τιμών της ασυμπτωτικής της προσέγγισης από τον N. Smirnov (*1939, "Estimate of deviation between empirical distribution functions in two independent samples (Russian)", Bull. Moscow Univ.*) και συνεχίστηκε με αρκετές δημοσιεύσεις πάνω στο θέμα αυτό, μέχρι να φτάσουμε σε κάποιο αρκετά ικανοποιητικό αποτέλεσμα (βλ. "*Marsaglia G, Tsang WW, Wang J (2003). Evaluating Kolmogorov's Distribution*"). Σήμερα πλέον, ο υπολογισμός (ακριβής ή προσεγγιστικός) της σ.κ. $F_n(x)$ του D_n ή της συμπληρωματικής της $\overline{F}_n(x)$ (συνάρτηση αξιοπιστίας) γίνεται με γρήγορο και αξιόπιστο τρόπο, για αυθαίρετες τιμές του x και του n. Πολλές μέθοδοι έχουν προταθεί και στα Σχ. 3.14 και 3.15 (*Simard, R. and L'Ecuyer, P., 2011. "Computing the two-sided Kolmogorov-Smirnov distribution"*) παρατηρούμε ότι η επιλογή της καλύτερης μεθόδου υπολογισμού είναι συνάρτηση του x και του n.

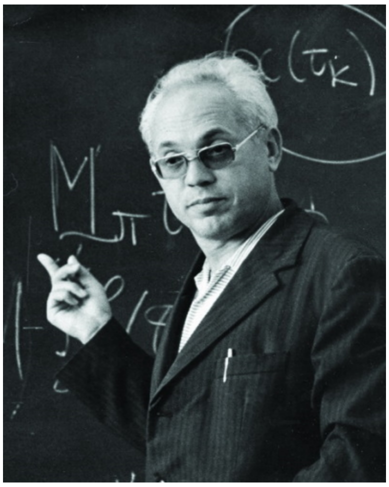


Σχημα 3.14: Επιλογή μεθόδου για να υπολογίσουμε (ή να προσεγγίσουμε) την $F_n(x)$ και τον ελάχιστο αριθμό δεκαδικών ψηφίων στην ακρίβεια της $F_n(x)$ ως συνάρτηση του (x, n) σε κάθε περιοχή του $[0, 1] \times \mathbb{N}^*$.



Σχημα 3.15: Επιλογή μεθόδου για να υπολογίσουμε (ή να προσεγγίσουμε) την $\overline{F}_n(x)$ και τον ελάχιστο αριθμό δεκαδικών ψηφίων στην ακρίβεια της $\overline{F}_n(x)$ ως συνάρτηση του (x, n) σε κάθε περιοχή του $[0, 1] \times \mathbb{N}^*$.

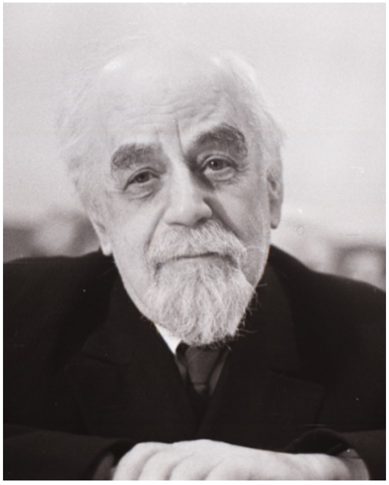
Ανάλογα με την εφαρμογή και το σκοπό που χρησιμοποιείται ο έλεγχος αυτός, η ακρίβεια υπολογισμού θα μπορούσε να παίζει μεγαλύτερο ή μικρότερο ρόλο. Ιστορικά, όταν η διαθεσιμότητα σε υπολογιστές (ή ακόμα και σε κομπιουτεράκια) ήταν εξαιρετικά μικρή, με πολύ μικρότερη υπολογιστική ισχύ, αλλά και το πλήθος των εφαρμογών πολύ περιορισμένο, οι απαιτήσεις σε ακρίβεια ήταν γενικά πολύ μικρότερες. Η δημοσίευση πινάκων με κρίσιμες τιμές (με λίγα δεκαδικά ψηφία) στα συνήθη ε.σ.σ. ήταν συνήθης πρακτική και προσεγγιστικές τιμές, που προκύπτουν είτε από ασυμπτωτικές προσεγγίσεις για μεγάλα δείγματα ή/και διάφορες διορθωμένες τιμές κριτηρίων για μικρότερα δείγματα (διορθώσεις κυρίως μεροληψίας και συνεχείας), ήταν σχετικά ικανοποιητικές για εξαγωγή στατιστικών συμπερασμάτων. Στις εξαιρετικά εξειδικευμένες σύγχρονες εφαρμογές οι απαιτήσεις αυτές έχουν αλλάξει, όπως βέβαια και ο τρόπος που δουλεύουν οι Στατιστικοί, αλλά και όλοι αυτοί που χρησιμοποιούν τα στατιστικά εργαλεία. Η χρήση λογισμικών πακέτων με τεράστιες δυνατότητες, με ήδη ενσωματωμένες πολλές διαφορετικές μεθόδους/συναρτήσεις/επιλογές είναι πλέον ευρέως διαδεδομένη και έτσι το ενδιαφέρον εστιάζεται στη σωστή τους χρήση και μεγαλύτερη ακρίβεια υπολογισμών. Η βαθύτερη όμως κατανόηση των μεθόδων είναι στενά συνδεδεμένη με την ιστορική εξέλιξη και το απαραίτητο μαθηματικό/στατιστικό υπόβαθρο, το οποίο μπορεί να αποτελέσει εφαλτήριο για περαιτέρω βελτιώσεις, αλλά ίσως το πιο σημαντικό, για δυνατότητα επέκτασης και ανανέωσης της στατιστικής μεθοδολογίας σε πολύ μοντέρνες εφαρμογές. Είναι σημαντικό λοιπόν, ακόμα και όταν διαθέτουμε έναν υπερσύγχρονο υπολογιστή, ένα πολύ καλό στατιστικό λογισμικό και το πιο ανανεωμένο πακέτο συναρτήσεων που μας υπολογίζει ακριβείς τιμές μιας ελεγχοσυνάρτησης, όπως αυτή του Kolmogorov-Smirnov, να επενδύσουμε χρόνο στην κατανόηση πολλών διαφορετικών τρόπων με τους οποίους μπορούμε να αντιμετωπίσουμε το πρόβλημα υπολογισμού/προσέγγισης της κατανομής της.



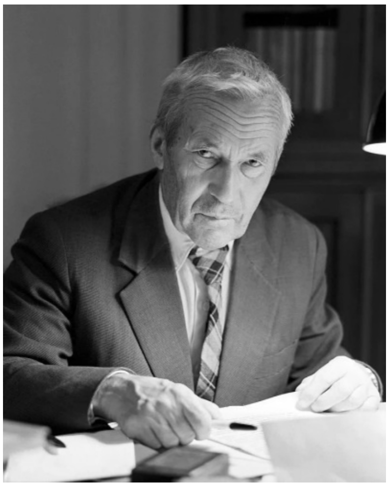
Σχημα 3.16: Anatoliy V. Skorokhod, 1930-2011, σημαντικός Ουκρανός μαθηματικός, εξέδωσε περισσότερες από 450 ερευνητικές εργασίες, κυρίως για στοχαστικές διαφορικές εξισώσεις, οριακά θεωρήματα στη θεωρία Πιθανοτήτων και στη μαθηματική Στατιστική.

Σχημα 3.17: Valery I. Glivenko, 1896-1940: Ρώσος μαθηματικός, η δουλειά του αφορούσε τα θεμέλια των Μαθηματικών, Πραγματική Ανάλυση, Θεωρία Πιθανοτήτων και Μαθηματική Στατιστική. Το μεγαλύτερο μέρος της δουλειάς του δημοσιεύτηκε στα γαλλικά.

Σχημα 3.18: Francesco P. Cantelli, 1875-1966: Ιταλός μαθηματικός, δούλεψε αρχικά στην αστρονομία και τη μηχανική και αργότερα στη Θεωρία Πιθανοτήτων. Ήταν ιδρυτής του Istituto Italiano degli Attuari για τις εφαρμογές των Μαθηματικών στα Οικονομικά.



Σχημα 3.19: Nikolai V. Smirnov: 1900-1966, Ρώσος μαθηματικός με συνεισφορά στη Θεωρία Πιθανοτήτων και τη Στατιστική, ειδικότερα στη Μη Παραμετρική και την Ασυμπτωτική Στατιστική.



Σχημα 3.20: Andrey N. Kolmogorov: 1903-1987, εμβληματικός Ρώσος μαθηματικός με σημαντική συνεισφορά στη Θεωρία Πιθανοτήτων (αξιωματική θεμελίωση μεταξύ άλλων), την Τοπολογία, την Υπολογιστική Πολυπλοκότητα, τη Θεωρία Πληροφορίας και την Κλασική Μηχανική.

Σχημα 3.21: Aryeh Dvoretzky: 1916-2008, Ισραηλινός μαθηματικός γεννημένος στη Ρωσία. Γνωστός για τη συνεισφορά του στη Συναρτησιακή Ανάλυση, τη Στατιστική και τη Θεωρία Πιθανοτήτων. Διετέλεσε πρόεδρος του Ινστιτούτου επιστήμης Weizmann.

Ασκήσεις

3.14.1 Να δειχθεί ότι η οικογένεια $\mathcal{C}^o$ των ανοικτών κυλίνδρων του $\mathbb{R}^T$

1. είναι βάση της τοπολογίας γινόμενο,
2. έχει ως υποβάση την $\mathcal{C}^{\,o;1} := \{\pi_t^{-1}(A_t) \mid t \in T, A_t$ ανοικτό στο $\mathbb{R}\}$,
3. παράγει τη σ–άλγεβρα γινόμενο $\mathcal{B}^T(\mathbb{R})$.

3.14.2 Να δειχθεί ότι για $T = \mathbb{R}$ η σχέση $\mathcal{B}^T(\mathbb{R}) \subsetneqq \mathcal{B}(\mathbb{R}^T)$ δεν προκύπτει από μία "μικρή" διαφορά συνόλων, αλλά υπάρχει διαφορά πληθαρίθμου μεταξύ των αντίστοιχων οικογενειών.

3.14.3 Αποδείξτε το αποτέλεσμα της Παρατήρησης **??**.

3.14.4 Αποδείξτε ότι

$$X \overset{a.s.}{=} Y \quad\Longrightarrow\quad X \overset{\text{mdf}}{=\!=} Y \quad\Longrightarrow\quad X \overset{d}{=} Y,$$

όπου η ισότητα $\overset{p.w.-a.s.}{=}$ αντιστοιχεί στη σχέση τροποποίησης μεταξύ δύο στοχαστικών διαδικασιών.

3.14.5 Αποδείξτε την Πρόταση 3.9.

3.14.6 Αποδείξτε ότι ο χώρος $\left(D_{[a,b]}\,,\, \|\cdot\|_\infty\right)$ είναι χώρος Banach.

3.14.7 Έστω (S, d) ένας διαχωρίσιμος μετρικός χώρος.

(α) Αποδείξτε ότι αν $R \subset S$, τότε το R είναι και αυτό διαχωρίσιμο (ως προς τη σχετική μετρική).

(β) Συμπεράνετε ότι ο χώρος $\left(D_{[a,b]}\,,\, \|\cdot\|_\infty\right)$ δεν είναι διαχωρίσιμος, εντοπίζοντας κατάλληλο μη διαχωρίσιμο υποσύνολό του.

3.14.8 Να επαναλάβετε τα παραδείγματα 3.10 και 3.11 για τις κατανομές

1. $\mathcal{N}(0, 1)$,
2. $\mathcal{L}(0, 1)$,
3. $Pois(5)$,
4. $Geo(0.6)$.

3.14.9 Να επαναλάβετε τα παραδείγματα 3.5 και **??** για δείγμα από την κατανομή $Exp(2)$.

Μέρος II

ΟΝΟΜΑ ΜΕΡΟΥΣ II

ΚΕΦΑΛΑΙΟ 4

Η ΕΜΠΕΙΡΙΚΗ ΔΙΑΔΙΚΑΣΙΑ

4.1 Εμπειρική διαδικασία και σύγκλιση της ο.π.δ.κ. της

Μία από τις πρώτες εφαρμογές των οριακών θεωρημάτων της προηγούμενης ενότητας αφορά την ασθενή σύγκλιση των πεπερασμένης διάστασης κατανομών της εμπειρικής συνάρτησης κατανομής. Στην Πρόταση **??** είχαμε βρει την ακριβή κατανομή, η οποία συνδέεται με την κατανομή των μερικών αθροισμάτων των συνιστωσών ενός τυχαίου διανύσματος που ακολουθεί πολυωνυμική κατανομή.

Στην επόμενη πρόταση προσδιορίζουμε την οριακή κατανομή των πεπερασμένης διάστασης κατανομών της εμπειρικής συνάρτησης κατανομής.

Πρόταση 4.1: Έστω $m \geq 1$ και I μία διατεταγμένη m-άδα πραγματικών αριθμών $t_1 < t_2 < \cdots < t_m$. Τότε, για κάθε $F \in \mathcal{F}$

$$\sqrt{n}\left(\mathbb{F}_{n;I} - F_I\right) \xrightarrow{d} \mathcal{N}_m\left(0_m\,, \Sigma_{F;I}\right), \tag{4.1}$$

όπου $\Sigma_{F;I} = \left(F(t_{i\wedge j}) - F(t_i)F(t_j)\right)_{i,j}$.

Απόδειξη: Έστω $m \geq 1$ και I η παραπάνω διατεταγμένη m-άδα. Τότε,

$$\mathbb{F}_{n;I} = s \circ \delta\left(\mathbb{F}_{n;I}\right) = s\left(\Delta\,\mathbb{F}_{n;I}\right). \tag{4.2}$$

Σάμης Τρέβεζας (2023). «Μη Παραμετρική Στατιστική».
Αθήνα: Κάλλιπος, Ανοικτές Ακαδημαϊκές Εκδόσεις.
Creative Commons Αναφορά Δημιουργού - Μη Εμπορική Χρήση - Παρόμοια Διανομή 4.0

Από την Πρόταση **??** έχουμε δείξει ότι

$$\Delta\,\mathbb{F}_{n;I} \sim \frac{1}{n}\mathcal{M}_m(n,\,\Delta F_I)$$

και από το πολυδιάστατο Κεντρικό Οριακό Θεώρημα ισχύει ότι

$$\sqrt{n}\,\big(\Delta\,\mathbb{F}_{n;I} - \Delta F_I\big) \xrightarrow{d} X \sim \mathcal{N}_m\big(0_m\,,\Sigma_{F;I}\big),$$

όπου

$$\Sigma_{F;I} = \mathrm{dg}\,(\Delta F_I) - (\Delta F_I)\,(\Delta F_I)^\top. \tag{4.3}$$

Από τη σχέση (4.2), τη γραμμικότητα της s και το Λήμμα Slutsky έχουμε

$$\sqrt{n}\,\big(\mathbb{F}_{n;I} - F_I\big) \xrightarrow{d} S\,X \sim \mathcal{N}_m\big(0_m\,,\,S\,\Sigma_{F;I}\,S^\top\big).$$

Αντικαθιστώντας τον $\Sigma_{F;I}$ από την (4.3), έχουμε

$$S\,\Sigma_{F;I}\,S^\top = S\,\mathrm{dg}\,(\Delta F_I)\,S^\top - S\,(\Delta F_I)\,(\Delta F_I)^\top\,S^\top. \tag{4.4}$$

Από το Λήμμα 2.2 που αφορά ιδιότητες του S, εύκολα φτάνουμε στη σχέση

$$S\,\mathrm{dg}(\Delta F_I)\,S^\top = \big(F(t_{i\wedge j})\big)_{i,j}, \tag{4.5}$$

και επιπλέον ότι

$$S\,(\Delta F_I)\,(\Delta F_I)^\top\,S^\top = \underbrace{S\,\Delta}_{I_m}\,F_I F_I^\top\,\underbrace{\Delta^\top S^\top}_{I_m} = F_I F_I^\top = \big(F(t_i)\,F(t_j)\big)_{i,j}. \tag{4.6}$$

Αντικαθιστώντας τις σχέσεις (4.5) και (4.6) στη σχέση (4.4), παίρνουμε το ζητούμενο. □

Το παραπάνω αποτέλεσμα μας δείχνει ότι οι πεπερασμένης διάστασης κατανομές της ε.σ.κ. κατάλληλα κανονικοποιημένες, σταθεροποιούνται οριακά σε κατανομές που θα θέλαμε με τη σειρά τους να αποτελούν την οικογένεια πεπερασμένης διάστασης κατανομών (ο.π.δ.κ.) μιας άλλης στοχαστικής διαδικασίας. Υπάρχει λοιπόν ένα κίνητρο μελέτης της αντίστοιχης ακολουθίας στοχαστικών διαδικασιών.

Ορισμός 4.1: Έστω $\{X_i\}_{i=1}^n$ τυχαίο δείγμα από συνάρτηση κατανομής F. Η στοχαστική διαδικασία $\mathbb{G}_n = \big(\mathbb{G}_n(t)\big)_{t\in\mathbb{R}}$, όπου για κάθε $t \in \mathbb{R}$

$$\mathbb{G}_n(t) = \sqrt{n}\,\big(\mathbb{F}_n(t) - F(t)\big),$$

καλείται *εμπειρική διαδικασία*.

Αναδιατυπώνοντας το αποτέλεσμα της προηγούμενης πρότασης συμπεραίνουμε ότι κάθε τυχαίο (πεπερασμένο) υποδιάνυσμα $\mathbb{G}_{n;I}$ της $\mathbb{G}_n$ συγκλίνει κατά κατανομή σε μία κανονική κατανομή και άρα, από γνωστό αποτέλεσμα, υπάρχει ένα τυχαίο διάνυσμα G_I, έτσι ώστε $\mathbb{G}_{n;I} \xrightarrow{d} G_I$. Τίθεται λοιπόν φυσιολογικά το ερώτημα αν μπορούμε από το παραπάνω αποτέλεσμα να συμπεράνουμε συνολικά ότι η $\mathbb{G}_n$ συγκλίνει κατά κατανομή σε κάποια στοχαστική διαδικασία $\mathbb{G}$ που να πληροί την ιδιότητα ότι $\mathbb{G}_I \stackrel{d}{=} G_I$ για κάθε I πεπερασμένο διάνυσμα δεικτών του $\mathbb{R}$. Δύο είναι τα προβλήματα που ανακύπτουν. Το πρώτο αφορά την ίδια την ύπαρξη μιας στοχαστικής διαδικασίας που να έχει ως ο.π.δ.κ. μία συγκεκριμένη οικογένεια, και η απάντηση είναι καταφατική (στη συγκεκριμένη περίπτωση) λόγω του θεωρήματος συνέπειας του Kolmogorov (θα το δούμε αργότερα). Επιπλέον, το είδος της στοχαστικής διαδικασίας που προκύπτει, ανήκει σε μία πολύ σημαντική κατηγορία στοχαστικών διαδικασιών, τις λεγόμενες Γκαουσιανές, που παίζουν τον ρόλο που παίζει η κανονική κατανομή όταν αναφερόμαστε σε τυχαία διανύσματα. Είναι, με άλλα λόγια, *η απειροδιάστατη γενίκευση των πολυδιάστατων κανονικών κατανομών* και αξίζει να αναφερθούμε σύντομα σε αυτές. Πριν όμως φτάσουμε εκεί, ας δούμε και το δεύτερο πρόβλημα, το οποίο ίσως να μην το περίμενε κανείς. Είδαμε ότι η κατανομή μιας σ.δ. καθορίζεται πλήρως από την ο.π.δ.κ. της. Θα ήταν λοιπόν λογικό να υποθέσει κανείς ότι αν μία ακολουθία σ.δ. είχε την ιδιότητα να συγκλίνει (ασθενώς) κάθε πεπερασμένης διάστασης κατανομή της (ως ακολουθία) στην αντίστοιχη πεπερασμένης διάστασης κατανομή μιας άλλης σ.δ., τότε αυτό θα ήταν αρκετό για να πούμε ότι συγκλίνει κατά κατανομή. Δυστυχώς, τα πράγματα δεν είναι τόσο απλά, όπως θα φανεί στο Παράδειγμα 4.1. Σχετικός είναι και ο επόμενος ορισμός.

Ορισμός 4.2: Έστω (X_n) μία ακολουθία στοχαστικών διαδικασιών. Θα λέμε ότι η (X_n) ***συγκλίνει π.δ.-κατά κατανομή*** σε μία σ.δ. X και θα το συμβολίζουμε $X_n \xrightarrow{fd} X$ (fd: finite dimensional), αν για κάθε πεπερασμένη επιλογή πραγματικών δεικτών I

$$X_{n;I} \xrightarrow{d} X_I. \tag{4.7}$$

Είναι εύκολο να συμπεράνουμε ότι από τη σύγκλιση κατά κατανομή παίρνουμε τη σύγκλιση π.δ.-κατά κατανομή.

