Κεφάλαιο 1

Χώροι Μέτρου και σ-άλγεβρες

1.1 Μοντελοποίηση των Τυχαίων Πειραμάτων – Πιθανότητα

Ας υποθέσουμε ότι κατά την εκτέλεση ενός πειράματος τύχης τα δυνατά εξαγόμενα είναι τα στοιχεία ενός συνόλου Ω, π.χ., ας πούμε ότι αυτά είναι τα $\omega_1, \ldots, \omega_\nu$. Το σύνολο των δυνατών εξαγομένων,

$$\Omega = \{\omega_1, \ldots, \omega_\nu\},$$

ονομάζεται **δειγματικός χώρος** (sample space). Θεωρούμε ότι κάθε φορά που εκτελείται το πείραμα, ένα (και μόνο ένα) από τα ω_j, $1 \leq j \leq \nu$, θα εμφανιστεί.

Ενώ λοιπόν ο δειγματικός χώρος Ω είναι απολύτως γνωστός, το εξαγόμενο ω_j, που εμφανίζεται κάθε φορά, είναι άγνωστο (απρόβλεπτο) στον παρατηρητή, και γίνεται γνωστό μετά την εκτέλεση του πειράματος.

Σε πρώτη προσέγγιση, ένα (τυχόν) υποσύνολο A του Ω καλείται **ενδεχόμενο** (event), και θεωρούμε ότι *το A εμφανίστηκε* όταν και μόνο όταν το εξαγόμενο $\omega_j \in A$. Όταν το $\omega_j \notin A$, τότε λέμε ότι *το A δεν εμφανίστηκε* ή, ισοδύναμα, λέμε ότι *εμφανίστηκε το συμπληρωματικό* (complement) *ενδεχόμενο* $A^c = \Omega \setminus A$. Θα δούμε αργότερα ότι υπάρχουν πολλές ενδιαφέρουσες περιπτώσεις στις οποίες, για τεχνικούς λόγους, δεν είναι δυνατόν να θεωρήσουμε ως ενδεχόμενο οποιοδήποτε υποσύνολο του δειγματικού χώρου.[1] Έτσι, θα εξαιρεθούν ορισμένα *προβληματικά* σύνολα, τα οποία, πάντως, δεν αντιστοιχούν σε ενδιαφέροντα ερωτήματα σχετικά με την έκβαση ενός πειράματος τύχης.

Η Θεωρία Πιθανοτήτων (Probability Theory) δημιουργήθηκε από την ανάγκη επιστημονικής μελέτης τυχαίων πειραμάτων. Είναι λογικό, αφού δεν είναι γνωστό αν θα εμφανιστεί ή όχι το ενδεχόμενο A, να προσπαθήσει κανείς να ορίσει μία ποσότητα, $\mathbb{P}(A)$, με την οποία θα ήταν δυνατόν να *μετρήσει* την εκ των προτέρων βεβαιότητα, όσον αφορά την εμφάνιση του A. Αυτός ο *βαθμός βεβαιότητος* ονομάζεται **πιθανότητα** (probability) του A, και, για προφανείς λόγους, μετριέται σε ποσοστά (από 0% έως 100%) ή, ισοδύναμα, σε απόλυτες μονάδες (από 0 έως 1).

Επειδή $\omega_j \in \Omega$ (ανεξαρτήτως του δείκτη j), θα πρέπει να

[1] π.χ., τυχαία εκλογή ενός αριθμού από τον δειγματικό χώρο $\Omega = (0, 1)$, ο οποίος είναι υπεραριθμήσιμος.

ορίσουμε $\mathbb{P}(\Omega) = 1$ (το Ω σίγουρα θα εμφανιστεί, και επομένως, οδηγούμαστε στο να του αποδώσουμε τον μέγιστο βαθμό βεβαιότητος). Παρομοίως, $\mathbb{P}(\varnothing) = 0$, αφού το $\varnothing$ δεν θα εμφανιστεί. Επίσης, θα ήταν λογικό να ισχύει $\mathbb{P}(A) \le \mathbb{P}(B)$, όταν[2] $A \subset B$, για τυχόντα ενδεχόμενα $A, B \in \mathcal{P}(\Omega)$, όπου $\mathcal{P}(\Omega)$ το σύνολο όλων των υποσυνόλων του Ω (**δυναμοσύνολο** του Ω), διότι ο βαθμός βεβαιότητος για ένα μικρότερο ενδεχόμενο θα πρέπει να είναι μικρότερος, ή το πολύ ίσος, του βαθμού βεβαιότητος ενός μεγαλύτερου ενδεχομένου. Επίσης, για το συμπληρωματικό ενδεχόμενο $A^c = \Omega \setminus A$ θα ήταν επιθυμητό να ισχύει η σχέση[3] $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$, αφού καθένα από τα A, A^c, εμφανίζεται όταν και μόνο όταν το άλλο δεν εμφανίζεται, διότι $A \cup A^c = \Omega$, και[4] $AA^c = \varnothing$.

Ο ορισμός του Laplace αντιστοιχεί στην περίπτωση της **κλασικής πιθανότητας**, κατά την οποία

---

[2]Ας σημειωθεί ότι στο παρόν κείμενο θα χρησιμοποιείται ο συμβολισμός $\subset$ για να δηλωθεί η σχέση *υποσύνολο* και όχι η σχέση *γνήσιο υποσύνολο*. Έτσι, οι δύο σχέσεις $A \subset B$ και $B \subset A$ ισοδυναμούν με την $A = B$, και, ως συνήθως, οι σχέσεις $A \subset B$ και $B \supset A$ δηλώνουν το ίδιο ακριβώς πράγμα.

[3]Σε όλο το κείμενο, το συμπληρωματικό ενός ενδεχομένου (συνόλου) A θα δηλώνεται με A^c. Εναλλακτικά, αν υπάρχει κίνδυνος σύγχυσης, θα χρησιμοποιείται και ο συμβολισμός $\Omega \setminus A$.

[4]Η τομή $A_1 \cap A_2 \cap \cdots \cap A_n$ των συνόλων $A_1, A_2, \ldots, A_n$, δηλαδή το σύνολο των στοιχείων που ανήκουν ταυτόχρονα στα $A_1, A_2, \ldots, A_n$, θα δηλώνεται και ως $A_1 A_2 \cdots A_n$ ή $\cap_{i=1}^n A_i$. Έτσι, $AB \equiv A \cap B$, $ABC \equiv A \cap B \cap C$, κ.ο.κ. Η ένωση $A_1 \cup A_2 \cup \cdots \cup A_n$ των συνόλων $A_1, A_2, \ldots, A_n$, δηλαδή το σύνολο των στοιχείων που ανήκουν σε τουλάχιστον ένα από τα $A_1, A_2, \ldots, A_n$, θα δηλώνεται και ως $\cup_{i=1}^n A_i$. Για τυχούσα συλλογή συνόλων $\{A_i, i \in I\}$, όπου I είναι οποιοδήποτε μη κενό σύνολο δεικτών, η ένωση $\cup_{i \in I} A_i$ παριστάνει το σύνολο εκείνων των στοιχείων που ανήκουν στο A_i για τουλάχιστον έναν δείκτη $i \in I$, ενώ η τομή $\cap_{i \in I} A_i$ παριστάνει το σύνολο εκείνων των στοιχείων που ανήκουν στο A_i για κάθε δείκτη $i \in I$.

$$\mathbb{P}(A) = \frac{\nu(A)}{\nu(\Omega)} = \frac{\nu(A)}{\nu}, \quad A \in \mathcal{P}(\Omega),$$

όπου $\nu(A) =$ *πλήθος στοιχείων του* A. Ο ορισμός αυτός ικανοποιεί τις παραπάνω αρχές, αλλά, επιπροσθέτως, υπονοεί ότι οι δυνατές εκβάσεις του τυχαίου πειράματος διέπονται από προφανή συμμετρία. Έτσι, τα στοιχειώδη ενδεχόμενα $\{\omega_1\}, \ldots, \{\omega_\nu\}$ είναι *ισοπίθανα* (έχουν πιθανότητα $1/\nu$). Αυτή η θεώρηση των πραγμάτων είναι ικανοποιητική όταν μιλάμε για συνήθη τυχερά παιχνίδια (π.χ., οι 6 πλευρές ενός συνηθισμένου ζαριού, οι 2 όψεις ενός αμερόληπτου νομίσματος, ή τα 52 φύλα μιας τράπουλας, αποτελούν στοιχειώδη ενδεχόμενα που αντιστοιχούν στους χώρους $\Omega_1 = \{\omega_1, \ldots, \omega_6\}$, $\Omega_2 = \{\omega_1, \omega_2\}$, και $\Omega_3 = \{\omega_1, \ldots, \omega_{52}\}$, αντίστοιχα). Τι γίνεται όμως όταν ο χώρος Ω έχει άπειρα στοιχεία, ή όταν τα στοιχειώδη ενδεχόμενα δεν είναι ισοπίθανα;

Μία λογική γενίκευση του ορισμού του Laplace, που στην ουσία περιγράφει όλους τους διακριτούς χώρους πιθανότητας (discrete probability spaces), είναι η εξής:

Έστω

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_n, \ldots\},$$

και $\{p_n, n \geq 1\}$ τυχούσα ακολουθία μη αρνητικών αριθμών, με

$$\sum_{n=1}^{\infty} p_n = 1.$$

Ορίζουμε τότε

$$\mathbb{P}(A) = \sum_{j:\ \omega_j \in A} p_j, \quad \text{για } A \in \mathcal{P}(\Omega). \tag{1.1}$$

Προφανώς, η συνολοσυνάρτηση $\mathbb{P} : \mathcal{P}(\Omega) \longrightarrow [0,1]$, αποδίδει πιθανότητα σε όλα τα $A \in \mathcal{P}(\Omega)$, ο δε ορισμός του Laplace ισοδυναμεί με την ειδική περίπτωση

$$p_j = \begin{cases} 1/\nu, & j = 1, 2, \ldots, \nu, \\ 0, & j = \nu+1, \nu+2, \ldots, \end{cases}$$

με την «φυσιολογική» αφαίρεση των δειγματικών σημείων ω_j για τα οποία $p_j = 0$.

Στην ουσία, η (1.1) επιτρέπει να ορίσουμε αυθαίρετα τις πιθανότητες των στοιχειωδών ενδεχομένων $\{\omega_j\}$, $j = 1, 2, \ldots$, και στην συνέχεια να υπολογίσουμε την πιθανότητα τυχόντος ενδεχομένου $A \in \mathcal{P}(\Omega)$, ως άθροισμα των πιθανοτήτων των στοιχειωδών ενδεχομένων που απαρτίζουν το A.

Το παραπάνω μοντέλο, αν και φαινομενικά είναι ικανοποιητικό, εντούτοις **αδυνατεί** να περιγράψει τα περισσότερα φαινόμενα που εμφανίζονται στην πράξη, διότι, συνήθως, τα τελευταία έχουν **συνεχή** χαρακτήρα. Για παράδειγμα, ο ακριβής χρόνος (σε second), που θα κάνει κάποιος αθλητής σε μία κούρσα 100m, είναι το εξαγόμενο ενός πειράματος τύχης, που ανήκει στον δειγματικό χώρο

$$\Omega = [0, +\infty],$$

ο οποίος, προφανώς, δεν μπορεί να γραφεί στην μορφή $\Omega = \{\omega_1, \omega_2, \ldots, \omega_\nu, \ldots\}$, διότι είναι υπεραριθμήσιμος. Το ίδιο ακριβώς φαινόμενο συναντάμε στο απλό πείραμα **τυχαίας εκλογής ενός αριθμού** από τον $\Omega = (0, 1)$. Αν επιχειρήσουμε να ορίσουμε «σημειακές» πιθανότητες στα στοιχεία ενός υπεραριθμήσιμου δειγματικού χώρου Ω (όπως στα προηγούμενα παραδείγματα), θα διαπιστώσουμε ότι είναι αδύνατον να «δουλέψει» η λογική της (1.1), εκτός αν δώσουμε πιθανότητα 0 *σχεδόν σε όλα* τα $\omega \in \Omega$, και συγκεκριμένα, σε όλα τα $\omega \in \Omega$ εκτός, ίσως, ενός αριθμησίμου πλήθους.

Πράγματι, ας θεωρήσουμε έναν υπεραριθμήσιμο χώρο Ω, και ένα υπεραριθμήσιμο υποσύνολό του, W, τέτοιο ώστε $\mathbb{P}(\{\omega\}) = p(\omega) > 0$ για κάθε $\omega \in W$, και $\mathbb{P}(\{\omega\}) = 0$ για κάθε $\omega \in W^c$. Ας υποθέσουμε ότι η πιθανότητα του A έχει οριστεί ως

$$\mathbb{P}(A) = \sum_{\omega \in A} p(\omega), \quad \text{για κάθε } A \in \mathcal{P}(\Omega),$$

όπως στην (1.1), όπου το σύμβολο $\sum_{\omega \in A}$ θα πρέπει να ερμηνευτεί ως *άθροισμα για όλα τα* $\omega \in A$. [Τέτοια «άθροιση», φυσικά, δεν έχει οριστεί αυστηρά, αλλά ας παραβλέψουμε προς στιγμήν αυτό το εμπόδιο και, χάριν απλότητος, ας θεωρήσουμε ότι διέπεται από τους συνήθεις νόμους των πεπερασμένων και απείρων αθροισμάτων μη αρνητικών αριθμών.][5] Επιχειρώντας την διαμέριση

[5] Ο συνήθης αυστηρός ορισμός που βρίσκει κανείς στην βιβλιογραφία είναι

$$W = \bigcup_{n=1}^{\infty} A_n,$$

με $A_n = \left\{\omega \in \Omega : \frac{1}{n} \leq p(\omega) < \frac{1}{n-1}\right\}$, $n = 1, 2, \ldots$ ($A_1 = \{\omega \in \Omega : p(\omega) \geq 1\}$), έχουμε προφανώς ότι $A_n \subset W$, και συνεπώς θα πρέπει

$$1 \geq \mathbb{P}(W) \geq \mathbb{P}(A_n) = \sum_{\omega \in A_n} p(\omega) \geq \sum_{\omega \in A_n} \frac{1}{n} = \frac{\nu(A_n)}{n},$$

όπου $\nu(A_n)$ = *πλήθος στοιχείων του* A_n. Άρα, $\nu(A_n) \leq n$, και επειδή

$$W = \bigcup_{n=1}^{\infty} A_n,$$

έπεται ότι το W είναι αριθμήσιμο, ως αριθμήσιμη ένωση των πεπερασμένων συνόλων A_n. Συνεπώς, το W **δεν μπορεί** να είναι υπεραριθμήσιμο, οπότε, η απόδοση «σημειακών» πιθανοτήτων στον Ω, ουσιαστικά, τον συρρικνώνει στον αριθμήσιμο W.

Μία εναλλακτική προσέγγιση του υπολογισμού της πιθανότητας ενός ενδεχομένου A δόθηκε από τον von Mises. Επειδή η προσέγγιση του von Mises γίνεται από την Στατιστική/Εμπειρική οπτική γωνία, καλείται **εμπειρική** (ή στατιστική) **πιθανότητα**, και περιγράφεται ως εξής: Ας θεωρήσουμε τον χώρο Ω, και τυχόν υποσύνολό του A, για το οποίο θέλουμε να υπολογίσουμε την πιθανότητα $\mathbb{P}(A)$. Επαναλαμβάνοντας το πείραμα n φορές, μπορούμε να παρατηρήσουμε πόσες φορές

$$\sum_{\omega \in W} p(\omega) = \sup_{N} \sum_{\omega \in N} p(\omega),$$

όπου το supremum λαμβάνεται ως προς όλα τα πεπερασμένα σύνολα N, με $N \subset W$.

εμφανίστηκε το A, και πόσες όχι. Αν $k_n(A)$ είναι το πλήθος φορών που εμφανίστηκε το A, τότε ορίζουμε

$$\mathbb{P}(A) = \lim_{n\to\infty} \frac{k_n(A)}{n},$$

με την προϋπόθεση ότι το όριο αυτό υπάρχει.

Τα μειονεκτήματα του ορισμού του von Mises είναι προφανή: Δεν γνωρίζουμε εάν υπάρχει το όριο, αλλά ακόμα και αν υποθέσουμε ότι υπάρχει, δεν έχουμε τρόπο να το υπολογίσουμε (μπορούμε μόνο να το προσεγγίσουμε εκτελώντας το πείραμα πολλές φορές), και, τέλος, δεν ξέρουμε εάν το όριο αυτό, εφ΄ όσον υπάρχει, θα είναι κοινό για όλους τους παρατηρητές που θα επαναλάβουν το πείραμα πολλές φορές. Μία, ακόμα, «τεχνική» δυσκολία, οφείλεται στο γεγονός ότι ενδέχεται να μην μπορούμε να εκτελέσουμε το πείραμα πολλές φορές. Για παράδειγμα, θα ήταν αδιανόητο (και επιστημονικά εσφαλμένο) να υποχρεώσουμε τον αθλητή των 100m, του προηγουμένου παραδείγματος, να τρέξει $n = 1000$ φορές την διαδρομή, προκειμένου να υπολογίσουμε, κατά προσέγγιση, την $\mathbb{P}(A)$ για $A = [10, 11]$, δηλαδή την πιθανότητα να φέρει χρόνο μεταξύ 10 και 11 sec.

1.2 Αναγκαιότητα Αυστηρού Ορισμού της Πιθανότητας

Έχοντας υπόψιν τα προηγούμενα, θεωρούμε το πείραμα τυχαίας εκλογής ενός σημείου ω, από τον δειγματικό χώρο

$$\Omega = (0, 1),$$

κατά τρόπο ώστε όλα τα σημεία του Ω να έχουν «την ίδια πιθανότητα να εκλεγούν». [Φυσικά, η έκφραση αυτή είναι καταχρηστική, διότι εύκολα αναγνωρίζεται ότι θα πρέπει $\mathbb{P}(\{\omega\}) = 0$, για κάθε $\omega \in \Omega$.] Σημειώνουμε στο σημείο αυτό ότι η επιλογή του συγκεκριμένου πειράματος έγινε για λόγους απλότητος, και ότι η προσέγγιση που θα επιχειρήσουμε γενικεύεται και σε οποιονδήποτε άλλον χώρο Ω, με την προϋπόθεση ότι ικανοποιούνται ανάλογες απαιτήσεις.

Η φυσικότερη ερμηνεία στην έννοια «ισοπίθανα» θα ήταν να αποδίδαμε πιθανότητα

$$\mathbb{P}(I) = b - a = \mu\acute{\eta}\kappa o\varsigma\ \tau o\upsilon\ I,$$

σε κάθε διάστημα της μορφής $I = (a, b) \subset (0, 1)$, διότι $\mathbb{P}(\Omega) = 1 =$ *μήκος του* Ω, και άρα, το τυχόν διάστημα I, που καταλαμβάνει, π.χ., το 50% του $(0, 1)$, θα πρέπει να έχει πιθανότητα $\frac{1}{2}$. Έτσι, καταλήγουμε φυσιολογικά στα κάτωθι ερωτήματα:

(Α) Μπορούμε να αποδώσουμε πιθανότητα (μήκος) σε όλα τα σύνολα του $\mathcal{P}(\Omega)$;

(Β) Αν ναι, πώς θα μετρήσουμε το μήκος τυχόντος συνόλου που δεν είναι διάστημα;

(Γ) Αν όχι, σε ποια μπορούμε και σε ποια δεν μπορούμε;

Αυτά τα προβλήματα, καθώς και άλλα σχετικά με την θεωρία ολοκλήρωσης, μελετήθηκαν από πολύ σπουδαίους μαθηματικούς του 19ου και του 20ου αιώνα, μεταξύ αυτών και οι Borel, Lebesgue, Καραθεοδωρή, κ.α., και τα αποτελέσματά τους απεδείχθησαν καθοριστικά για τα σύγχρονα μαθηματικά και την θεωρία πιθανοτήτων. Αναπόφευκτα λοιπόν, μία ολοκληρωμένη αξιωματική θεμελίωση των πιθανοτήτων θα πρέπει να αντλήσει πληροφορίες και αποτελέσματα από την θεωρία μέτρου.

Επιχειρούμε μία διαισθητική προσέγγιση της μεθόδου υπολογισμού του μήκους (μέτρου) ενός συνόλου, που χρησιμοποιήθηκε από τον Καραθεοδωρή – ο αναγνώστης που επιθυμεί περισσότερη αυστηρότητα και λεπτομέρειες παραπέμπεται στο βιβλίο των Κουμουλλή και Νεγρεπόντη (1988), Κεφ. 1–4.

