Εισαγωγή στην Θεωρία Μέτρου, την Θεωρία της Ολοκλήρωσης με εφαρμογές στην Θεωρία Πιθανοτήτων

Α. Ν. Γιαννακόπουλος,
Τμήμα Στατιστικής
Οικονομικό Πανεπιστήμιο Αθηνών

8 Οκτωβρίου 2018

Περιεχόμενα

Κεφάλαιο 1

Εισαγωγικές έννοιες

1.1 Εισαγωγή

Στην ενότητα αυτή θα εισάγουμε ορισμένες βασικές έννοιες απο την θεωρία των ακολουθιών πραγματικών αριθμών και τις σειρές οι οποίες αποτελούν μία βασική εισαγωγή στις πιο προχωρημένες έννοιες της πραγματικής ανάλυσης. Επίσης, οι έννοιες αυτές έχουν αρκετό ενδιαφέρον και απο μόνες τους εφόσον βρισκουν σημαντικές εφαρμογές στις πιθανότητες και στην στατιστική, όπως θα δούμε με λεπτομέρεια σε μία σειρά παραδειγμάτων που θα ακολουθήσουν.

1.2 Σύνολα και πράξεις συνόλων.

Ένα **σύνολο** είναι μία συλλογή ομοειδών πραγμάτων ή εννοιών. Κάθε ένα απο αυτά τα πράγματα ή τις έννοιες λέμε οτι ανήκει στον σύνολο αυτό, και το συμβολίζουμε με το $\in$. Αν κάποιο στοιχείο δεν ανήκει σε κάποιο σύνολο τότε αυτό το συμβολίζουμε με το $\notin$. Ένα σύνολο το οποίο δεν περιέχει κανένα στοιχείο ονομάζεται το **κενό σύνολο** και συμβολίζεται $\emptyset$.

Παράδειγμα 1.2.1 *Η συλλογή των αριθμών* $2, 4, 6$ *αποτελεί ένα σύνολο το οποίο θα συμβολίζουμε* $A = \{2, 4, 6\}$*. Μπορούμε συνεπώς να γράψουμε* $2 \in A$, $4 \in A$, $6 \in A$*. Όμως, π.χ.* $8 \notin A$.

Οι βασικές πράξεις μεταξύ των συνόλων είναι η **ένωση** και η **τομή**.
Η ένωση δύο συνόλων A και B μας δίνει ένα καινούργιο σύνολο το οποίο συμβολίζουμε με $A \cup B$ και περιέχει τα στοιχεία του A ή του B. Άρα,

$$c \in A \cup B \iff c \in A \text{ ή } c \in B$$

Η τομή δύο συνόλων A και B μας δίνει ένα καινούργιο σύνολο το οποίο συμβολίοζυμε με $A \cap B$ και περιέχει τα **κοινά στοιχεία** των A και B. Άρα,

$$c \in A \cap B \iff c \in A \text{ και } c \in B$$

Δύο σύνολα A και B ονομάζονται **ξένα** μεταξύ τους αν $A \cap B = \emptyset$.

Παράδειγμα 1.2.2 *Ας θεωρήσουμε τα σύνολα*

$$A = \{2, 4, 6\}$$
$$B = \{1, 3, 5\}$$

Τότε,

$$\begin{aligned} A \cup B &= \{1, 2, 3, 4, 5, 6\} \\ A \cap B &= \emptyset \end{aligned}$$

Η ένωση και η τομή συνόλων μπορεί να οριστεί και για περισσότερα απο δύο σύνολα ακριβώς με τον ίδιο τρόπο όπως για δύο. Συνεπώς αν έχουμε τα σύνολα $A_1, A_2, \cdots, A_n$ μπορούμε να ορίσουμε την ένωση και την τομή

$$A_1 \cup A_2 \cup \cdots \cup A_n,$$
$$A_1 \cap A_2 \cap \cdots \cap A_n$$

Η ένωση και η τομή συνόλων μπορεί να οριστεί και για άπειρα σύνολα,

$$\bigcup_{n=1}^{\infty} A_n := A_1 \cup A_2 \cup \cdots,$$
$$\bigcap_{n=1}^{\infty} A_n := A_1 \cap A_2 \cap \cdots$$

Λέμε ότι ένα σύνολο A είναι **υποσύνολο** ενός συνόλου S, και το συμβολίζουμε $A \subset S$, αν κάθε στοιχείο του A είναι και στοιχείο του S αλλά δεν ισχύει το ανάποδο. Δηλαδή,

$$A \subset S \iff \{a \in A \implies a \in S\}$$

Για ένα σύνολο $A \subset S$ μπορούμε να ορίσουμε το **συμπληρωματικό** του σύνολο, το οποίο συμβολίζουμε A^c, και το οποίο περιέχει όλα τα στοιχεία του S τα οποία **δεν** περιέχονται στο S. Συνεπώς,

$$b \in A^c \iff b \in S \text{ και } b \notin A$$

Κάνοντας χρήση της έννοιας του συμπληρωματικού συνόλου μπορούμε να ορίσουμε την διαφορά δύο συνόλων A και B τα οποία είναι και τα δύο υποσύνολα του ίδιου συνόλου S. Η διαφορά του A απο το B συμβολίζεται με $A \setminus B$ και το σύνολο αυτό περιέχει όλα τα στοιχεία του A που δεν ανήκουν στο B. Συνεπώς

$$c \in A \setminus B \iff c \in A \textbf{ αλλά } c \notin B$$

Είναι προφανές ότι $A \setminus B = A \cap B^c$.

Πολλές φορές είναι χρήσιμο να ορίσουμε και την συμμετρική διαφορά δύο συνόλων A και B,

$$A \triangle B = (A \cap B^c) \cup (A^c \cap B).$$

Η συμμετρική διαφορά περιέχει τα στοιχεία τα οποία ανήκουν σε ένα απο τα A και B αλλά όχι και στα δύο.

Οι ακόλουθοι κανόνες, γνωστοί και ως νόμοι του De Morgan , συνδέουν την ένωση, την τομή και την συμπλήρωση, και είναι πολύ χρήσιμοι,

$$(\cup_{i=1}^{n} A_i)^c = \cap_{i=1}^{n} A_i^c$$
$$(\cap_{i=1}^{n} A_i)^c = \cup_{i=1}^{n} A_i^c$$

Οι νόμοι του De Morgan μπορεί να γενικευθούν και για άπειρες ενώσεις και τομές.

Μία τελευταία ιδιότητα της τομής και της ένωσης είναι η ακόλουθη,

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

Θα κλείσουμε την σύντομη αυτή παρουσίαση των συνόλων, με ορισμένες έννοιες σύγκλισης ακολουθίων συνόλων οι οποίες είναι πολύ χρήσιμες τοσο στην ανάλυση όσο και στην θεωρία πιθανοτήτων.

Ας υποθέσουμε ότι έχουμε άπειρα σύνολα A_i. Μπορούμε να ορίσουμε την τομή ή και την ένωση των συνόλων αυτών, ακριβώς με τον ίδιο τρόπο που ορίζουμε την ένωση και την τομή πεπερασμένου πλήθους συνόλων. Σε πολλές περιπτώσεις έχει νόημα να ρωτήσουμε την ερώτηση,

> Ποιό είναι το σύνολο των κοινών στοιχείων όλων των A_i απο κάποιο i και πάνω·

ή την ερώτηση

> Ποιό είναι το σύνολο των κοινών στοιχείων άπείρων το πλήθος A_i·

Το πρώτο σύνολο το ονομάζουμε το **κάτω όριο** την ακολουθίας συνόλων $\{A_i\}$ ενώ το δεύτερο το ονομάζουμε το **άνω όριο** της ακολουθίας συνόλων $\{A_i\}$.

Ορισμός 1.2.1 (*i*) *Το* ***κάτω όριο*** *της ακολουθίας συνόλων* $\{A_i\}$ *είναι το σύνολο*

$$\lim_{i \to \infty} \inf A_i = \cup_{n=1}^{\infty} \left[\cap_{k=n}^{\infty} A_k\right]$$

(*ii*) *Το* **άνω όριο** *της ακολουθίας συνόλων* $\{A_i\}$ *είναι το σύνολο*

$$\lim_{i\to\infty}\sup A_i = \cap_{n=1}^{\infty}[\cup_{k=n}^{\infty}A_k]$$

Παράδειγμα 1.2.3 *Δίνεται η αύξουσα ακολουθία συνόλων* $E_n \subset E_{n+1}$. *Απο αυτή κατασκευάζουμε την ακολουθία* $B_1 = E_1$, $B_2 = E_2 \setminus E_1$, $\cdots$, $B_n = E_n \setminus E_{n-1}$, $n \in \mathbb{N}$.
Δείξτε ότι $B_i \cap B_j = \emptyset$ *για κάθε* $i, j \in \mathbb{N}$ *και ότι*

$$E_n = \bigcup_{i=1}^{n} B_i,\ n \in \mathbb{N}.$$

Παράδειγμα 1.2.4 *Δίνεται η φθίνουσα ακολουθία συνόλων* $E_{n+1} \subset E_n$. *Δείξτε ότι*

$$\bigcup_{n=1}^{\infty}(E_1 \setminus E_n) = E_1 \setminus \left(\bigcap_{n=1}^{\infty} E_n\right)$$

Ορισμός 1.2.2 *Έστω η ακολουθία συνόλων* A_n. *Λέμε ότι η ακολουθία* A_n *τείνει μονότονα (φθίνει) στο* $\emptyset$ *και συμβολίζουμε* $A_n \downarrow \emptyset$, *αν* $A_{n+1} \subset A_n$ *για κάθε* n *και αν* $\bigcap_{n=1}^{\infty} A_n = \emptyset$.

Ορισμός 1.2.3 *Έστω η ακολουθία συνόλων* $A_n \subset \Omega$. *Λέμε ότι η ακολουθία* A_n *τείνει μονότονα (αυξάνει) στο* Ω *και συμβολίζουμε* $A_n \uparrow \Omega$ *αν* $A_n \subset A_{n+1}$ *για κάθε* n *και* $\bigcup_{n=1}^{\infty} A_n = \Omega$.

1.3 Ορισμένα χρήσιμα σύνολα.

Άλλα σύνολα τα οποία μας ενδιαφέρουν είναι τα ακόλουθα:

(*i*) Το σύνολο των φυσικών αριθμών
$$\mathbb{N} = \{1, 2, \cdots\}$$

(*ii*) Το σύνολο των ακεραίων αριθμών
$$\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$$

(*iii*) Το σύνολο των ρητών αριθμών,
$$\mathbb{Q} = \left\{\frac{a}{b},\ a, b \in \mathbb{Z} \setminus \{0\}\right\}$$

Επίσης θα χρησιμοποιήσουμε τους συμβολισμούς $[a, b]$, (a, b), $(a, b]$, $[a, b)$ για τα ακόλουθα υποσύνολα του $\mathbb{R}$,

$$[a, b] = \{x \in \mathbb{R} \ \mid\ a \le x \le b\},$$

$$(a, b) = \{x \in \mathbb{R} \ \mid\ a \le x < b\},$$

$$(a, b] = \{x \in \mathbb{R} \ \mid\ a < x \le b\},$$

$$[a, b) = \{x \in \mathbb{R} \ \mid\ a \le x < b\}.$$

Παράδειγμα 1.3.1 *Δείξτε ότι*

$$[a, b] = \bigcap_{n=1}^{\infty}\left(a - \frac{1}{n}, b + \frac{1}{n}\right)$$

1.4 Αριθμήσιμα και μη αριθμήσιμα σύνολα

Ορισμός 1.4.1 *Ένα σύνολο ονομάζεται* **αριθμήσιμο** *αν μπορεί να βρεθεί μία 1-1 απεικόνιση μεταξύ των στοιχείων του συνόλου αυτού και του* $\mathbb{N}$.
Ένα σύνολο για το οποία δεν είναι αυτό δυνατό, ονομάζεται **μη αριθμήσιμο**.

Παράδειγμα 1.4.1 *Το σύνολο των αρτίων* $A = \{2, 4, 6, \cdots\}$ *είναι αριθμήσιμο επειδή υπάρχει η* $1-1$ *αντιστοιχία* $m = 2\,n$ *μεταξύ κάθε* $m \in A$ *και κάθε* $n \in \mathbb{N}$.

Παράδειγμα 1.4.2 *Όμοια το σύνολο των περιττών* $B = \{1, 3, 5, \cdots\}$.

Παράδειγμα 1.4.3 *Το σύνολο των ακεραίων* $\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$ *είναι αριθμήσιμο σύνολο*

Παράδειγμα 1.4.4 *Το σύνολο των ρητών* $\mathbb{Q} = \{\frac{n}{m} \mid n, m \in \mathbb{N}\}$ *είναι αριθμήσιμο σύνολο.*

Παράδειγμα 1.4.5 *Το σύνολο των πραγματικών αριθμών* $\mathbb{R}$ *δεν είναι αριθμήσιμο σύνολο. Το αποτέλεσμα αυτο αποδείχθηκε απο τον μεγάλο μαθηματικό και θεμελιωτή της σύγχρονης θεωρίας των συνόλων Georg Cantor .*

Παράδειγμα 1.4.6 *Το σύνολο των αρρήτων αριθμών,* $\mathbb{R} \setminus \mathbb{Q}$ **δεν** *είναι αριθμήσιμο.*

Παράδειγμα 1.4.7 *Τα σύνολα* $[a, b]$, (a, b), $(a, b]$,$[a, b)$, **δεν** *είναι αριθμήσιμα.*

1.5 Η απόλυτη τιμή.

Ορισμός 1.5.1 *Η* **απόλυτη τιμή** *ενός πραγματικού αριθμού* x *συμβολίζεται με* $\mid x \mid$ *και ορίζεται ως*

$$\mid x \mid = \begin{cases} x & \text{αν} \quad x \geq 0 \\ -x & \text{αν} \quad x < 0 \end{cases}$$

Οι ακόλουθες ιδιότητες της απόλυτης τιμής είναι ιδιαίτερα χρήσιμες,

$$\begin{aligned} &\mid x \mid \leq c \Longleftrightarrow -c \leq x \leq c \\ &\mid x\, y \mid = \mid x \mid \mid y \mid \\ &\mid x \mid - \mid y \mid \leq \mid x - y \mid \leq \mid x \mid + \mid y \mid \end{aligned}$$

Αν αναπαραστήσουμε τους πραγματικούς αριθμούς x_1, x_2 με σημεία της πραγματικής ευθείας, τότε η απόλυτη τιμή της διαφοράς $\mid x_1 - x_2 \mid$ παίζει τον ρόλο της **απόστασης** μεταξύ των σημείων αυτών.

1.6 *sup* και *inf*

Θα ξεκινήσουμε με τις έννοιες του άνω και του κάτω φράγματος για υποσύνολα του $\mathbb{R}$.

Ορισμός 1.6.1 *Ένα ύποσύνολο* X *του* $\mathbb{R}$ *λέμε ότι είναι* **άνω φραγμένο** *αν υπάρχει κάποιο* $C \in \mathbb{R}$ *τέτοιο ώστε* $x \leq C$ *για κάθε* $x \in X$.
Ο πραγματικός αριθμός C *ονομάζεται ένα* **άνω φράγμα** *του συνόλου* X.

Ορισμός 1.6.2 *Ένα ύποσύνολο* X *του* $\mathbb{R}$ *λέμε ότι είναι* **κάτω φραγμένο** *αν υπάρχει κάποιο* $c \in \mathbb{R}$ *τέτοιο ώστε* $c \leq x$ *για κάθε* $x \in X$.
Ο πραγματικός αριθμός c *ονομάζεται ένα* **κάτω φράγμα** *του συνόλου* X.

Προφανώς τα άνω και κάτω φράγματα για κάποιο σύνολο **δεν** είναι μοναδικά.

Παράδειγμα 1.6.1 *Έστω* $X = [a, b] \subset \mathbb{R}$, *με* $a < b$.
Οποιοσδήποτε πραγματικός αριθμός $C \geq b$ *είναι ένα άνω φράγμα του* X.
Οποιοσδήποτε πραγματικός αριθμός $c \leq a$ *είναι ένα κάτω φράγμα του* X.

Θα ορίσουμε τώρα τις έννοιες του ελάχιστου άνω φράγματος και του μέγιστου κάτω φράγματος.

Ορισμός 1.6.3 *Έστω* X *ένα υποσύνολο του* $\mathbb{R}$ *το οποίο ειναι μη κενό. Ο πραγματικός αριθμός* M *ότι είναι ένα ελάχιστο άνω φράγμα (sup) για το* X *αν*

(*i*) *για κάθε πραγματικό αριθμό* x *που ανήκει στο* X *ισχύει ότι* $x \leq M$ (*δηλαδή το* M *είναι ένα άνω φράγμα για το* X, *και*

(*ii*) *αν* y *είναι ένας πραγματικός αριθμός τέτοιος ώστε* $x \leq y$ *για κάθε* $x \in X$ (*δηλαδή το* y *είναι ένα άνω φράγμα για το* X) *τότε* $M \leq y$.

Ορισμός 1.6.4 *Έστω X ένα υποσύνολο του $\mathbb{R}$ το οποίο ειναι μη κενό. Ο πραγματικός αριθμός m ότι είναι ένα μέγιστο κάτω φράγμα (inf) για το X αν*

(i) για κάθε πραγματικό αριθμό x που ανήκει στο X ισχύει ότι $m \leq x$ (δηλαδή το m είναι ένα κάτω φράγμα για το X, και

(ii) αν y είναι ένας πραγματικός αριθμός τέτοιος ώστε $y \leq x$ για κάθε $x \in X$ (δηλαδή το y είναι ένα κάτω φράγμα για το X) τότε $y \leq m$.

Από τον παραπάνω ορισμό, μπορούμε να δούμε ότι το M είναι το sup του συνόλου Q αν για κάθε $\epsilon > 0$ υπάρχει κάποιο $x \in X$ τέτοιο ώστε $M - \epsilon < x \leq M$.
Αντίστοιχα, το m είναι το inf του συνόλου X αν για κάθε $\epsilon > 0$ υπάρχει κάποιο $x \in X$ τέτοιο ώστε $m \leq x < m+\epsilon$.

Πρόταση 1.6.1 *Το ελάχιστο άνω φράγμα και το μέγιστο κάτω φράγμα για ένα σύνολο X είναι* **μοναδικά**.

Απόδειξη: Θα το αποδείξουμε για το μέγιστο κάτω φράγμα και αφήνουμε την απόδειξη για το ελάχιστο άνω φράγμα σαν άσκηση.
Έστω m_1, m_2 δύο μέγιστα κάτω φράγματα του X. Θα δείξουμε ότι $m_1 = m_2$.
Απο τον ορισμό, εφοσον m_1 είναι μέγιστο κάτω φράγμα του X θα έχουμε ότι το m_1 είναι κάτω φράγμα για το σύνολο X και οποιοδήποτε άλλο κάτω φράγμα του X θα είναι μικρότερο απο αυτό. Συνεπώς, επειδή το m_2 είναι κάτω φράγμα για το X (επειδή έχουμε δεχθεί ότι είναι μέγιστο κάτω φράγμα) θα πρέπει να ισχύει ότι

$$m_2 \leq m_1.$$

Επίσης απο τον ορισμό, εφόσον το m_2 είναι μέγιστο κάτω φράγμα του X θα έχουμε ότι το m_2 είναι κάτω φράγμα για το σύνολο X και οποιοδήποτε άλλο κάτω φράγμα του X θα είναι μικρότερο απο αυτό. Συνεπώς, επειδή το m_1 είναι κάτω φράγμα για το X (επειδή έχουμε δεχθεί ότι είναι μέγιστο κάτω φράγμα) θα πρέπει να ισχύει ότι

$$m_1 \leq m_2.$$

Οι δύο αυτές ανισότητες μας δείχνουν ότι $m_1 = m_2$, άρα το μέγιστο κάτω φράγμα είναι μοναδικό. □

Παράδειγμα 1.6.2 *Ας υποθέσουμε ότι $X = [a, b]$ για $a, b \in \mathbb{R}$, $a \leq b$. Τότε $\sup X = b$ και $\inf X = a$.*

Το ελάχιστο άνω φράγμα και το μέγιστο κάτω φράγμα ενός συνόλου X, δεν είναι απαραίτητο να ανήκουν στο σύνολο X.

Παράδειγμα 1.6.3 *Ας υποθέσουμε ότι $X = (a, b)$ για $a, b \in \mathbb{R}$, $a \leq b$. Τότε $\sup X = b$ και $\inf X = a$, τα οποία δεν ανήκουν στο X.*

Παράδειγμα 1.6.4 *Έστω*

$$X = \left\{ \frac{1}{n},\ n \in \mathbb{N}^+ \right\}$$

Τότε $\sup\ X = 1$ *και* $\inf\ X = 0$.

Η ακόλουθη ιδιότητα του $\mathbb{R}$ είναι πολύ σημαντική.

Πρόταση 1.6.2 *Ένα άνω φραγμένο και μη κενό υποσύνολο του $\mathbb{R}$ έχει ελάχιστο άνω φράγμα.*

Σχόλιο 1.6.1 *Η ιδιότητα αυτή που ισχύει για τα υποσύνολα του $\mathbb{R}$ δεν είναι αυτονόητη για κάθε σύνολο. Πολλές φορές λοιπόν, όταν εργαζομαστε σε γενικότερα σύνολα, χρειάζεται να θέσουμε την παραπάνω πρόταση σαν αξίωμα.*

1.7 Η αρχή της επαγωγής

Η επαγωγή είναι μία πολύ χρήσιμη αποδεικτική διαδικασία, η οποία μας επιτρέπει να ελέγχουμε την ορθότητα ορισμένων προτάσεων. Χρησιμοποιείται αρκετά σαν εργαλείο στην μαθηματική ανάλυση, οπότε την υπενθυμίζουμε εδώ. Θα παρουσιάσουμε μόνο μία ειδική μορφή της αρχής της επαγωγής, την οποία θα χρησιμοποιήσουμε αρκετά συχνά.

Θεώρημα 1.7.1 *Ας υποθέσουμε ότι έχουμε μία πρόταση* $S(n)$, $n \in \mathbb{N}$.
Αν

(i) *η πρόταση* $S(1)$ *είναι αληθής,*

(ii) *αν η πρόταση* $S(n)$ *είναι αληθής τότε και η πρόταση* $S(n+1)$ *είναι επίσης αληθής,*

τότε η πρόταση $S(n)$ *είναι αληθής για κάθε* $n \in \mathbb{N}$.

Παράδειγμα 1.7.1 *Κάνοντας χρήση της αρχής της επαγωγής δείξτε ότι*

$$1 + 2 + \cdots + n = \frac{n\,(n+1)}{2},$$

για κάθε $n \in \mathbb{N}$.

Η πρόταση $S(n)$ *είναι η προταση που θέλουμε να αποδείξουμε.*
Εύκολα βλέπουμε ότι η $S(1)$ *είναι αληθής, εφόσον* $1 = \frac{1\,(1+1)}{2}$.
Ας υποθέσουμε τώρα ότι η $S(n)$ *είναι αληθής. Τότε θα ισχύει*

$$1 + 2 + \cdots + n = \frac{n\,(n+1)}{2}$$

Ας ελέγξουμε το κατά πόσον η $S(n+1)$ *είναι αληθής.*
Έχουμε ότι

$$1 + 2 + \cdots + n + (n+1) = (1 + 2 + \cdots + n) + (n+1) = \frac{n\,(n+1)}{2} + (n+1) =$$
$$\frac{(n+1)\,(n+2)}{2} = \frac{(n+1)\,((n+1)+1)}{2}.$$

Άρα, η $S(n+1)$ *είναι επίσης αληθής και απο την αρχή της επαγωγής, η* $S(n)$ *είναι αληθής για κάθε* n.

1.8 Βασικές έννοιες απο τις ακολουθίες και τις σειρές

1.8.1 Ακολουθίες

Θα χρειαστούμε πολύ συχνά τα ακόλουθα αποτέλεσματα για τις πραγματικές ακολουθίες, τα οποία παραθέτουμε χωρίς απόδειξη.

Πρόταση 1.8.1 *Μία μονότονη και φραγμένη πραγματική ακολουθία* a_n *είναι συγκλίνουσα.*

Πρόταση 1.8.2 *Μια φραγμένη πραγματική ακολουθία έχει συγκλίνουσα υποακολουθία*

Πρόταση 1.8.3 *Ο* $\mathbb{R}$ *είναι πλήρης μετρικός χώρος, δηλαδή κάθε ακολουθία Cauchy στον* $\mathbb{R}$ *συγκλίνει στο* $\mathbb{R}$.

1.8.2 Σειρές

Θεωρούμε δεδομένα τα πιο βασικά κριτήρια σύγκλισης σειρών τα οποία παραθέτουμε χωρίς απόδειξη.

Πρόταση 1.8.4 *Ας θεωρήσουμε τη σειρά* $\sum_{n=1}^{\infty} a_n$.

(1) Ας θεωρήσουμε το όριο

$$R = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|$$

[(i)] *Αν $R < 1$ τότε η σειρά συγκλίνει απόλυτα.*

[(ii)] *Αν $R > 1$ η σειρά αποκλίνει.*

[(iii)] *Αν $R = 1$ δεν μπορούμε να αποφανθούμε.*

(2) Ας θεωρήσουμε το όριο

$$r = \lim \sup_{n\to\infty} \mid a_n \mid^{\frac{1}{n}}$$

[(i)] *Αν $r < 1$ τότε η σειρά συγκλίνει απόλυτα.*

[(ii)] *Αν $r > 1$ η σειρά αποκλίνει.*

(3) Αν η $\{a_n\}_{n\in\mathbb{N}}$ είναι μία φθίνουσα ακολουθία μη αρνητικών αριθμών, η σειρά $\sum_{n=1}^{\infty} a_n$ συγκλίνει αν και μόνο αν η σειρά $\sum_{n=1}^{\infty} 2^n a_{2^n}$ συγκλίνει.

1.8.3 Διπλές σειρές

Οι διπλές σειρές έχουν αρκετά μεγάλο ενδιαφέρον, τόσο απο θεωρητικής άποψης όσο και απο πλευράς εφαρμογών.

Ορισμός 1.8.1 *Μία ακολουθία με δύο δείκτες είναι μία απεικόνιση απο το $\mathbb{N}\times\mathbb{N} \to \mathbb{R}$. Για κάθε ζεύγος $(n,m)\in\mathbb{N}\times\mathbb{N}$ παίρνουμε ένα πραγματικό αριθμό που τον συμβολίζουμε με $a_{n,m}$.*
Μία διπλή ακολουθία μπορεί να συμβολίζεται σαν $\{a_{m,n}\}$ ή σαν ένας άπειρος πίνακας.

Το ακόλουθο θεώρημα ονομάζεται η διαγώνιος μέθοδος και οφείλεται στον Cantor

Θεώρημα 1.8.1 *Ας θεωρήσουμε μια διπλή σειρά $a_{n,m}$ τους όρους της οποίας διατάσσουμε σαν ένα άπειρο πίνακα*

$$\begin{array}{cccc} a_{1,1} & a_{1,2} & a_{1,3} & \cdots \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots \\ \vdots & \vdots & \vdots & \end{array}$$

Ας υποθέσουμε επίσης ότι κάθε γραμμή του πίνακα αυτού είναι μια φραγμένη ακολουθία πραγματικών αριθμών. Τότε υπάρχει μια υποακολουθία a_{n,m_k} η οποία συγκλίνει για κάθε $n = 1,2,\cdots$ (δηλαδή το $\lim_k a_{n,m_k}$ υπάρχει για κάθε n).

Σχόλιο 1.8.1 *Το παραπάνω δεν είναι μια τόσο απλή εφαρμογή του θεωρήματος Heine-Borel όσο φαίνεται. Σύμφωνα με αυτό για κάθε γραμμή του πίνακα r θα μπορούσαμε να επιλέξουμε μια υποακολουθία $m_k(r)$ τέτοια ώστε να υπάρχει το $\lim_k a_{n,m_k(r)}$. Εν γένει η $m_k(r)$ είναι διαφορετική για διαφορετικά r. Το διαγώνιο επιχείρημα του Cantor μας εξασφαλίζει ότι μπορούμε να επιλέξουμε την* **ίδια** *υποακολουθία m_k για όλες τις γραμμές του πίνακα.*

Απο μία διπλή ακολουθία μπορεί κανείς να ορίσει διάφορες σειρές.
Μία κατηγορία τέτοιων σειρών είναι οι σειρές $S_n = \sum_{m=1}^{\infty} a_{m,n}$, $n = 1,2,\cdots$.
Μία άλλη κατηγορία τέτοιων σειρών είναι οι σειρές $\bar{S}_m = \sum_{n=1}^{\infty} a_{m,n}$, $m = 1,2,\cdots$.
Τέλος μπορούμε να πάρουμε και τα διπλά αθροίσματα $\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{m,n}$ και $\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{m,n}$.
Σε πολλές εφαρμογές είναι ενδιαφέρον να γνωρίζουμε πότε τα αθροίσματα αυτά ορίζονται και πότε μπορούμε να εναλλάξουμε την σειρά των αθροίσεων.

Έχουμε το επόμενο θεώρημα το οποίο μας εξασφαλίζει το πότε τα παραπάνω ισχύουν.

Πρόταση 1.8.5 *Αν υποθέσουμε ότι υπάρχει κάποιο $C < \infty$ τέτοιο ώστε να ισχύει*

$$\sum_{m=1}^{k}\sum_{n=1}^{k} \mid a_{m,n} \mid < C, \ \ \text{για κάθε } k \in \mathbb{N}$$

τότε η διπλή σειρά συγκλίνει. Αυτό σημαίνει ότι

$$\sum_{n=1}^{\infty} S_n = \sum_{n=1}^{\infty}\left(\sum_{m=1}^{\infty} a_{n,m}\right) = \sum_{m=1}^{\infty}\left(\sum_{n=1}^{\infty} a_{n,m}\right) = \sum_{m=1}^{\infty} \bar{S}_m$$

καθώς και ότι οποιαδήποτε **αναδιάταξη** *της διπλής σειράς συγκλίνει.*

Για διπλές σειρές με θετικούς όρους έχουμε το ακόλουθο αποτέλεσμα

Πρόταση 1.8.6 *Αν $a_{nm} \geq 0$ για κάθε n, m τότε*

$$\sum_{n=0}^{\infty}\left(\sum_{m=0}^{\infty} a_{n,m}\right) = \sum_{m=0}^{\infty}\left(\sum_{n=0}^{\infty} a_{n,m}\right)$$

όπου φυσικά αν η διπλή σειρά δεν συγκλίνει τότε και οι δύο όροι της ισότητας αυτής είναι ίσοι με ∞.

Το γινόμενο δύο σειρών μπορεί πολλές φορές να εκφραστεί σαν μια διπλή σειρά ειδικής μορφής.

Πρόταση 1.8.7 *Έστω δύο σειρές $\sum_{n=0}^{\infty} a_n$ και $\sum_{n=0}^{\infty} b_n$ οι οποίες είναι απόλυτα συγκλίνουσες. Τότε η διπλή σειρά $\sum_{n=0}^{\infty}\left(\sum_{j=0}^{n} a_{n-j}\, b_j\right)$, η οποία ονομάζεται άθροισμα του Cauchy , συγκλίνει και έχουμε ότι*

$$\left(\sum_{i=0}^{\infty} a_i\right)\left(\sum_{j=0}^{\infty} b_j\right) = \sum_{n=0}^{\infty}\left(\sum_{j=0}^{n} a_{n-j}\, b_j\right)$$

1.8.4 Εναλλαγή ορίου και αθροίσματος

Πολλές φορές, θα χρειαστεί να εναλλάξουμε ένα όριο με ένα άπειρο άθροισμα. Αυτό πρέπει να γίνεται με πολύ προσοχή, γιατί δεν είναι πάντοτε δυνατό. Στις περιπτώσεις που μας επιτρέπεται να το κάνουμε όμως, είναι ένα πολύ χρήσιμο εργαλείο για την απόδειξη διαφόρων βασικών αποτελεσμάτων π.χ. στις πιθανότητες ή στη στατιστική.

Η παρακάτω πρόταση μας εξασφαλίζει τις προυποθέσεις κάτω απο τις οποίες αυτό είναι εφικτό.

Πρόταση 1.8.8 *Ας υποθέσουμε ότι έχουμε μία διπλή ακολουθία $\{a_{n,m}\}$ όλα τα στοιχεία της οποίας*

$$\{a_{0,m}, a_{1,m}, a_{2,m}, \cdots\}$$

έχουν το ίδιο πρόσημο για κάθε m, και ότι $\mid a_{n+1,m} \mid \geq \mid a_{n,m} \mid$ για κάθε n, m. Αν υπάρχει $C < \infty$ τέτοιο ώστε $\sum_{m=0}^{n} \mid a_{n,m} \mid \leq C$ για κάθε $n \geq 0$ τότε

$$\lim_{n\to\infty} \sum_{m=0}^{\infty} a_{n,m} = \sum_{m=0}^{\infty} \lim_{n\to\infty} a_{n,m}$$

Προτάσεις του τύπου αυτου προκύπτουν πολύ συχνά στις εφαρμογές. Για τον λόγο αυτό θα δώσουμε και μια ακόμη παραλλαγή του αποτελέσματος αυτού, που θα χρειαστούμε στο μέλλον.

Πρόταση 1.8.9 *Έστω $\{a_{n,m}\}$ μία διπλή ακολουθία πραγματικών αριθμών, για την οποια υπάρχει μια ακολουθία πραγματικών αριθμών $\{b_m\}$ έτσι ώστε να ισχύει $\mid a_{n,m} \mid \leq b_m$.*
Ας υποθέσουμε επιπλέον ότι $\sum_{m=0}^{\infty} b_m < \infty$.
Αν $\lim_{n\to\infty} a_{n,m} = a_m$ για κάθε $m \geq 0$ τότε

$$\lim_{n\to\infty} \sum_{m=0}^{\infty} a_{n,m} = \sum_{m=0}^{\infty} \lim_{n\to\infty} a_{n,m} = \sum_{m=0}^{\infty} a_m$$

Τέλος έχει μεγάλο ενδιαφερον και το ακόλουθο λήμμα που ειναι μια ειδική περίπτωση του λήμματος του Fatou

Πρόταση 1.8.10 *Έστω $a_{n,m} \geq 0$. Τότε*

$$\sum_{m=1}^{\infty} \liminf_{n\to\infty} a_{n,m} \leq \liminf_{n\to\infty} \sum_{m=1}^{\infty} a_{n,m}$$

1.9 Ανοιχτά και κλειστά σύνολα του $\mathbb{R}$

Η έννοια του ανοιχτού και κλειστού συνόλου σχετίζεται με το πως ένα σύνολο τοποθετείται μέσα σε ένα μεγαλύτερο σύνολο.

Ορισμός 1.9.1 *Έστω $X \subset \mathbb{R}$.*
Ένα σημείο του $x \in \mathbb{R}$ είναι ένα **οριακό σημείο** *ή* **σημείο συσσώρευσης** *του X αν υπάρχει μία ακολουθία $x^{(n)} \in X$ τέτοια ώστε $x^{(n)} \to x$.*
Ένα υποσύνολο $X \subset \mathbb{R}$ είναι **κλειστό** *αν κάθε οριακό σημείο του X ανήκει στο X.*

Παράδειγμα 1.9.1 *Το $[a, b]$ είναι κλειστο σύνολο στο $\mathbb{R}$. Το $[a, b)$ δεν είναι κλειστό.*

Η τομή και η ένωση κλειστών συνόλων παραμένει κλειστό σύνολο. Το παρακάτω παρατίθεται χωρίς απόδειξη.

Πρόταση 1.9.1 *Έστω $X_i \subset \mathbb{R}$, $i = 1, \cdots, m$, υποσύνολα του $\mathbb{R}$ τα οποία είναι κλειστά.*
Τότε,

1. *$\cup_{i=1}^{m} X_i$ είναι κλειστό (m πεπερασμένο)*
2. *$\cap_{i=1}^{m} X_i$ είναι κλειστό (το $m = \infty$ επιτρέπεται)*

Το αποτέλεσμα της πρότασης 1.9.1(1) δεν μπορεί να γενικευθεί για $m = \infty$ όπως δείχνει το παρακάτω παράδειγμα.

Παράδειγμα 1.9.2 *Παρατηρείστε ότι μπορούμε να γράψουμε*

$$(0, 1) = \bigcup_{n=2}^{\infty} X_n = \bigcup_{n=2}^{\infty} \left[\frac{1}{n}, \frac{n}{n+1}\right]$$

Παρότι τα $X_n = \left[\frac{1}{n}, \frac{n}{n+1}\right]$ είναι κλειστά υποσύνολα του $\mathbb{R}$ η άπειρη ένωση τους είναι $(0, 1)$ το οποίο δεν είναι κλειστό υποσύνολο του $\mathbb{R}$.

Ορισμός 1.9.2 *Έστω $X \subset \mathbb{R}$. Θα λέμε ότι το X είναι ανοιχτό στο $\mathbb{R}$ αν για κάθε $x \in X$ υπάρχει μία ανοιχτή μπάλα $B_\epsilon(x) := \{y \in \mathbb{R} \mid\, \mid y - x \mid < \epsilon\}$ τέτοια ώστε $B_\epsilon(x) \subset X$.*

Με άλλα λόγια μπορεί κανείς να πει ότι ένα υποσύνολο X του $\mathbb{R}$ είναι ανοιχτό, αν για κάθε σημείο x του X μπορεί κανείς να βρει μία ανοιχτή μπάλα (ακτίνας της επιλογής μας αλλά αυστηρά μεγαλύτερης του 0) με κέντρο το x η οποία να περιέχεται εξ ολοκλήρου στο X.

Παράδειγμα 1.9.3 *Έστω $X = (a, b)$. Το X είναι ανοιχτό στο $\mathbb{R}$. Πράγματι για κάθε $x \in (a, b)$ μπορούμε να κατασκευάσουμε μία ανοιχτή μπάλα με κέντρο το x και ακτίνα $\epsilon = min(b - x, x - a)$ η οποία περιέχεται εξ ολοκλήρου στο X.*

Πρόταση 1.9.2 *Έστω $X_i \subset \mathbb{R}$, $i = 1, \cdots, m$, υποσύνολα του $\mathbb{R}$ τα οποία είναι ανοιχτά.*
Τότε,

1. *$\cup_{i=1}^{m} X_i$ είναι ανοιχτό στο M (το $m = \infty$ επιτρέπεται(*
2. *$\cap_{i=1}^{m} X_i$ είναι ανοιχτό στο M (m πεπερασμένο).*

Το αποτέλεσμα της πρότασης 1.9.2(2) δεν μπορεί να γενικευθεί για $m = \infty$ όπως δείχνει το παρακάτω παράδειγμα.

Παράδειγμα 1.9.4 *Παρατηρείστε ότι μπορούμε να γράψουμε*

$$\{0\} = \bigcap_{n=1}^{\infty} X_n = \bigcap_{n=1}^{\infty} \left(-\frac{1}{n}, \frac{1}{n}\right)$$

Παρότι τα $X_n = \left(-\frac{1}{n}, \frac{1}{n}\right)$ είναι ανοιχτά υποσύνολα του $\mathbb{R}$ η άπειρη τομή τους είναι το $\{0, \}$ το οποίο δεν είναι ανοιχτό υποσύνολο του $\mathbb{R}$.

Πρόταση 1.9.3 *Το X είναι ανοιχτό αν και μόνο αν το X^c είναι κλειστό.*

Το ακόλουθο αποτέλεσμα είναι θεμελιώδες για την ανάλυση και την θεωρία μέτρου.

Πρόταση 1.9.4 *(Heine-Borel) Οποιοδήποτε κλειστό και φραγμένο διαστήμα στο $\mathbb{R}$ μπορεί να καλυφθεί απο πεπερασμένο αριθμό ανοιχτών διαστημάτων, δηλαδή μπορεί να γραφεί σαν την ένωση πεπερασμένων ανοιχτών διαστημάτων.*

Το θεώρημα Heine-Borel εξασφαλίζει την **συμπάγεια** των κλειστών και φραγμένων διαστημάτων του $\mathbb{R}$ και ισοδύναμα μπορεί να γραφεί με την μορφή: αν $[a,b] \subseteq \bigcup_{k=1}^{\infty}(a_k,b_k)$ τότε υπάρχει n τέτοιο ώστε $[a,b] \subseteq \bigcup_{k=1}^{n}(a_k,b_k)$. Το ότι οποιαδήποτε φραγμένη ακολουθία στον $\mathbb{R}$ έχει συγκλίνουσα υποακολουθία είναι μια ισοδύναμη μορφή του θεωρήματος αυτού.