Πρόταση 4.2: $X_n \xrightarrow{d} X \implies X_n \xrightarrow{fd} X.$

Απόδειξη: Έστω $X_n \xrightarrow{d} X$ και I ένα πεπερασμένο διάνυσμα δεικτών του T. Η προβολή

$$\pi_I : S^T \to \mathbb{R}^{|I|}, \text{ όπου } \pi_I(x) = x_I,$$

είναι συνεχής συνάρτηση και άρα από την υπόθεση και το Θεώρημα Συνεχούς Απεικόνισης συμπεραίνουμε ότι

$$X_{n;I} = \pi_I(X_n) \xrightarrow{d} \pi_I(X) = X_I.$$

Από το παραπάνω συνάγουμε ότι $X_n \xrightarrow{fd} X$.

□

Θα δείξουμε τώρα ότι η $\xrightarrow{fd}$ δεν αρκεί για να συμπεράνουμε την $\xrightarrow{d}$.

Παράδειγμα 4.1: Έστω (X_n) μία ακολουθία στοχαστικών διαδικασιών (ορισμένων στον ίδιο χώρο πιθανότητας) με τιμές στον χώρο Banach των συνεχών (και φραγμένων) συναρτήσεων $\left(C[0,1],\ \|\cdot\|_\infty\right)$, όπου για κάθε $\omega \in \Omega$:

$$X_n(t,\omega) = \begin{cases} nt, & \text{αν} \quad t \in [0, 1/n), \\ 2 - nt, & \text{αν} \quad t \in [1/n, 2/n), \\ 0, & \text{αν} \quad t \in [2/n, 1), \end{cases}$$

και άρα όλες οι πραγματοποιήσεις της (X_n) συμπίπτουν με την παραπάνω ακολουθία "ντετερμινιστικών" συναρτήσεων. Θα δείξουμε πρώτα ότι $X_n \xrightarrow{fd} X = 0$, όπου η $X = 0$ αναφέρεται στη μηδενική στοχαστική διαδικασία, όπου όλα τα μονοπάτια της αντιστοιχούν στη μηδενική συνάρτηση. Είναι σαφές ότι η (X_n) συγκλίνει κατά σημείο στη μηδενική συνάρτηση, και άρα για κάθε επιλογή πεπερασμένων δεικτών I, η ακολουθία των τυχαίων διανυσμάτων $X_{n;I} \xrightarrow{a.s.} X_I$ που συνεπάγεται τη σύγκλιση κατά κατανομή. Το αποτέλεσμα αυτό μας δείχνει ότι η $X_n \xrightarrow{fd} X = 0$. Αν υποθέσουμε τώρα ότι $X_n \xrightarrow{d} X$, τότε από το θεώρημα της συνεχούς απεικόνισης και το γεγονός ότι κάθε νόρμα είναι (ομοιόμορφα) συνεχής συνάρτηση (στον δικό της χώρο) συμπεραίνουμε ότι

$$1 = \|X_n\|_\infty \xrightarrow{d} \|X\|_\infty = 0,$$

το οποίο είναι βέβαια άτοπο και άρα η X_n δε συγκλίνει κατά κατανομή στη X.

4.2 Γκαουσιανή στοχαστική διαδικασία

Ας επιστρέψουμε τώρα στην προσπάθεια γενίκευσης των κανονικών κατανομών. Πριν ορίσουμε αυστηρά τις Γκαουσιανές στοχαστικές διαδικασίες θα χρειαστούμε την έννοια της μεσοσυνάρτησης και του πυρήνα συνδιακύμανσης μιας στοχαστικής διαδικασίας.

Ορισμός 4.3: Έστω $X = \left(X(t)\right)_{t \in T}$ μία στοχαστική διαδικασία για την οποία $\mathbb{E}\,|X(t)| < +\infty$, για κάθε $t \in T$. Η συνάρτηση $\mu \colon T \to \mathbb{R}$ με

$$\mu(t) = \mathbb{E}\left[X(t)\right],$$

καλείται *μεσοσυνάρτηση* της σ.δ. X. Αν, επιπλέον, $\mathbb{E}\left[X^2(t)\right] < +\infty$, για κάθε $t \in T$, τότε η συνάρτηση $\Sigma : T \times T \to \mathbb{R}$ με

$$\Sigma(s,t) = \mathbb{C}\left[X(s), X(t)\right],$$

καλείται *συνάρτηση συνδιακύμανσης* ή *πυρήνας συνδιακύμανσης* της σ.δ. X.

Παρατήρηση 4.1. Αν $I = t_{1:d}$ είναι μία επιλογή d δεικτών του T, και θέσουμε

$$\Sigma_I := \big(\Sigma(t_i, t_j)\big)_{i,j},$$

τότε είναι φανερό ότι $\Sigma_I = \mathbb{V}(X_I)$ και άρα ο πίνακας Σ_I είναι θετικά ημιορισμένος πίνακας. Όταν μία συνάρτηση Σ ικανοποιεί την παραπάνω ιδιότητα για κάθε επιλογή I όπως παραπάνω, τότε καλείται *θετικά ημιορισμένη*. Έτσι προκύπτει ότι ένας πυρήνας συνδιακύμανσης είναι αναγκαστικά συμμετρικός, αλλά και θετικά ημιορισμένος.

Ορισμός 4.4: Έστω $X = \big(X(t)\big)_{t\in T}$ μία στοχαστική διαδικασία, $\mu : T \to \mathbb{R}$ μία τυχούσα συνάρτηση και $\Sigma : T \times T \to \mathbb{R}$ μία συμμετρική και θετικά ημιορισμένη συνάρτηση. Θα λέμε ότι η X είναι *Γκαουσιανή διαδικασία* με μεσοσυνάρτηση μ και πυρήνα συνδιακύμανσης Σ, και θα γράφουμε $X \sim \mathbb{G}(\mu, \Sigma)$, αν για κάθε πεπερασμένη επιλογή δεικτών I του T:

$$X_I \sim \mathcal{N}_{|I|}(\mu_I, \Sigma_I).$$

Όπως μία κανονική κατανομή καθορίζεται πλήρως από το διάνυσμα των μέσων τιμών και τον πίνακα συνδιακύμανσης, έτσι και σε μία Γκαουσιανή στοχαστική διαδικασία η γνώση της μεσοσυνάρτησης και του πυρήνα συνδιακύμανσης καθορίζει πλήρως την κατανομή της. Από την άλλη μεριά, όλα θα ήταν κενά περιεχομένου αν δεν μπορούσαμε να εξασφαλίσουμε την ύπαρξη τέτοιων στοχαστικών διαδικασιών. Το επόμενο θεώρημα κινείται σε αυτή την κατεύθυνση, εξασφαλίζοντάς μας ότι οι αναγκαίες συνθήκες που πρέπει να πληροί ο Σ είναι και ικανές για την ύπαρξή της.

Θεώρημα 4.1: Έστω $T \subseteq \mathbb{R}$, $\mu\colon T \to \mathbb{R}$ μία τυχούσα συνάρτηση και $\Sigma : T\times T \to \mathbb{R}$ μία συμμετρική και θετικά ημιορισμένη συνάρτηση. Τότε υπάρχει χώρος πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$ και τυχαίες μεταβλητές $X_t : \Omega \to \mathbb{R}$, $\forall t \in T$, έτσι ώστε η στοχαστική διαδικασία $X = (X_t)_{t\in T} \sim \mathbb{G}(\mu, \Sigma)$.

Η απόδειξη του θεωρήματος είναι στην ουσία εφαρμογή ενός γενικότερου θεωρήματος, που κατοχυρώνει την ύπαρξη στοχαστικών διαδικασιών στην περίπτωση που η ο.π.δ.κ. τους ικανοποιεί κάποιες "φυσιολογικές" απαιτήσεις συνέπειας. Στην κατεύθυνση αυτή δίνουμε τον παρακάτω ορισμό.

Ορισμός 4.5: Έστω T ένα δεικτοσύνολο, $I = t_{1:n}$ ένα πεπερασμένο διάνυσμα δεικτών του T για κάποιο $n \geq 1$ και $A = A_1 \times A_2 \times \cdots \times A_n$ ένα μετρήσιμο ορθογώνιο του $\mathbb{R}^n$. Αν σ είναι μία μετάθεση του $\{1, 2, \ldots, n\}$, τότε συμβολίζουμε με

$$
\begin{aligned}
I_\sigma &:= \left(t_{\sigma(1)}, t_{\sigma(2)}, \ldots, t_{\sigma(n)}\right),\\
A_\sigma &:= A_{\sigma(1)} \times A_{\sigma(2)} \times \cdots \times A_{\sigma(n)}.
\end{aligned}
$$

Ας υποθέσουμε λοιπόν ότι έχουμε μία στοχαστική διαδικασία $X = \big(X(t)\big)_{t\in T}$ με τιμές στον $\left(\mathbb{R}^T, \mathcal{B}^T\right)$, όπου $\mathcal{B}^T$ είναι η σ-άλγεβρα γινόμενο (ή κυλινδρική σ-άλγεβρα του $\mathbb{R}^T$). Είναι φανερό λοιπόν ότι για κάθε πεπερασμένο διάνυσμα δεικτών I του T, μετάθεση σ και μετρήσιμο ορθογώνιο A, έχουμε

$$
\mathbb{P}(X_I \in A) = \mathbb{P}\left(\bigcap_{i=1}^{n} \{X_{t_i} \in A_i\}\right) = \mathbb{P}\left(\bigcap_{i=1}^{n} \left\{X_{t_{\sigma(i)}} \in A_{\sigma(i)}\right\}\right) = \mathbb{P}\left(X_{I_\sigma} \in A_\sigma\right). \tag{4.8}
$$

Επιπλέον, αν $I' = t_{1:n+1}$ είναι μία επέκταση του I με ένα επιπλέον στοιχείο t_{n+1}, τότε

$$
\mathbb{P}_{X_I}(A) = \mathbb{P}(X_I \in A) = \mathbb{P}\left(X_I \in A, X_{t_{n+1}} \in \mathbb{R}\right) = \mathbb{P}\left(X_{I'} \in A \times \mathbb{R}\right) = \mathbb{P}_{X_{I'}}(A \times \mathbb{R}). \tag{4.9}
$$

Αναδιατυπώνοντας τις (4.8) και (4.9) με τη βοήθεια των κατανομών $\mu_I := \mathbb{P}_{X_I}$, έχουμε

$$
\mu_I(A) = \mu_{I_\sigma}(A_\sigma) \qquad \text{και} \qquad \mu_I(A) = \mu_{I'}(A \times \mathbb{R}) \tag{4.10}
$$

Οι παραπάνω συνθήκες ικανοποιούνται από την οικογένεια μέτρων πιθανότητας $\left\{\mathbb{P}_{X_I}\right\}_I$ της στοχαστικής διαδικασίας X και άρα είναι αναγκαίες συνθήκες που πρέπει να ικανοποιεί μία οικογένεια μέτρων πιθανότητας για να αντιστοιχεί σε ο.π.δ.κ. μίας στοχαστικής διαδικασίας. Είναι γνωστές με το όνομα *συνθήκες συνέπειας*.

Ορισμός 4.6: Έστω $T \subset \mathbb{R}$ και $\mathcal{I}$ η συλλογή όλων των πεπερασμένων υποδιανυσμάτων I που μπορούμε να σχηματίσουμε με στοιχεία του T (μπορούν να επιλεγούν και με διακεκριμένα στοιχεία του T). Αν $\{\mu_I\}_{I\in\mathcal{I}}$ είναι μία οικογένεια μέτρων πιθανότητας που ικανοποιεί τις σχέσεις (4.10), τότε αυτή καλείται *συνεπής*.

Σύμφωνα με το παρακάτω θεώρημα οι συνθήκες συνέπειας είναι και ικανές συνθήκες ύπαρξης στοχαστικής διαδικασίας που να έχει μία δεδομένη συνεπή οικογένεια μέτρων πιθανότητας ως ο.π.δ.κ. της.

Θεώρημα 4.2 (*Daniell-Kolmogorov; συνέπειας/ύπαρξης του Kolmogorov*): Αν $(\mu_I)_{I\in\mathcal{I}}$ είναι μία συνεπής οικογένεια μέτρων πιθανότητας, τότε υπάρχει χώρος πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$ και στοχαστική διαδικασία $X = (X_t)_{t\in T}$ ορισμένη σε αυτόν, που έχει την παραπάνω οικογένεια ως ο.π.δ.κ. της, δηλαδή $X_I \sim \mu_I$, για κάθε $I \in \mathcal{I}$.

4.3 Θεωρήματα Καραθεοδωρή*

Τα εξωτερικά μέτρα, που είναι συνολοσυναρτήσεις με λιγότερες απαιτήσεις από τα μέτρα, μας δίνουν τη δυνατότητα κατασκευής μέτρων με ενδιαφέρουσες ιδιότητες, με σημαντικότερο παράδειγμα την κατασκευή του μέτρου Lebesgue. Η παρεμβολή των εξωτερικών μέτρων εδώ γίνεται σε σχέση με τη δυνατότητα κατασκευής μέτρων πιθανότητας πάνω σε απειροδιάστατους χώρους όπως είναι οι κατανομές των στοχαστικών διαδικασιών.

Ορισμός 4.7 (άλγεβρα συνόλων): Μία οικογένεια $\mathcal{A}$ υποσυνόλων ενός συνόλου S καλείται *άλγεβρα* στο S αν ικανοποιεί τις εξής ιδιότητες:

1.
(ii) αν $A \in \mathcal{A}$, τότε $A^c := S \setminus A \in \mathcal{A}$,
(iii) αν $A_1, A_2, \ldots, A_n \in \mathcal{A}$, $n \in \mathbb{N}$, τότε $\bigcap_{i=1}^{n} A_i \in \mathcal{A}$.

Με τις ιδιότητες (ii) και (iii) απαιτούμε η $\mathcal{A}$ να είναι κλειστή στα συμπληρώματα και τις πεπερασμένες τομές αντίστοιχα. Η κλειστότητα στις πεπερασμένες ενώσεις έπεται άμεσα από αυτές τις δύο ιδιότητες και μπορεί να αντικαταστήσει τη συνθήκη (iii) του ορισμού. Κάθε σ–άλγεβρα είναι άλγεβρα, ενώ κάθε άλγεβρα δεν είναι γενικά σ–άλγεβρα αφού δεν απαιτεί κλειστότητα σε αριθμήσιμες ενώσεις ή τομές.

Ορισμός 4.8 (εξωτερικό μέτρο): Μία συνολοσυνάρτηση $\phi\colon \mathcal{P}(S) \to [0, +\infty]$ καλείται *εξωτερικό μέτρο* στο S αν ικανοποιεί τις εξής ιδιότητες:

1.
(ii) αν $A \subset B \subset S$, τότε $\phi(A) \leq \phi(B)$ (*μονοτονία*),
(iii) αν (A_n) είναι ακολουθία υποσυνόλων του S, τότε $\phi\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \phi(A_n)$ (*αριθμήσιμη υποπροσθετικότητα ή σ-υποπροσθετικότητα*).

Ενώ ένα εξωτερικό μέτρο ορίζεται πάνω σε όλα τα υποσύνολα του χώρου, μπορούμε μέσα από αυτά να ξεχωρίσουμε μία συλλογή που επικοινωνεί ϕ-καλά με όλα τα υπόλοιπα, με την έννοια της πεπερασμένης προσθετικότητας.

Ορισμός 4.9 (ϕ-μετρήσιμα σύνολα): Έστω $\phi : \mathcal{P}(S) \to [0, +\infty]$ ένα εξωτερικό μέτρο. Ένα σύνολο $B \subset S$ καλείται ϕ-*μετρήσιμο*, αν για κάθε $A \subset S$

$$\phi(A) = \phi(A \cap B) + \phi(A \setminus B).$$

Συμβολίζουμε με $\mathcal{M}_\phi$ την οικογένεια όλων των ϕ−μετρήσιμων υποσυνόλων του S.

Αν περιοριστούμε μόνο στα ϕ−μετρήσιμα υποσύνολα του χώρου, τότε αυτό δίνει τη δυνατότητα κατασκευής ενός μέτρου. Μάλιστα ο περιορισμός αυτός είναι όσο πιο οικονομικός γίνεται (έχουν εξαιρεθεί τα λιγότερα δυνατά σύνολα), υπό την έννοια ότι το μέτρο που προκύπτει δεν μπορεί να επεκταθεί άλλο κατά έναν ϕ−φυσιολογικό τρόπο. Έτσι το μέτρο που προκύπτει είναι πλήρες, δηλ., περιέχει όλα τα ϕ−μηδενικά σύνολα (βλ. Ορισμό 1.27).

Θεώρημα 4.3 (Καραθεοδωρή): Έστω ϕ ένα εξωτερικό μέτρο στο S. Τότε,

1. η οικογένεια $\mathcal{M}_\phi$ είναι σ-άλγεβρα στο S,
2. ο περιορισμός $\phi|\mathcal{M}_\phi$ του ϕ στη $\mathcal{M}_\phi$ είναι πλήρες μέτρο.

Το παρακάτω αποτέλεσμα είναι γνωστό ως θεώρημα επέκτασης του Καραθεοδωρή.

Θεώρημα 4.4 (επέκτασης του Καραθεοδωρή): Έστω $\mathcal{A}$ μία άλγεβρα σε ένα σύνολο S και $\mu : \mathcal{A} \to [0, +\infty]$ μία συνολοσυνάρτηση που ικανοποιεί: (α) $\mu(\varnothing) = 0$ και
(β) αριθμήσιμη προσθετικότητα για ακολουθίες $(A_n)+ \in \mathcal{A}$, τέτοιες ώστε $\cup A_n \in \mathcal{A}$.
Θέτουμε $\phi : \mathcal{P}(S) \to [0, \infty]$ με

$$\phi(A) = \inf\left\{ \sum_{n=1}^{\infty} \mu(A_n) : \; A_n \in \mathcal{A}, \;\; n \geq 1 \text{ και } A \subset \bigcup_{n=1}^{\infty} A_n \right\}.$$

Τότε,

-
(ii) $\sigma(\mathcal{A}) \subset \mathcal{M}_\phi$,
(iii) $\phi|\mathcal{A} = \mu$.

Συμπεραίνουμε ότι το $\widetilde{\mu} := \phi|\,\sigma(\mathcal{A})$ είναι μέτρο στο μετρήσιμο χώρο $(S, \sigma(\mathcal{A}))$ που επεκτείνει το μ. Επιπλέον, αν $\nu\colon \sigma(\mathcal{A}) \to [0, \infty]$ είναι ένα μέτρο ώστε $\nu|\mathcal{A} = \mu$, τότε,

(iv) $\widetilde{\mu}(A) \geq \nu(A)$ για κάθε $A \in \sigma(\mathcal{A})$,
(v) αν υπάρχει (A_n) στην $\mathcal{A}$, τέτοια ώστε $S = \bigcup_{n=1}^{\infty} A_n$ και $\mu(A_n) < \infty$, τότε $\widetilde{\mu} = \nu$ και άρα η επέκταση είναι μοναδική.

4.4 Στοχαστική διαδικασία Lévy

Η κατανομή μίας στοχαστικής διαδικασίας καθορίζεται μόνο από την οικογένεια των κατανομών των τυχαίων διανυσμάτων της. Το ίδιο βέβαια συμβαίνει και με τις Γκαουσιανές στοχαστικές διαδικασίες, οι οποίες έχουν ο.π.δ.κ. που αντιστοιχεί εξ'ολοκλήρου σε κανονικές κατανομές. Εκτός όμως από το είδος των κατανομών, σημαντικό ενδιαφέρον παρουσιάζεται στη συμπεριφορά των μονοπατιών μίας στοχαστικής διαδικασίας. Θα θέλαμε να υπολογίζουμε πιθανότητες "υποψήφιων" ενδεχομένων της μορφής "$X(t)$ συνεχής στο T", "$X(t)$ συνεχής στο $t_0 \in T$" ή "$X(t)$ φραγμένη στο T" και πολλά άλλα ενδεχόμενα που σχετίζονται με τη συμπεριφορά των μονοπατιών της. Μας ενδιαφέρουν επίσης "υποψήφιες" τυχαίες μεταβλητές όπως $\sup_{t\in T}|X(t)|$ ή $\int_T X_t^2\, dt$. Δυστυχώς, όμως, όταν η X θεωρείται με τιμές στον $\left(\mathbb{R}^T, \mathcal{B}^T\right)$, τότε τα παραπάνω σύνολα δεν είναι γενικά ενδεχόμενα (δηλ. μετρήσιμα σύνολα) και οι παραπάνω συναρτήσεις δεν είναι γενικά τυχαίες μεταβλητές (δηλ. μετρήσιμες συναρτήσεις). Αυτά τα προβλήματα εμφανίζονται όταν το T είναι υπεραριθμήσιμο, και σε τέτοιες περιπτώσεις η πληροφορία στην οποία θέλουμε να έχουμε πρόσβαση εξαρτάται από τη συμπεριφορά των μονοπατιών σε ένα υπεραριθμήσιμο πλήθος σημείων. Εμείς όμως με τη σ−άλγεβρα γινόμενο $\mathcal{B}^T$ μπορούμε να ελέγξουμε τη συμπεριφορά των μονοπατιών μόνο σε ένα αριθμήσιμο πλήθος σημείων και έτσι μπορούμε να αποδίδουμε πιθανότητα μόνο σε σύνολα που εμπλέκουν το πολύ ένα αριθμήσιμο πλήθος από τις τυχαίες μεταβλητές (X_t). Από τα παραπάνω είναι σαφές ότι ο ορισμός αυτός είναι ανεπαρκής, καθώς δε μπορούμε να απαντήσουμε πλήρως σε ερωτήματα που εμφανίζονται σε πολλές εφαρμογές, και αφορούν για παράδειγμα την εξέλιξη ενός χρηματιστηριακού δείκτη, την κίνηση Brown, την ροή κυκλοφορίας σε ένα οδικό δίκτυο σε συνεχή χρόνο και πολλά άλλα μοντέρνα προβλήματα. Μία πληρέστερη ανάλυση των προβλημάτων αυτών είναι εφικτή με τη μελέτη της συμπεριφοράς των μονοπατιών μίας στοχαστικής διαδικασίας. Οι πιο χρήσιμες στοχαστικές διαδικασίες συνεχούς χρόνου ικανοποιούν ιδιότητες στη φύση των κατανομών τους, αλλά και στη συμπεριφορά των μονοπατιών τους, που τους επιτρέπουν την ανάπτυξη μίας πλούσιας μαθηματικής θεωρίας, αλλά και μεγαλύτερης εφαρμοσιμότητας.