Ορισμός 1.1 (εξωτερικό μέτρο) Έστω Ω τυχόν, μη κενό, σύνολο, και φ μία μη αρνητική συνολοσυνάρτηση, ορισμένη στον $\mathcal{P}(\Omega)$, $\varphi : \mathcal{P}(\Omega) \longrightarrow [0, +\infty]$. Η φ θα λέγεται εξωτερικό μέτρο, εάν ικανοποιεί τις εξής τρεις ιδιότητες:

(i) $\varphi(\varnothing) = 0$.

(ii) Για κάθε $A, B \in \mathcal{P}(\Omega)$ με $A \subset B$, έπεται ότι $\varphi(A) \leq \varphi(B)$.
[Δηλαδή η φ είναι **μονότονη**.]

(iii) Αν η $\{A_n,\ n \geq 1\}$ είναι τυχούσα ακολουθία στον $\mathcal{P}(\Omega)$,

τότε
$$\varphi\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \varphi(A_n).$$
[Δηλαδή η φ είναι **αριθμήσιμα υποπροσθετική** ή **σ-υποπροσθετική**.]

Ορισμός 1.2 (εξωτερικό μέτρο του Lebesgue στον (0,1)) Το εξωτερικό μέτρο του Lebesgue (στο διάστημα $(0,1)$) είναι η συνολοσυνάρτηση $\lambda^* : \mathcal{P}((0,1)) \longrightarrow [0,1]$, που ορίζεται ως
$$\lambda^*(A) = \inf\left\{\sum_{n=1}^{\infty}(b_n - a_n),\ A \subset \bigcup_{n=1}^{\infty}(a_n, b_n),\ 0 \leq a_n \leq b_n \leq 1,\ n = 1, 2, \ldots\right\},$$
για κάθε $A \subset (0,1)$.

Επειδή κάθε $A \in \mathcal{P}((0,1))$ μπορεί να καλυφθεί από την ένωση διαστημάτων $(0,1) \cup \varnothing \cup \varnothing \cup \cdots$ (όπου, π.χ., $\varnothing = (\frac{1}{2}, \frac{1}{2})$), και $\sum_{n=1}^{\infty}(b_n - a_n) = 1 + 0 + 0 + \ldots = 1$, έπεται ότι το σύνολο, του οποίου λαμβάνουμε το infimum, περιέχει το 1 (και προφανώς περιέχει μόνο μη αρνητικούς αριθμούς). Άρα, $\lambda^*(A) \in [0,1]$ για κάθε $A \subset (0,1)$.

Είναι διαισθητικά προφανές ότι ισχύει η σχέση
$$\lambda^*((a,b)) = \lambda^*((a,b]) = \lambda^*([a,b)) = \lambda^*([a,b]) = b - a,$$
για οποιοδήποτε διάστημα $I \subset (0,1)$, δηλ. της μορφής $I = (a,b)$, ή $I = (a,b]$, ή $I = [a,b)$, ή $I = [a,b]$. [Με την παραδοχή ότι $(a,a) = (a,a] = [a,a) = \varnothing$, $[a,a] = \{a\}$.] Δηλαδή ισχύει ότι

$$\lambda^*(I) = \mu\acute{\eta}\kappa o\varsigma\ \tau o\upsilon\ I,$$

για κάθε διάστημα $I \subset (0,1)$. Επίσης, μπορεί να δειχθεί (και είναι επίσης διαισθητικά προφανές) ότι αν το $A \subset (0,1)$ είναι ένωση, ξένων ανά δύο, ανοικτών, διαστημάτων, δηλ. είναι της μορφής

$$A = \bigcup_{n=1}^{\infty} (a_n, b_n), \tag{1.2}$$

με $0 \leq a_n \leq b_n \leq 1$, και $(a_n, b_n) \cap (a_m, b_m) = \varnothing$ για $m \neq n$, τότε

$$\lambda^*(A) = \sum_{n=1}^{\infty} (b_n - a_n).$$

Είναι γνωστό ότι κάθε ανοικτό σύνολο $U \subset (0,1)$ μπορεί να γραφεί στην μορφή (1.2), και συνεπώς, το $\lambda^*(U)$ παριστάνει το γενικευμένο μήκος (μέτρο) τυχόντος ανοικτού συνόλου U. Τελικά αποδεικνύεται[6] ότι η $\lambda^* : \mathcal{P}((0,1)) \to [0,1]$ είναι εξωτερικό μέτρο, σύμφωνα με τον προηγούμενο ορισμό των εξωτερικών μέτρων, και επομένως,

$$\begin{aligned} 1 &= \lambda^*((0,1)) = \lambda^*(A \cup A^c \cup \varnothing \cup \varnothing \cup \cdots) \\ &\leq \lambda^*(A) + \lambda^*(A^c) + \lambda^*(\varnothing) + \lambda^*(\varnothing) + \ldots \\ &= \lambda^*(A) + \lambda^*(A^c). \end{aligned}$$

Άρα, για τυχόν $A \subset (0,1)$, ισχύει ότι

$$\lambda^*(A) + \lambda^*(A^c) \geq 1, \tag{1.3}$$

[6] βλ. Κουμουλλής και Νεγρεπόντης (1988)

ενώ εμείς θα αναμέναμε ότι ισχύει πάντα η ισότητα στην (1.3). Αυτό, δυστυχώς, **δεν είναι αληθές** για όλα τα σύνολα του $\mathcal{P}((0,1))$, και ακριβώς αυτά που ικανοποιούν την (1.3) ως ισότητα (και, άρα, είναι «καλά» σύνολα), ονομάζονται Lebesgue μετρήσιμα σύνολα (ή σύνολα Lebesgue, ή μετρήσιμα σύνολα), και είναι θεμελιώδη στην θεωρία μέτρου.

Ορισμός 1.3 (Lebesgue μετρήσιμο σύνολο) Ένα σύνολο $A \subset (0,1)$ θα ονομάζεται Lebesgue μετρήσιμο όταν

$$\lambda^*(A) + \lambda^*(A^c) = 1.$$

Η οικογένεια των Lebesgue μετρησίμων συνόλων του $(0,1)$ συμβολίζεται με

$$\mathcal{M}_{\lambda^*}\left((0,1)\right),$$

ή, απλά, $\mathcal{M}_{\lambda^*}$, όταν το διάστημα $(0,1)$ εννοείται.

Σύμφωνα με την παραπάνω ανάλυση, η $\mathcal{M}_{\lambda^*}$ περιέχει το $\varnothing$ και το $\Omega = (0,1)$, καθώς και κάθε ανοικτό υποσύνολο του $(0,1)$, και, προφανώς, περιέχει και τα συμπληρώματα αυτών, αφού $(A^c)^c = A$. Τελικά, η $\mathcal{M}_{\lambda^*}$ είναι μία οικογένεια υποσυνόλων του $(0,1)$, για την οποία ισχύει το εξής θεώρημα.

Θεώρημα 1.4 (Καραθεοδωρή)[7] Για την $\mathcal{M}_{\lambda^*}$ ισχύουν οι εξής τρεις ιδιότητες:

(i) $\varnothing \in \mathcal{M}_{\lambda^*}$.

[7] βλ. Κουμουλλής και Νεγρεπόντης (1988), Θεώρημα 3.9

(ii) Αν $A \in \mathcal{M}_{\lambda^*}$ τότε $A^c \in \mathcal{M}_{\lambda^*}$.
[κλειστότητα ως προς τα συμπληρώματα]

(iii) Αν η $\{A_n,\ n \geq 1\}$ είναι τυχούσα ακολουθία στην $\mathcal{M}_{\lambda^*}$, τότε
$$A = \bigcup_{n=1}^{\infty} A_n \in \mathcal{M}_{\lambda^*}.$$
[κλειστότητα ως προς τις αριθμήσιμες ενώσεις, ή σ-κλειστότητα]

Μία οικογένεια υποσυνόλων του Ω, που ικανοποιεί τις συνθήκες (i), (ii), και (iii), του Θεωρήματος 1.4, ονομάζεται **σ-άλγεβρα** (σ-field), σύμφωνα με τον επόμενο ορισμό:

Ορισμός 1.5 (σ-άλγεβρα) Έστω Ω τυχόν, μη κενό, σύνολο. Μία οικογένεια $\mathcal{A}$, υποσυνόλων του Ω [δηλ. $\mathcal{A} \subset \mathcal{P}(\Omega)$], καλείται **σ-άλγεβρα** στον Ω (ή σ-άλγεβρα επί του Ω ή, απλά, σ-άλγεβρα, όταν το Ω εννοείται), όταν η $\mathcal{A}$ περιέχει το $\varnothing$, και είναι κλειστή ως προς τα συμπληρώματα και τις αριθμήσιμες ενώσεις, δηλ. όταν:

(i) $\varnothing \in \mathcal{A}$.

(ii) Αν $A \in \mathcal{A}$ τότε $A^c \in \mathcal{A}$. [κλειστότητα ως προς τα συμπληρώματα]

(iii) Αν η $\{A_n,\ n \geq 1\}$ είναι τυχούσα ακολουθία στην $\mathcal{A}$, τότε $A = \bigcup\limits_{n=1}^{\infty} A_n \in \mathcal{A}$.
[κλειστότητα ως προς τις αριθμήσιμες ενώσεις, ή σ-κλειστότητα]

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

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

Τελικά διαπιστώνουμε ότι το Θεώρημα 1.4 του Καραθεοδωρή απλώς μας διαβεβαιώνει ότι «η $\mathcal{M}_{\lambda^*}$ είναι μία σ-άλγεβρα».

Παράδειγμα 1.6

(i) Για τυχόν μη κενό σύνολο Ω, οι οικογένειες $\mathcal{A}_1 = \{\varnothing, \Omega\}$ και $\mathcal{A}_2 = \mathcal{P}(\Omega)$ είναι σ-άλγεβρες, η μικρότερη και η μεγαλύτερη σ-άλγεβρα του Ω, αντίστοιχα. Προφανώς, μία (τυχούσα) σ-άλγεβρα $\mathcal{A}$, επί του Ω, ικανοποιεί την

$$\mathcal{A}_1 \subset \mathcal{A} \subset \mathcal{A}_2.$$

(ii) Έστω $\mathbb{N} = \{1, 2, \ldots\}$ και

$$\mathcal{A} = \{A \subset \mathbb{N} : A \textit{ πεπερασμένο, ή } A^c \textit{ πεπερασμένο}\}.$$

[Το A είναι «πεπερασμένο» όταν περιέχει πεπερασμένα το πλήθος στοιχεία, ή είναι το $\varnothing$.]

Εύκολα διαπιστώνεται ότι $\varnothing \in \mathcal{A}$, ότι αν $A \in \mathcal{A}$ τότε $A^c \in \mathcal{A}$ και, τέλος, ότι αν $A, B \in \mathcal{A}$ τότε $A \cup B \in \mathcal{A}$. Όμως η $\mathcal{A}$ δεν είναι σ-κλειστή, αφού, π.χ.,

$$\bigcup_{n=1}^{\infty}\{2n\} = \{2,4,6,\ldots\} \notin \mathcal{A},$$

ενώ τα σύνολα $A_n = \{2n\} \in \mathcal{A}$, για κάθε $n = 1, 2, \ldots$.
$\square$

Η $\mathcal{A}$ του Παραδείγματος 1.6(ii) είναι, απλά, **άλγεβρα**, σύμφωνα με τον εξής ορισμό:

Ορισμός 1.7 (άλγεβρα) Έστω Ω τυχόν μη κενό σύνολο. Μία οικογένεια υποσυνόλων, έστω $\mathcal{A}$, του μη κενού συνόλου Ω, λέγεται **άλγεβρα** στον Ω, όταν:

(i) $\varnothing \in \mathcal{A}$.

(ii) Αν $A \in \mathcal{A}$ τότε $A^c \in \mathcal{A}$.

(iii) Αν $A, B \in \mathcal{A}$ τότε $A \cup B \in \mathcal{A}$.
[κλειστότητα ως προς τις **πεπερασμένες ενώσεις**, ή **πεπερασμένη κλειστότητα**]

Όπως και στις σ-άλγεβρες, εύκολα διαπιστώνεται ότι η (i) ισοδυναμεί με την (i)$'$: $\Omega \in \mathcal{A}$, ενώ η (iii) ισοδυναμεί με την (iii)$'$: Αν $A, B \in \mathcal{A}$ τότε $AB \in \mathcal{A}$. [Διότι $AB = (A^c \cup B^c)^c$.]

Σε μία άλγεβρα μπορεί εύκολα να δειχθεί, με επαγωγή, ότι αν $A_1, \ldots, A_n \in \mathcal{A}$ τότε $A_1 \cup \cdots \cup A_n \in \mathcal{A}$ (ισοδύναμα, $\bigcap_{j=1}^{n} A_j \in \mathcal{A}$), δηλαδή η $\mathcal{A}$ είναι κλειστή ως προς τις **πεπερασμένες** ενώσεις (τομές). Προφανώς μία σ-άλγεβρα είναι πάντα άλγεβρα, διότι

$A \cup B = A \cup B \cup \varnothing \cup \varnothing \cup \cdots \in \mathcal{A}$, ενώ το αντίστροφο δεν ισχύει, όπως γίνεται αντιληπτό από το Παράδειγμα 1.6(ii).

Είναι φανερό ότι οι σ-άλγεβρες (και οι άλγεβρες, είναι η αλήθεια) έχουν κατάλληλη δομή για μία επαρκή αξιωματική θεμελίωση της Θεωρίας Πιθανοτήτων. Έτσι, φυσιολογικά, καταλήγουμε στον αξιωματικό ορισμό της πιθανότητας του Kolmogorov (1930).

Ορισμός 1.8 (χώρος πιθανότητας, Kolmogorov, 1930) Έστω Ω τυχόν, μη κενό, σύνολο (δειγματικός χώρος), και $\mathcal{A}$ μία σ-άλγεβρα στον Ω. Τότε, **οποιαδήποτε** μη αρνητική συνολοσυνάρτηση $\mathbb{P} : \mathcal{A} \longrightarrow [0,1]$ θα ονομάζεται **πιθανότητα** (probability), εάν ικανοποιεί τις εξής δύο ιδιότητες:

(i) $\mathbb{P}(\Omega) = 1$, και

(ii) Για κάθε ακολουθία $\{A_n,\ n \ge 1\}$ στην $\mathcal{A}$, με $A_iA_j = \varnothing$ για $i \neq j$,

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

[Δηλαδή η $\mathbb{P}$ είναι μία **αριθμήσιμα προσθετική (σ-προσθετική)** συνολοσυνάρτηση.]

Τα στοιχεία της $\mathcal{A}$, και μόνον αυτά, ονομάζονται **ενδεχόμενα** (events), και η τριάδα $(\Omega, \mathcal{A}, \mathbb{P})$ καλείται **χώρος πιθανότητας** (probability space).

Επειδή, προφανώς, $\Omega = \Omega \cup \varnothing \cup \varnothing \cup \cdots$, έπεται ότι $\mathbb{P}(\varnothing) = 0$. Άρα η συνολοσυνάρτηση $\mathbb{P}$ είναι, στην ουσία, ένα **μέτρο**[8] στον **μετρήσιμο χώρο** $(\Omega, \mathcal{A})$, με την ιδιότητα $\mathbb{P}(\Omega) = 1$.

Γενικά έχουμε τον παρακάτω ορισμό:

Ορισμός 1.9 (μετρήσιμος χώρος – μέτρο) Έστω Ω τυχόν, μη κενό, σύνολο, και $\mathcal{A}$ τυχούσα σ-άλγεβρα επί του Ω. Το ζεύγος $(\Omega, \mathcal{A})$ καλείται **μετρήσιμος χώρος**. Αν η μη αρνητική συνολοσυνάρτηση $\mu : \mathcal{A} \longrightarrow [0, +\infty]$ ικανοποιεί τις:

(i) $\mu(\varnothing) = 0$, και

(ii) Για κάθε ακολουθία $\{A_n,\ n \geq 1\}$, ξένων ανά δύο συνόλων στην $\mathcal{A}$ (δηλ. $A_i A_j = \varnothing$ για $i \neq j$),

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

[δηλαδή η συνολοσυνάρτηση μ είναι **σ-προσθετική**]

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

Χρήσιμος είναι και ο επόμενος ορισμός.

Ορισμός 1.10 (σ-πεπερασμένο, πεπερασμένο, πλήρες μέτρο) Έστω $(\Omega, \mathcal{A}, \mu)$ ένας χώρος μέτρου.

[8] Αφού ο Ορισμός 1.8 αποτελεί (πολύ) ειδική περίπτωση μέτρου (που προϋπήρχε για πολλά χρόνια – βλ. Ορισμό 1.9), δεν είναι καθόλου προφανής η **σπουδαιότητα του αξιωματικού ορισμού** του Kolmogorov. Ο αναγνώστης θα πρέπει να κάνει «υπομονή» μέχρι το Κεφ. ;; (νόμοι μεγάλων αριθμών), οπότε και θα λάμψει η αξία του ορισμού, μέσα από την ίδια του την **απλότητα**!

(i) Το μ καλείται **σ-πεπερασμένο**, αν υπάρχει ακολουθία $\{A_n, n \geq 1\}$ στην $\mathcal{A}$, με $\mu(A_n) < \infty$ για κάθε n, τέτοια ώστε $\Omega = \bigcup_{n=1}^{\infty} A_n$.

(ii) Το μ καλείται **πεπερασμένο** αν $\mu(\Omega) < \infty$.

(iii) Το μ καλείται **πλήρες** αν για κάθε $B \subset \Omega$ με την ιδιότητα:

$$\text{Υπάρχουν } A, C \in \mathcal{A}, \text{ με } A \subset B \subset C, \text{ και } \mu(C \setminus A) = 0,$$

[όπου $C \setminus A \stackrel{\text{ορ.}}{=\!=} CA^c$] έπεται ότι $B \in \mathcal{A}$.

[Και τότε, αναγκαστικά, $\mu(B) = \mu(A) = \mu(C)$.]

Σύμφωνα με τα παραπάνω, η τριάδα $(\Omega, \mathcal{A}, \mathbb{P})$ λέγεται **χώρος πιθανότητας**, όταν είναι ένας χώρος πεπερασμένου μέτρου, με $\mathbb{P}(\Omega) = 1$. Σε αυτήν την περίπτωση, το μέτρο $\mu = \mathbb{P}$ καλείται **μέτρο πιθανότητας** (probability measure).