Παράδειγμα 1.9.5 *Πολλά υποσύνολα του $\mathbb{R}$ μπορεί να γραφούν σαν άπειρη ένωση ανοιχτών διαστημάτων*

$$
\begin{aligned}
[a,b] &= \bigcap_{n=1}^{\infty}(a-\frac{1}{n}, b+\frac{1}{n}) \\
(a,b] &= \bigcap_{n=1}^{\infty}(a, b+\frac{1}{n})
\end{aligned}
$$

Οι συνεχείς συναρτήσεις έχουν χρήσιμες ιδιότητες ως προς τα κλειστά και τα ανοίχτά σύνολα.

Πρόταση 1.9.5 *Έστω $X \subset \mathbb{R}$ και $f : X \to \mathbb{R}$ μία συνάρτηση. Η f είναι συνεχής στο X αν και μόνο αν η f^{-1} απεικονίζει κλειστά υποσύνολα του $\mathbb{R}$ σε κλειστά υποσύνολα του X.*

Πρόταση 1.9.6 *Έστω $X \subset \mathbb{R}$ και $f : X \to \mathbb{R}$ μία συνάρτηση. Η f είναι συνεχής στο X αν και μόνο αν η f^{-1} απεικονίζει ανοιχτά υποσύνολα του $\mathbb{R}$ σε ανοιχτά υποσύνολα του X.*

Κεφάλαιο 2

Διακριτά μέτρα

2.1 Εισαγωγή

Στο κεφάλαιο αυτό θα ξεκινήσουμε την μελέτη της θεωρίας μέτρου με την πιο απλή περίπτωση μέτρων, τα διακριτά μέτρα. Αυτά είναι μέτρα τα οποία ορίζονται επάνω σε αριθμήσιμα σύνολα, και αυτό τα κάνει όπως θα δούμε πιο εύκολα στον χειρισμό. Σχετίζονται άμεσα με τις διακριτές τυχαίες μεταβλητές και μας επιτρέπουν να πάρουμε μια καλή ιδέα για αρκετά απο τα θέματα που θα μας απασχολήσουν χωρίς να χρειαστεί να μπούμε σε δύσκολες τεχνικές λεπτομέρειες.

2.2 Διακριτά μέτρα σε πεπερασμένα σύνολα

2.2.1 Το μέτρο αρίθμησης

Θα ξεκινήσουμε με την απλούστερη μορφή μέτρου, το μέτρο αρίθμησης.

Ας υποθέσουμε ότι έχουμε ένα πεπερασμένο σύνολο το οποίο αποτελείται απο διακριτά στοιχεία, π.χ. $\Omega = \{\omega_1, \omega_2, \cdots, \omega_N\}$. Ο πληθάριθμος του συνόλου αυτού δηλαδή ο αριθμός των στοιχείων του συνόλου αυτού είναι $card(\Omega) = N$. Για κάθε υποσύνολο $A \subset \Omega$ μπορούμε να ορίσουμε μία απεικόνιση μ απο το σύνολο όλων των υποσυνόλων του Ω στο $[0,1]$ ως εξής

$$\mu(A) = \frac{card(A)}{card(\Omega)}$$

Η απεικόνιση αυτή είναι ένα μέτρο το οποίο ονομάζεται μέτρο αρίθμησης.

Ας δούμε τις ιδιότητες της απεικόνισης αυτής.

Πρόταση 2.2.1 *Τό μέτρο αρίθμησης έχει τις ακόλουθες ιδιότητες:*

1. $\mu(A) \in [0,1]$ *για κάθε* $A \subset \Omega$
2. $\mu(\emptyset) = 0$
3. *Για κάθε συλλογή* $A_i \subset \Omega$ *τέτοια ώστε* $A_i \cap A_j = \emptyset$ *ισχύει ότι*

$$\mu\left(\bigcup_n A_i\right) = \sum_i \mu(A_i)$$

Η απόδειξη των ιδιοτήτων αυτών για το μέτρο αρίθμησης αφήνεται σαν άσκηση.

Το μέτρο αρίθμησης μπορεί να ερμηνευθεί σαν πιθανότητα, δηλαδή σαν ένας τρόπος να ποσοτικοποιήσουμε πόσο 'εύκολο' ή 'δύσκολο' είναι να συμβεί κάποιο γεγονός.

Παράδειγμα 2.2.1 *Ας πάρουμε το πείραμα που αποτελείται απο την ρίψη ενός ζαριού. Ο δειγματικός χώρος είναι το πεπερασμένο και διακριτό σύνολο*

$$\Omega = \{1,2,3,4,5,6\}$$

Το μέτρο αρίθμησης στο σύνολο αυτό μπορεί να ερμηνευθεί σαν η πιθανότητα να έρθουν τα διαφορετικά αποτελέσματα του ζαριού. Για παράδειγμα

$$\mu(\{2,4,6\}) = \frac{3}{6} = \frac{1}{2} = P(\text{ το αποτέλεσμα είναι ζυγός })$$

2.2.2 Γενικά διακριτά μέτρα

Ορισμός 2.2.1 *Μία οποιαδήποτε απεικόνιση απο ένα σύνολο υποσυνόλων του Ω στο $[0,\infty)$ που ικανοποιεί τις παραπάνω ιδιότητες ονομάζεται* **μέτρο**.

Μία ενδιαφέρουσα ερώτηση είναι:

> Αν θεωρήσουμε την μ σαν μία απεικόνιση απο το σύνολο των υποσυνόλων του Ω στο $[0,1]$ ποιό θα είναι το πεδίο ορισμού της;

Ορισμός 2.2.2 *Για τα πεπερασμένα διακριτά σύνολα Ω, μας αρκεί να επιλέξουμε σαν πεδίο ορισμού ένα σύνολο υποσυνόλων του Ω το οποίο είναι κλειστό ως προς την πράξη της συμπλήρωσης και την πράξη της ένωσης, δηλαδή ένα σύνολο υποσυνόλων $\mathcal{F}$ το οποίο να έχει τις ιδιότητες*

1. *Αν $A \in \mathcal{F}$ τότε και $A^c \in \mathcal{F}$*
2. *Αν $A_i \in \mathcal{F}$ τότε και $\bigcup_n A_i \in \mathcal{F}$, όπου τα A_i είναι πεπερασμένα το πλήθος.*
3. *$\Omega \in \mathcal{F}$.*

Το μέτρο αρίθμησης δεν είναι ο μόνος τρόπος να ορίσουμε ένα μέτρο επάνω σε ένα πεπερασμένο αριθμήσιμο σύνολο.

Ας θεωρήσουμε το σύνολο $\Omega = \{\omega_1, \cdots, \omega_n\}$ και το διάνυσμα $p = (p_1, \cdots, p_n)$, τέτοιο ώστε $p_i \in [0,1]$ και $\sum_{i=1}^n p_i = 1$.

Για κάθε υποσύνολο $A \subset \Omega$ ας ορίσουμε την δείκτρια συνάρτηση του, $\mathbf{1}_A$ η οποία ικανοποιεί το

$$\mathbf{1}_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A. \end{cases}$$

Αν ορίσουμε την απεικόνιση $\mu : 2^\Omega \to [0,1]$ συμφωνα με τον κανόνα

$$\mu(A) = \sum_{i=1}^n p_i\, \mathbf{1}_A(\omega_i), \;\; A \subset \Omega$$

τότε η μ ικανοποιεί τις ιδιότητες της Πρότασης 2.2.1 άρα είναι μέτρο.

Παράδειγμα 2.2.2 *Ας πάρουμε π.χ. $A_1 = \{\omega_1, \omega_2, \omega_3\}$. Σύμφωνα με τον παραπάνω ορισμό $\mu(A_1) = p_1 + p_2 + p_3$.*

Και το παραπάνω μέτρο μπορεί να έχει εφαρμογές στις πιθανότητες.

Παράδειγμα 2.2.3 *Ας πάρουμε και πάλι $\Omega = \{1,2,3,4,5,6\}$ και $p_1 = p_3 = p_5 = \frac{1}{12}$, $p_2 = p_4 = p_6 = \frac{3}{12}$. Μπορούμε με τον τρόπο αυτό να ορίσουμε ένα μέτρο επάνω στο Ω το οποίο αντιστοιχεί στην μοντελοποίηση της ρίψης ενος ζαριού, το οποίο μπορεί να φέρει τους αριθμούς $1,3,5$ με πιθανότητα $\frac{1}{12}$ αντιστοίχως και τους αριθμούς $2,4,6$ με πιθανότητα $\frac{3}{12}$ αντιστοίχως.*
Αν $A = \{1,3,5\}$ τότε $\mu(A) = p_1 + p_3 + p_5 = \frac{1}{4}$, και αυτό μπορεί να ερμηνευθεί σαν η πιθανότητα το ζάρι αυτό να φέρει περιττό αποτέλεσμα.
Αν $B = \{2,4,6\}$ τότε $\mu(B) = p_2 + p_4 + p_6 = \frac{3}{4}$, και αυτό μπορεί να ερμηνευθεί σαν η πιθανότητα το ζάρι αυτό να φέρει άρτιο αριθμό σαν αποτέλεσμα.

Οι ιδιότητες τις οποίες ζητάμε να ικανοποιεί το σύνολο υποσυνόλων $\mathcal{F}$ που είναι το πεδίο ορισμού του μ είναι ενδιαφέρουσες απο την πλευρά της πιθανής τους ερμηνείας στην θεωρία πιθανοτήτων. Το $\mathcal{F}$ μπορεί να θεωρήσουμε ότι περιέχει όλες τις πιθανές ερωτήσεις που μπορεί να κάνει κάποιος για κάποιο πείραμα. Για παράδειγμα, η ιδιότητα 1 του ορισμού 2.2.2 μας εξασφαλίζει ότι αν μπορούμε να αναρωτηθούμε για την 'ευκολία' με την οποία μπορεί να συμβεί το γεγονός A, θα μπορούμε να αναρωτηθούμε και για την 'ευκολία' του να μη συμβεί το γεγονός A, δηλαδή το A^c. Η ιδιότητα 2 μας εξασφαλίζει το ότι αν μπορούμε να περιγράψουμε την 'ευκολία' του να συμβούν (χωριστά) τα γεγονότα, $A_1, A_2, \cdots$ κλπ. μπορούμε να περιγράψουμε την 'ευκολία' του να συμβεί το σύνθετο γεγονός A_1ή$A_2 \cdots$.

Πρόταση 2.2.2 *Για το γενικό αυτό μέτρο ισχύουν τα παρακάτω:*

1. *Για οποιαδήποτε $A, B \subset \Omega$ έχουμε ότι*
$$\mu(A \setminus B) = \mu(A) - \mu(A \cap B)$$

2. *Για οποιαδήποτε $A, B \subset \Omega$ έχουμε ότι*
$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$$

3. *Αν $A \subset B$ τότε $\mu(A) \leq \mu(B)$.*

2.3 Διακριτά μέτρα σε άπειρα αλλά αριθμήσιμα σύνολα.

Ας υποθέσουμε ότι το σύνολο Ω είναι άπειρο αλλά αριθμήσιμο, π.χ.
$$\Omega = \{\omega_1, \omega_2, \cdots\}.$$
Εδώ π.χ. η έννοια του μέτρου αρίθμησης, δεν έχει νόημα αλλά μπορούμε σε κάθε περίπτωση να ζητήσουμε την ύπαρξη απεικονίσεων απο τα υποσύνολα του Ω στο $[0,1]$ οι οποίες έχουν τις παρόμοιες ιδιότητες με το μέτρο αρίθμησης, δηλαδή τις ιδιότητες που περιγράφουμε στην Προταση 2.2.1.

Θα ξεκινήσουμε με τον ορισμό της σ-άλγεβρας, η οποία είναι και το πεδίο ορισμού του μέτρου.

Ορισμός 2.3.1 *Έστω $\mathcal{F}$ μία συλλογή υποσυνόλων του Ω που ικανοποιεί τις ιδιότητες*

1. *$\emptyset \in \mathcal{F}$ και $\Omega \in \mathcal{F}$,*
2. *Αν $A \in \mathcal{F}$ τότε και $A^c \in \mathcal{F}$,*
3. *Αν $A_i \in \mathcal{F}$, $i \in \mathbb{N}$ τότε και $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.*

*Η $\mathcal{F}$ ονομάζεται $\sigma-$**άλγεβρα**.*

Η σ-άλγεβρα μπορεί να θεωρηθεί ότι περιέχει σύνθετα γεγονότα σχετικά με κάποιο πείραμα. Επίσης μπορεί να θεωρηθεί ότι περιέχει όλες τις πιθανές ερωτήσεις που μπορεί καποιος να θελήσει να ρωτήσει σχετικά με κάποιο πείραμα.

Παράδειγμα 2.3.1 *Η σ-άλγεβρα περιέχει γεγονότα της μορφής*
$$\limsup_n A_n = \{A_n \ i.o\} = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$
Τα γεγονότα της μορφής αυτής περιέχουν αυτά τα $\omega \in \Omega$ τα οποία εμφανίζονται σε άπειρα το πλήθος A_n, γι΄ αυτό και αναφέρονται πολλές φορές στην θεωρία πιθανοτήτων σαν A_n άπειρα συχνά (infinitely often i.o.).

Η σ-άλγεβρα περιέχει επίσης γεγονότα της μορφής
$$\liminf_n A_n = \{A_n \ a.a\} = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$$
Τα γεγονότα της μορφής αυτής περιέχουν αυτά τα $\omega \in \Omega$ τα οποία εμφανίζονται σε όλα εκτός από πεπερασμένα το πλήθος A_n, γι΄ αυτό και αναφέρονται πολλές φορές στην θεωρία πιθανοτήτων σαν A_n σχεδόν πάντοτε (almost always a.a.).

Γεγονότα της μορφής αυτής είναι απαραίτητα για να απαντήσουμε ερωτήματα σχετικά με την ασυμπτωτική συμπεριφορά ενός πειράματος.

Παράδειγμα 2.3.2 *Για κάποιο σύνολο Ω μπορούμε να ορίσουμε πολλές διαφορετικές $\sigma-$άλγεβρες. Π.χ. οι*
$$\begin{aligned} \mathcal{F}_0 &= \{\emptyset, \Omega\} \\ \mathcal{F}_1 &= \{\emptyset, \Omega, \{\omega_1\}, \{\omega_1\}^c\} \\ \mathcal{F}_2 &= \{\emptyset, \Omega, E, E^c\} \end{aligned}$$
όπου E οποιοδήποτε υποσύνολο του Ω, είναι παραδείγματα τέτοιων $\sigma-$αλγεβρών.

Ορισμός 2.3.2 *Έστω Ω ένα σύνολο και $\mathcal{F}$ μία $\sigma-$άλγεβρα επάνω σε αυτό. Το ζεύγος $(\Omega, \mathcal{F})$ ονομάζεται* **μετρήσιμος χώρος**.

Παράδειγμα 2.3.3 *Στα πλαίσια του Παραδείγματος 2.3.2 τα ζεύγη $(\Omega, \mathcal{F}_0)$, $(\Omega, \mathcal{F}_1)$ και $(\Omega, \mathcal{F}_2)$ είναι μετρήσιμοι χώροι.*

Ορισμός 2.3.3 *Η απεικόνιση $\mu : \mathcal{F} \to [0,1]$ ονομάζεται* **μέτρο πιθανότητας** *αν ικανοποιεί τις παρακάτω ιδιότητες*

1. $\mu(\emptyset) = 0$, *και* $\mu(\Omega) = 1$.
2. *Για κάθε ακολουθία $A_i \in \mathcal{F}$, $i \in \mathbb{N}$, τα μέλη της οποίας είναι ανα δύο ξένα, δηλαδή $A_i \cap A_j = \emptyset$ για κάθε $i, j \in \mathbb{N}$, ισχύει*
$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

Σχόλιο 2.3.1 *Αν θεωρήσουμε ότι η μ παίρνει τιμές στο $[0, \infty)$ και αφαιρέσουμε την συνθήκη $\mu(\Omega) = 1$ τότε η απεικόνιση μ ονομάζεται απλά* **μέτρο** *και όχι μέτρο πιθανότητας. Αν $\mu(\Omega) < \infty$ ονομάζεται* **πεπερασμένο μέτρο**.

Ορισμός 2.3.4 *Έστω Ω ένα σύνολο, $\mathcal{F}$ μία $\sigma-$άλγεβρα και μ ένα μέτρο. Η τριάδα $(\Omega, \mathcal{F}, \mu)$ ονομάζεται* **χώρος μέτρου**.

Ας δούμε τώρα μία μεθοδολογία κατασκευής ενός μέτρου σε ένα άπειρο αλλά αριθμήσιμο σύνολο.

Πρόταση 2.3.1 *Ας πάρουμε το*
$$\Omega = \{\omega_1, \omega_2, \cdots\}$$
και μία ακολουθία $p_1, p_2, \cdots \in [0,1]$ τέτοια ώστε $\sum_{i=1}^{\infty} p_i = 1$.
Για κάθε $A \subset \Omega$ μπορούμε να ορίσουμε την απεικόνιση
$$\mu(A) = \sum_{i=1}^{\infty} p_i \, \mathbf{1}_A(\omega_i)$$
Η απεικόνιση αυτή ορίζει ένα μέτρο.

Για να το δούμε αυτό, αρκεί να ελέγξουμε ότι ικανοποιούνται οι ιδιότητες του ορισμού. Οι πρώτες δύο ιδιότητες είναι προφανείς, οπότε αρκεί να ελέγξουμε την αθροιστικότητα.
Ας πάρουμε μία ακολουθία υπόσυνόλων $A_i \subset \Omega$ τα οποία είναι ξένα μεταξύ τους, δηλαδή $A_i \cap A_j = \emptyset$. Παρατηρούμε πρώτα ότι αν
$$E = \bigcup_{i=1}^{\infty} A_i,$$
τότε
$$\mathbf{1}_E(\omega_j) = \sum_{i=1}^{\infty} \mathbf{1}_{A_i}(\omega_j) \tag{2.1}$$
για οποιοδήποτε j.
Πραγματικά, ας πάρουμε οποιοδήποτε $\omega_j \in \Omega$. Έχουμε ότι $\mathbf{1}_E(\omega_j) = 1$ αν $\omega_j \in E = \bigcup_{i=1}^{\infty} A_i$, και επειδή τα A_i είναι ξένα μεταξύ του αυτό σημαίνει ότι υπάρχει κάποιο (μοναδικό) k τέτοιο ώστε να ισχύει $\omega_j \in A_k$. Συνεπώς, $\mathbf{1}_{A_k}(\omega_j) = 1$, ενώ, $\mathbf{1}_{A_i}(\omega_j) = 0$, για κάθε $i \neq k$. Αυτό όμως σημαίνει ότι
$$\sum_{i=1}^{\infty} \mathbf{1}_{A_i}(\omega_j) = 1$$
οπότε και ισχύει η ισότητα (2.1). Αν τώρα $\mathbf{1}_E(\omega_j) = 0$ τότε $\omega_j \notin E = \bigcup_{i=1}^{\infty} A_i$ άρα $\omega_j \notin A_i$ για όλα τα i, οπότε και $\mathbf{1}_{A_i}(\omega_j) = 0$, για κάθε i. Συνεπώς,
$$\sum_{i=1}^{\infty} \mathbf{1}_{A_i}(\omega_j) = 0$$

οπότε και ισχύει η ισότητα (2.1).
Έχοντας τώρα υπ'οψιν την (2.1) μπορούμε να γράψουμε

$$\begin{aligned}
\mu\left(\bigcup_{i=1}^{\infty} A_i\right) &= \sum_{j=1}^{\infty} p_j \, \mathbf{1}_{\bigcup_{i=1}^{\infty} A_i}(\omega_j) \\
&= \sum_{j=1}^{\infty} p_i \left(\sum_{i=1}^{\infty} \mathbf{1}_{A_i}(\omega_j)\right) \\
&= \sum_{i=1}^{\infty} \left(\sum_{j=1}^{\infty} p_j \mathbf{1}_{A_i}(\omega_j)\right) \\
&= \sum_{i=1}^{\infty} \mu(A_i)
\end{aligned}$$

οπότε και αποδείχθηκε η αθροιστικότητα. Για να καταλήξουμε στο αποτέλεσμα αυτό χρειάστηκε να χρησιμοποιήσουμε ιδιότητες σύγκλισης των διπλών αθροισμάτων.

Η παραπάνω κατασκευή είναι ιδιαίτερα χρήσιμη στις διακριτές κατανομές.

Παράδειγμα 2.3.4 *Ας πάρουμε* $\Omega = \mathbb{N}$ *δηλαδή* $\omega_i = i$, $i = 0, 1, \cdots$, *και* $p_i = e^{-\lambda}\frac{\lambda^i}{i!}$. *Έτσι π.χ. αν* $A_i = \{i\}$ *ο ορισμός του διακριτού μέτρου δίνει* $\mu(A_i) = p_i$ *και βλέπουμε ότι το μέτρο αυτό αντιστοιχεί στην κατανομή Poisson.*

Παράδειγμα 2.3.5 *Ας πάρουμε* $\Omega = \mathbb{N}^+$ *δηλαδή* $\omega_i = i$, $i = 1, 2, \cdots$, *και* $p_i = (1-p)^{i-1}p$. *Έτσι π.χ. αν* $A_i = \{i\}$ *ο ορισμός του διακριτού μέτρου δίνει* $\mu(A_i) = p_i$ *και βλέπουμε ότι το μέτρο αυτό αντιστοιχεί στην γεωμετρική κατανομή.*

Παράδειγμα 2.3.6 *Ας θεωρήσουμε το παράδειγμα των επαναλαμβανόμενων ρίψεων νομισμάτων. Αν συμβολίσουμε με 1 την κορώνα και με 0 τα γράμματα, το* Ω *θα είναι το σύνολο των ακολουθιών που αποτελούνται απο 0 και 1 δηλαδή*

$$\Omega = \{(r_1 r_2 r_3 \cdots) \;\mid\; r_i = 0 \text{ ή } 1\}$$

Τα γεγονότα που μας ενδιαφέρουν είναι γεγονότα της μορφής 'τα m *πρώτα νομίσματα έφεραν ένα συγκεκριμένο αποτέλεσμα* $a_1 a_2 \cdots a_m$*'. Τα γεγονότα αυτά μπορεί να περιγραφούν ως*

$$A_{a_1 a_2 \cdots a_m} = \{a_1 a_2 \cdots a_m r_{m+1} r_{m+2} \cdots, \;\mid\; r_i = 0 \text{ ή } 1,\; i = m+1, m+2, \cdots\}$$

Για την κατασκευή του μέτρου που θα περιγράφει το πείραμα των άπειρα επαναλαμβανόμενων ρίψεων νομισμάτων μπορούμε να γράψουμε

$$\mu(A_{a_1 a_2 \cdots a_m}) = \frac{1}{2^m}$$

Η παραπάνω σχέση μας ορίζει την τιμή που θα πάρει το μέτρο αυτό σε ένα συγκεκριμένο γεγονός, το γεγονός να φέρουν τα πρώτα m *νομίσματα το συγκεκριμένο αποτέλεσμα* $a_1 a_2 \cdots a_m$. *Επειδή μπορεί να θέλουμε να ασχοληθούμε και με πιο γενικά ερωτήματα, όπως π.χ. την πιθανότητα να έχουμε σύνθετα γεγονότα θα πρέπει να ορίσουμε το πεδίο ορισμού του* μ *δηλαδή μια σ-άλγεβρα η οποία να περιέχει όλα τα πιθανά υποσύνολα που περιέχουν στοιχεία του τύπου* $A_{a_1 a_2 \cdots a_m}$, *για όλα τα* $m \in \mathbb{N}$, *τις ενώσεις τους, τις τομές τους κλπ.*

Η κατασκευή μια τέτοιας σ-άλγεβρας δεν είναι εύκολη υπόθεση. Για την κατασκευή του μέτρου μ *λοιπόν χρειάζεται να ορίσουμε την απεικόνιση με τις ιδιότητες τις οποίες θέλουμε σε απλούστερα σύνολα γεγονότων και μετά χρησιμοποιώντας ένα θεωρητικό αποτέλεσμα να το επεκτείνουμε στην σ-άλγεβρα.Παρόμοια φαινόμενα συμβαίνουν και σε άλλα ενδιαφέροντα παραδείγματα απο την θεωρία πιθανοτήτων και αυτό καταδεικνύει την ανάγκη να μελετήσουμε βαθύτερα και σε πιο μεγάλη λεπτομέρεια την κατασκευή του μέτρου.*

2.4 Τυχαίες μεταβλητές

Ας πάρουμε μία απεικόνιση $X : \Omega \to \mathbb{R}$. Μια τέτοια απεικόνιση ονομάζεται **τυχαία μεταβλητή**.

Παράδειγμα 2.4.1 *Ας πάρουμε $\Omega = \{H_1H_2, H_1T_2, T_1H_2, T_1T_2\}$ και το μέτρο αρίθμησης μ. Αυτο το πλαίσιο μπορεί να είναι ένα μοντέλο για το πείραμα της ρίψης δύο δίκαιων νομισμάτων.*
Αν κάποιος ποντάρει στις κορώνες, και κερδίζει 1 ευρώ αν έρθει κορώνα και χάνει -1 ευρώ αν έρθει γράμματα, το συνολικό του κέρδος X είναι η τυχαία μεταβλητή για την οποία

$$X(H_1H_2) = 2,\ X(H_1T_2) = 0,\ X(T_1H_2) = 0,\ X(T_1T_2) = -2$$

Μπορούμε να ορίσουμε την πιθανότητα P η τυχαία αυτή μεταβλητή να παίρνει διάφορες τιμές

$$P(X=2) = \mu(\{H_1H_2\}) = \frac{1}{4},\ P(X=0) = \mu(\{H_1T_2, T_1H_2\}) = \frac{1}{2},$$
$$P(X=-2) = \mu(\{T_1T_2\}) = \frac{1}{4},\ P(X \neq 2, 0, -2) = P(X \in \mathbb{R} \setminus \{-2, 0, 2\}) = \mu(\emptyset) = 0$$

Εφόσον η πιθανότητα να πάρει η τυχαία μεταβλητή X κάποιες τιμές στο $\mathbb{R}$, δίνεται μέσω του μέτρου πιθανότητας μ το οποίο είναι μία απεικόνιση απο τα υποσύνολα του Ω στο $[0,1]$, για να μιλήσουμε για την πιθανότητα αυτή θα πρέπει να γνωρίζουμε την αντίστροφη απεικόνιση X^{-1}. Η απεικόνιση αυτή είναι μία απεικόνιση απο το $\mathbb{R}$ σε ένα σύνολο υποσυνόλων του Ω, αλλά όχι απαραίτητα στο σύνολο που περιέχει όλα τα πιθανά υποσύνολα του Ω.

Παράδειγμα 2.4.2 *Ας προσπαθήσουμε να βρούμε το X^{-1} για το Παράδειγμα 2.4.1.*
Ακολουθώντας το παράδειγμα αυτό βλέπουμε ότι η X^{-1} θα πρέπει να περιέχει τα υποσύνολα $\{H_1H_2\}$, $\{H_1T_2, T_1H_2\}$, $\{T_1T_2\}$. Για να μπορούμε να χειριστούμε σωστά τις πιθανές ερωτήσεις για τις τιμές που μπορεί να πάρει η τυχαία μεταβλητή X θα πρέπει να κατασκευάσουμε την μικρότερη $\sigma-$άλγεβρα που περιέχει τα υποσύνολα αυτά. Αυτή είναι η

$$\mathcal{F}_1 = \{\emptyset, \Omega, \{H_1H_2\}, \{H_1T_2, T_1H_2\}, \{T_1T_2\}, \{H_1H_2\}^c, \{T_1T_2\}^c, \{H_1T_2, T_1H_2\}^c\}$$

Αυτή είναι μικρότερη απο την σ-άλγεβρα $\mathcal{F}$ η οποία περιεχει όλα τα υποσύνολα του Ω. Για παράδειγμα, η $\mathcal{F}$ θα περιέχει το $\{H_1T_2\}$ το οποίο δεν περιέχει η $\mathcal{F}_1$. Άρα $\mathcal{F}_1 \subset \mathcal{F}$ και ο εγκλεισμός είναι αυστηρός.
Η διαισθητική ερμηνεία του X^{-1} και της μικρότερης σ-άλγεβρας που το περιέχει είναι η εξής: Αν έχουμε γνώση των τιμών που παίρνει η τυχαία μεταβλητή X τι γνώση έχουμε σχετικά με τα γεγονότα που συνέβησαν, δηλαδή, πόσο η παρατήρηση των τιμών της X μας επιτρέπει να συνάγουμε συμπεράσματα σχετικά με το ποια γεγονότα του πειράματος πραγματοποιήθηκαν. Συνεπώς, αν ξέρω ότι $X = 0$ τότε γνωρίζω ότι πραγματοποιήθηκε είτε το γεγονός $\{H_1T_2\}$ είτε το γεγονός $\{T_1H_2\}$ δηλαδή το $\{H_1T_2, T_1H_2\}$ αλλά δεν μπορώ να ξεχωρίσω μετάξύ αυτών. Συνεπώς, έχω το υποσύνολο $\{H_1T_2, T_1H_2\}$ αλλά όχι το $\{H_1T_2\}$ ούτε το $\{T_1H_2\}$.

Το ερώτημα σχετικά με το σε ποιο σύνολο υποσυνόλων του Ω, δηλαδή σε ποιά $\sigma-$άλγεβρα βρίσκεται το X^{-1}, σχετίζεται με την **μετρησιμότητα** ή όχι της απεικόνισης X, ως προς συγκεκριμένες $\sigma-$άλγεβρες.

Ορισμός 2.4.1 *Ας πάρουμε ένα σύνολο Ω, μία σ-άλγεβρα απο υποσύνολα του $\mathcal{F}$ και μία απεικόνιση $X : \Omega \to \mathbb{R}$. Η X ονομάζεται μετρήσιμη ως προς την $\mathcal{F}$ αν όλα τα υποσύνολα του Ω που αποτελούν την X^{-1} ανήκουν στην $\mathcal{F}$.*

Παράδειγμα 2.4.3 *Ας πάρουμε $\Omega = \{H_1H_2, H_1T_2, T_1H_2, T_1T_2\}$ και τις σ-άλγεβρες*

$$\begin{aligned}
\mathcal{F}_1 &= \{\emptyset, \Omega, \{H_1H_2\}, \{H_1T_2, T_1H_2\}, \{T_1T_2\}, \{H_1H_2\}^c, \{T_1T_2\}^c, \{H_1T_2, T_1H_2\}^c\} \\
\mathcal{F}_2 &= \{\emptyset, \Omega, \{H_1H_2\}, \{H_1H_2\}^c\} \\
\mathcal{F}_3 &= \{\emptyset, \Omega, \{H_1T_2\}, \{H_1T_2\}^c\}
\end{aligned}$$

και $\mathcal{F}$ την σ-άλγεβρα που περιέχει όλα τα υποσύνολα του Ω.
Η απεικόνιση X του Παραδείγματος 2.4.1 είναι μετρήσιμη ως προς τις $\mathcal{F}_3$ και $\mathcal{F}$ αλλά δεν είναι μετρήσιμη ως προς τις $\mathcal{F}_2$ και $\mathcal{F}_3$.

Ας υποθέσουμε τώρα ότι έχουμε ένα χώρο πιθανοτήτων $(\Omega, \mathcal{F}, \mu)$ και μια τυχαία μεταβλητή $X : \Omega \to \mathbb{Z}$. Αντι να ρωτάμε τις ερωτήσεις 'ποιά η πιθανότητα να συμβεί το γεγονός $A \subset \Omega$' μπορεί να χρειαστεί να ρωτήσουμε την ερώτηση 'ποιά η πιθανότητα η τυχαία μεταβλητή X να πάρει τιμές στο σύνολο $B \subset \mathbb{Z}$'. Για να απαντήσουμε την ερώτηση αυτή θα πρέπει να ορίσουμε ένα μέτρο στα υποσύνολα του $\mathbb{Z}$ (δηλαδή στα υποσύνολα του πεδίου τιμών της X) και όχι στα υποσύνολα του δειγματοχώρου Ω.

Ορισμός 2.4.2 *Έστω $\mathcal{B} = 2^{\mathbb{Z}}$ το σύνολο που αποτελείται απο όλα τα υποσύνολα του $\mathbb{Z}$. Μπορούμε να ορίσουμε την απεικόνιση $\mu_X : \mathcal{B} \to [0,1]$ ως εξής*

$$\mu_X(B) = \mu(X^{-1}(B)) = \mu(\{\omega : X(\omega) \in B\}), \;\; \forall\, B \in \mathcal{B}$$

Το μ_X ονομάζεται το **επαγόμενο μέτρο** *απο την τυχαία μεταβλητή X.*

Παράδειγμα 2.4.4 *Αν Ω είναι το συνολο των ενδεχομένων για την ρίψη δύο τίμιων νομισμάτων και X είναι η τυχαία μεταβλητή η οποία δίνει το κέρδος ενός παίκτη απο το παιχνίδι αυτό, βρείτε το επαγόμενο μέτρο μ_X.*

Σχόλιο 2.4.1 *Το επαγόμενο μέτρο μπορεί να είναι ένα διακριτό μέτρο, ακόμα και αν το σύνολο των ενδεχομένων Ω δεν είναι διακριτό σύνολο, αν η τυχαία μεταβλητή X είναι διακριτή τυχαία μεταβλητή δηλαδή αν μπορεί να πάρει διακριτές τιμές x_i με πιθανότητα p_i. Αν $(\Omega, \mathcal{F}, \mu)$ είναι ο χώρος πιθανότητας και $(B, \mathcal{B}, \mu_X)$ είναι ο επαγόμενος χώρος πιθανότητας απο την διακριτή τυχαία μεταβλητή X τότε $p_i = \mu_X(X = x_i) = \mu(\{\omega,\ X(\omega) = x_i\})$. Σαν παράδειγμα αυτού μπορείτε να φανταστείτε ότι παίρνετε ένα δείγμα ω απο την ομοιόμορφη κατανομή στο διάστημα $[0,1]$ και αν $\omega \leq \frac{1}{2}$ τότε $X(\omega) = 1$ ενώ αν $\omega > \frac{1}{2}$ τότε $X(\omega) = -1$.*

2.5 Μέση τιμή και ολοκλήρωση επάνω σε μέτρα

Ας υποθέσουμε ότι X είναι μία τυχαία μεταβλητή σε ένα (διακριτό) χώρο πιθανοτήτων $(\Omega, \mathcal{F}, \mu)$. Εφόσον ο χώρος πιθανοτήτων είναι διακριτός $\Omega = \{\omega_1, \omega_2, \cdots\}$ η τυχαία μεταβλητή X θα παίρνει διακριτές τιμές $x_i = X(\omega_i)$. Θεωρούμε ότι $\mu(\omega_i) = p_i$.

Ορισμός 2.5.1 *Η μέση τιμή της τυχαίας μεταβλητής X ορίζεται ως*

$$\mathbb{E}_\mu[X] = \sum_{i=1}^{\infty} p_i \, x_i$$

Η μέση τιμή της τυχαίας μεταβλητής $f(X)$ ορίζεται ως

$$\mathbb{E}_\mu[f(X)] = \sum_{i=1}^{\infty} p_i \, f(x_i)$$

Η ύπαρξη ή όχι της μέσης τιμής μιας τυχαίας μεταβλητής εξαρτάται απο το αν το άπειρο άθροισμα που την ορίζει συγκλίνει ή όχι. Αυτό εξαρτάται απο το επαγόμενο μέτρο πιθανότητας απο την τυχαία αυτή μεταβλητή.

Παράδειγμα 2.5.1 *Δώστε ένα παράδειγμα μιας διακριτής τυχαίας μεταβλητής X για την οποία δεν ορίζεται η μέση τιμή.*

Η μέση τιμή μπορεί να θεωρηθεί σαν ένα άθροισμα επάνω σε όλα τα στοιχεία του συνόλου Ω αλλά ζυγισμένα με το μέτρο δηλαδή εναλλακτικά μπορούμε να έχουμε

Ορισμός 2.5.2 *Η μέση τιμή της τυχαίας μεταβλητής X μπορεί να γραφεί σαν*

$$\mathbb{E}_\mu[X] = \sum_{\omega \in \Omega} X(\omega)\, \mu(\omega) =: \int_\Omega X(\omega)\, d\mu(\omega)$$

Με τον ίδιο τρόπο για την τυχαία μεταβλητή $f(X)$ έχουμε ότι

$$\mathbb{E}_\mu[f(X)] = \sum_{\omega \in \Omega} f(X(\omega))\, \mu(\omega) =: \int_\Omega f(X(\omega))\, d\mu(\omega)$$

Σχόλιο 2.5.1 *Χρησιμοποιούμε τον συμβολισμό $\int_\Omega X(\omega)\, d\mu(\omega) := \sum_{\omega \in \Omega} X(\omega)\, \mu(\omega)$ για τα διακριτά μέτρα γιατί αργότερα θα ορίσουμε το ολοκλήρωμα επάνω σε ένα γενικό μέτρο σαν κατάλληλο όριο μιας προσέγγισης απο διακριτά μέτρα.*

Η μέση τιμή μιας τυχαίας μεταβλητής X μπορεί να γραφεί κάνοντας χρήση του επαγόμενου μέτρου P_X αντί για το μέτρο μ. Αν η τυχαία μεταβλητή X μπορεί να γραφεί με την αναπαράσταση

$$X(\omega) = x_1 \, \mathbf{1}_{A_1}(\omega) + \cdots + x_n \, \mathbf{1}_{A_n}(\omega)$$

όπου $x_i \in \mathbb{R}$ και $A_i \subset \Omega$, A_i ξένα μεταξύ τους, τότε μπορούμε να δούμε ότι

$$\begin{aligned} \mathbb{E}_\mu[X] &= x_1 \, \mu(A_1) + \cdots + x_n \, \mu(A_n) \\ &= x_1 \, \mu_X(\{X = x_1\}) + \cdots + x_n \, \mu_X(\{X = x_n\}) \end{aligned}$$

Αν λοιπόν

$$B = \{x \in \mathbb{R}, \text{ τέτοια ώστε } \exists \omega \in \Omega, X(\omega) = x\} = X^{-1}(\mathbb{R})$$

μπορούμε να γράψουμε

$$\mathbb{E}_\mu[X] = \sum_{x_i \in B} x_i \, \mu_X(\{X = x_i\}) =: \int_B X \, d\mu_X$$

Συνεπώς

$$\mathbb{E}_\mu[X] = \sum_{\omega \in \Omega} X(\omega)\mu(\omega) = \sum_{x_i \in B} x_i \, \mu_X(\{X = x_i\})$$

ή χρησιμοποιώντας τον συμβολισμό με τα ολοκληρώματα

$$\mathbb{E}_\mu[X] = \int_\Omega X(\omega) \, d\mu(\omega) = \int_B x \, d\mu_X(x)$$

Προσέξτε ότι το πρώτο άθροισμα (ολοκλήρωμα) είναι επάνω σε όλα τα ενδεχόμενα δηλαδή επάνω σε όλα τα $\omega \in \Omega$, ενώ το δεύτερο είναι επάνω σε όλες τις τιμές που μπορεί να πάρει η τυχαία μεταβλητή X.

Παράδειγμα 2.5.2 *Υπολογίστε την μέση τιμή της τυχαίας μεταβλητής X που εκφράζει το κέρδος απο παιχνίδι κορώνα-γράμματα στην ρίψη δύο τίμιων νομισμάτων με τους δύο παραπάνω τρόπους.*

2.6 Ιδιότητες της μέσης τιμής

Η μέση τιμή όπως την ορίσαμε για τα διακριτά μέτρα έχει τις ακόλουθες πολύ χρήσιμες ιδιότητες

Πρόταση 2.6.1 *Ισχύουν τα ακόλουθα:*

1. *Η μέση τιμή είναι ένας γραμμικός τελεστής δηλαδή*

$$\mathbb{E}_\mu[\lambda_1 \, X_1(\omega) + \lambda_2 \, X_2(\omega)] = \lambda_1 \mathbb{E}_\mu[X_1(\omega)] + \lambda_2 \mathbb{E}_\mu[X_2(\omega)]$$

   για οποιαδήποτε $\lambda_1, \lambda_2 \in \mathbb{R}$ (όχι τυχαίες μεταβλητές) και X_1, X_2 τυχαίες μεταβλητές.