Μία πολύ σημαντική κατηγορία στοχαστικών διαδικασιών είναι οι λεγόμενες διαδικασίες Lévy, οι οποίες χαρακτηρίζονται μεταξύ άλλων από ανεξάρτητες και στάσιμες προσαυξήσεις.

Ορισμός 4.10 (ανεξάρτητες και στάσιμες προσαυξήσεις):
Μία στοχαστική διαδικασία $X = (X_t)_{t\geq 0}$, λέμε ότι έχει:

- για κάθε $n \geq 1$ και επιλογή δεικτών
  $0 = t_0 < t_1 < t_2 < \ldots < t_n$, οι τυχαίες μεταβλητές

$$\left\{X_{t_1} - X_0,\, X_{t_2} - X_{t_1},\, \ldots\, , X_{t_n} - X_{t_{n-1}}\right\} \tag{4.11}$$

  είναι ανεξάρτητες,

(ii) *στάσιμες προσαυξήσεις*, αν για κάθε $0 < s < t$

$$X_t - X_s \stackrel{d}{=} X_{t-s} - X_0. \tag{4.12}$$

Σε στοχαστικές διαδικασίες συνεχούς χρόνου, πολλές φορές είναι επιθυμητό να ελέγχουμε επαρκώς τις μεταβολές της X_t σε μικρές περιοχές γύρω από το t. Αυτό επιτυγχάνεται με την εισαγωγή μίας έννοιας στοχαστικής συνέχειας που εξασφαλίζει μεταξύ άλλων την ύπαρξη μονοπατιών με πολύ καλές ιδιότητες.

Ορισμός 4.11 (στοχαστική συνέχεια): Μία στοχαστική διαδικασία $X = (X_t)_{t\geq 0}$, καλείται *στοχαστικά συνεχής* (ή *συνεχής κατά πιθανότητα*) στο $t \geq 0$, αν για κάθε $\varepsilon > 0$

$$\lim_{h\to 0} \mathbb{P}\left(|X_{t+h} - X_t| > \varepsilon\right) = 0, \tag{4.13}$$

ή ισοδύναμα,

$$X_{t+h} \xrightarrow[h\to 0]{p} X_t, \tag{4.14}$$

και αν η X είναι στοχαστικά συνεχής για κάθε $t \geq 0$, τότε καλείται *στοχαστικά συνεχής*.

Η μορφή αυτή της συνέχειας είναι βέβαια μία από πολλές διαφορετικές μορφές συνέχειας που θα μπορούσαμε να διατυπώσουμε, όπως συνέχεια με πιθανότητα 1, συνέχεια κατά μέσο τετράγωνο ή συνέχεια κατά κατανομή. Η παρακάτω μορφή συνέχειας που είναι ισχυρότερη από όλες τις παραπάνω αναφέρεται και ως δειγματική συνέχεια.

Ορισμός 4.12 (δειγματική συνέχεια): Μία στοχαστική διαδικασία $X = (X_t)_{t\geq 0}$, καλείται *δειγματικά συνεχής*, ή απλά *συνεχής*, αν έχει συνεχή μονοπάτια πιθανότητα 1, δηλαδή,

$$\mathbb{P}\left(X \in C\left[0, +\infty\right)\right) = \mathbb{P}\left(\left\{\omega \mid X(\omega) \text{ είναι συνεχής}\right\}\right) = 1.$$

Η παραπάνω μορφή συνέχειας είναι βέβαια πολλές φορές επιθυμητή, αλλά είναι πολύ ισχυρή και δεν μπορούμε να την εξασφαλίσουμε πάντα. Όπως θα δούμε στο επόμενο θεώρημα, μία ασθενέστερη μορφή, η στοχαστική συνέχεια, δίνει τη δυνατότητα να δουλεύουμε με στοχαστικές διαδικασίες που έχουν αρκετά ομαλά μονοπάτια, χωρίς να χάνουμε τίποτα από τις στοχαστικές ιδιότητες της διαδικασίας, δηλαδή ιδιότητες που σχετίζονται με την κατανομή τους. Πιο συγκεκριμένα, εξασφαλίζει την ύπαρξη τροποποιήσεων με cadlag μονοπάτια.

Θεώρημα 4.5: Μία στοχαστική διαδικασία X είναι στοχαστικά συνεχής αν και μόνο αν έχει μία τροποποίηση Y που είναι cadlag.

Είμαστε τώρα σε θέση να ορίσουμε τις διαδικασίες *Lévy*.

Ορισμός 4.13 (διαδικασία Lévy): Μία στοχαστική διαδικασία $L = (L_t)_{t \geq 0}$, καλείται *διαδικασία Lévy* αν:

-
- (ii) έχει ανεξάρτητες προσαυξήσεις,
- (iii) έχει στάσιμες προσαυξήσεις,
- (iv) είναι στοχαστικά συνεχής.

Είναι αρκετά ενδιαφέρον να παρατηρήσει κανείς ότι οι διαδικασίες Lévy συμπεριφέρονται σαν συνεχούς χρόνου γενικεύσεις των αθροισμάτων ανεξάρτητων και ισόνομων τυχαίων μεταβλητών. Πράγματι, έστω $0 \leq s < t$, $n \geq 1$ και $t_k = s + \frac{k}{n}(t-s)$, $0 \leq k \leq n$, τα διαδοχικά διαιρετικά σημεία μίας ομοιόμορφης διαμέρισης του $[s,t]$ σε n διαστήματα. Τότε, η προσαύξηση της L στο $[s,t]$ περιγράφεται ως ένα άθροισμα ανεξάρτητων και ισόνομων τυχαίων μεταβλητών μέσω του τηλεσκοπικού αθροίσματος

$$L_t - L_s = (L_{t_1} - L_s) + (L_{t_2} - L_{t_1}) + \ldots + (L_t - L_{t_{n-1}}).$$

Καθώς το n αυξάνεται, είναι φανερό ότι οι μεταβολές της L, άρα και η ίδια η L_t ($L_0 = 0$) πρέπει να αποτυπώνουν τη γενίκευση τέτοιων αθροισμάτων στο συνεχές.

4.5 Στοχαστική διαδικασία Wiener - κίνηση Brown

Το σημαντικότερο παράδειγμα στοχαστικής διαδικασίας είναι η διαδικασία Wiener, περισσότερο γνωστή ως κίνηση Brown.

Ορισμός 4.14 (διαδικασία Wiener): Μία στοχαστική διαδικασία $W = (W_t)_{t \geq 0}$, καλείται *διαδικασία Wiener*, ή *κίνηση Brown*, αν είναι διαδικασία Lévy με τις ενισχυμένες συνθήκες:

- (i)' $W_t \sim \mathcal{N}(0,t)$,
- (iv)' είναι δειγματικά συνεχής, δηλαδή η W έχει με πιθανότητα 1 συνεχή μονοπάτια.

Η διαδικασία Wiener είναι η πιο μελετημένη στοχαστική διαδικασία. Καθιερώθηκε με το όνομα *κίνηση Brown* από τις παρατηρήσεις του Άγγλου βοτανολόγου *R. Brown*, ο οποίος μελέτησε με μικροσκόπιο την άτακτη κίνηση ενός πλήθους σωματιδίων που προέρχονται από κόκκους γύρης όταν είναι βυθισμένα στο νερό (1827). Ήταν ο *A. Einstein* όμως που μοντελοποίησε την κίνηση αυτή ως το αποτέλεσμα της σύγκρουσης των σωματιδίων αυτών με τα μόρια του νερού (η ύπαρξη των οποίων ήταν ακόμα σε αμφισβήτηση). Το μέρος της θεωρίας που μας ενδιαφέρει είναι η διατύπωση μίας *εξίσωσης διάχυσης* (diffusion equation) που αφορά τις προσαυξήσεις των θέσεων των σωματιδίων στον χρόνο, η οποία τελικά θα μας δώσει την εξέλιξη της κατανομής των θέσεων των σωματιδίων ως συνάρτηση του χρόνου. Οι βασικές υποθέσεις του μοντέλου σε διάσταση 1 είναι οι παρακάτω:

Υποθέσεις Κίνησης Brown:

(α) η κίνηση είναι συνεχής,

(β) σε ξένα μεταξύ τους χρονικά διαστήματα, οι μεταβολές της θέσης των σωματιδίων είναι αμοιβαία ανεξάρτητες,

(γ) σε μικρό χρονικό διάστημα τ, το ποσοστό των σωματιδίων που θα υποστούν μία (προσημασμένη) μετατόπιση από δ μέχρι $\delta + d\delta$ θα είναι περίπου $\phi_\tau(\delta)\, d\delta$, όπου ϕ_τ είναι μία συνάρτηση πυκνότητας μίας τ.μ. Δ_τ με $\mathbb{E}(\Delta_\tau) = 0$ και $\mathbb{V}(\Delta_\tau) = 2D \cdot \tau$, όπου ο $D \geq 0$ καλείται συντελεστής διάχυσης.

Παρατήρηση 4.2. Περιγράφουμε εδώ μία ευρετική (heuristic) μέθοδο λύσης του προβλήματος. Παρακάτω, η συνάρτηση $f(x,t)$ αντιστοιχεί στη συνάρτηση πυκνότητας πιθανότητας της τ.μ. W_t. Κατ'αρχάς,

$$f(x, t+\tau) \simeq f(x,t) + \tau \cdot \frac{\partial f}{\partial t}(x,t). \tag{4.15}$$

Επιπλέον,

$$\begin{aligned} f(x,t+\tau) &= \int_{-\infty}^{+\infty} f(x-\delta,t)\, \phi_\tau(\delta)\, d\delta \\ &\simeq \int_{-\infty}^{+\infty} \left[f(x,t) - \delta \cdot \frac{\partial f}{\partial x}(x,t) + \frac{1}{2}\,\delta^2 \cdot \frac{\partial^2 f}{\partial x^2}(x,t) \right] \phi_\tau(\delta)\, d\delta \\ &= f(x,t) - \frac{\partial f}{\partial x}(x,t)\, \mathbb{E}(\Delta_\tau) + \frac{1}{2}\, \mathbb{E}(\Delta_\tau^2)\, \frac{\partial^2 f}{\partial x^2}(x,t) \\ &= f(x,t) + D \cdot \tau \cdot \frac{\partial^2 f}{\partial x^2}(x,t), \end{aligned} \tag{4.16}$$

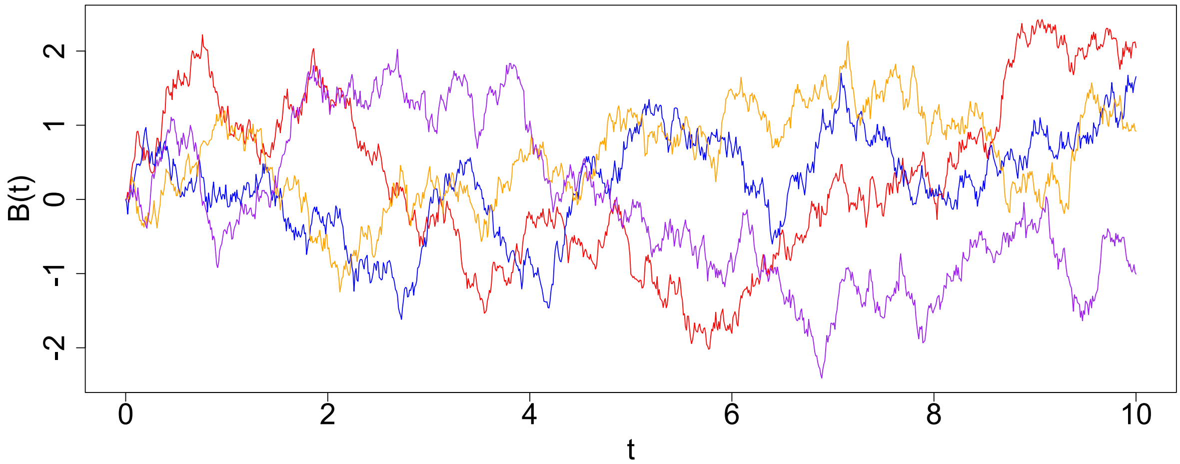
όπου στην τελευταία ισότητα χρησιμοποιήσαμε ότι $\mathbb{E}(\Delta_\tau) = 0$ και ότι $\mathbb{E}(\Delta_\tau^2) = \mathbb{V}(\Delta_\tau) = 2D \cdot \tau$. Εξισώνοντας τις (4.15) και (4.16) προκύπτει η εξίσωση θερμότητας (heat equation)

$$\frac{\partial f}{\partial t} = D\, \frac{\partial^2 f}{\partial x^2},$$

που είναι ειδική περίπτωση εξίσωσης διάχυσης (diffusion equation), με λύση

$$f(x,t) = \frac{1}{\sqrt{4\pi D t}}\, \exp\left\{ -\frac{x^2}{4Dt} \right\}, \qquad t \geq 0.$$

Αυτό οδηγεί σε τυχαίες μεταβλητές $W_t \sim \mathcal{N}(0,\, 2D\, t)$, $\forall t \geq 0$. Είναι αυτή η θεώρηση που οδήγησε σε περαιτέρω μελέτη της προκύπτουσας στοχαστικής διαδικασίας, η οποία και θεμελιώθηκε αυστηρά από τον *N. Wiener* (αλλά και από άλλους), όταν πλέον η θεωρία μέτρου είχε αναπτυχθεί αρκετά. Στο Σχήμα 4.1 παρουσιάζονται 4 μονοπάτια μίας κίνησης Brown.



Σχημα 4.1: 4 μονοπάτια μία κίνησης Brown.

Θεώρημα 4.6 (Κριτήριο συνεχείας του Kolmogorov): Έστω $T > 0$ και $X = (X_t)_{t\in[0,T]}$ μία στοχαστική διαδικασία ορισμένη στο χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$. Αν υποθέσουμε ότι υπάρχουν σταθερές $a, \beta, K > 0$ τέτοιες ώστε

$$\mathbb{E}\left|X_t - X_s\right|^a \leq K\left|t - s\right|^{1+\beta}, \tag{4.17}$$

για κάθε $s, t \in [0, T]$, τότε έχει μία τροποποίηση που είναι (δειγματικά) συνεχής.

Υπόδειξη: Η απόδειξη μπορεί να ακολουθήσει τα ακόλουθα 4 βήματα:

1. Δείξτε ότι από την (4.17) έπεται ότι η X είναι στοχαστικά συνεχής.
2. η X είναι με πιθ. 1 ομοιόμορφα συνεχής σε ένα αριθμήσιμο πυκνό υποσύνολο $D \subset [0, T]$.
3. Επεκτείνετε τη X από το D σε μία συνεχή σ.δ. Y σε όλο το $[0, T]$.
4. Δείξτε ότι η Y είναι καλά ορισμένη σ.δ. που αποτελεί συνεχή τροποποίηση της X.

Για ευκολία υποθέστε ότι $T = 1$ και $D = \bigcup\limits_{n=1}^{\infty} D_n$, όπου $D_n = \left\{\dfrac{k}{2^n} \,:\, k = 0, 1, \ldots, 2^n\right\}$.

Πόρισμα 4.1: Η κίνηση Brown υπάρχει.

Υπόδειξη: Θεωρήστε $X = (X_t)_{t \geq 0}$ μία κατάλληλη Γκαουσιανή στοχαστική διαδικασία. Δείξτε ότι πληροί το κριτήριο συνεχείας του Kolmogorov. Για κάθε $T > 0$ φτιάξτε κατάλληλες συνεχείς τροποποιήσεις X^T και επαληθεύστε ότι η $W_t = \sum\limits_{n=1}^{\infty} X_t^n \mathbf{1}_{[n-1,n)}(t)$ είναι κίνηση Brown.

Πρόταση 4.3: Μία συνεχής Γκαουσιανή διαδικασία W είναι κίνηση Brown αν, και μόνο αν, $\mu(t) = \mathbb{E}(W_t) = 0$ και $\Sigma(s,t) = \mathbb{C}\left(W_s, W_t\right) = s \wedge t$.

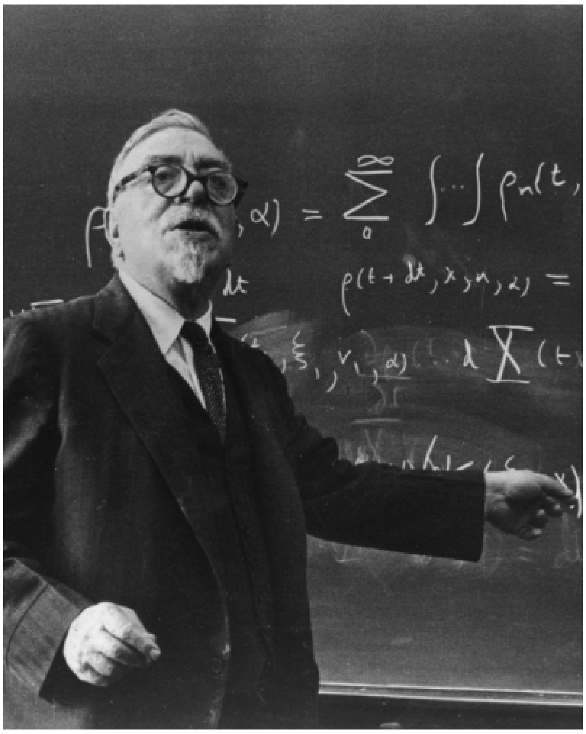
Η κίνηση Brown απολαμβάνει κάποιες πολύ χρήσιμες ιδιότητες.

Θεώρημα 4.7: Έστω W μία κίνηση Brown σε έναν χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$. Τότε, οι επόμενες στοχαστικές διαδικασίες είναι επίσης κινήσεις Brown:

1. *χρονική ομογένεια:* για κάθε $s \geq 0$ η μετατοπισμένη διαδικασία
$$W_{+s} = \big(W_{t+s} - W_s\big)_{t \geq 0},$$
2. *συμμετρία:* η στοχαστική διαδικασία $-W = (-W_t)_{t \geq 0}$,
3. *κλίμακας:* για κάθε $a > 0$ η στοχαστική διαδικασία W^a που ορίζεται από
$$W_t^a = a^{-1/2}\, W_{at},$$
4. *χρονική αντιστροφή:* η σ.δ. $X = (X_t)_{t \geq 0}$ με $X_0 = 0$ και $X_t = t\, W_{1/t}$, $t > 0$.

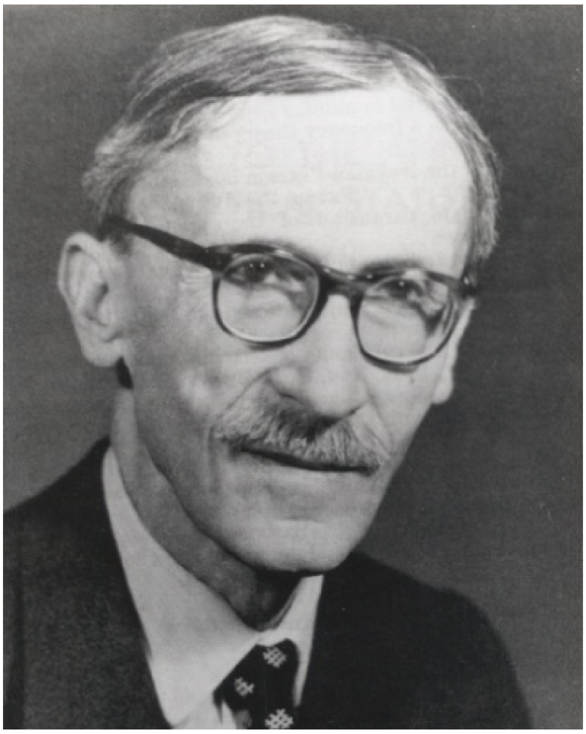
4.6 Ιστορικά στοιχεία - Περαιτέρω μελέτη

Σχημα 4.2: Percy John Daniell: 1889-1946, Αγγλο-Χιλιανός μαθηματικός, ανέπτυξε μία γενική θεωρία παραγώγισης και ολοκλήρωσης (*ολοκλήρωμα Daniell*).