Είναι φανερό ότι κάθε πεπερασμένο μέτρο (και συνεπώς, κάθε μέτρο πιθανότητας) είναι και σ-πεπερασμένο, διότι μπορούμε να γράψουμε $\Omega = \Omega \cup \varnothing \cup \varnothing \cup \cdots$, ενώ το αντίστροφο δεν ισχύει:

Παράδειγμα 1.11 Έστω $\Omega = \mathbb{N} = \{1, 2, \ldots\}$, $\mathcal{A} = \mathcal{P}(\mathbb{N})$, και $\mu(A) =$ *πλήθος στοιχείων του* A, για κάθε $A \in \mathcal{P}(\mathbb{N})$. Τότε ο $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$ είναι χώρος μέτρου. Το συγκεκριμένο αυτό μέτρο ονομάζεται **αριθμητικό μέτρο**, ή **μέτρο απαρίθμησης** (counting measure), στον μετρήσιμο χώρο $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Προφανώς,

$\mu(\mathbb{N}) = +\infty$ (άρα το μ δεν είναι πεπερασμένο), αλλά

$$\mathbb{N} = \bigcup_{n=1}^{\infty} A_n, \ \textit{όπου}\, A_n = \{n\}, \ \textit{με}\, \mu(A_n) = 1 < \infty, \ \textit{για κάθε}\, n.$$

[Και άρα το μ είναι σ-πεπερασμένο.] $\square$

Σχετικά με τον ορισμό της πληρότητας σημειώνουμε τα εξής: Όταν για τυχόν σύνολο $B \in \mathcal{P}(\Omega)$, υπάρχουν μετρήσιμα σύνολα $A, C \in \mathcal{A}$, με[9]

$$A \subset B \subset C, \quad \text{και} \quad \mu(C \setminus A) = 0,$$

θα ήταν λογικό να ίσχυε η ισότητα

$$\mu(A) = \mu(B) = \mu(C),$$

διότι τα A και C αποτελούν μία «κάτω» και «άνω» προσέγγιση του B, αντίστοιχα, και $\mu(A) = \mu(C)$. Παρατηρήστε ότι η σχέση $\mu(A) = \mu(C)$ προκύπτει εύκολα από τις ιδιότητες του μέτρου, αφού το C γράφεται ως αριθμήσιμη ξένη ένωση συνόλων της $\mathcal{A}$,

$$C = (C \setminus A) \cup A \cup \varnothing \cup \cdots,$$

και συνεπώς,

$$\mu(C) = \mu(C \setminus A) + \mu(A) + \mu(\varnothing) + \ldots = 0 + \mu(A) + 0 + \ldots = \mu(A).$$

Όμως, δεν γνωρίζουμε εξ' αρχής εάν $B \in \mathcal{A}$, δηλ. αν το B είναι μετρήσιμο. Αυτό εξαρτάται από την επιλογή της σ-άλγεβρας

[9] γιατί ισχύει ότι $C \setminus A \in \mathcal{A}$;

$\mathcal{A}$, **και όχι μόνο** από την εκλογή του μέτρου μ. Αν η σ-άλγεβρα είναι «αρκετά πλούσια», τότε θα περιέχει το B, και συνεπώς το $\mu(B)$ θα είναι καλά ορισμένο, και ίσο με το $\mu(C)$, διότι $C = B \cup (C \setminus B) \cup \varnothing \cup \cdots$, όπως πριν. Αν όμως η $\mathcal{A}$ δεν είναι «αρκετά πλούσια», τότε ένα B της παραπάνω μορφής ενδέχεται να μην ανήκει στην $\mathcal{A}$ (να μην είναι μετρήσιμο), και σε αυτήν την περίπτωση **δεν έχει νόημα** να μιλάμε για το μέτρο του, $\mu(B)$. Εκείνο που μπορούμε να κάνουμε είναι να επεκτείνουμε την σ-άλγεβρα $\mathcal{A}$ σε μία μεγαλύτερη, $\mathcal{A}_\mu$, ώστε η τελευταία να περιέχει **όλα** τα σύνολα B με την παραπάνω ιδιότητα. Πιο συγκεκριμένα, η κατασκευή αυτή γίνεται ως εξής:

Συναρτήσει του «παλιού» μέτρου μ ορίζουμε ένα «καινούριο» μέτρο, έστω $\bar{\mu}$, έτσι ώστε:

(i) $\bar{\mu}(B) = \mu(B)$, αν $B \in \mathcal{A}$, και

(ii) $\bar{\mu}(B) = \mu(C)$, αν $B \subset \Omega$, $B \notin \mathcal{A}$, αλλά για το B ικανοποιείται η ιδιότητα:

Υπάρχουν $A, C \in \mathcal{A}$, *με* $A \subset B \subset C$, *και* $\mu(C \setminus A) = 0$.

Η $\mathcal{A}_\mu$ περιέχει, εξ' ορισμού, όλα τα σύνολα της $\mathcal{A}$, καθώς, επίσης, και όλα τα σύνολα $B \in \mathcal{P}(\Omega) \setminus \mathcal{A}$, τα οποία ικανοποιούν το (ii).

Αποδεικνύεται ότι ο χώρος $(\Omega, \mathcal{A}_\mu, \bar{\mu})$ είναι πλήρης χώρος μέτρου, και καλά ορισμένος (μονοσήμαντα ορισμένος). Το μέτρο $\bar{\mu}$ ονομάζεται **πλήρωση** του μ στην $\mathcal{A}$. Η επέκταση αυτή της $\mathcal{A}$ στην $\mathcal{A}_\mu$ γίνεται κατά τρόπο μοναδικό, και η $\mathcal{A}_\mu$ είναι, πράγματι, σ-άλγεβρα. Τέλος, αποδεικνύεται ότι ο χώρος μέτρου $(\Omega, \mathcal{A}, \mu)$ είναι πλήρης αν και μόνο αν $\mathcal{A} = \mathcal{A}_\mu$.

Επανερχόμαστε τώρα στον μετρήσιμο χώρο $((0,1), \mathcal{M}_{\lambda^*})$, και διατυπώνουμε το εξής πολύ χρήσιμο θεώρημα.

Θεώρημα 1.12 (Καραθεοδωρή)[10] Η συνολοσυνάρτηση $\lambda : \mathcal{M}_{\lambda^*} \longrightarrow [0,1]$, με τύπο

$$\lambda(A) = \lambda^*(A), \quad \textit{για κάθε } A \in \mathcal{M}_{\lambda^*},$$

ορίζει έναν **πλήρη χώρο πιθανότητας**, $((0,1), \mathcal{M}_{\lambda^*}, \lambda)$, με την ιδιότητα

$$\lambda(I) = \textit{μήκος του } I,$$

για κάθε διάστημα $I \subset (0,1)$.

Με απλά λόγια, το μέτρο λ είναι ο περιορισμός της λ^* στην $\mathcal{M}_{\lambda^*}$. [Υπενθυμίζεται ότι η συνολοσυνάρτηση λ^*, ως εξωτερικό μέτρο, ορίζεται σε **ολόκληρο** το δυναμοσύνολο του $(0,1)$, $\mathcal{P}((0,1))$.] Το πλήρες μέτρο λ του Θεωρήματος 1.12 ονομάζεται **μέτρο Lebesgue**[11] (στον $(0,1)$), και η σ-άλγεβρα $\mathcal{M}_{\lambda^*}$ ονομάζεται **σ-άλγεβρα του Lebesgue**, ή **σ-άλγεβρα των Lebesgue μετρησίμων συνόλων** (του $(0,1)$).

---

[10] βλ. Κουμουλλής και Νεγρεπόντης (1988), Θεώρημα 3.9.

[11] Για την μοναδικότητα του μέτρου Lebesgue βλ. Παράδειγμα 2.21.

Συνοψίζοντας, το αρχικό πρόβλημά μας, περί ορισμού πιθανότητας στο τυχαίο πείραμα εκλογής σημείου από το $(0,1)$, επιδέχεται την εξής λύση:

(Α) Θα πρέπει να ονομάσουμε ενδεχόμενα **μόνο** τα μετρήσιμα σύνολα, δηλ. τα σύνολα της $\mathcal{M}_{\lambda^*}$.

(Β) Θα πρέπει να ορίσουμε την πιθανότητα ως $\mathbb{P}(A) = \lambda(A)$, για κάθε $A \in \mathcal{M}_{\lambda^*}$.

(Γ) Θα πρέπει **να μην ορίσουμε πιθανότητα** για τα υπόλοιπα σύνολα.

Με αυτόν τον τρόπο ικανοποιείται η απαίτηση $\mathbb{P}(A) = \textit{μήκος του } A$, για κάθε σύνολο $A \subset (0,1)$, για το οποίο μπορούμε να ορίσουμε το μήκος του, και το τυχαίο πείραμα εκλογής ενός σημείου ω από τον $(0,1)$ περιγράφεται από τον χώρο πιθανότητας

$$(\Omega, \mathcal{A}, \mathbb{P}) = ((0,1), \mathcal{M}_{\lambda^*}, \lambda)\,.$$

Η αλήθεια είναι ότι μέχρι στιγμής δεν έχουμε δώσει (και ούτε πρόκειται να το κάνουμε στην συνέχεια!) παράδειγμα συνόλου $A \subset (0,1)$, τέτοιο ώστε $A \notin \mathcal{M}_{\lambda^*}$. Τέτοια «παθολογικά» σύνολα υπάρχουν – έχουν βρεθεί από τους Vitali και Bernstein – αλλά η αλήθεια είναι ότι είναι πολύ δύσκολο, έως αδύνατο, να τα φανταστούμε και να τα κατασκευάσουμε.[12] Σε αυτά τα

[12] Έχει αποδειχθεί ότι η ύπαρξή τους εξαρτάται **ουσιωδώς** από το **Αξίωμα της Επιλογής** (Axiom of Choice).

σύνολα δεν μπορεί να οριστεί πιθανότητα, διότι

$$\lambda^*(A) + \lambda^*(A^c) > 1 = \lambda^*(A \cup A^c).$$

[Μην ξεχνάμε ότι η λ^* είναι υποπροσθετική.] Φυσικά, η εξαίρεση αυτών των «αφύσικων» συνόλων από την αξιωματική θεμελίωση των πιθανοτήτων δεν δημιουργεί ιδιαίτερα προβλήματα, αφού δεν συναντάμε «συχνά» τέτοιου είδους «ενδεχόμενα» ....

1.3 Πόσο πλούσια είναι η σ-άλγεβρα του Lebesgue;

Όπως είδαμε στα προηγούμενα, η σ-άλγεβρα $\mathcal{M}_{\lambda^*}$ περιέχει τα διαστήματα και τα ανοικτά σύνολα, και, φυσικά, περιέχει και τα κλειστά σύνολα, αφού, εξ ορισμού, μία σ-άλγεβρα είναι κλειστή ως προς τα συμπληρώματα. Είναι χρήσιμο να ορίσουμε την έννοια της **παραγόμενης σ-άλγεβρας**, για να γίνει πιο κατανοητή η υφή της $\mathcal{M}_{\lambda^*}$.

Ορισμός 1.13 (παραγόμενη σ-άλγεβρα) Έστω Ω τυχόν, μη κενό, σύνολο, και $\mathfrak{D}$ τυχούσα, μη κενή, συλλογή (κλάση) υποσυνόλων του Ω. Τότε, οποιαδήποτε σ-άλγεβρα $\mathcal{A}$, για την οποία ικανοποιούνται οι εξής δύο ιδιότητες:

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

(ii) Αν η $\mathcal{A}'$ είναι σ-άλγεβρα στον Ω και $\mathfrak{D} \subset \mathcal{A}'$ τότε $\mathcal{A} \subset \mathcal{A}'$, ονομάζεται **σ-άλγεβρα παραγόμενη** από την $\mathfrak{D}$, και συμβολίζεται με $\sigma(\mathfrak{D})$.

Πρόταση 1.14 Για Ω και $\mathfrak{D}$ όπως στον Ορισμό 1.13, υπάρχει πάντα η $\sigma(\mathfrak{D})$, και είναι μοναδική.

Απόδειξη: Ας θεωρήσουμε την συλλογή σ-αλγεβρών

$$\mathbf{\Sigma} = \{\mathcal{A}' : \textit{η } \mathcal{A}' \textit{ είναι σ-άλγεβρα στον } \Omega,\ \textit{με } \mathfrak{D} \subset \mathcal{A}'\}\,.$$

Προφανώς $\mathcal{P}(\Omega) \in \mathbf{\Sigma}$, και άρα το $\mathbf{\Sigma}$ είναι μη κενό. Θεωρούμε την

$$\mathcal{A} = \bigcap_{\mathcal{A}' \in \mathbf{\Sigma}} \mathcal{A}',$$

και διαπιστώνουμε ότι η $\mathcal{A}$ είναι σ-άλγεβρα στον Ω, και $\mathfrak{D} \subset \mathcal{A}$ (δείξτε το). Είναι προφανές τώρα ότι αν και η $\mathcal{A}'$ είναι σ-άλγεβρα στον Ω και $\mathfrak{D} \subset \mathcal{A}'$, τότε $\mathcal{A}' \in \mathbf{\Sigma}$, και συνεπώς, $\mathcal{A} \subset \mathcal{A}'$. Επομένως, η $\mathcal{A}$ ικανοποιεί τα (i) και (ii) του Ορισμού 1.13.

Αν τώρα υποθέσουμε ότι και μία άλλη σ-άλγεβρα, έστω $\mathcal{B}$, επίσης ικανοποιεί τα (i) και (ii) του Ορισμού 1.13, τότε θα είναι $\mathcal{A} \subset \mathcal{B}$ (γιατί $\mathfrak{D} \subset \mathcal{B}$), αλλά και $\mathcal{B} \subset \mathcal{A}$ (γιατί $\mathfrak{D} \subset \mathcal{A}$). Επομένως, $\mathcal{A} = \mathcal{B}$, η $\mathcal{A}$ είναι η **μοναδική** σ-άλγεβρα που παράγεται από την $\mathfrak{D}$, και συνεπώς, $\sigma(\mathfrak{D}) = \mathcal{A}$. $\square$

Θεωρούμε τώρα τις εξής κλάσεις υποσυνόλων του[13] $(0,1)$:

[13] Η $\mathfrak{D}_1$ ονομάζεται «οικογένεια των δεξιά κλειστών ημιευθειών» (του $(0,1)$), και θα διαδραματίσει σπουδαίο ρόλο στα Κεφάλαια 2–;;.

(i) $\mathfrak{D}_1 = \{(0, y],\ 0 \le y < 1\}$,

(ii) $\mathfrak{D}_2 = \{(x, y],\ 0 \le x \le y < 1\}$,

(iii) $\mathfrak{D}_3 = \{(x, y),\ 0 \le x \le y \le 1\}$,

(iv) $\mathcal{G} = \{U : U\ \ \textit{ανοικτό},\ U \subset (0,1)\}$,

(v) $\mathcal{F} = \{F : F\ \ \textit{κλειστό},\ F \subset (0,1)\}$.

Τότε ισχύει η εξής:

Πρόταση 1.15 $\sigma(\mathfrak{D}_1) = \sigma(\mathfrak{D}_2) = \sigma(\mathfrak{D}_3) = \sigma(\mathcal{G}) = \sigma(\mathcal{F})$.

Απόδειξη: Αφού η $\sigma(\mathcal{G})$ περιέχει τα ανοικτά σύνολα, και είναι σ-άλγεβρα, θα περιέχει και τα κλειστά σύνολα, δηλ. $\mathcal{F} \subset \sigma(\mathcal{G})$. Άρα, $\sigma(\mathcal{F}) \subset \sigma(\mathcal{G})$.

Παρατηρούμε τώρα ότι $\mathfrak{D}_1 \subset \mathcal{F}$, διότι τα σύνολα της μορφής $(0, x]$ είναι κλειστά στο $(0, 1)$, αφού τα συμπληρώματά τους, $(x, 1)$, είναι ανοικτά, και συνεπώς, $\sigma(\mathfrak{D}_1) \subset \sigma(\mathcal{F})$.

Από την ταυτότητα $(x, y] = (0, y] \setminus (0, x] = (0, y] \cap (0, x]^c = ((0, y]^c \cup (0, x])^c = ((0, x] \cup (0, y]^c \cup \varnothing \cup \cdots)^c$ έπεται ότι $(x, y] \in \sigma(\mathfrak{D}_1)$. [Διότι, προφανώς, τα σύνολα $\varnothing, (0, x], (0, y]$ ανήκουν στην $\sigma(\mathfrak{D}_1)$, οπότε και το $(0, y]^c$ ανήκει στην $\sigma(\mathfrak{D}_1)$. Έτσι, $(0, x] \cup (0, y]^c \cup \varnothing \cup \cdots \in \sigma(\mathfrak{D}_1)$, και επομένως, $((0, x] \cup (0, y]^c \cup \varnothing \cup \cdots)^c \in \sigma(\mathfrak{D}_1)$.] Συνεπώς, $\mathfrak{D}_2 \subset \sigma(\mathfrak{D}_1)$, και, δεδομένου ότι η $\sigma(\mathfrak{D}_1)$ είναι μία σ-άλγεβρα που περιέχει την $\mathfrak{D}_2$, έπεται ότι και $\sigma(\mathfrak{D}_2) \subset$

$\sigma(\mathfrak{D}_1)$, αφού, εξ' ορισμού, η $\sigma(\mathfrak{D}_2)$ είναι η ελάχιστη σ-άλγεβρα με αυτήν την ιδιότητα.

Παρατηρούμε ότι $(x,y) = \bigcup_{n=1}^{\infty} \big(x, y-\frac{1}{n}\big] \in \sigma(\mathfrak{D}_2)$, διότι $\big(x, y-\frac{1}{n}\big] \in \mathfrak{D}_2 \subset \sigma(\mathfrak{D}_2)$. Επομένως, προκύπτει ότι $\mathfrak{D}_3 \subset \sigma(\mathfrak{D}_2)$, και συνεπώς, $\sigma(\mathfrak{D}_3) \subset \sigma(\mathfrak{D}_2)$.

Έχοντας λοιπόν δείξει ότι

$$\sigma(\mathfrak{D}_3) \subset \sigma(\mathfrak{D}_2) \subset \sigma(\mathfrak{D}_1) \subset \sigma(\mathcal{F}) \subset \sigma(\mathcal{G}),$$

μένει να αποδειχθεί ότι και $\sigma(\mathcal{G}) \subset \sigma(\mathfrak{D}_3)$, οπότε η απόδειξη θα είναι πλήρης. Αυτό έπεται άμεσα από το γεγονός ότι κάθε ανοικτό σύνολο $U \subset (0,1)$ γράφεται ως αριθμήσιμη ένωση ανοικτών διαστημάτων,[14]

$$U = \bigcup_{n=1}^{\infty} (a_n, b_n), \ \mu\epsilon\ 0 \le a_n \le b_n \le 1, \ n = 1, 2, \ldots,$$

οπότε $U \in \sigma(\mathfrak{D}_3)$, και συνεπώς, $\mathcal{G} \subset \sigma(\mathfrak{D}_3)$, οπότε και $\sigma(\mathcal{G}) \subset \sigma(\mathfrak{D}_3)$. $\square$

Ορισμός 1.16 (σ-άλγεβρα του Borel σε τυχόντα μετρικό χώρο Ω) Έστω (Ω, ρ) τυχόν μετρικός χώρος (ρ η μετρική του). Η σ-άλγεβρα που παράγεται από τα ανοικτά υποσύνολα του Ω ονομάζεται **σ-άλγεβρα του Borel** (στον Ω), και συμβολίζεται ως $\mathcal{B}(\Omega)$. Τα σύνολα που περιέχει η $\mathcal{B}(\Omega)$ ονομάζονται **σύνολα Borel**, ή **Borel μετρήσιμα σύνολα** (του Ω).

[14] αρκεί να θεωρήσουμε την (αριθμήσιμη) ένωση όλων των ανοικτών διαστημάτων που περιέχονται στο U, και τα οποία έχουν ρητά άκρα

Σύμφωνα με την Πρόταση 1.15, ισχύει ότι[15]

$$\mathcal{B}((0,1)) \overset{\text{ορ.}}{=\!=} \sigma(\mathcal{G}) = \sigma(\mathfrak{D}_1).$$

Επειδή, προφανώς, τα σύνολα της $\mathfrak{D}_1$ είναι Lebesgue μετρήσιμα, αφού $\lambda((0,y]) = y$, έπεται ότι $\mathfrak{D}_1 \subset \mathcal{M}_{\lambda^*}$, και συνεπώς,

$$\mathcal{B}((0,1)) = \sigma(\mathfrak{D}_1) \subset \mathcal{M}_{\lambda^*}, \tag{1.4}$$

διότι η $\mathcal{M}_{\lambda^*}$ είναι σ-άλγεβρα. Επομένως, ένα Borel σύνολο είναι **πάντα** Lebesgue σύνολο. Το αντίστροφο **δεν ισχύει**, δηλ. η ανισότητα στην (1.4) είναι γνήσια. Αυτό όμως δεν είναι καθόλου προφανές, και αποδεικνύεται με βάση το γεγονός ότι ο πληθάριθμος της $\mathcal{B}((0,1))$ είναι ο πληθάριθμος του συνεχούς, c, ενώ ο πληθάριθμος της $\mathcal{M}_{\lambda^*}$ είναι ο γνήσια μεγαλύτερος πληθάριθμος του δυναμοσυνόλου των πραγματικών αριθμών, $\mathcal{P}(\mathbb{R})$, ο οποίος συμβολίζεται με 2^c. Άρα,

$$\mathcal{B}((0,1)) \subsetneqq \mathcal{M}_{\lambda^*}.$$

---

[15] Η $\mathfrak{D}_1$ ονομάζεται **οικογένεια των δεξιά κλειστών ημιευθειών**. Όταν αναφερόμαστε στο διάστημα $(0,1)$, η $\mathfrak{D}_1$ περιέχει τα «δεξιά κλειστά ημιδιαστήματα», αλλά, αν θεωρήσουμε ως χώρο αναφοράς το σύνολο όλων των πραγματικών αριθμών, $\mathbb{R}$, τότε η $\mathfrak{D}_1$ περιέχει τις ημιευθείες $(-\infty, x]$, $x \in \mathbb{R}$. Αντίστοιχοι ορισμοί της $\mathfrak{D}_1$ μπορούν να δοθούν στην πολυδιάστατη περίπτωση. Για παράδειγμα, στον $\mathbb{R}^n$ η $\mathfrak{D}_1$ περιέχει τα υποσύνολα της μορφής $(-\infty, x_1] \times \cdots \times (-\infty, x_n]$, $x_1, \ldots, x_n \in \mathbb{R}$, ενώ στον χώρο $(0,1)^n$ περιέχει τα υποσύνολα της μορφής $(0, x_1] \times \cdots \times (0, x_n]$, $x_1, \ldots, x_n \in [0,1)$. Σε όλες τις περιπτώσεις, πάντως, η $\mathfrak{D}_1$ παράγει την αντίστοιχη σ-άλγεβρα του Borel, όταν τα ανοικτά ορίζονται ως προς την (συνήθη) Ευκλείδεια μετρική (Άσκηση 1.10), η δε απόδειξη δεν διαφέρει από αυτήν της Πρότασης 1.15, για τον χώρο $(0,1)$.