2. *Η μέση τιμή είναι θετικός τελεστής, δηλαδή $\mathbb{E}_\mu[X] \geq 0$ αν $X \geq 0$ και αν $\mathbb{E}_\mu[X = 0]$ και $X \geq 0$ τότε $\mu_X(\{X \neq 0\}) = 0$.*

Ας θεωρήσουμε τώρα ακολουθίες απο τυχαίες μεταβλητές $X^{(n)}$ σε κάποιο χώρο πιθανοτήτων $(\Omega, \mathcal{F}, \mu)$. Μια ακολουθία απο τυχαίες μεταβλητές δεν είναι τίποτε άλλο απο μια οικογένεια απο τυχαίες μεταβλητές, όπου σε κάθε μέλος της οικογενείας δίνεται ένα όνομα με τον δείκτη $n \in \mathbb{N}$. Εφόσον οι τυχαίες μεταβλητές είναι διακριτές το μέλος της οικογένειας αυτής που αντιστοιχεί στην τιμή n, θα παίρνει τις τιμές $X^{(n)} = \{x_{n,1}, x_{n,2}, \cdots\}$, όπου αν π.χ. το Ω είναι διακριτό σύνολο θα συμβολίζουμε $x_{n,i} = X^{(n)}(\omega_i)$. Βλέπουμε λοιπόν ότι μια ακολουθία τυχαίων μεταβλητών σε ένα διακριτό χώρο πιθανότητων ή απλώς μια ακολουθία διακριτών τυχαίων μεταβλητών είναι ισοδύναμη με μια διπλή ακολουθία.

Μπορούμε τώρα να ορίσουμε την μέση τιμή κάθε μέλους της ακολουθίας δηλαδή την ποσότητα $\mathbb{E}_\mu[X^{(n)}]$ ως εξής

$$\mathbb{E}_\mu[X^{(n)}] = \sum_{m=0}^{\infty} p_m \, x_{n,m} = \sum_{\omega \in \Omega} X^{(n)}(\omega)\mu(\omega)$$

Σε πολλές περιπτώσεις η ακολουθία των τυχαίων μεταβλητών $X^{(n)}$ είναι τέτοια ώστε $\lim_{n\to\infty} x_{n,m} = x_m$ για κάθε $m \in \mathbb{N}$. Θα μας ενδιέφερε για αρκετές εφαρμογές να γνωρίζουμε πότε μπορούμε να γράψουμε

$$\lim_{n\to\infty} \sum_{m=0}^{\infty} p_m\, x_{n,m} = \sum_{m=0}^{\infty} \lim_{n\to\infty} p_m\, x_{n,m} = \sum_{m=0}^{\infty} p_m\, x_m$$

ή αν γράψουμε $\lim_{n\to\infty} X^{(n)} = X$ όπου $X = (x_1, x_2, \cdots)$

$$\lim_{n\to\infty} \mathbb{E}_\mu[X^{(n)}] = \mathbb{E}_\mu[\lim_{n\to\infty} X^{(n)}] = \mathbb{E}_\mu[X]$$

Απο τις ιδιότητες των διπλών σειρών έχουμε τα ακολουθα

Πρόταση 2.6.2 *1. (Μονότονη σύγκλιση) Έστω $X, X^{(n)}$ τέτοιες ώστε $x_{n,m} \le x_{n+1,m}$ και $x_{n,m} \to x_m$. Τότε $\sum_{\omega\in\Omega} X^{(n)}(\omega)\mu(\omega) \to \sum_{\omega\in\Omega} X(\omega)\mu(\omega)$ ή ισοδύναμα $\mathbb{E}_\mu[X^{(n)}] \to \mathbb{E}_\mu[X]$.*

2. (Κυριαρχημένη σύγκλιση) Έστω $X^{(n)}, X$ τέτοιες ώστε $x_{n,m} \to x_m$ και $\mid x_{n,m} \mid \le y_i$ για κάθε m, με $\sum_{\omega\in\Omega} Y(\omega)\mu(\omega)$. Τότε, $\sum_{\omega\in\Omega} X^{(n)}(\omega)\mu(\omega) \to \sum_{\omega\in\Omega} X(\omega)\mu(\omega)$ ή ισοδύναμα $\mathbb{E}_\mu[X^{(n)}] \to \mathbb{E}_\mu[X]$.

3. (Fatou) Ισχύει ότι

$$\liminf_{n\to\infty} \mathbb{E}_\mu[X^{(n)}] \le \mathbb{E}_\mu[\liminf_{n\to\infty} X^{(n)}]$$

Παράδειγμα 2.6.1 *Σαν ένα παράδειγμα του θεωρήματος κυριαρχημένης σύγκλισης για τα διακριτά μέτρα μπορούμε να δώσουμε την συνέχεια της χαρακτηριστικής συνάρτησης ή της ροπογεννήτριας. Αν X είναι μια διακριτή τυχαία μεταβλητή τέτοια ώστε $P(X = x_k) = \mu_X(\{X = x_k\}) = p_k$, $k \in \mathbb{N}$, η χαρακτηριστική συνάρτηση της είναι η*

$$\mu_X(\lambda) = \mathbb{E}_\mu[\exp(i\,\lambda\, X)] = \sum_k p_k\, \exp(i\,\lambda\, X_k)$$

όπου i είναι η φανταστική μονάδα, $i^2 = -1$. Η χαρακτηριστική συνάρτηση είναι συνεχής συνάρτηση της μεταβλητής λ, δηλαδή για κάθε ακολουθία $\lambda_n \to \lambda$ έχουμε ότι $\mu_X(\lambda_n) \to \mu_X(\lambda)$. Η απόδειξη αυτής της πολύ σημαντικής ιδιότητας είναι επακόλουθο της κυριαρχημένης συγκλισης για διακριτά μέτρα.

Παράδειγμα 2.6.2 *Αν $X^{(n)}$ είναι μία ακολουθία διακριτών μη αρνητικών τυχαίων μεταβλητών τότε ισχύει ότι*

$$\mathbb{E}_\mu[\sum_{n=1}^{\infty} X^{(n)}] = \sum_{n=1}^{\infty} \mathbb{E}_\mu[X^{(n)}]$$

Αυτό είναι μια εφαρμογή του θεωρήματος μονότονης σύγκλισης για την μονότονη ακολουθία $S^{(n)} = \sum_{k=1}^{n} X^{(k)}$.

Αν φύγουμε απο τα διακριτά μέτρα, η μέση τιμή δεν μπορεί πλέον να ερμηνευθεί σαν άθροισμα επάνω σε όλα τα στοιχεία του συνόλου Ω αλλά σαν ένα ολοκλήρωμα. Σκοπός αυτού του μαθήματος είναι να ορίσουμε επακριβώς και να κατανοήσουμε τις ιδιότητες ενός ολοκληρώματος $\int_\Omega X(\omega)\, dP(\omega)$ όπου P είναι ένα μέτρο πιθανότητας. Επίσης θα δείξουμε ότι οι παραπάνω ιδιότητες συνεχίζουν να ισχύουν αν ερμηνεύσουμε σωστα ανισότητες όπως $X(\omega) \ge 0$ ή την σχέση $P(X \ne 0) = \mu_X(\{X \ne 0\}) = 0$.

2.7 Χώροι ολοκληρώσιμων τυχαίων μεταβλητών

Θα μελετήσουμε τώρα το σύνολο των διακριτών τυχαίων μεταβλητών X για τις οποίες ισχύει $\mathbb{E}_\mu[\mid X \mid] < \infty$. Θα συμβολίζουμε τον σύνολο αυτό με $\ell^1(\mu)$ δηλαδή

$$\ell^1(\mu) = \{X \mid \mathbb{E}_\mu[\mid X \mid] < \infty\}$$

ή ισοδύναμα αν χαρακτηρίσουμε την τυχαία αυτή μεταβλητή με τα $X(\omega_i) = x_i$ και το μέτρο μ με την ακολουθία $(p_1, p_2, \cdots)$, μπορούμε να χαρακτηρίσουμε το $\ell^1(\mu)$ σαν το σύνολο

$$\ell^1(\mu) = \{(x_1, x_2, \cdots) \mid \sum_{i=1}^{\infty} \mid x_i \mid p_i < \infty\}$$

Το σύνολο αυτό περιέχει ακολουθίες $\{x_i\}$ οι οποίες είναι αθροίσιμες αν κάθε όρος σταθμιστεί με τα $\{p_i\}$. Το σύνολο αυτό είναι ένας **διανυσματικός χώρος** αν ορίσουμε την πράξη της πρόσθεσης και την πράξη του πολλαπλασιασμού με πραγματικό αριθμό ως ακολούθως

1. $(X+Y)(\omega_i) = X(\omega_i) + Y(\omega_i)$,
2. $(\lambda X)(\omega_i) = \lambda X(\omega_i)$

Επίσης, αν ορίσουμε την απεικόνιση $|| \cdot ||_{\ell^1(\mu)}$: $\ell^1(\mu) \to [0, \infty)$ με τον ακόλουθο τρόπο

$$|| X ||_{\ell^1(\mu)} = \mathbb{E}_\mu[| X |] = \sum_{i=1}^{\infty} | x_i | \ p_i$$

η απεικόνιση αυτή είναι μία νόρμα για τον χώρο $\ell^1(\mu)$.
Με αυτό εννοούμε ότι ικανοποιεί τις ιδιότητες

1. $|| X ||_{\ell^1(\mu)} = 0$ αν και μόνο αν $X(\omega_i) = 0$ για κάθε i τέτοιο ώστε $p_i > 0$.
2. $\lambda || X ||_{\ell^1(\mu)} = | \lambda | \ || X ||_{\ell^1(\mu)}$.
3. $|| X_1 + X_2 ||_{\ell^1(\mu)} \leq || X_1 ||_{\ell^1(\mu)} + || X_2 ||_{\ell^1(\mu)}$

Ο χώρος $\ell^1(\mu)$ εφοδιασμένος με την νόρμα $|| \cdot ||_{\ell^1(\mu)}$ ονομάζεται ο χώρος των ολοκληρώσιμων τυχαίων μεταβλητών κάτω απο το διακριτό μέτρο μ.

Μπορούμε να ορίσουμε την έννοια της σύγκλισης ακολουθίων καθώς και την έννοια της ακολουθίας Cauchy στον χώρο αυτό.

Ορισμός 2.7.1 *(ι) Λέμε ότι η ακολουθία $X^{(n)} \to X$ στον $\ell^1(\mu)$ αν $|| X^{(n)} - X ||_{\ell^{(1)}(\mu)} \to 0$.*

(ιι) Λέμε ότι η ακολουθία $X^{(n)}$ είναι ακολουθία Cauchy στον $\ell^1(\mu)$ αν για κάθε $\epsilon > 0$ υπάρχει $N \in \mathbb{N}$ τέτοιο ώστε $|| X^{(n)} - X^{(m)} ||_{\ell^{(1)}(\mu)} < \epsilon$ για $n, m \geq N$.

Ο χώρος αυτός έχει την πολύ σημαντική ιδιότητα του ότι είναι **πλήρης** δηλαδή είναι ένας χώρος στο οποίο κάθε ακολουθία Cauchy είναι συγκλίνουσα.

Πρόταση 2.7.1 *Ας πάρουμε μία ακολουθία $X^{(n)} \in \ell^1(\mu)$ τέτοια ώστε για κάθε $\epsilon > 0$ να υπάρχει N έτσι ώστε $|| X^{(n)} - X^{(m)} ||_{\ell^1(\mu)} < \epsilon$, αν $n, m > N$. Τότε υπάρχει κάποιο $X \in \ell^1(\mu)$ τέτοιο ώστε $\lim_{n\to\infty} X^{(n)} = X$ στον $\ell^1(\mu)$, όπου με το τελευταίο εννοούμε ότι για κάθε $\epsilon > 0$ υπάρχει N τέτοιο ώστε $|| X^{(n)} - X ||_{\ell^1(\mu)} < \epsilon$ για $n > N$.*

Απόδειξη: Βλ. παράρτημα Παράγραφος 2.12.1.□

Τα παραπάνω γενικεύονται και για διακριτές τυχαίες μεταβλητές οι οποίες έχουν ροπές μεγαλύτερης τάξης της πρώτης. Μπορούμε να ορίσουμε τους χώρους $\ell^p(\mu)$ ως τα σύνολα

$$\ell^p(\mu) = \{(x_1, x_2, \cdots) \ | \ \sum_{i=1}^{\infty} | x_i |^p \ p_i < \infty\}$$

Οι χώροι αυτοί για $p \geq 1$ αν εφοδιαστούν με την νόρμα

$$|| X ||_{\ell^p(\mu)} = \mathbb{E}_\mu[| X |^p] = \sum_{i=1}^{\infty} | x_i |^p \ p_i \tag{2.2}$$

και για την προφανή γενίκευση του Ορισμού 2.7.1 μπορεί να αποδειχθεί ότι είναι πλήρεις μετρικοί χώροι, δηλαδή είναι χώροι Banach. Στην ειδική περίπτωση όπου $p = 2$ η νόρμα (2.2) παράγεται απο το εσωτερικό γινόμενο

$$(X, Y) = \sum_{i=1}^{\infty} p_i \, x_i \, y_i = \mathbb{E}_\mu[X \, Y]$$

όπου με $X Y$ συμβολίζουμε την τυχαία μεταβλητή $(x_1 y_1, x_2 y_2, \cdots)$. Στην περίπτωση $p = 2$ ο χώρος $\ell^2(P)$ είναι χώρος Hilbert και οι ιδιότητες του είναι πολύ χρήσιμες και θεμελιώδεις για τις πιθανότητες και την στατιστική.

2.8 Δεσμευμένη πιθανότητα και δεσμευμένα μέτρα

Ορισμός 2.8.1 *Έστω $(\Omega, \mathcal{F}, \mu)$ ένας χώρος πιθανοτήτων και $A, B \in \mathcal{F}$ δυο γεγονότα, για τα οποία ισχύει ότι $\mu(B) > 0$. Η δεσμευμένη πιθανότητα του A δεδομένου του B είναι η $\mu(A \mid B) = \frac{\mu(A\cap B)}{\mu(B)}$.*

Πρόταση 2.8.1 *Η απεικόνιση $A \to \mu(A \mid B)$ απο το $\mathcal{F} \to [0,1]$ ορίζει ένα νέο μέτρο πιθανότητας μ_B στο $\mathcal{F}$ το οποίο ονομάζεται δεσμευμένο μέτρο πιθανότητας δεδομένου του B.*

Για να δείξουμε ότι είναι μέτρο αρκεί να δείξουμε την ιδιότητα της αθροισιμότητας. Πράγματι αν υποθέσουμε ότι παίρνουμε μία ακολουθία γεγονότων $A_i \in \mathcal{F}$, $i \in \mathbb{N}$ που είναι ανα δύο ξένα. Θα δείξουμε ότι

$$\mu_B\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu_B(A_i).$$

Απο τον ορισμό του μ_B έχουμε ότι

$$\mu_B\left(\bigcup_{i=1}^{\infty} A_i\right) = \mu\left(\bigcup_{i=1}^{\infty} A_i \,\middle|\, B\right) = \frac{\mu\left(\left(\bigcup_{i=1}^{\infty} A_i\right)\cap B\right)}{\mu(B)} = \frac{\mu\left(\bigcup_{i=1}^{\infty}(A_i \cap B)\right)}{\mu(B)} \tag{2.3}$$

Όμως, εφόσον τα A_i είναι ανα δύο ξένα, και τα $E_i := (A_i \cap B)$ είναι επίσης ανα δύο ξένα, συνεπώς χρησιμοποιώντας τις ιδιότητες του μέτρου μ καταλήγουμε ότι

$$\mu\left(\left(\bigcup_{i=1}^{\infty} A_i\right)\cap B\right) = \sum_{i=1}^{\infty} \mu(A_i \cap B)$$

Αντικαθιστώντας αυτό στην (2.3) βλέπουμε ότι

$$\mu_B\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \frac{\mu(A_i \cap B)}{\mu(B)} = \sum_{i=1}^{\infty} \mu(A_i \mid B) = \sum_{i=1}^{\infty} \mu_B(A_i)$$

Το δεσμευμένο μέτρο που ορίσαμε παραπάνω μπορεί να χρησιμοποιηθεί για την μελέτη των δεσμευμένων πιθανοτήτων καθώς και των διαφόρων δεσμευμένων μέσων τιμών.

Επίσης μπορούμε να δούμε το παραπάνω παράδειγμα πιο γενικά υπο το πρίσμα του περιορισμού ενός μέτρου σε μία μικρότερη $\sigma-$άλγεβρα απο την αρχική.

Ορισμός 2.8.2 *Έστω $(\Omega, \mathcal{F}, \mu)$ ένας χώρος μέτρου και $B \in \mathcal{F}$. Το μέτρο μ_B το οποίο ορίζεται απο την σχέση $\mu_B(A) = \mu(A \cap B)$ για κάθε $A \in \mathcal{F}$ ονομάζεται ο* **περιορισμός** *του μ στο B.*

Σχόλιο 2.8.1 *Για να γίνει το μ_B του Ορισμού 2.8.2 μέτρο πιθανότητας πρέπει να διαιρέσουμε με το $\mu(B)$.*

Το μέτρο του Ορισμού 2.8.2 ορίζεται σε όλο το $\mathcal{F}$, απλά δίνει την τιμή 0 στα $A \in \mathcal{F}$ για τα οποία ισχύει $A \cap B = \emptyset$. Υπάρχει και ένας διαφορετικός τρόπος να ορίσουμε τον περιορισμό ενός μέτρου σε κάποιο $B \in \mathcal{F}$.

Ορισμός 2.8.3 *Έστω $(\Omega, \mathcal{F})$ ένας μετρήσιμος χώρος και $B \in \mathcal{F}$. Ο περιορισμός της $\sigma-$άλγεβρας $\mathcal{F}$ στο B, ορίζεται σαν το*

$$\mathcal{F} \rceil B := \{A \cap B \mid A \in \mathcal{F}\} = \{A \in \mathcal{F} \mid A \subset B\}$$

Ορισμός 2.8.4 *Ας ορίσουμε το μέτρο $\mu \rceil B : \mathcal{F} \rceil B \to [0, \infty)$ με τον ακόλουθο τρόπο*

$$(\mu \rceil B)(A) = \mu(A), \;\; A \in \mathcal{F} \rceil B$$

Το μέτρο $\mu \rceil B$ στον μετρήσιμο χώρο $(B, \mathcal{F} \rceil B)$ ονομάζεται ο περιορισμός του μέτρου μ στην $\mathcal{F} \rceil B$.

Ο περιορισμός αυτός απλά αγνοεί όλα τα σύνολα A, για τα οποία $A \cap B = \emptyset$ εφόσον ορίζεται στην $\sigma-$άλγεβρα $\mathcal{F} \rceil B$. Φυσικά και οι δύο αυτοί περιορισμοί δίνουν τις ίδιες τιμές στα σύνολα $A \in \mathcal{F}$ τα οποία περιέχονται στο B.

Παράδειγμα 2.8.1 *Κάντε τις κατασκευές των δύο παραπάνω ορισμών για το παράδειγμα της ρίψης δύο νομισμάτων στο οποίο χρησιμοποιείτε ως B το γεγονός $B = \{H_1H_2, H_1T_2\}$, δηλαδή το γεγονός ότι το πρώτο νόμισμα έχει φέρει κορώνα.*

Παράδειγμα 2.8.2 *Ας πάρουμε το παράδειγμα της διαδοχικής ρίψης τριών νομισμάτων αξία 10, 20 και 50 λεπτών αντίστοιχα. Ας υποθέσουμε ότι παίζουμε το ακόλουθο παιχνίδι: Το συνολικό μας κέρδος είναι το άθροισμα της αξίας των νομισμάτων που έφεραν κορώνα. Έστω X η τυχαία μεταβλητή που δίνει το κέρδος απο το παιχνίδι αυτό. Ποια είναι η πιθανότητα κερδίσουμε 0,10,30, 50, 70, 80 λεπτά αντιστοίχως δεδομένου ότι δύο νομίσματα (χωρίς να γνωρίζουμε ακριβώς ποιά) έχουν φέρει κορώνα;*
Το γεγονός {δύο νομίσματα έχουν φέρει κορώνα} ισοδυναμεί με το γεγονός $B = \{H_1H_2T_3, H_1T_2H_3, T_1H_2H_3\}$.
Η $\sigma-$άλγεβρα $\mathcal{F} \rceil B$ είναι η

$$\mathcal{F} \rceil B = \{\emptyset, B, \{H_1H_2T_3\}, \{H_1T_2H_3, T_1H_2H_3\}, \{H_1T_2H_3\}, \{H_1H_2T_3, T_1H_2H_3\}, \{T_1H_2H_3\}, \{H_1H_2T_3, H_1T_2H_3\}, \}$$

Το γεγονός π.χ. $\{\omega : X(\omega) = 50\}$ είναι ισοδύναμο με το γεγονός $A = \{T_1T_2H_3\}$. Σύμφωνα με τον ορισμό της δεσμευμένης πιθανότητας

$$\mu(A \mid B) = \frac{\mu(A \cap B)}{\mu(B)} = \frac{\mu(\emptyset)}{\mu(B)} = 0$$

Αν χρησιμοποιούσαμε τον περιορισμό του μέτρου $\mu \rceil B$ το ερώτημα αυτό δεν έχει νόημα εφόσον $A \notin \mathcal{F} \rceil B$.
Το γεγονός π.χ. $\{\omega : X(\omega) = 70\}$ είναι ισοδύναμο με το γεγονός $A = \{T_1H_2H_3\}$. Σύμφωνα με τον ορισμό της δεσμευμένης πιθανότητας

$$\mu(A \mid B) = \frac{\mu(A \cap B)}{\mu(B)} = \frac{\mu(A)}{\mu(B)} = \frac{\frac{1}{8}}{\frac{3}{8}} = \frac{1}{3}$$

Αν χρησιμοποιούσαμε τον περιορισμό του μέτρου $\mu \rceil B$ το ερώτημα αυτό τώρα έχει νόημα εφόσον $A \in \mathcal{F} \rceil B$. Σύμφωνα με τον ορισμό του μέτρου αυτού $\mu \rceil B(A) = \mu(A) = \frac{1}{8}$. Για να γίνει όμως αυτό μέτρο πιθανοτήτων πρέπει να κανονικοποιήσουμε διαιρώντας ως προς $\mu(B) = \frac{3}{8}$. Αυτό δίνει και το ζητούμενο αποτέλεσμα.

2.9 Ανεξαρτησία και γινόμενο μέτρο

Ας ξεκινήσουμε πρώτα με την έννοια της ανεξαρτησίας γεγονότων.

Ορισμός 2.9.1 *Ας πάρουμε ένα χώρο πιθανοτήτων $(\Omega, \mathcal{F}, \mu)$ και δύο γεγονότα $A, B \in \mathcal{F}$. Τα γεγονότα ονομάζονται ανεξάρτητα αν $\mu(A \cap B) = \mu(A)\,\mu(B)$.*

Η διαισθητική ερμηνεία αυτού είναι ότι το γεγονός A **δεν** επηρεάζει το γεγονός B και αντίστροφα.

Παράδειγμα 2.9.1 *Ας πάρουμε το πείραμα της ρίψης δύο νομισμάτων. Ο χώρος γεγονότων μπορεί να είναι το*

$$\Omega = \{H_1H_2, H_1T_2, T_1H_2, T_1T_2\}$$

Σαν μέτρο πιθανότητας ας πάρουμε το μέτρο αρίθμησης στο Ω.
Ας πάρουμε τώρα τα υποσύνολα

$$\begin{aligned} A &= \{H_1H_2, H_1T_2\} \\ B &= \{H_1H_2, T_1H_2\} \end{aligned}$$

Το γεγονός A είναι το γεγονός το πρώτο νόμισμα φέρνει κορώνα, ενώ το γεγονός B είναι το γεγονός το δεύτερο νόμισμα φέρνει κορώνα. Τα A και B είναι ανεξάρτητα γεγονότα.

Το παράδειγμα αυτό μπορούμε να το δούμε με ένα διαφορετικό πρίσμα.

Παράδειγμα 2.9.2 *Ας πάρουμε τα σύνολα $\Omega_1 = \{H_1, T_1\}$, $\Omega_2 = \{H_2, T_2\}$ και ας ορίσουμε το $\Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2), \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$. Αν ακολουθήσουμε τον συμβολισμό $(\omega_1, \omega_2) \equiv \omega_1\omega_2$, όπου $\omega_1 = H_1, T_1$, $\omega_2 = H_2, T_2$ βλέπουμε ότι $\Omega = \Omega_1 \times \Omega_2$.*
Θα λέμε ότι το σύνολο $A \subset \Omega_1 \times \Omega_2$ αν $A = A_1 \times A_2$ για $A_1 \subset \Omega_1$, $A_2 \subset \Omega_2$. Για παράδειγμα $A = \{H_1T_2, H_1H_2\} \subset \Omega_1 \times \Omega_2$ εφόσον $A = \{H_1\} \times \{H_2, T_2\}$.
Ας εφοδιάσουμε τώρα το Ω_1 με το μέτρο αρίθμησης μ_1 σύμφωνα με το οποίο $\mu_1(\{H_1\}) = \mu_1(\{T_1\}) = \frac{1}{2}$ και το Ω_2 με το μέτρο αρίθμησης μ_2 σύμφωνα με το οποίο $\mu_2(\{H_2\}) = \mu_2(\{T_2\}) = \frac{1}{2}$.
Μπορούμε να ορίσουμε μία νέα απεικόνιση μ απο τα υποσύνολα $\Omega = \Omega_1 \times \Omega_2$ στο $[0, 1]$ με την σχέση

$$\mu(A_1 \times A_2) = \mu_1(A_1)\,\mu_2(A_2)$$

και να δείξουμε ότι είναι ένα μέτρο.
Απο τον ορισμό αυτό μπορούμε να περιγράψουμε πλήρως το μέτρο μ *ως*

$$\begin{aligned}
\mu(\{H_1H_2\}) &= \mu_1(\{H_1\})\mu_2(\{H_2\}) = \frac{1}{2}\frac{1}{2} = \frac{1}{4} \\
\mu(\{H_1T_2\}) &= \mu_1(\{H_1\})\mu_2(\{T_2\}) = \frac{1}{2}\frac{1}{2} = \frac{1}{4} \\
\mu(\{T_1H_2\}) &= \mu_1(\{T_1\})\mu_2(\{H_2\}) = \frac{1}{2}\frac{1}{2} = \frac{1}{4} \\
\mu(\{T_1T_2\}) &= \mu_1(\{T_1\})\mu_2(\{T_2\}) = \frac{1}{2}\frac{1}{2} = \frac{1}{4}
\end{aligned}$$

Το μέτρο αυτό ονομάζεται **γινόμενο μέτρο** *και όπως βλέπουμε σχετίζεται άμεσα με την ανεξαρτησία.*

Ορισμός 2.9.2 *Το* **γινόμενο μέτρο** *μπορεί να οριστεί για οποιουσδήποτε σύνολα και μέτρα* (Ω_1, μ_1),(Ω_2, μ_2) *με τον παραπάνω ορισμό*

$$\mu(A_1 \times A_2) = \mu_1(A_1)\, \mu_2(A_2)$$

Το πεδίο ορισμού του γινόμενου μέτρου είναι η $\sigma-$*άλγεβρα γινόμενο* $\mathcal{F}_1 \times \mathcal{F}_2$ *η οποία ορίζεται ως η μικρότερη* $\sigma-$*άλγεβρα η οποία περιέχει σύνολα της μορφής* $A_1 \times A_2$*, όπου* $A_1 \in \mathcal{F}_1$ *και* $A_2 \in \mathcal{F}_2$.

Το γινόμενο μέτρο μπορούμε να το δούμε και λίγο διαφορετικά.

Ορισμός 2.9.3 *Ας πάρουμε ένα οποιοδήποτε υποσύνολο* $C \subset \Omega_1 \times \Omega_2$. *Για κάθε* $\omega_1 \in \Omega_1$ *μπορούμε να ορίζουμε την* ω_1- *τομή του* C *ως*

$$C_{\omega_1} = \{\omega_2 : (\omega_1, \omega_2) \in C \subset \Omega_1 \times \Omega_2\}$$

Με παρόμοιο τρόπο ορίζεται και η ω_2- *τομή του* C.

Πρόταση 2.9.1 *Ας πάρουμε δύο χώρους μέτρου* $(\Omega_1, \mathcal{F}_1, \mu_1)$, $(\Omega_2, \mathcal{F}_2, \mu_2)$ *και έστω κάποιο οποιοδήποτε* $C \in \mathcal{F}_1 \times \mathcal{F}_2$. *Τότε,*

1. *Κάθε* $C_{\omega_1} \in \mathcal{F}_2$ *και κάθε* $C_{\omega_2} \in \mathcal{F}_1$.

2. *Ας ορίσουμε την απεικόνιση*

$$\phi(C) = \int_{\Omega_2} \mu_1(C_{\omega_2})\, d\mu_2(\omega_2) = \int_{\Omega_1} \mu_2(C_{\omega_1})\, d\mu_1(\omega_1)$$

   όπου φυσικά $\int_{\Omega_2} \mu_1(C_{\omega_2})\, d\mu_2(\omega_2) := \sum_{\omega_2 \in \Omega_2} \mu_1(C_{\omega_2})\, \mu_2(\omega_2)$.
   Η απεικόνιση αυτή είναι ένα μέτρο και μάλιστα $\phi = \mu_1 \times \mu_2$.

Ας θεωρήσουμε τώρα μία τυχαία μεταβλητή $X : \Omega_1 \times \Omega_2 \to \mathbb{R}$.

Ορισμός 2.9.4 *Για κάθε* $\omega_1 \in \Omega_1$ *η τυχαία μεταβλητή* $X_\omega : \Omega_2 \to \mathbb{R}$ *η οποία ορίζεται ως* $X_{\omega_1}(\omega_2) = X(\omega_1\omega_2)$ *ονομάζεται η* $\omega-$*τομή της* X.

Έχουμε λοιπόν το παρακάτω βασικό αποτέλεσμα για τα διακριτά μέτρα γινόμενα.

Πρόταση 2.9.2 *Ας πάρουμε δύο χώρους μέτρου* $(\Omega_1, \mathcal{F}_1, \mu_1)$, $(\Omega_2, \mathcal{F}_2, \mu_2)$ *και το γίνομενο μέτρο* $\phi = \mu_1 \times \mu_2$. *Ας υποθέσουμε ότι η* $X(\omega_1, \omega_2)$ *είναι ολοκληρώσιμη ως προς το μέτρο* ϕ.
Τότε,

1. *Οι τομές* $X_{\omega_2}(\cdot, \omega_2)$ *είναι* μ_1-*ολοκληρώσιμες*

2. *Η* $h(\omega_2) := \int_{\Omega_1} X_{\omega_2}(\omega_1)\, d\mu_1(\omega_1)$ *είναι* μ_2-*ολοκληρώσιμη*

3. *Ισχύει ότι*

$$\int_{\Omega_1 \times \Omega_2} X(\omega_1, \omega_2)\, d\phi(\omega_1, \omega_2) = \int_{\Omega_2} \left[\int_{\Omega_1} X(\omega_1, \omega_2)\, d\mu_1(\omega_1)\right] d\mu_2(\omega_2) = \int_{\Omega_2} h(\omega_2)\, d\mu_2(\omega_2)$$

4. *Ισχύει ότι*

$$\int X\,d\phi = \int \left[\int X\,d\mu_1 \right] d\mu_2 = \int \left[\int X\,d\mu_2 \right] d\mu_1$$

όπου φυσικά με τον συμβολισμό $\int X\,d\mu$ *εννοούμε το αντίστοιχο άθροισμα.*

Η κατασκευή αυτή επαγωγικά μπορεί να γενικευθεί σε οποιαδήποτε αριθμό μέτρων ακόμα και σε άπειρο, και με την χρήση του γινόμενου αυτού μέτρου και των ιδιοτήτων του μπορούμε να καταλάβουμε σε βάθος ερωτήματα τα οποία έχουν σχέση με επαναλαμβανόμενα πειράματα, π.χ. νόμους μεγάλων αριθμών, συγκλίσεις κλπ. Τέτοιου τύπου ερωτήματα είναι θεμελιώδη για την στατιστική επιστήμη. Επίσης, αν μπορέσουμε να ορίσουμε το ολοκλήρωμα επάνω στο γινόμενο μέτρο μπορούμε να καταλάβουμε τις ιδιότητες των μέσων τιμών ανεξάρτητων τυχαίων μεταβλητών.

Ένας λοιπόν απο τους σκοπούς του μαθήματος αυτού είναι η κατασκευή και η κατανόηση της έννοιας του γινόμενου μέτρου, και της σχέσης του με την ανεξαρτησία, σε περιπτώσεις όπου μας ενδιαφέρουν ερωτήματα που σχετίζονται με γενικότερα μέτρα απο τα διακριτά μέτρα. Θα επανέλθουμε σε τέτοια ερωτήματα σε ερχόμενα κεφάλαια.

2.10 Δεσμευμένη μέση τιμή

Ας υποθέσουμε ένα χώρο πιθανοτήτων $(\Omega, \mathcal{F}, \mu)$ και ας επιλέξουμε ένα γεγονός $B \in \mathcal{F}$ τέτοιο ώστε $\mu(B) > 0$, και μία τυχαία μεταβλητή $X : \Omega \to \mathbb{R}$.

Ορισμός 2.10.1 *Η δεσμευμένη μέση τιμή της X δεδομένου του γεγονότος B, είναι η ποσότητα*

$$\mathbb{E}_\mu[X \mid \mathcal{B}] = \frac{1}{\mu(B)} \sum_{\omega \in B} X(\omega)\,\mu(\omega)$$

Μπορούμε να δούμε ότι αυτό μπορεί να ξαναγραφεί ως

$$\mathbb{E}_\mu[X \mid \mathcal{B}] = \sum_{\omega \in \Omega \cap B} X(\omega)\,\frac{\mu(\omega)}{\mu(B)} = \sum_{\omega \in \Omega} X(\omega)\,\mu_B(\omega) = \sum_{\omega \in B} X(\omega)\,(\mu \rceil B)(\omega)$$

δηλαδη η δεσμευμένη μέση τιμή μπορεί να γραφεί σαν μέση τιμή επάνω στα περιορισμένα μέτρα. Συνεπώς, όλες οι ιδιότητες της μέσης τιμής ισχύουν **και** για την δεσμευμένη μέση τιμή.
Ισχύει η ακόλουθη ιδίοτητα:

$$\mathbb{E}_\mu[X \mid B]\mu(B) = \sum_{\omega \in B} X(\omega)\,\mu(\omega) = \mathbb{E}_\mu[X\,\mathbf{1}_B]$$

Μπορούμε τώρα να επεκτείνουμε την έννοια της δεσμευμένης μέσης τιμής ως προς ένα γεγονός, στην έννοια της δεσμευμένης μέσης τιμής ως προς μία συλλογή γεγονότων, δηλαδή ως προς μία $\sigma-$άλγεβρα. Ας πάρουμε μία $\sigma-$άλγεβρα $\mathcal{G} \subset \mathcal{F}$. Η $\mathcal{G} = \{\omega, B_1, B_2, \cdots\}$ όπου $B_i \in \mathcal{F}$.

Ορισμός 2.10.2 *Η δεσμευμένη μέση τιμή της τυχαίας μεταβλητής X ως προς την σ-άλγεβρα $\mathcal{G}$ είναι η* **τυχαία μεταβλητή** $\mathbb{E}_\mu[X \mid \mathcal{G}] : \Omega \to \mathbb{R}$ *η οποία ορίζεται βάσει του κανονα*

$$\mathbb{E}_\mu[X \mid \mathcal{G}](\omega) = \mathbb{E}_\mu[X \mid B_i], \text{ αν } \omega \in B_i$$

Η τυχαία αυτή μεταβλητή έχει τις εξής ιδιότητες:

Πρόταση 2.10.1 *Έστω $(\Omega, \mathcal{F}, \mu)$ ένας χώρος πιθανοτήτων, $X : \Omega \to \mathbb{R}$ μία τυχαία μεταβλητή και $\mathcal{G} \subset \mathcal{F}$ μία σ-άλγεβρα. Η τυχαία μεταβλητή $\mathbb{E}[X \mid \mathcal{G}]$ ικανοποιεί τις ακόλουθες ιδιότητες:*

1. *Είναι μετρήσιμη ως προς την $\sigma-$άλγεβρα $\mathcal{G}$.*

2. *Για κάθε $B_i \in \mathcal{G}$ ισχύει*

$$\sum_{\omega \in B_i} \mathbb{E}_\mu[X \mid \mathcal{G}](\omega)\,\mu(\omega) = \sum_{\omega \in B_i} X(\omega)\,\mu(\omega) \tag{2.4}$$

Η ιδιότητα 2 είναι πολύ χρήσιμη και μπορεί να χρησιμοποιηθεί σαν ο ορισμός της δεσμευμένης μέσης τιμής ως προς μία γενική $\sigma-$άλγεβρα.

Ας δούμε την απόδειξη του 2. Εφόσον $\omega \in B_i$ έχουμε ότι

$$\mathbb{E}_\mu[X \mid \mathcal{G}](\omega) = \mathbb{E}_\mu[X \mid B_i] = \frac{1}{\mu(B_i)} \sum_{\omega \in B_i} X(\omega)\,\mu(\omega).$$

Κατά συνέπεια

$$\begin{aligned}\sum_{\omega \in B_i} \mathbb{E}_\mu[X \mid \mathcal{G}](\omega)\,\mu(\omega) &= \left(\frac{1}{\mu(B_i)} \sum_{\omega \in B_i} X(\omega)\,\mu(\omega)\right) \sum_{\omega \in B_i} \mu(\omega) \\ &= \left(\frac{1}{\mu(B_i)} \sum_{\omega \in B_i} X(\omega)\,\mu(\omega)\right) \mu(B_i) = \sum_{\omega \in B_i} X(\omega)\,\mu(\omega)\end{aligned}$$

Η $\mathbb{E}_\mu[X \mid \mathcal{G}]$ είναι μια $\mathcal{G}$-μετρήσιμη τυχαία μεταβλητή. Πως σχετίζεται αυτή με όλες τις πιθανές $\mathcal{G}$-μετρήσιμες τυχαίες μεταβλητές ως προς την προσέγγιση της στην X;

Πρόταση 2.10.2 *Ας πάρουμε ένα χώρο πιθανοτήτων $(\Omega, \mathcal{F}, \mu)$, μία τυχαία μεταβλητή $X : \Omega \to \mathbb{R}$ για την οποία ορίζεται η δεύτερη ροπή και μία $\sigma-$άλγεβρα $\mathcal{G} \subset \mathcal{F}$.*
Τότε υπάρχει μία τυχαία μεταβλητή Y, μετρήσιμη ως προς την $\mathcal{G}$ και η οποία λύνει το πρόβλημα

$$\mathbb{E}_\mu[(X - Y)^2] = \inf_{Z \in m\mathcal{G}} [(X - Z)^2]$$

Η τυχαία μεταβλητή Y είναι η $Y = \mathbb{E}_\mu[X \mid \mathcal{G}]$

Συνεπώς, η δεσμευμένη μέση τιμή μπορεί να θεωρηθεί σαν η καλύτερη εκτιμήτρια της X δεδομένης της πληροφορίας που περιέχεται στην σ-άλγεβρα $\mathcal{G}$.

Η απόδειξη χρησιμοποιεί την ιδιότητα της δεσμευμένης μέσης τιμής σύμφωνα με την οποία

$$\sum_{\omega \in \Omega} (X(\omega) - \mathbb{E}_\mu[X \mid \mathcal{G}](\omega))\,Z(\omega) = 0$$

για κάθε τυχαία μεταβλητή $Z \in m\mathcal{G}$. Η έκφραση αυτή μπόρει να γραφεί και ως

$$\mathbb{E}_\mu[(X - \mathbb{E}[X \mid \mathcal{G}])\,Z] = 0,\ \forall Z \in m\mathcal{G}$$

αρκεί βέβαια η Z να είναι τετραγωνικά ολοκληρώσιμη. Βλέπουμε λοιπόν ότι η $\mathbb{E}_\mu[X \mid \mathcal{G}]$ είναι η **προβολή** της X στον υπόχωρο των τετραγωνικά ολοκληρώσιμων $\mathcal{G}$-μετρήσιμων τυχαίων μεταβλητών.

Παράδειγμα 2.10.1 *Ένας παίκτης παίζει ένα παιχνίδι κορώνα γράμματα στο οποίο ποντάρει 1 εύρω στην κορώνα. Ας υποθέσουμε ότι ο παίκτης έχει παρακολουθήσει το παιχνίδι μέχρι την πρώτη ρίψη και θέλει να βρεί την καλύτερη πρόβλεψη για το κέρδος του στην δεύτερη ρίψη, δεδομένης της πληροφορίας την οποία έχει στην διάθεση του μετά την πραγματοποίηση της πρώτης ρίψης. Ποια θα είναι αυτή;*

Το σύνολο των γεγονότων είναι το $\Omega = \{HH, HT, TH, TT\}$ και η $\sigma-$άλγεβρα η οποία σχετίζεται με αυτό είναι η

$$\mathcal{F}_2 = 2^\Omega = \{\emptyset\Omega, \{HH\}, \{HT\}, \{TH\}, \{TT\}, \cdots\}$$

Ας ονομάσουμε $\mathcal{F}_1$ την σ-άλγεβρα που περίεχει την πληροφορία για την πρώτη ρίψη

$$\mathcal{F}_1 = \{\emptyset, \Omega, \{HH, HT\}, \{TH, TT\}\}$$

Η καλύτερη πρόβλεψη θα είναι μια τυχαία μεταβλητή Y η οποία θα είναι μετρήσιμη ως προς την $\mathcal{F}_1$. Η γενικότερη τέτοια τυχαία μεταβλητή θα είναι της μορφής

$$Y(\omega) = \left\{\begin{array}{ll} a & \omega = \{HH, HT\} \\ b & \omega = \{TH, TT\} \end{array}\right.$$

για $a, b \in \mathbb{R}$. Με άλλα λόγια μπορούμε να ταυτίσουμε την οποιαδήποτε $Y \in m\mathcal{F}_1$ με το διάνυσμα $(a, b) \in \mathbb{R}^2$.