Σχημα 4.3: Norbert Wiener, 1894-1964: Αμερικανός μαθηματικός, ασχολήθηκε με τις στοχαστικές διαδικασίες με στόχο κυρίως προβλήματα για ηλεκτρονικές επικοινωνίες και συστήματα ελέγχου. Το έργο του περιλαμβάνει, μεταξύ άλλων, το Φίλτρο Wiener, την εξίσωση Wiener, την άλγεβρα Wiener και τα θεωρήματα Paley-Wiener και Wiener-Khinchin.

Σχημα 4.4: Κωνσταντίνος Καραθεοδωρής, 1873-1950: κορυφαίος Έλληνας μαθηματικός με συνεισφορά επίσης στη Φυσική και (ακόμα και) στην Αρχαιολογία. Έπειτα από πιέσεις του πατέρα του, σπούδασε και εργάστηκε για μερικά χρόνια ως μηχανικός. Στα 27 του αποφάσισε να στραφεί στα μαθηματικά και σπούδασε για 2 χρόνια στο Πανεπιστήμιο του Βερολίνου. Εκεί παρακολούθησε μαθήματα από τους σπουδαίους μαθηματικούς Schwarz και Frobenius. Έπειτα πήγε στο Πανεπιστήμιο του Γκέτινγκεν, όπου συνάντησε τους κορυφαίους μαθηματικούς Hilbert και Klein, και έκανε τη διδακτορική του διατριβή υπό την επίβλεψη του Minkowski. Είχε τεράστια συνεισφορά στους τομείς της Πραγματικής Ανάλυσης, της Συναρτησιακής Ανάλυσης και της Θεωρίας Μέτρου.



Σχημα 4.5: Paul Pierre Lévy, 1886-1971: Γάλλος μαθηματικός με πολύ σημαντική συνεισφορά στη Θεωρία Πιθανοτήτων και Στοχαστικών διαδικασιών. Στον Α' Παγκόσμιο Πόλεμο εργάστηκε σε εφαρμογές των μαθηματικών για το γαλλικό πυροβολικό. Το 1948 έγραψε τη σπουδαία εργασία *Processus stochastiques et mouvement brownien*, στην οποία εισήγαγε πληθώρα νέων εννοιών. Είναι πολύ γνωστός για τις στοχαστικές διαδικασίες που φέρουν το όνομά του και τις ευσταθείς κατανομές (stable distributions).

ΚΕΦΑΛΑΙΟ 5

Ο ΕΛΕΓΧΟΣ χ^2

Ο έλεγχος χ^2 του Karl Pearson [12] είναι ο απλούστερος και πιο δημοφιλής έλεγχος καλής προσαρμογής. Μας επιτρέπει να ελέγξουμε με έναν πολύ απλό τρόπο αν μπορούμε να υποθέσουμε ότι μία δεδομένη κατανομή (ή μία οικογένεια κατανομών) βρίσκεται πίσω από την παραγωγή των δεδομένων ή, απλούστερα, αν μία υποθετική κατανομή προσαρμόζεται καλά στα δεδομένα. Για την πραγματοποίηση του ελέγχου, η συνήθης τακτική συνίσταται στην ομαδοποίηση (ή κατηγοριοποίηση) των τιμών ενός χαρακτηριστικού και στην απόδοση ενός μέτρου απόκλισης του πλήθους των παρατηρήσεων του δείγματος που θα βρεθούν μέσα στην κάθε ομάδα από την θεωρητικά αναμενόμενη αν η υποθετική κατανομή ήταν αυτή που παρήγαγε πραγματικά τα δεδομένα. Αυτό επιτυγχάνεται με τη βοήθεια του στατιστικού χ^2, όπως θα δούμε παρακάτω.

Στο κεφάλαιο αυτό παρουσιάζουμε τη χρήση του ελέγχου χ^2 σε διάφορες περιπτώσεις ενδιαφέροντος, που περιλαμβάνουν επίσης τη χρήση του ως έλεγχο ομοιογένειας ανεξάρτητων δειγμάτων, αλλά και ως έλεγχο ανεξαρτησίας δύο χαρακτηριστικών όταν διαθέτουμε ένα ανεξάρτητο τυχαίο δείγμα που αποτελείται από ζεύγη τιμών αυτών των δύο χαρακτηριστικών.

Σάμης Τρέβεζας (2023). «Μη Παραμετρική Στατιστική».
Αθήνα: Κάλλιπος, Ανοικτές Ακαδημαϊκές Εκδόσεις.
BY-NC-SA Creative Commons Αναφορά Δημιουργού - Μη Εμπορική Χρήση - Παρόμοια Διανομή 4.0

5.1 Καλή προσαρμογή μίας κατανομής χωρίς ομαδοποίηση

Ξεκινάμε αρχικά με την υπόθεση ότι διαθέτουμε ένα τυχαίο δείγμα $X_1, X_2, \ldots, X_n$ από την κατανομή ενός χαρακτηριστικού X που παίρνει d διακεκριμένες τιμές. Για απλότητα υποθέτουμε ότι $X \in \{1, 2, \ldots, d\}$ και οι τιμές αυτές είναι γενικά συμβολικές, χωρίς η αριθμητική τους αξία να παίζει κάποιον ιδιαίτερο ρόλο. Αν δεν κάνουμε κάποια υπόθεση παραμετρικής φύσης για την κατανομή της X, τότε

$$\mathbb{P}(X = i) = p_i > 0, \quad \forall i \in \{1, 2, \ldots, d\},$$

όπου το $p = (p_1, p_2, \ldots, p_d)$ είναι ένα γενικά άγνωστο διάνυσμα πιθανοτήτων, με μόνο περιορισμό ότι αθροίζει στη μονάδα, που μας επιτρέπει να κινούμαστε στο χώρο όλων των δυνατών κατανομών με στήριγμα το $\{1, 2, \ldots, d\}$. Συχνά λοιπόν, δεν ενδιαφερόμαστε να εκτιμήσουμε την παράμετρο, αλλά εστιάζουμε στην ισχύ μιας συγκεκριμένης υπόθεσης που αφορά την κατανομή της X. Η απλούστερη περίπτωση αντιστοιχεί στην υπόθεση ότι $X \sim F_0$ ή με όρους συναρτήσεων κατανομής $F = F_0$. Εφόσον η τιμή του p καθορίζει πλήρως και μονοσήμαντα την κατανομή, αυτό είναι ισοδύναμο με την παραμετρική υπόθεση $p = p_0$. Έτσι φτάνουμε στη διατύπωση του ελέγχου υποθέσεων που αντιστοιχεί στον έλεγχο καλής προσαρμογής:

$$H_0 : p = p_0 \quad \text{vs} \quad H_1 : p \neq p_0.$$

Παρατηρούμε εδώ ότι η εναλλακτική υπόθεση περιλαμβάνει όλες τις δυνατές κατανομές, εκτός από αυτήν που καθορίζεται από τη μηδενική υπόθεση. Πώς όμως μπορούμε να πραγματοποιήσουμε αυτόν τον έλεγχο; Ποια είναι η κατάλληλη ελεγχοσυνάρτηση και πώς θα βρούμε την κατανομή της; Η απάντηση στο τελευταίο ερώτημα έχει εξήγηση ασυμπτωτικού χαρακτήρα και θα δοθεί αμέσως μετά τον επόμενο ορισμό, που καθορίζει τη μορφή της ελεγχοσυνάρτησης. Ο συμβολισμός που χρησιμοποιούμε εδώ είναι και ο πιο συνηθισμένος στη βιβλιογραφία.

Ορισμός 5.1: Έστω $X_1, X_2, \ldots, X_n$ ένα τυχαίο δείγμα από την κατανομή ενός χαρακτηριστικού $X \in \{1, 2, \ldots, d\}$. Θέτουμε O_i (O: Observed) το πλήθος των παρατηρήσεων του δείγματος που έχουν την τιμή i και E_i (E: Expected) το αντίστοιχο αναμενόμενο πλήθος για $i = 1, 2, \ldots, d$ κάτω από τη μηδενική υπόθεση H_0. Η στατιστική συνάρτηση

$$Q_n = \sum_{i=1}^{d} \frac{(O_i - E_i)^2}{E_i}, \tag{5.1}$$

λέγεται *στατιστικό* χ^2 και αντιστοιχεί στην ελεγχοσυνάρτηση του (ασυμπτωτικού) ελέγχου χ^2 καλής προσαρμογής.

Παρατήρηση 5.1. Παρατηρούμε ότι στις ποσότητες που εμπλέκονται στον υπολογισμό του Q_n, δηλαδή τα O_i και E_i, δεν εμφανίζεται η εξάρτηση στο μέγεθος του δείγματος n. Αυτό είναι καθαρά για λόγους απλοποίησης του συμβολισμού και δεν θα πρέπει να ξεχνάμε ότι η εξάρτηση αυτή υπάρχει. Επιπλέον, είναι σαφές ότι μπορούμε να θυμόμαστε ότι το στατιστικό χ^2 εκφράζεται μέσα από αθροίσματα κατάλληλα σταθμισμένων τετραγωνικών αποκλίσεων των παρατηρούμενων από τις αναμενόμενες τιμές. Μεγάλες τιμές του στατιστικού θα οδηγούν σε απόρριψη της μηδενικής υπόθεσης, υποδηλώνοντας κακή προσαρμογή της υποθετικής κατανομής στα δεδομένα. Η διαδικασία αυτή θα καθοριστεί αυστηρά από την ασυμπτωτική κατανομή του Q_n και το επίπεδο στατιστικής σημαντικότητας (ε.σ.σ.).

Θα υπολογίσουμε τώρα την ασυμπτωτική κατανομή του Q_n. Θέτουμε

$$S_n = (S_{n,1}\ S_{n,2}\ \cdots\ S_{n,d})^\top,$$

όπου

$$S_{n,i} = O_i = \sum_{k=1}^{n} \mathbf{1}_{\{X_k=i\}},$$

αντιστοιχεί στο πλήθος εμφανίσεων του i στο δείγμα. Έχουμε δείξει ότι αν p είναι η πραγματική τιμή της παραμέτρου, τότε $S_n \sim \mathcal{M}_d(n,p)$ και ιδιαιτέρως $S_{n,i} \sim \mathrm{Bin}(n,p_i)$. Από αυτό συμπεραίνουμε ότι $E_i = \mathbb{E}(S_{n,i}) = np_i$. Το επόμενο αποτέλεσμα προκύπτει άμεσα από μία εφαρμογή του πολυδιάστατου Κ.Ο.Θ. και του θεωρήματος συνεχούς απεικόνισης.

Πρόταση 5.1: Αν p είναι η πραγματική τιμή της παραμέτρου, τότε

$$\left(\sqrt{\mathrm{dg}(p)}\right)^{-1} \frac{S_n - n\,p}{\sqrt{n}} \xrightarrow{d} \mathcal{N}_d\Big(0_d,\, I_d - \sqrt{p}\big(\sqrt{p}\big)^\top\Big), \tag{5.2}$$

όπου $\sqrt{\mathrm{dg}(p)} = \begin{pmatrix} \sqrt{p_1} & 0 & \cdots & 0 \\ 0 & \sqrt{p_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{p_d} \end{pmatrix}$ και $\sqrt{p} = \begin{pmatrix} \sqrt{p_1} \\ \sqrt{p_2} \\ \vdots \\ \sqrt{p_d} \end{pmatrix}$.

Απόδειξη: Έχουμε δείξει από το πολυδιάστατο Κ.Ο.Θ. ότι

$$\frac{S_n - n\,p}{\sqrt{n}} \xrightarrow{d} \mathcal{N}_d\Big(0_d,\, \mathrm{dg}(p) - pp^\top\Big),$$

άρα από το Θ.Σ.Α.

$$\left(\sqrt{\mathrm{dg}(p)}\right)^{-1} \frac{S_n - n\,p}{\sqrt{n}} \xrightarrow{d} \left(\sqrt{\mathrm{dg}(p)}\right)^{-1} \mathcal{N}_d\Big(0_d,\, \mathrm{dg}(p) - pp^\top\Big).$$

Όμως ο $\mathrm{dg}(p)$ είναι συμμετρικός, και άρα

$$\left(\sqrt{\mathrm{dg}(p)}\right)^{-1} \mathcal{N}_d\Big(0_d,\, \mathrm{dg}(p) - pp^\top\Big) = \mathcal{N}_d\left(0_d,\, \left(\sqrt{\mathrm{dg}(p)}\right)^{-1} \left(\mathrm{dg}(p) - pp^\top\right) \left(\sqrt{\mathrm{dg}(p)}\right)^{-1}\right).$$

Τέλος, παρατηρούμε ότι

$$p=\begin{pmatrix}p_1\\ \vdots\\ p_d\end{pmatrix}=\begin{pmatrix}\sqrt{p_1}\cdot\sqrt{p_1}\\ \vdots\\ \sqrt{p_d}\cdot\sqrt{p_d}\end{pmatrix}=\begin{pmatrix}\sqrt{p_1} & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \sqrt{p_d}\end{pmatrix}\begin{pmatrix}\sqrt{p_1}\\ \vdots\\ \sqrt{p_d}\end{pmatrix}=\sqrt{\mathrm{dg}(p)}\sqrt{p},$$

και άρα

$$\left(\sqrt{\mathrm{dg}(p)}\right)^{-1}\left(\mathrm{dg}(p)-\left(\sqrt{\mathrm{dg}(p)}\sqrt{p}\right)\left(\sqrt{\mathrm{dg}(p)}\sqrt{p}\right)^{\top}\right)\left(\sqrt{\mathrm{dg}(p)}\right)^{-1}=I_d-\sqrt{p}\left(\sqrt{p}\right)^{\top}.$$

□

Είμαστε τώρα σε θέση να προσδιορίσουμε την ασυμπτωτική κατανομή του στατιστικού χ^2 υπό τη μηδενική υπόθεση.

Πόρισμα 5.1: Αν ισχύει η $H_0 : p = p_0$, τότε

$$Q_n \xrightarrow[n\to\infty]{d} Q \sim \chi^2_{d-1}.$$

Συμπερασματικά, αν q_n είναι η παρατηρούμενη τιμή της ελεγχοσυνάρτησης Q_n, τότε απορρίπτουμε τη μηδενική υπόθεση H_0 σε (ασυμπτωτικό) ε.σ.σ. α αν, και μόνο αν, $q_n > \chi^2_{d-1;\alpha}$ ή, ισοδύναμα, αν p-value $= \mathbb{P}(Q_n \geqslant q_n) \approx \mathbb{P}(Q \geqslant q_n) < \alpha$.

Απόδειξη: Έχουμε ότι $Q_n = \sum_{i=1}^{d}\left(\frac{S_{n,i}-n\,p_{0,i}}{\sqrt{n}\sqrt{p_{0,i}}}\right)^2 = \sum_{i=1}^{d} X_{n,i}^2 = \|X_n\|^2$, όπου

$$X_n=\left(\sqrt{\mathrm{dg}(p_0)}\right)^{-1}\frac{S_n-n\,p_0}{\sqrt{n}} \xrightarrow{d} X\sim\mathcal{N}_d\Big(0_d,\,I_d-\underbrace{\sqrt{p_0}\left(\sqrt{p_0}\right)^{\top}}_{\Pi_1=B_1B_1^{\top}}\Big)=(I_d-\Pi_1)\,Z,$$

όπου $Z\sim\mathcal{N}_d\big(0_d,\,I_d\big)$ και Π_1 είναι ο πίνακας ορθογώνιας προβολής στον διανυσματικό υπόχωρο $V_1=\langle\sqrt{p_0}\rangle$ (διάστασης 1) του $\mathbb{R}^d$. Συμπεραίνουμε ότι ο $I_d-\Pi_1$ αντιστοιχεί στον πίνακα ορθογώνιας προβολής στον ορθογώνιο υπόχωρο $V_1^{\perp}$ του V_1, που έχει διάσταση $d-1$ και άρα από το ΘΣΑ για τη συνεχή συνάρτηση $\|\cdot\|^2:\mathbb{R}^d\to\mathbb{R}$ και το θεώρημα Cochran για $\mu=0_n$ και $\sigma^2=1$ έχουμε

$$Q_n=\|X_n\|^2 \xrightarrow{d} \|X\|^2 =: Q\sim\chi^2_{d-1}.$$

□

Η πιο άμεση εφαρμογή αφορά τον έλεγχο ομοιομορφίας μιας διακριτής κατανομής σε ένα πεπερασμένο σύνολο δυνατών παρατηρήσεων, για απλότητα το $\{1,2,\ldots,d\}$. Τα παρακάτω ερωτήματα κινούνται σε αυτήν την κατεύθυνση.

- Είναι ένα νόμισμα δίκαιο ή έχουμε ενδείξεις ότι είναι κίβδηλο ($d = 2$);
- Είναι ένα ζάρι αμερόληπτο ($d = 6$) ή έχουμε ενδείξεις για το αντίθετο;
- Παράγει ένας γεννήτορας ψευδοτυχαίων αριθμών μία σειρά από δυαδικά ψηφία με εντελώς τυχαίο τρόπο ή έχουμε ενδείξεις ότι ευνοεί κάποιο μοτίβο ($d = 2^k$, για κάποιο k);

Παρακάτω δίνουμε δύο παραδείγματα εφαρμογής της μεθόδου.

Παράδειγμα 5.1: Θέλουμε να εξετάσουμε σε ε.σ.σ. $\alpha = 0.05$ αν ένα ζάρι είναι αμερόληπτο. Ελέγξτε τον ισχυρισμό αυτόν:

1. όταν σε 120 ανεξάρτητες ρίψεις ενός ζαριού καταγράψαμε τις εξής συχνότητες εμφάνισης των αποτελεσμάτων 1 έως 6: 10, 10, 15, 15, 30, 40,
2. για τις διαφορετικές δυνατές τιμές του x, όταν σε 60 ανεξάρτητες ρίψεις ενός ζαριού οι συχνότητες εμφάνισης των αποτελεσμάτων 1 έως 6 είναι : x, $20 - x$, 10, 10, x, $20 - x$.

Λύση.
(i): Έστω X η τυχαία μεταβλητή που περιγράφει το αποτέλεσμα μίας ρίψης του ζαριού. Υπό τη μηδενική υπόθεση της αμεροληψίας του ζαριού, δηλαδή την $H_0 : X \sim \mathcal{U}\{1, 2, \ldots, 6\}$, γνωρίζουμε ότι $\mathbb{P}(X = i) = \frac{1}{6}$ για $i = 1, 2, \ldots, 6$. Κατασκευάζουμε τον πίνακα συχνοτήτων των δεδομένων:

Πινακας 5.1: Δεδομένα Παραδείγματος 5.1–(i)

i	1	2	3	4	5	6	Άθροισμα
O_i	10	10	15	15	30	40	**120**
$p_{i,0}$	1/6	1/6	1/6	1/6	1/6	1/6	**1**
E_i	20	20	20	20	20	20	**120**

Σχετικά με την ελεγχοσυνάρτηση του ελέγχου χ^2, γνωρίζουμε ότι:

$$Q_n = \sum_{i=1}^{6} \frac{(O_i - E_i)^2}{E_i} \xrightarrow{d} Q \sim \chi^2_5.$$

Υπολογίζοντας την παρατηρούμενη τιμή της ελεγχοσυνάρτησης, παίρνουμε ότι

$$q_n = 37.5 > 11.071 = \chi^2_{5;0.05} \quad \text{και} \quad \text{p-value} = \mathbb{P}(Q_n \geqslant q_n) \approx \mathbb{P}(Q \geqslant q_n) = 4.755 \cdot 10^{-7} < 0.05.$$

Επομένως, απορρίπτουμε την H_0 σε ε.σ.σ. $\alpha = 0.05$. Η επαλήθευση των παραπάνω αποτελεσμάτων υλοποιείται ως εξής:

```r
O = c(10, 10, 15, 15, 30, 40)
a = 0.05 ; df= length(O)-1
E = rep(20, times=6)
```

```
q = sum((O-E)^2/E)
qa = qchisq(a,df,lower.tail=FALSE)
p = pchisq(q ,df,lower.tail=FALSE)
c(q ,qa, p)
```

Η απευθείας πραγματοποίηση του παραπάνω ελέγχου είναι επίσης εφικτή:

```
O = c(10, 10, 15, 15, 30, 40)
> chisq.test(O)
Chi-squared test for given probabilities
data:  O
X-squared = 37.5, df = 5, p-value = 4.754e-07
```

(ii) Κατασκευάζουμε τον πίνακα συχνοτήτων των δεδομένων:

Πινακας 5.2: Δεδομένα Παραδείγματος 5.1–(ii)

i	1	2	3	4	5	6	Άθροισμα
O_i	x	$20-x$	10	10	x	$20-x$	60
$p_{i,0}$	1/6	1/6	1/6	1/6	1/6	1/6	1
E_i	10	10	10	10	10	10	60

Σχετικά με την ελεγχοσυνάρτηση του ελέγχου χ^2, γνωρίζουμε ότι:

$$Q_n = \sum_{j=1}^{6} \frac{(O_i - E_i)^2}{E_i} \xrightarrow{d} Q \sim \chi^2_5.$$

Υπολογίζοντας την παρατηρούμενη τιμή της ελεγχοσυνάρτησης, παίρνουμε ότι:

$$q_n = 2 \cdot \frac{(x-10)^2}{10} + 2 \cdot \frac{(20-x-10)^2}{10} = \frac{4(x-10)^2}{10} > \chi^2_{5;0.05} \Leftrightarrow$$

$$|x-10| > \frac{1}{2}\sqrt{10\chi^2_{5;0.05}} = 5.261 \Leftrightarrow x > 15.261 \text{ ή } x < 4.739.$$

Συναληθεύοντας με τον προφανή περιορισμό $0 \le x \le 20$, συμπεραίνουμε ότι απορρίπτουμε την H_0 σε ε.σ.σ. $\alpha = 0.05$ αν και μόνο αν $x \in \{0, 1, 2, 3, 4\}$ ή $x \in \{16, 17, 18, 19, 20\}$.