Τελικά αποδεικνύεται ότι ο χώρος $((0,1), \mathcal{M}_{\lambda^*}, \lambda)$ είναι η **πλήρωση** του χώρου

$$((0,1), \mathcal{B}((0,1)), \lambda).$$

[Χρησιμοποιούμε και για τους δύο χώρους τον συμβολισμό λ, με την παραδοχή ότι στον δεύτερο χώρο το μέτρο Lebesgue λ ορίζεται **μόνο** στα Borel σύνολα.]

Στην ουσία, το μόνο περισσότερο που περιέχει ο χώρος $((0,1), \mathcal{M}_{\lambda^*}, \lambda)$, είναι ορισμένα σύνολα B, τα οποία δεν είναι Borel, και για τα οποία υπάρχουν Borel σύνολα A, C, με

$$A \subset B \subset C, \quad \text{και} \quad \lambda(C \setminus A) = 0.$$

Επειδή η σ-άλγεβρα του Borel είναι πολύ πλούσια, διότι περιέχει, π.χ., όλα τα ανοικτά, κλειστά, F_σ (αριθμήσιμες ενώσεις κλειστών), και G_δ (αριθμήσιμες τομές ανοικτών) σύνολα, καθώς και αριθμήσιμες ενώσεις και τομές αυτών, τελικά κρίνεται επαρκής **για όλα** τα «λογικά» ερωτήματα, για τα οποία θα ενδιαφερόμασταν να υπολογίσουμε την πιθανότητά τους (μήκος). Επομένως, τα «πολύ περίεργα» σύνολα που ανήκουν στην οικογένεια

$$\mathcal{M}_{\lambda^*} \setminus \mathcal{B}((0,1)),$$

αποτελούν μία «μικρή ενόχληση» που δεν θα μας απασχολήσει ιδιαίτερα.

Τέλος, θα πρέπει να σημειωθεί ότι, εν γένει, μία σ-άλγεβρα περιέχει **όλους** τους «ενδιαφέροντες» συνδυασμούς συνόλων της. Για παράδειγμα, μπορεί να δειχθεί εύκολα (δείξτε το) ότι αν $A, B, A_1, A_2, \ldots$ είναι στοιχεία μιας σ-άλγεβρας $\mathcal{A}$, τότε και τα σύνολα που ορίζονται στις (i)–(vii) επίσης είναι:

(i) $A \cup B$ [ένωση]

(ii) AB [τομή]

(iii) $\bigcap_{n=1}^{\infty} A_n$ [αριθμήσιμη τομή]

(iv) $A \setminus B \stackrel{\text{ορ.}}{=\!=} AB^c$ [διαφορά]

(v) $A \oplus B \stackrel{\text{ορ.}}{=\!=} (A \setminus B) \cup (B \setminus A)$ [συμμετρική διαφορά]

(vi) $\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$ [κατώτερο όριο]

(vii) $\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$ [ανώτερο όριο]

Ειδικότερα, τα δύο τελευταία σύνολα (ενδεχόμενα) ονομάζονται **κατώτερο όριο** (limit inferior) και **ανώτερο όριο** (limit superior) της ακολουθίας συνόλων (ενδεχομένων) $\{A_n, n \geq 1\}$, και συμβολίζονται ως

$$\liminf_{n\to\infty} A_n \stackrel{\text{ορ.}}{=\!=} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k \qquad (\text{ή} \quad \liminf A_n, \quad \text{ή και} \quad \underline{\lim} A_n)$$

και

$$\limsup_{n\to\infty} A_n \stackrel{\text{ορ.}}{=\!=} \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \quad \big(\text{ή} \quad \limsup A_n, \quad \text{ή και} \quad \overline{\lim} A_n\big) .$$

Τα οριακά αυτά σύνολα συμπεριφέρονται εντελώς ανάλογα με τα αντίστοιχα «ενδιαφέροντα όρια» μίας ακολουθίας πραγματικών αριθμών $\{a_n,\ n \geq 1\}$, δηλ. το $\liminf a_n$ και το $\limsup a_n$. Για παράδειγμα,

$$\liminf A_n \subset \limsup A_n.$$

Επίσης, εύκολα αποδεικνύεται (Άσκηση 1.1) ότι

$$\liminf A_n = \{\omega \in \Omega : \omega \in A_n \text{ τελικά για κάθε } n\},$$

και

$$\limsup A_n = \{\omega \in \Omega : \omega \in A_n \text{ για άπειρα } n\}.$$

Η έκφραση «η πρόταση P_n ισχύει τελικά για κάθε n» σημαίνει, ως συνήθως, ότι «υπάρχει κάποιος δείκτης n_0, τέτοιος ώστε η P_n να ισχύει για κάθε $n \geq n_0$», ενώ η έκφραση «η P_n ισχύει για άπειρα n» σημαίνει ότι «υπάρχει υπακολουθία $\{n_k,\ k = 1, 2, \ldots\} \subset \mathbb{N}$, τέτοια ώστε η P_{n_k} να ισχύει για κάθε k».

Τέλος, λέμε ότι το όριο της ακολουθίας $\{A_n,\ n \geq 1\}$ **υπάρχει**, όταν

$$\limsup A_n \subset \liminf A_n,$$

(οπότε $\limsup A_n = \liminf A_n$). Όταν γράφουμε $\lim A_n = A$ (ή $\lim_n A_n = A$ ή, ακριβέστερα, $\lim_{n\to\infty} A_n = A$), εννοούμε ότι

$$\limsup A_n = \liminf A_n = A \quad (= \lim A_n).$$

Φυσικά, όποτε υπάρχει, $\lim A_n \in \mathcal{A}$.

Ασκήσεις Κεφ. 1:

1.1. Έστω $(\Omega, \mathcal{A})$ μετρήσιμος χώρος, και $\{A_n,\ n \geq 1\}$ ακολουθία της $\mathcal{A}$. Δείξτε ότι:

(i) $\liminf A_n = \{\omega \in \Omega : \omega \in A_n$ τελικά για κάθε $n\}$.
[Στην γλώσσα των πιθανοτήτων, το $\liminf A_n$ παριστάνει το ενδεχόμενο: $\{$*να εμφανιστούν όλα τα* A_n *από ένα* n_0 *και πάνω*$\}$, δηλ. $\{$*να εμφανιστούν τα* A_n *«τελικά για κάθε* n*»*$\}$.]

(ii) $\limsup A_n = \{\omega \in \Omega : \omega \in A_n$ για άπειρα $n\}$.
[Εξ' αιτίας αυτού, για το $\limsup A_n$ συναντάμε στην βιβλιογραφία τον συμβολισμό $\{A_n,$ i.o.$\}$, που προέρχεται από το i.o. = infinitely often = απείρως συχνά, επειδή το $\{$*να συμβούν «άπειρα το πλήθος»* $A_n\}$ ισοδυναμεί με το $\{$*να συμβεί το* A_n *«για άπειρα* n*»*$\}$.]

(iii) Αν η $\{A_n,\ n \geq 1\}$ είναι αύξουσα, δηλ.

$$A_n \subset A_{n+1}, \ \text{για } n = 1, 2, \ldots,$$

[συμβολισμός: $A_n \nearrow$] τότε το $\lim A_n$ υπάρχει, και μάλιστα,

$$\lim A_n = \bigcup_{n=1}^{\infty} A_n.$$

[Συμβολισμός: $A_n \nearrow A$, όπου $A = \bigcup_{n=1}^{\infty} A_n$.]

(iv) Αντίστοιχα, αν η $\{A_n,\ n \geq 1\}$ είναι φθίνουσα, δηλ.

$$A_n \supset A_{n+1}, \ \text{για } n = 1, 2, \ldots,$$

[συμβολισμός: $A_n \searrow$] τότε το $\lim A_n$ υπάρχει, και μάλιστα,

$$\lim A_n = \bigcap_{n=1}^{\infty} A_n.$$

[Συμβολισμός: $A_n \searrow A$, όπου $A = \bigcap_{n=1}^{\infty} A_n$.]

(v) Δώστε παράδειγμα ακολουθίας $\{A_n,\ n \geq 1\}$, για την οποία

$$\liminf A_n = \varnothing, \quad \text{και} \quad \limsup A_n = \Omega.$$

(vi) Αποδείξτε ότι

$$(\liminf A_n)^c = \limsup A_n^c, \ \text{ και } \ (\limsup A_n)^c = \liminf A_n^c.$$

(vii) Δώστε παράδειγμα ακολουθίας $\{A_n,\ n \geq 1\}$ που δεν είναι τελικά μονότονη (δηλ. δεν υπάρχει $m \geq 1$, τέτοιο ώστε η $\{A_{n+m},\ n \geq 1\}$ να είναι αύξουσα ή φθίνουσα), και το $\lim A_n$ υπάρχει.

1.2. Έστω $(\Omega, \mathcal{A})$ ένας μετρήσιμος χώρος, και $\{A_n,\ n \geq 1\}$ τυχούσα ακολουθία στην $\mathcal{A}$. Δείξτε ότι

$$I_{\limsup A_n} = \limsup I_{A_n}, \quad \text{και} \quad I_{\liminf A_n} = \liminf I_{A_n},$$

όπου I_A η **δείκτρια συνάρτηση** του A, με τιμές $I_A(\omega) = 0$ ή 1, ανάλογα με το αν $\omega \notin A$ ή $\omega \in A$, αντίστοιχα.

[Μερικές φορές χρησιμοποιείται και ο συμβολισμός $I(\omega \in A)$, ή $I_{\{\omega\in A\}}$, ή $1(\omega \in A)$, ή $\mathbf{1}(\omega \in A)$, ή $1_{\{\omega\in A\}}$, ή $\mathbf{1}_{\{\omega\in A\}}$, ή $1_A(\omega)$, ή $\mathbf{1}_A(\omega)$, καθώς και ο $\chi_A(\omega)$, αντί του $I_A(\omega)$.][16]

1.3. Δείξτε ότι αν $A_n \to A$ (δηλ. $\lim_{n\to\infty} A_n = A$), και $B_n \to B$, τότε

$$A_n \cup B_n \to A \cup B, \ \text{ και } \ A_n B_n \to AB.$$

1.4. Δείξτε ότι ισχύουν τα εξής:

(i) $(\limsup A_n) \cap (\limsup B_n) \supset \limsup(A_n B_n)$.

(ii) $(\limsup A_n) \cup (\limsup B_n) = \limsup(A_n \cup B_n)$.

(iii) $(\liminf A_n) \cap (\liminf B_n) = \liminf(A_n B_n)$.

(iv) $(\liminf A_n) \cup (\liminf B_n) \subset \liminf(A_n \cup B_n)$.

Μπορούν οι ανισότητες (i) και (iv) να γίνουν γνήσιες;

1.5. Έστω $(\Omega, \mathcal{A})$ μετρήσιμος χώρος, και $\{\mathbb{P}_n,\ n \geq 1\}$ τυχούσα ακολουθία πιθανοτήτων στον $(\Omega, \mathcal{A})$, έτσι ώστε ο $(\Omega, \mathcal{A}, \mathbb{P}_n)$ να είναι χώρος πιθανότητας για κάθε $n \geq 1$. Δείξτε ότι ο $(\Omega, \mathcal{A}, \mathbb{P})$, με

$$\mathbb{P}(A) = \sum_{n=1}^{\infty} \frac{1}{2^n} \mathbb{P}_n(A), \quad \text{για κάθε } A \in \mathcal{A},$$

[16] Υπάρχουν τόσοι πολλοί συμβολισμοί ακριβώς επειδή οι δείκτριες συναρτήσεις αποτελούν τον «θεμέλιο λίθο», τόσο της Θεωρίας Πιθανοτήτων, όσο και της Θεωρίας Μέτρου και Ολοκλήρωσης.

είναι χώρος πιθανότητας.

1.6. Έστω $\mathcal{A}$ τυχούσα σ-άλγεβρα στον $\Omega \neq \varnothing$, και $B \in \mathcal{A}$, με $B \neq \varnothing$. Τότε η

$$\mathcal{A}_B = \{AB,\ A \in \mathcal{A}\}$$

είναι σ-άλγεβρα επί του B.
[Η $\mathcal{A}_B$ ονομάζεται σ-άλγεβρα **ίχνος** της $\mathcal{A}$ στο B.]

1.7. (συνέχεια) Αν, επιπλέον, ο $(\Omega, \mathcal{A}, \mathbb{P})$ είναι χώρος πιθανότητας, και αν $\mathbb{P}(B) > 0$, τότε ο χώρος $(B, \mathcal{A}_B, \mathbb{P}_B)$, με

$$\mathbb{P}_B(A) = \frac{\mathbb{P}(A)}{\mathbb{P}(B)}, \quad \text{για κάθε } A \in \mathcal{A}_B,$$

είναι, επίσης, χώρος πιθανότητας.
[Η πιθανότητα $\mathbb{P}_B$ λέγεται **πιθανότητα ίχνος** της $\mathbb{P}$ στο B.][17]

1.8. Δείξτε ότι $\mathcal{A}$ και $\mathcal{B}$ είναι σ-άλγεβρες στον Ω, τότε η $\mathcal{A} \cup \mathcal{B}$ μπορεί να μην είναι σ-άλγεβρα στον Ω, και η

$$\mathcal{A} \times \mathcal{B} \overset{\text{ορ.}}{=\!=} \{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\}$$

μπορεί να μην είναι σ-άλγεβρα στον $\Omega^2 \overset{\text{ορ.}}{=\!=} \Omega \times \Omega$.

[17] Η πιθανότητα $\mathbb{P}_B$ είναι στενά συνδεδεμένη με την έννοια της **δεσμευμένης πιθανότητας**, βλ. Κεφάλαια 2, ;;, ;;.

Γενικά, για τυχούσες κλάσεις υποσυνόλων του Ω, έστω $\mathfrak{D}_1$ και $\mathfrak{D}_2$,

$$\sigma(\mathfrak{D}_1) \cup \sigma(\mathfrak{D}_2) \subset \sigma(\mathfrak{D}_1 \cup \mathfrak{D}_2) \subset \sigma(\sigma(\mathfrak{D}_1) \cup \sigma(\mathfrak{D}_2)),$$

και

$$\sigma(\mathfrak{D}_1) \times \sigma(\mathfrak{D}_2) \subset \sigma(\sigma(\mathfrak{D}_1) \times \sigma(\mathfrak{D}_2)),$$

χωρίς να αποκλείεται να είναι γνήσιες οι παραπάνω ανισότητες.

Ισχύει γενικά ότι

$$\sigma(\mathfrak{D}_1 \cap \mathfrak{D}_2) = \sigma(\mathfrak{D}_1) \cap \sigma(\mathfrak{D}_2) \ ;$$

1.9. Δείξτε ότι κάθε αριθμήσιμο σύνολο $N \subset (0,1)$ είναι Borel, και $\lambda(N) = 0$. Επομένως, κάθε σύνολο M που είναι το συμπλήρωμα ενός αριθμήσιμου συνόλου N, $M = N^c$, είναι Borel, και $\lambda(M) = 1$. Συμπεράνατε ότι ένας τυχαία επιλεγμένος αριθμός του $(0,1)$ είναι, με πιθανότητα 1, **υπερβατικός**.

1.10. **(δεξιά κλειστές ημιευθείες)**

(i) Θέτουμε $\mathfrak{D}_1 = \{(-\infty, y],\ y \in \mathbb{R}\}$. Να δείξετε ότι ισχύει η σχέση $\mathcal{B}(\mathbb{R}) = \sigma(\mathfrak{D}_1)$.

(ii) Θέτουμε $\mathfrak{D}_1 = \{(-\infty, y_1] \times \cdots \times (-\infty, y_n],\ y_1, \ldots, y_n \in \mathbb{R}\}$. Να δείξετε ότι ισχύει η σχέση $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathfrak{D}_1)$.

(iii) Θέτουμε $\mathfrak{D}_1 = \{(0, y_1] \times \cdots \times (0, y_n],\ y_1, \ldots, y_n \in [0,1)\}$. Να δείξετε ότι ισχύει η σχέση $\mathcal{B}((0,1)^n) = \sigma(\mathfrak{D}_1)$.

[**Υπόδειξη:** Για τον $\mathbb{R}$ η απόδειξη είναι ακριβώς ίδια με αυτήν της Πρότασης 1.15. Για να αποδείξετε την ανισότητα $\sigma(\mathfrak{D}_2) \subset \sigma(\mathfrak{D}_1)$ στον $(0,1)^n$, χρησιμοποιήστε την ταυτότητα $(x_1, y_1] \times \cdots \times (x_n, y_n] = B_0 \cap B_1^c \cap \cdots \cap B_n^c$, όπου $B_0 = (0, y_1] \times \cdots \times (0, y_n]$, και $B_j = (0, y_1] \times \cdots \times (0, y_{j-1}] \times (0, x_j] \times (0, y_{j+1}] \times \cdots \times (0, y_n]$, $j = 1, 2, \ldots, n$. Ομοίως και για τον $\mathbb{R}^n$.]

Κεφάλαιο 2

Ιδιότητες της Πιθανότητας – Ανεξαρτησία

2.1 Βασικές Ιδιότητες της Πιθανότητας

Στο κεφάλαιο αυτό θα μελετήσουμε τις στοιχειώδεις ιδιότητες που προκύπτουν από τον ορισμό της πιθανότητας, και στην συνέχεια θα ορίσουμε και θα μελετήσουμε την έννοια της **(στοχαστικής) ανεξαρτησίας ενδεχομένων**, και τις συνέπειές της.

Θεώρημα 2.1 Έστω $(\Omega, \mathcal{A}, \mathbb{P})$ τυχόν χώρος πιθανότητας, και $A, B, A_1, \ldots, A_n \in \mathcal{A}$. Τότε:

(i) $\mathbb{P}(\varnothing) = 0$.

(ii) Αν $A_i A_j = \varnothing$ για $i \neq j$, $1 \leq i, j \leq n$, τότε $\mathbb{P}\big(\bigcup_{j=1}^{n} A_j\big) = \sum_{j=1}^{n} \mathbb{P}(A_j)$.

(iii) $\mathbb{P}(B \smallsetminus A) = \mathbb{P}(B) - \mathbb{P}(AB)$. [Ειδικότερα, αν $A \subset B$ τότε $\mathbb{P}(B \smallsetminus A) = \mathbb{P}(B) - \mathbb{P}(A)$.]

(iv) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

(v) Αν $A \subset B$ τότε $\mathbb{P}(A) \leq \mathbb{P}(B)$. [μονοτονία]

(vi) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(AB)$.

(vii) $\mathbb{P}(A_1 \cup \cdots \cup A_n)$
$= \sum_{j=1}^{n} P(A_j) - \sum\sum_{1 \leq j_1 < j_2 \leq n} \mathbb{P}(A_{j_1} A_{j_2}) + \ldots + (-1)^{n-1} \mathbb{P}(A_1 A_2 \cdots A_n)$.
[τύπος Poincaré, ή αρχή εγκλεισμού-αποκλεισμού]

Απόδειξη: Κατ΄ αρχήν θα πρέπει να σημειωθεί ότι όλα τα ενδεχόμενα, των οποίων λαμβάνουμε την πιθανότητα, όπως τα $\varnothing$, $A \cup B$, $A_1 \cup \cdots \cup A_n$, $B \setminus A$, κ.λπ., ανήκουν στην $\mathcal{A}$.

(i) Αφού $\Omega = \Omega \cup \varnothing \cup \varnothing \cup \cdots$, έπεται ότι

$$1 = \mathbb{P}(\Omega \cup \varnothing \cup \varnothing \cup \cdots) = 1 + \mathbb{P}(\varnothing) + \mathbb{P}(\varnothing) + \ldots,$$

και επομένως, $\mathbb{P}(\varnothing) = 0$.