Ποιά θα είναι αυτή η Y η οποία θα ελαχιστοποιεί την $\mathbb{E}[(X-Y)^2]$ όπου X είναι το κέρδος του παίκτη απο το παιχνίδι αυτό;

Η τυχαία μεταβλητή X θα είναι

$$X(\omega) = \begin{cases} 2 & \omega = \{HH\} \\ 0 & \omega = \{HT\} \\ 0 & \omega = \{TH\} \\ -2 & \omega = \{TT\} \end{cases}$$

Κατά συνέπεια

$$\mathbb{E}[(X-Y)^2] = \frac{1}{4}(2-a)^2 + \frac{1}{4}(0-a)^2 + \frac{1}{4}(0-b)^2 + \frac{1}{4}(-2-b)^2$$

Η ελαχιστοποίηση της συνάρτησης

$$\phi(a,b) = \frac{1}{4}(2-a)^2 + \frac{1}{4}a^2 + \frac{1}{4}b^2 + \frac{1}{4}(2+b)^2$$

θα μας καθορίσει την τυχαία μεταβλητή Y. Οι συνθήκες πρώτης τάξης δίνουν

$$\begin{aligned} \frac{\partial \phi}{\partial a} &= -\frac{1}{2}(2-a) + \frac{1}{2}a = 0 \\ \frac{\partial \phi}{\partial b} &= \frac{1}{2}(2+b) + \frac{1}{2}b = 0 \end{aligned}$$

δηλαδή

$$a = 1,\ b = -1$$

2.11 Παραγώγιση μέτρων και πιθανοφάνεια.

Ας θεωρήσουμε και πάλι το άπειρο αλλά αριθμήσιμο σύνολο $\Omega = \{\omega_1, \omega_2, \cdots\}$ και το διακριτό μέτρο που ορίζεται απο την ακολουθία $(p_1, p_2, \cdots)$.

Αν θεωρήσουμε την απεικόνιση $X : \Omega \to \mathbb{R}^+$ η οποία ορίζεται με τον σύνηθη τρόπο $X(\omega_i) = x_i \geq 0$, τότε η απεικονιση ν τέτοια ώστε

$$\nu(A) = \sum_{\omega \in A} X(\omega)\, \mu(\omega) \tag{2.5}$$

είναι και αυτή ένα μέτρο (όχι απαραίτητα μέτρο πιθανότητας). Βέβαια στην περίπτωση που την κανονικοποιήσουμε ως

$$\nu(A) = \frac{1}{\mathbb{E}_\mu[X]} \sum_{\omega \in A} X(\omega)\, \mu(\omega)$$

τότε έχουμε ένα μέτρο πιθανότητας.

Αν υποθέσουμε ότι υπάρχουν κάποια $i \in \mathbb{I} \subset \mathbb{N}$ για τα οποία $p_i = 0$ τότε αν $A = \{\omega_{i_1}, \omega_{i_2}, \cdots\}$, $i_1, i_2, \cdots \in \mathbb{I}$ έχουμε πως $\mu(A) = 0$. Ένα τέτοιο A ονομάζεται μηδενοσύνολο. Φυσικά αν $p_i > 0$ δεν υπάρχει κανένα μηδενοσύνολο εκτός του κενού.

Απο τον ορισμό του καινούργιου μέτρου ν παρατηρούμε ότι αν $\mu(A) = 0$ τότε και $\nu(A) = 0$. Θα δούμε την ιδιότητα αυτή στο μέλλον για γενικότερα μέτρα.

Ορισμός 2.11.1 *Αν για δύο μέτρα μ και ν ισχύει $\nu(A) = 0$ αν $\mu(A) = 0$ θα λέμε ότι το μέτρο ν είναι απόλυτα συνεχές ως προς το μ και θα συμβολίζουμε $\nu << \mu$.*

Ένας ισοδύναμος ορισμός των απόλυτα συνεχών μέτρων που είναι χρήσιμος για ότι ακολουθήσει είναι και ακόλουθος

Ορισμός 2.11.2 *Θα λεμε ότι το μέτρο ν είναι απόλυτα συνεχές ως προς το μ και θα συμβολίζουμε $\nu << \mu$, αν για κάθε $\epsilon > 0$ υπάρχει $\delta > 0$ τέτοιο ώστε για κάθε $A \in \mathcal{F}$, $\nu(A) < \epsilon$ για $\mu(A) < \delta$.*

Τα απόλυτα συνεχή μέτρα μπορεί να χαρακτηριστούν με την βοήθεια μιάς σχέσης σαν την (2.5).

Πρόταση 2.11.1 *Ας υποθέσουμε ότι έχουμε δύο μέτρα ν και μ τέτοια ώστε $0 \le \nu(A) \le \mu(A)$ για κάθε $A \in \mathcal{F}$. Τότε υπάρχει μία μη αρνητική απεικόνιση $h : \Omega \to \mathbb{R}^+$ η οποία είναι μετρήσιμη ως προς την $\mathcal{F}$ τέτοια ώστε $\nu(A) = \sum_{\omega \in A} h(\omega)\, \mu(A)$. Η h ονομάζεται πυκνότητα του μέτρου ν ως προς το μ ή λόγος πιθανοφάνειας.*

Στην απλή αυτή περίπτωση, ουσιαστικά δεν έχουμε τίποτα περίπλοκο να αποδείξουμε. Αν υποθέσουμε ότι η $\mathcal{F}$ παράγεται απο μία πεπερασμένη διαμέριση $\{A_1, A_2, \cdots, A_n\}$ του Ω τότε μπορούμε να κατασκευάσουμε την h ως εξής

$$h(\omega) = \begin{cases} \frac{\nu(A_i)}{\mu(A_i)}, & \omega \in A_i \quad \text{αν}\, \mu(A_i) > 0 \\ 0 & \text{αλλιώς} \end{cases}$$

Η h αυτή είναι η απεικόνιση που θέλουμε εφόσον π.χ. αν $F = \bigcup_{i \in J} A_i$, για κάποιο σύνολο δεικτών, έχουμε ότι

$$\begin{aligned} \nu(A) &= \sum_{i \in J} \nu(A_i) = \sum_{i \in J} h(\omega) \mathbf{1}_{A_i}(\omega) \mu(A_i) \\ &= \sum_{\omega \in A} h(\omega)\, \mu(\omega) \end{aligned}$$

Η παράγωγος Radon-Nikodym σχετίζεται άμεσα με την έννοια της δεσμευμένης μέσης τιμής. Έστω $(\Omega, \mathcal{F}, \mu)$ ένας χώρος πιθανοτήτων και έστω $\mathcal{G}$ μία υποάλγεβρα της $\mathcal{F}$. Για οποιαδήποτε μη αρνητική τυχαία μεταβλητή X η απεικόνιση $\nu(A) = \mathbb{E}[X\, \mathbf{1}_A]$ για κάθε $A \in \mathcal{G}$ ορίζει ένα μέτρο στο $(\Omega, \mathcal{G})$ και ισχύει $\nu(A) = 0$ αν $\mu(A) = 0$. Συνεπώς, υπάρχει η τυχαία μεταβλητή $Y := \frac{d\nu}{d\mu}$ η οποία είναι μετρήσιμη ως προς την $\mathcal{G}$. Επίσης ισχύει ότι για κάθε $B \in \mathcal{G}$ ότι

$$\mathbb{E}[Y\, \mathbf{1}_B] = \nu(B) = \mathbb{E}[Y\, \mathbf{1}_B]$$

συνεπώς η $Y := \frac{d\nu}{d\mu}$ είναι μία εκδοχή της δεσμευμένης μέσης τιμής.

Η κατασκευή αυτή μας επιτρέπει να ορίσουμε την δεσμευμένη μέση τιμή για τυχαίες μεταβλητές οι οποίες δεν είναι απαραίτητα τετραγωνικά ολοκληρώσιμες αλλά μόνο ολοκληρώσιμες.

2.12 Παράρτημα: Αποδείξεις ορισμένων προτάσεων.

2.12.1 Απόδειξη της Πρότασης 2.7.1

Ας υποθέσουμε ότι $p_i > 0$ για κάθε i και ας πάρουμε μια ακολουθία $X^{(n)} \in \ell^1(\mu)$. Είναι προφανές ότι την ακολουθία $X^{(n)}$ θα πρέπει να την κατανοούμε σαν μια διπλή ακολουθία ή σαν μια ακολουθία ακολουθιών

$$X^{(n)} = (x_1^{(n)}, x_2^{(n)}, \cdots) = (x_{1,n}, x_{2,n}, \cdots)$$

Εφόσον η $X^{(n)}$ είναι Cauchy στον $\ell^1(\mu)$, ισχύει ότι για κάθε ϵ

$$\| X^{(n)} - X^{(m)} \|_{\ell^1(\mu)} = \sum_{i=1}^{\infty} \mid x_i^{(n)} - x_i^{(m)} \mid p_i < \epsilon$$

αν $n, m \ge N$.

Παρατηρούμε ότι για κάθε i,

$$p_i \mid x_i^{(n)} - x_i^{(m)} \mid \le \sum_{i=1}^{\infty} \mid x_i^{(n)} - x_i^{(m)} \mid p_i < \epsilon$$

συνεπώς για κάθε $\epsilon' > 0$ υπάρχει N τέτοιο ώστε για την ακολουθία $x_i^{(n)} \in \mathbb{R}$ να ισχύει

$$\mid x_i^{(n)} - x_i^{(m)} \mid < \epsilon'$$

αν $n, m > N$ συνεπώς η πραγματική ακολουθία $x_i^{(n)}$ είναι Cauchy. Επειδή ο $\mathbb{R}$ είναι πλήρης, η ακολουθία $x_i^{(n)}$ συγκλίνει για κάθε i. Συνεπώς, για κάθε i υπάρχει $x_i \in \mathbb{R}$ τέτοιο ώστε $x_i^{(n)} \to x_i$ (με την έννοια της σύγκλισης στο $\mathbb{R}$).

Ορίζουμε την τυχαία μεταβλητή X, ως την τυχαία μεταβλητή για την οποία ισχύει $X(\omega_i) = x_i$. Θα δείξουμε πρώτα ότι $X \in \ell^1(\mu)$.
Η τριγωνική ανισότητα μας εξασφαλίζει ότι

$$\sum_{i=1}^{\infty} p_i \mid x_i^{(n)} \mid = \sum_{i=1}^{\infty} p_i \mid x_i^{(n)} - x_i^{(N)} + x_i^{(N)} \mid \leq \sum_{i=1}^{\infty} p_i \mid x_i^{(n)} - x_i^{(N)} \mid + \sum_{i=1}^{\infty} p_i \mid x_i^{(N)} \mid \leq \epsilon + \sum_{i=1}^{\infty} p_i \mid x_i^{(N)} \mid$$

όπου στην τελευταία ανισότητα χρησιμοποιήσαμε την ιδιότητα Cauchy .
Αν κόψουμε τα άπειρα αθροίσματα μέχρι έναν πεπερασμένο όρο π.χ k έχουμε ότι

$$\sum_{i=1}^{k} p_i \mid x_i^{(n)} \mid < M$$

όπου $M = \epsilon + \sum_{i=1}^{\infty} p_i \mid x_i^{(N)} \mid$. Αυτό που κάναμε είναι ισοδύναμο με το να περιοριστούμε στην ακολουθία $\bar{x}^{(n)} = (x_1^{(n)}, x_2^{(n)}, \cdots, x_k^{(n)}) \in \mathbb{R}^k$. Αυτή είναι μια ακολουθία στον πεπερασμένης διάστασης χώρο $\mathbb{R}^k$ για τον οποίο είναι γνωστό ότι αν $x_i^{(n)} \to x_i$, $i = 1, 2, \cdots, k$ τότε $\bar{x}^{(n)} \to \bar{x}$ στον $\mathbb{R}^k$ όπου $\bar{x} = (x_1, x_2, \cdots, x_k)$. Αυτό δεν ισχύει αν γένει για χώρους άπειρης διάστασης όπως π.χ. ο $\ell^1(\mu)$, για τον λόγο αυτό και αναγκαζόμαστε να πάρουμε την πεπερασμένης διάστασης προσέγγιση της αρχικής ακολουθίας $x^{(n)}$. Αν πάρουμε το όριο $n \to \infty$ παίρνουμε ότι

$$\sum_{i=1}^{k} p_i \mid x_i \mid < M$$

Η ακολουθία $y_k = \sum_{i=1}^{k} p_i \mid x_i \mid$ είναι μονότονη και φραγμένη άρα συγκλίνει και το όριο ικανοποιεί το ίδιο φράγμα, συνεπώς, παίρνοντας $k \to \infty$ έχουμε ότι

$$\sum_{i=1}^{\infty} p_i \mid x_i \mid < M$$

Άρα $X \in \ell^1(\mu)$.
Θα δείξουμε τώρα ότι $\lim_{n\to\infty} X^{(n)} = X$ στον $\ell^1(\mu)$, ή ισοδύναμα ότι για κάθε $\epsilon > 0$ υπάρχει $N > 0$ τέτοιο ώστε $\| X^{(n)} - X \|_{\ell^1(\mu)} < \epsilon$ για $n \geq N$.
Η ιδιότητα Cauchy μας εξασφαλίζει ότι για κάθε k έχουμε ότι

$$\sum_{i=1}^{k} p_i \mid x_i^{(n)} - x_i^{(m)} \mid < \epsilon$$

για $n, m \geq N$. Ας πάρουμε το όριο $m \to \infty$. Αυτό μας δίνει

$$\sum_{i=1}^{k} p_i \mid x_i^{(n)} - x_i \mid < \epsilon$$

Μετά παίρνουμε το όριο $k \to \infty$ οπότε χρησιμοποιώντας τις ιδιότητες των μονότονων και φραγμένων ακολουθιών καταλήγουμε στο ότι

$$\sum_{i=1}^{\infty} p_i \mid x_i^{(n)} - x_i \mid < \epsilon$$

συνεπώς $\lim_{n\to\infty} X^{(n)} = X$ στον $\ell^1(\mu)$.

Κεφάλαιο 3

Το μέτρο *Lebesgue*

3.1 Εισαγωγή

Τα διακριτά μέτρα έχουν ενδιαφέρον αλλά οι δυνατότητες τους για την περιγραφή των τυχαίων πειραμάτων είναι περιορισμένες, στην περιγραφή πειραμάτων για τα οποία ο χώρος πιθανοτήτων είναι διακριτός ή για την περιγραφή και μελέτη τυχαίων μεταβλητών οι οποίες είναι διακριτές. Είναι προφανές ότι όλα τα τυχαία πειράματα και όλες οι τυχαίες μεταβλητές που θα χρειαστεί να μελετήσουμε δεν θα είναι της μορφής αυτής, οπότε είναι απαραίτητο να γενικεύσουμε την έννοια του μέτρου εγκαταλείποντας τον περιορισμό των διακριτών χώρων πιθανοτήτων. Αυτό θα μας οδηγήσει σε μια θεωρία για το μέτρο η οποία θα είναι ένα χρήσιμο εργαλείο για να μπορέσουμε να καταλάβουμε την θεωρία πιθανοτήτων στο πιο γενικό της πλαίσιο, όπου οι τυχαίες μεταβλητές θα μπορεί πλέον να είναι συνεχείς τυχαίες μεταβλητές, που παίρνουν τιμές στο $\mathbb{R}$, το $\mathbb{R}^n$ ή ακόμα και σε σύνολα με πιο περίπλοκες δομές, π.χ. γενικότεροι μετρικοί χώροι.

Στο κεφάλαιο αυτό θα προσπαθήσουμε να γενικεύσουμε αυτά που είδαμε σχετικά με τα διακριτά μέτρα στην περίπτωση όπου το συνολο Ω δεν είναι πλέον ένα αριθμήσιμο σύνολο αλλά είναι το $\mathbb{R}$ ή κάποιο υποσύνολο του π.χ. ένα διάστημα. Στην προσπάθεια να κάνουμε αυτό θα δούμε ότι πλέον, σε αντίθεση με το ότι συμβαίνει με τα διακριτά μέτρα θα μπορούμε να έχουμε μη κενά υποσύνολα $A \subset \Omega$, για τα οποία να ισχύει $\mu(A) = 0$.

Θα παρουσιάσουμε την κατασκευή του απλούστερου δυνατού μέτρου στο $\mathbb{R}$, το οποίο είναι το μέτρο Lebesgue.

3.2 Μηδενοσύνολα

Η έννοια του μέτρου σχετίζεται με την έννοια του μεγέθους ενός συνόλου. Αν το σύνολο αυτό είναι ένα υποσύνολο της ευθείας των πραγματικών αριθμών, μία λογική έννοια του μεγέθους είναι αυτή του μήκους.

Ορισμός 3.2.1 *Αν I είναι ένα φραγμένο διάστημα $I = [a, b]$, $I = (a, b]$, $I = [a, b)$, $I = (a, b)$ τότε ορίζουμε σαν μήκος του I το $\ell(I) = b - a$. Αν $I = \{a\} = [a, a]$ τότε $\ell(I) = 0$.*

Σχόλιο 3.2.1 *Το μήκος ℓ μπορούμε να το δούμε σαν μία απεικόνιση απο το σύνολο των διαστημάτων στο σύνολο των πραγματικών αριθμών. Παρατηρείστε ότι η απεικόνιση αυτή δεν ξεχωρίζει μεταξύ των ανοικτών και των κλειστών διαστημάτων ως προς τη εικόνα. Για τον λόγο αυτό στα παρακάτω θα δεν θα παίζει ρόλο η διακριση μεταξύ ανοιχτών και κλειστών διαστημάτων.*

Είδαμε ότι το μήκος ℓ ενός σημείο είναι 0. Ποια άλλα διαστήματα θα έχουν μήκος μηδέν; Η γενίκευση αυτή θα μας δώσει την έννοια του μηδενοσύνολου.

Ορισμός 3.2.2 *Έστω ένα σύνολο $A \subset \mathbb{R}$ το οποίο μπορεί να καλυφθεί απο μία συλογή διαστημάτων $\{I_n \mid n \geq 1\}$ δηλαδή*

$$A \subseteq \bigcup_{n=1}^{\infty} I_n$$

Αν ισχύει ότι για κάθε $\epsilon > 0$ έχουμε

$$\sum_{n=1}^{\infty} \ell(I_n) < \epsilon$$

τότε λέμε ότι το σύνολο A είναι **μηδενοσύνολο**.

Παράδειγμα 3.2.1 *Οποιοδήποτε αριθμήσιμο σύνολο $A = \{x_1, x_2, \cdots\}$ είναι μηδενοσύνολο.*
Ας προσπαθήσουμε να δούμε αν ο παραπάνω ορισμός μπορεί να εφαρμοστεί. Μπορούμε να καλύψουμε το A χρησιμοποιώντας τα $I_n = (x_n - \epsilon\frac{1}{2^{n+1}}, x_n + \epsilon\frac{1}{2^{n+1}})$. Ειναι προφανές ότι $A \subseteq \bigcup_{n=1}^{\infty} I_n$. Συνεπώς, γύρω απο κάθε σημείο x_i που περιέχεται στο σύνολο A, τοποθετούμε ένα ανοιχτό διάστημα μήκους $\ell(x_i) = \frac{\epsilon}{2^i}$. Παρατηρούμε ότι το μήκος μικραίνει καθώς πάμε στα σημεία του συνόλου με μεγαλύτερο δείκτη. Έχουμε ότι

$$\sum_{n=1}^{\infty} \ell(I_n) = \epsilon \sum_{n=1}^{\infty} \frac{1}{2^n} = \epsilon$$

Συνεπώς το σύνολο αυτό είναι μηδενοσύνολο. Φυσικά το n μπορούσε να ήταν και πεπερασμένο.

Πρόταση 3.2.1 *Η ένωση μηδενοσυνόλων παραμένει μηδενοσύνολο, δηλαδή αν A_i, $i \in \mathbb{N}$ είναι μηδενοσύνολο, τότε $A = \bigcup_{i=1}^{\infty} A_i$ είναι επίσης μηδενοσύνολο.*

Απόδειξη: Βλ. Παράρτημα, Παράγραφος 3.11.1. □

Τα μηδενοσύνολα, δεν είναι απαραίτητο να είναι αριθμήσιμα σύνολα.

Παράδειγμα 3.2.2 *Θεωρείστε το σύνολο του Cantor το οποίο κατασκευάζεται ως εξής. Απο το $[0, 1]$ αφού το χωρίσετε σε 3 ίσα μέρη, αφαιρέστε το μεσαίο. Αυτό που απομένει, ας το ονομάσουμε C_1 δηλαδή $C_1 = [0, 1/3] \cup [2/3, 1]$. Στην δεύτερη επανάληψη, απο τα διαστήματα που έχουν απομείνει, αφαιρούμε το μεσαίο τρίτο, δηλαδή παίρνουμε το $C_2 = [0, 1/9] \cup [2/9.1/3] \cup [2/3, 7/9] \cup [8/9, 1]$. Συνεχίζουμε με τον ίδιο τρόπο.*
Ας ορίσουμε ως $C = \bigcap_{n=1}^{\infty} C_n$, το σύνολο του Cantor. Μπορεί κανείς να δείξει ότι το C είναι **μη αριθμήσιμο** *αλλά* **μηδενοσύνολο**.

3.3 Το εξωτερικό μέτρο *Lebesgue*.

Θα προσπαθήσουμε τώρα να γενικεύσουμε την έννοια του μήκους για υποσύνολα του $\mathbb{R}$ που είναι γενικότερης μορφής απο τα διαστήματα. Φυσικά, στην προσπάθεια μας θα ξεκινήσουμε απο το γεγονός ότι μπορούμε να καλύψουμε οποιοδήποτε υποσύνολο του $\mathbb{R}$ απο μία συλλογή ανοιχτών διαστημάτων.

Ορισμός 3.3.1 *Το εξωτερικό μέτρο Lebesgue ενός συνόλου $A \subset \mathbb{R}$, $\mu^*(A)$, είναι το μέγιστο κάτω φράγμα (inf) των μηκών όλων των πιθανών καλυμάτων του A. Δηλαδή αν*

$$Z_A = \{\sum_{n=1}^{\infty} \ell(I_n) \ \mid \ A \subseteq \bigcup_{n=1}^{\infty} I_n\}$$

όπου I_n διαστήματα, έχουμε ότι

$$\mu^*(A) = \inf Z_A$$

Παράδειγμα 3.3.1 *Βρείτε το εξωτερικό μέτρο Lebesgue του συνόλου $A = \{x\}$ για οποιοδήποτε $x \in \mathbb{R}$.*
Για οποιοδήποτε $\epsilon > 0$ μπορούμε να βρούμε ένα κάλυμα του A που αποτελείται απο μόνο ένα διάστημα της μορφής $I = (x - \frac{\epsilon}{2}, x + \frac{\epsilon}{2})$. Το μήκος του καλύματος αυτού ικανοποιεί το $\ell(I) = \epsilon$ το οποίο μπορεί να γίνει οσοδήποτε μικρό θέλουμε, αλλά πρέπει πάντοτε να είναι θετικό. Συνεπώς το $\inf Z_A = 0$, άρα $\mu^*(A) = 0$.

Το εξωτερικό μέτρο Lebesgue έχει ορισμένες πολύ χρήσιμες ιδιότητες

Πρόταση 3.3.1 *Για το εξωτερικό μέτρο Lebesgue μ^* ισχύουν τα ακόλουθα:*

1. *Αν A μηδενοσύνολο τότε $\mu^*(A) = 0$ και αντίστροφα.*
2. *Αν $A \subset B$ τότε $\mu^*(A) \leq \mu^*(B)$, δηλαδή η απεικόνιση μ^* είναι μονότονη.*
3. *Αν I ένα διάστημα του $\mathbb{R}$, τότε $\mu^*(I) = \ell(I)$.*

4. Αν E_n μία ακολουθία υποσυνόλων του $\mathbb{R}$, $E_n \subset \mathbb{R}$, τότε

$$\mu^*\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} \mu^*(E_n)$$

Απόδειξη: Βλ. Παράρτημα , Παράγραφος 3.11.2 □

Η τελευταία ιδιότητα, της υποαθροιστικότητας είναι πολύ σημαντική.

Ας υποθέσουμε ότι τα E_n είναι ξένα μεταξύ τους δηλαδή ότι $E_i \cap E_j = \emptyset$ για κάθε i, j. Στην περίπτωση αυτή αν το $\mu^*(A)$ το καταλάβουμε σαν το 'μήκοσ' του A, θα πρέπει το μήκος του A να ισούται με το άθροισμα των μηκών του κάθε E_n, δηλαδή η ανισότητα θα έπρεπε να γίνει ισότητα. Δυστυχώς, το εξωτερικό μέτρο Lebesgue **δεν** ικανοποιεί την ιδιότητα αυτή. Θα πρέπει λοιπόν να ραφινάρουμε λιγάκι τον ορισμό μας, και για να γίνει αυτό θα πρέπει να πάρουμε το εξωτερικό μέτρο και να περιοριστούμε μόνο σε υποσύνολα του $\mathbb{R}$ τα οποία έχουν συγκεκριμένες ιδιότητες.

Μπορεί κανείς να κατασκευάσει υποσύνολα $A_i \subset \mathbb{R}$ για τα οποία να ισχύει $A_i \cap A_j = \emptyset$ για κάθε $i, j \in \mathbb{N}$ αλλά

$$\mu^*\left(\bigcup_{i=1}^{\infty} A_i\right) < \sum_{i=1}^{\infty} \mu^*(A_i)$$

δηλαδή να κατασκευάσει σύνολα που το μήκος τους είναι μικρότερο απο το συνολικό άθροισμα των μηκών των (ξένων μεταξύ τους) υποσυνόλων που τα απαρτίζουν. Τέτοια σύνολα δεν είναι παρά πολύ εύκολο να κατασκευαστούν αλλά υπάρχουν. Η ακόλουθη κατασκευή ενός τέτοιου συνόλου οφείλεται στον Vitali.

Παράδειγμα 3.3.2 Τα σύνολα του Vitali
Ας πάρουμε δύο πραγματικούς αριθμούς x, y τους οποίους ταυτίζουμε αν $x - y$ είναι ένας ρητός αριθμός. Θα συμβολίζουμε $x \sim y$ αν $x - y = r$ όπου $r \in \mathbb{Q}$. Κατασκευάζουμε λοιπόν με τον τρόπο αυτό μια κλάση ισοδύναμίας $[x] = \{y \; : \; x \sim y\}$. Απο κάθε τέτοια κλάση ισοδυναμίας επιλέγουμε έναν αριθμό ο οποίος να βρίσκεται στο διάστημα $[0,1]$. Για παράδειγμα ο αριθμός $\sqrt{2}$ και ο αριθμός $\sqrt{2} - 1$ ανήκουν στην ίδια κλάση ισοδυναμίας, την $[\sqrt{2}]$ και ο $\sqrt{2} - 1 \in [0,1]$. Το σύνολο που φτιάχνουμε με τον τρόπο αυτό ονομάζεται σύνολο του Vitali και θα το συμβολίζουμε V.

Αν τώρα για κάθε ρητό αριθμό $r_k \in [-1,1]$ κατασκευάσουμε το σύνολο $V_k = V + r_k = \{v + r_k, \; v \in V\}$ τότε μπορούμε να δούμε ότι $V_i \cap V_j = \emptyset$, $i, j \in \mathbb{N}$ αλλά

$$\mu^*\left(\bigcup_{i=1}^{\infty} V_i\right) < \sum_{i=1}^{\infty} \mu^*(V_i)$$

Ας υποθέσουμε ότι ισχύει η ισότητα αυτή και θα δούμε ότι σύντομα θα οδηγηθούμε σε άτοπο.

Μπορούμε εύκολα να δούμε ότι

$$(0,1) \subseteq \bigcup_{k \in \mathbb{Q} \cap (-1,1)} V_k \subseteq (-1,2)$$

Απο τις ιδιότητες του εξωτερικού μέτρου Lebesgue

$$\mu^*((0,1)) \leq \mu^*\left(\bigcup_{k \in \mathbb{Q} \cap (-1,1)} V_k\right) \leq \mu^*((-1,2))$$

άρα και

$$1 \leq \mu^*\left(\bigcup_{k \in \mathbb{Q} \cap (-1,1)} V_k\right) \leq 3 \tag{3.1}$$

Απο την υποαθροιστικότητα του εξωτερικού μέτρου έχουμε ότι

$$\mu^*\left(\bigcup_{k \in \mathbb{Q} \cap (-1,1)} V_k\right) \leq \sum_{k \in \mathbb{Q} \cap (-1,1)} \mu^*(V_k) = \sum_{k \in \mathbb{Q} \cap (-1,1)} \mu^*(V)$$

όπου στην τελευταία ισότητα χρησιμοποιήσαμε την ιδιότητα του εξωτερικού μέτρου Lebesgue να είναι αμετάβλητο ως προς τις μεταφορές (translation invariance) που προκύπτει με την σειρά της απο την ιδιότητα του μήκους ενός διαστήματος να είναι αμετάβλητο ως προς τις μεταφορές. Εφόσον

$$1 \leq \sum_{k \in \mathbb{Q} \cap (-1,1)} \mu^*(V)$$

καταλήγουμε στο ότι $\mu^*(V) > 0$. Γνωρίζουμε ότι στο διάστημα $(-1,1)$ έχουμε άπειρους το πλήθος ρητούς αριθμούς. Αν προσθέσουμε άπειρές φορές τον θετικό αριθμό $\mu^*(V)$ θα καταλήξουμε στο ότι

$$\sum_{k \in \mathbb{Q} \cap (-1,1)} \mu^*(V) = \infty.$$

Ας υποθέσουμε τώρα ότι

$$\mu^* \left(\bigcup_{k \in \mathbb{Q} \cap (-1,1)} V_k \right) = \sum_{k \in \mathbb{Q} \cap (-1,1)} \mu^*(V_k)$$

Τότε η (3.1) μας δίνει ότι

$$\infty = \sum_{k \in \mathbb{Q} \cap (-1,1)} \mu^*(V_k) = \mu^* \left(\bigcup_{k \in \mathbb{Q} \cap (-1,1)} V_k \right) \leq 3$$

το οποίο είναι προφανώς άτοπο.

Συνεπώς κατασκευάσαμε με τον τρόπο αυτό ένα παράδειγμα συνόλων για τα οποία ισχύει ότι το εξωτερικό μέτρο Lebesgue δεν είναι αριθμήσιμα αθροίσιμο.

3.4 Μετρήσιμα σύνολα.

Για να έχει το εξωτερικό μέτρο Lebesgue την ιδιότητα της αριθμήσιμης αθροισιμότητας, θα πρέπει να το περιορίσουμε σε μία ειδική κλάση απο υποσύνολα του $\mathbb{R}$. Τα υποσύνολα αυτά θα τα ονομάσουμε **μετρήσιμα** υποσύνολα του $\mathbb{R}$ (ως προς το μέτρο Lebesgue).

Ορισμός 3.4.1 *Ένα υποσύνολο $E \subset \mathbb{R}$ ονομάζεται* **μετρήσιμο** *ως προς το εξωτερικό μέτρο Lebesgue αν για κάθε $A \subset \mathbb{R}$ ισχύει ότι*

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$$

Το σύνολο A ονομάζεται **σύνολο δοκιμής***. Το σύνολο των μετρήσιμων υποσυνόλων του $\mathbb{R}$ συμβολίζεται $\mathcal{M}$.*

Παράδειγμα 3.4.1 *Η συνθήκη του ορισμού 3.4.1 μπορεί να γραφεί ισοδύναμα*

$$\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \cap E^c)$$

Πράγματι, εφόσον για κάθε $A, E \subset \mathbb{R}$ ισχύει ότι $A = (A \cap E) \cup (A \cap E^c)$ λόγω της υποαθροιστικότητας του εξωτερικού μέτρου έχουμε ότι

$$\mu^*(A) = \mu^*(\,(A \cap E) \cup (A \cap E^c)\,) \leq \mu^*(A \cap E) + \mu^*(A \cap E^c)$$

Αν λοιπόν ισχύει και

$$\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \cap E^c)$$

τότε οι δύο αυτές ανισότητες εξασφαλίζουν την ισότητα του ορισμού 3.4.1.

Πολλές φορές θα χρησιμοποιούμε την ισοδύναμη αυτή συνθήκη για να αποδείξουμε ότι κάποια σύνολα είναι μετρήσιμα ως προς το μ^*.

Λήμμα 3.4.1 *Η ιδιότητα της μετρησιμότητας μας εξασφαλίζει αυτό που θέλουμε, για πεπερασμένο αριθμό συνόλων, δηλαδή γιατί αν* $E_i \in \mathcal{M}$ *και* $E_i \cap E_j = \emptyset$ *για κάθε* i, j *τότε* $\mu^*(\bigcup_{i=1}^m E_i) = \sum_{i=1}^m \mu^*(E_i)$ *για κάθε* $m \in \mathbb{N}$.

Απόδειξη: Ας πάρουμε πρώτα την περίπτωση δύο συνόλων $E_1, E_2 \in \mathcal{M}$ και οποιοδήποτε $B \subset \mathbb{R}$. Αν επιλεξουμε ως $A = B \cap (E_1 \cup E_2)$ και εφαρμόσουμε τον ορισμό της μετρησιμότητας για το E_2 καταλήγουμε ότι

$$\begin{aligned} \mu^*(B \cap (E_1 \cup E_2)) &= \mu^*(B \cap (E_1 \cup E_2) \cap E_2) + \mu^*(B \cap (E_1 \cup E_2) \cap (E_2)^c) \\ &= \mu^*(B \cap E_1) + \mu^*(B \cap E_2). \end{aligned}$$

Ας πάρουμε τώρα την περίπτωση τριών συνόλων $E_1, E_2, E_3 \in \mathcal{M}$ και οποιοδήποτε $B \subset \mathbb{R}$. Αν επιλεξουμε ως $A = B \cap (E_1 \cup E_2 \cup E_3)$ και εφαρμόσουμε τον ορισμό της μετρησιμότητας για το E_3 καταλήγουμε ότι

$$\begin{aligned} \mu^*(B \cap (E_1 \cup E_2 \cup E_3)) &= \mu^*(B \cap (E_1 \cup E_2 \cup E_3) \cap E_3) + \mu^*(B \cap (E_1 \cup E_2 \cup E_3) \cap (E_3)^c) \\ &= \mu^*(B \cap E_3) + \mu^*(B \cap (E_1 \cup E_2)) \\ &= \mu^*(B \cap E_1) + \mu^*(B \cap E_2) + \mu^*(B \cap E_3) \end{aligned}$$

Κάνοντας χρήση της μεθόδου της επαγωγής μπορούμε να δείξουμε οτι αν $F_m = \bigcup_{k=1}^m E_k$ τότε

$$\mu^*(B \cap F_m) = \sum_{k=1}^m \mu^*(E_k)$$

Αυτό ολοκληρώνει και την απόδειξη. ■

Τι γίνεται με τις άπειρες ενώσεις;

Πρόταση 3.4.1 *Αν* $E_i \in \mathcal{M}$, E_i *ξένα μεταξύ τους, τότε*

$$\mu^*\left(\bigcup_{i=1}^{\infty}\right) = \sum_{i=1}^{\infty} \mu^*(E_i).$$

Απόδειξη: Για κάθε $m \in \mathbb{N}$ ισχύει ότι

$$\sum_{i=1}^m \mu^*(E_i) = \mu^*\left(\bigcup_{i=1}^m E_i\right) \tag{3.2}$$

Επειδή $\bigcup_{i=1}^m E_i \subset \bigcup_{i=1}^\infty E_i$ απο την μονοτονία του εξωτερικού μέτρου Lebesgue έχουμε ότι

$$\mu^*\left(\bigcup_{i=1}^m E_i\right) \leq \mu^*\left(\bigcup_{i=1}^\infty E_i\right)$$

και αντικαθιστώντας στην (3.2) παίρνουμε

$$\sum_{i=1}^m \mu^*(E_i) \leq \mu^*\left(\bigcup_{i=1}^\infty E_i\right) \tag{3.3}$$

Παίρνουμε τώρα το όριο $m \to \infty$ στην (3.3) και καταλήγουμε στο ότι

$$\sum_{i=1}^\infty \mu^*(E_i) \leq \mu^*\left(\bigcup_{i=1}^\infty E_i\right) \tag{3.4}$$

Η υποαθροιστικότητα του εξωτερικού μέτρου Lebesgue μας εξασφαλίζει ότι

$$\mu^*\left(\bigcup_{i=1}^\infty E_i\right) \leq \sum_{i=1}^\infty \mu^*(E_i) \tag{3.5}$$

Οι (3.4) και (3.5) μας εξασφαλίζουν την ισότητα που θέλουμε να αποδείξουμε. ■

Ας δούμε τώρα τι είδους σύνολα μπορεί να ικανοποιούν τις ιδιότητες που ζητάμε απο τα μετρήσιμα σύνολα.

Παράδειγμα 3.4.2 *Δείξτε ότι ένα μηδενοσύνολο E είναι μετρήσιμο ως προς το μ^*.*
Ισχύει ότι $A \cap E \subseteq E$ για οποιοδήποτε $A \in \mathbb{R}$, συνεπώς $\mu^*(A \cap E) \leq \mu^*(E) = 0$ εφόσον το E είναι μηδενοσύνολο.
Επίσης, $A \cap E^c \subseteq A$ συνεπώς $\mu^*(A) \cap E^c) \leq \mu^*(A)$.
Αν αθροίσουμε τις ανισότητες αυτές βρίσκουμε ότι $\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \cap E^*)$ οπότε απο το Παράδειγμα 3.4.1 καταλήγουμε ότι το $E \in \mathcal{M}$.

Παράδειγμα 3.4.3 *Δείξτε ότι ένα διάστημα $I = [a, b]$ είναι μετρήσιμο σύνολο ως προς το μ^*.*
Ας θεωρήσουμε οποιοδήποτε $A \subset \mathbb{R}$ και ας πάρουμε ένα κάλυμα του A, $\{I_n\}$ τέτοιο ώστε $A \subseteq \bigcup_{n=1}^{\infty} I_n$. Απο τον ορισμό του μ^*, ως το μέγιστο κάτω φράγμα του συνολικού μήκους όλων των καλύμάτων, έχουμε ότι για κάθε $\epsilon > 0$ ισχύει

$$\mu^*(A) \leq \sum_{n=1}^{\infty} \ell(I_n) \leq \mu^*(A) + \epsilon$$

Ας πάρουμε τώρα το σύνολο $A \cap I$. Μπορούμε να δούμε ότι το $\{I_n \cap I\}$ είναι ένα κάλυμα του $A \cap I$. Απο τον ορισμό του μ^* έχουμε ότι

$$\mu^*(A \cap I) \leq \sum_{n=1}^{\infty} \ell(I_n \cap I)$$

Τέλος, τα διαστήματα $\{I_n \cap (-\infty, a)\}$ και $\{I_n \cap (b, \infty)\}$ αποτελούν ένα κάλυμα του $A \cap I^c$. Απο τον ορισμό του μ^* έχουμε ότι

$$\mu^*(A \cap I^c) \leq \sum_{n=1}^{\infty} \ell(I_n \cap (-\infty, a)) + \sum_{n=1}^{\infty} \ell(I_n \cap (b, \infty))$$

Οι ανισότητες αυτές μας εξασφαλίζουν ότι

$$\mu^*(A) \geq \mu^*(A \cap I) + \mu^*(A \cap I^c)$$

άρα και την μετρησιμότητα του I.

Ας δούμε τώρα τις ιδιότητες των μετρήσιμων συνόλων ως προς τις ενώσεις, τις τομές και τα συμπληρώματα.