5.2 Καλή προσαρμογή μίας οικογένειας κατανομών

Στην προηγούμενη ενότητα πραγματοποιήσαμε έναν έλεγχο μίας συγκεκριμένης υποψήφιας κατανομής έναντι όλων των άλλων πιθανών κατανομών που αντιστοιχούν σε ένα τυχαίο δείγμα $X_1, X_2, \ldots, X_n$ πάνω στο σύνολο των δυνατών τιμών τους $\{1, 2, \ldots, d\}$. Θα μπορούσε βέβαια ο έλεγχος να επεκταθεί από μία μόνο κατανομή σε μία οικογένεια κατανομών, κάνοντας κάποιες υποθέσεις παραμετρικής φύσης. Σε αυτήν την περίπτωση βέβαια είμαστε υποχρεωμένοι να εκτιμήσουμε την άγνωστη παράμετρο πριν πραγματοποιήσουμε τον έλεγχο. Επειδή ο έλεγχος γίνεται στην ουσία για μία ολόκληρη οικογένεια κατανομών, αυτό επιφέρει αλλαγές στην κατανομή της ελεγχοσυνάρτησης. Διαισθητικά, για το ίδιο ε.σ.σ., πρέπει να ελαττώνεται η κρίσιμη τιμή, αφού τώρα δεν περιλαμβάνει μόνο μία κατανομή, αλλά μία ολόκληρη οικογένεια κατανομών, επιφέροντας συνεπακόλουθα καλύτερη προσαρμογή του μοντέλου στα δεδομένα και, άρα, ελάττωση των τιμών της ελεγχοσυνάρτησης. Έτσι λοιπόν αποδεικνύεται ότι η κατανομή της ελεγχοσυνάρτησης χ^2 χάνει σε τυπικές περιπτώσεις ασυμπτωτικά τόσους βαθμούς ελευθερίας όση είναι και η διάσταση της παραμετρικής οικογένειας την οποία ελέγχουμε υπό την H_0, ελαττώνοντας την κρίσιμη τιμή του ελέγχου.

Ορισμός 5.2 (έλεγχος χ^2 για οικογένεια κατανομών):
Έστω $p = (p_1, \ldots, p_d)$, $\theta = (\theta_1, \ldots, \theta_{d_0}) \in \Theta \subset \mathbb{R}^{d_0}$, $d_0 < d$, και $f\colon \Theta \to \mathbb{R}^d$ μία γνωστή συνάρτηση. Ένας *έλεγχος χ^2 για οικογένεια κατανομών* είναι ένας έλεγχος της μορφής:

$$H_0 : p = f(\theta), \text{ για κάποιο } \theta \in \Theta, \quad \text{vs} \quad H_1 : p \neq f(\theta), \text{ για κάθε } \theta \in \Theta,$$

όπου

$$p = f(\theta) \iff p_i = f_i(\theta), \text{ για κάθε } i \in \{1, \ldots, d\}.$$

Αν

$$Q_n(\theta) = \sum_{i=1}^{d} \frac{\big(S_{n,i} - nf_i(\theta)\big)^2}{nf_i(\theta)}, \tag{5.3}$$

είναι η ελεγχοσυνάρτηση του ελέγχου χ^2 που αντιστοιχεί σε ένα δεδομένο $\theta \in \Theta$ και $\widehat{\theta}_n$ είναι μία εκτιμήτρια του $\theta \in \Theta$, τότε η $Q_n(\widehat{\theta}_n)$ είναι μία ελεγχοσυνάρτηση του παραπάνω ελέγχου.

Στο παρακάτω θεώρημα δίνεται η ασυμπτωτική κατανομή της ελεγχοσυνάρτησης χ^2 σε έναν τέτοιο έλεγχο.

Θεώρημα 5.1: Αν ισχύει η $H_0 : p = f(\theta)$ για κάποιο $\theta \in \Theta \subset \mathbb{R}^{d_0}$ και $\widehat{\theta}_n$ είναι μία αποδοτική εκτιμήτρια του θ, τότε υπό κάποιες συνθήκες ομαλότητας (βλ. Lehmann & Romano, Testing Statistical Hypothesis, Κεφ. 14)

$$Q_n(\widehat{\theta}_n) \xrightarrow[n\to\infty]{d} \chi^2_{d-d_0-1}. \tag{5.4}$$

Παρατήρηση 5.2. Όπως και στην περίπτωση της απλής μηδενικής υπόθεσης, έτσι και εδώ με σύνθετη μηδενική, μπορούμε να απλοποιήσουμε την έκφραση του στατιστικού, θέτοντας $O_i = S_{n,i}$ και $\widehat{E}_i = nf_i(\widehat{\theta}_n)$ για την εκτιμώμενη μέση τιμή. Έτσι, μπορούμε να γράφουμε για απλότητα

$$Q_n = \sum_{i=1}^{d} \frac{(O_i - \widehat{E}_i)^2}{\widehat{E}_i}, \tag{5.5}$$

όπου $\widehat{E}_i$ αντιστοιχεί τώρα στην εκτιμώμενη μέση τιμή.

Ας δούμε ένα παράδειγμα σε αυτήν την κατεύθυνση.

Παράδειγμα 5.2: Όπως στο Παράδειγμα 5.1, υποθέτουμε ότι προέκυψαν οι εξής συχνότητες εμφάνισης των αποτελεσμάτων 1 έως 6: 10, 10, 15, 15, 30, 40, σε 120 ανεξάρτητες ρίψεις ενός ζαριού. Αν X είναι η τυχαία μεταβλητή που περιγράφει το αποτέλεσμα μίας ρίψης του ζαριού, τότε ένας ερευνητής υποστηρίζει ότι $\mathbb{P}(X = 1) = \mathbb{P}(X = 2) = p$, $\mathbb{P}(X = 3) = \mathbb{P}(X = 4) = 2p$, $\mathbb{P}(X = 5) = \mathbb{P}(X = 6) = 0.5 - 3p$. Να ελεγχθεί σε ε.σ.σ. $\alpha = 0.05$ αν ισχύει ο ισχυρισμός του ερευνητή.

Λύση.

Η μηδενική υπόθεση είναι σύνθετη και επιλέγουμε ως εκτιμήτρια του p την εκτιμήτρια μέγιστης πιθανοφάνειας που αντιστοιχεί στο τυχαίο δείγμα $X_1, X_2, \ldots, X_n$. Υπολογίζουμε τη συνάρτηση πιθανοφάνειας των δεδομένων υπό τον ισχυρισμό του ερευνητή, δηλαδή αν ισχύει το υπομοντέλο που καθορίζεται από τη μηδενική υπόθεση. Εδώ, η συνάρτηση πιθανότητας που καθορίζεται από την εκφώνηση είναι:

$$f(x;p) = p^{\delta_{1x}+\delta_{2x}}(2p)^{\delta_{3x}+\delta_{4x}}(0.5-3p)^{\delta_{5x}+\delta_{6x}}, \quad x \in \{1,2,\ldots,6\},$$

όπου δ_{ij} είναι το δέλτα του Kronecker, και έτσι η συνάρτηση πιθανοφάνειας διαμορφώνεται ως

$$L(p) = \prod_{k=1}^{120} f(X_k;p) = p^{S_{n,1}+S_{n,2}}(2p)^{S_{n,3}+S_{n,4}}(0.5-3p)^{S_{n,5}+S_{n,6}}, \quad p \in [0, 1/6],$$

όπου το πεδίο μεταβολής του p καθορίζεται σύμφωνα με τους περιορισμούς που θέτει η μηδενική υπόθεση. Στο συγκεκριμένο δείγμα λοιπόν παίρνουμε

$$L(p) = p^{20} \cdot (2p)^{30} \cdot (0.5-3p)^{70} = 2^{30} \cdot p^{50} \cdot (0.5-3p)^{70}.$$

Από το παραπάνω προκύπτει ότι η συνάρτηση λογαριθμοπιθανοφάνειας

$$\ell(p) = 30\log 2 + 50\log p + 70\log(0.5-3p) \Rightarrow$$

$$\ell^{'}(p) = \frac{50}{p} - \frac{210}{0.5-3p}.$$

Λύνοντας την εξίσωση πιθανοφάνειας έχουμε

$$\ell'(p) = 0 \quad \Leftrightarrow \quad 25 - 150\,p = 210\,p \ \Rightarrow\ p^* = \frac{5}{72} < \frac{1}{6}, \quad \text{όπου:}$$

$$\ell''(p) = -\frac{50}{p^2} - \frac{630}{(0.5-3p)^2} < 0, \quad \forall p \in (0,1).$$

Δηλαδή, η $L(p)$ είναι γνησίως κοίλη συνάρτηση στο $(0,1)$ και, άρα, το μοναδικό στάσιμο σημείο της είναι μοναδικό ολικό μέγιστο. Συμπεραίνουμε λοιπόν ότι $\widehat{p} = p^*$. Για την πραγματοποίηση του ελέγχου κατασκευάζουμε τον πίνακα συχνοτήτων των δεδομένων:

Πινακας 5.3: Δεδομένα Παραδείγματος 5.1–(i)

i	1	2	3	4	5	6	Άθροισμα
O_i	10	10	15	15	30	40	120
$p_{i,0}$	$\frac{5}{72}$	$\frac{5}{72}$	$\frac{10}{72}$	$\frac{10}{72}$	$\frac{21}{72}$	$\frac{21}{72}$	1
E_i	$\frac{25}{3}$	$\frac{25}{3}$	$\frac{50}{3}$	$\frac{50}{3}$	35	35	120

Εφόσον έχουμε εκτιμήσει μία άγνωστη παράμετρο p και άρα $d_0 = 1$, παίρνουμε ότι:

$$Q_n = \sum_{j=1}^{6} \frac{\left(O_i - \widehat{E}_i\right)^2}{\widehat{E}_i} \xrightarrow{d} Q \sim \chi^2_4.$$

Υπολογίζοντας την παρατηρούμενη τιμή της ελεγχοσυνάρτησης, παίρνουμε ότι

$$q_n = 2.429 < 9.488 = \chi^2_{4;0.05} \quad \text{και} \quad \text{p-value} = \mathbb{P}(Q_n \geqslant q_n) \approx \mathbb{P}(Q \geqslant q_n) = 0.658 > 0.05.$$

Επομένως, δεν απορρίπτουμε τον ισχυρισμό του ερευνητή σε ε.σ.σ. $\alpha = 0.05$.

Παράδειγμα 5.3: Στα πλαίσια μίας σύντομης έρευνας, ζητήθηκε από 246 φοιτητές να επιλέξουν τυχαία έναν ακέραιο αριθμό από το 0 μέχρι το 9. Οι απόλυτες συχνότητες των απαντήσεων ήταν 11, 9, 17, 44, 27, 30, 31, 44, 8 και 25, αντίστοιχα. Θέλουμε να εξετάσουμε σε ε.σ.σ. $a = 0.05$ αν η επιλογή του αριθμού είναι όντως τυχαία.

Λύση.

Έστω X η τυχαία μεταβλητή που περιγράφει τον αριθμό που επιλέχθηκε. Υπό την μηδενική υπόθεση της τυχαιότητας έχουμε $X \sim \mathcal{U}\{0,1,\ldots,9\}$, δηλαδή $\mathbb{P}(X = i) = 0.1$ για $i = 0,1,\ldots,9$. Κατασκευάζουμε τον πίνακα συχνοτήτων των δεδομένων:

Πινακας 5.4: Δεδομένα Παραδείγματος 5.3

i	0	1	2	3	4	5	6	7	8	9	Άθροισμα
O_i	11	9	17	44	27	30	31	44	8	25	246
$p_{i,0}$	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1
E_i	24.6	24.6	24.6	24.6	24.6	24.6	24.6	24.6	24.6	24.6	246

Σχετικά με την ελεγχοσυνάρτηση του ελέγχου χ^2, γνωρίζουμε ότι:

$$Q_n = \sum_{i=0}^{9} \frac{(O_i - E_i)^2}{E_i} \xrightarrow{d} Q \sim \chi_9^2.$$

Υπολογίζοντας την παρατηρούμενη τιμή της ελεγχοσυνάρτησης, παίρνουμε ότι

$$q_n = 64.65 > 16.91 = \chi^2_{9;0.05} \quad \text{και} \quad \text{p-value} = \mathbb{P}(Q_n \geqslant q_n) \approx \mathbb{P}(Q \geqslant q_n) = 1.687 \cdot 10^{-10} < 0.05.$$

Επομένως, απορρίπτουμε την H_0 σε ε.σ.σ. $\alpha = 0.05$.

Παράδειγμα 5.4: Στα πλαίσια μίας σύντομης έρευνας, ζητήθηκε από 246 φοιτητές να επιλέξουν τον αγαπημένο τους ακέραιο αριθμό από το 0 μέχρι το 9. Οι απόλυτες συχνότητες των απαντήσεων ήταν 10, 14, 30, 21, 16, 25, 17, 67, 20 και 26, αντίστοιχα. Θέλουμε να εξετάσουμε σε ε.σ.σ. $a = 0.05$ αν η επιλογή του αριθμού είναι τυχαία.
Λύση.
Έστω X η τυχαία μεταβλητή που περιγράφει τον αριθμό που επιλέχθηκε. Υπό την μηδενική υπόθεση της τυχαιότητας έχουμε $X \sim \mathcal{U}\{0, 1, \ldots, 9\}$, δηλαδή $\mathbb{P}(X = i) = 0.1$ για $i = 0, 1, \ldots, 9$. Κατασκευάζουμε τον πίνακα συχνοτήτων των δεδομένων:

Πινακας 5.5: Δεδομένα Παραδείγματος 5.4

i	0	1	2	3	4	5	6	7	8	9	Άθροισμα
O_i	10	14	30	21	16	25	17	67	20	26	246
$p_{i,0}$	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1
E_i	24.6	24.6	24.6	24.6	24.6	24.6	24.6	24.6	24.6	24.6	246

Σχετικά με την ελεγχοσυνάρτηση του ελέγχου χ^2, γνωρίζουμε ότι:

$$Q_n = \sum_{i=0}^{9} \frac{(O_i - E_i)^2}{E_i} \xrightarrow{d} Q \sim \chi_9^2.$$

Υπολογίζοντας την παρατηρούμενη τιμή της ελεγχοσυνάρτησης, παίρνουμε ότι

$$q_n = 94.33 > 16.91 = \chi^2_{9;0.05} \quad \text{και} \quad \text{p-value} = \mathbb{P}(Q_n \geqslant q_n) \approx \mathbb{P}(Q \geqslant q_n) = 2.199{\cdot}10^{-16} < 0.05.$$

Επομένως, απορρίπτουμε την H_0 σε ε.σ.σ. $\alpha = 0.05$.

Ενδιαφέρον παρουσιάζει η υψηλή συχνότητα επιλογής του αριθμού 7. Επαναλαμβάνουμε τον έλεγχο ανάμεσα στους υπόλοιπους 9 αριθμούς. Υπολογίζοντας την παρατηρούμενη τιμή της ελεγχοσυνάρτησης, παίρνουμε ότι

$$q_n = 16.23 > 15.51 = \chi^2_{8;0.05} \quad \text{και} \quad \text{p-value} = \mathbb{P}(Q_n \geqslant q_n) \approx \mathbb{P}(Q \geqslant q_n) = 0.039 < 0.05.$$

Επομένως, απορρίπτουμε την H_0 σε ε.σ.σ. $\alpha = 0.05$. Παρότι η τελική απάντηση παραμένει ίδια, είναι σημαντικό να παρατηρήσουμε ότι η βεβαιότητα με την οποία απορρίπτουμε την αρχική υπόθεση διαφέρει σημαντικά στους δύο ελέγχους. Για μικρότερη τιμή του α (π.χ. $\alpha = 0.01$), στον πρώτο έλεγχο θα απορρίπταμε, ενώ στον δεύτερο θα δεχόμασταν την H_0.

Παράδειγμα 5.5: Έστω X η τυχαία μεταβλητή που περιγράφει το αποτέλεσμα ενός πειράματος με 4 δυνατά αποτελέσματα. Χρησιμοποιώντας ένα δείγμα μεγέθους $n = 1000$ με απόλυτες συχνότητες 206, 503, 156 και 135, αντίστοιχα, θέλουμε να ελέγξουμε την υπόθεση $X \sim \mathcal{C}\boldsymbol{at}(0.2, 0.5, 0.15, 0.15)$.

Λύση.

Έστω X η τυχαία μεταβλητή που περιγράφει τον αριθμό που επιλέχθηκε. Υπό την μηδενική υπόθεση της τυχαιότητας έχουμε $X \sim \mathcal{C}\boldsymbol{at}(0.2, 0.5, 0.15, 0.15)$. Κατασκευάζουμε τον πίνακα συχνοτήτων των δεδομένων:

Πινακας 5.6: Δεδομένα Παραδείγματος 5.5

i	1	2	3	4	Άθροισμα
O_i	206	503	156	135	1000
$p_{i,0}$	0.2	0.5	0.15	0.15	1
E_i	200	500	150	150	1000

Σχετικά με την ελεγχοσυνάρτηση του ελέγχου χ^2, γνωρίζουμε ότι:

$$Q_n = \sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i} \xrightarrow{d} Q \sim \chi^2_3.$$

Υπολογίζοντας την παρατηρούμενη τιμή της ελεγχοσυνάρτησης, παίρνουμε ότι

$$q_n = 1.94 < 7.81 = \chi^2_{3;0.05} \quad \text{και} \quad \text{p-value} = \mathbb{P}(Q_n \geqslant q_n) \approx \mathbb{P}(Q \geqslant q_n) = 0.59 > 0.05.$$

Επομένως, δεν απορρίπτουμε την H_0 σε ε.σ.σ. $\alpha = 0.05$.

Παράδειγμα 5.6: Έχουμε στη διάθεσή μας 2 νομίσματα. Θέλουμε να εξετάσουμε σε ε.σ.σ. $\alpha = 0.1$ αν κάθε νόμισμα είναι αμερόληπτο. Ελέγξτε τον ισχυρισμό αυτό όταν σε 30 ανεξάρτητες ρίψεις:

1. το νόμισμα 1 έφερε 18 κορώνες και 12 γράμματα,
2. το νόμισμα 2 έφερε 10 κορώνες και 20 γράμματα.

Λύση.
Έστω $X \in \{0, 1\}$ η τυχαία μεταβλητή που περιγράφει το αποτέλεσμα μίας ρίψης του νομίσματος. Θέτουμε $\{X = 1\}$ και $\{X = 0\}$ τα ενδεχόμενα εμφάνισης κορώνας και γραμμάτων αντίστοιχα. Υπό τη μηδενική υπόθεση της αμεροληψίας έχουμε ότι αυτά είναι ισοπίθανα.

(i): Κατασκευάζουμε τον πίνακα συχνοτήτων των δεδομένων για το νόμισμα 1:

Πινακας 5.7: Δεδομένα Παραδείγματος 5.6 – (i)

i	0	1	Άθροισμα
O_i	12	18	30
$p_{i,0}$	1/2	1/2	1
E_i	15	15	30

Σχετικά με την ελεγχοσυνάρτηση του ελέγχου χ^2, γνωρίζουμε ότι:

$$Q_n = \sum_{i=0}^{1} \frac{(O_i - E_i)^2}{E_i} \xrightarrow{d} Q \sim \chi^2_1.$$

Υπολογίζοντας την παρατηρούμενη τιμή της ελεγχοσυνάρτησης, παίρνουμε ότι

$$q = 1.20 < 2.71 = \chi^2_{1;0.1} \quad \text{και} \quad \text{p-value} = \mathbb{P}(Q_n \geq q) \approx \mathbb{P}(Q \geqslant q) = 0.27 > 0.1.$$

Επομένως, δεν απορρίπτουμε την H_0 σε ε.σ.σ. $\alpha = 0.1$.