(ii) Είναι $A_1 \cup \cdots \cup A_n = B_1 \cup B_2 \cup \cdots$, όπου

$$B_j = \begin{cases} A_j, & \text{αν} \quad j \leq n, \\ \varnothing, & \text{αν} \quad j > n, \end{cases} \qquad j = 1, 2, \ldots,$$

και αφού $B_i B_j = \varnothing$ για $i \neq j$, $i, j = 1, 2, \ldots$, έπεται ότι

$$\mathbb{P}(A_1 \cup \cdots \cup A_n) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(B_n) = \mathbb{P}(A_1) + \ldots + \mathbb{P}(A_n).$$

(iii) Είναι $B = (B \setminus A) \cup BA$ με $(B \setminus A) \cap (BA) = \varnothing$. Από το (ii) έπεται ότι

$$\mathbb{P}(B) = \mathbb{P}(B \setminus A) + \mathbb{P}(BA), \text{ δηλ. } \mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(BA).$$

[Φυσικά, αν $A \subset B$ τότε $AB = A$, και $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$.]

(iv) Αφού $\Omega = A \cup A^c$ και $AA^c = \varnothing$, από το (ii) έπεται ότι

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c).$$

(v) Αν $A \subset B$ τότε λόγω του (iii),

$$\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A),$$

και επομένως, $\mathbb{P}(B) \geq \mathbb{P}(A)$, αφού $\mathbb{P}(B \setminus A) \geq 0$.

(vi) Είναι $A \cup B = A \cup (B \setminus A)$ και $A \cap (B \setminus A) = \varnothing$. Άρα,

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) = \mathbb{P}(A) + [\mathbb{P}(B) - \mathbb{P}(AB)],$$

λόγω των (ii) και (iii).

(vii) Η απόδειξη προκύπτει με επαγωγή στο n, βλ., π.χ., Χαραλαμπίδης (1990), Θεώρημα 7.2, σελ. 54. [Για $n = 2$ είναι το (vi).]

Θεώρημα 2.2 (στοιχειώδεις ανισότητες) Έστω $A_1, A_2, \ldots$ μία ακολουθία ενδεχομένων στον χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$. Τότε:

$$\left.\begin{array}{l} \text{(i)}\ \mathbb{P}\left(\bigcup_{j=1}^{n} A_j\right) \leq \sum_{j=1}^{n} \mathbb{P}(A_j). \\ \\ \text{(ii)}\ \mathbb{P}\left(\bigcup_{j=1}^{\infty} A_j\right) \leq \sum_{j=1}^{\infty} \mathbb{P}(A_j). \end{array}\right\} \quad \text{[ανισότη-}$$

τα Boole]

(iii) $\mathbb{P}\left(\bigcap_{j=1}^{n} A_j\right) \geq \sum_{j=1}^{n} \mathbb{P}(A_j) - n + 1.$ [ανισότητα Bonferroni]

Απόδειξη: (ii) Θέτουμε $B_1 = A_1$, και $B_j = A_j \setminus \left(\bigcup_{i=1}^{j-1} A_j\right)$, για $j = 2, 3, \ldots$. Τότε $B_j \in \mathcal{A}$, $B_i B_j = \varnothing$ για $i \neq j$, $i, j = 1, 2, \ldots$, και $\bigcup_{j=1}^{\infty} A_j = \bigcup_{j=1}^{\infty} B_j$. Επίσης, $B_j \subset A_j$ για $j = 1, 2, \ldots$, και συνεπώς, $\mathbb{P}\left(\bigcup_{j=1}^{\infty} A_j\right) = \mathbb{P}\left(\bigcup_{j=1}^{\infty} B_j\right) = \sum_{j=1}^{\infty} \mathbb{P}(B_j) \leq \sum_{j=1}^{\infty} \mathbb{P}(A_j)$.

(i) Προκύπτει άμεσα από το (ii), θέτοντας $A_j = \varnothing$ για $j > n$.

(iii) Έχουμε

$$\begin{aligned}
\mathbb{P}\Big(\bigcap_{j=1}^{n} A_j\Big) &= 1 - \mathbb{P}\Big(\big(\bigcap_{j=1}^{n} A_j\big)^c\Big) = 1 - \mathbb{P}\Big(\bigcup_{j=1}^{n} A_j^c\Big) \\
&\geq 1 - \sum_{j=1}^{n} \mathbb{P}(A_j^c) = 1 - \sum_{j=1}^{n} [1 - \mathbb{P}(A_j)] \\
&= \sum_{j=1}^{n} \mathbb{P}(A_j) - n + 1,
\end{aligned}$$

όπου η ανισότητα προκύπτει από την (i), εφαρμοζόμενη στα ενδεχόμενα A_j^c, $j = 1, 2, \ldots, n$. $\square$

Πρόταση 2.3 Εάν η $\{A_n,\ n \geq 1\}$ είναι μονότονη ακολουθία ενδεχομένων (σε χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$), τότε[1]

$$\mathbb{P}(\lim A_n) = \lim \mathbb{P}(A_n).$$

Απόδειξη: Από την Άσκηση 1.1 γνωρίζουμε ότι το $\lim A_n$ υπάρχει, και μάλιστα,

$$\lim A_n = \begin{cases} \bigcup_{n=1}^{\infty} A_n, & \text{αν η } A_n \text{ είναι αύξουσα,} \\ \bigcap_{n=1}^{\infty} A_n, & \text{αν η } A_n \text{ είναι φθίνουσα,} \end{cases}$$

και συνεπώς, $\lim A_n \in \mathcal{A}$. Επίσης, επειδή η $\{\mathbb{P}(A_n),\ n \geq 1\}$ είναι μονότονη (λόγω της μονοτονίας της $\mathbb{P}$), και φραγμένη, έπεται ότι το $\lim \mathbb{P}(A_n)$ υπάρχει, και μάλιστα,

$$0 \leq \lim \mathbb{P}(A_n) \leq 1.$$

Άρα, οι $\mathbb{P}(\lim A_n)$ και $\lim \mathbb{P}(A_n)$ είναι καλά ορισμένοι πραγματικοί αριθμοί του $[0, 1]$ (όταν η $\{A_n,\ n \geq 1\}$ είναι μονότονη), και αρκεί να δείξουμε ότι είναι και ίσοι.

Ας υποθέσουμε ότι η $\{A_n,\ n \geq 1\}$ είναι αύξουσα, οπότε $A_j \subset A_{j+1}$ για $j \geq 1$. Θέτοντας $B_1 = A_1$, και $B_j = A_j \smallsetminus A_{j-1}$ για $j = 2, 3, \ldots$, έχουμε ότι $B_i B_j = \varnothing$ για $i \neq j$, $B_1 \cup \cdots \cup B_n = A_1 \cup \cdots \cup A_n = A_n$, και, τέλος, $\bigcup_{j=1}^{\infty} B_j = \bigcup_{j=1}^{\infty} A_j =$

---

[1] Σημειώνεται ότι για **αύξουσες ακολουθίες** $\{A_n,\ n \geq 1\}$, η σχέση $\mu(\lim A_n) = \lim \mu(A_n)$ ισχύει για οποιοδήποτε μέτρο μ, ακόμα και αν το μ δεν είναι πεπερασμένο. Η απόδειξη είναι ακριβώς ίδια με αυτήν που δίδεται εδώ για μέτρα πιθανότητας, $\mu = \mathbb{P}$. Θα πρέπει να σημειωθεί ότι η σχέση $\mu(\lim A_n) = \lim \mu(A_n)$ **μπορεί να μην ισχύει** όταν η ακολουθία είναι φθίνουσα και το μέτρο δεν είναι πεπερασμένο, όπως, για παράδειγμα, αν πάρουμε ως μ το μέτρο Lebesgue στον $\mathbb{R}$, και ως $A_n = [n, +\infty)$.

$\lim A_n$. Άρα,

$$\begin{aligned}\mathbb{P}(\lim A_n) &= \mathbb{P}\left(\bigcup_{j=1}^{\infty} A_j\right) = \mathbb{P}\left(\bigcup_{j=1}^{\infty} B_j\right) = \sum_{j=1}^{\infty} \mathbb{P}(B_j) \\ &= \lim \sum_{j=1}^{n} \mathbb{P}(B_j) = \lim \mathbb{P}(B_1 \cup \cdots \cup B_n) = \lim \mathbb{P}(A_n).\end{aligned}$$

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

$$\mathbb{P}(\lim A_n^c) = \lim \mathbb{P}(A_n^c). \tag{2.1}$$

Όμως,

$$\lim A_n^c = \bigcup_{n=1}^{\infty} A_n^c = \left(\bigcap_{n=1}^{\infty} A_n\right)^c = (\lim A_n)^c, \text{ και}$$

$$\lim \mathbb{P}(A_n^c) = \lim [1 - \mathbb{P}(A_n)] = 1 - \lim \mathbb{P}(A_n).$$

Χρησιμοποιώντας τις δύο παραπάνω σχέσεις και την (2.1), προκύπτει ότι

$$1-\lim \mathbb{P}(A_n) = \lim \mathbb{P}(A_n^c) = \mathbb{P}(\lim A_n^c) = \mathbb{P}((\lim A_n)^c) = 1-\mathbb{P}(\lim A_n),$$

και το συμπέρασμα έπεται και για φθίνουσες ακολουθίες. □

Είμαστε τώρα σε θέση να δείξουμε το εξής βασικό θεώρημα.

Θεώρημα 2.4 (i) Για οποιαδήποτε ακολουθία ενδεχομένων $\{A_n,\ n \geq 1\}$, ενός χώρου πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, ισχύει η σχέση:[2]

$$\mathbb{P}(\liminf A_n) \leq \liminf \mathbb{P}(A_n) \leq \limsup \mathbb{P}(A_n) \leq \mathbb{P}(\limsup A_n).$$

[2] Η πρώτη ανισότητα είναι ειδική περίπτωση του Λήμματος Fatou, Πρόταση ;;(iii).

(ii) Αν, επιπλέον, $A_n \to A$ (δηλαδή $\limsup A_n = \liminf A_n = A$), τότε[3] και $\mathbb{P}(A_n) \to \mathbb{P}(A)$, καθώς $n \to \infty$. [Αυτό σημαίνει ότι η συνολοσυνάρτηση $\mathbb{P} : \mathcal{A} \longrightarrow [0,1]$ είναι **συνεχής**.]

Απόδειξη: Προφανώς το (ii) προκύπτει από το (i). Όσον αφορά το (i), αρκεί να δειχθεί η πρώτη ανισότητα,

$$\mathbb{P}(\liminf A_n) \le \liminf \mathbb{P}(A_n), \tag{2.2}$$

διότι τότε η τρίτη ανισότητα προκύπτει από την (2.2), εφαρμοζόμενη στην ακολουθία $\{A_n^c,\ n \ge 1\}$, ενώ η δεύτερη ανισότητα ισχύει για κάθε ακολουθία πραγματικών αριθμών. [Ας σημειωθεί ότι ισχύουν οι σχέσεις $\liminf A_n^c = (\limsup A_n)^c$ και $\liminf\,[1 - \mathbb{P}(A_n)] = 1 - \limsup \mathbb{P}(A_n)$.]

Έστω λοιπόν $B_n = \bigcap_{k=n}^{\infty} A_k$. Προφανώς η $\{B_n,\ n \ge 1\}$ είναι αύξουσα, και $\lim B_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k = \liminf A_n$. Άρα, από την Πρόταση 2.3,

$$\mathbb{P}(\liminf A_n) = \mathbb{P}(\lim B_n) = \lim \mathbb{P}(B_n).$$

Όμως, εξ ορισμού, $B_n \subset A_n$ για κάθε n, και άρα $\mathbb{P}(B_n) \le \mathbb{P}(A_n)$. Συνεπώς,

$$\lim \mathbb{P}(B_n) = \liminf \mathbb{P}(B_n) \le \liminf \mathbb{P}(A_n),$$

και η απόδειξη είναι πλήρης. □

[3]Είναι απλουστευμένη μορφή του Θεωρήματος Κυριαρχημένης Σύγκλισης, Θεώρημα ;;.

Ένα πολύ σημαντικό αποτέλεσμα, χρήσιμο στη μελέτη της οριακής συμπεριφοράς ακολουθίας ενδεχομένων, δίδεται στο επόμενο θεώρημα.

Θεώρημα 2.5 (πρώτο Λήμμα Borel-Cantelli)[4] Για τυχούσα ακολουθία ενδεχομένων $\{A_n,\ n \geq 1\}$ στον χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, η συνθήκη

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$$

μας εξασφαλίζει ότι

$$\mathbb{P}(\limsup A_n) = 0.$$

[Άρα, με πιθανότητα 1, τα A_n^c θα εμφανιστούν τελικά για κάθε n ή, ισοδύναμα, με πιθανότητα 1 θα εμφανιστούν το πολύ πεπερασμένα το πλήθος από τα A_n.]

Απόδειξη: Θέτοντας $B_n = \bigcup_{k=n}^{\infty} A_k$, έχουμε ότι η ακολουθία $\{B_n,\ n \geq 1\}$ είναι φθίνουσα, και

$$\lim_{n\to\infty} B_n = \bigcap_{n=1}^{\infty} B_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \limsup_{n\to\infty} A_n.$$

Το ζητούμενο αποτέλεσμα προκύπτει άμεσα από την Πρόταση 2.3 και την ανισότητα Boole, αφού

$$\begin{aligned}\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) &= \mathbb{P}\left(\lim_{n\to\infty} B_n\right)\\ &= \lim_{n\to\infty} \mathbb{P}(B_n) = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k=n}^{\infty} A_k\right) \leq \lim_{n\to\infty} \sum_{k=n}^{\infty} \mathbb{P}(A_k) = 0,\end{aligned}$$

[4] Βασικά είναι του Borel.

επειδή, από την υπόθεση, η σειρά $\sum_{n=1}^{\infty} \mathbb{P}(A_n)$ συγκλίνει, και επομένως, η ουρά της τείνει στο 0, δηλ. $\lim_{n\to\infty} \sum_{k=n}^{\infty} \mathbb{P}(A_k) = 0$. □

Παράδειγμα 2.6 Στον χώρο πιθανότητας $((0,1), \mathcal{B}((0,1)), \lambda)$, ας θεωρήσουμε τα ενδεχόμενα $A_n = [c_n - \frac{1}{n^2}, c_n + \frac{1}{n^2}]$, όπου οι αυθαίρετοι πραγματικοί αριθμοί c_n ικανοποιούν την $\frac{1}{n^2} < c_n < 1 - \frac{1}{n^2}$ για $n = 2, 3, \ldots$. Τότε, όπως και να εκλεγούν τα c_n,

$$\sum_{n=2}^{\infty} \mathbb{P}(A_n) = \sum_{n=2}^{\infty} \frac{2}{n^2} < \infty,$$

και συνεπώς, εκλέγοντας τυχαία έναν αριθμό στο $(0,1)$, υπάρχει πιθανότητα 1 όπως αυτός συμπεριλαμβάνεται το πολύ σε πεπερασμένα το πλήθος διαστήματα A_n. □

2.2 Στοχαστική Ανεξαρτησία Κλάσεων Ενδεχομένων

Ορισμός 2.7 (δεσμευμένη πιθανότητα του A δοθέντος του B) Για δύο ενδεχόμενα $A, B \in \mathcal{A}$ με $\mathbb{P}(B) > 0$ (σε έναν χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$), ορίζουμε ως δεσμευμένη πιθανότητα (conditional probability) του A, δοθέντος του B, την ποσότητα

$$\mathbb{P}(A|B) \overset{\text{ορ.}}{=\!=} \frac{\mathbb{P}(AB)}{\mathbb{P}(B)}.$$

Μια διαισθητική εξήγηση του γεγονότος ότι το πηλίκο $\mathbb{P}(AB)/\mathbb{P}(B)$ παριστάνει, πράγματι, δεσμευμένη πιθανότητα, μπορεί να δοθεί εύκολα, αν περιοριστούμε στην περίπτωση της **κλασικής πιθανότητας**,

$$\mathbb{P}(A) = \frac{\text{«ευνοϊκές περιπτώσεις»}}{\text{«δυνατές περιπτώσεις»}}, \quad \text{δηλ.} \ \ \mathbb{P}(A) = \frac{\nu(A)}{\nu(\Omega)},$$

όπου $\nu(A) =$ πλήθος στοιχείων του A. Τότε, δοθέντος ότι το ενδεχόμενο B πραγματοποιήθηκε, οι «ευνοϊκές περιπτώσεις» για το A περιορίζονται σε $\nu(AB)$, ενώ οι «δυνατές περιπτώσεις» σε $\nu(B)$, και έτσι, θα ισχύει ότι $\mathbb{P}(A|B) = \nu(AB)/\nu(B)$, πηλίκο το οποίο οδηγεί στον Ορισμό 2.7, διαιρώντας αριθμητή και παρονομαστή με $\nu(\Omega)$. [Για σχετικά παραδείγματα βλ. Χαραλαμπίδης (1990), σελ. 77–80.] Σημειώνουμε ότι στην γενική περίπτωση αυθαίρετου μέτρου πιθανότητας, είναι δύσκολο να ερμηνευθεί διαισθητικά ο Ορισμός 2.7, και, κατά συνέπεια, τον δεχόμαστε **αξιωματικά**.

Εξ′ ορισμού,

$$\mathbb{P}(AB) = \mathbb{P}(A|B)\mathbb{P}(B), \tag{2.3}$$

και επαγωγικά, εάν $\mathbb{P}(A_1A_2\cdots A_{n-1}) > 0$, τότε

$$\mathbb{P}(A_1A_2\cdots A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\cdots\mathbb{P}(A_n|A_1A_2\cdots A_{n-1}).$$

[πολλαπλασιαστικός νόμος των πιθανοτήτων]

Ας σημειωθεί ότι όταν $\mathbb{P}(B) = 0$ τότε και $\mathbb{P}(AB) = 0$ (αφού $AB \subset B$), και η (2.3) ισχύει αν η $\mathbb{P}(A|B)$ ορισθεί αυθαίρετα. Αυτό επιτρέπει να οριστούν δεσμευμένες πιθανότητες και στην περίπτωση που $\mathbb{P}(B) = 0$, αλλά δεν θα μας απασχολήσει περισσότερο στο σημείο αυτό.[5]

Όταν $B \subset \bigcup_{n=1}^{\infty} A_n$, και $A_i A_j = \varnothing$ για $i \neq j$, τότε

$$\mathbb{P}(B) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)\mathbb{P}(B|A_n),$$ [Θεώρημα/τύπος Ολικής Πιθανότητας]

και, επομένως, έχουμε την έκφραση

$$\mathbb{P}(A_j|B) = \frac{\mathbb{P}(B|A_j)\mathbb{P}(A_j)}{\sum_{n=1}^{\infty} \mathbb{P}(A_n)\mathbb{P}(B|A_n)},$$ [τύπος Bayes]

η οποία μας επιτρέπει να υπολογίσουμε την **εκ των υστέρων** πιθανότητα, $\mathbb{P}(A_j|B)$, όταν γνωρίζουμε τις $\mathbb{P}(B|A_n)$, και $\mathbb{P}(A_n)$, για κάθε n.

[Εδώ, φυσικά, υποτίθεται ότι $\mathbb{P}(A_n) > 0$ για κάθε n, και $\mathbb{P}(B) > 0$, ώστε να έχουν νόημα οι $\mathbb{P}(B|A_n)$, και $\mathbb{P}(A_j|B)$.]

Ορισμός 2.8 (ανεξαρτησία πεπερασμένου πλήθους ενδεχομένων) Έστω A και B ενδεχόμενα στον χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$. Τα A, B λέγονται **(στοχαστικά) ανεξάρτητα**[6] όταν

$$\mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B).$$

[Αυτό είναι ισοδύναμο με να απαιτούμε $\mathbb{P}(A|B) = \mathbb{P}(B)$, στην ειδική περίπτωση που $\mathbb{P}(B) > 0$, αλλά αυτός ο ορισμός

[5] βλ. Κεφάλαια ;;, ;;.

[6] (stochastically) independent

δεν απαιτεί $\mathbb{P}(B) > 0$, και είναι συμμετρικός ως προς τα A και B.] Γενικότερα, τα ενδεχόμενα $A_1, A_2, \ldots, A_n \in \mathcal{A}$ καλούνται (πλήρως) ανεξάρτητα όταν

$$\mathbb{P}(A_{i_1} A_{i_2} \cdots A_{i_k}) = \mathbb{P}(A_{i_1}) \mathbb{P}(A_{i_2}) \cdots \mathbb{P}(A_{i_k}),$$

για όλα τα $k \in \{2, \ldots, n\}$, και για όλες τις επιλογές δεικτών $1 \leq i_1 < i_2 < \cdots < i_k \leq n$. Άρα, στην ουσία, η ανεξαρτησία n ενδεχομένων ορίζεται με βάση τις $2^n - n - 1$ παραπάνω ισότητες, οι οποίες μπορούν να γραφούν σε συμπαγή μορφή ως

$$\mathbb{P}(B_1 \cdots B_n) = \mathbb{P}(B_1) \cdots \mathbb{P}(B_n), \tag{2.4}$$

για κάθε επιλογή των $B_j \in \{A_j, \Omega\}$, $j = 1, 2, \ldots, n$.