Παράδειγμα 3.4.4 *Δείξτε ότι αν $E_1, E_2 \in \mathcal{M}$ τότε $E_1 \cup E_2 \in \mathcal{M}$.*

Χωρίς βλάβη της γενικότητας ας θεωρήσουμε ότι $E_1 \cap E_2 = \emptyset$.
Εφόσον το $E_1 \in \mathcal{M}$ για οποιοδήποτε $A \subset \mathbb{R}$ ισχύει ότι

$$\mu^*(A) = \mu^*(A \cap E_1) + \mu^*(A \cap E_1^c) \tag{3.6}$$

Ας πάρουμε τώρα το $A \cap E_1^c$ στην θέση του A και ας εφαρμόσουμε την ιδιότητα της μετρησιμότητας για το σύνολο E_2,

$$\mu^*(A \cap E_1^c) = \mu^*((A \cap E_1^c) \cap E_2) + \mu^*(A \cap E_1^c) \cap E_2^c) \tag{3.7}$$

Παρατηρείστε ότι $A \cap E_1 = A \cap (E_1 \cup E_2) \cap E_1$. Αντικαθιστούμε αυτό στην (3.6) και παίρνουμε

$$\mu^*(A) = \mu^*(A \cap (E_1 \cup E_2) \cap E_1) + \mu^*(A \cap E_1^c) \tag{3.8}$$

Παρατηρείστε επίσης ότι $A \cap (E_1^c \cap E_2) = A \cap (E_1 \cup E_2) \cap E_1^c$. Αντικαθιστούμε αυτό στην (3.7) και παίρνουμε

$$\mu^*(A \cap E_1^c) = \mu^*(A \cap (E_1 \cup E_2) \cap E_1^c) + \mu^*((A \cap E_1^c) \cap E_2^c) \tag{3.9}$$

Οι σχέσεις αυτές μας δίνουν

$$\mu^*(A) = \mu^*(A \cap (E_1 \cup E_2) \cap E_1) + \mu^*(A \cap (E_1 \cup E_2) \cap E_1^c) + \mu^*(A \cap E_1^c) \cap E_2^c) \tag{3.10}$$

Η μετρήσιμότητα του E_1 αν πάρουμε στην θέση του A το $A \cap (E_1 \cup E_2)$ μας δίνει ότι

$$\mu^*(A \cap (E_1 \cup E_2)) = \mu^*(A \cap (E_1 \cup E_2) \cap E_1) + \mu^*(A \cap (E_1 \cup E_2) \cap E_1^c)$$

το οποίο αν αντικατασταθεί στην (3.10) δίνει

$$\begin{aligned}\mu^*(A) &= \mu^*(A\cap(E_1\cup E_2)) + \mu^*(A\cap E_1^c)\cap E_2^c) \\ &= \mu^*(A\cap(E_1\cup E_2)) + \mu^*(A\cap(E_1\cup E_2)^c)\end{aligned}$$

όπου χρησιμοποιήσαμε τον νόμο του De Morgan. Η τελευταία όμως σχέση μας δείχνει οτι το $E_1\cup E_2\in\mathcal{M}$*. Κάνοντας χρήση της επαγωγής μπορούμε να δείξουμε ότι η πεπερασμένη ένωση στοιχείων του* $\mathcal{M}$ *ανήκει στο* $\mathcal{M}$*.*

Παράδειγμα 3.4.5 *Δείξτε ότι αν* $E_1, E_2\in\mathcal{M}$ *τότε και* $E_1\cap E_2\in\mathcal{M}$*.*
Αν $E_1, E_2\in\mathcal{M}$ *τότε και* $E_1^c, E_2^c\in\mathcal{M}$ *όπως εύκολα φαίνεται από τον ορισμό. Από το Παράδειγμα 3.4.4* $E_1^c\cup E_2^c\in\mathcal{M}$ *οπότε απο τον κανόνα του De Morgan* $(E_1\cap E_2)^c\in\mathcal{M}$ *και συνεπώς* $E_1\cap E_2\in\mathcal{M}$*.*
Κάνοντας χρήση της επαγωγής μπορούμε να δείξουμε ότι η πεπερασμένη τομή στοιχείων του $\mathcal{M}$ *ανήκει στο* $\mathcal{M}$*.*

Παράδειγμα 3.4.6 *Δείξτε ότι αν τα σύνολα* $E_i\in\mathcal{M}$ *τότε και η ένωση τους* $\bigcup_{i=1}^{\infty}E_i\in\mathcal{M}$*.*

Στο παράδειγμα 3.4.4 δείξαμε τον ισχυρισμό αυτό για την περίπτωση των πεπερασμένων ενώσεων δηλαδή για ενώσεις την μορφής $\bigcup_{i=1}^{m}E_i$, $m\in\mathbb{N}$ *κάνοντας χρήση της μεθόδου της επαγωγής. Θα δείξουμε τώρα ότι η ιδιότητα αυτή ισχύει και για την ένωση άπειρων συνόλων. Η μέθοδος που θα χρησιμοποιήσουμε αποτελεί πολύ συχνή τεχνική για την απόδειξη αποτελεσμάτων στην θεωρία μέτρου.*
Χωρίς βλάβη της γενικότητας μπορούμε να θεωρήσουμε ότι τα E_i *είναι ξένα μεταξύ τους. Αν όχι μπορούμε να ορίσουμε την ακολουθία συνόλων* $B_i = E_i\cap E_1^c\cap\cdots\cap E_{i-1}^c$*. Παρατηρούμε ότι η ακολουθία* B_i *αποτελείται απο σύνολα που είναι ξένα μεταξύ τους και επιπλέον* $\bigcup E_i = \bigcup B_i$*. Συνεπώς αν αποδείξουμε τον ισχυρισμό για σύνολα ξένα μεταξύ τους, χρησιμοποιώντας την παρατήρηση αυτή μπορούμε να δείξουμε ότι ισχύει γενικά.*
Ας θέσουμε $F_m=\bigcup_{i=1}^{m}E_i$*. Εφόσον* F_m *μετρήσιμο ισχύει ότι*

$$\begin{aligned}\mu^*(A) &= \mu^*(A\cap F_m) + \mu^*(A\cap F_m^c) \\ &= \sum_{i=1}^{m}\mu^*(A\cap E_i) + \mu^*(A\cap F_m^c)\end{aligned} \tag{3.11}$$

όπου χρησιμοποιήσαμε την αθροιστικότητα του εξωτερικού μέτρου Lebesgue για πεπερασμένες ενώσεις συνόλων $E_i\in\mathcal{M}$ *τα οποία είναι ξένα μεταξύ τους.*
Παρατηρούμε ότι $F_m=\bigcup_{i=1}^{m}E_i\subset\bigcup_{i=1}^{\infty}E_i$ *συνεπώς για τα συμπληρωματικά σύνολα ισχύει*

$$\left(\bigcup_{i=1}^{\infty}E_i\right)^c\subset F_m^c$$

και απο την μονοτονία του εξωτερικού μέτρου Lebesgue έχουμε ότι

$$\mu^*\left(\left(\bigcup_{i=1}^{\infty}E_i\right)^c\right)\le\mu^*(F_m^c).$$

Αντικαθιστώντας στην (3.11) παίρνουμε την ανισότητα

$$\mu^*(A)\ge\sum_{i=1}^{m}\mu^*(A\cap E_i)+\mu^*\left(A\cap\left(\bigcup_{i=1}^{\infty}E_i\right)^c\right) \tag{3.12}$$

Τώρα παίρνουμε το όριο $m\to\infty$ *στην (3.12), η οποία δίνει*

$$\mu^*(A)\ge\sum_{i=1}^{\infty}\mu^*(A\cap E_i)+\mu^*\left(A\cap\left(\bigcup_{i=1}^{\infty}E_i\right)^c\right) \tag{3.13}$$

(προσπαθήστε να εξηγήσετε γιατί;).
Όμως, γνωρίζουμε ότι το εξωτερικό μέτρο Lebesgue είναι υποαθροιστική συνάρτηση, συνεπώς ισχύει ότι

$$\mu^*\left(A\cap\left(\bigcup_{i=1}^{\infty}E_i\right)\right)\le\sum_{i=1}^{\infty}\mu^*(A\cap E_i)$$

και αντικαθιστώντας στην (3.13) παίρνουμε την ανισότητα

$$\mu^*(A) \geq \mu^*\left(A \cap \left(\bigcup_{i=1}^{\infty} E_i\right)\right) + \mu^*\left(A \cap \left(\bigcup_{i=1}^{\infty} E_i\right)^c\right)$$

Η ανισότητα αυτή μας εξασφαλίζει ότι $\bigcup_{i=1}^{\infty} E_i \in \mathcal{M}$.

Τα μετρήσιμα σύνολα έχουν τις παρακάτω πολύ ενδιαφέρουσες ιδιότητες.

Πρόταση 3.4.2 *Για το σύνολο $\mathcal{M}$ των μετρήσιμων συνόλων ισχύουν ότι*

1. *Το $\mathbb{R}$ και το $\emptyset$ είναι μετρήσιμα σύνολα.*
2. *Αν υποθέσουμε ότι το E είναι μετρήσιμο σύνολο τότε και το E^c είναι επίσης μετρήσιμο.*
3. *Αν υποθέσουμε ότι τα σύνολα E_i, $i \in \mathbb{N}$ είναι μετρήσιμα, τότε και το $\bigcup_{i=1}^{\infty} E_i$ είναι επίσης μετρήσιμο.*

Απόδειξη: Βλ. Παράρτημα, Παράγραφος 3.11.3. □

Ορισμός 3.4.2 *Μία οικογένεια συνόλων που ικανοποιεί τις ιδιότητες 1-3 της πρότασης 3.4.2 ονομάζεται σ−άλγεβρα.*

Οι ιδιότητες αυτές θα χρησιμοποιηθούν σαν αξιώματα όταν θελήσουμε να γενικεύσουμε την έννοια του μέτρου σε πιο γενικά σύνολα απο το $\mathbb{R}$.

3.5 Το μέτρο *Lebesgue*.

Ορισμός 3.5.1 *Ας υποθέσουμε ότι $E \subset \mathbb{R}$ είναι ένα μετρήσιμο σύνολο. Η απεικόνιση $\mu : \mathcal{F} \to \mathbb{R}$ που ορίζεται σαν $\mu(E) = \mu^*(E)$ για κάθε E μετρήσιμο, ονομάζεται* **μέτρο Lebesgue**. *Κατά συνέπεια, το μέτρο Lebesgue δεν είναι τίποτε άλλο παρά το εξωτερικό μέτρο Lebesgue περιορισμένο στα μετρήσιμα σύνολα.*

Το μέτρο Lebesgue έχει τις εξής χρήσιμες ιδιότητες:

Πρόταση 3.5.1 *Αν μ είναι το μέτρο Lebesgue ισχύει ότι*

1. *Έστω ότι $A \subset B$. Τότε, $\mu(A) \leq \mu(B)$. Αν $\mu(A) < \infty$ τότε $\mu(B \setminus A) = \mu(B) - \mu(A)$*
2. *Έστω ότι E_n μετρήσιμα για κάθε n και ότι $E_n \subset E_{n+1}$ για κάθε n. Τότε*

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mu(E_n)$$

3. *Έστω ότι E_n μετρήσιμα για κάθε n και ότι $E_{n+1} \subset E_n$ για κάθε n. Τότε*

$$\mu\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mu(E_n)$$

   εφόσον $\mu(E_1) < \infty$.

Απόδειξη: Βλ. Παράρτημα, Παράγραφο 3.11.4. □

Παράδειγμα 3.5.1 *Δείξτε ότι το μέτρο Lebesgue είναι συνεχές στο $\emptyset$ δηλαδή αν μία ακολουθία συνόλων φθίνει στο $\emptyset$ τότε $\lim_{n\to\infty} \mu(E_n) = 0$.*
Λέμε ότι η ακολουθία E_n φθίνει στο $\emptyset$ αν $E_{n+1} \subset E_n$ και $\bigcap_{n=1}^{\infty} E_n = \emptyset$. Απο το Θεώρημα 3.5.1 4. έχουμε ότι

$$\lim_{n\to\infty} \mu(E_n) = \mu\left(\bigcap_{n=1}^{\infty} E_n\right) = \mu(\emptyset) = 0$$

οπότε και αποδεικνύεται η συνέχεια του μ.

3.6 Η αλγεβρα *Borel* και τα σύνολα *Borel*.

Στο σημείο αυτό θα εισάγουμε μία πολύ σημαντική $\sigma-$ άλγεβρα, η οποία ονομάζεται η άλγεβρα Borel.

Ορισμός 3.6.1 *Ας υποθέσουμε ότι $\mathcal{F}$ είναι μία $\sigma-$ άλγεβρα τέτοια ώστε $A \in \mathcal{F}$. Λέμε ότι η $\mathcal{F}$ περιέχει το σύνολο A. Μπορούν να υπάρχουν πολλές $\sigma-$ άλγεβρες που περιέχουν ένα σύνολο A. Η*

$$\sigma(A) = \bigcap \{\mathcal{F} \ \mid \ \mathcal{F} \text{ σ-αλγεβρα που περιέχει το } A\}$$

είναι η μικρότερη $\sigma-$άλγεβρα που περιέχει το A.

Παράδειγμα 3.6.1 *Έστω A ένα σύνολο.*
Οι $\mathcal{F}_1 = \{\emptyset, \mathbb{R}, A, A^c\}$ και $\mathcal{F}_2 = \{\emptyset, \mathbb{R}, A, A^c, B, B^c, A \cup B, (A \cup B)^c, (A \cup B^c), (A \cup B^c)^c, (A^c \cup B), (A^c \cup B)^C\}$ είναι δύο $\sigma-$ άλγεβρες οι οποίες περιέχουν το σύνολο A. Απο αυτές, η $\mathcal{F}_1$ είναι η μικρότερη $\sigma-$άλγεβρα που περιέχει το A, άρα $\sigma(A) = \mathcal{F}_1$.

Ορισμός 3.6.2 *Η μικρότερη $\sigma-$ άλγεβρα που περιέχει όλα τα ανοιχτά διαστήματα, ονομάζεται άλγεβρα Borel και συμβολίζεται $\mathcal{B}$. Συμβολικά $\mathcal{B} = \bigcap \{\mathcal{F} \ \mid \ \mathcal{F}$ σ-αλγεβρα που περιέχει όλα τα διαστήματα$\}$*

Παράδειγμα 3.6.2 *Η άλγεβρα Borel περιέχει οποιοδήποτε κλειστό διάστημα $[a, b]$.*
Πράγματι, μπορούμε να γράψουμε

$$[a, b] = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, b + \frac{1}{n} \right)$$

Τα διαστήματα $I_n = \left(a - \frac{1}{n}, b + \frac{1}{n}\right)$ είναι ανοιχτά διαστήματα, οπότε ανήκουν στην $\mathcal{B}$, και λόγω του ότι η $\mathcal{B}$ είναι σ-άλγεβρα και οι τομές τους θα ανήκουν επίσης στην $\mathcal{B}$.

Η άλγεβρα Borel **δεν** ταυτίζεται με τα μετρήσιμα σύνολα του $\mathbb{R}$, για την ακρίβεια, αν ένα σύνολο είναι σύνολο Borel τότε είναι και μετρήσιμο αλλά υπάρχουν μετρήσιμα σύνολα τα οποία δεν είναι σύνολα Borel.

Το ακόλουθο παράδειγμα το οποίο οφείλεται στον Lusin μας δίνει ένα σύνολο το οποίο είναι μετρήσιμο αλλά δεν είναι Borel

Παράδειγμα 3.6.3 *Κάθε άρρητος αριθμός x μπορεί να γραφεί με* **μοναδικό** *τρόπο σαν ένα συνεχιζόμενο κλάσμα της μορφής*

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots}}}}$$

όπου a_0 είναι κάποιος ακέραιος και $a_k \in \mathbb{Z}^+$, $k = 1, 2, \cdots$. Το ανάπτυγμα αυτό πολλές φορές συμβολίζεται για πιο απλά ως $x = [a_0; a_1, a_2, \cdots]$. Υπάρχουν αλγόριθμοι για τον υπολογισμό αυτών των αναπτυγμάτων. Πολλές φορές άρρητοι αριθμοί εμφανίζουν πολύ ενδιαφέρουσες δομές στα αναπτύγματα τους ως επαναλαμβανόμενα κλάσματα π.χ.

$$e = [2; 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, 1, 1, 10, 1, \cdots]$$

Ας ορίσουμε τώρα το σύνολο A, των αρρήτων αριθμών για τους οποίους ισχύει η ιδιότητα ότι τα αναπτύγματα τους αντιστοιχούν σε ακολουθίες $(a_0, a_1, \cdots)$ για τις οποίες υπάρχει μια άπειρη υποακολουθία (a_{n_k}) τέτοια ώστε το κάθε στοιχείο της να είναι διαιρέτης του επόμενου στοιχείου. Μπορεί κανείς να δείξει ότι το σύνολο A δεν είναι σύνολο Borel παρότι είναι μετρήσιμο κατά Lebesgue.

Η σχέση του $\mathcal{B}$ και του $\mathcal{M}$ ειναι η ακόλουθη. Έχουμε ότι $\mathcal{B} \subset \mathcal{M}$ με αυστηρό εγκλεισμό. Μπορεί κανείς όμως να δείξει ότι το μέτρο Lebesgue **δεν** μπορεί να ξεχωρίζει τα σύνολα Borel απο τα μετρήσιμα σύνολα, με την έννοια ότι για κάθε μετρήσιμο σύνολο E μπορούμε να βρούμε ένα σύνολο Borel B, τέτοιο ώστε $E \subset B$ και $\mu(E) = \mu(B)$ ή ισοδύναμα $\mu(B \setminus E) = 0$. Συνεπώς, μπορούμε να πούμε ότι η άλγεβρα Borel περιέχει πρακτικά όλα τα υποσύνολα του $\mathbb{R}$ τα οποία μας ενδιαφέρουν.

Επίσης όλα τα μηδενοσύνολα κάτω απο το μέτρο Lebesgue δεν περιέχονται απαραίτητα στο $\mathcal{B}$, δηλαδή μπορεί να βρεί κάποιος σύνολα F τέτοια ώστε $\mu(F) = 0$ αλλά $F \notin \mathcal{B}$. Παραδείγματα τέτοιων συνόλων είναι ορισμένα υποσύνολα του συνόλου του Cantor (το οποίο ανήκει στο $\mathcal{B}$). Βέβαια, τα σύνολα αυτά είναι μετρήσιμα κατά

Lebesgue. Το παραπάνω μας λέει ότι υπάρχουν μηδενοσύνολα N ως προς το μέτρο Lebesgue για τα οποία $N \notin \mathcal{B}$ αλλά $N \in \mathcal{M}$. Αφού το $\mathcal{B}$ **δεν** περιέχει όλα τα μηδενοσύνολα του μέτρου Lebesgue θα λέμε ότι το μέτρο Lebesgue δεν είναι πλήρες ως προς την σ-άλγεβρα $\mathcal{B}$.

Μπορεί όμως κανείς να δείξει οτι το $\mathcal{M}$ περιέχει όλα τα μηδενοσύνολα του ως προς το μέτρο Lebesgue και ακόμα ότι περιέχει και όλα τα υποσύνολα των μηδενοσυνόλων, για την ακρίβεια η $\mathcal{M}$ είναι η μικρότερη σ-άλγεβρα με την ιδιότητα αυτή.

Αν πάρουμε οποιοδήποτε μηδενοσύνολο κάτω απο το μέτρο Lebesgue , N, το οποίο ανήκει στο $\mathcal{M}$ και μετά πάρουμε τις ενώσεις $F \cup N$ για οποιαδήποτε $F \in \mathcal{B}$ τότε φυσικά $F \cup N \in \mathcal{M}$. Μπορεί κανείς να δείξει ότι η μικρότερη $\sigma-$άλγεβρα η οποία περιέχει όλα σύνολα της μορφής $F \cup N$ ταυτίζεται με την $\mathcal{M}$.

Πρόταση 3.6.1 *Ισχύει ότι*

$$\sigma(\{F \cup N \ | \ F \in \mathcal{B}, N \in \mathcal{M}, \mu(N) = 0\}) = \mathcal{M}$$

Λέμε ότι το $\mathcal{M}$ είναι η πλήρωση του $\mathcal{B}$ (ως προς το μέτρο Lebesgue).

3.7 Τό μέτρο *Lebesgue* σε υποσύνολα του $\mathbb{R}$

. Μπορούμε με ακριβώς τον ίδιο τρόπο όπως εργαστήκαμε για όλο το $\mathbb{R}$ να ορίσουμε το μέτρο Lebesgue σε υποσύνολα του $\mathbb{R}$. Ιδιαίτερο ενδιαφέρον παρουσιάζει για τις εφαρμογές, ο ορισμός του μέτρου Lebesgue σε διαστήματα $[a, b]$. Η κατασκευή αφήνεται σαν άσκηση. Επίσης μπορεί να οριστεί και η άλγεβρα Borel $\mathcal{B}([a, b])$ η οποία ειναι η μικρότερη $\sigma-$άλγεβρα που περιέχει όλα τα ανοιχτά διαστήματα του $[a, b]$.

3.8 Το θεώρημα επέκτασης

Ας θεωρήσουμε ξανά την κατασκευή του μέτρου Lebesgue στο διάστημα $(0, 1]$. Ας πάρουμε το σύνολο Σ_0 το οποίο περιέχει όλα τα $F \subseteq (0, 1]$ τα οποία μπορεί να γραφούν στην μορφή

$$F = (a_1, b_1] \cup \cdots \cup (a_k, b_k]$$

όπου $k \in \mathbb{N}$ και $0 \leq a_1 \leq b_1 \leq \cdots \leq a_k \leq b_k$.

Μπορεί να δείξει κανείς ότι το σύνολο Σ_0 το οποίο είναι ένα σύνολο απο υποσύνολα του $\Omega = (0, 1]$ έχει τις ιδιότητες

1. $\Omega \in \Sigma_0$
2. $A \in \Sigma_0$ τότε και $A^c = (0, 1] \setminus A \in \Sigma_0$
3. $A, B \in \Sigma_0$ τότε και $A \cup B \in \Sigma_0$.

Μια συλλογή υποσυνόλων ενός συνόλου Ω με τις ιδιότητες αυτές ονομάζεται **άλγεβρα**. Παρατηρούμε ότι αν έχουμε μια άλγεβρα τότε αυτή είναι κλειστή ως προς τις πεπερασμένες ενώσεις των στοιχείων της δηλαδή

$$A_1, \cdots, A_k \in \Sigma_0 \implies \bigcup_{j=1}^{k} A_j \in \Sigma_0$$

αλλα δεν είναι απαραίτητο ότι μπορούμε να πάρουμε $k = \infty$. Αν μπορούμε να πάρουμε και $k = \infty$ τότε το Σ_0 θα είναι μια σ-άλγεβρα.

Μπορούμε τώρα να θεωρήσουμε την μικρότερη σ-άλγεβρα η οποία περιέχει την άλγεβρα Σ_0. Αυτή θα συμβολίζεται με $\sigma(\Sigma_0)$. Στην περίπτωση που μελετάμε $\sigma(\Sigma_0) = \mathcal{B}((0, 1])$.

Ένα μέτρο θα πρέπει να ορίζεται σε μια σ-άλγεβρα και όχι σε μια άλγεβρα. Είναι όμως πολλές φορές ευκολότερο να ορίσουμε το μέτρο σε μία άλγεβρα Σ_0. Για παράδειγμα, μπορεί κανείς να ορίσει την τιμή του μέτρου Lebesgue σε κάθε στοιχείο $A \in \Sigma_0$. Εφόσον $A \in \Sigma_0$, απο τον ορισμό του Σ_0 γνωρίζουμε ότι $A = (a_1, b_1] \cup \cdots \cup (a_k, b_k]$ για κάποιο $k \in \mathbb{N}$ και κάποια $a_i, b_i \in (0, 1]$, $i = 1, \cdots, k$, τέτοια ώστε $0 \leq a_1 \leq b_1 \leq \cdots \leq a_k \leq b_k$. Συνεπώς μπορούμε να ορίσουμε

$$\bar{\mu}(A) = \sum_{i=1}^{k} (b_i - a_i)$$

Η απεικόνιση $\bar{\mu} : \Sigma_0 \to [0,1]$ είναι μια απεικόνιση η οποία ικανοποιεί όλες τις ιδιότητες ενός μέτρου με την μοναδική διαφορά ότι το πεδίο ορισμού της είναι μια άλγεβρα αντι για μια σ-άλγεβρα. Αν λοιπόν μπορούσαμε να **επεκτείνουμε** την απεικονιση $\bar{\mu}$ σε μία απεικόνιση $\mu : \sigma(\Sigma_0) \to [0,1]$ η οποία έχει τις ιδιότητες του μέτρου και είναι τέτοια ώστε $\mu(A) = \bar{\mu}(A)$ για όλα τα $A \in \Sigma_0$, δηλαδή **ταυτίζεται** με την $\bar{\mu}$ αν περιοριστεί στο $\Sigma_0 \subset \sigma(\Sigma_0)$ τότε ουσιαστικά έχουμε εξασφαλίσει την ύπαρξη του μέτρου Lebesgue.

Την ύπαρξη μιας τέτοιας επέκτασης εξασφαλίζει το θεώρημα επέκτασης το οποίο οφείλεται στον Έλληνα μαθηματικό Κωσταντίνο Καραθεοδωρή (C. Caratheodory).

Θεώρημα 3.8.1 *Έστω Σ_0 μια άλγεβρα απο υποσύνολα ενός συνόλου Ω και έστω $\sigma(\Sigma_0)$ η σ-άλγεβρα που παράγεται απο την Σ_0. Αν έχουμε ορίσει μια απεικόνιση $\bar{\mu} : \Sigma_0 \to [0,1]$ η οποία έχει τις ιδιότητες του μετρου (δηλαδή είναι σ-αθροίσιμη) τότε υπάρχει ένα μέτρο $\mu : \sigma(\Sigma_0) \to [0,1]$ το οποίο είναι η επέκταση του $\bar{\mu}$ στο $\sigma(\Sigma_0)$. Το μέτρο μ είναι* **μοναδικό**.

3.9 Μία πρώτη επαφή με τις πιθανότητες

Ας υποθέσουμε ότι επιλέγουμε τυχαίους αριθμούς απο το διάστημα $[0,1]$ οι οποίοι είναι ομοιόμορφα κατανεμημένοι, δηλαδή έχουμε την τυχαία μεταβλητή $X \sim \mathcal{U}[0,1]$. Γνωρίζουμε ότι $P(X \leq x) = x$.

Ένας τρόπος να περιγράψουμε το πείραμα αυτό είναι με την χρήση του μέτρου Lebesgue που μόλις ορίσαμε.

Ας πάρουμε σαν δειγματικό χώρο το $\Omega = [0,1]$ και ας θεωρήσουμε την άλγεβρα Borel $\mathcal{B}$ η οποία είναι η μικρότερη $\sigma-$άλγεβρα που περιέχει όλα τα υποσύνολα του $[0,1]$. Μπορούμε να θεωρήσουμε ότι η $\mathcal{B}$ περιέχει όλες τις πιθανές ερωτήσεις που μπορεί κανείς να κάνει για το πείραμα αυτό.

Παράδειγμα 3.9.1 *Το υποσύνολο $A = [0,x] \in \mathcal{B}$ αντιστοιχεί στην ερώτηση 'Η τυχαία μεταβλητή X παίρνει τιμές $0 \leq X \leq x$.*

Παράδειγμα 3.9.2 *Το υποσύνολο $A = [0,\frac{1}{3}] \cup [\frac{2}{3},1] \in \mathcal{B}$ αντιστοιχεί στην ερώτηση 'Η τυχαία μεταβλητή X παίρνει τιμές είτε $0 \leq X \leq \frac{1}{3}$ είτε $\frac{2}{3} \leq X \leq 1$'.*

Το ότι $P(X \leq x) = x$ μπορεί να ξαναγραφεί με το παρακάτω τρόπο $P(X \leq x) = \mu([0,x]) = x$ όπου μ είναι το μέτρο Lebesgue στο διάστημα $[0,1]$. Συνεπώς η ομοιόμορφη κατανομή μπορεί να εκφραστεί με την χρήση του μέτρου Lebesgue. Αυτό μας επιτρέπει το να μπορούμε να χειριστούμε μία σειρά απο ενδιαφέροντα ερωτήματα σχετικά με την τυχαία μεταβλητή.

Παράδειγμα 3.9.3 *Αν $X \sim \mathcal{U}[0,1]$ υπολογίστε την πιθανότητα*

$$P\left(0 \leq X \leq \frac{1}{3} \text{ ή } \frac{2}{3} \leq X \leq 1\right).$$

Σύμφωνα με τα παραπάνω αν ορίσουμε $A = [0,\frac{1}{3}] \cup [\frac{2}{3},1] \in \mathcal{B}$ τότε

$$P\left(0 \leq X \leq \frac{1}{3} \text{ ή } \frac{2}{3} \leq X \leq 1\right) = \mu(A)$$

όπου μ το μέτρο Lebesgue.
Το A είναι η ένωση δύο διαστημάτων που είναι ξένα μεταξύ τους, οπότε απο τις ιδιότητες του μέτρου έχουμε ότι

$$\mu(A) = \mu\left(\left[0,\frac{1}{3}\right]\right) + \mu\left(\left[\frac{2}{3},1\right]\right) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}.$$

Συνεπώς η πιθανότητα που ζητάμε είναι ίση με $\frac{2}{3}$ όπως άλλωστε μας υποδεικνύει και η εμπειρία μας.

Παράδειγμα 3.9.4 *Αν $X \sim \mathcal{U}[0,1]$ υπολογίστε την πιθανότητα $P(X = \frac{1}{3} \text{ ή } X = \frac{1}{2})$.*
Η πιθανότητα αυτή μπορεί να εκφραστεί με την βοήθεια του μέτρου Lebesgue ως εξής:

$$P\left(X = \frac{1}{3} \text{ ή } X = \frac{1}{2}\right) = \mu\left(\left\{\frac{1}{3}\right\} \cup \left\{\frac{1}{2}\right\}\right) = 0$$

εφόσον το σύνολο αυτό είναι μηδενοσύνολο.

Σχήμα 3.1: Ο Henri Lebesgue

Παράδειγμα 3.9.5 *Ποια η πιθανότητα $P(X = \text{ ρητός})$;*
Η πιθανότητα αυτή μπορεί να εκφραστεί με την βοήθεια του μέτρου Lebesgue ως

$$P(X = \text{ ρητός}) = \mu(\mathbb{Q}).$$

Όμως, το σύνολο των ρητών είναι μηδενοσύνολο ως προς το μέτρο Lebesgue άρα η πιθανότητα αυτή είναι μηδενική.

3.10 Παράρτημα 1: Μερικά ιστορικά στοιχεία

Στο παράρτημα αυτό παραθέτουμε ορισμένα ιστορικά στοιχεία για τους πρωτεργάτες της θεωρίας μέτρου.

3.10.1 H. Lebesgue, 1875-1941

Ο Henri Lebesgue ήταν ένας μεγάλος Γάλλος μαθηματικός ο οποίος θεωρείται και ένας απο τους σημαντικότερους θεμελιωτές της θεωρίας μέτρου και έδωσε το όνομα του στο μέτρο Lebesgue. Σπούδασε στην περίφημη Ecole Normale Supérieure και στην διδακτορική του διατριβή με τον τίτλο Integrale, Longuer, Aire (Ολοκλήρωμα, Μήκος και Εμβαδό) έθεσε τις βάσεις της θεωρίας μέτρου και της ολοκλήρωσης όπως τις χρησιμοποιούμε ακόμα και σήμερα. Στην διατριβή αυτή γενίκευσε την έννοια του ολοκληρώματος Riemann-Stieltjes διατυπώνοντας την ιδέα του ολοκληρώματος με την χρήση του μέτρου, το οποίο και πήρε το όνομα του. Η γενίκευση αυτή οδήγησε σε πολύ ενδιαφέρουσες εφαρμογές και παράκαμψε ορισμένες αδυναμίες της θεωρίας της ολοκλήρωσης των Riemann-Stieltjes όπως π.χ. θέματα σχετικά με το ποιές συναρτήσεις είναι ολοκληρώσιμες, ή θέματα που σχετίζονται με τις εναλλαγές της πράξης του ορίου και της ολοκλήρωσης. Ο Lebesgue ασχολήθηκε και με άλλα θέματα της ανάλυσης και της τοπολογίας αλλά έμεινε στην ιστορία της σύγχρονης ανάλυσης για την αναμόρφωση της θεωρίας της ολοκλήρωσης που πρότεινε.

3.10.2 Emile Borel, 1871-1956

Ο Emile Borel ήταν ένας μεγάλος Γάλλος μαθηματικός ο οποίος είχε σημαντικές συμβολές στην θεωρία μέτρου και στην εφαρμογή της στην θεωρία πιθανοτήτων, καθώς και στη θεωρία παιγνίων. Το σύνολο Borel και η άλγεβρα Borel ονομάστηκαν έτσι προς τιμή του. Είναι επίσης γνωστός για το λήμμα Borel-Cantelli στις πιθανότητες, το οποίο ξεκαθαρίζει την συμπεριφορά οριακών γεγονότων, το θεώρημα Heine-Borel το οποίο χαρακτηρίζει τα συμπαγή

Σχήμα 3.2: Ο Emile Borel

υποσύνολα μετρικών χώρων κ.α. Εκτός απο τα μαθηματικά ασχολήθηκε και με την πολιτική, ενώ κατά την διάρκεια της Γερμανικής κατοχής ήταν μέλος της Αντίστασης. Ένας κρατήρας στην Σελήνη φέρει το όνομα του.

3.10.3 Κωσταντίνος Καραθεοδωρή, 1873-1950

Ο Κωσταντίνος Καραθεοδωρή ήταν ένας Έλληνας μαθηματικός (απο οικογένεία Φαναριωτών) ο οποίος γενήθηκε, στην Γερμανία, σπούδασε αρχικά μηχανικός σε στρατιωτική σχολή στο Βέλγιο, εργάστηκε σαν μηχανικός στην Αίγυπτο και το 1900 ξεκίνησε σπουδές μαθηματικών στο Πανεπιστήμιο του Βερολίνου τις οποίες συνέχισε το Πανεπιστήμιο του Göttingen το οποίο ήταν και ένα απο τα μεγαλύτερα κέντρα για την μαθηματική έρευνα την εποχή αυτή στην Ευρώπη αν όχι σε ολόκληρο τον κόσμο. Στο Göttingen ολοκλήρωσε την διδακτορική του διατριβή επάνω στον λογισμό των μεταβολών. Στην συνέχεια εργάστηκε σε διάφορα Γερμανικά πανεπιστήμια (Βόνη, Ανόβερο, Μπρεσλάου) ώσπου ξαναγύρισε το 1913 ως καθηγητής των Εφαρμοσμένων Μαθηματικών στο Göttingen όπου και έμεινε 5 χρόνια για να αποδεχθεί μια θέση στο Βερολίνο την οποία και άφησε για να επιστρέψει στην Ελλάδα μετά απο πρόσκληση της Ελληνικής κυβέρνησης για να οργανώσει ένα δεύτερο πανεπιστήμιο στην Σμύρνη. Παράλληλα με αυτό ήταν καθηγητής στο Εθνικό Καποδιστριακό Πανεπιστήμιο Αθηνών και μετά την Μικρασιατική καταστροφή στο Εθνικό Μετσόβιο Πολυτεχνείο το οποίο εγκατέλειψε το 1924 για να ξαναγυρίσει στο Μόναχο απο όπου και συνταξιοδοτήθηκε.

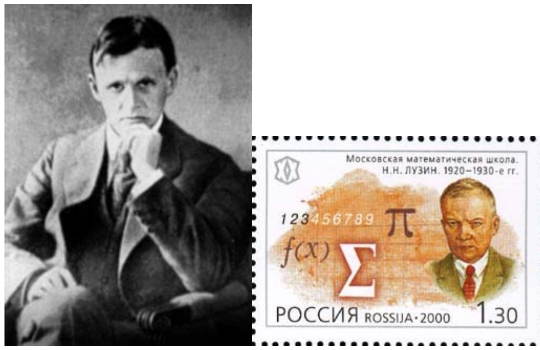
Ο Καραθεοδωρή ήταν ένας μαθηματικός διεθνής φήμης, η οποία ξεπερνούσε τα όρια της Ευρώπης και είχε σημαντικές συμβολές στον λογισμό των μεταβολών, την θεωρία μέτρου, την μιγαδική ανάλυση, τις διαφορικές εξισώσεις, την κυρτή ανάλυση και την μαθηματική φυσική. Στην θεωρία μέτρου οφείλεται σε αυτό το θεώρημα επέκτασης το οποίο είναι ιδιαίτερα σημαντικό για την κατασκευή γενικών μετρων.

3.10.4 Giuseppe Vitali, 1875-1932

Ο Giuseppe Vitali ήταν ένας Ιταλός μαθηματικός ο οποίος αποφοίτησε απο την Scuola Normale Superiore στην Pisa εργάστηκε μαζί με τον μεγάλο Ιταλό μαθηματικό Dini σαν βοηθός του και μετά απασχολήθηκε στην μέση εκπαίδευση. Ασχολήθηκε επίσης με την πολιτική ως μέλος του Σοσιαλιστικού Κόμματος, μετά την διάλυση του οποίου επέστρεψε στα μαθηματικά, και αργότερα στην ανώτατη εκπαίδευση ως καθηγητής ανάλυσης στα πανεπιστήμια της Padua αρχικά και μετά της Bologna. Ο Vitali είναι γνωστός για την δουλειά του στην θεωρία μέτρου, και ειδικότερα για την κατασκευή ενός παραδείγματος μη μετρήσιμου συνόλου, καθώς και για ορισμένα αποτελέσματα στην θεωρία της σύγκλισης ακολουθιών συναρτήσεων.

Σχήμα 3.3: Ο Κωσταντίνος Καραθεοδωρή

Σχήμα 3.4: Ο Giuseppe Vitali



Σχήμα 3.5: Ο Nikolai Luzin

3.10.5 Nikolai Luzin, 1883-1950

Ο Nikolai Luzin ήταν ένας Ρώσος μαθηματικός ο οποίος είναι ιδιαίτερα γνωστός για τη συμβολή του στην θεωρία των συναρτήσεων και την θεωρία μέτρου. Συνεργάστηκε με τον μεγάλο Ρώσο μαθηματικό Egorov ο οποίος ήταν και δάσκαλος του, και μετά την εκπόνηση της διδακτορικής του διατριβής με θέμα τις τριγωνομετρικές σειρές έγινε καθηγητής καθαρών μαθηματικών στο πανεπιστήμιο της Μοσχας, το 1917, όπου για 10 έτη μαζί με τον Egorov έχτισε μια πολύ ισχυρή μαθηματική ομάδα με φοιτητές μεταξύ των οποίων ήταν ο Kolmogorov , ο Aleksandrov κ.α. Εκτός απο την θεωρία μέτρου ασχολήθηκε με την συνολοθεωρητική τοπολογία, και τις αναλυτικές συναρτήσεις. Στην θεωρία μέτρου είναι γνωστός για την κατασκευή ενός συνόλου το οποίο είναι μετρήσιμο αλλά όχι Borel , την θεωρία των χώρων Luzin (που θα αναφέρουμε συνοπτικά αργότερα) και το θεώρημα του Luzin σύμφωνα με το οποίο μια μετρήσιμη συνάρτηση είναι συνεχής σχεδόν σε όλο το πεδίο ορισμού της.

3.11 Παράρτημα: Αποδείξεις ορισμένων προτάσεων

3.11.1 Απόδειξη της Πρότασης 3.2.1

Εφόσον το A_i είναι μηδενοσύνολο για κάθε i, έχουμε ότι μπορούμε για κάθε $\epsilon = \epsilon_i$ να επιλέξουμε ένα κάλυμα $I_n^{(i)}$ τέτοιο ώστε

$$A_i \subset \bigcup_{n=1}^{\infty} I_n^{(i)}$$

και

$$\sum_{n=1}^{\infty} \ell(I_n^{(i)}) < \epsilon_i.$$

Θα επιλέξουμε $\epsilon_i = \frac{\epsilon_0}{2^i}$ για οποιοδήποτε $\epsilon_0 > 0$.

Θα αναδιατάξουμε τώρα την διπλή ακολουθία διαστημάτων $I_n^{(i)}$ σε μία ακολουθία με ένα δείκτη μόνο την J_k. Για οποιοδήποτε $k \in \mathbb{N}$ θα υπάρχει ένα ζεύγος $(n, i) \in \mathbb{N} \times \mathbb{N}$ έτσι ώστε $J_k = I_n^{(i)}$. Π.χ. $J_1 = I_1^{(1)}, J_2 = I_2^{(1)}$ κ.ο.κ.