(ii) Κατασκευάζουμε τον πίνακα συχνοτήτων των δεδομένων για το νόμισμα 2:

Πινακας 5.8: Δεδομένα Παραδείγματος 5.6 – (i)

i	0	1	Άθροισμα
O_i	20	10	30
$p_{i,0}$	1/2	1/2	1
E_i	15	15	30

Σχετικά με την ελεγχοσυνάρτηση του ελέγχου χ^2, γνωρίζουμε ότι:

$$Q_n = \sum_{i=0}^{1} \frac{(O_i - E_i)^2}{E_i} \xrightarrow{d} Q \sim \chi^2_1.$$

Υπολογίζοντας την παρατηρούμενη τιμή της ελεγχοσυνάρτησης, παίρνουμε ότι

$$q = 3.33 > 2.71 = \chi^2_{1;0.1} \quad \text{και} \quad \text{p-value} = \mathbb{P}(Q_n \geq q) \approx \mathbb{P}(Q \geq q) = 0.068 < 0.1.$$

Επομένως, απορρίπτουμε την H_0 σε ε.σ.σ. $\alpha = 0.1$ αποφασίζοντας ότι το νόμισμα είναι μεροληπτικό.

5.3 Έλεγχος χ^2 με ομαδοποίηση τιμών

Είναι πολύ συχνές οι περιπτώσεις που ο έλεγχος χ^2 εφαρμόζεται με ομαδοποίηση των τιμών του υπό μελέτη χαρακτηριστικού X. Η ανάγκη ομαδοποίησης προκύπτει σε περιπτώσεις όπου το σύνολο τιμών της X είναι:

- πεπερασμένο, αλλά αρκετά μεγάλο,
- πεπερασμένο, αλλά το μέγεθος του δείγματος δεν είναι επαρκώς μεγάλο,
- αριθμησίμως άπειρο για διακριτές τυχαίες μεταβλητές,
- κάποιο διάστημα του $\mathbb{R}$ για συνεχείς τυχαίες μεταβλητές,

Δεν θα πρέπει να ξεχνάμε εξάλλου ότι η ισχύς του κριτηρίου χ^2 είναι ασυμπτωτική και, έτσι, χρειαζόμαστε ένα επαρκώς μεγάλο μέγεθος δείγματος.

Σε περιπτώσεις όπως οι παραπάνω, το σύνολο των δυνατών τιμών S της X ομαδοποιείται σε ένα πεπερασμένο πλήθος ομάδων (κλάσεων, κελιών), έτσι ώστε να μπορέσει να υπολογιστεί το στατιστικό χ^2. Οι τιμές E_i που αντιστοιχούν στο αναμενόμενο πλήθος των παρατηρήσεων κάθε ομάδας υπολογίζονται υπό την ισχύ της μηδενικής υπόθεσης ή εκτιμώνται αν η μηδενική υπόθεση είναι σύνθετη με τη βοήθεια κάποιας εκτιμήτριας, όπως περιγράψαμε στην προηγούμενη παράγραφο. Τυπικά λοιπόν η διαδικασία υπολογισμού είναι ίδια όπως πριν, αντικαθιστώντας τα ενδεχόμενα $\{X = i\}$ με ενδεχόμενα της μορφής $\{X \in A_i\}$, όπου τα $\{A_i\}_{i=1}^d$ συνιστούν μία διαμέριση του S, δηλαδή $S = \bigcup_{i=1}^d A_i$ και $A_i A_j = \emptyset$ για $i \neq j$.

Είναι φανερό ότι μία κακή ομαδοποίηση μπορεί να μειώσει την ισχύ του ελέγχου, που έτσι και αλλιώς εξετάζει πλέον την προσαρμογή της υποθετικής κατανομής των ομαδοποιημένων τιμών πάνω στα ομαδοποιημένα δεδομένα, και χρειάζεται προσοχή στον τρόπο επιλογής των ομάδων. Για τους δικούς μας σκοπούς παρουσίασης της μεθόδου θα αρκεστούμε σε μία πρακτική συμβουλή να επιλέγουμε το πλήθος των κλάσεων με τέτοιο τρόπο ώστε $E_i \geq 5$ και στην περιπτώση συνεχών κατανομών να επιλέγουμε τα διαιρετικά σημεία της διαμέρισης με τρόπο ώστε οι κλάσεις να γίνονται ισοπίθανες. Δίνουμε τώρα ένα παράδειγμα.

Παράδειγμα 5.7: Τα πλήθη των πελατών που επισκέφθηκαν ένα συνοικιακό κατάστημα ανά ημέρα για ένα δείγμα 20 εργάσιμων ημερών ήταν:

$$2,\quad 3,\quad 0,\quad 4,\quad 2,\quad 2,\quad 3,\quad 1,\quad 0,\quad 2,\quad 1,\quad 3,\quad 3,\quad 4,\quad 2,\quad 4,\quad 0,\quad 1,\quad 2,\quad 3.$$

Να ελεγχθεί σε ε.σ.σ. $\alpha = 0.05$ αν τα δεδομένα παρέχουν ενδείξεις ότι ο αριθμός πελατών ανά ημέρα ακολουθεί κατανομή Poisson της μορφής: (i) $\mathcal{P}(1)$, (ii) $\mathcal{P}(\lambda)$, $\lambda > 0$.

Λύση.

(i) Έστω X το πλήθος των πελατών που επισκέπτονται το κατάστημα ανά ημέρα. Υπό την $H_0 : X \sim \text{Poisson}(1)$, και άρα

$$\mathbb{P}(X = x) = e^{-1} \cdot \frac{1}{x!}, \quad x = 0, 1, 2, \dots$$

Ορίζουμε αρχικά $A_1 = \{0\}$, $A_2 = \{1\}$, $A_3 = \{2\}$, $A_4 = \{3\}$, $A_5 = \{4\}$, $A_6 = \{5, 6, \dots\}$ και έτσι προκύπτει ο παρακάτω πίνακας:

Πίνακας 5.9: Δεδομένα Παραδείγματος 5.7 – (i)

i	0	1	2	3	4	5+	Άθροισμα
O_i	3	3	6	5	3	0	**20**
$p_{i,0}$	0.368	0.368	0.184	0.061	0.015	0.004	1
E_i	7.358	7.358	3.679	1.227	0.307	0.073	20

Παρατηρούμε ότι $E_i < 5$ για $i \geq 3$, οπότε ορίζουμε εκ νέου $A_1 = \{0\}$, $A_2 = \{1\}$, $A_3 = \{2, 3, \dots\}$ και παίρνουμε τον καινούργιο πίνακα συχνοτήτων:

Πίνακας 5.10: Δεδομένα Παραδείγματος 5.7 – (i) με ομαδοποίηση

i	0	1	2+	Άθροισμα
O_i	3	3	14	**20**
$p_{i,0}$	0.368	0.368	0.264	1
E_i	7.358	7.358	5.285	20

Η ελεγχοσυνάρτηση του ελέγχου χ^2 ικανοποιεί:

$$Q_n = \sum_{i=1}^{3} \frac{(O_i - E_i)^2}{E_i} \xrightarrow{d} Q \sim \chi^2_2.$$

Έτσι, σύμφωνα με τον παραπάνω πίνακα προκύπτει ότι

$$q = 11.167 > 5.992 = \chi^2_{2;0.05} \quad \text{και} \quad \text{p-value} = \mathbb{P}(Q_n \geq q) \approx \mathbb{P}(Q \geqslant q) = 0.004 < 0.05.$$

Επομένως, οδηγούμε στην απόρριψη της $H_0 : X \sim \text{Poisson}(1)$ σε ε.σ.σ. $\alpha = 0.05$.

(ii) Στην περίπτωση αυτή η μηδενική υπόθεση είναι σύνθετη. Υπό την $H_0 : X \sim \text{Poisson}(\lambda)$, όπου υπενθυμίζουμε ότι

$$\mathbb{P}(X = x) = e^{-\lambda} \cdot \frac{\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots, \quad \lambda > 0.$$

Η παράμετρος λ σε αυτήν την περίπτωση πρέπει να εκτιμηθεί. Η εκτιμήτρια μέγιστης πιθανοφάνειας του λ όταν οι παρατηρήσεις δεν είναι ομαδοποιημένες είναι $\widetilde{\lambda} = \overline{X}$. Έτσι η ε.μ.π. στο δείγμα αυτό δίνει $\widetilde{\lambda} = \overline{x} = 2.1$ Ορίζουμε αρχικά $A_1 = \{0\}$, $A_2 = \{1\}$, $A_3 = \{2\}$, $A_4 = \{3\}$, $A_5 = \{4\}$, $A_6 = \{5, 6, \ldots\}$ και κατασκευάζουμε τον πίνακα συχνοτήτων των δεδομένων:

Πίνακας 5.11: Δεδομένα Παραδείγματος 5.7 – (ii)

A_i	0	1	2	3	4	5+	Άθροισμα
O_i	3	3	6	5	3	0	**20**
$\widetilde{p}_{i,0}$	0.123	0.257	0.270	0.189	0.099	0.062	1
$\widetilde{E}_i$	2.449	5.143	5.400	3.780	1.985	1.243	20

Παρατηρούμε ότι $\widetilde{E}_i < 5$ για $i = 0$ και $i \geq 3$, οπότε ορίζουμε εκ νέου $A_1 = \{0, 1\}$, $A_2 = \{2\}$, $A_3 = \{3, 4, \ldots\}$ και παίρνουμε τον καινούργιο πίνακα συχνοτήτων:

Πίνακας 5.12: Δεδομένα Παραδείγματος με ε.μ.π. ατομικών παρατηρήσεων 5.7 – (ii)

A_i	$0, 1$	2	3+	Άθροισμα
O_i	6	6	8	**20**
$\widetilde{p}_{i,0}$	0.380	0.270	0.350	1
$\widetilde{E}_i$	7.592	5.400	7.007	20

Εφόσον έχουμε εκτιμήσει μία άγνωστη παράμετρο λ, θα περίμενε κανείς ότι:

$$\widetilde{Q}_n = \sum_{i=1}^{3} \frac{\left(O_i - \widetilde{E}_i\right)^2}{\widetilde{E}_i} \xrightarrow{d} Q \sim \chi^2_1.$$

Αν το παραπάνω αποτέλεσμα ήταν σωστό, θα έπρεπε λοιπόν $F^{-1}_{\widetilde{Q}_n}(0.95) \approx \chi^2_{1;0.05}$

(ανεξάρτητα τιμής λ_0) για επαρκώς μεγάλο n. Με τη μέθοδο Monte Carlo μπορούμε λοιπόν να εκτιμήσουμε το $\chi^2_{1;0.05}$ με το $\widetilde{q}_{n;0.05} \equiv F^{-1}_{\widetilde{Q}_n}(0.95)$ και να το επαληθεύσουμε. Προσομοιώσαμε 5×10^4 φορές την τιμή του στατιστικού $\widetilde{Q}_n$ για $n = 10^4$ με $\lambda = 2$. Η τιμή που πήραμε ήταν $\widetilde{q}_{n;0.05} = 4.046$, που διαφέρει από το $\chi^2_{1;0.05} = 3.841$. Επαναλάβαμε το πείραμα αυτό για $\lambda = 1$ και η τιμή που πήραμε ήταν $\widetilde{q}_{n;0.05} = 4.107$. Μπορεί οι αποκλίσεις αυτές να είναι αποτέλεσμα της μεταβλητότητας που οφείλεται στο Monte Carlo ή υπάρχει κάποιο λάθος στη θεωρία; Η αύξηση του Monte Carlo δείγματος απλώς επιβεβαίωσε τον υπολογισμό με ακρίβεια τουλάχιστον 3 δεκαδικών ψηφίων. Από την άλλη μεριά, μία προσεκτική ματιά στις υποθέσεις του ελέγχου χ^2 μας αποκαλύπτει την πηγή του προβλήματος. Η εκτιμήτρια μέγιστης πιθανοφάνειας $\widehat{\lambda}_n$ του υπομοντέλου που προκύπτει από τις ομαδοποιημένες παρατηρήσεις διαφέρει γενικά από την ε.μ.π. του μοντέλου των ατομικών παρατηρήσεων και, έτσι, παρόλο που και οι δύο εκτιμήτριες είναι τυπικά συνεπείς, δεν έχουν υποχρεωτικά ίδιες ασυμπτωτικές κατανομές. Επομένως, ο υπολογισμός του στατιστικού χ^2 θα πρέπει να γίνεται με την ε.μ.π. που προκύπτει από το αντίστοιχο πολυωνυμικό μοντέλο μετά τη διαδικασία της ομαδοποίησης. Συνήθως, αυτό θα οδηγεί σε ε.μ.π. που δεν έχουν λύση σε κλειστή μορφή και, έτσι, θα είναι απαραίτητο να ενσωματωθεί κάποιος αλγόριθμος αριθμητικής βελτιστοποίησης για να βρεθεί η ε.μ.π.. Αυτό βέβαια δεν αποτελεί πρόβλημα, αφού σε λογισμικά υπάρχουν πλέον έτοιμες συναρτήσεις βελτιστοποίησης. Το μόνο που χρειάζεται να καθοριστεί είναι η μορφή της συνάρτησης και, ενδεχομένως, και οι μερικές παράγωγοι για μεγαλύτερη ακρίβεια υπολογισμού. Η συνάρτηση πιθανοφάνειας λοιπόν που αντιστοιχεί στο υπομοντέλο θα είναι της μορφής

$$L(p) \propto \left[e^{-\lambda}(1+\lambda)\right]^{S_{n,1}} \left[e^{-\lambda}(\lambda^2/2)\right]^{S_{n,2}} \left[1 - e^{-\lambda}(1+\lambda+\lambda^2/2)\right]^{S_{n,3}},$$

όπου $S_{n,i}$ είναι το πλήθος των παρατηρήσεων στην ομάδα i, με $A_1 = \{0,1\}$, $A_2 = \{2\}$ και $A_3 = \{3,4,\ldots\}$. Με τη μέθοδο BFGS πήραμε ως ε.μ.π. $\widehat{\lambda}_{20} = 2.342$, ενώ ο καινούργιος πίνακας συχνοτήτων είναι:

Πινακας 5.13: Δεδομένα Παραδείγματος 5.7 – (ii) με ε.μ.π. ομαδοποιημένων παρατηρήσεων

A_i	$0,1$	2	$3+$	Άθροισμα
O_i	6	6	8	**20**
$\widetilde{p}_{i,0}$	0.321	0.264	0.415	1
$\widetilde{E}_i$	6.425	5.273	8.302	20

Υπολογίζοντας την τιμή της ελεγχοσυνάρτησης στο δείγμα προκύπτει:

$$q_{20} = 0.139 < 3.842 = \chi^2_{1;0.05} \quad \text{και} \quad \text{p-value} = \mathbb{P}(Q_{20} \geqslant q_{20}) \approx \mathbb{P}(Q \geqslant q_{20}) = 0.709 > 0.05,$$

και άρα δεν απορρίπτουμε την $H_0 : X \sim \mathcal{P}(\lambda)$ σε ε.σ.σ. $\alpha = 0.05$.

5.4 Έλεγχος Ομοιογένειας χ^2

Ο έλεγχος χ^2 μπορεί να χρησιμοποιηθεί για να ελεγχθεί το κατά πόσον ανεξάρτητα δείγματα προέρχονται από τον ίδιο πληθυσμό. Τέτοιες δοκιμασίες είναι γνωστές ως έλεγχοι ομοιογένειας.

Ορισμός 5.3 (έλεγχος ομοιογένειας χ^2):
Έστω $\{X_{11}, X_{12}, \ldots, X_{1n_{1\bullet}}\}$, $\{X_{21}, X_{22}, \ldots, X_{2n_{2\bullet}}\}, \ldots, \{X_{k1}, X_{k2}, \ldots, X_{k\,n_{k\bullet}}\}$ k ανεξάρτητα τυχαία δείγματα μεγεθών $n_{1\bullet}, n_{2\bullet}, \ldots, n_{k\bullet}$ από κατανομές k ανεξάρτητων χαρακτηριστικών $X_1, X_2, \ldots, X_k$ που παίρνουν τιμές σε κοινό σύνολο S. Υποθέτουμε ότι $\{A_1, A_2, \ldots, A_\ell\}$ είναι μία διαμέριση του S και $p_{ij} := \mathbb{P}(X_i \in A_j) > 0$ για κάθε $1 \leq i \leq k$ και $1 \leq j \leq \ell$. Ο έλεγχος υποθέσεων:

$$\begin{cases} H_0 : p_{1j} = p_{2j} = \cdots = p_{kj} \text{ για κάθε } j \in \{1, 2, \ldots, \ell\} \text{ vs.} \\ H_1 : p_{ij} \neq p_{rj} \text{ για κάποια } i \neq r \text{ και } j \in \{1, 2, \ldots, \ell\}, \end{cases}$$

καλείται *έλεγχος ομοιογένειας*. Αν

$$O_{ij} = \sum_{r=1}^{n_{i\bullet}} \mathbf{1}_{\{X_{ir} \in A_j\}}, \qquad 1 \leq i \leq k, \quad 1 \leq j \leq \ell,$$

το πλήθος των παρατηρήσεων του i-οστού δείγματος που ανήκουν στο A_j και ως ελεγχοσυνάρτηση επιλεγεί η

$$Q_n = \sum_{i=1}^{k} \sum_{j=1}^{\ell} \frac{(O_{ij} - \widehat{E}_{ij})^2}{\widehat{E}_{ij}} = \sum_{i=1}^{k} \sum_{j=1}^{\ell} \frac{(O_{ij} - \frac{n_{i\bullet} n_{\bullet j}}{n})^2}{\frac{n_{i\bullet} n_{\bullet j}}{n}},$$

τότε ο παραπάνω έλεγχος καλείται *έλεγχος ομοιογένειας* χ^2.

Παρατήρηση 5.3. Δίνουμε εδώ μία αιτιολόγηση της χρήσης του παραπάνω στατιστικού. Ας υποθέσουμε ότι $p_{1j} = p_{2j} = \cdots = p_{kj} = p_j$, για κάθε $1 \leq j \leq \ell$, όπου το διάνυσμα πιθανότητας $p = (p_1, p_2, \ldots, p_\ell)$ είναι γνωστό. Τότε, η μηδενική υπόθεση είναι απλή και μπορούμε να χρησιμοποιήσουμε το στατιστικό χ^2:

$$Q_n(p) = \sum_{i=1}^{k} \sum_{j=1}^{\ell} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \sum_{i=1}^{k} \sum_{j=1}^{\ell} \frac{(O_{ij} - n_{i\bullet} p_j)^2}{n_{i\bullet} p_j},$$

καθώς το αναμενόμενο πλήθος παρατηρήσεων του i-οστού δείγματος που ανήκουν στο A_j είναι:

$$E_{ij} = \mathbb{E}(O_{ij}) = \sum_{r=1}^{n_{i\bullet}} \mathbb{P}\left(X_{ir} \in A_j\right) = \sum_{r=1}^{n_{i\bullet}} p_j = n_{i\bullet} p_j.$$

Όμως, η μηδενική υπόθεση ομοιογένειας δεν προσδιορίζει μία συγκεκριμένη τιμή του p, αλλά το αφήνει ελεύθερο να κινείται σε όλα τα δυνατά διανύσματα πιθανοτήτων. Έτσι η H_0 είναι μία σύνθετη μηδενική υπόθεση και το διάνυσμα p πρέπει να εκτιμηθεί. Μπορούμε λοιπόν άμεσα να υπολογίσουμε τις εκτιμήτριες μέγιστης πιθανοφάνειάς τους, που δίνονται ως εξής:

$$\widehat{p}_j = \frac{n_{\bullet j}}{n}, \quad \text{όπου} \quad n_{\bullet j} = \sum_{i=1}^{k} O_{ij} \quad \text{και} \quad n = \sum_{i=1}^{k} n_{i\bullet} = \sum_{j=1}^{\ell} n_{\bullet j},$$

όπου το n αντιστοιχεί στο συνολικό πλήθος παρατηρήσεων που προκύπτουν συγχωνεύοντας όλα τα δείγματα.

Στο παρακάτω θεώρημα δίνεται η ασυμπτωτική κατανομή της ελεγχοσυνάρτησης χ^2 σε έναν τέτοιο έλεγχο ομοιογένειας. Η απόδειξη παραλείπεται.