Όταν τα ενδεχόμενα A, B είναι ανεξάρτητα, είναι εύκολο να διαπιστώσουμε ότι, π.χ., και τα A^c, B είναι ανεξάρτητα, διότι

$$\mathbb{P}(B) = \mathbb{P}(AB) + \mathbb{P}(A^c B) = \mathbb{P}(A) \mathbb{P}(B) + \mathbb{P}(A^c B),$$

οπότε

$$\mathbb{P}(A^c B) = \mathbb{P}(B) - \mathbb{P}(A) \mathbb{P}(B) = [1 - \mathbb{P}(A)] \mathbb{P}(B) = \mathbb{P}(A^c) \mathbb{P}(B).$$

Έτσι, η ανεξαρτησία ορισμένων ενδεχομένων μπορεί να μας εξασφαλίσει (υπό ορισμένες συνθήκες) την ανεξαρτησία περισσοτέρων ενδεχομένων. Αν π.χ., $\varnothing \neq A, B \neq \Omega$, οπότε

$$\sigma(\{A\}) = \{\varnothing, A, A^c, \Omega\}, \quad \sigma(\{B\}) = \{\varnothing, B, B^c, \Omega\},$$

η υπόθεση ότι τα A, B είναι ανεξάρτητα μας εξασφαλίζει ότι τα A', B' είναι ανεξάρτητα **για κάθε επιλογή** $A' \in \sigma(\{A\})$, $B' \in \sigma(\{B\})$. Σε αυτήν την περίπτωση λέμε ότι οι κλάσεις ενδεχομένων $\sigma(\{A\})$ και $\sigma(\{B\})$ είναι **(στοχαστικά) ανεξάρτητες.**

Γενικά, έχουμε τον παρακάτω ορισμό.

Ορισμός 2.9 (ανεξαρτησία πεπερασμένου πλήθους κλάσεων ενδεχομένων) Έστω $\mathfrak{D}_1, \mathfrak{D}_2, \ldots, \mathfrak{D}_n$ μη κενές κλάσεις ενδεχομένων ενός χώρου πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$ (δηλαδή $\varnothing \neq \mathfrak{D}_i \subset \mathcal{A}$ για κάθε $i = 1, 2, \ldots, n$). Λέμε ότι οι $\mathfrak{D}_1, \ldots, \mathfrak{D}_n$ είναι **(στοχαστικά) ανεξάρτητες** όταν τα ενδεχόμενα $A_1, \ldots, A_n$ είναι ανεξάρτητα (δηλ. ικανοποιούν την (2.4)), για κάθε επιλογή των $A_i \in \mathfrak{D}_i$, $i = 1, 2, \ldots, n$. Ισοδύναμα, όταν ισχύει η σχέση

$$\mathbb{P}(B_1 \cdots B_n) = \mathbb{P}(B_1) \cdots \mathbb{P}(B_n),$$

για κάθε $B_1 \in \mathfrak{D}_1 \cup \{\Omega\}, \ldots, B_n \in \mathfrak{D}_n \cup \{\Omega\}$.

Επομένως, η ανεξαρτησία των ενδεχομένων $A_1, \ldots, A_n$ μπορεί να θεωρηθεί ως ανεξαρτησία των κλάσεων $\{A_1\}, \ldots, \{A_n\}$.

Είναι φανερό ότι αν οι $\mathfrak{D}_i$, $i = 1, 2, \ldots, n$, είναι ανεξάρτητες, τότε και οι $\mathfrak{D}'_i$, $i = 1, 2, \ldots, n$, με $\varnothing \neq \mathfrak{D}'_i \subset \mathfrak{D}_i$, είναι ανεξάρτητες. Ενδιαφέρον παρουσιάζει το αντίστροφο φαινόμενο, δηλ. κατά πόσον είναι εφικτό από την ανεξαρτησία των (μικροτέρων) κλάσεων $\mathfrak{D}'_i$ να μπορούμε (κάτω από ορισμένες συνθήκες) να συμπεράνουμε την ανεξαρτησία των (μεγαλυτέρων) κλάσεων $\mathfrak{D}_i$. Αυτό ακριβώς συνέβη προηγουμένως, όπου από την ανεξαρτησία των κλάσεων $\{A\}$, $\{B\}$ (δηλ. των ενδεχομένων A, B) συμπεράναμε την ανεξαρτησία των $\sigma(\{A\})$, $\sigma(\{B\})$.

Από τον ορισμό προκύπτουν άμεσα αρκετές άλλες προφανείς ιδιότητες. Για παράδειγμα, όταν οι $\mathfrak{D}_1, \ldots, \mathfrak{D}_n$ είναι ανεξάρτητες, τότε και οι $\mathfrak{D}_1, \ldots, \mathfrak{D}_{n-1}$ είναι ανεξάρτητες, όπως και οι $\mathfrak{D}_1 \cup \mathfrak{D}_2, \mathfrak{D}_3, \ldots, \mathfrak{D}_n$, κ.ο.κ.

Η ανεξαρτησία οποιουδήποτε πλήθους κλάσεων ορίζεται ως εξής:

Ορισμός 2.10 (ανεξαρτησία οποιουδήποτε πλήθους κλάσεων) Εάν το I είναι αυθαίρετο σύνολο δεικτών, και οι $\mathfrak{D}_i$, $i \in I$, είναι μη κενές κλάσεις ενδεχομένων ενός χώρου πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, δηλ. $\varnothing \neq \mathfrak{D}_i \subset \mathcal{A}$ για κάθε $i \in I$, τότε οι κλάσεις $\mathfrak{D}_i$, $i \in I$, καλούνται **ανεξάρτητες** όταν για οποιονδήποτε ακέραιο n, οι κλάσεις $\mathfrak{D}_{i_1}, \ldots, \mathfrak{D}_{i_n}$ είναι ανεξάρτητες, για οποιαδήποτε επιλογή δεικτών $i_1, \ldots, i_n \in I$, με $i_j \neq i_k$ για $j \neq k$.

Φυσικά, τα ενδεχόμενα A_i, $i \in I$, είναι ανεξάρτητα όταν (εξ' ορισμού) οι κλάσεις $\{A_i\}$, $i \in I$, είναι ανεξάρτητες. Η πιο ενδιαφέρουσα (ειδική) περίπτωση είναι η εξής:

Παρατήρηση 2.11 Η ακολουθία ενδεχομένων $\{A_n,\ n \geq 1\}$ είναι

ανεξάρτητη όταν τα ενδεχόμενα $A_1, \ldots, A_n$ είναι ανεξάρτητα για κάθε $n = 2, 3, \ldots$.

2.3 Θεωρήματα Επέκτασης της Ανεξαρτησίας – Dynkin

Το πρώτο αποτέλεσμα επέκτασης της ανεξαρτησίας που θα αποδείξουμε χρησιμοποιεί τις **παραγόμενες κλάσεις Dynkin** $\delta(\mathfrak{D}_i)$, $i \in I$. Για τον σκοπό αυτό δίνουμε πρώτα τον παρακάτω ορισμό.

Ορισμός 2.12 (κλάση Dynkin) Μία οικογένεια $\mathfrak{D}$ υποσυνόλων του (μη κενού) Ω λέγεται **κλάση Dynkin** στον Ω, όταν:

(i) $\Omega \in \mathfrak{D}$.

(ii) Αν $A, B \in \mathfrak{D}$ και $A \subset B$ τότε $B \smallsetminus A \in \mathfrak{D}$.

(iii) Αν η $\{A_n,\ n \geq 1\}$ είναι αύξουσα ακολουθία στην $\mathfrak{D}$, τότε
$$\bigcup_{n=1}^{\infty} A_n \, (= \lim A_n) \in \mathfrak{D}.$$

Προφανώς μία σ-άλγεβρα είναι κλάση Dynkin, αφού τα (ii) και (iii) ικανοποιούνται για κάθε $A, B \in \mathfrak{D}$, και για κάθε ακολουθία $\{A_n,\ n \geq 1\} \subset \mathfrak{D}$, σε μία σ-άλγεβρα $\mathfrak{D}$. Το αντίστροφο δεν ισχύει, αφού, π.χ., η κλάση
$$\mathfrak{D} = \{\varnothing, \{1,2\}, \{1,3\}, \{1,4\}, \{2,3\}, \{2,4\}, \{3,4\}, \{1,2,3,4\}\}$$
είναι κλάση Dynkin στον $\Omega = \{1, 2, 3, 4\}$, αλλά δεν είναι σ-άλγεβρα (ούτε καν άλγεβρα), διότι, π.χ., $\{1, 2\} \cup \{1, 3\} \notin \mathfrak{D}$.

Πολύ εύκολα αποδεικνύεται η παρακάτω πρόταση.

Πρόταση 2.13 (παραγόμενη κλάση Dynkin) Για τυχούσα οικογένεια $\mathfrak{D} \subset \mathcal{P}(\Omega)$, υπάρχει η ελάχιστη κλάση Dynkin που περιέχει την $\mathfrak{D}$, η οποία ονομάζεται **κλάση Dynkin** που **παράγεται** από την $\mathfrak{D}$, και συμβολίζεται με $\delta(\mathfrak{D})$.

Απόδειξη: Όπως στην Πρόταση 1.14 για σ-άλγεβρες, αφού η τομή οποιουδήποτε πλήθους κλάσεων Dynkin είναι κλάση Dynkin, και η $\mathcal{P}(\Omega)$ είναι κλάση Dynkin που περιέχει την $\mathfrak{D}$. $\square$

Επειδή για τυχούσα κλάση $\mathfrak{D} \subset \mathcal{P}(\Omega)$, η $\sigma(\mathfrak{D})$ είναι κλάση Dynkin που περιέχει την $\mathfrak{D}$, έπεται ότι γενικά

$$\delta(\mathfrak{D}) \subset \sigma(\mathfrak{D}).$$

Το αντίστροφο δεν ισχύει φυσικά (βλ. και το παράδειγμα μετά τον Ορισμό 2.12 της κλάσης Dynkin), εκτός αν υποτεθούν περισσότερες συνθήκες στην $\mathfrak{D}$.

Έχουμε τώρα το εξής σημαντικό θεώρημα.

Θεώρημα 2.14 (ανεξαρτησία των κλάσεων Dynkin) Εάν οι κλάσεις ενδεχομένων $\mathfrak{D}_i,\ i \in I$, ενός χώρου πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, είναι ανεξάρτητες, τότε και οι

$$\delta(\mathfrak{D}_i),\ i \in I,$$

είναι ανεξάρτητες.

Απόδειξη: Βάσει των ορισμών, αρκεί να δείξουμε ότι αν οι $\mathfrak{D}_1, \ldots, \mathfrak{D}_n$ είναι ανεξάρτητες, τότε και οι $\delta(\mathfrak{D}_1), \ldots, \delta(\mathfrak{D}_n)$ είναι ανεξάρτητες. Επομένως, αρκεί να αποδείξουμε ότι αν οι $\mathfrak{D}_1, \ldots, \mathfrak{D}_n$ είναι ανεξάρτητες, τότε είναι ανεξάρτητες και οι

$$\delta(\mathfrak{D}_1), \mathfrak{D}_2, \ldots, \mathfrak{D}_n.$$

[Οπότε, εφαρμόζοντάς το επαγωγικά θα προκύψει η αποδεικτέα.]

Θεωρούμε την οικογένεια (των «καλών συνόλων»)

$$\begin{aligned}\mathcal{D}_1 = \{A \in \mathcal{A} : \mathbb{P}(AA_2 \cdots A_n) = \mathbb{P}(A)\mathbb{P}(A_2) \cdots \mathbb{P}(A_n) \\ \text{για κάθε } A_2 \in \mathfrak{D}_2 \cup \{\Omega\}, \ldots, A_n \in \mathfrak{D}_n \cup \{\Omega\}\}.\end{aligned}$$

Δηλ. η $\mathcal{D}_1$ είναι η κλάση που περιέχει όλα τα ενδεχόμενα που είναι ανεξάρτητα από τις $\mathfrak{D}_2, \ldots, \mathfrak{D}_n$, και άρα η $\mathcal{D}_1$ είναι, στην ουσία, η μέγιστη κλάση για την οποία οι $\mathcal{D}_1, \mathfrak{D}_2, \ldots, \mathfrak{D}_n$ είναι ανεξάρτητες.

Εξ' ορισμού, $\mathfrak{D}_1 \subset \mathcal{D}_1$. Επίσης παρατηρούμε ότι η $\mathcal{D}_1$ είναι κλάση Dynkin, διότι

(i) $\Omega \in \mathcal{D}_1$. [Αφού $\mathbb{P}(\Omega A_2 \cdots A_n) = \mathbb{P}(A_2) \cdots \mathbb{P}(A_n)$ για κάθε $A_2 \in \mathfrak{D}_2 \cup \{\Omega\}, \ldots, A_n \in \mathfrak{D}_n \cup \{\Omega\}$.]

(ii) Αν $A, B \in \mathcal{D}_1$ με $A \subset B$, τότε για κάθε $A_2 \in \mathfrak{D}_2 \cup$

$\{\Omega\}, \ldots, A_n \in \mathfrak{D}_n \cup \{\Omega\}$,

$$\begin{aligned} \mathbb{P}(AA_2 \cdots A_n) &= \mathbb{P}(A)\mathbb{P}(A_2)\cdots\mathbb{P}(A_n), \text{ και} \\ \mathbb{P}(BA_2 \cdots A_n) &= \mathbb{P}(B)\mathbb{P}(A_2)\cdots\mathbb{P}(A_n). \end{aligned}$$

Άρα,

$$\begin{aligned} \mathbb{P}((B \setminus A)A_2 \cdots A_n) &= \mathbb{P}(BA_2 \cdots A_n \setminus AA_2 \cdots A_n) \\ &= \mathbb{P}(BA_2 \cdots A_n) - \mathbb{P}(AA_2 \cdots A_n) \qquad [\text{διότι } AA_2 \cdots A_n \subset BA_2 \cdots A_n] \\ &= \mathbb{P}(B)\mathbb{P}(A_2)\cdots\mathbb{P}(A_n) - \mathbb{P}(A)\mathbb{P}(A_2)\cdots\mathbb{P}(A_n) \\ &= [\mathbb{P}(B) - \mathbb{P}(A)]\,\mathbb{P}(A_2)\cdots\mathbb{P}(A_n) \\ &= \mathbb{P}(B \setminus A)\mathbb{P}(A_2)\cdots\mathbb{P}(A_n), \end{aligned}$$

και συνεπώς, $B \setminus A \in \mathcal{D}_1$.

(iii) Αν η $\{B_m,\ m \geq 1\}$ είναι αύξουσα ακολουθία της $\mathcal{D}_1$, τότε για κάθε $A_2 \in \mathfrak{D}_2 \cup \{\Omega\}, \ldots, A_n \in \mathfrak{D}_n \cup \{\Omega\}$,

$$\mathbb{P}(B_m A_2 \cdots A_n) = \mathbb{P}(B_m)\mathbb{P}(A_2)\cdots\mathbb{P}(A_n),\ \ m = 1, 2, \ldots\ .$$

Τότε έπεται ότι

$$\begin{aligned}
\mathbb{P}\left(\left(\bigcup_{m=1}^{\infty} B_m\right) A_2 \cdots A_n\right) &= \mathbb{P}\left(\bigcup_{m=1}^{\infty} (B_m A_2 \cdots A_n)\right) \\
&= \mathbb{P}\Big[\lim_{m\to\infty} (B_m A_2 \cdots A_n)\Big] \qquad [\text{διότι η } B_m A_2 \cdots A_n \text{ είναι αύξουσα}] \\
&= \lim_{m\to\infty} \mathbb{P}(B_m A_2 \cdots A_n) \\
&= \lim_{m\to\infty} [\mathbb{P}(B_m)\mathbb{P}(A_2)\cdots \mathbb{P}(A_n)] \\
&= \mathbb{P}(A_2)\cdots \mathbb{P}(A_n) \lim_{m\to\infty} \mathbb{P}(B_m) \\
&= \mathbb{P}(A_2)\cdots \mathbb{P}(A_n)\mathbb{P}\Big(\lim_{m\to\infty} B_m\Big) \\
&= \mathbb{P}(A_2)\cdots \mathbb{P}(A_n)\mathbb{P}\left(\bigcup_{m=1}^{\infty} B_m\right).
\end{aligned}$$

Επομένως, $\bigcup_{m=1}^{\infty} B_m \in \mathcal{D}_1$.

Άρα η $\mathcal{D}_1$ είναι κλάση Dynkin, και $\mathfrak{D}_1 \subset \mathcal{D}_1$. Συνεπώς, $\delta(\mathfrak{D}_1) \subset \mathcal{D}_1$, επειδή η $\delta(\mathfrak{D}_1)$ είναι, εξ' ορισμού, η ελάχιστη κλάση Dynkin που περιέχει την $\mathfrak{D}_1$. Επομένως, για κάθε $A \in \delta(\mathfrak{D}_1), A_2 \in \mathfrak{D}_2 \cup \{\Omega\}, \ldots, A_n \in \mathfrak{D}_n \cup \{\Omega\}$,

$$\mathbb{P}(AA_2 \cdots A_n) = \mathbb{P}(A)\mathbb{P}(A_2)\cdots \mathbb{P}(A_n),$$

που σημαίνει ότι οι κλάσεις $\delta(\mathfrak{D}_1), \mathfrak{D}_2, \ldots, \mathfrak{D}_n$ είναι ανεξάρτητες, και το θεώρημα απεδείχθη. $\square$

Η παραπάνω τεχνική απόδειξης είναι **κλασική** στην θεωρία μέτρου, και θα χρησιμοποιηθεί αρκετές φορές στην συνέχεια. Σε γενικές γραμμές, μπορεί να περιγραφεί ως εξής:

Κλασική τεχνική απόδειξη της θεωρίας μέτρου. *Ας υποθέσουμε ότι θέλουμε να αποδείξουμε την ισχύ μιας πρότασης για μία μεγάλη οικογένεια συνόλων, $\mathcal{A}$, όταν είναι γνωστό ότι η πρόταση ισχύει σε μία μικρότερη οικογένεια, $\mathfrak{D}$, με $\mathfrak{D} \subset \mathcal{A}$. Τότε, θεωρούμε την οικογένεια των* «καλών συνόλων»*, $\mathcal{B}$, με $\mathcal{B} \subset \mathcal{A}$, στα οποία* (εξ' ορισμού) *αληθεύει η πρόταση, οπότε $\mathfrak{D} \subset \mathcal{B} \subset \mathcal{A}$. Πολλές φορές είναι εύκολο να αποδείξουμε ότι η $\mathcal{B}$ είναι κλάση με γνωστές,* «καλές»*, ιδιότητες* (συνήθως σ-άλγεβρα ή κλάση Dynkin)*, ας πούμε κ-κλάση* (από το $\kappa =$ «καλή»)*. Αν μπορεί επίσης να δειχθεί, ή είναι γνωστό εκ' των προτέρων, ότι η ελάχιστη κ-κλάση που περιέχει την $\mathfrak{D}$ είναι η $\mathcal{A}$, δηλ. $\kappa(\mathfrak{D}) = \mathcal{A}$, τότε η αλήθεια της ζητούμενης πρότασης προκύπτει από το γεγονός ότι $(\mathcal{A} =)\ \kappa(\mathfrak{D}) \subset \mathcal{B}$, η οποία ισχύει επειδή η $\mathcal{B}$ είναι κ-κλάση που περιέχει την $\mathfrak{D}$, ενώ η $\kappa(\mathfrak{D})$ είναι, εξ' ορισμού, η ελάχιστη κ-κλάση που περιέχει την $\mathfrak{D}$.*[7]

Όπως τονίσαμε προηγουμένως, μία σ-άλγεβρα είναι πάντα κλάση Dynkin, χωρίς να ισχύει γενικά το αντίστροφο. Η αλήθεια είναι ότι το μόνο που χρειαζόμαστε επιπλέον, ώστε μία

[7] Ο αναγνώστης δεν θα δυσκολευτεί να παρατηρήσει ότι η ίδια τεχνική απόδειξης μπορεί να εφαρμοστεί και σε κάπως γενικότερο πλαίσιο, χωρίς να απαιτείται η $\mathcal{A}$ να είναι κ-κλάση. Συγκεκριμένα, δεν χρειάζεται να υποθέσουμε ότι $\mathfrak{D} \subset \mathcal{A}$, ούτε καν ότι η $\mathcal{B}$, η κλάση των «καλών συνόλων» (με $\mathfrak{D} \subset \mathcal{B}$), περιέχεται στην $\mathcal{A}$. Το μόνο που απαιτείται είναι να ισχύει η σχέση $\mathcal{A} \subset \kappa(\mathfrak{D})$, και το ότι η $\mathcal{B}$ είναι κ-κλάση. Σημειώνεται, πάντως, ότι η γενικευμένη αυτή τεχνική απόδειξης δεν θα μας χρειαστεί στα επόμενα.