Είναι προφανές ότι

$$A = \bigcup_{i=1}^{\infty} A_i \subset \bigcup_{i=1}^{\infty} \bigcup_{n=1}^{\infty} I_n^{(i)} = \bigcup_{k=1}^{\infty} J_k$$

άρα η συλλογή J_k, $k \in \mathbb{N}$ είναι ένα κάλυμα του A. Για να δείξουμε ότι το A είναι μηδενοσύνολο απομένει να υπολογίσουμε το συνολικό μήκος του $\bigcup_{k=1}^{\infty} J_k$ και να δείξουμε ότι μπορεί να γίνει μικρότερο απο οποιοδήποτε ϵ_0.

Έχουμε ότι

$$\sum_{k=1}^{\infty} \ell(J_k) = \sum_{n=1}^{\infty} \sum_{i=1}^{\infty} \ell(I_n^{(i)})$$

το οποίο είναι μία διπλή σειρά. Η διπλή αυτή σειρά λόγω του ότι οι όροι της είναι θετικοί μπορεί να γραφεί ως επαναλαμβανόμενο άθροισμα της μορφής

$$\sum_{n=1}^{\infty}\sum_{i=1}^{\infty}\ell(I_n^{(i)}) = \sum_{i=1}^{\infty}\left(\sum_{n=1}^{\infty}\ell(I_n^{(i)}\right) < \sum_{i=1}^{\infty}\frac{\epsilon_0}{2^i} = \epsilon_0$$

Συνεπώς το A είναι μηδενοσύνολο.

3.11.2 Απόδειξη της Πρότασης 3.3.1

1. Ας υποθέσουμε πρωτα ότι A μηδενοσύνολο. Θα πρέπει να δείξουμε ότι $\mu^*(A) = \inf Z_A = 0$.
Εφόσον $\mu^*(A) > 0$ για κάθε $A \subset \mathbb{R}$ αρκεί να δείξουμε ότι για κάθε $\epsilon > 0$ μπορούμε να βρούμε κάποιο $z \in Z_A$ τέτοιο ώστε $z < \epsilon$. Αυτό σημαίνει ότι μπορούμε να βρούμε ένα κάλυμα του A, I_n τέτοιο ώστε $\sum_{n=1}^{\infty}\ell(I_n) < \epsilon$. Απο τον ορισμό όμως του μηδενοσυνόλου αυτό ισχύει.
Ας υποθέσουμε τώρα ότι για κάποιο $A \subset \mathbb{R}$ ισχύει $\mu^*(A) = 0$. Αυτό σημαίνει ότι $\inf Z_A = 0$ και επειδή $Z_A \geq 0$ έχουμε ότι για κάθε $\epsilon > 0$ θα υπάρχει $z \in Z_A$ τέτοιο ώστε $z < \epsilon$ δηλαδή θα υπάρχει κάλυμα του A απο διαστήματα συνολικού μήκους μικρότερο του ϵ. Αυτός όμως είναι και ο ορισμός του μηδενοσύνολου.
2. Εφόσον $A \subset B$ έχουμε ότι αν $\{I_n\}$ είναι ένα κάλυμα του B τότε θα είναι και ένα κάλυμα του A ενώ αν $\{J_n\}$ είναι ένα κάλυμα του A δεν ισχύει απαραίτητα ότι θα είναι και κάλυμα του B. Απο το επιχείρημα αυτό βλέπουμε ότι αν έχουμε ένα κάλυμα του A μπορούμε να το συμπληρώσουμε με μερικά ακόμα διαστήματα έτσι ώστε να πάρουμε ένα κάλυμα του B. Συμπληρώνοντας το όμως αυξάνουμε το συνολικό του μήκος. Συνεπώς, $Z_B \subset Z_A$ οπότε και $\inf Z_A \leq \inf Z_B$ δηλαδή $\mu^*(A) \leq \mu^*(B)$.
3. Θα επικεντρώσουμε την προσοχή μας σε φραγμένα διαστήματα.
Θα δείξουμε πρώτα ότι $\mu^*(I) \leq \ell(I)$. Αν πάρουμε το απλούστερο κάλυμα του I, το οποίο αποτελείται απο το I μόνο, βλέπουμε ότι $\ell(I) \in Z_I$. Εφόσον, $\mu^*(I) = \inf Z_I$ έχουμε ότι $\mu^*(I) \leq \ell(I)$.
Θα δείξουμε κατόπιν ότι $\mu^*(I) \geq \ell(I)$. Ας θεωρήσουμε ότι το διάστημα που μας ενδιαφέρει είναι το κλειστό διάστημα $I = [a, b]$. Αρκεί να δείξουμε ότι για κάθε $\epsilon > 0$ ισχύει $\ell(I) \leq \mu^*(I) + \epsilon$.
Από τον ορισμό του εξωτερικού μέτρου έχουμε ότι $\mu^*(I) = \inf Z_I$ δηλαδή ότι μπορούμε να βρούμε κάποιο κάλυμα του I, $\{I_n\}$ τέτοιο ώστε

$$\sum_{n=1}^{\infty}\ell(I) \leq \mu^*(I) + \frac{\epsilon}{2} \tag{3.14}$$

Για κάθε ένα απο τα διαστήματα I_n μπορούμε να βρούμε ένα ανοιχτό διάστημα J_n τέτοιο ώστε $I_n \subset J_n$, και με τέτοιο τρόπο ώστε το συνολικό μήκος των J_n να είναι οσοδήποτε κοντά θέλουμε στο συνολικό μήκος των I_n. Πράγματι αν $I_n = [a_n, b_n]$ αρκεί να πάρουμε $J_n = (a_n - \frac{\epsilon}{2^{n+2}}, b_n + \frac{\epsilon}{2^{n+2}})$ για να βρούμε ότι

$$\sum_{n=1}^{\infty}\ell(J_n) \leq \sum_{n=1}^{\infty}\ell(I_n) + \frac{\epsilon}{2} \tag{3.15}$$

Συνδυάζοντας τις ανισότητες (3.14) και (3.15) καταλήγουμε στο ότι

$$\sum_{n=1}^{\infty}\ell(J_n) \leq \mu^*(I) + \epsilon \tag{3.16}$$

Είναι προφανές ότι

$$I \subset \bigcup_{n=1}^{\infty} J_n.$$

Θα ανατρέξουμε τώρα στο Θεώρημα Heine-Borel σύμφωνα με το οποίο οποιοδήποτε ανοιχτό κάλυμα ενός κλειστού διαστήματος $I = [a, b]$ πρέπει να έχει πεπερασμένο υποκάλυμα. Εφόσον το $\{J_n\}$ είναι ένα τέτοιο ανοιχτό κάλυμα θα πρέπει να έχει και πεπερασμένο υποκάλυμα, το οποίο ας υποθέσουμε ότι είναι το $\{J_k\}$, $k = 1, \cdots m$. Συνεπώς,

$$I \subset \bigcup_{k=1}^{m} J_k \tag{3.17}$$

Ας συμβολίσουμε ως $J_n = (\bar{a}_n, \bar{b}_n)$ τα διαστήματα αυτά και ας πάρουμε $\bar{a} = min(\bar{a}_1, \cdots, \bar{a}_m)$ και $\bar{b} = max(\bar{b}_1, \cdots, \bar{b}_m)$. Εφόσον ισχύει η (3.17) είναι προφανές ότι $\bar{a} < a$ και $b < \bar{b}$ συνεπώς $I = [a, b] \subset (\bar{a}, \bar{b})$. Απο τις ιδιότητες του μήκους όμως

$$\ell(I) = \ell([a, b]) \leq \ell((\bar{a}, \bar{b})) = \bar{b} - \bar{a}.$$

Εφόσον τα διαστήματα J_n έχουν αλληλοεπικαλύψεις έχουμε ότι

$$\ell((\bar{a}, \bar{b})) = \bar{b} - \bar{a} < \sum_{k=1}^{m} \ell(J_k)$$

συνεπώς επειδή $\sum_{k=1}^{m} \ell(J_k) < \sum_{k=1}^{\infty} \ell(J_k)$ ισχύει ότι $\ell(I) \leq \mu^*(I) + \epsilon$.
4. Αρκεί να δείξουμε ότι για κάθε $\epsilon > 0$ ισχύει

$$\mu^*\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} \mu^*(E_n) + \epsilon \tag{3.18}$$

Αν στην ανισότητα αυτή πάρουμε το όριο $\epsilon \to 0$ καταλήγουμε στο ζητούμενο αποτέλεσμα.
Ας προσπαθήσουμε να δείξουμε την ανισότητα (3.18). Απο τον ορισμό του $\mu^*(E_n) = \inf Z_{E_n}$ για κάθε ϵ_n μπορούμε να βρούμε ένα κάλυμα του E_n, $\{I_k^{(n)}\}$ τέτοιο ώστε

$$\sum_{k=1}^{\infty} \ell(I_k^{(n)}) \leq \mu^*(E_n) + \epsilon_n.$$

Αν επιλέξουμε τα ϵ_n κατάλληλα, π.χ. $\epsilon_n = \frac{\epsilon}{2^n}$ μπορούμε να αθροισουμε τις ανισότητες αυτές κατά μέλη και να πάρουμε οτι

$$\sum_{n=1}^{\infty}\left(\sum_{k=1}^{\infty} \ell(I_k^{(n)})\right) \leq \sum_{n=1}^{\infty} \mu^*(E_n) + \epsilon$$

Συνεπώς,

$$\sum_{n,k} \ell(I_k^{(n)}) = \sum_{n=1}^{\infty}\left(\sum_{k=1}^{\infty} \ell(I_k^{(n)})\right)$$

Εφόσον η διπλή ακολουθία $\{I_k^{(n)}\}$ καλύπτει το $\bigcup_{n=1}^{\infty} E_n$ έχουμε ότι

$$\mu^*\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n,k} \ell(I_k^{(n)}) \leq \sum_{n=1}^{\infty} \ell(I_k^{(n)}) + \epsilon$$

3.11.3 Απόδειξη της Πρότασης 3.4.2.

1. Ας υποθέσουμε ότι $E = \emptyset$. Τότε $E^c = \mathbb{R}$ οπότε για οποιοδήποτε $A \subset \mathbb{R}$ έχουμε ότι $A \cap E = A \cap \emptyset = \emptyset$ και $A \cap E^c = A \cap \mathbb{R} = A$. Εφόσον $\mu^*(\emptyset) = 0$ είναι προφανές ότι ισχύει η συνθήκη της μετρησιμότητας $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$. Όμοια αν $E = \mathbb{R}$.
2. Εφόσον $E \in \mathcal{M}$ για κάθε $A \subset \mathbb{R}$ ισχύει ότι $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$. Επειδή $(E^c)^c) = E$ η σχέση αυτή ισχύει και για το E^c οπότε $E^c \in \mathcal{M}$.
3. Ας υποθέσουμε πρώτα ότι τα $\{E_n\}$ είναι ανά δύο ξένα, δηλαδή ότι $E_i \cap E_j = \emptyset$.
Ας πάρουμε πρώτα την περίπτωση δύο συνόλων $E_1, E_2 \in \mathcal{M}$ και οποιοδήποτε $B \subset \mathbb{R}$. Αν επιλεξουμε ως $A = B \cap (E_1 \cup E_2)$ και εφαρμόσουμε τον ορισμό της μετρησιμότητας για το E_2 καταλήγουμε ότι

$$\begin{aligned}\mu^*(B \cap (E_1 \cup E_2)) &= \mu^*(B \cap (E_1 \cup E_2) \cap E_2) + \mu^*(B \cap (E_1 \cup E_2) \cap (E_2)^c) \\ &= \mu^*(B \cap E_1) + \mu^*(B \cap E_2).\end{aligned}$$

Ας πάρουμε τώρα την περίπτωση τριών συνόλων $E_1, E_2, E_3 \in \mathcal{M}$ και οποιοδήποτε $B \subset \mathbb{R}$. Αν επιλεξουμε ως $A = B \cap (E_1 \cup E_2 \cup E_3)$ και εφαρμόσουμε τον ορισμό της μετρησιμότητας για το E_3 καταλήγουμε ότι

$$\begin{aligned}\mu^*(B \cap (E_1 \cup E_2 \cup E_3)) &= \mu^*(B \cap (E_1 \cup E_2 \cup E_3) \cap E_3) + \mu^*(B \cap (E_1 \cup E_2 \cup E_3) \cap (E_3)^c) \\ &= \mu^*(B \cap E_3) + \mu^*(B \cap (E_1 \cup E_2)) \\ &= \mu^*(B \cap E_1) + \mu^*(B \cap E_2) + \mu^*(B \cap E_3)\end{aligned}$$

Κάνοντας χρήση της μεθόδου της επαγωγής μπορούμε να δείξουμε οτι αν $F_n = \bigcup_{k=1}^{n} E_k$ τότε

$$\mu^*(B \cap F_n) = \sum_{k=1}^{n} \mu^*(B \cap E_k)$$

Εφόσον η πεπερασμένη ένωση F_n είναι μετρήσιμη ως προς το εξωτερικό μέτρο μ^* έχουμε ότι

$$\begin{aligned} \mu^*(B) &= \mu^*(B \cap F_n) + \mu^*(B \cap F_n^c) \\ &= \sum_{k=1}^{n} \mu^*(B \cap E_k) + \mu^*(B \cap F_n^c) \end{aligned} \tag{3.19}$$

Ας πάρουμε τώρα το όριο $n \to \infty$. Μπορούμε να δούμε πολύ εύκολα ότι $B \cap E^c \subset B \cap F_n^c$. Από τις ιδιότητες του εξωτερικού μέτρου Lebesgue έχουμε ότι

$$\mu^*(B \cap E^c) \leq \mu^*(B \cap F_n^c).$$

Αντικαθιστώντας αυτό στην εξίσωση (3.19) παίρνουμε ότι

$$\mu^*(B) \geq \sum_{k=1}^{n} \mu^*(B \cap E_k) + \mu^*(B \cap E^c) \tag{3.20}$$

Μπορούμε λοιπόν τώρα να πάρουμε το όριο $n \to \infty$ στην (3.20) και να καταλήξουμε στο ότι

$$\mu^*(B) \geq \sum_{k=1}^{\infty} \mu^*(B \cap E_k) + \mu^*(B \cap E^c) \tag{3.21}$$

Όμως

$$\bigcup_{k=1}^{\infty} (B \cap E_k) = B \cap \left(\bigcup_{k=1}^{\infty} E_k \right) = B \cap E$$

οπότε απο την υποαθροιστική ιδιότητα του εξωτερικού μέτρου Lebesgue έχουμε ότι

$$\mu^*(B \cap E) = \mu^* \left(\bigcup_{k=1}^{\infty} (B \cap E_k) \right) \leq \sum_{k=1}^{\infty} \mu^*(B \cap E_k) \tag{3.22}$$

Συνδυάζοντας αυτό με την ανισότητα (3.21) καταλήγουμε στο ότι

$$\mu^*(B) \geq \mu^*(B \cap E) + \mu^*(B \cap E^c)$$

Απο το Παράδειγμα 3.4.1 αυτό είναι αρκετό για να δούμε ότι το $E = \bigcup_{k=1}^{\infty} E_k$ είναι μετρήσιμο ως προς το μ^*.

3.11.4 Απόδειξη της Πρότασης 3.5.1

1. Απο τις ιδιότητες του εξωτερικού μέτρου Lebesgue έχουμε ότι $\mu^*(A) \leq \mu^*(B)$. Επειδή το μέτρο Lebesgue είναι απλά ο περιορισμός του μ^* στο $\mathcal{M}$ έχουμε ότι $\mu(A) \leq \mu(B)$.
Ας γράψουμε τώρα το B σαν την ένωση $B = A \cup (B \setminus A)$. Επειδή τα B και $B \setminus A$ είναι ξένα μεταξύ τους σύνολα απο την αθροιστικότητα του μέτρου Lebesgue έχουμε ότι

$$\mu(B) = \mu(A) + \mu(B \setminus A)$$

απ' όπου αποδεικνύεται το ζητούμενο.
2. Ας θέσουμε $B_1 = E_1$ και $B_n = E_n \setminus E_{n-1}$ για $n > 1$.
Ισχύει ότι

$$\bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} B_n$$

και τα B_n είναι ξένα μεταξύ τους.
Απο την αθροιστικότητα του μέτρου Lebesgue έχουμε ότι

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(B_n) = \lim_{n\to\infty} \sum_{i=1}^{n} \mu(B_i) = \lim_{n\to\infty} \mu\left(\bigcup_{i=1}^{n} B_i\right) = \lim_{n\to\infty} \mu(E_n)$$

αφού $E_n = \bigcup_{i=1}^{n} B_i$.
3. Ας ορίσουμε την ακολουθία συνόλων B_n ως εξής: $B_1 = E_1 \setminus E_1 = \emptyset$, $B_2 = E_1 \setminus E_2$, $\cdots$, $B_n = E_1 \setminus E_n$ κλπ.
Είναι προφανές ότι $B_n \subset B_{n+1}$ για κάθε n.
Συνεπώς, απο το 2 έχουμε ότι

$$\mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} \mu(B_n)$$

Αλλά

$$\mu(B_n) = \mu(E_1 \setminus E_n) = \mu(E_1) - \mu(E_n).$$

Όμως,

$$\mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \mu\left(\bigcup_{n=1}^{\infty} (E_1 \setminus E_n)\right) = \mu\left(E_1 \setminus \left(\bigcap_{n=1}^{\infty} E_n\right)\right) = \mu(E_1) - \mu\left(\bigcap_{n=1}^{\infty} E_n\right)$$

Συνδυάζοντας τις σχέσεις αυτές καταλήγουμε στο ότι

$$\mu\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mu(E_n)$$

Κεφάλαιο 4

Γενίκευση της έννοιας του μέτρου

4.1 Εισαγωγή

Το μέτρο Lebesgue που μελετήσαμε στο προηγούμενο κεφάλαιο, μπορεί να γενικευθεί. Θα δούμε στο κεφάλαιο αυτό πως μπορούμε να ορίσουμε απεικονίσεις που είναι μέτρα επάνω στο $\mathbb{R}$ τα οποία όμως είναι γενικότερα απο το μέτρο Lebesgue.

4.2 $\sigma-$άλγεβρες

Έστω Ω ένα σύνολο. Στο σύνολο αυτό μπορούμε να πάρουμε συλλογές απο υποσύνολα του.

Ορισμός 4.2.1 *Έστω Ω ένα σύνολο και $\mathcal{F}$ μία συλλογή απο υποσύνολα του Ω, η οποία ικανοποιεί τις παρακάτω ιδιότητες*

1. $\Omega \in \mathcal{F}$ *και* $\emptyset \in \mathcal{F}$
2. *Αν* $F \in \mathcal{F}$ *τότε και* $F^c \in \mathcal{F}$.
3. *Αν* $F_n \in \mathcal{F}$, $n = 1, 2, \cdots$ *τότε* $\bigcup_{n=1}^{\infty} F_n \in \mathcal{F}$

ονομάζεται $\sigma-$άλγεβρα.

Ορισμός 4.2.2 *Έστω Ω ένα σύνολο και $\mathcal{F} m'\iota\alpha \sigma-$άλγεβρα επάνω σε αυτό. Το ζεύγος $(\Omega, \mathcal{F})$ ονομάζεται* **μετρήσιμος χώρος**.

Ορισμός 4.2.3 *Έστω $(\Omega, \mathcal{F})$ ένας μετρήσιμος χώρος. Μία απεικόνιση $\mu : \mathcal{F} \to \mathbb{R}$ με τις ιδιότητες*

1. $\mu(\emptyset) = 0$
2. *Αν* $E_i \in \mathcal{F}$, *ξένα μεταξύ τους, τότε* $\mu(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu(E_i)$

ονομάζεται μέτρο.

Ορισμός 4.2.4 *Πλήρες μέτρο*

4.3 Κατασκευή μέτρων και το Θέωρημα επέκτασης του Καραθεοδωρή

Στο σημείο αυτό θα ασχοληθούμε με το πως αν μας δώσει κανείς ένα μετρήσιμο χώρο, μπορούμε να κατασκευάσουμε ένα μέτρο επάνω στον χώρο αυτό. Η κατασκευή αυτή χρησιμοποιεί την γενίκευση της έννοιας του εξωτερικού μέτρου που είδαμε όταν κατασκευάζαμε το μέτρο Lebesgue .

Ορισμός 4.3.1 *Έστω $(\Omega, \mathcal{F})$ ένας μετρήσιμος χώρος. Μία απεικόνιση $\mu^* : \mathcal{F} \to [0, \infty)$ η οποία ικανοποιεί τις ιδιότητες*

1. $\mu^*(\emptyset) = 0$
2. $A \subset B \implies \mu^*(A) \leq \mu^*(B)$ *(μονότονη)*
3. $\mu^*(\bigcup_{i=1}^{\infty} E_i) \leq \sum_{i=1}^{\infty} \mu^*(E_i)$ *(αριθμήσιμα υποαθροιστική)*

ονομάζεται **εξωτερικό μέτρο**.

Για να μπορέσουμε να κατασκευάζουμε ένα μέτρο απο ένα εξωτερικό μέτρο θα πρέπει να Ένας τρόπος κατασκευή-ς ενός εξωτερικού μέτρου είναι ο ακόλουθος

$$\mu^*(A) = \inf \left\{ \sum_{i=1}^{\infty} \rho(E_i) \ \mid \ A \subset \bigcup_{i=1}^{infty} E_i, \ E_i \in \mathcal{F} \right\}$$

όπου ρ είναι μία μη αρνητική συνάρτηση που ορίζεται στο $\mathcal{F}$ τέτοια ώστε $\rho(\emptyset) = 0$.
Στην περίπτωση του εξωτερικού μέτρου Lebesgue είχαμε επιλέξει $\rho([a, b]) = b - a$.

Το εξωτερικό μέτρο, για να μπορέσει να γίνει μέτρο πρέπει να αποκτήσει και την συμπληρωματική ιδιότητα

$$\mu^* \left(\bigcup_{i=1}^{\infty} E_i \right) = \sum_{i=1}^{\infty} \mu^*(E_i)$$

αν ισχύει $E_i \cap E_j = \emptyset$.

Αυτό δεν ισχύει για όλα τα $E_i \in \mathcal{F}$ αλλά μόνο για ορισμένα απο αυτά τα οποία ικανοποιούν την ιδιότητα της μετρήσιμότητας.

Ορισμός 4.3.2 *Ένα σύνολο $A \in \mathcal{F}$ ονομάζεται μετρήσιμο ως προς το εξωτερικό μέτρο μ^* αν ικανοποιεί την ιδιότητα*

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

για κάθε $E \in \mathcal{F}$. Το σύνολο των μετρήσιμων συνόλων ως προς το εξωτερικό μέτρο μ^* θα συμβολίζεται ως $\mathcal{M}$.

Ορισμός 4.3.3 *Ένα μέτρο ονομάζεται* **πλήρες** *αν κάθε μηδενοσύνολο, και συνεπώς και κάθε υποσύνολο αυτού, είναι μετρήσιμα σύνολα.*

Πρόταση 4.3.1 *Το $\mathcal{M}$ είναι μία $\sigma-$ άλγεβρα. Επιπλέον, το μ^* περιορισμένο στο $\mathcal{M}$ είναι μέτρο. Το μέτρο που επάγεται απο το μ^* είναι πλήρες, δηλαδή κάθε E για το οποίο $\mu^*(E) = 0$ ανήκει στο $\mathcal{M}$.*

Κεφάλαιο 5

Το ολοκλήρωμα κατά *Lebesgue*

5.1 Εισαγωγή

Στην ενότητα αυτή θα εισάγουμε το ολοκλήρωμα του Lebesgue . Η θεωρία της ολοκλήρωσης του Lebesgue χρησιμοποιεί την έννοια του μέτρου που ορίσαμε στις προηγούμενες ενότητες, και αποτελεί γενίκευση της θεωρίας της ολοκλήρωσης του Riemann παρακάμπτωντας αρκετές από τις δυσκολίες της. Αν θεωρήσουμε ότι οι χώροι μέτρου στους οποίους εργαζόμαστε είναι χώροι πιθανοτήτων, τότε η θεωρία της ολοκλήρωσης του Lebesgue μας δίνει μια πολύ καλό θεωρητικό πλαίσιο για την μελέτη των ροπών τυχαίων μεταβλητών καθώς και των ροπών ακολουθίων τυχαίων μεταβλητών και των ορίων τους.

5.2 Μετρήσιμες συναρτήσεις

Ορισμός 5.2.1 *Έστω $(\mathbb{X}, \mathcal{F}, \mu)$ ένας χώρος μέτρου και $f : \mathbb{X} \to \mathbb{R}$ μία συνάρτηση. Η συνάρτηση f ονομάζεται μετρήσιμη αν $f^{-1}(G) \in \mathcal{F}$ για κάθε ανοικτό σύνολο G του $\mathbb{R}$.*

Σχόλιο 5.2.1 (1) Θα μπορούσαμε ισοδύναμα να γράψουμε στον ορισμο 5.2.1, για κάθε $G \in \mathcal{B}(\mathbb{R})$.
(2) Η μετρησιμότητα δεν εξαρτάται απο το μέτρο μ αλλά μόνο απο την $\sigma-$άλγεβρα $\mathcal{F}$.

Παράδειγμα 5.2.1 *Η σταθερή συνάρτηση $f : \mathbb{X} \to \mathbb{R}$, $f(x) = c$ για κάθε $x \in \mathbb{X}$ είναι μετρήσιμη.*

Παράδειγμα 5.2.2 *Ας πάρουμε την $\sigma-$άλγεβρα $\mathcal{F} = \{\mathbb{X}, \emptyset, E, E^c\}$ όπου $E \subset \mathbb{X}$ και την συνάρτηση*

$$f(x) = \mathbf{1}_G(x) := \begin{cases} 1 & \text{αν} \quad x \in G \\ 0 & \text{αν} \quad x \notin G \end{cases}$$

όπου $G \subset \mathbb{X}$. Η συνάρτηση f είναι μετρήσιμη ως προς την $\mathcal{F}$ εάν $F = E$ και μη μετρήσιμη ως προς αυτή σε κάθε άλλη περίπτωση.

Πρόταση 5.2.1 *Οι συνεχείς συναρτήσεις είναι και μετρήσιμες ως προς την $\sigma-$άλγεβρα $\mathcal{B}(\mathbb{X})$ που περιέχει όλα τα ανοιχτά υποσύνολα του $\mathbb{X}$.*

Απόδειξη: Απο την συνέχεια της f, έχουμε ότι $f^{-1}(G)$ είναι ανοιχτό υποσύνολο του $\mathbb{X}$ αν G είναι ανοιχτό υποσύνολο του $\mathbb{R}$. Τότε εξ ορισμού, για κάθε ανοιχτό υποσύνολο G του $\mathbb{R}$ το $f^{-1}(G)$ ανήκει στο $\mathcal{B}(\mathbb{X})$ συνεπώς η f είναι μετρήσιμη. ∎

Πρόταση 5.2.2 *Έστω f, g μετρήσιμες συναρτήσεις $f, g : \mathbb{X} \to \mathbb{R}$. Τότε και οι συναρτήσεις λf, $f + g$, fg, $max(f, g)$, $min(f, g)$ είναι επίσης μετρήσιμες.*

Μια ειδική περίπτωση της έννοιας της μετρήσιμης συνάρτησης είναι και η έννοια της τυχαίας μεταβλητής. Θα χρησιμοποιούμε τον συμβολισμό X για τις τυχαίες μεταβλητές αντί του συμβολισμού f που χρησιμοποιούμε για τις συναρτήσεις. Η τιμή που θα πάρει μια τυχαία μεταβλητή για κάποιο ενδεχομενο ω θα συμβολίζεται $X(\omega)$.

Ορισμός 5.2.2 *Έστω $(\Omega, \mathcal{F}, \mu)$ ένας χώρος πιθανοτήτων και $X : \Omega \to \mathbb{R}$ μία συνάρτηση η οποία είναι μετρήσιμη ως προς την $\mathcal{F}$. Τότε η X ονομάζεται* **τυχαία μεταβλητή**.

Για μία τυχαία μεταβλητή η μετρησιμότητα μας λέει σε ποιά υποσύνολα του Ω ζούν οι αντίστροφες εικόνες της X, ή με άλλα λόγια, αν θέλουμε να απαντήσουμε ναι ή όχι στην ερώτηση η τυχαία μεταβλητή X παίρνει τιμές στο σύνολο $A^{'} \subset \mathbb{R}$ ποιά γεγονότα θα πρέπει να γνωρίζουμε.

Θα χρειαστούμε και την έννοια της ακολουθίας συναρτήσεων ή τυχαίων μεταβλητών.

Ορισμός 5.2.3 *Η οικογένεια συναρτήσεων f_n, $f_n : \mathbb{X} \to \mathbb{R}$, $n \in \mathbb{N}$ ονομάζεται* **ακολουθία συναρτήσεων**. *Για κάθε δεδομένο $n = n_1$, έχουμε και μια (διαφορετική) συνάρτηση την f_{n_1}. Με $f_{n_1}(x)$ θα συμβολίζουμε την τιμή που παίρνει το μέλος της οικογένειας που αντιστοιχεί στην επιλογή $n = n_1$ για το $x \in \mathbb{X}$.*

Παράδειγμα 5.2.3 *Έστω $\mathbb{X} = [-\pi, \pi]$. Οι συναρτήσεις $f_n : [-\pi, \pi] \to \mathbb{R}$ που ορίζονται απο $f_n(x) = \sin(n\,x)$ είναι ένα παράδειγμα μιας ακολουθίας συναρτήσεων.*

Παράδειγμα 5.2.4 *Έστω $\mathbb{X} = [0, 1]$. Οι συναρτήσεις $f_n : [0, 1] \to \mathbb{R}$ που ορίζονται απο $f_n(x) = x^n$ είναι ένα παράδειγμα μιας ακολουθίας συναρτήσεων.*

Ορισμός 5.2.4 *Έστω $f_n : \mathbb{X} \to \mathbb{R}$ μια ακολουθία συναρτήσεων. Θα λέμε ότι $f_n \to f$ αν για κάθε $x \in \mathbb{X}$, $f_n(x) \to f(x)$ (με την έννοια της σύγκλισης στο $\mathbb{R}$). Δηλαδή, για κάθε $\epsilon > 0$, υπάρχει $N \in \mathbb{N}$ τέτοιο ώστε $\mid f_n(x) - f(x) \mid < \epsilon$, για $n > N$. Η ανισότητα αυτή ισχύει για κάθε $x \in \mathbb{X}$ αλλά το N μπορει να εξαρτάται απο το x.*

Παράδειγμα 5.2.5 *Έστω $\mathbb{X} = [0, 1]$, $f_n, f : [0, 1] \to \mathbb{R}$, $f_n(x) = x^n$ και*

$$f(x) = \begin{cases} 0, & x \in [0, 1) \\ 1, & x = 1 \end{cases}$$

Σύμφωνα με τον ορισμό 5.2.4 ισχύει ότι $f_n \to f$.

Η έννοια της σύγκλισης όπως δίνεται απο τον ορισμό 5.2.4 είναι πολύ περιοριστικη! Θα δούμε στα πλαίσια της θεωρίας μέτρου εναλλακτικούς ορισμούς σχετικά με την σύγκλιση που ξεπερνούν τα προβλήματα που σχετίζονται με τον ορισμό αυτό και μπορεί να μας φανούν πολύ χρήσιμοι σε διάφορες εφαρμογές.

Με παρόμοιο τρόπο μπορούμε να ορίσουμε και την έννοια των ακολουθίων τυχαίων μεταβλητών.

Ορισμός 5.2.5 *Η οικογένεια τυχαίων μεταβλητών X_n, $X_n : \Omega \to \mathbb{R}$, $n \in \mathbb{N}$ ονομάζεται* **ακολουθία τυχαίων μεταβλητών**. *Για κάθε δεδομένο $n = n_1$, έχουμε και μια (διαφορετική) τυχαία μεταβλητή την X_{n_1}. Με $X_{n_1}(\omega)$ θα συμβολίζουμε την τιμή που παίρνει το μέλος της οικογένειας που αντιστοιχεί στην επιλογή $n = n_1$ για το $\omega \in \Omega$, δηλαδή την πραγματοποίηση της τυχαίας μεταβλητής X_n αν συμβεί το γεγονός ω.*

Ο ορισμός 5.2.4 μπορεί να μετατραπεί άμεσα σε ένα ισοδύναμο ορισμό για την σύγκλιση ακολουθίων τυχαίων μεταβλητών αλλά και στην περίπτωση αυτή ο ορισμός αυτός είναι πολύ περιοριστικός και θα χρειαστεί γενικεύσεις για να μας φανεί χρήσιμος στις πιθανότητες και την στατιστική.

Αν υποθέσουμε οτι μια ακολουθία συναρτήσεων f_n είναι μετρήσιμη τι συμβαίνει με το όριο της όταν αυτό υπάρχει; Είναι και αυτό μια μετρήσιμη συνάρτηση; Όμοιως, αν υποθέσουμε ότι μια ακολουθία τυχαίων μεταβλητών X_n είναι μετρήσιμη ως προς μια σ-άλγεβρα $\mathcal{F}$ και υπάρχει το όριο X της καθώς $n \to \infty$, είναι και η X τυχαία μεταβλητή;

Για να απαντήσουμε το ερώτημα αυτό θα πρέπει να εκφράσουμε τους παραπάνω ορισμούς σχετικά με την σύγκλιση ακολουθίων συναρτήσεων ή τυχαίων μεταβλητών με λίγο διαφορετικό (αλλά ισοδύναμο) τρόπο.

Ορισμός 5.2.6 *Έστω $f_n : \mathbb{X} \to \mathbb{R}$ μια ακολουθία συναρτήσεων. Μπορούμε να ορίσουμε ως*

$$\limsup_{n\to\infty} f_n = \inf_{n\geq 1} \left(\sup_{m\geq n} f_m \right)$$

$$\liminf_{n\to\infty} f_n = \sup_{n\geq 1} \left(\inf_{m\geq n} f_m \right)$$

Αν οι δύο παραπάνω συναρτήσεις συμπίπτουν τοτε λέμε ότι υπάρχει το όριο της ακολουθίας συναρτήσεων

$$\lim_{n\to\infty} f_n = \limsup_{n\to\infty} f_n = \liminf_{n\to\infty} f_n \tag{5.1}$$

Ο ορισμός αυτός είναι ισοδύναμος με τον ορισμό 5.2.4.

Σχόλιο 5.2.2 Οι συναρτήσεις $\limsup_n f_n$ και $\liminf_n f_n$ ορίζονται **σημειακά**, δηλαδή είναι η συνάρτηση που η τιμή της για κάθε $x \in \mathbb{X}$ ορίζεται απο τις σχέσεις

$$\limsup_{n\to\infty} f_n(x) = \inf_{n\geq 1}\left(\sup_{m\geq n} f_m(x)\right)$$
$$\liminf_{n\to\infty} f_n(x) = \sup_{n\geq 1}\left(\inf_{m\geq n} f_m(x)\right)$$

Με όμοιο τρόπο μπορούμε να εργαστούμε και για τις ακολουθίες τυχαίων μεταβλητών.

Ορισμός 5.2.7 *Έστω $X_n : \Omega \to \mathbb{R}$ μια ακολουθία τυχαίων μεταβλητών. Μπορούμε να ορίσουμε ως*

$$\limsup_{n\to\infty} X_n = \inf_{n\geq 1}\left(\sup_{m\geq n} X_m\right)$$
$$\liminf_{n\to\infty} X_n = \sup_{n\geq 1}\left(\inf_{m\geq n} X_m\right)$$

Αν οι δύο παραπάνω τυχαίες μεταβλητές συμπίπτουν τοτε λέμε ότι υπάρχει το όριο της ακολουθίας τυχαίων μεταβλητών

$$\lim_{n\to\infty} X_n = \limsup_{n\to\infty} X_n = \liminf_{n\to\infty} X_n \tag{5.2}$$

Ο ορισμός αυτός είναι ισοδύναμος με τον αντίστοιχο ορισμό 5.2.4 όπως βέβαια αυτός θα πρέπει να γραφτεί κατάλληλα για τυχαίες μεταβλητές.

Σχόλιο 5.2.3 Οι τυχαίες μεταβλητές $\limsup_n X_n$ και $\liminf_n X_n$ ορίζονται **σημειακά**, δηλαδή είναι η συνάρτηση που η τιμή της για κάθε $\omega \in \Omega$ ορίζεται απο τις σχέσεις

$$\limsup_{n\to\infty} X_n(\omega) = \inf_{n\geq 1}\left(\sup_{m\geq n} X_m(\omega)\right)$$
$$\liminf_{n\to\infty} X_n(\omega) = \sup_{n\geq 1}\left(\inf_{m\geq n} X_m(\omega)\right)$$

Έχουμε τώρα το ακόλουθο πολύ χρήσιμο αποτέλεσμα.

Πρόταση 5.2.3

1. *Έστω $f_n : \mathbb{X} \to \mathbb{R}$ μία ακολουθία συναρτήσεων η οποία είναι μετρήσιμη ως προς την σ-άλγεβρα $\mathcal{F}$. Τότε $\limsup_n f_n$ και $\liminf_n f_n$ είναι μετρήσιμες συναρτήσεις ως προς την $\mathcal{F}$. Όμοια και το $\lim_n f_n$ αν υπάρχει.*
2. *Έστω $X_n : \Omega \to \mathbb{R}$ μία ακολουθία τυχαίων μεταβλητών η οποία είναι μετρήσιμη ως προς την σ-άλγεβρα $\mathcal{F}$. Τότε $\limsup_n X_n$ και $\liminf_n X_n$ είναι μετρήσιμες τυχαίες μεταβλητές ως προς την $\mathcal{F}$. Όμοια και το $\lim_n X_n$ αν υπάρχει.*

Απόδειξη: (1) Μπορεί κανείς να παρατηρήσει ότι αν G ένα ανοιχτό υποσύνολο του $\mathbb{R}$ έχουμε ότι

$$\left\{x \in \mathbb{X} \ \mid \ \left(\inf_{n\geq 1} f_n(x)\right) \in G\right\} = \bigcap_{n=1}^{\infty}\{x \in \mathbb{X} \ \mid \ f_n(x) \in G\}$$
$$\left\{x \in \mathbb{X} \ \mid \ \left(\sup_{n\geq 1} f_n(x)\right) \in G\right\} = \bigcup_{n=1}^{\infty}\{x \in \mathbb{X} \ \mid \ f_n(x) \in G\}$$

Επειδή f_n μετρήσιμες $A_n := \{x \in \mathbb{X} \ | \ f_n(x) \in G\} = f_n^{-1}(G) \in \mathcal{F}$ για κάθε n οπότε επειδή $\mathcal{F}$ είναι σ-άλγεβρα θα ισχύει ότι $\bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$ και $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$. Όμως $\bigcap_{n=1}^{\infty} A_n = (\inf_{n\geq 1} f_n)^{-1}(G)$ και $\bigcup_{n=1}^{\infty} A_n = (\sup_{n\geq 1} f_n)^{-1}(G)$. Αυτό μας εξασφαλίζει το ότι οι συναρτήσεις $\inf_n f_n$ και $\sup_n f_n$ είναι μετρήσιμες. Αν εργαστούμε με παρόμοιο τρόπο εξασφαλίζουμε και την μετρησιμότητα των συναρτήσεων $\liminf_n f_n$ και $\limsup_n f_n$. Η περίπτωση των τυχαίων μεταβλητών μπορεί να αποδειχθεί παρομοίως. ■

Για την μελέτη της σύγκλισης ακολουθίων συναρτήσεων ή τυχαίων μεταβλητών δεν είναι πάντοτε απαραίτητο να θεωρήσουμε ότι τα $\limsup_n f_n$, $\liminf_n f_n$ ή $\limsup_n X_n$, $\liminf_n X_n$ αντίστοιχα, ορίζονται σημειακά δηλαδή για κάθε $x \in \mathbb{X}$ ή κάθε $\omega \in \Omega$. Μπορούμε να 'αφαιρέσουμε' απο τον σημειακό ορισμό κάποια x ή ω αρκεί να ισχύει το ότι αν G είναι το σύνολο των x ή των ω που δεν λαμβάνουμε υπόψιν, τότε $\mu(G) = 0$. Απο τον ορισμό του μέτρου, γνωρίζουμε ότι τα x η τα ω αυτά μπορεί να είναι άπειρα το πλήθος! Αν αντι του σημειακού ορισμού υιοθετήσουμε την παραπάνω έννοια, δηλαδή ορίσουμε τα $\liminf_n f_n$, $\limsup_n f_n$ για κάθε $x \in \mathbb{X} \setminus G$, $\mu(G) = 0$, τότε θα μιλαμε για το όριο της ακολουθίας συναρτήσεων σχεδόν παντού (ς.π) ως προς το μέτρο μ. Στην περίπτωση των ακολουθιών τυχαίων μεταβλητών αν αντι του σημειακού ορισμού υιοθετήσουμε την παραπάνω έννοια, δηλαδή ορίσουμε τα $\liminf_n X_n$, $\limsup_n X_n$ για κάθε $\omega \in \Omega \setminus G$, $\mu(G) = 0$, τότε θα μιλαμε για το όριο της ακολουθίας τυχαίων μεταβλητών σχεδόν βέβαια (ς.β) ως προς το μέτρο μ.