Θεώρημα 5.2: Αν ισχύει η H_0 του ελέγχου ομοιογένειας στον Ορισμό 5.3, τότε

$$Q_n\left(\widehat{p}_n\right) = \sum_{i=1}^{k}\sum_{j=1}^{\ell} \frac{\left(O_{ij} - \widehat{E}_{ij}\right)^2}{\widehat{E}_{ij}} \xrightarrow[\substack{n_{i\bullet}\to\infty \\ 1\le i\le n}]{d} Q \sim \chi^2_{(k-1)(\ell-1)}, \quad \text{όπου} \quad \widehat{E}_{ij} = n_{i\bullet}\,\widehat{p}_j = \frac{n_{i\bullet}\, n_{\bullet j}}{n}.$$

Συμπερασματικά, αν q_n είναι η παρατηρούμενη τιμή της ελεγχοσυνάρτησης Q_n, τότε απορρίπτουμε τη μηδενική υπόθεση H_0 σε (ασυμπτωτικό) ε.σ.σ. α αν, και μόνο αν, $q_n > \chi^2_{(k-1)(\ell-1);\alpha}$ ή, ισοδύναμα, αν p-value $= \sup_{H_0} \mathbb{P}\left(Q_n(\widehat{p}_n) \geqslant q_n\right) \approx \mathbb{P}(Q \geqslant q_n) < \alpha$.

Παρατήρηση 5.4. Η διαίσθηση πίσω από το πλήθος των βαθμών ελευθερίας που προκύπτει στην παραπάνω ασυμπτωτική κατανομή χ^2 είναι η εξής: αν το διάνυσμα πιθανοτήτων p είναι γνωστό (απλή μηδενική), τότε το πλήθος των βαθμών ελευθερίας προκύπτει άμεσα ως το άθροισμα των $\ell - 1$ βαθμών, τόσες φορές όσα και τα ανεξάρτητα δείγματα και άρα συνολικά $k(\ell-1)$. Στην περίπτωση όμως που το p είναι άγνωστο, πρέπει να αφαιρέσουμε τόσους βαθμούς, όσο και η διάσταση του παραμετρικού χώρου που προκύπτει για την εκτίμηση του p. Καθώς το p εκτιμάται "μη παραμετρικά", δηλ. με μόνο περιορισμό ότι οι συνιστώσες του αθροίζουν στη μονάδα, η διάσταση είναι $\ell - 1$. Έτσι, τελικά, προκύπτουν $(k-1)(\ell-1)$ βαθμοί ελευθερίας.

Παρατήρηση 5.5. Για να είναι έγκυρος ο έλεγχος θα πρέπει τα μεγέθη των δειγμάτων $\{n_{i\bullet}\}_{i=1}^{k}$, $1 \le i \le k$, να είναι επαρκώς μεγάλα ώστε να εξασφαλίζεται η ισχύς του πολυδιάστατου κεντρικού οριακού θεωρήματος (κατάλληλης επέκτασης με πολυδείκτες) και της προκύπτουσας χ^2 κατανομής. Επιπλέον, η επιλογή της διαμέρισης $\{A_j\}_1^\ell$ είναι μεν υποκειμενική, αλλά θα πρέπει να φροντίζουμε να τηρείται ο περιορισμός $\widehat{E}_{ij} \geq 5$.

Παράδειγμα 5.8: Σε 2 ανεξάρτητα δείγματα 20 και 30 ατόμων αντίστοιχα δόθηκε να δοκιμάσουν μία νέα μάρκα απορρυπαντικού και να το συγκρίνουν με αυτό που χρησιμοποιούσαν. Οι απαντήσεις τους συνοψίζονται στον παρακάτω πίνακα.

Πινακας 5.14: Δεδομένα Παραδείγματος 5.8

O_{ij}	Καμία Διαφορά	Χειρότερο	Καλύτερο	$n_{i\bullet}$
Δείγμα 1	4	7	9	20
Δείγμα 2	11	6	13	30
$n_{\bullet j}$	15	13	22	50

Ζητείται να διερευνηθεί αν υπάρχει στατιστικά σημαντική διαφορά στις απαντήσεις των 2 δειγμάτων σε ε.σ.σ. $\alpha = 0.05$.

Λύση.

Ο έλεγχος αυτός αντιστοιχεί σε έναν έλεγχο ομοιογένειας των δειγμάτων. Υπολογίζουμε τον πίνακα με τα αναμενόμενα πλήθη των παρατηρήσεων του i-οστού δείγματος που ανήκουν στο A_j για $i = 1, 2$ και $j = 1, 2, 3$ υπό την H_0. Σχετικά με την ελεγχοσυνάρτηση του ελέγχου ομοιογένειας χ^2, έχουμε $(k-1)(l-1) = 2$ και άρα

$$Q_n = \sum_{i=1}^{2} \sum_{j=1}^{3} \frac{\left(O_{ij} - \widehat{E}_{ij}\right)^2}{\widehat{E}_{ij}} \xrightarrow{d} Q \sim \chi_2^2.$$

Υπολογίζοντας την παρατηρούμενη τιμή της ελεγχοσυνάρτησης, παίρνουμε ότι

$$q_n = 2.157 < 5.992 = \chi^2_{2;0.05} \quad \text{και} \quad \text{p-value} = \sup_{H_0} \mathbb{P}(Q_n \geqslant q_n) \approx \mathbb{P}(Q \geqslant q_n) = 0.340 > 0.05.$$

Επομένως, δεν απορρίπτουμε τη μηδενική υπόθεση ότι τα δύο δείγματα προέρχονται από την ίδια κατανομή σε (ασυμπτωτικό) ε.σ.σ. $\alpha = 0.05$.

Παράδειγμα 5.9: Μία εταιρεία διεξάγει τρεις διαφορετικές έρευνες για την προτίμηση των καταναλωτών ανάμεσα σε τέσσερα προϊόντα, Α, Β, Γ, Δ. Τα αποτελέσματα των ερευνών σε παρουσιάζονται στον παρακάτω πίνακα.

Πινακας 5.15: Δεδομένα Παραδείγματος 5.9

O_{ij}	Α	Β	Γ	Δ	$n_{i\bullet}$
Έρευνα 1	201	507	153	139	1000
Έρευνα 2	47	92	31	30	200
Έρευνα 3	28	78	20	24	150
$n_{\bullet j}$	276	677	204	193	1350

Ζητείται να διερευνηθεί αν υπάρχει στατιστικά σημαντική διαφορά στις απαντήσεις των 3 δειγμάτων σε ε.σ.σ. $\alpha = 0.05$.
Λύση.
Ο έλεγχος αυτός αντιστοιχεί σε έναν έλεγχο ομοιογένειας 3 ανεξάρτητων δειγμάτων. Υπολογίζουμε τον πίνακα του αναμενόμενου πλήθους καταναλωτών που προτιμούν κάθε κατηγορία προϊόντος για κάθε δείγμα ξεχωριστά υπό την H_0. Σχετικά με την ελεγχοσυνάρτηση του ελέγχου ομοιογένειας χ^2, έχουμε $(k-1)(l-1) = 2 * 3 = 6$ και άρα

$$Q_n = \sum_{i=1}^{3}\sum_{j=1}^{4} \frac{\left(O_{ij} - \widehat{E}_{ij}\right)^2}{\widehat{E}_{ij}} \xrightarrow{d} Q \sim \chi^2_6.$$

Υπολογίζοντας την παρατηρούμενη τιμή της ελεγχοσυνάρτησης, παίρνουμε ότι

$$q = 2.89 < 12.59 = \chi^2_{6;0.05} \quad \text{και} \quad \text{p-value} = \sup_{H_0} \mathbb{P}(Q_n \geqslant q) \approx \mathbb{P}(Q \geqslant q) = 0.822 > 0.05.$$

Επομένως, δεν απορρίπτουμε τη μηδενική υπόθεση ότι τα τρία δείγματα προέρχονται από την ίδια κατανομή σε (ασυμπτωτικό) ε.σ.σ. $\alpha = 0.05$.

Παράδειγμα 5.10: Μία εταιρεία διεξάγει τρεις διαφορετικές έρευνες για την πρόθεση ψήφου των πολιτών ενός δήμου ανάμεσα σε πέντε υποψήφιους δημάρχους, Α, Β, Γ, Δ, Ε. Τα αποτελέσματα των ερευνών παρουσιάζονται στον παρακάτω πίνακα.

Πινακας 5.16: Δεδομένα Παραδείγματος 5.10

O_{ij}	Α	Β	Γ	Δ	Ε	$n_{i\bullet}$
Έρευνα 1	54	24	14	22	36	150
Έρευνα 2	65	53	42	35	55	250
Έρευνα 3	61	40	50	14	35	200
$n_{\bullet j}$	180	117	106	71	126	600

Ζητείται να διερευνηθεί αν υπάρχει στατιστικά σημαντική διαφορά στις απαντήσεις των 3 δειγμάτων σε ε.σ.σ. $\alpha = 0.01$.
Λύση.
Ο έλεγχος αυτός αντιστοιχεί σε έναν έλεγχο ομοιογένειας 3 ανεξάρτητων δειγμάτων. Υπολογίζουμε τον πίνακα του αναμενόμενου αριθμού ψηφοφόρων που προτιμούν κάθε υποψήφιο δήμαρχο για κάθε δείγμα ξεχωριστά υπό την H_0. Σχετικά με την ελεγχοσυνάρτηση του ελέγχου ομοιογένειας χ^2, έχουμε $(k-1)(l-1) = 2 \cdot 4 = 8$ και άρα

$$Q_n = \sum_{i=1}^{3} \sum_{j=1}^{5} \frac{\left(O_{ij} - \widehat{E}_{ij}\right)^2}{\widehat{E}_{ij}} \xrightarrow{d} Q \sim \chi^2_8.$$

Υπολογίζοντας την παρατηρούμενη τιμή της ελεγχοσυνάρτησης, παίρνουμε ότι

$$q = 24.47 > 15.51 = \chi^2_{8;0.05} \quad \text{και} \quad \text{p-value} = \sup_{H_0} \mathbb{P}(Q_n \geqslant q) \approx \mathbb{P}(Q \geqslant q) = 0.0019 < 0.01.$$

Επομένως, απορρίπτουμε τη μηδενική υπόθεση ότι τα τρία δείγματα προέρχονται από την ίδια κατανομή σε (ασυμπτωτικό) ε.σ.σ. $\alpha = 0.01$. Η επαλήθευση των παραπάνω αποτελεσμάτων υλοποιείται ως εξής:

```
O = matrix(c(54,65,61,24,53,40,14,42,50,22,35,14,36,55,35), 3,5)
a = 0.05 ; df= (3-1)*(5-1)
E = rowSums(O) %*% t(colSums(O))/sum(O)
q = sum((O-E)^2/E)
qa = qchisq(a,df,lower.tail=FALSE)
p = pchisq(q ,df,lower.tail=FALSE)
c(q ,qa, p)
```

Η απευθείας πραγματοποίηση του παραπάνω ελέγχου είναι επίσης εφικτή:

```
O = matrix(c(54, 65, 61, 24, 53, 40, 14, 42, 50, 22, 35, 14, 36,
   55, 35), nrow=3, ncol=5)
> chisq.test(O)
```

```
Pearson Chi-squared test
data:  O
X-squared = 24.465, df = 8, p-value = 0.001914
```

5.5 Έλεγχος McNemar

Ο έλεγχος McNemar μας επιτρέπει να ελέγξουμε κατά πόσον μία αναλογία αναφορικά με την παρουσία ή την απουσία κάποιου χαρακτηριστικού ενός πληθυσμού παραμένει σταθερή μετά τη μεσολάβηση κάποιου γεγονότος όταν χρησιμοποιηθεί ένα τυχαίο δείγμα από ζευγαρωτές παρατηρήσεις. Προκύπτει ως ειδική περίπτωση ενός ελέγχου ομοιογένειας χ^2 που αφορά όμως δύο δίτιμες κατηγορικές τυχαίες μεταβλητές που βρίσκονται σε αντιστοιχία.

Ορισμός 5.4 (έλεγχος McNemar):
Έστω $\{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ ένα τυχαίο δείγμα ζευγαρωτών παρατηρήσεων από την κατανομή ενός δισδιάστατου χαρακτηριστικού (X, Y). Υποθέτουμε ότι το σύνολο τιμών των X και Y είναι κοινό και $\{A_1, A_2\}$ αποτελεί μία διαμέρισή του. Αν $\mathbb{P}(X \in A_i, Y \in A_j) = p_{ij} > 0$, για $1 \leq i, j \leq 2$, τότε ο έλεγχος υποθέσεων:

$$H_0 : p_{12} = p_{21} \qquad \text{vs.} \qquad H_1 : p_{12} \neq p_{21}.$$

με χρήση της ελεγχοσυνάρτησης

$$Q_n = \frac{(O_{12} - O_{21})^2}{O_{12} + O_{21}},$$

όπου

$$O_{ij} = \sum_{r=1}^{n} \mathbf{1}_{\{X_r \in A_i, Y_r \in A_j\}}, \quad 1 \leq i, j \leq 2,$$

καλείται *έλεγχος McNemar*.

Παρατήρηση 5.6. Δίνουμε εδώ μία αιτιολόγηση της χρήσης του παραπάνω στατιστικού. Υπό την ισχύ της $H_0 : p_{12} = p_{21}$, έχουμε

$$\begin{aligned} \mathbb{P}(X \in A_1) = p_{11} + p_{12} &\overset{H_0}{=} p_{11} + p_{21} = \mathbb{P}(Y \in A_1), \\ \mathbb{P}(X \in A_2) = 1 - \mathbb{P}(X \in A_1) &\overset{H_0}{=} 1 - \mathbb{P}(Y \in A_1) = \mathbb{P}(Y \in A_2), \end{aligned}$$

και έτσι προκύπτει μία ειδική μορφή ελέγχου ομοιογένειας. Με χρήση του στατιστικού χ^2 σε έλεγχο ομοιογένειας που δίνεται στην (5.9), έχουμε

$$Q_n = \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(O_{ij} - \widehat{E}_{ij})^2}{\widehat{E}_{ij}}, \quad \text{όπου} \quad \widehat{E}_{ij} = n\,\widehat{p}_{ij}, \tag{5.6}$$

και $\widehat{p}_{ij}$ εκτιμάται μέσω μέγιστης πιθανοφάνειας. Υπό την ισχύ της $H_0 : p_{12} = p_{21}$, και της προφανούς σχέσης $p_{11} + p_{12} + p_{21} + p_{22} = 1$ έχουμε 2 άγνωστες παραμέτρους και μέσω της μεγιστοποίησης της πιθανοφάνειας κάτω από αυτούς τους περιορισμούς προκύπτει ότι:

$$\widehat{p}_{11} = \frac{O_{11}}{n}, \quad \widehat{p}_{22} = \frac{O_{22}}{n} \quad \text{και} \quad \widehat{p}_{12} = \widehat{p}_{21} = \frac{O_{12} + O_{21}}{2n}. \tag{5.7}$$

Αντικαθιστώντας τη μορφή (5.7) των ε.μ.π. στη σχέση (5.6) προκύπτει τελικά ότι

$$Q_n = \frac{(\cancel{O_{11}} - \cancel{O_{11}})^2}{O_{11}} + \frac{(O_{12} - O_{21})^2}{2\,(O_{12} + O_{21})} + \frac{(O_{21} - O_{12})^2}{2\,(O_{12} + O_{21})} + \frac{(\cancel{O_{22}} - \cancel{O_{22}})^2}{O_{22}} = \frac{(O_{12} - O_{21})^2}{O_{12} + O_{21}}.$$

Στο παρακάτω θεώρημα δίνεται η ασυμπτωτική κατανομή της ελεγχοσυνάρτησης McNemar.

Θεώρημα 5.3: Αν ισχύει η $H_0 : p_{12} = p_{21}$ του ελέγχου McNemar που περιγράφεται στον Ορισμό 5.4, τότε

$$Q_n = \frac{(O_{12} - O_{21})^2}{O_{12} + O_{21}} \xrightarrow[n\to\infty]{d} Q \sim \chi^2_1.$$

Συμπερασματικά, αν q_n είναι η παρατηρούμενη τιμή της ελεγχοσυνάρτησης Q_n, τότε απορρίπτουμε τη μηδενική υπόθεση H_0 σε (ασυμπτωτικό) ε.σ.σ. α αν, και μόνο αν, $q_n > \chi^2_{1;\alpha}$ ή, ισοδύναμα, αν p-value $= \sup_{H_0} \mathbb{P}\left(Q_n(\widehat{p}_n) \geqslant q_n\right) \approx \mathbb{P}(Q \geqslant q_n) < \alpha$.

Παρατήρηση 5.7. Η διαίσθηση πίσω από το πλήθος των βαθμών ελευθερίας που προκύπτει στην παραπάνω ασυμπτωτική κατανομή χ^2 είναι η εξής: αν το διάνυσμα πιθανοτήτων $p = (p_{ij})_{ij}$ είναι γνωστό (απλή μηδενική), τότε το πλήθος των βαθμών ελευθερίας που προκύπτει είναι 3, καθώς ο μόνος περιορισμός είναι ότι αθροίζουν στη μονάδα. Ο παραμετρικός χώρος που προκύπτει υπό την H_0 έχει διάσταση 2. Έτσι οι συνολικοί βαθμοί ελευθερίας που προκύπτουν για την ασυμπτωτική κατανομή είναι $3 - 2 = 1$.

Παρατήρηση 5.8. Για να είναι έγκυρος ο έλεγχος θα πρέπει το μέγεθος του δείγματος n να είναι επαρκώς μεγάλο ώστε να μπορεί να εφαρμοστεί η ασυμπτωτική προσέγγιση. Επιπλέον, θα πρέπει να πληρείται η συνθήκη $O_{12} + O_{21} \geq 25$. Στην περίπτωση που δεν ισχύουν οι παραπάνω συνθήκες, δηλ., για μικρά δείγματα ή όταν $O_{12} + O_{21} < 25$, έχει προταθεί μία διόρθωση που είναι γνωστή με το όνομα *Yates*. το διορθωμένο στατιστικό υπολογίζεται από τη σχέση:

$$Q_{\text{Yates}} = \frac{(|O_{12} - O_{21}| - 1)^2}{O_{12} + O_{21}}, \tag{5.8}$$

Η μέθοδος τροποποίησης του στατιστικού προέρχεται από τη *διόρθωση συνεχείας* για διακριτές τυχαίες μεταβλητές που εισάγεται στις στοιχειώδεις Πιθανότητες. Η χρησιμοποίηση του Q_{Yates} δίνει λιγότερα στατιστικά σημαντικά αποτελέσματα, ισχυροποιώντας τον έλεγχο όταν ικανοποιούνται οι παραπάνω συνθήκες. Επομένως, εκτός βέβαια αν $_{12} = {}_{21}$, θα έχουμε $Q_{\text{Yates}} < Q_n$.

Παράδειγμα 5.11: Ας υποθέσουμε ότι ζητείται από 50 άτομα να δηλώσουν ποιον από τους δύο υποψήφιους δημάρχους A και B προτιμούν. Μετά από μία τηλεοπτική αντιπαράθεση των δύο υποψηφίων, τα ίδια άτομα ξαναρωτήθηκαν για την προτίμησή τους μεταξύ των δύο υποψηφίων. Το πλήθος των ατόμων που άλλαξαν προτίμηση από τον A στον B είναι 28, από τον B στον A είναι 8, ενώ 8 άτομα παρέμειναν σταθεροί στον A και 6 άτομα παρέμειναν σταθεροί στον B. Να ελεγχθεί σε ε.σ.σ. $\alpha = 0.05$ αν υπάρχουν σημαντικές αλλαγές προτίμησης των δύο υποψηφίων μετά την τηλεοπτική τους αντιπαράθεση.

Λύση.

Κατασκευάζουμε αρχικά τον πίνακα συχνοτήτων των δεδομένων.

		Μετά		
	O_{ij}	A	B	$n_{i.}$
Πριν	A	8	28	36
	B	8	6	14
	$n_{.j}$	16	34	50

Σχετικά με την ελεγχοσυνάρτηση του ελέγχου McNemar, γνωρίζουμε ότι:

$$Q_n = \frac{(O_{12} - O_{21})^2}{O_{12} + O_{21}} \xrightarrow{d} Q \sim \chi^2_1.$$

Υπολογίζοντας την τιμή q_n, από τα δεδομένα που δίνονται έχουμε:

$$q_n = 11.111 > 3.842 = \chi^2_{1;0.05} \quad \text{ή} \quad \text{p-value} = \sup_{H_0} \mathbb{P}(Q_n \geqslant q_n) \approx \mathbb{P}(Q \geqslant q_n) = 8.581{\cdot}10^{-4} < 0.05.$$

Επομένως, απορρίπτουμε τη μηδενική υπόθεση ότι δεν υπάρχουν σημαντικές αλλαγές προτίμησης των δύο υποψηφίων σε ε.σ.σ. $\alpha = 0.05$.

5.6 Έλεγχος Ανεξαρτησίας χ^2

Ο έλεγχος χ^2 μπορεί να χρησιμοποιηθεί και για τον έλεγχο ανεξαρτησίας μεταξύ τυχαίων μεταβλητών.