κλάση Dynkin να είναι και σ-άλγεβρα είναι η ιδιότητα της κλειστότητας των πεπερασμένων τομών, δηλαδή:

$$\text{Αν } A, B \in \mathfrak{D}, \text{ τότε } AB \in \mathfrak{D}. \tag{π}$$

Ορισμός 2.15 (π-σύστημα) Μία οικογένεια $\mathfrak{D}$ που ικανοποιεί την (π) ονομάζεται π-σύστημα.

Γενικά ισχύει η παρακάτω πρόταση.

Πρόταση 2.16 Μία τυχούσα κλάση $\mathfrak{D} \subset \mathcal{P}(\Omega)$ είναι σ-άλγεβρα όταν και μόνο όταν είναι κλάση Dynkin και π-σύστημα.

Απόδειξη: Προφανώς, αν η $\mathfrak{D}$ είναι σ-άλγεβρα, τότε είναι κλάση Dynkin και π-σύστημα. Αντιστρόφως, αν η $\mathfrak{D}$ είναι κλάση Dynkin και π-σύστημα, τότε

(i) $\Omega \in \mathfrak{D}$, και άρα $\varnothing = \Omega \setminus \Omega \in \mathfrak{D}$. [Επειδή $\Omega \in \mathfrak{D}$, και $\Omega \subset \Omega$.]

(ii) $A^c \in \mathfrak{D}$ για κάθε $A \in \mathfrak{D}$, διότι $A^c = \Omega \setminus A$, $A \subset \Omega$, και $A, \Omega \in \mathfrak{D}$.

(iii) Αν τώρα η $\{A_n,\ n \geq 1\}$ είναι τυχούσα ακολουθία στην $\mathfrak{D}$, τότε θέτοντας $B_n = \bigcup_{j=1}^n A_j = (\bigcap_{j=1}^n A_j^c)^c$, έχουμε ότι η $\{B_n,\ n \geq 1\}$ είναι αύξουσα, και ότι $B_n \in \mathfrak{D}$ για κάθε $n = 1, 2, \ldots$, διότι από το (ii), $A_j^c \in \mathfrak{D}$, και άρα $A_1^c \cap \cdots \cap A_n^c \in \mathfrak{D}$ (αφού η $\mathfrak{D}$ είναι π-σύστημα), οπότε, και πάλι από το (ii), $(A_1^c \cap \cdots \cap A_n^c)^c \in \mathfrak{D}$. Επομένως, $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n \in \mathfrak{D}$, διότι η $\mathfrak{D}$ είναι κλάση Dynkin, και η $\{B_n,\ n \geq 1\}$ είναι αύξουσα ακολουθία της $\mathfrak{D}$. $\square$

Είμαστε τώρα σε θέση να δείξουμε το εξής σημαντικό αποτέλεσμα.

Θεώρημα 2.17 (Dynkin) Αν η $\mathfrak{D}$ είναι π-σύστημα τότε $\delta(\mathfrak{D}) = \sigma(\mathfrak{D})$.

Απόδειξη: Επειδή πάντα ισχύει ότι $\delta(\mathfrak{D}) \subset \sigma(\mathfrak{D})$, είναι αρκετό να δείξουμε ότι και $\sigma(\mathfrak{D}) \subset \delta(\mathfrak{D})$, όταν η $\mathfrak{D}$ είναι π-σύστημα. Επομένως είναι αρκετό να δείξουμε ότι η $\delta(\mathfrak{D})$ είναι επίσης π-σύστημα, διότι τότε, σύμφωνα με την Πρόταση 2.16, η $\delta(\mathfrak{D})$ θα είναι π-σύστημα και κλάση Dynkin, και άρα η $\delta(\mathfrak{D})$ θα είναι σ-άλγεβρα που περιέχει την $\mathfrak{D}$, οπότε θα προκύψει ότι και $\sigma(\mathfrak{D}) \subset \delta(\mathfrak{D})$, από τον ορισμό της $\sigma(\mathfrak{D})$.

Για τον σκοπό αυτό, θεωρούμε πρώτα την οικογένεια

$$\overline{\mathfrak{D}} = \{A \subset \Omega : \text{για κάθε } C \in \mathfrak{D} \text{ ισχύει ότι } AC \in \delta(\mathfrak{D})\}.$$

Προφανώς $\mathfrak{D} \subset \overline{\mathfrak{D}}$, διότι αν $\Delta \in \mathfrak{D}$, τότε για κάθε $C \in \mathfrak{D}$, $\Delta C \in \mathfrak{D} \subset \delta(\mathfrak{D})$, αφού η $\mathfrak{D}$ είναι, από υπόθεση, π-σύστημα.

Η $\overline{\mathfrak{D}}$ είναι κλάση Dynkin στον Ω, διότι:

(i) $\Omega \in \overline{\mathfrak{D}}$. [Αφού για κάθε $C \in \mathfrak{D}$, $\Omega C = C \in \mathfrak{D} \subset \delta(\mathfrak{D})$.]

(ii) Αν $A, B \in \overline{\mathfrak{D}}$ και $A \subset B$, τότε για κάθε $C \in \mathfrak{D}, AC \in \delta(\mathfrak{D})$, και $BC \in \delta(\mathfrak{D})$. Συνεπώς, αφού $AC \subset BC$ και η $\delta(\mathfrak{D})$ είναι κλάση Dynkin, $BC \setminus AC \in \delta(\mathfrak{D})$, δηλ. για κάθε $C \in \mathfrak{D}$, $(B \setminus A)C \in \delta(\mathfrak{D})$, που σημαίνει ότι $B \setminus A \in \overline{\mathfrak{D}}$.

(iii) Αν τώρα η $\{A_n,\ n \geq 1\}$ είναι αύξουσα ακολουθία της $\overline{\mathfrak{D}}$, τότε για κάθε $C \in \mathfrak{D}$, $A_nC \in \delta(\mathfrak{D})$, $n = 1, 2, \dots$. Επομένως, για κάθε $C \in \mathfrak{D}$, η ακολουθία $\{A_nC,\ n \geq 1\}$ είναι αύξουσα ακολουθία της $\delta(\mathfrak{D})$, και συνεπώς, αφού η $\delta(\mathfrak{D})$ είναι κλάση Dynkin, $\bigcup_{n=1}^{\infty} (A_nC) \in \delta(\mathfrak{D})$. Αυτό σημαίνει ότι $\left(\bigcup_{n=1}^{\infty} A_n\right) C \in \delta(\mathfrak{D})$ (για κάθε $C \in \mathfrak{D}$), και άρα, από τον ορισμό της $\overline{\mathfrak{D}}$, $\bigcup_{n=1}^{\infty} A_n \in \overline{\mathfrak{D}}$.

Από τα παραπάνω προκύπτει ότι $\mathfrak{D} \subset \overline{\mathfrak{D}}$, και η $\overline{\mathfrak{D}}$ είναι κλάση Dynkin, που σημαίνει ότι και $\delta(\mathfrak{D}) \subset \overline{\mathfrak{D}}$. Άρα, αν $\Delta \in \delta(\mathfrak{D})$, τότε για κάθε $C \in \mathfrak{D}$, $\Delta C \in \delta(\mathfrak{D})$. [Από τον ορισμό της $\overline{\mathfrak{D}}$.] Με άλλα λόγια, έχει μέχρι στιγμής αποδειχθεί το εξής:

Όταν ένα από τα A, B ανήκει στην $\mathfrak{D}$, και το άλλο στην $\delta(\mathfrak{D})$, τότε $AB \in \delta(\mathfrak{D})$. $(\star)$

Θεωρούμε τώρα την οικογένεια

$$\overline{\delta(\mathfrak{D})} = \{A \subset \Omega : \text{για κάθε } C \in \delta(\mathfrak{D}) \text{ ισχύει ότι } AC \in \delta(\mathfrak{D})\},$$

οπότε, σύμφωνα με τα προηγούμενα, $\mathfrak{D} \subset \overline{\delta(\mathfrak{D})}$, διότι, για κάθε $\Delta \in \mathfrak{D}$, και για κάθε $C \in \delta(\mathfrak{D})$, $\Delta C \in \delta(\mathfrak{D})$, σύμφωνα με την ($\star$). Όμως η $\overline{\delta(\mathfrak{D})}$ είναι επίσης κλάση Dynkin, διότι:

(i) $\Omega \in \overline{\delta(\mathfrak{D})}$. [Αφού για κάθε $C \in \delta(\mathfrak{D})$, $\Omega C = C \in \delta(\mathfrak{D})$.]

(ii) Αν $A, B \in \overline{\delta(\mathfrak{D})}$ με $A \subset B$, τότε, για τυχόν $C \in \delta(\mathfrak{D})$, $AC \in \delta(\mathfrak{D})$ και $BC \in \delta(\mathfrak{D})$, οπότε $BC \setminus AC = (B \setminus A)C \in \delta(\mathfrak{D})$ (αφού $AC \subset BC$ και η $\delta(\mathfrak{D})$ είναι κλάση Dynkin), και συνεπώς, $B \setminus A \in \overline{\delta(\mathfrak{D})}$.

(iii) Τέλος, αν η $\{A_n,\ n \geq 1\}$ είναι αύξουσα ακολουθία της $\overline{\delta(\mathfrak{D})}$, τότε, για οποιοδήποτε (σταθερό) $C \in \delta(\mathfrak{D})$, η $\{A_nC,\ n \geq 1\}$ είναι αύξουσα ακολουθία της $\delta(\mathfrak{D})$, και συνεπώς, $\bigcup_{n=1}^{\infty} (A_nC) = \left(\bigcup_{n=1}^{\infty} A_n\right) C \in \delta(\mathfrak{D})$ (επειδή η $\delta(\mathfrak{D})$ είναι κλάση Dynkin), που σημαίνει ότι $\bigcup_{n=1}^{\infty} A_n \in \overline{\delta(\mathfrak{D})}$.

Αφού λοιπόν $\mathfrak{D} \subset \overline{\delta(\mathfrak{D})}$ και η $\overline{\delta(\mathfrak{D})}$ είναι κλάση Dynkin, συμπεραίνουμε ότι $\delta(\mathfrak{D}) \subset \overline{\delta(\mathfrak{D})}$. Αυτό σημαίνει ότι αν $\varDelta \in \delta(\mathfrak{D})$ τότε $\varDelta \in \overline{\delta(\mathfrak{D})}$. Δηλαδή, από τον ορισμό της $\overline{\delta(\mathfrak{D})}$, προκύπτει το εξής:

Αν $\varDelta \in \delta(\mathfrak{D})$ τότε για κάθε $C \in \delta(\mathfrak{D})$, $\varDelta C \in \delta(\mathfrak{D})$.

Επομένως, η κλάση Dynkin $\delta(\mathfrak{D})$ είναι κλειστή ως προς τις πεπερασμένες τομές (δηλ. π-σύστημα), και άρα είναι σ-άλγεβρα, σύμφωνα με την Πρόταση 2.16. $\square$

Το παραπάνω θεώρημα έχει σπουδαίες εφαρμογές σε θεωρήματα επέκτασης. Για να γίνει κατανοητή η σπουδαιότητά του, δίνουμε πρώτα τον ορισμό της λ-κλάσης, και στην συνέχεια αποδεικνύουμε την μοναδικότητα της επέκτασης ενός μέτρου πιθανότητας στην σ-άλγεβρα που παράγεται από ένα π-σύστημα.

Ορισμός 2.18 (λ-κλάση) Μία οικογένεια $\mathfrak{D} \subset \mathcal{P}(\Omega)$ ονομάζεται **λ-κλάση** όταν ικανοποιεί τα εξής:

(i) $\Omega \in \mathfrak{D}$.

(ii) Αν $A \in \mathfrak{D}$ τότε $A^c \in \mathfrak{D}$.

(iii) Αν η $\{A_n,\ n \geq 1\}$ είναι ακολουθία υποσυνόλων της $\mathfrak{D}$, με $A_i A_j = \varnothing$ για $i \neq j$, τότε $\bigcup_{n=1}^{\infty} A_n \in \mathfrak{D}$.

Είναι φανερό ότι μία σ-άλγεβρα είναι πάντοτε λ-κλάση, ενώ το αντίστροφο δεν ισχύει πάντα (βρείτε αντιπαράδειγμα). Επίσης, ισχύει η εξής πρόταση:

Πρόταση 2.19 Μία λ-κλάση είναι πάντα κλάση Dynkin.

Απόδειξη: Άσκηση 2.1. □

Επομένως, από το Θεώρημα του Dynkin έχουμε αμέσως το εξής πόρισμα.

Πόρισμα 2.20 Αν η σ-άλγεβρα $\mathcal{A}$ παράγεται από ένα π-σύστημα

$\mathfrak{D}$, δηλ. $\mathcal{A} = \sigma(\mathfrak{D})$, και δύο μέτρα πιθανότητας $\mathbb{P}_1$, $\mathbb{P}_2$ συμφωνούν στην $\mathfrak{D}$, τότε θα συμφωνούν και στην $\mathcal{A}$ (δηλ. $\mathbb{P}_1 \equiv \mathbb{P}_2$).

Απόδειξη: Σύμφωνα με την κλασική τεχνική, θεωρούμε την κλάση των «καλών συνόλων»

$$\mathcal{D} = \{A \in \mathcal{A} : \mathbb{P}_1(A) = \mathbb{P}_2(A)\} \subset \mathcal{A}.$$

Από την υπόθεση, $\mathfrak{D} \subset \mathcal{D}$ και η $\mathfrak{D}$ είναι π-σύστημα. Είναι εύκολο να δείξουμε ότι η $\mathcal{D}$ είναι λ-κλάση. Πράγματι,

(i) $\Omega \in \mathcal{D}$, αφού $\Omega \in \mathcal{A}$ και $\mathbb{P}_1(\Omega) = \mathbb{P}_2(\Omega) = 1$.

(ii) Αν $A \in \mathcal{D}$ τότε $A \in \mathcal{A}$ και $\mathbb{P}_1(A) = \mathbb{P}_2(A)$, οπότε $\mathbb{P}_1(A^c) = 1 - \mathbb{P}_1(A) = 1 - \mathbb{P}_2(A) = \mathbb{P}_2(A^c)$, που σημαίνει ότι $A^c \in \mathcal{D}$.

(iii) Αν $\{A_n,\ n \geq 1\} \subset \mathcal{D}$ και $A_i A_j = \varnothing$ για $i \neq j$, τότε $A_n \in \mathcal{A}$ για κάθε n, και $\mathbb{P}_1(A_n) = \mathbb{P}_2(A_n)$. Άρα, $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ και

$$\mathbb{P}_1\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}_1(A_n) = \sum_{n=1}^{\infty} \mathbb{P}_2(A_n) = \mathbb{P}_2\left(\bigcup_{n=1}^{\infty} A_n\right),$$

που σημαίνει ότι $\bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$.

Αφού η $\mathcal{D}$ είναι λ-κλάση, θα είναι και κλάση Dynkin (Πρόταση 2.19), και επειδή $\mathfrak{D} \subset \mathcal{D}$, έπεται ότι και $\delta(\mathfrak{D}) \subset \mathcal{D}$. Όμως $\delta(\mathfrak{D}) = \sigma(\mathfrak{D})$ από το Θεώρημα 2.17 του Dynkin, που σημαίνει ότι $\mathcal{A} = \sigma(\mathfrak{D}) \subset \mathcal{D}$, δηλ. $\mathcal{D} = \mathcal{A}$, και συνεπώς, $\mathbb{P}_1(A) = \mathbb{P}_2(A)$ για κάθε $A \in \mathcal{A}$. $\square$

Παράδειγμα 2.21 Έστω $((0,1), \mathcal{B}((0,1)))$ ο μετρήσιμος χώρος των Borel υποσυνόλων του $(0,1)$. Τότε,

$$\mathcal{B}((0,1)) = \sigma(\mathfrak{D}_1),$$

όπου $\mathfrak{D}_1 = \{(0,y],\ 0 \le y < 1\}$ (βλ. Πρόταση 1.15). Επειδή προφανώς η $\mathfrak{D}_1$ είναι π-σύστημα (διότι $(0,y_1] \cap (0,y_2] = (0, \min\{y_1,y_2\}] \in \mathfrak{D}_1$), και το μέτρο Lebesgue $\mathbb{P}_1 = \lambda$ ικανοποιεί την $\lambda((0,y]) = y$, έπεται ότι οποιοδήποτε άλλο μέτρο, έστω $\mathbb{P}_2$, ορισμένο στην $\mathcal{B}((0,1))$, με $\mathbb{P}_2((0,y]) = y$, θα ταυτίζεται αναγκαστικά με το μέτρο Lebesgue στον $((0,1), \mathcal{B}((0,1)))$. Άρα το μέτρο Lebesgue **είναι το μοναδικό μέτρο πιθανότητας** που θα μπορούσαμε να ορίσουμε στον μετρήσιμο χώρο $((0,1), \mathcal{B}((0,1)))$, έτσι ώστε κάθε διάστημα να έχει πιθανότητα ίση με το μήκος του. $\square$

Πόρισμα 2.22 Αν οι $\mathfrak{D}_i,\ i \in I$, είναι ανεξάρτητες κλάσεις ενδεχομένων και π-συστήματα (για κάθε $i \in I$), τότε και οι $\sigma(\mathfrak{D}_i),\ i \in I$, είναι ανεξάρτητες.

Απόδειξη: Από το Θεώρημα 2.14 οι $\delta(\mathfrak{D}_i),\ i \in I$, είναι ανεξάρτητες, που, λόγω του θεωρήματος Dynkin, ισοδυναμεί με το ότι οι $\sigma(\mathfrak{D}_i),\ i \in I$, είναι ανεξάρτητες (επειδή οι $\mathfrak{D}_i$ είναι π-συστήματα). $\square$

Πόρισμα 2.23 Θεωρούμε τα ανεξάρτητα ενδεχόμενα

$$\begin{array}{cccc} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ & \vdots & & \\ A_{n1} & A_{n2} & A_{n3} & \cdots \\ & \vdots & & \end{array}$$

(όπου κάθε γραμμή μπορεί να είναι πεπερασμένη ή αριθμήσιμα άπειρη και υπάρχουν πεπερασμένες ή αριθμήσιμα άπειρες γραμμές). Θεωρούμε τις κλάσεις

$$\mathcal{A}_n = \sigma\left(\{A_{n1}, A_{n2}, \ldots\}\right), \quad n = 1, 2, \ldots,$$

έτσι ώστε η $\mathcal{A}_n$ να είναι η σ-άλγεβρα που παράγεται από την n-οστή γραμμή. Τότε οι σ-άλγεβρες $\mathcal{A}_n$, $n \geq 1$, είναι ανεξάρτητες.

Απόδειξη: Θέτουμε

$$\mathfrak{D}_n = \{A : A = A_{nj_1} A_{nj_2} \cdots A_{nj_k} \text{ για κάποιο } k, \text{ και για } 1 \leq j_1 < \cdots < j_k\},$$

δηλ. $\mathfrak{D}_n$ είναι η οικογένεια όλων των πεπερασμένων τομών ενδεχομένων της n-οστής γραμμής. Προφανώς κάθε $\mathfrak{D}_n$ είναι π-σύστημα, και επειδή $\{A_{n1}, A_{n2}, \ldots\} \subset \mathfrak{D}_n$ και $\mathfrak{D}_n \subset \sigma\left(\{A_{n1}, A_{n2}, \ldots\}\right)$, έπεται ότι

$$\sigma(\mathfrak{D}_n) = \sigma\left(\{A_{n1}, A_{n2}, \ldots\}\right).$$

Αν λοιπόν αποδείξουμε ότι οι $\mathfrak{D}_n,\ n \geq 1$, που είναι π-συστήματα, είναι ανεξάρτητες, τότε θα προκύψει, από το Θεώρημα 2.14 και το Θεώρημα του Dynkin, ότι και οι $\delta(\mathfrak{D}_n) = \sigma(\mathfrak{D}_n),\ n \geq 1$, είναι ανεξάρτητες.