Εν γένει μπορούμε να δώσουμε τον ακόλουθο ορισμό.

Ορισμός 5.2.8 *1. Έστω $(\mathbb{X}, \mathcal{F}, \mu)$ χώρος μέτρου. Θα λέμε ότι η ιδιότητα C ισχύει σχεδόν παντού ως προς το μ αν*

$$\mu(\{x \in \mathbb{X}, \mid C \text{ δεν ισχύει}\}) = 0$$

2. Έστω $(\Omega, \mathcal{F}, \mu)$ χώρος πιθανότητων. Θα λέμε ότι η ιδιότητα C ισχύει σχεδόν βέβαια ως προς το μ αν

$$\mu(\{\omega \in \Omega, \mid C \text{ δεν ισχύει}\}) = 0$$

Οι ιδιότητες της μετρησιμότητας των ορίων ισχύουν ακόμα και όταν παίρνουμε την έννοια του ορίου σχεδόν παντού ή σχεδόν βέβαια.

5.3 Απλές συναρτήσεις και απλές τυχαίες μεταβλητές

Ορισμός 5.3.1 *Έστω $(\mathbb{X}, \mathcal{F}, \mu)$ ένας χώρος μέτρου. Η συνάρτηση $f : \mathbb{X} \to \mathbb{R}$ ονομάζεται* **απλή συνάρτηση** *αν μπορεί να γραφεί*

$$f(x) = \sum_{i=1}^{n} c_i \, \mathbf{1}_{A_i}(x) \tag{5.3}$$

για πεπερασμένο το πλήθος n, $c_i \in \mathbb{R}$ και $A_i \subset \mathbb{X}$ όπου $A_i \in \mathcal{F}$.

Μια απλή συνάρτηση είναι λοιπόν μια συνάρτηση η οποία παίρνει διακριτές και πεπερασμένες το πλήθος τιμές. Το σύνολο τιμών της απλής συνάρτησης (5.3) είναι το σύνολο $R(f) = \{c_1, \cdots, c_n\}$. Είναι προφανές ότι μια απλή συνάρτηση είναι μετρήσιμη ως προς την $\mathcal{F}$. Η μικρότερη όμως σ-άλγεβρα η οποία κάνει μετρήσιμη την απλή συνάρτηση (5.3) είναι η $\sigma(A_1, \cdots, A_n)$ δηλαδή η σ-άλγεβρα η οποία παράγεται απο τα σύνολα $A_1, \cdots, A_n \in \mathcal{F}$.

Στην ειδική περίπτωση που ο χώρος μέτρου είναι ένας χώρος πιθανοτήτων η έννοια της απλής συνάρτησης μπορεί να μεταφερθεί στην έννοια της απλής τυχαίας μεταβλητής.

Ορισμός 5.3.2 *Έστω $(\Omega, \mathcal{F}, \mu)$ ένας χώρος χώρος πιθανοτήτων. Η τυχαία μεταβλητή X ονομάζεται* **απλή τυχαία μεταβλητή** *αν μπορεί να γραφεί στην μορφή*

$$X = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i} \tag{5.4}$$

για $A_i \in \mathcal{F}$ και n πεπερασμένο.

Το σύνολο των τιμών της τυχαίας μεταβλητής (5.4) είναι το πεπερασμένο σύνολο $\{a_1, \cdots, a_n\}$. Υπο μία έννοια οι απλές τυχαίες μεταβλητές είναι διακριτές τυχαίες μεταβλητές. Η απλή τυχαία μεταβλητή (5.4) είναι μετρήσιμη ως προς την σ-άλγεβρα $\mathcal{F}$ αλλά η μικρότερη σ-άλγεβρα που καθιστά την τυχαία μεταβλητή αυτή μετρήσιμη είναι η $\sigma(A_1, \cdots, A_n)$ δηλαδή η σ-άλγεβρα που παράγεται από τα ενδεχόμενα $A_1, \cdots, A_n$.

Παράδειγμα 5.3.1 *Έστω X η τυχαία μεταβλητή που αντιστοιχεί στο κέρδος απο το παιχνίδι στην ρίψη δύο νομισμάτων αν ποντάρουμε 1 ευρώ στο αν θα έρθει κορώνα ή όχι. Το κέρδος είναι μία απλή τυχαία μεταβλητή που μπορεί να γραφεί στην μορφή*

$$X(\omega) = 2 \; \mathbf{1}_{\{H_1H_2\}}(\omega) + 0 \; \mathbf{1}_{\{H_1T_2, T_1H_2\}}(\omega) + (-2) \; \mathbf{1}_{\{T_1T_2\}}(\omega)$$

Φυσικά, μπορώ να ορίσω απλές συναρτήσεις (τυχαίες μεταβλητές) ακόμα και σε περιπτώσεις όπου το Ω δεν είναι διακριτό.

Παράδειγμα 5.3.2 *Έστω $\Omega = [0,1]$, $\mathcal{F} = \mathcal{B}([0,1])$. Η τυχαία μεταβλητή η οποία παίρνει την τιμή 10 αν $\omega \in [0, 0.1)$, την τιμή 5 αν $\omega \in [0.1, 0.5)$ και την τιμή 0 αν $\omega \in [0.5, 1]$ είναι μία απλή τυχαία μεταβλητή της μορφής*

$$X(\omega) = 10\,\mathbf{1}_{[0,0.1)}(\omega) + 5\,\mathbf{1}_{[0.1,0.5)}(\omega)$$

5.4 Το ολοκλήρωμα για απλές συναρτήσεις ή απλές τυχαίες μεταβλητές

Θα εισάγουμε αρχικά το ολοκλήρωμα για απλές συναρτήσεις.

Ορισμός 5.4.1 *Έστω $(\mathbb{X}, \mathcal{F}, \mu)$ ένας χώρος μέτρου και $f : \mathbb{X} \to \mathbb{R}$ μια απλή συνάρτηση, δηλαδή μια συνάρτηση της μορφής (5.3). Ορίζουμε*

$$\int_{\mathbb{X}} f\,d\mu := \sum_{i=1}^{n} c_i\,\mu(A_i)$$

Η ποσότητα $\int_{\mathbb{X}} f\,d\mu$ ονομάζεται το **ολοκλήρωμα** *της f επάνω στο $\mathbb{X}$ ως προς το μέτρο μ.*

Στην περίπτωση των τυχαίων μεταβλητών ο παραπάνω ορισμός παίρνει την ακόλουθη μορφή.

Ορισμός 5.4.2 *Έστω $(\Omega, \mathcal{F}, \mu)$ ένας χώρος πιθανοτήτων και X μία απλή $\mathcal{F}$-μετρήσιμη συνάρτηση (τυχαία μεταβλητή). Αν*

$$X = \sum_{i=1}^{n} a_i\,\mathbf{1}_{A_i}, \ \ A_i \in \mathcal{F}$$

το **ολοκλήρωμα της** *X* **ως προς το μέτρο** *μ (η μέση τιμή της X ως προς το μέτρο μ) ορίζεται ως*

$$\int_{\Omega} X\,d\mu := \sum_{i=1}^{n} a_i\,\mu(A_i)$$

Το ολοκλήρωμα αυτό συμβολίζεται και ως $\mathbb{E}_\mu[X]$.

Είναι πολυ σημαντικό να παρατηρήσουμε ότι αν εκφράσουμε την ίδια συνάρτηση (τυχαία μεταβλητή) X με διαφορετικό τρόπο, το ολοκλήρωμα της **δεν** μεταβάλλεται!

Πρόταση 5.4.1 *Το ολοκλήρωμα μιας απλης συνάρτησης ή μιας απλής τυχαίας μεταβλητής είναι ανεξάρτητο της αναπαράστασης της.*

1. *Έστω $(\mathbb{X}, \mathcal{F}, \mu)$ χώρος μέτρου και $f = \sum_{i=1}^{n} c_i\,\mathbf{1}_{A_i} = \sum_{i=1}^{m} d_i\,\mathbf{1}_{B_i}$ μία απλή συνάρτηση. Τότε*

$$\int_{\mathbb{X}} f\,d\mu = \sum_{i=1}^{n} c_i\,\mu(A_i) = \sum_{i=1}^{m} d_i\,\mu(B_i)$$

2. *Έστω $(\Omega, \mathcal{F}, \mu)$ χώρος πιθανότητων και $X = \sum_{i=1}^{n} a_i\,\mathbf{1}_{A_i} = \sum_{i=1}^{m} b_i\,\mathbf{1}_{B_i}$ μία απλή τυχαία μεταβλητή. Τότε*

$$\mathbb{E}_\mu[X] = \int_{\Omega} X\,d\mu = \sum_{i=1}^{n} a_i\,\mu(A_i) = \sum_{i=1}^{m} b_i\,\mu(B_i)$$

Πολλές φορές επίσης μπορεί να χρειαστεί να ορίσουμε το ολοκλήρωμα μιας απλης συνάρτησης f (τυχαίας μεταβλητής X) σε ένα υποσύνολο A του $\mathbb{X}$ (ή Ω αντίστοιχα). Αυτό γίνεται ως εξής

$$\int_{A} f\,d\mu = \int_{\mathbb{X}} f\,\mathbf{1}_A\,d\mu$$

ή

$$\int_A X\, d\mu := \int_\Omega X\, \mathbf{1}_A\, d\mu$$

αντίστοιχα.

Οι παρακάτω ιδιοτητες του ολοκληρώματος (μέσης τιμής) για απλές συναρτήσεις (τυχαίες μεταβλητές) είναι πολύ σημαντικές

Πρόταση 5.4.2 *Έστω $(\mathbb{X}, \mathcal{F}, \mu)$ χώρος μέτρου και f, g απλές συναρτήσεις, $f, g : \mathbb{X} \to \mathbb{R}$.*

1. *Το ολοκλήρωμα είναι γραμμικό δηλαδή αν f, g απλές συναρτήσεις και $\lambda_1, \lambda_2 \in \mathbb{R}$ τότε*

$$\int_{\mathbb{X}} (\lambda_1 f + \lambda_2 g)\, d\mu = \lambda_1 \int_{\mathbb{X}} f\, d\mu + \lambda_2 \int_{\mathbb{X}} g\, d\mu$$

2. *Το ολοκλήρωμα είναι θετικός τελεστής, δηλαδή αν f, g απλές συναρτήσεις (τυχαίες μεταβλητές) τέτοιες ώστε $f \geq g$ τότε ισχύει*

$$\int_{\mathbb{X}} f\, d\mu \geq \int_{\mathbb{X}} g\, d\mu$$

Σχόλιο 5.4.1 Ας ονομάσουμε $\mathbb{F}$ το σύνολο όλων των απλών συναρτήσεων απο το $\mathbb{X} \to \mathbb{R}$,

$$\mathbb{F} = \{ f : \mathbb{X} \to \mathbb{R} \;\mid\; f \text{ της μορφής } (5.3) \}$$

Μπορούμε να θεωρήσουμε το ολοκλήρωμα σαν την απεικόνιση $I : \mathbb{F} \to \mathbb{R}$ η οποία ορίζεται ως

$$I(f) = \int_{\mathbb{X}} f\, d\mu = \sum_{i=1}^{n} c_i\, \mu(A_i)$$

για κάθε $f \in \mathbb{F}$. Η πρόταση 5.4.2 μας λέει οτι η απεικόνιση I είναι
(i) γραμμική απεικόνιση, δηλαδή ότι $I(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 I(f_1) + \lambda_2 I(f_2)$ για κάθε $\lambda_1, \lambda_2 \in \mathbb{R}$ και $f_1, f_2 \in \mathbb{F}$ και (ii) θετική απεικόνιση (τελεστής) δηλαδή ότι αν η $f \in \mathbb{F}$ είναι τέτοια ώστε $f \geq 0$ τότε $I(f) \geq 0$ (με την συνήθη έννοια της ανισότητας στους πραγματικούς αριθμούς). Επειδή η f είναι μια συνάρτηση $f : \mathbb{X} \to \mathbb{R}$ με τον συμβολισμό $f \geq 0$ εννοούμε ότι $f(x) \geq 0$ (με την έννοια της ανισότητας στους πραγματικούς αριθμούς) για κάθε $x \in \mathbb{X}$.

Πρόταση 5.4.3 *Έστω $(\Omega, \mathcal{F}, \mu)$ χώρος πιθανοτήτων και X, Y απλές τυχαίες μεταβλητές, $X, Y : \Omega \to \mathbb{R}$.*

1. *Το ολοκλήρωμα (μέση τιμή) είναι γραμμικό δηλαδή αν X, Y απλές τυχαίες μεταβλητές τότε για κάθε $\lambda_1, \lambda_2 \in \mathbb{R}$ ισχύει ότι*

$$\int_\Omega (\lambda_1 X + \lambda_2 Y)\, d\mu = \lambda_1 \int_\Omega X\, d\mu + \lambda_2 \int_\Omega Y\, d\mu$$

ή ισοδύναμα

$$\mathbb{E}_\mu[\lambda_1 X + \lambda_2 Y] = \lambda_1 \mathbb{E}_\mu[X] + \lambda_2 \mathbb{E}_\mu[Y].$$

2. *Το ολοκλήρωμα (μέση τιμή) είναι θετικός τελεστής, δηλαδή αν X, Y απλές τυχαίες μεταβλητές τέτοιες ώστε $f \geq g$ ισχύει*

$$\int_\Omega X\, d\mu \geq \int_\Omega Y\, d\mu$$

ή ισοδύναμα

$$\mathbb{E}_\mu[X] \geq \mathbb{E}_\mu[Y].$$

Σχόλιο 5.4.2 Ας ονομάσουμε $\mathcal{L}$ το σύνολο όλων των απλών τυχαίων μεταβλητών απο το $\Omega \to \mathbb{R}$,

$$\mathcal{L} = \{X : \Omega \to \mathbb{R} \ | \ X \text{ της μορφής } (5.4)\}$$

Μπορούμε να θεωρήσουμε το ολοκλήρωμα σαν την απεικόνιση $I : \mathcal{L} \to \mathbb{R}$ η οποία ορίζεται ως

$$I(X) := \int_\Omega X\, d\mu = \mathbb{E}_\mu[X] = \sum_{i=1}^{n} a_i\, \mu(A_i)$$

για κάθε $X \in \mathcal{L}$. Η πρόταση 5.4.3 μας λέει οτι η απεικόνιση I είναι
(i) γραμμική απεικόνιση, δηλαδή ότι $I(\lambda_1 X_1 + \lambda_2 X_2) = \lambda_1 I(X_1) + \lambda_2 I(X_2)$ για κάθε $\lambda_1, \lambda_2 \in \mathbb{R}$ και $X_1, X_2 \in \mathcal{L}$ και
(ii) θετική απεικόνιση (τελεστής) δηλαδή ότι αν η $X \in \mathcal{L}$ είναι τέτοια ώστε $X \geq 0$ τότε $I(X) \geq 0$ (με την συνήθη έννοια της ανισότητας στους πραγματικούς αριθμούς). Επειδή η X είναι μια τυχαία μεταβλητή $X : \Omega \to \mathbb{R}$ με τον συμβολισμό $X \geq 0$ εννοούμε ότι $X(\omega) \geq 0$ (με την έννοια της ανισότητας στους πραγματικούς αριθμούς) για κάθε $\omega \in \Omega$, δηλαδή ότι για κάθε ενδεχόμενο $\omega \in \Omega$ η πραγματοποίηση $X(\omega)$ της τυχαίας μεταβλητής X είναι θετικός πραγματικός αριθμός.

Σχόλιο 5.4.3 Επειδή οι ορισμοί και οι ιδιότητες του ολοκλήρώματος για (απλές) συναρτήσεις $f : \mathbb{X} \to \mathbb{R}$ και για (απλές) τυχαίες μεταβλητές $X : \Omega \to \mathbb{R}$ είναι παρόμοιοι, για οικονομία χώρου θα δώσουμε παρακάτω τις ιδιότητες και τους ορισμούς στην περίπτωση των τυχαίων μεταβλητών, με την έννοια ότι αυτά που παρουσιάζουμε γενικεύονται και στην περίπτωση των συναρτήσεων. Συνεπώς απο εδώ και κάτω θα χρησιμοποιούμε τον συμβολισμό $\mathbb{X} = \Omega$, $f = X$.

5.5 Το ολοκλήρωμα για θετικές συναρτήσεις ή τυχαίες μεταβλητές

Ας υποθέσουμε ότι $f : \mathbb{X} \to \mathbb{R}$ είναι μια θετική συνάρτηση η οποία δεν είναι απλή. Η στρατηγική μας για να ορίσουμε το ολοκλήρωμα $\int_\mathbb{X} f\, d\mu$ είναι να χρησιμοποιήσουμε την προσέγγιση της f απο απλές συναρτήσεις.

Θα δώσουμε δύο διαφορετικούς αλλά ισοδύναμους ορισμούς.

Ορισμός 5.5.1 *(I) Ας θεωρήσουμε το σύνολο $\mathbb{F}_f$ όλων των απλών συναρτήσεων g για τις οποίες ισχύει $0 \leq g \leq f$, υπό την έννοια ότι σχεδόν παντού $0 \leq g(x) \leq f(x)$. Ας ορισουμε σαν I_f το σύνολο των ολοκληρωμάτων αυτών των g επάνω στο μέτρο μ. Συνεπώς,*

$$\begin{aligned} \mathbb{F}_f &= \{g \in \mathbb{F} \ | \ 0 \leq g \leq f\} \subset \mathbb{F}, \\ I_f &= \{\int_\mathbb{X} g\, d\mu \ | \ g \in \mathbb{F}_f\} \subset \mathbb{R} \end{aligned}$$

Ορίζουμε σαν το ολοκλήρωμα της f επάνω στο μέτρο μ το $\sup$ του συνόλου I_f, δηλαδή

$$\int_\mathbb{X} f\, d\mu := \sup I_f$$

Σχόλιο 5.5.1 *Παρατηρείστε οτι το σύνολο I_f ορίζεται πολύ καλά εφόσον ο ορισμός του χρειάζεται μόνο το ολοκλήρωμα απλών συναρτήσεων το οποίο ορίζεται ως γνωστό με την βοήθεια πεπερασμένων αθροισμάτων. Επίσης, επιτρέπουμε το ολοκλήρωμα να παίρνει και την τιμή ∞, άρα το sup του συνόλου I_f υπάρχει πάντοτε.*

Ορισμός 5.5.2 *(II) Ας θεωρησουμε $\{f_n\}$ μία ακολουθία απλών συναρτήσεων που συγκλινει μονότονα στην f, $f_n \uparrow f$. Παίρνουμε και την ακολουθία πραγματικών αριθμών $r_n := \int_\mathbb{X} f_n\, d\mu$ η οποία είναι καλά ορισμένη. Σαν ολοκλήρωμα της f επάνω στο μέτρο μ ορίζουμε το όριο της ακολουθίας r_n, δηλαδή*

$$\int_\mathbb{X} f\, d\mu := \lim_{n \to \infty} \int_\mathbb{X} f_n\, d\mu$$

όπου και πάλι επιτρέπουμε την προοπτική να είναι το όριο αυτο άπειρο.

Σχόλιο 5.5.2 Με τον συμβολισμό $f_n \uparrow f$ εννοούμε εκτός της σύγκλισης ισχύει και ότι $f_n \leq f_{n+1}$ με την έννοια ότι $f_n(x) \leq f_{n+1}(x)$ σχεδόν παντού.

Μπορούμε πάντοτε να βρούμε μία τέτοια ακολουθία απλών συναρτήσεων για να προσεγγίσουμε οποιαδήποτε συνάρτηση. Η επιλογή της ακολουθίας **δεν** είναι μοναδική.

Παράδειγμα 5.5.1 *Έστω $\mathbb{X} = [a,b]$ και f μία συνάρτηση $f : [a,b] \to \mathbb{R}^+$. Η συνάρτηση αυτή μπορεί να προσεγγιστεί απο την ακολουθία απλών συναρτήσεων*

$$f_n(x) = \begin{cases} k\,2^{-n} & k\,2^{-n} \le f(x) < (k+1)\,2^{-n} \quad 0 \le k \le n\,2^n - 1 \\ n & f(x) \ge n \end{cases}$$

υπό την έννοια ότι $f_n \uparrow f$. Προσπαθείστε να σχεδιάσετε τις συναρτήσεις f_1, f_2, f_3 για να κατανοήσετε την έννοια με την οποία η ακολουθία f_n προσεγγίζει την f.

Μπορεί να δείξει κανείς ότι οι ορισμοί I και II είναι ισοδύναμοι.

Ας υποθέσουμε τώρα ότι X είναι μία $\mathcal{F}-$μετρήσιμη τυχαία μεταβλητή για την οποία ισχύει $X \ge 0$ αλλά η X **δεν** είναι απλή. Η στρατηγική μας για να ορίσουμε το ολοκλήρωμα (μεση τιμή) $\int_\Omega Q\,d\mu = \mathbb{E}_\mu[X]$ είναι να χρησιμοποιήσουμε την προσέγγιση της X απο απλές τυχαίες μεταβλητές.

Θα δώσουμε δύο διαφορετικούς αλλά ισοδύναμους ορισμούς.

Ορισμός 5.5.3 *(I) Ας θεωρήσουμε το σύνολο $\mathcal{L}_X$ όλων των απλών τυχαίων μεταβλητών Y για τις οποίες ισχύει $0 \le Y \le X$, και ας οριζουμε σαν L_X το σύνολο των ολοκληρωμάτων αυτών των Y επάνω στο μέτρο μ. Συνεπώς,*

$$\begin{aligned} \mathcal{L}_X &= \{Y \in \mathcal{L} \mid 0 \le Y \le X\} \subset \mathcal{L}, \\ L_X &= \{\int_\Omega Y\,d\mu \mid Y \in \mathcal{L}_X\} \subset \mathbb{R} \end{aligned}$$

Ορίζουμε σαν το ολοκλήρωμα της X επάνω στο μέτρο μ το sup του συνόλου L_X, δηλαδή

$$\mathbb{E}_\mu[X] = \int_\Omega X\,d\mu := \sup L_X$$

Σχόλιο 5.5.3 *Παρατηρείστε οτι το σύνολο L_X ορίζεται πολύ καλά εφόσον ο ορισμός του χρειάζεται μόνο το ολοκλήρωμα απλών συναρτήσεων (τυχαίων μεταβλητών) το οποίο ορίζεται ως γνωστό με την βοήθεια πεπερασμένων αθροισμάτων. Επίσης, επιτρέπουμε το ολοκλήρωμα να παίρνει και την τιμή ∞, άρα το sup του συνόλου L_X υπάρχει πάντοτε.*

Ορισμός 5.5.4 *(II) Ας θεωρησουμε $\{X_n\}$ μία ακολουθία απλών τυχαίων μεταβλητων που συγκλινει μονότονα στην X, $X_n \uparrow X$. Παίρνουμε και την ακολουθία πραγματικών αριθμών $r_n := \int_\Omega X_n\,d\mu$ η οποία είναι καλά ορισμένη. Σαν ολοκλήρωμα της X επάνω στο μέτρο μ ορίζουμε το όριο της ακολουθίας r_n, δηλαδή*

$$\mathbb{E}_\mu[X] = \int_\Omega X\,d\mu := \lim_{n\to\infty} \int_\Omega X_n\,d\mu$$

όπου και πάλι επιτρέπουμε την προοπτική να είναι το όριο αυτο άπειρο.

Σχόλιο 5.5.4 Με τον συμβολισμό $X_n \uparrow X$ εννοούμε εκτός της σύγκλισης ισχύει και ότι $X_n \le X_{n+1}$ με την έννοια ότι $X_n(\omega) \le X_{n+1}(\omega)$ σχεδόν βέβαια.

Μπορούμε πάντοτε να βρούμε μία τέτοια ακολουθία απλών τυχαίων μεταβλητών για να προσεγγίσουμε οποιαδήποτε τυχαία μεταβλητή. Η επιλογή της ακολουθίας δεν είναι μοναδική.

Παράδειγμα 5.5.2 *Έστω X μία τυχαία μεταβλητή $X : \Omega \to \mathbb{R}^+$. Η τυχαία μεταβλητή αυτή μπορεί να προσεγγιστεί απο την ακολουθία απλών τυχαίων μεταβλητών*

$$X_n(\omega) = \begin{cases} k\,2^{-n} & k\,2^{-n} \le X(\omega) < (k+1)\,2^{-n} \quad 0 \le k \le n\,2^n - 1 \\ n & X(\omega) \ge n \end{cases}$$

υπό την έννοια ότι $X_n \uparrow X$.

Οι ορισμοί I και II είναι ισοδύναμοι.

5.6 Το ολοκλήρωμα για μια οποιαδήποτε συνάρτηση ή τυχαία μεταβλητή

Έστω τώρα $f : \mathbb{X} \to \mathbb{R}$ μία συνάρτηση όχι απαραίτητα θετική. Μπορούμε να ορίσουμε τις συναρτήσεις f^+, f^- με βάση τους τύπους

$$(f^+)(x) = (f(x))^+ = max(f(x), 0) = f(x) \vee 0, \ \ \forall\, x \in \mathbb{X},$$
$$(f^-)(x) = (f(x))^- = max(-f(x), 0) = (-f(x)) \vee 0, \ \ \forall\, x \in \mathbb{X}.$$

Οι συναρτήσεις f^+, f^- ονομάζονται το θετικό και το αρνητικό μέρος, αντίστοιχα, της f. Μπορεί κανείς να δει εύκολα ότι οι συναρτήσεις f^+, f^- είναι θετικές συναρτήσεις και

$$f = f^+ - f^-$$

Ορισμός 5.6.1 *Αν $f : \mathbb{X} \to \mathbb{R}$ μια οποιαδήποτε μετρήσιμη συνάρτηση*

$$\int_{\mathbb{X}} f\, d\mu := \int_{\mathbb{X}} f^+\, d\mu - \int_{\mathbb{X}} f^-\, d\mu$$

Ορισμός 5.6.2 *Η συνάρτηση $f : \mathbb{X} \to \mathbb{R}$ ονομάζεται* **ολοκληρώσιμη** *αν $\int_{\mathbb{X}} \mid f \mid \, d\mu < \infty$.*

Με παρόμοιο τρόπο μπορούμε να ορίσουμε το ολοκλήρωμα (μέση τιμή) για οποιαδήποτε τυχαία μεταβλητή $X : \Omega \to \mathbb{R}$. Συγκεκριμένα ορίζουμε τις τυχαίες μεταβλητές X^+, X^- απο τις σχέσεις

$$(X^+)(\omega) := (X(\omega))^+ = max(X(\omega), 0) = X(\omega) \vee 0, \ \omega \in \Omega,$$
$$(X^-)(\omega) := (X(\omega))^- = max(-X(\omega), 0) = (-X(\omega)) \vee 0, \ \omega \in \Omega,$$

όπου X^+, X^- θετικές τυχαίες μεταβλητές και

$$X = X^+ - X^-$$

Ορισμός 5.6.3 *Αν $X : \Omega \to \mathbb{R}$ μια οποιαδήποτε τυχαία μεταβλητή*

$$\mathbb{E}_\mu[X] = \int_\Omega X\, d\mu := \int_\Omega X^+\, d\mu - \int_\Omega X^-\, d\mu = \mathbb{E}_\mu[X^+] - \mathbb{E}_\mu[X^-]$$

Ορισμός 5.6.4 *Η τυχαία μεταβλητή $X : \Omega \to \mathbb{R}$ ονομάζεται* **ολοκληρώσιμη** *αν $\mathbb{E}_\mu[\mid X \mid] = \int_\Omega \mid X \mid \, d\mu < \infty$.*

5.7 Ιδιότητες του ολοκληρώματος

Θα μελετήσουμε τώρα ορισμένες βασικές ιδιότητες του ολοκληρώματος επάνω σε ένα μέτρο.

5.7.1 Γενικές ιδιότητες του ολοκληρώματος

Η ακολουθη πρόταση συνοψίζει τις ιδιότητες του ολοκληρώματος.

Πρόταση 5.7.1 *Έστω $(\mathbb{X}, \mathcal{F}, \mu)$ χώρος μέτρου και $f, g : \mathbb{X} \to \mathbb{R}$ μετρήσιμες συναρτήσεις.*

1. *Το ολοκλήρωμα είναι γραμμικό, δηλαδή,*

$$\int_{\mathbb{X}} (\lambda_1\, f + \lambda_2\, g)\, d\mu = \lambda_1 \int_{\mathbb{X}} f\, d\mu + \lambda_2 \int_{\mathbb{X}} g\, d\mu$$

για κάθε $\lambda_1, \lambda_2 \in \mathbb{R}$.

2. *Το ολοκλήρωμα είναι θετικός τελεστής δηλαδή αν $f \geq g$ υπο την έννοια ότι $f(x) \geq g(x)$ σχεδόν παντού τότε*

$$\int_{\mathbb{X}} f\, d\mu \geq \int_{\mathbb{X}} g\, d\mu$$

3. *Για οποιαδήποτε ολοκληρώσιμη συνάρτηση ισχύει ότι*

$$\left|\int_{\mathbb{X}} f\, d\mu\right| \leq \int_{\mathbb{X}} \mid f \mid \, d\mu$$

4. *Αν $A, B \subset \mathbb{X}$ τέτοια ώστε $A \cap B = \emptyset$ τότε*

$$\int_{A\cup B} f\, d\mu = \int_A f\, d\mu + \int_B f\, d\mu$$

5. *Αν $A, B \subset \mathbb{X}$ τέτοια ώστε $\mu(A) \leq \mu(B)$ και f θετική συνάρτηση τότε*

$$\int_A f\, d\mu \leq \int_B f\, d\mu$$

Οι ιδιότητες αυτές είναι πολύ χρήσιμες και ο τρόπος απόδειξης τους ακολουθεί τον ορισμό του ολοκληρώματος. Για παράδειγμα, ως προς την 1 ή την 2 θυμηθείτε ότι οι ιδιότητες αυτές ισχύουν για απλές συναρτήσεις. Αν λοιπόν πάρουμε τις ακολουθίες απλών συναρτήσεων f_n και g_n που προσεγγίζουν αντιστοίχως την f και g οι ιδιότητες αυτές ισχύουν για κάθε n. Πηγαίνοντας στο όριο παίρνουμε τα ολοκληρώματα των f και g και μπορούμε να δείξουμε ότι η ιδιότητα αυτή ισχύει και για αυτά. Θέλει όμως λίγο προσοχή το πως μπορούμε να πάμε απο το όριο των ακολουθιών συναρτήσεων στο όριο των ολοκληρωμάτων. Αυτό μας το εξασφαλίζει το θεώρημα μονότονης σύγκλισης το οποίο θα μελετηθεί στην επόμενη παράγραφο.

Η ιδιότητα 4 μπορεί να αποδειχθεί πολύ εύκολα λαμβάνοντας υπόψιν ότι $\mathbf{1}_{A\cup B} = \mathbf{1}_A + \mathbf{1}_B$ και τη γραμμικότητα του ολοκληρώματος. Η ιδιότητα 5 είναι επίσης πολύ σημαντική γιατί δίνει την μονοτονία του ολοκληρώματος ως προς το μέγεθος του συνόλου επάνω στο οποίο ολοκληρώνουμε, όπως αυτό καθορίζεται απο το μέτρο του.

Η πρόταση 5.7.1 έχει την προφανή της γενίκευση για τις μέσες τιμές τυχαίων μεταβλητών.

Πρόταση 5.7.2 *Έστω $(\Omega, \mathcal{F}, \mu)$ χώρος μέτρου και $X, Y : \Omega \to \mathbb{R}$ τυχαίες μεταβλητές.*

1. *Το ολοκλήρωμα (μέση τιμή) είναι γραμμικό δηλαδή*

$$\int_\Omega (\lambda_1 X + \lambda_2 Y)\, d\mu = \lambda_1 \int \Omega X\, d\mu + \lambda_2 \int \Omega Y\, d\mu,$$

ή ισοδύναμα

$$\mathbb{E}_\mu[\lambda_1 X + \lambda_2 Y] = \lambda_1\, \mathbb{E}_\mu[X] + \lambda_2\, \mathbb{E}_\mu[Y],$$

για κάθε $\lambda_1, \lambda_2 \in \mathbb{R}$.

2. *Το ολοκλήρωμα (μέση τιμή) είναι θετικός τελεστής δηλαδή αν $X \geq Y$ σχεδόν βέβαια τότε*

$$\int_\Omega X\, d\mu \geq \int_\Omega Y\, d\mu$$

ή ισοδύναμα

$$\mathbb{E}_\mu[X] \geq \mathbb{E}_\mu[Y]$$

3. *Ισχύει ότι*

$$\left|\int_\Omega X\, d\mu\right| \leq \int_\Omega \mid X \mid \, d\mu$$

ή ισοδύναμα

$$\mid \mathbb{E}_\mu[X] \mid \leq \mathbb{E}_\mu[\mid X \mid]$$

4. *Αν $A, B \subset \Omega$ τέτοια ώστε $A \cap B = \emptyset$ τότε*

$$\int_{A\cup B} X\, d\mu = \int_A X\, d\mu + \int_B X\, d\mu$$

ή ισοδύναμα

$$\mathbb{E}_\mu[X\, \mathbf{1}_{A\cup B}] = \mathbb{E}_\mu[X\, \mathbf{1}_A] + \mathbb{E}_\mu[X\, \mathbf{1}_B]$$

5. *Αν $A, B \subset \mathbb{X}$ τέτοια ώστε $\mu(A) \leq \mu(B)$ και X θετική τυχαία μεταβλητή τότε*

$$\int_A X\, d\mu \leq \int_B X\, d\mu$$

ή ισοδύναμα

$$\mathbb{E}_\mu[X\, \mathbf{1}_A] \leq \mathbb{E}_\mu[X\, \mathbf{1}_B]$$

5.7.2 Το ολοκλήρωμα δεν καταλαβαίνει σύνολα μέτρου 0

Ας υποθέσουμε ότι έχουμε δύο μετρήσιμες συναρτήσεις f, g για τις οποίες δεν ισχύει $f(x) = g(x)$ για κάθε $x \in \mathbb{X}$ αλλά επιτρέπουμε η ισότητα αυτή να παραβιάζεται για κάποια x αρκεί το μέτρο αυτών των x να είναι 0. Δηλαδή αν $A = \{x \in \mathbb{X} \mid f(x) \neq g(x)\}$ τότε $\mu(A) = 0$. Θα λέμε ότι $f = g$ σχεδόν παντού ως προς το μέτρο μ.

Αντίστοιχα, ας υποθέσουμε ότι έχουμε δύο τυχαίες μεταβλητές X και Y για τις οποίες δεν ισχύει $X(\omega) = Y(\omega)$ για κάθε $\omega \in \Omega$ αλλά επιτρέπουμε η ισότητα αυτή να παραβιάζεται για κάποια ω αρκεί το σύνολο όλων αυτών των ω για τα οποία παραβιάζεται αυτή η ισότητα να είναι μέτρου 0. Δηλαδή αν $A = \{\omega : X(\omega) \neq Y(\omega)\}$ τότε $\mu(A) = 0$. Τότε λέμε ότι η ισότητα αυτή $X = Y$ ισχύει $\mu-$σχεδόν βέβαια.

Παράδειγμα 5.7.1 *Έστω $\Omega = [0,1]$ και μ το μέτρο Lebesgue στο $[0,1]$. Ας ορίσουμε $X(\omega) = \omega$ και*

$$Y(\omega) = \begin{cases} \omega & \omega \in \mathbb{R} \setminus \mathbb{Q} \\ 1 & \omega \in \mathbb{Q} \end{cases}$$

Τότε $X = Y$ σχεδόν παντού ή σχεδόν βέβαια ως προς το μέτρο Lebesgue. Αυτό γιατί διαφέρουν μόνο αν $\omega \in \mathbb{Q}$ και ως γνωστό $\mu(\mathbb{Q}) = 0$.

Πρόταση 5.7.3 1. *Έστω $(\mathbb{X}, \mathcal{F}, \mu)$ χώρος μέτρου και $f, g : \mathbb{X} \to \mathbb{R}$ δύο μετρήσιμες συναρτήσεις για τις οποίες ισχύει $f = g$ μ-σχεδόν παντού. Τότε*

$$\int_{\mathbb{X}} f\, d\mu = \int_{\mathbb{X}} g\, d\mu$$

δηλαδή το μέτρου 0 σύνολο A στο οποίο οι τιμές των f και g διαφέρουν δεν συνεισφέρει καθόλου στην ολοκλήρωση.

2. *Έστω $(\Omega, \mathcal{F}, \mu)$ χώρος μέτρου (χώρος πιθανοτήτων) και $X, Y : \Omega \to \mathbb{R}$ δύο τυχαίες μεταβλητές για τις οποίες ισχύει οτι $X = Y$ σχεδον βέβαια ως προς το μέτρο μ. Τότε*

$$\int_\Omega X\, d\mu = \int_\Omega Y\, d\mu$$

ή ισοδύναμα

$$\mathbb{E}_\mu[X] = \mathbb{E}_\mu[Y],$$

δηλαδή το μέτρου 0 σύνολο A στο οποίο οι τυχαίες μεταβλητές διαφέρουν δεν συνεισφέρει στην μέση τιμή.

Απόδειξη: Θα αποδείξουμε το (2). Η απόδειξη του (1) είναι ακριβώς η ίδια.
Έστω $A = \{\omega : X(\omega) \neq Y(\omega)\}$. Απο την υπόθεση $\mu(A) = 0$. Παρατηρείστε οτι $Y = Y\, \mathbf{1}_A + Y\, \mathbf{1}_{A^c}$. Συνεπώς,

$$\int_\Omega Y\, d\mu = \int_\Omega Y \mathbf{1}_A\, d\mu + \int_\Omega Y \mathbf{1}_{A^c}\, d\mu = \int_\Omega Y \mathbf{1}_A\, d\mu + \int_\Omega X\, \mathbf{1}_{A^c}\, d\mu$$

Ας πάρουμε τώρα μία ακολουθία απο απλές συναρτήσεις (τυχαίες μεταβλητές) Y_n, τέτοια ώστε $Y_n \uparrow Y$. Τότε $Y_n\, \mathbf{1}_A$ είναι απλή και $Y_n\, \mathbf{1}_A \uparrow Y\, \mathbf{1}_A$. Εφόσον η Y_n είναι απλή, είναι και φραγμένη απο τον αριθμό N. Συγκεκριμένα εφόσον Y_n απλή μπορεί να γραφει στην μορφή $Y_n = \sum_{j=1}^m a_j\, \mathbf{1}_{A_j}$ και $N = max\{a_1, \cdots, a_m\}$.

Κατα συνέπεια,

$$0 \leq \int_\Omega Y_n\, \mathbf{1}_A\, d\mu \leq \int_\Omega N\, \mathbf{1}_A\, d\mu = N \int_\Omega \mathbf{1}_A\, d\mu = N\, \mu(A) = 0$$

Η αριστερή ανισότητα προκύπτει επειδή $Y_n \geq 0$.

Απο την παραπάνω ανισότητα λοιπόν βλέπουμε ότι $\int_\Omega Y_n \mathbf{1}_A \, d\mu = 0$ για κάθε n, και συνεπώς αν πάρουμε και το όριο της ακολουθίας αυτής θα έχουμε ότι και το όριο είναι 0 άρα $\int_\Omega Y \mathbf{1}_A \, d\mu = 0$.

Δουλεύοντας ομοια και για την X έχουμε ότι $\int_\Omega X \mathbf{1}_A \, d\mu = 0$.

Συνεπώς,

$$\begin{aligned}\int_\Omega Y \, d\mu &= \int_\Omega Y \mathbf{1}_A \, d\mu + \int_\Omega Y \mathbf{1}_{A^c} \, d\mu = \int_\Omega Y \mathbf{1}_A \, d\mu + \int_\Omega X \mathbf{1}_{A^c} \, d\mu = \int_\Omega X \mathbf{1}_{A^c} \, d\mu \\ &= \int_\Omega X \mathbf{1}_A \, d\mu + \int_\Omega X \mathbf{1}_{A^c} \, d\mu = \int_\Omega (X\mathbf{1}_A + X\mathbf{1}_{A^c}) \, d\mu = \int_\Omega X \, d\mu\end{aligned}$$

και ο ισχυρισμός αποδείχθηκε. ■

5.8 Ακολουθίες συναρτήσεων ή τυχαίων μεταβλητών και ολοκλήρωση

Ας υποθέσουμε ότι έχουμε μια ακολουθία συναρτήσεων $f_n : \mathbb{X} \to \mathbb{R}$ για την οποία ισχύει

$$f_n \to f, \text{ σχεδόν παντού ως προς το μέτρο } \mu.$$

Μπορούμε να ισχυριστούμε γενικά ότι

$$\int_\mathbb{X} f_n \, d\mu \to \int_\mathbb{X} f \, d\mu \, ?$$

Η απάντηση είναι γενικά **όχι**! Είναι λοιπόν πολύ σημαντικό να γνωρίζουμε κάτω απο ποιές συνθήκες η σύγκλιση της ακολουθίας συναρτήσεων μπορεί να εξασφαλίζει και την σύγκλιση των ολοκληρωμάτων.