Ορισμός 5.5 (έλεγχος ανεξαρτησίας χ^2):
Έστω $\{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ ένα τυχαίο δείγμα ζευγαρωτών παρατηρήσεων από την κατανομή ενός δισδιάστατου χαρακτηριστικού (X, Y). Υποθέτουμε ότι $\{A_i\}_{i=1}^{k}$ και $\{B_j\}_{j=1}^{\ell}$ είναι δύο διαμερίσεις του πεδίου τιμών της X και της Y αντίστοιχα, και

$$\begin{aligned}
\mathbb{P}(X \in A_i) &= p_{i\bullet} > 0, && i \in \{1, 2, \ldots, k\}, \\
\mathbb{P}(Y \in B_j) &= p_{\bullet j} > 0, && j \in \{1, 2, \ldots, \ell\}, \\
\mathbb{P}(X \in A_i, Y \in B_j) &= p_{ij} > 0, && i \in \{1, 2, \ldots, k\},\ j \in \{1, 2, \ldots, \ell\}.
\end{aligned}$$

Ο έλεγχος υποθέσεων:

$$\begin{cases}
H_0 : p_{ij} = p_{i\bullet} \cdot p_{\bullet j} \text{ για } i \in \{1, 2, \ldots, k\} \text{ και } j \in \{1, 2, \ldots, \ell\} \text{ vs.} \\
H_1 : p_{ij} \neq p_{i\bullet} \cdot p_{\bullet j} \text{ για κάποια } i \in \{1, 2, \ldots, \ell\} \text{ και } j \in \{1, 2, \ldots, k\},
\end{cases}$$

καλείται *έλεγχος ανεξαρτησίας*. Αν ορίσουμε

$$O_{ij} = \sum_{r=1}^{n} \mathbf{1}_{\{X_r \in A_i, Y_r \in B_j\}}$$

την απαριθμήτρια τυχαία μεταβλητή που εκφράζει το πλήθος των ζευγών $\{(X_r, Y_r)\}_{r=1}^{n}$, που θα βρεθούν εντός του ορθογωνίου $A_i \times B_j$ και ως ελεγχοσυνάρτηση επιλεγεί η

$$Q_n = \sum_{i=1}^{k} \sum_{j=1}^{\ell} \frac{(O_{ij} - \widehat{E}_{ij})^2}{\widehat{E}_{ij}}, \quad \text{όπου} \quad \widehat{E}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n},\ n_{i\bullet} = \sum_{j=1}^{\ell} O_{ij},\ n_{\bullet j} = \sum_{i=1}^{k} O_{ij}, \tag{5.9}$$

τότε ο παραπάνω έλεγχος καλείται *έλεγχος ανεξαρτησίας* χ^2.

Παρατήρηση 5.9. Η αιτιολόγηση χρήσης του παραπάνω στατιστικού είναι απλή. Ας υποθέσουμε ότι ελέγχουμε την ανεξαρτησία για μία συγκεκριμένη επιλογή παραμέτρων $p = (p_{ij})_{i,j}$ έτσι ώστε να πληρείται η σχέση $p_{ij} = p_{i\bullet}\, p_{\bullet j}$. Τότε, η μηδενική υπόθεση είναι απλή και μπορούμε να χρησιμοποιήσουμε το στατιστικό χ^2:

$$Q_n(p) = \sum_{i=1}^{k} \sum_{j=1}^{\ell} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \sum_{i=1}^{k} \sum_{j=1}^{\ell} \frac{(O_{ij} - n\, p_{i\bullet} p_{\bullet j})^2}{n\, p_{i\bullet} p_{\bullet j}}, \tag{5.10}$$

καθώς το αναμενόμενο πλήθος παρατηρήσεων που θα βρεθούν εντός του ορθογωνίου $A_i \times B_j$ υπό την H_0 είναι:

$$E_{ij} = \mathbb{E}(O_{ij}) = \sum_{r=1}^{n} \mathbb{P}\left(X_r \in A_i, Y_r \in B_j\right) = \sum_{r=1}^{n} \mathbb{P}\left(X_r \in A_i\right) \mathbb{P}\left(Y_r \in B_j\right) = n\, p_{i\bullet} p_{\bullet j}.$$

Όμως, η μηδενική υπόθεση ανεξαρτησίας δεν προσδιορίζει μία συγκεκριμένη τιμή του $p = (p_{ij})_{ij}$, αλλά το αφήνει ελεύθερο να κινείται σε όλα τα δυνατά στοχαστικά διανύσματα που ικανοποιούν τους περιορισμούς της ανεξαρτησίας. Έτσι η H_0 είναι μία σύνθετη μηδενική υπόθεση και το διάνυσμα p πρέπει εκτιμηθεί. Οι εκτιμήτριες μέγιστης πιθανοφάνειάς τους, υπολογίζονται άμεσα και δίνονται από τις σχέσεις:

$$\widehat{p}_{i\bullet} = \frac{n_{i\bullet}}{n} \quad \text{και} \quad \widehat{p}_{\bullet j} = \frac{n_{\bullet j}}{n},$$

όπου

$$n_{i\bullet} = \sum_{r=1}^{n} \mathbf{1}_{\{X_r \in A_i\}} = \sum_{j=1}^{\ell} O_{ij} \quad \text{και} \quad n_{\bullet j} = \sum_{r=1}^{n} \mathbf{1}_{\{Y_r \in B_j\}} = \sum_{i=1}^{k} O_{ij},$$

εκφράζοντας το συνολικό πλήθος των παρατηρήσεων $\{X_r\}_{r=1}^n$ που βρίσκονται εντός του A_i και των παρατηρήσεων $\{Y_r\}_{r=1}^n$ που βρίσκονται εντός του B_j αντίστοιχα. Τελικά, η μορφή του στατιστικού χ^2 που δίνεται στη σχέση (5.9) προκύπτει από την (5.10), αντικαθιστώντας τα $p_{i\bullet}$ και $p_{\bullet j}$ με τις παραπάνω εκτιμήτριες.

Στο παρακάτω θεώρημα δίνεται η ασυμπτωτική κατανομή της ελεγχοσυνάρτησης χ^2 σε έναν τέτοιο έλεγχο ανεξαρτησίας. Η απόδειξη παραλείπεται.

Θεώρημα 5.4: Αν ισχύει η H_0 του ελέγχου ανεξαρτησίας στον Ορισμό 5.5, τότε

$$Q_n(\widehat{p}_n) = \sum_{i=1}^{k} \sum_{j=1}^{\ell} \frac{(O_{ij} - \widehat{E}_{ij})^2}{\widehat{E}_{ij}} \xrightarrow{d} Q \sim \chi^2_{(k-1)(\ell-1)}, \quad \text{όπου} \quad \widehat{E}_{ij} = n\,\widehat{p}_{i\bullet}\,\widehat{p}_{\bullet j} = \frac{n_{i\bullet} n_{\bullet j}}{n}.$$

Συμπερασματικά, αν q_n είναι η παρατηρούμενη τιμή της ελεγχοσυνάρτησης Q_n, τότε απορρίπτουμε τη μηδενική υπόθεση H_0 σε (ασυμπτωτικό) ε.σ.σ. α αν, και μόνο αν, $q_n > \chi^2_{(k-1)(\ell-1);\alpha}$ ή, ισοδύναμα, αν p-value $= \sup_{H_0} \mathbb{P}\left(Q_n(\widehat{p}_n) \geqslant q_n\right) \approx \mathbb{P}(Q \geqslant q_n) < \alpha$.

Παρατήρηση 5.10. Η διαίσθηση πίσω από το πλήθος των βαθμών ελευθερίας που προκύπτει στην παραπάνω ασυμπτωτική κατανομή χ^2 είναι η εξής: αν το διάνυσμα πιθανοτήτων $p = (p_{ij})_{ij}$ είναι γνωστό (απλή μηδενική), τότε το πλήθος των βαθμών ελευθερίας που προκύπτει είναι $kl-1$. Υπό την H_0, η εκτίμηση των $p_{i\bullet}$ και $p_{\bullet j}$, με μόνο περιορισμό της στοχαστικότητάς τους, δεσμεύει $k+l-2$ παραμέτρους. Άρα οι συνολικοί βαθμοί ελευθερίας που προκύπτουν για την ασυμπτωτική κατανομή είναι $kl - 1 - (k + l - 2) = (k - 1)(l - 1)$.

Παρατήρηση 5.11. Για να είναι έγκυρος ο έλεγχος θα πρέπει το μέγεθος του δείγματος n να είναι επαρκώς μεγάλο ώστε να μπορεί να εφαρμοστεί η ασυμπτωτική προσέγγιση. Επιπλέον, η επιλογή των A_i και B_j γίνεται μεν υποκειμενικά, αλλά είναι ασφαλέστερο να εξασφαλίζεται η συνθήκη $\widehat{E}_{ij} \geq 5$ για $1 \leq i \leq k$ και $1 \leq j \leq \ell$.

Παράδειγμα 5.12: Ο παρακάτω πίνακας συνάφειας παρουσιάζει δεδομένα που αφορούν 141 ασθενείς με κύστη στο συκώτι, οι οποίοι ταξινομήθηκαν με βάση τη θέση και το μέγεθος της κύστης. Να ελεγχθεί αν η θέση και το μέγεθος της κύστης είναι ανεξάρτητα σε ε.σ.σ. $\alpha = 0.05$.

		Θέση			
	O_{ij}	Α	Β	Γ	$n_{i\bullet}$
Μέγεθος	Ι	23	9	6	38
	ΙΙ	21	4	3	28
	ΙΙΙ	34	24	17	75
	$n_{\bullet j}$	78	37	26	141

Λύση.
Σχηματίζουμε τον πίνακα συνάφειας με τα αναμενόμενα πλήθη παρατηρήσεων που ανήκουν στο A_i και στο B_j για $i, j = 1, 2, 3$ υπό την H_0.

		Θέση			
	$\widehat{E}_{ij}$	Α	Β	Γ	$n_{i\cdot}$
Μέγεθος	Ι	21.021	9.972	7.007	38
	ΙΙ	15.489	7.3475	5.163	28
	ΙΙΙ	41.489	19.681	13.830	75
	$n_{\cdot j}$	78	37	26	141

Σχετικά με την ελεγχοσυνάρτηση του ελέγχου ανεξαρτησίας χ^2, γνωρίζουμε ότι:

$$Q_n = \sum_{i=1}^{3}\sum_{j=1}^{3} \frac{\left(O_{ij} - \widehat{E}_{ij}\right)^2}{\widehat{E}_{ij}} \xrightarrow{d} Q \sim \chi^2_4.$$

Υπολογίζοντας την παρατηρούμενη τιμή της ελεγχοσυνάρτησης, παίρνουμε ότι

$$q_n = 7.844 < 9.488 = \chi^2_{4;0.05} \text{ και p-value} = \sup_{H_0} \mathbb{P}(Q_n \geqslant q_n) \approx \mathbb{P}(Q \geq q_n) = 0.098 > 0.05.$$

Επομένως, δεν απορρίπτουμε τη μηδενική υπόθεση ότι το μέγεθος και η θέση του όγκου είναι ανεξάρτητα σε (ασυμπτωτικό) ε.σ.σ. $\alpha = 0.05$.

Παράδειγμα 5.13: Ο παρακάτω πίνακας συνάφειας παρουσιάζει δεδομένα που αφορούν 200 μεταλλικά καλώδια ίσων διαστάσεων, τα οποία ταξινομήθηκαν με βάση το υλικό (Α, Β, Γ, Δ, Ε) και την ηλεκτρική αγωγιμότητα. Να ελεγχθεί αν το υλικό του μετάλλου και η ηλεκτρική αγωγιμότητα είναι ανεξάρτητα σε ε.σ.σ. $\alpha = 0.05$.

		Υλικό					
	O_{ij}	Α	Β	Γ	Δ	Ε	$n_{i\bullet}$
Ηλ. Αγ.	I	16	5	10	10	9	50
	II	17	6	15	9	13	60
	III	4	6	8	13	9	40
	IV	11	6	16	11	6	50
	$n_{\bullet j}$	48	23	49	43	37	200

Λύση.
Σχηματίζουμε τον πίνακα συνάφειας με τα αναμενόμενα πλήθη παρατηρήσεων που ανήκουν στο A_i και στο B_j για $i = 1, 2, 3, 4,\ j = 1, 2, 3, 4, 5$ υπό την H_0.

		Υλικό					
	$\widehat{E}_{ij}$	Α	Β	Γ	Δ	Ε	$n_{i\bullet}$
Ηλ. Αγ.	I	12	5.75	12.25	10.75	9.25	50
	II	14.4	6.90	14.70	12.90	11.10	60
	III	9.60	4.60	9.80	8.60	7.40	40
	IV	12.0	5.75	12.25	10.75	9.25	50
	$n_{\bullet j}$	48	23	49	43	37	200

Σχετικά με την ελεγχοσυνάρτηση του ελέγχου ανεξαρτησίας χ^2, γνωρίζουμε ότι:

$$Q_n = \sum_{i=1}^{4} \sum_{j=1}^{5} \frac{\left(O_{ij} - \widehat{E}_{ij}\right)^2}{\widehat{E}_{ij}} \xrightarrow{d} Q \sim \chi^2_{12}.$$

Υπολογίζοντας την παρατηρούμενη τιμή της ελεγχοσυνάρτησης, παίρνουμε ότι

$$q_n = 13.01 < 21.03 = \chi^2_{12;0.05} \text{ και p-value} = \sup_{H_0} \mathbb{P}(Q_n \geqslant q_n) \approx \mathbb{P}(Q \geq q_n) = 0.368 > 0.05.$$

Επομένως, δεν απορρίπτουμε τη μηδενική υπόθεση ότι το υλικό και η ηλεκτρική αγωγιμότητα είναι ανεξάρτητα σε (ασυμπτωτικό) ε.σ.σ. $\alpha = 0.05$.

5.7 Έλεγχος χ^2 - Δυαδικές Ταξινομήσεις

Ο έλεγχος χ^2 που αφορά δυαδικές ταξινομήσεις είναι ένα κριτήριο συσχέτισης ποιοτικών χαρακτηριστικών και προκύπτει ως ειδική περίπτωση του ελέγχου ανεξαρτησίας χ^2 όταν τα υπό μελέτη χαρακτηριστικά χωρίζονται σε 2 μόνο κατηγορίες. Συνήθως, αυτό αφορά την παρουσία ή την απουσία κάποιου χαρακτηριστικού και η ειδική αυτή μορφή οδηγεί σε κάποιες απλοποιήσεις στον υπολογισμό της τιμής της ελεγχοσυνάρτησης χ^2. Η διαδικασία υπολογισμού του διευκολύνεται με τη χρήση του λεγόμενου *τετράπτυχου πίνακα*:

		2ο χαρακτηριστικό		
	O_{ij}	Παρόν	Απόν	r_i
1ο χαρακτηριστικό	Παρόν	a	b	$a+b=r_1$
	Απόν	c	d	$c+d=r_2$
	c_j	$a+c=c_1$	$b+d=c_2$	n

Με τους παραπάνω συμβολισμούς ο υπολογισμός της ελεγχοσυνάρτησης χ^2 απλοποιείται στη μορφή:

$$Q_n = \frac{(\det A)^2 \cdot n}{r_1 r_2 c_1 c_2}. \tag{5.11}$$

Έχει προταθεί και μία διόρθωση για μικρά δείγματα ή όταν $\widehat{E}_{ij} < 5$ που είναι γνωστή με το όνομα *διόρθωση Yates*:

$$Q_{\text{Yates}} = \frac{\left(|\det A| - \frac{n}{2}\right)^2 \cdot n}{r_1 r_2 c_1 c_2}.$$

Η χρησιμοποίηση του Q_{Yates} δίνει λιγότερα στατιστικά σημαντικά αποτελέσματα, ισχυροποιώντας τον έλεγχο όταν ικανοποιούνται οι παραπάνω συνθήκες. Επομένως, αναμένουμε $Q_{\text{Yates}} < Q_n$ όταν εφαρμόζεται σωστά.

Παράδειγμα 5.14: Θέλουμε να ελέγξουμε αν υπάρχει στατιστικά σημαντική σχέση ανάμεσα στον εμβολιασμό και την προσβολή από μία νόσο. Πήραμε ένα δείγμα μεγέθους $n = 81$ και τα αποτελέσματα συνοψίζονται στον παρακάτω τετράπτυχο πίνακα:

		Νόσος		
	O_{ij}	Παρόν	Απόν	r_i
Εμβολιασμός	Παρόν	12	33	45
	Απόν	18	18	36
	c_j	30	51	81

Αν $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 12 & 33 \\ 18 & 18 \end{pmatrix}$, τότε $Q_n = \dfrac{(\det A)^2 \cdot n}{r_1 r_2 c_1 c_2} = 4.67 > \chi^2_{1,0.05} = 3.84.$

Όμως,

$$Q_{\text{Yates}} = 3.71 < 3.84.$$

και έτσι ενώ *χωρίς διόρθωση* το αποτέλεσμα του ελέγχου βγαίνει *στατιστικά σημαντικό* σε ε.σ.σ. 0.05, υποδηλώνοντας στατιστική εξάρτηση μεταξύ εμβολιασμού και νόσου, *με διόρθωση Yates* το αποτέλεσμα βγαίνει *μη στατιστικά σημαντικό*.

5.8 Ιστορικά Στοιχεία - Περαιτέρω Μελέτη

Σχημα 5.1: Karl Pearson, 1857-1936: Βρετανός μαθηματικός και βιοστατιστικός, θεωρείται ο πατέρας της Μαθηματικής Στατιστικής. Ήταν ο άνθρωπος που ίδρυσε το πρώτο Τμήμα Στατιστικής παγκοσμίως, στο πανεπιστήμιο University College London, το 1911. Μαζί κυρίως με τον Ronald Fisher, και τις εργασίες του τελευταίου τη δεκαετία 1915-1925 έθεσαν τα θεμέλια της μοντέρνας στατιστικής επιστήμης.

Ασκήσεις

5.8.1 Στα πλαίσια μίας σύντομης έρευνας, ζητήθηκε από 246 φοιτητές να καταγράψουν το τελευταίο ψηφίο του κινητού τους τηλεφώνου. Οι απόλυτες συχνότητες των απαντήσεων ήταν 16, 20, 32, 23, 28, 24, 20, 28, 28 και 27 αντίστοιχα. Να εξετάσετε σε ε.σ.σ. $a = 0.05$ αν τα διαφορετικά ψηφία εμφανίζονται ισοπίθανα.

5.8.2 Μία διαφημιστική εταιρεία διεξάγει έρευνα σε τρεις διαφορετικές ηλικιακές ομάδες I, II, III για την προτίμηση ανάμεσα σε τέσσερεις διαφημίσεις, A, B, Γ, Δ. Τα αποτελέσματα της έρευνας παρουσιάζονται στον παρακάτω πίνακα.

Πινακας 5.17: Δεδομένα Άσκησης ??

O_{ij}	Α	Β	Γ	Δ	$n_{i\bullet}$
Έρευνα 1	21	48	19	12	100
Έρευνα 2	103	20	39	38	200
Έρευνα 3	102	44	126	28	300
$n_{\bullet j}$	226	112	184	78	500

Ζητείται να διερευνηθεί αν υπάρχει στατιστικά σημαντική διαφορά στις απαντήσεις των 3 δειγμάτων σε ε.σ.σ. $\alpha = 0.05$.

5.8.3 Διδάσκουμε τρεις διαφορετικές μεθόδους επίλυσης (I, II, III) του ίδιου προβλήματος σε μαθητές ηλικιών Α, Β, Γ, Δ και καταγράφουμε τη μέθοδο που κάθε μαθητής βρήκε πιο κατανοητή. Ο παρακάτω πίνακας συνάφειας παρουσιάζει τα δεδομένα της έρευνας.

Πινακας 5.18: Δεδομένα Άσκησης ??

O_{ij}	Α	Β	Γ	Δ	$n_{i\bullet}$
I	243	56	135	66	500
II	116	114	73	97	400
III	19	26	19	36	100
$n_{\bullet j}$	378	196	227	199	1000

Να μελετηθεί η ανεξαρτησία των δύο μεταβλητών σε ε.σ.σ. $\alpha = 0.05$.

5.8.4 Χρησιμοποιώντας τα δεδομένα του παραδείγματος 5.6, να ελέγξετε σε ε.σ.σ. $\alpha = 5\%$ αν τα δύο δείγματα προέρχονται από την ίδια κατανομή.

Πινακας 5.19: Δεδομένα Παραδείγματος ??

O_{ij}	0	1	$n_{i\bullet}$
Νόμισμα 1	18	12	30
Νόμισμα 2	20	10	30
$n_{\bullet j}$	38	22	60

Μέρος III

ΠΑΡΑΡΤΗΜΑΤΑ

ΟΔΗΓΙΕΣ ΓΙΑ ΣΥΓΓΡΑΦΕΙΣ XƎLaTeX

ΚΑΛΛΙΠΟΣ 2021
ISBN 978-618-85370-X-X