Έστω λοιπόν αυθαίρετοι δείκτες $1 \leq n_1 < \cdots < n_k$. Θεωρούμε ένα $B_{n_1} \in \mathfrak{D}_{n_1}$, ένα $B_{n_2} \in \mathfrak{D}_{n_2}$, ..., ένα $B_{n_k} \in \mathfrak{D}_{n_k}$. Τότε, το B_{n_j} θα είναι της μορφής

$$B_{n_j} = A_{n_j i_1(j)} \cap A_{n_j i_2(j)} \cap \cdots \cap A_{n_j i_{s(j)}(j)},$$

για $s(j)$ δείκτες $1 \leq i_1(j) < i_2(j) < \cdots < i_{s(j)}(j)$, δηλαδή

$$B_{n_j} = \bigcap_{r=1}^{s(j)} A_{n_j i_r(j)}, \quad j = 1, 2, \ldots, k.$$

Επομένως,

$$\begin{aligned} \mathbb{P}\left(\bigcap_{j=1}^{k} B_{n_j}\right) &= \mathbb{P}\left(\bigcap_{j=1}^{k}\bigcap_{r=1}^{s(j)} A_{n_j i_r(j)}\right) \\ &= \prod_{j=1}^{k}\prod_{r=1}^{s(j)} \mathbb{P}\left(A_{n_j i_r(j)}\right) \qquad [\text{λόγω ανεξαρτησίας}] \\ &= \prod_{j=1}^{k} \mathbb{P}(B_{n_j}), \end{aligned}$$

και συνεπώς, οι κλάσεις $\mathfrak{D}_n,\ n \geq 1$, είναι ανεξάρτητες. □

2.4 Οριακή Συμπεριφορά Ανεξάρτητης Ακολουθίας Ενδεχομένων

Η οριακή συμπεριφορά ($n \to \infty$) μιας ακολουθίας ενδεχομένων $\{A_n,\ n \geq 1\}$ περιγράφεται από την τελική τους σ-άλγεβρα:

Ορισμός 2.24 (τελική σ-άλγεβρα ακολουθίας ενδεχομένων) Έστω $\{A_n,\ n \geq 1\}$ μία (αυθαίρετη) ακολουθία ενδεχομένων στον χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$. Τότε, η σ-άλγεβρα

$$\mathcal{T}\left(\{A_n,\ n \geq 1\}\right) \stackrel{\text{ορ.}}{=\!=} \bigcap_{n=1}^{\infty} \sigma\left(\{A_n, A_{n+1}, \ldots\}\right)$$

ονομάζεται **τελική σ-άλγεβρα** (tail σ-field) της ακολουθίας $\{A_n,\ n \geq 1\}$.

Παρατηρούμε ότι $\mathcal{T}\left(\{A_n,\ n \geq 1\}\right) \subset \mathcal{A}$, και ότι η $\mathcal{T}\left(\{A_n,\ n \geq 1\}\right)$, ως τομή σ-αλγεβρών, είναι πράγματι σ-άλγεβρα. Επειδή

$$\sigma\left(\{A_{n+1}, A_{n+2}, \ldots\}\right) \subset \sigma\left(\{A_n, A_{n+1}, \ldots\}\right),$$

έπεται ότι

$$\mathcal{T}\left(\{A_n,\ n \geq 1\}\right) = \lim_{n \to \infty} \sigma\left(\{A_n, A_{n+1}, \ldots\}\right).$$

Ο όρος *τελική* προέρχεται από το γεγονός ότι για δύο ακολουθίες $\{A_n,\ n \geq 1\}$ και $\{B_n,\ n \geq 1\}$, με $A_n = B_n$ για $n \geq n_0$,

$$\mathcal{T}\left(\{A_n,\ n \geq 1\}\right) = \mathcal{T}\left(\{B_n,\ n \geq 1\}\right),$$

και συνεπώς, η τελική σ-άλγεβρα εξαρτάται μόνο από την **τελική συμπεριφορά** της ακολουθίας ενδεχομένων. Είναι φανερό ότι ένα ενδεχόμενο $A \in \mathcal{A}$ περιέχεται στην $\mathcal{T}\left(\{A_n,\ n \geq 1\}\right)$ αν και μόνο αν το $A \in \sigma\left(\{A_n, A_{n+1}, \ldots\}\right)$ για κάθε $n = 1, 2, \ldots$.

Τα οριακά σύνολα $\liminf A_n$ και $\limsup A_n$ (και, φυσικά, το $\lim A_n$ όταν υπάρχει) ανήκουν στην $\mathcal{T}(\{A_n,\ n \geq 1\})$, διότι, π.χ., $\limsup A_n = \bigcap_{n=m}^{\infty} \bigcup_{k=n}^{\infty} A_k$ για κάθε $m = 1, 2, \ldots$, και συνεπώς, $\limsup A_n \in \sigma(\{A_m, A_{m+1}, \ldots\})$ για κάθε $m = 1, 2, \ldots$, δηλ. $\limsup A_n \in \mathcal{T}(\{A_n,\ n \geq 1\})$.

Για τα ενδεχόμενα της τελικής σ-άλγεβρας μιας **ανεξάρτητης** ακολουθίας ενδεχομένων έχουμε το εξής θεμελιώδες αποτέλεσμα.

Θεώρημα 2.25 (νόμος 0–1 του Kolmogorov για ενδεχόμενα) Εάν η ακολουθία $\{A_n,\ n \geq 1\}$ του χώρου πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$ είναι ανεξάρτητη, τότε για κάθε $A \in \mathcal{T}(\{A_n,\ n \geq 1\})$ ισχύει

$$\mathbb{P}(A) = 0 \quad ή \quad \mathbb{P}(A) = 1.$$

Απόδειξη: Αφού τα ενδεχόμενα

$$\begin{array}{l} A_1 \\ A_2 \\ \vdots \\ A_n \\ A_{n+1}, A_{n+2}, \ldots \end{array}$$

είναι ανεξάρτητα, έπεται από το Πόρισμα 2.23 ότι και οι κλάσεις

$$\sigma(\{A_1\}), \ldots, \sigma(\{A_n\}),\ \ \sigma(\{A_{n+1}, A_{n+2}, \ldots\})$$

είναι ανεξάρτητες. Όμως, για τυχόν $A \in \mathcal{T}\left(\{A_n,\ n \geq 1\}\right)$, έπεται ότι

$$A \in \sigma\left(\{A_{n+1}, A_{n+2}, \ldots\}\right),$$

και συνεπώς, τα ενδεχόμενα

$$A, A_1, \ldots, A_n$$

είναι ανεξάρτητα. Αφού αυτό ισχύει για κάθε n, έπεται ότι η ακολουθία $A, A_1, A_2, \ldots$ είναι ανεξάρτητη, επειδή η ανεξαρτησία ακολουθίας ενδεχομένων είναι ισοδύναμη με την ανεξαρτησία των n πρώτων ενδεχομένων της, για κάθε n – Παρατήρηση 2.11. Συνεπώς, τα ενδεχόμενα

$$\begin{array}{l} A \\ A_1, A_2, \ldots \end{array}$$

είναι ανεξάρτητα, οπότε, λόγω του Πορίσματος 2.23, έπεται ότι οι κλάσεις $\sigma(\{A\})$ και $\sigma(\{A_1, A_2, \ldots\})$ είναι ανεξάρτητες. Εφόσον όμως $A \in \mathcal{T}\left(\{A_n,\ n \geq 1\}\right)$, έπεται ότι $A \in \sigma\left(\{A_1, A_2, \ldots\}\right)$, οπότε τα ενδεχόμενα $A,\ A$ είναι ανεξάρτητα. Αυτό σημαίνει ότι $\mathbb{P}(AA) = \mathbb{P}(A)\mathbb{P}(A)$, δηλ. $\mathbb{P}(A) = \mathbb{P}^2(A)$, που ισοδυναμεί με την αποδεικτέα. $\square$

Παρατήρηση 2.26 Στην πραγματικότητα αποδείξαμε ότι η $\mathcal{T}\left(\{A_n,\ n \geq 1\}\right)$ είναι ανεξάρτητη από τον εαυτό της, αφού η ανεξαρτησία των A και B, για οποιαδήποτε ενδεχόμενα $A \in \mathcal{T}\left(\{A_n,\ n \geq 1\}\right)$

και $B \in \sigma(\{A_1, A_2, \ldots\})$, συνεπάγεται την ανεξαρτησία των $\mathcal{T}(\{A_n,\ n \geq 1\})$ και $\sigma(\{A_1, A_2, \ldots\})$, και άρα την ανεξαρτησία της $\mathcal{T}(\{A_n,\ n \geq 1\})$ από τον εαυτό της, αφού $\mathcal{T}(\{A_n,\ n \geq 1\}) \subset \sigma(\{A_1, A_2, \ldots\})$.

Ο νόμος 0–1 του Kolmogorov είναι πολύ χρήσιμος όταν μελετάμε οριακά ενδεχόμενα. Μας εξασφαλίζει, π.χ., ότι η $\mathbb{P}($*να συμβούν άπειρα το πλήθος* $A_n)$ είναι 0 ή 1, όταν τα $A_n,\ n \geq 1$, είναι ανεξάρτητα. Το ερώτημα είναι πώς θα αποφασίσουμε εάν η πιθανότητα αυτή είναι 0 ή 1. Το πρώτο Λήμμα Borel-Cantelli (Θεώρημα 2.5) μας δίνει απάντηση όταν $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$. Τότε η πιθανότητα αυτή είναι 0, ακόμα και αν τα A_n δεν είναι ανεξάρτητα. Γενικά, όμως, είναι δύσκολο να αποφασίσουμε εάν $\mathbb{P}(A) = 0$ ή $\mathbb{P}(A) = 1$. Προς αυτήν την κατεύθυνση, σχετικό είναι το επόμενο θεώρημα.

Θεώρημα 2.27 (δεύτερο Λήμμα Borel-Cantelli)[8] Εάν η $\{A_n,\ n \geq 1\}$ είναι ανεξάρτητη ακολουθία ενδεχομένων στον χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, και εάν

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = +\infty,$$

τότε $\mathbb{P}(\limsup A_n) = 1$.

---

[8] Βασικά είναι του Cantelli, και αποτελεί μερικό αντίστροφο του πρώτου Λήμματος Borel-Cantelli, Θεώρημα 2.5.

Απόδειξη: Επειδή $\mathbb{P}(\limsup A_n) = 1 - \mathbb{P}(\liminf A_n^c) = 1 - \mathbb{P}(\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^c)$, αρκεί να δείξουμε ότι

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^c\right) = 0.$$

Θέτοντας $B_n = \bigcap_{k=n}^{\infty} A_k^c$, έχουμε ότι η ακολουθία $\{B_n,\ n \geq 1\}$ είναι αύξουσα, και άρα,

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^c\right) = \mathbb{P}\left(\lim_n B_n\right) = \lim_n \mathbb{P}(B_n).$$

Όμως, $B_n = \bigcap_{k=n}^{\infty} A_k^c \subset \bigcap_{k=n}^{N} A_k^c$ για κάθε $N \geq n$. Από την ανεξαρτησία των ενδεχομένων A_n^c, $n \geq 1$, προκύπτει ότι

$$\mathbb{P}(B_n) \leq \mathbb{P}\left(\bigcap_{k=n}^{N} A_k^c\right) = \prod_{k=n}^{N} \mathbb{P}(A_k^c) = \prod_{k=n}^{N} [1 - \mathbb{P}(A_k)] \leq \exp\left[-\sum_{k=n}^{N} \mathbb{P}(A_k)\right],$$

όπου η τελευταία ανισότητα ικανοποιείται επειδή $1 - x \leq e^{-x}$ για κάθε $x \in [0,1]$. Εφόσον η $\sum_{k=1}^{\infty} \mathbb{P}(A_k)$ αποκλίνει στο $+\infty$, έπεται ότι και

$$\lim_{N \to \infty} \sum_{k=n}^{N} \mathbb{P}(A_k) = +\infty, \quad (\text{για κάθε } n \text{ σταθερό}),$$

και άρα,

$$\mathbb{P}(B_n) \leq \exp\left[-\sum_{k=n}^{N} \mathbb{P}(A_k)\right] \to \exp(-\infty) = 0, \ \text{καθώς } N \to \infty.$$

Συνεπώς, $\mathbb{P}(B_n) = 0$ για κάθε n, και άρα

$$\lim_{n \to \infty} \mathbb{P}(B_n) = 0,$$

όπως έπρεπε να δειχθεί. $\square$

Ασκήσεις Κεφ. 2:

2.1. Αποδείξτε ότι μία λ-κλάση είναι πάντα κλάση Dynkin. Εξετάστε εάν ισχύει και το αντίστροφο.

2.2. Έστω $(\Omega, \mathcal{A}, \mathbb{P})$ ένας χώρος πιθανότητας με $\Omega = \mathbb{N} = \{1, 2, \ldots\}$, $\mathcal{A} = \mathcal{P}(\mathbb{N})$, και $\mathbb{P}$ τέτοια ώστε

$$\mathbb{P}(\{n\}) = \frac{1}{n(n+1)}, \quad n = 1, 2, \ldots \,.$$

Δείξτε ότι για την ακολουθία $\{A_n,\ n \geq 1\}$, όπου $A_n = \{n, n+1, \ldots\}$, ισχύει ότι

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = +\infty, \quad \text{και} \quad \mathbb{P}(\limsup A_n) = 0.$$

Υπάρχει αντίφαση με τα Λήμματα Borel-Cantelli;

2.3. Αποδείξτε ότι εάν $\mathbb{P}(A_n) \geq \alpha > 0$ για άπειρες τιμές του n, τότε και

$$\mathbb{P}(\limsup A_n) \geq \alpha.$$

2.4. (i) Εάν $\mathbb{P}(A_n) = 1$ για κάθε n, τότε $\mathbb{P}\Big(\bigcap_{n=1}^{\infty} A_n\Big) = 1$.

(ii) Εάν $\mathbb{P}(A_n) = 0$ για κάθε n, τότε $\mathbb{P}\Big(\bigcup_{n=1}^{\infty} A_n\Big) = 0$.

(iii) Δώστε παράδειγμα χώρου πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, και συνόλου ενδεχομένων $\{A_i,\ i \in I\} \subset \mathcal{A}$, έτσι ώστε $\mathbb{P}(A_i) = 0$ για κάθε $i \in I$, και

$$\mathbb{P}\Big(\bigcup_{i \in I} A_i\Big) = 1.$$

2.5. (i) Έστω $\{a_n,\ n \geq 1\}$ ακολουθία στο $\mathbb{R}$, με $0 \leq a_n \leq 1$ για κάθε n. Τότε:

(α) Εάν $\sum_{n=1}^{\infty} a_n = \infty$ τότε $\prod_{n=1}^{\infty} (1 - a_n) = 0$.

(β) $\sum_{n=1}^{\infty} a_n < \infty$ αν και μόνο αν $\prod_{n=1}^{\infty} (1 + a_n) < \infty$.

(γ) Εξετάστε πότε ισχύει το αντίστροφο του (α).

(ii) Αν τα ενδεχόμενα $A_n,\ n \geq 1$, είναι ανεξάρτητα, δείξτε ότι

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \prod_{n=1}^{\infty} \mathbb{P}(A_n), \text{ και } \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = 1 - \prod_{n=1}^{\infty} [1 - \mathbb{P}(A_n)].$$

(iii) Δείξτε το δεύτερο Λήμμα Borel-Cantelli χρησιμοποιώντας τα (i) και (ii).

2.6. **(Θεώρημα Halmos)** Μία οικογένεια $\mathcal{M}$ υποσυνόλων του Ω λέγεται **μονότονη κλάση** όταν περιέχει τα όρια των μονότονων ακολουθιών της, δηλαδή όταν ισχύουν τα εξής:

(i) Αν $A_n \in \mathcal{M}$ και η $\{A_n,\ n \geq 1\}$ είναι αύξουσα, τότε $\bigcup_{n=1}^{\infty} A_n \in \mathcal{M}$.

(ii) Αν $A_n \in \mathcal{M}$ και η $\{A_n,\ n \geq 1\}$ είναι φθίνουσα, τότε $\bigcap_{n=1}^{\infty} A_n \in \mathcal{M}$.

Αποδείξτε ότι αν μία άλγεβρα $\mathfrak{D}$ περιέχεται στην μονότονη κλάση $\mathcal{M}$, τότε και $\sigma(\mathfrak{D}) \subset \mathcal{M}$.

[**Υπόδειξη:** Θεωρήστε ως $m(\mathfrak{D})$ την ελάχιστη μονότονη κλάση που περιέχει την $\mathfrak{D}$, αφού αποδείξετε ότι υπάρχει τέτοια κλάση φυσικά, και αποδείξτε ότι η κλάση $\mathcal{E} = \{A : A^c \in m(\mathfrak{D})\}$ είναι μονότονη, και περιέχει την $\mathfrak{D}$. Στην συνέχεια θεωρήστε την κλάση $\mathcal{E}_1 = \{A : \text{για κάθε } B \in \mathfrak{D},\ A \cup B \in m(\mathfrak{D})\}$, και δείξτε ότι $\mathfrak{D} \subset \mathcal{E}_1$, και ότι η $\mathcal{E}_1$ είναι μονότονη κλάση. Τέλος, θεωρήστε την κλάση $\mathcal{E}_2 = \{A : \text{για κάθε } B \in m(\mathfrak{D}),\ A \cup B \in m(\mathfrak{D})\}$, και δείξτε ότι $\mathfrak{D} \subset \mathcal{E}_2$, και ότι η $\mathcal{E}_2$ είναι μονότονη κλάση. Το θεώρημα συνάγεται από το γεγονός ότι $m(\mathfrak{D}) \subset \mathcal{E}_2$.]

2.7. Έστω $\mathfrak{D}$ μία, μη κενή, οικογένεια υποσυνόλων του Ω. Δείξτε ότι για κάθε $\mathcal{P} \subset \delta(\mathfrak{D})$, η

$$\overline{\mathcal{P}} = \{A \subset \Omega : \text{για κάθε } C \in \mathcal{P},\ AC \in \delta(\mathfrak{D})\}$$

είναι κλάση Dynkin (πρβλ. απόδειξη Θεωρήματος Dynkin).

2.8. Οι κλάσεις $\mathfrak{D}_1$ και $\mathfrak{D}_2$, με $\varnothing \neq \mathfrak{D}_1 \subset \mathfrak{D}_2 \subset \mathcal{A}$, είναι ανεξάρτητες (στον χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$), αν και μόνο αν για κάθε $A \in \mathfrak{D}_1$, $\mathbb{P}(A) \in \{0, 1\}$.

2.9. Σε ακολουθία ανεξαρτήτων ρίψεων ενός νομίσματος, καθεμία με πιθανότητα επιτυχίας $p \neq \frac{1}{2}$ $(0 < p < 1)$,[9] έστω

$$A_{2n} = \{\text{το νόμισμα έφερε } n \text{ επιτυχίες στις πρώτες } 2n \text{ δοκιμές}\},\quad n = 1, 2, \ldots .$$

[9] δοκιμές Bernoulli(p)

Αποδείξτε ότι, με πιθανότητα 1, μόνο ένα πεπερασμένο πλήθος από τα A_{2n} θα εμφανιστεί.

[**Υπόδειξη**: Αν και υπάρχει απλούστερος τρόπος, μπορείτε να χρησιμοποιήσετε τον τύπο του Stirling,

$$\lim \frac{n!e^n}{n^{n+1/2}} = \sqrt{2\pi}$$

(βλ. π.χ., Χαραλαμπίδης (1996), σελ. 109).]

2.10. Έστω $\{A_n,\ n \geq 1\}$ τυχούσα ακολουθία ενδεχομένων στον χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$. Θέτουμε $A^* = \limsup A_n$ και $A_* = \liminf A_n$. Δείξτε ότι:

(i) $\lim\limits_{n\to\infty} \mathbb{P}\Big(\liminf\limits_{k\to\infty}(A_n A_k^c)\Big) = 0.$

(ii) $\mathbb{P}(A_n \setminus A^*) \to 0$ και $\mathbb{P}(A_* \setminus A_n) \to 0$, καθώς $n \to \infty$.

(iii) Αν $A_n \to A$ (καθώς $n \to \infty$) τότε $\mathbb{P}(A \oplus A_n) \to 0$.
$[A \oplus B \stackrel{\text{ορ.}}{=\!=} (A \setminus B) \cup (B \setminus A)]$