Αντίστοιχα, ας υποθέσουμε ότι έχουμε μια ακολουθία τυχαίων μεταβλητών $X_n : \Omega \to \mathbb{R}$ για την οποία ισχύει

$$X_n \to X, \text{ σχεδόν βέβαια ως προς το μέτρο } \mu.$$

Μπορούμε να ισχυριστούμε γενικά ότι

$$\mathbb{E}_\mu[X_n] \to \mathbb{E}_\mu[X] \, ?$$

Η απάντηση είναι γενικά **όχι**! Είναι λοιπόν πολύ σημαντικό να γνωρίζουμε κάτω απο ποιές συνθήκες η σύγκλιση της ακολουθίας τυχαίων μεταβλητών μπορεί να εξασφαλίζει και την σύγκλιση των μέσων τιμών.

Παράδειγμα 5.8.1 *Έστω $\Omega = [0,1]$ και μ το μέτρο Lebesgue στο $[0,1]$. Αν πάρουμε την ακολουθία συναρτήσεων (τυχαίων μεταβλητών) $X_n = n \mathbf{1}_{[0,1/n]}$ τοτε $X_n \to X = 0$ αλλά $\int X_n \, d\mu = 1$ για κάθε n, συνεπώς $\int X_n \, d\mu \not\to \int X \, d\mu = 0$.*

Τα θεωρήματα που δίνουμε στην παράγραφο αυτή μας δίνουν συνθήκες που μας εξασφαλίζουν το ότι η σύγκλιση της ακολουθίας συναρτήσεων ή τυχαίων μεταβλητών οδηγεί και σε σύγκλιση των ολοκληρωμάτων.

5.8.1 Το θεώρημα μονότονης σύγκλισης

Η παρακάτω πρόταση μας δίνει συνθήκες που εξασφαλίζουν την σύγκλιση αυτή.

Πρόταση 5.8.1 Μονότονη σύγκλιση, Beppo Levi
Έστω $f_n : \mathbb{X} \to \mathbb{R}$ ακολουθία συναρτήσεων τέτοια ώστε $f_n \uparrow f$ σχεδόν παντού. Τότε

$$\int_\mathbb{X} f_n \, d\mu \to \int_\mathbb{X} f \, d\mu.$$

Πρόταση 5.8.2 Μονότονη σύγκλιση, Beppo Levi
Έστω $X_n : \Omega \to \mathbb{R}$ τυχαίων μεταβλητών τέτοια ώστε $X_n \uparrow X$ σχεδόν βέβαια. Τότε

$$\int_\Omega X_n \, d\mu \to \int_\Omega X \, d\mu$$

ή ισοδύναμα

$$\mathbb{E}_\mu[X_n] \to \mathbb{E}_\mu[X].$$

Σχόλιο 5.8.1 Θυμηθείτε ότι αν X_n είναι ακολουθία απο απλές τυχαίες μεταβλητές η οποία είναι μονότονη και $X_n \uparrow X$ τότε απο τον ορισμό του ολοκληρώματος το παραπάνω ισχύει. Το θεώρημα μονότονης σύγκλισης όμως γενικεύει αυτό το αποτέλεσμα, και για μονότονες ακολουθίες που δεν αποτελούνται απο απλές τυχαίες μεταβλητές!

Σχόλιο 5.8.2 Το θεώρημα μονότονης σύγκλισης μπορεί να συμπληρωθεί και για να μας δώσει ένα κριτήριο σύγκλισης της ακολουθίας συναρτήσεων ή τυχαίων μεταβλητών. Αν f_n μονότονη και μη αρνητική ακολουθία συναρτήσεων τέτοια ώστε $\int_{\mathbb{X}} f_n \, d\mu < C$ για κάθε n, όπου C ανεξάρτητο το n τότε η ακολουθία f_n συγκλίνει σχεδόν παντού σε μια πεπερασμένη συνάρτηση f. Παρόμοιο αποτέλεσμα, τηρουμένων των αναλογιών, ισχύει και για τυχαίες μεταβλητές. Το αποτέλεσμα αυτό μπορεί να θεωρηθεί ως μια γενίκευση του θεωρήματος που γνωρίζουμε απο την πραγματική ανάλυση ότι μονότονες και φραγμένες ακολουθίες είναι συγκλίνουσες.

Το θεώρημα αυτό αποδείχθηκε απο τον Ιταλό μαθηματικό Beppo Levi το 1906. Είναι ένα πολύ χρήσιμο αποτέλεσμα και βοηθάει στο να αποδείξουμε πολλά και ενδιαφέροντα αποτελέσματα σχετικά με το ολοκλήρωμα Lebesgue, την προσέγγιση των ολοκληρώσιμων συναρτήσεων κ.α.

Παράδειγμα 5.8.2 *Στο παράδειγμα 5.8.1 δεν μπορούμε να εφαρμόσουμε την πρόταση αυτή επειδή η ακολουθία X_n* **δεν** *είναι μονότονη.*

Παράδειγμα 5.8.3 *Η ακολουθία συναρτήσεων $f_n(x) : [0,1] \to \mathbb{R}$, $f_n = \frac{n}{n+x^a}$, $a < 1$ είναι μονότονη. Αυτό μας εξασφαλίζει ότι $\int_0^1 f_n(x)\, dx \to 0$.*

5.8.2 Λήμμα του Fatou

Πρόταση 5.8.3 Fatou *Έστω $f_n \geq 0$, $f_n : \mathbb{X} \to \mathbb{R}$ μια ακολουθία συναρτήσεων. Ισχύει ότι*

$$\int_{\mathbb{X}} \liminf_{n\to\infty} f_n \, d\mu \leq \liminf_{n\to\infty} \int_{\mathbb{X}} f_n \, d\mu.$$

Πρόταση 5.8.4 Fatou *Έστω $X_n \geq 0$, $X_n : \Omega \to \mathbb{R}$ μια ακολουθία τυχαίων μεταβλητών. Ισχύει ότι*

$$\int_{\Omega} \liminf_{n\to\infty} X_n \, d\mu \leq \liminf_{n\to\infty} \int_{\Omega} X_n \, d\mu,$$

ή ισοδύναμα

$$\mathbb{E}_\mu \left[\liminf_{n\to\infty} X_n\right] \leq \liminf_{n\to\infty} \mathbb{E}_\mu[X_n].$$

Το λήμμα αυτό αποδείχθηκε απο τον Γάλλο μαθηματικό και αστρονόμο Pierre Fatou το 1906.

Παράδειγμα 5.8.4 *Ας θεωρήσουμε την ακολουθία συναρτήσεων $f_n(x) = n\,\mathbf{1}_{[0,\frac{1}{n}]}(x)$. Μπορούμε εύκολα να δούμε ότι*

$$\liminf_{n\to\infty} f_n(x) = f(x) = 0$$

ενώ $\int_0^1 f_n(x)\, dx = 1$ για κάθε n. Συνεπώς

$$\int_0^1 \liminf_{n\to\infty} f_n(x)\, dx = 0 \leq \liminf_{n\to\infty} \int_0^1 f_n(x)\, dx = 1$$

5.8.3 Το θεώρημα κυριαρχημένης σύγκλισης

Ίσως το πιο ευέλικτο κριτήριο για να μπορέσουμε να εγγυηθούμε αν μια ακολουθία συναρτήσεων συγκλίνει σχεδόν παντού τότε τα ολοκληρώματα των μελών της ακολουθίας συγκλίνουν στο ολοκλήρωμα του ορίου είναι το θεώρημα κυριαρχημένης σύγκλισης του Lebesgue.

Πρόταση 5.8.5 *Έστω $f_n : \mathbb{X} \to \mathbb{R}$ ακολουθία συναρτήσεων τέτοια ώστε $f_n \to f$ σχεδον παντού και έστω ότι υπάρχει συνάρτηση g τέτοια ώστε $\mid f_n \mid \leq g$ σχεδον παντού για κάθε n, και επιπλεον $\int_{\mathbb{X}} \mid g \mid \, d\mu \leq \infty$.*
Τότε,

$$\int_{\mathbb{X}} f_n \, d\mu \to \int_{\mathbb{X}} f \, d\mu.$$

Πρόταση 5.8.6 *Έστω $X_n : \Omega \to \mathbb{R}$ ακολουθία τυχαίων μεταβλητών τέτοια ώστε $X_n \to X$ σχεδόν βέβαια και έστω ότι υπάρχει τυχαία μεταβλητή Y τέτοια ώστε $\mid X_n \mid \leq Y$ σχεδον βέβαια για κάθε n, και επιπλεον $\mathbb{E}_\mu[Y] = \int_\Omega \mid Y \mid \, d\mu \leq \infty$.*
Τότε,

$$\int_\Omega X_n \, d\mu \to \int_\Omega X \, d\mu$$

ή ισοδύναμα

$$\mathbb{E}_\mu[X_n] \to \mathbb{E}_\mu[X].$$

Παράδειγμα 5.8.5 *Ας πάρουμε την ακολουθία συναρτήσεων $f_n(x) = \frac{n\sqrt{x}}{1+n^2x^2}$. Για κάθε n ισχύει ότι $f_n(x) \leq \frac{1}{2\sqrt{x}} =: g(x)$. Επειδή g ολοκληρώσιμη και $f_n(x) \to 0$ απο το θεώρημα κυριαρχημένης σύγκλισης καταλήγουμε ότι $\int_{\mathbb{X}} f_n(x) \, dx \to 0$.*

Παράδειγμα 5.8.6 *Το θεώρημα κυριαρχημένης σύγκλισης εξασφαλίζει ότι αν $\mid X_n \mid < C$ σχεδόν βέβαια και $X_n \to X$ σχεδόν βέβαια τότε $\mathbb{E}[X_n] \to \mathbb{E}[X]$.*

Παράδειγμα 5.8.7 *Ας πάρουμε την ακολουθία $f_n(x) = n\,\mathbf{1}_{[0,\frac{1}{n}]}(x)$. Μπορούμε να δούμε ότι $f_n(x) \to 0$ αλλά $\int_0^1 f_n(x) \, dx = 1$. Το θεώρημα κυριαρχημένης σύγκλισης δεν μπορεί να εφαρμοστεί εδώ.*

Παράδειγμα 5.8.8 *Έστω X μια τυχαία μεταβλητή. Μπορούμε να ορίσουμε την χαρακτηριστική της συνάρτηση*

$$\phi_X(u) := \mathbb{E}[e^{i\,u\,X}] := \mathbb{E}[\cos(u\,X)] + i\,\mathbb{E}[\sin(u\,X)]$$

όπου $i^2 = -1$ είναι η μιγαδική μονάδα. Χρησιμοποιώντας το θεώρημα κυριαρχημένης σύγκλισης για την ακολουθία τυχαίων μεταβλητών $X_n = e^{i\,u_n\,X}$ όπου u_n είναι μια πραγματική ακολουθία $u_n \to u$ μπορούμε να δείξουμε ότι η συνάρτηση $\phi_X(u)$ είναι συνεχής για κάθε $u \in \mathbb{R}$ δηλαδή $\phi_X(u_n) \to \phi_X(u)$.

5.9 Η σχέση του ολοκληρώματος *Lebesgue* με το ολοκλήρωμα *Riemann*

Μια ενδιαφέρουσα ειδική περίπτωση του ολοκληρώματος Lebesgue είναι η περίπτωση όπου $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}(\mathbb{R})$, $\mu = Leb(\mathbb{R})$, ή η περίπτωση όπου $\Omega = [a,b]$, $\mathcal{F} = \mathcal{B}([a,b])$, $\mu = Leb([a,b])$, όπου με $Leb(\mathbb{R})$ ή $Leb([a,b])$ συμβολίζουμε το μέτρο Lebesgue στο $\mathbb{R}$ ή στο $[a,b]$ αντίστοιχα. Στην περίπτωση αυτή η $X : \Omega \to \mathbb{R}$ μπορεί να κατανοηθεί (μπορεί να ταυτιστεί) με μια πραγματική συνάρτηση $f : \mathbb{R} \to \mathbb{R}$ ή $f : [a,b] \to \mathbb{R}$ αντίστοιχα. Κατά συνέπεια στην περίπτωση αυτή θα γράφουμε

$$\mathbb{E}_\mu[X] = \int_{\mathbb{R}} f(x)\mu(dx)$$

ή

$$\mathbb{E}_\mu[X] = \int_a^b f(x)\mu(dx)$$

ανάλογα με την περίπτωση που μας ενδιαφέρει.

Η έννοια της μετρήσιμης συνάρτησης στο πλαίσιο αυτό μπορεί να εκφραστεί με τον ακόλουθο ορισμό

Ορισμός 5.9.1 *Μια πραγματική συνάρτηση ονομάζεται μετρήσιμη (κατά Borel) αν το σύνολο $\{x \in [a,b], \mid f(x) > c\} \in \mathcal{B}([a,b])$ για κάθε $c \in \mathbb{R}$.*

Μπορούμε να παρατηρήσουμε ότι ο ορισμος αυτός είναι ισοδύναμος με τον ορισμό 5.2.1 εφοσον στην περίπτωση αυτή $\mathcal{F} = \mathcal{B}([a,b])$ και το σύνολο $I_c = \{x \in [a,b], \mid f(x) > c\} \subset \Omega = [a,b]$ δεν είναι άλλο απο το το $f^{-1}((-c,\infty))$ όπου φυσικά $(-c,\infty) \in \mathcal{B}([a,b])$.

Ισχύει η ακόλουθη σχέση μεταξύ των μετρήσιμων κατά Lebesgue συναρτήσεων στο διάστημα $[a,b]$ και των ολοκληρώσιμων κατά Riemann συναρτήσεων

Πρόταση 5.9.1 *Κάθε συνάρτηση που είναι ολοκληρώσιμη κατά Riemann είναι και μετρήσιμη στο $[a,b]$.*

Σχήμα 5.1: O Pierre Fatou

Η ολοκλήρωση του Lebesgue γενικεύει την έννοια της ολοκλήρωσης του Riemann .

Πρόταση 5.9.2 *Αν μια συνάρτηση* $f : [a, b] \to \mathbb{R}$ *είναι ολοκληρώσιμη κατά Riemann τότε είναι και ολοκληρώσιμη κατά Lebesgue. Το αντίστροφο δεν ισχύει πάντα.*
Αν μια συνάρτηση είναι και ολοκληρώσιμη κατά Riemann και ολοκληρώσιμη κατά Lebesgue τα δύο ολοκληρώματα ταυτίζονται.

Παράδειγμα 5.9.1 *Η συνάρτηση Dirichlet είναι ολοκληρώσιμη κατά Lebesgue αλλά δεν είναι ολοκληρώσιμη κατά Riemann.*

5.10 Παράρτημα 1: Μερικά ιστορικά στοιχεία

Στο παράρτημα αυτό παραθέτουμε ορισμένα ιστορικά στοιχεία για τους πρωτεργάτες της θεωρίας μέτρου.

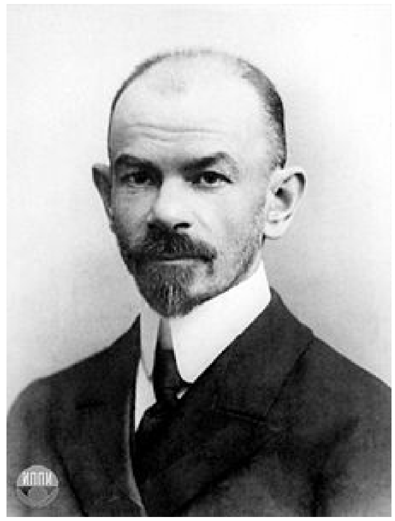
5.10.1 Pierre Fatou, 1875-1941

O Pierre Fatou ... Σπούδασε στην περίφημη Ecole Normale Supérieure και στην διδακτορική του διατριβή με τον τίτλο Integrale, Longuer, Aire (Ολοκλήρωμα, Μήκος και Εμβαδό) έθεσε τις βάσεις της θεωρίας μέτρου και της ολοκλήρωσης όπως τις χρησιμοποιούμε ακόμα και σήμερα. Στην διατριβή αυτή γενίκευσε την έννοια του ολοκληρώματος Riemann-Stieltjes διατυπώνοντας την ιδέα του ολοκληρώματος με την χρήση του μέτρου, το οποίο και πήρε το όνομα του. Η γενίκευση αυτή οδήγησε σε πολύ ενδιαφέρουσες εφαρμογές και παράκαμψε ορισμένες αδυναμίες της θεωρίας της ολοκλήρωσης των Riemann-Stieltjes όπως π.χ. θέματα σχετικά με το ποιές συναρτήσεις είναι ολοκληρώσιμες, ή θέματα που σχετίζονται με τις εναλλαγές της πράξης του ορίου και της ολοκλήρωσης. Ο Lebesgue ασχολήθηκε και με άλλα θέματα της ανάλυσης και της τοπολογίας αλλά έμεινε στην ιστορία της σύγχρονης ανάλυσης για την αναμόρφωση της θεωρίας της όλοκλήρωσης που πρότεινε.

5.10.2 Dimitri Egorov

5.10.3 Beppo Levi

5.10.4 Georg Riemann



Σχήμα 5.2: Ο Dimitri Egorov

Σχήμα 5.3: Ο Beppo Levi

Σχήμα 5.4: Ο Georg Riemann

Κεφάλαιο 6

Χώροι L^p

6.1 Εισαγωγή

6.2 Ο χώρος $L^1(\mu)$

Ας υποθέσουμε τον χώρο μέτρου $(\Omega, \mathcal{F}, \mu)$. Θα μελετήσουμε το σύνολο όλων των συναρτήσεων X οι οποίες είναι μετρήσιμες ώς προς την $\mathcal{F}$ και η απόλυτη τιμή τους είναι ολοκληρώσιμη ως προς το μέτρο μ, δηλαδή τις συναρτήσεις για τις οποίες ισχύει $\int_\Omega \mid X(\omega) \mid, d\mu < \infty$. Στην περίπτωση που μπορούμε να ερμηνεύσουμε τον χώρο μέτρου σαν χώρο πιθανοτήτων,οι συναρτήσεις X μπορεί να θεωρηθούν ως τυχαίες μεταβλητές. Στην περίπτωση αυτή η συνθήκη γίνεται $\mathbb{E}_\mu[\mid X \mid] < \infty$.

Ορισμός 6.2.1 *Έστω χώρος μέτρου $(\Omega, \mathcal{F}, \mu)$. Θα συμβολίζουμε με $L^1(\mu)$ το σύνολο των συναρτήσεων $X : \Omega \to \mathbb{R}$ οι οποίες είναι μετρήσιμες ως προς την σ-άλγεβρα $\mathcal{F}$ και για τις οποίες ισχύει*

$$\int_\Omega \mid X(\omega) \mid, d\mu < \infty$$

Η απεικόνιση $\|\cdot\|_{L^1(\mu)}: L^1(\mu) \to \mathbb{R}^+$, που ορίζεται ως $\| X \|_{L^1(\mu)} := \int_\Omega \mid X(\omega) \mid, d\mu$ ονομάζεται νόρμα του χώρου $L^1(\mu)$.
Ο χώρος αυτός μπορεί να εφοδιαστεί με την μετρική $d(X_1, X_2) = \int_\Omega \mid X_1 - X_2 \mid, d\mu = \| X_1 - X_2 \|_{L^1(\mu)}$.

Είναι ενδιαφέρον να δούμε τι σημαίνει η έκφραση $X_1 = X_2$ για δύο στοιχεία του $L^1(\mu)$. Την ισότητα αυτή στον $L^1(\mu)$ θα πρέπει να την ερμηνεύσουμε ως $d(X_1, X_2) = 0$. Αυτό απο τις ιδιότητες του ολοκληρώματος μας λέει ότι $\mu(\{\omega : X_1(\omega) \neq X_2(\omega)\}) = 0$ δηλαδή οι X_1 και X_2 θεωρούνται σαν ίσες στον $L^1(\mu)$ αν είναι ίσες σε υποσύνολα του A του Ω για τα οποία ισχύει ότι $\mu(A) = \mu(\Omega)$. Στην περίπτωση που το μ είναι μέτρο πιθανότητας αυτό μεταφράζεται στο $\mu(A) = 1$.

Η έννοια της ισότητας στον $L^1(\mu)$ είναι αρκετά ασθενής, εφόσον με την έννοια της ισότητας αυτή, ταυτίζουμε στοιχεία τα οποία μπορεί να είναι πολύ διαφορετικά.

Παράδειγμα 6.2.1 *Ας θεωρησουμε $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}(\mathbb{R})$ και μ το μέτρο Lebesgue. Οι $X_1(\omega) = 0$, για κάθε $\omega \in \mathbb{R}$ και $X_2 = \delta_{\omega,k}$, $k \in \mathbb{N}$ είναι ίσες ως προς την έννοια της ισότητας στον $L^1(\mu)$. Αυτό επειδή, διαφέρουν μόνο στο σύνολο των φυσικών αριθμών, δηλαδή αν $\omega \in \mathbb{N}$ και ως γνωστό το σύνολο αυτό είναι μηδενοσύνολο για το μέτρο Lebesgue . Αν βέβαια διάλεγαμε κάποιο άλλο διακριτό μέτρο και όχι το μέτρο Lebesgue αυτό δεν θα ίσχυε.*

Τα στοιχεία του $L^1(\mu)$ θα λέμε ότι είναι **κλάσεις ισοδυναμίας** στοιχείων, για τον λόγο που περιγράψαμε παραπάνω και που φαίνεται και πιο καθαρά στο παράδειγμα 6.2.1.

Πρόταση 6.2.1 *Ο χώρος αυτός έχει την ιδιοτητα να είναι γραμμικός, δηλαδή αν $X_1, X_2 \in L^1(\mu)$ και $\lambda_1, \lambda_2 \in \mathbb{R}$ τότε και $\lambda_1 X_1 + \lambda_2 X_2 \in L^1(\mu)$.*
Ισχύει μάλιστα ότι

$$\| \lambda_1 X_1 + \lambda_2 X_2 \|_{L^1(\mu)} < \mid \lambda_1 \mid \| X_1 \|_{L^1(\mu)} + \mid \lambda_2 \mid \| X_2 \|_{L^1(\mu)}$$

δηλαδή ικανοποιείται η τριγωνική ανισότητα.

Μπορούμε τώρα να ορίσουμε την έννοια της σύγκλισης μιας ακολουθίας στοιχείων του $L^1(\mu)$.

Ορισμός 6.2.2 *Έστω $X^{(n)}$ μία ακολουθία στοιχείων του $L^1(\mu)$. Θα λέμε ότι η ακολουθία $X^{(n)}$ συγκλίνει στο στοιχείο $X \in L^1(\mu)$ αν $\| X^{(n)} - X \|_{L^1(\mu)} \to 0$ δηλαδή αν για κάθε $\epsilon > 0$ μπορεί να βρεθεί $N \in \mathbb{N}$ τέτοιο ώστε και κάθε $n > N$ ισχύει $\| X^{(n)} - X \|_{L^1(\mu)} < \epsilon$. Αυτό μεταφράζεται ισοδύναμα στο $\int_\Omega | X^{(n)}(\omega) - X(\omega) | \, d\mu(\omega) < \epsilon$ για $n > N$.*

Παράδειγμα 6.2.2 *Έστω $(\Omega, \mathcal{F}, \mu) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), Leb)$ και ας θεωρήσουμε την ακολουθία στοιχείων του $L^1(\mu)$,*

$$X^{(n)}(\omega) = f^{(n)}(x) = \begin{cases} 1 - n\,x & 0 \le x \le 1/n \\ 1 + n\,x & -1/n \le x \le 0 \\ 0 & \text{αλλιώς} \end{cases}$$

Μπορούμε να δείξουμε χρησιμοποιώντας τον ορισμό ότι $X^{(n)} \to 0$ στον $L^1(\mu)$.

Ορισμός 6.2.3 *Η ακολουθία $X^{(n)} \in L^1(\mu)$ ονομάζεται Cauchy αν για κάθε $\epsilon > 0$ υπάρχει $N \in \mathbb{N}$ τέτοιο ώστε $\| X^{(n)} - X^{(m)} \|_{L^1(\mu)} < \epsilon$ για $n, m > N$.*

Μία συγκλίνουσα ακολουθία είναι πάντοτε Cauchy. Όμως, δεν είναι σίγουρο το αν μία ακολουθία Cauchy συγκλίνει. Το αν θα ισχύει αυτό ή όχι εξαρτάται απο τις ιδιότητες του χώρου στο οποίο ζούν τα στοιχεία της ακολουθίας, καθώς και απο την έννοια της μετρικής (απόσταση) που έχουμε ορίσει επάνω στον χώρο αυτό.

Ορισμός 6.2.4 *Ένας μετρικός χώρος M για τον οποίο ισχύει ότι κάθε ακολουθία Cauchy είναι συγκλινουσα τον χώρο αυτό, δηλαδή ότι αν $X^{(n)} \in M$ και $X^{(n)}$ Cauchy τότε υπάρχει $X \in M$ τέτοιο ώστε $X^{(n)} \to X$ ονομάζεται* **πλήρης μετρικός χώρος**.

Ένα παράδειγμα πλήρους μετρικού χώρου είναι ο $\mathbb{R}$.

Πρόταση 6.2.2 *Ο χώρος $L^1(\mu)$ είναι πλήρης μετρικός χώρος.*

6.3 Ο χώρος $L^2(\mu)$

Θα ζητήσουμε τώρα απο τις συναρτήσεις μας (τυχαίες μεταβλητές) ακόμα πιο ισχυρές ιδιότητες, δηλαδή δεν θέλουμε μόνο αν υπάρχει το ολοκλήρωμα (μέση τιμή) της $| X |$ αλλά και της X^2.

Ορισμός 6.3.1 *Έστω χώρος μέτρου $(\Omega, \mathcal{F}, \mu)$. Θα συμβολίζουμε με $L^2(\mu)$ το σύνολο των συναρτήσεων $X : \Omega \to \mathbb{R}$ οι οποίες είναι μετρήσιμες ως προς την σ-άλγεβρα $\mathcal{F}$ και για τις οποίες ισχύει*

$$\int_\Omega | X(\omega) |^2, d\mu < \infty$$

Η απεικόνιση $\| \cdot \|_{L^2(\mu)}$: $L^2(\mu) \to \mathbb{R}^+$, που ορίζεται ως $\| X \|_{L^2(\mu)} := \int_\Omega | X(\omega) |^2, d\mu$ ονομάζεται νόρμα του χώρου $L^2(\mu)$.
Ο χώρος αυτός μπορεί να εφοδιαστεί με την μετρική $d(X_1, X_2) = \int_\Omega | X_1 - X_2 |^2, d\mu = \| X_1 - X_2 \|_{L^2(\mu)}$.

Όσα είπαμε σχετικά με την έννοια της ισότητας και της κλάσης ισοδυναμίας για τις συναρτήσεις (τυχαίες μεταβλητές) στον $L^1(\mu)$ ισχύουν και στον $L^2(\mu)$. Αν τα X αντιστοιχούν σε τυχαίες μεταβλητές, τότε ο χώρος $L^2(\mu)$ είναι η κλάση ισοδυναμίας των τυχαίων μεταβλητών για τις οποίες ορίζεται η διασπορά.

Παράδειγμα 6.3.1 *Έστω X τυχαία μεταβλητή $X \sim N(a, \sigma^2)$ σε ένα χώρο πιθανοτήτων $(\Omega, \mathcal{F}, \mu)$. Ισχύει ότι $X \in L^2(\mu)$.*

Παράδειγμα 6.3.2 *Στα πλαίσια του παραδείγματος 6.3.5, μπορούμε να δείξουμε ότι $X^{(n)} \in L^2(\mu)$ για κάθε n.*

Πρόταση 6.3.1 *Ο χώρος αυτός έχει την ιδιοτητα να είναι γραμμικός, δηλαδή αν $X_1, X_2 \in L^1(\mu)$ και $\lambda_1, \lambda_2 \in \mathbb{R}$ τότε και $\lambda_1 X_1 + \lambda_2 X_2 \in L^1(\mu)$.*
Ισχύει μάλιστα ότι

$$\| \lambda_1 X_1 + \lambda_2 X_2 \|_{L^1(\mu)} < | \lambda_1 | \| X_1 \|_{L^1(\mu)} + | \lambda_2 | \| X_2 \|_{L^1(\mu)}$$

δηλαδή ικανοποιείται η τριγωνική ανισότητα.

Ο χώρος L^2 έχει μία ειδική ιδιότητα. Στον χώρο αυτό η νόρμα μπορεί να οριστεί με την βοήθεια μίας άλλης απεικονισης που ονομάζεται **εσωτερικό γινόμενο**.

Θα ξεκινήσουμε με τον γενικό ορισμό της έννοιας του εσωτερικού γινομένου για οποιοδήποτε γραμμικό χώρο.

Ορισμός 6.3.2 *Έστω M ένας γραμμικός χώρος και η απεικόνιση $< \cdot, \cdot >: M \times M \to \mathbb{R}$ με τις ιδιότητες*

1. $< x, x > \geq 0$ *και* $< x, x >= 0$ *αν και μόνο αν* $x = 0$.

2. $< x_1, x_2 >=< x_2, x_1 >$

3. $< \lambda_1 x_1 + \lambda_2 x_2, x_3 >= \lambda_1 < x_1, x_3 > + \lambda_2 < x_2, x_3 >$ *για κάθε* $x_1, x_2, x_3 \in M$ *και* $\lambda_1, \lambda_2 \in \mathbb{R}$.

Η απεικόνιση $< \cdot, \cdot >$ ονομάζεται **εσωτερικό γινόμενο** *στον M.*

Το εσωτερικό γινόμενο είναι η γενίκευση της έννοιας του εσωτερικού γινομένου όπως την έχουμε συναντήσει στον ευκλείδιο χώρο $\mathbb{R}^n$.

Στον χώρο $L^2(\mu)$ σαν εσωτερικό γινόμενο μπορούμε να ορίσουμε την απεικόνιση

$$< X, Y >= \int_\Omega X(\omega) Y(\omega)\, d\mu(\omega)$$

Αν ο χώρος $(\Omega, \mathcal{F}, \mu)$ είναι χώρος πιθανοτήτων τότε το εσωτερικό γινόμενο μπορεί να εκφραστεί και ως

$$< X, Y >= \mathbb{E}_\mu[X\, Y]$$

Μπορούμε να δούμε ότι η νορμα του χώρου $L^2(\mu)$ μπορεί να εκφραστεί με την βοήθεια του εσωτερικού γινομένου ως εξής

$$\| X \|_{L^2(\mu)} = \{< X, X >\}^{1/2}$$

Χώροι με την ιδιότητα αυτή ονομάζονται χώροι εσωτερικού γινομένου, και θυμίζουν πολλές απο τις ιδιότητες των ευκλειδίων χώρων. Για παράδειγμα μπορούμε τώρα να ορίσουμε την έννοια της καθετότητας για στοιχεία του $L^2(\mu)$.

Ορισμός 6.3.3 *Έστω M ένας χώρος εσωτερικού γινομένου με εσωτερικό γινόμενο $< \cdot, \cdot >$.*
Τα στοιχεία x και y του M ονομάζονται **κάθετα** *αν $< x, y >= 0$.*

Παράδειγμα 6.3.3 *Ας υποθέσουμε $(\Omega, \mathcal{F}, \mu) = ([-\pi, \pi], \mathcal{B}([-\pi, \pi]), Leb)$ και ας πάρουμε τα στοιχεία $X = \sin(n\, x)$ και $Y = \sin(m\, x)$. Τα στοιχεία αυτά είναι κάθετα για $n \neq m$ εφόσον*

$$< X, Y >= \int_{-\pi}^{\pi} \sin(n\, x)\, \sin(m\, x)\, dx = 0$$

Παράδειγμα 6.3.4 *Ας υποθέσουμε $(\Omega, \mathcal{F}, \mu)$ χώρος πιθανοτήτων και X, Y τυχαίες μεταβλητές για τις οποίες ορίζεται η διασπορά, έχουν μέση τιμή 0 και είναι ανεξάρτητες. Τότε οι X και Y είναι κάθετες.*

Μπορούμε τώρα να ορίσουμε την έννοια της σύγκλισης μιας ακολουθίας στοιχείων του $L^2(\mu)$.

Ορισμός 6.3.4 *Έστω $X^{(n)}$ μία ακολουθία στοιχείων του $L^2(\mu)$. Θα λέμε ότι η ακολουθία $X^{(n)}$ συγκλίνει στο στοιχείο $X \in L^2(\mu)$ αν $\| X^{(n)} - X \|_{L^2(\mu)} \to 0$ δηλαδή αν για κάθε $\epsilon > 0$ μπορεί να βρεθεί $N \in \mathbb{N}$ τέτοιο ώστε και κάθε $n > N$ ισχύει $\| X^{(n)} - X \|_{L^2(\mu)} < \epsilon$. Αυτό μεταφράζεται ισοδύναμα στο $\int_\Omega | X^{(n)}(\omega) - X(\omega) |^2 \; d\mu(\omega) < \epsilon$ για $n > N$.*

Παράδειγμα 6.3.5 *Έστω $(\Omega, \mathcal{F}, \mu) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), Leb)$ και ας θεωρήσουμε την ακολουθία στοιχείων του $L^2(\mu)$,*

$$X^{(n)}(\omega) = f^{(n)}(x) = \left\{ \begin{array}{cc} 1 - n\, x & 0 \leq x \leq 1/n \\ 1 + n\, x & -1/n \leq x \leq 0 \\ 0 & \text{αλλιώς} \end{array} \right.$$

Μπορούμε να δείξουμε χρησιμοποιώντας τον ορισμό ότι $X^{(n)} \to 0$ στον $L^2(\mu)$.

Ορισμός 6.3.5 *Η ακολουθία $X^{(n)} \in L^2(\mu)$ ονομάζεται Cauchy αν για κάθε $\epsilon > 0$ υπάρχει $N \in \mathbb{N}$ τέτοιο ώστε $|| X^{(n)} - X^{(m)} ||_{L^2(\mu)} < \epsilon$ για $n, m > N$.*

Μπορούμε να δείξουμε ότι

Πρόταση 6.3.2 *Ο χώρος $L^2(\mu)$ είναι πλήρης μετρικός χώρος.*

Ορισμός 6.3.6 *Ένας πλήρης χώρος εσωτερικού γινομένου ονομάζεται χώρος Hilbert .*

Πρόταση 6.3.3 *Ο χώρος $L^2(\mu)$ είναι χώρος Hilbert .*

Η ακόλουθη ανισότητα είναι μία γενική ανισότητα που συνδέει του χώρους $L^2(\mu)$ με τους χώρους $L^1(\mu)$.

Πρόταση 6.3.4 *(Cauchy-Schwartz)*
Έστω $X, Y \in L^2(\mu)$. Τότε το γινόμενο $XY \in L^1(\mu)$ και μάλιστα

$$\int_{\Omega} | XY | \, d\mu \leq \left\{ \int_{\Omega} | X |^2 \, d\mu \right\}^{\frac{1}{2}} \left\{ \int_{\Omega} | Y |^2 \, d\mu \right\}^{\frac{1}{2}}$$

Η ισότητα ισχύει αν και μόνο αν υπάρχει $\lambda \in \mathbb{R}$ τέτοιο ώστε $Y = \lambda X$.

6.4 Το θεώρημα της προβολής

Το θεώρημα της προβολής είναι ένα γενικό αποτέλεσμα το οποίο ισχύει στους χώρους Hilbert γενικά και όχι μόνο στους χώρους $L^2(\mu)$. Η εφαρμογή του όμως στους χώρους αυτούς είναι ιδιαίτερα χρήσιμη στην θεωρία πιθανοτήτων και την στατιστική όπως θα δούμε παρακάτω.

Στην παράγραφο αυτή θα παρουσιάσουμε το θεώρημα προβολής στην γενική του μορφή, για οποιοδήποτε χώρο Hilbert , και θα εμβαθύνουμε στις εφαρμογές αργότερα.

Θα ξεκινήσουμε δίνοντας την έννοια του υποχώρου.

Ορισμός 6.4.1 *Έστω H ένας χώρος Hilbert και $M \subset H$ ένα υποσύνολο του με τις ιδιότητες*

1. *Αν $x_1, x_2 \in M$ τότε και $\lambda_1 x_1 + \lambda_2 x_2 \in M$ για κάθε $\lambda_1, \lambda_2 \in \mathbb{R}$.*
2. *Έστω $x^{(n)} \in M$ μία οποιαδήποτε ακολουθία στοιχείων του M. Αν $x^{(n)} \to x$ με την έννοια της σύγκλισης στον H (δηλαδή $|| x^{(n)} - x ||_H \to 0$) τότε $x \in M$, δηλαδή το οριο κάθε συγκλίνουσας ακολουθίας του M ανήκει επίσης στο M.*

Ο M ονομάζεται **υπόχωρος** *του H.*

Παράδειγμα 6.4.1 *Έστω $(\Omega, \mathcal{F}, \mu)$ ένας χώρος πιθανοτήτων και $\mathcal{G} \subset \mathcal{F}$. Ας ορίσουμε ως $H = L^2(\mu, \mathcal{F})$ τον χώρο των συναρτήσεων $X : \Omega \to \mathbb{R}$ οι οποίες είναι μετρήσιμες ως προς την $\sigma-$άλγεβρα $\mathcal{F}$. Αν ορίσουμε ως $M = L^2(\mu, \mathcal{G})$ τον χώρο των συναρτήσεων $X^{'} : \Omega \to \mathbb{R}$ οι οποίες είναι μετρήσιμες ως προς την $\sigma-$άλγεβρα $\mathcal{G}$, τότε ο M είναι υπόχωρος του H.*
Αυτό μπορεί να φανεί ως εξής. Κατ' αρχήν αν $X^{'} \in m\mathcal{G}$ τότε και $X^{'} \in m\mathcal{F}$ εφόσον $\mathcal{G}\mathcal{F}$. Μετά απο τις ιδιότητες της μετρησιμότητας μπορούμε να δούμε ότι οι ιδιότητες του ορισμού 6.4.1 ισχύουν.

Ας θεωρήσουμε ένα χώρο Hilbert H και ένα υπόχωρο του M. Ας υποθεσουμε οτι επιλεγουμε ένα οποιοδήποτε στοιχείο $x \in H$ και θέτουμε το ερώτημα

> Μπορεί να βρεθεί ένα στοιχείο $m^* \in M$ το οποίο να απέχει την μικρότερη απόσταση απο το σημείο x αν ως αποσταση θεωρήσουμε την νόρμα της διαφοράς $x - m^*$ στον χώρο H;

Είναι προφανές ότι αν $x \in M$ τότε το στοιχείο αυτό ταυτίζεται με το x, δηλαδή $m^* = x$ το οποίο μάλιστα και μηδενίζει την απόσταση. Στην περίπτωση όπου $x \notin M$ τότε η απόσταση $|| x - m^* ||_H$ **δεν** γίνεται 0 αλλά παίρνει την ελάχιστη, μη μηδενική τιμή της. Κατα συνέπεια το παραπάνω πρόβλημα μπορεί να ξαναδιατυπωθεί μαθηματικά σαν ένα πρόβλημα βελτιστοποίησης

Έστω $x \in H$. Μπορεί να βρεθεί ένα στοιχείο $m^* \in M$ τέτοιο ώστε

$$\inf_{m \in M} \| x - m \|_H = \| x - m^* \|_H$$

Παράδειγμα 6.4.2 *Για να δούμε το πόσο ενδιαφέρον είναι αυτό το ερώτημα για τις πιθανότητες και την στατιστική ας θεωρήσουμε στα πλαίσια του Παραδείγματος 6.4.1, τους χώρους H και M. Στα πλαίσια αυτά $x = X$ και $m = X'$, δηλαδή x είναι μία τυχαία μεταβλητή η οποία είναι μετρήσιμη ως προς την $\sigma-$ άλγεβρα $\mathcal{F}$ και X' είναι μία τυχαία μεταβλητη η οποία είναι μετρήσιμη ως προς την $\sigma-$ άλγεβρα $\mathcal{G}$. Συνεπώς, X είναι μια τυχάια μεταβλητή η οποία για να περιγραφεί χρειάζεται περισσότερη πληροφορία απο ότι χρειάζεται για να περιγραφεί η X'. Απο τον ορισμό της νόρμας στους χώρους αυτούς έχουμε ότι*

$$\| x - m \|_H = \left\{ \int_\Omega | X - X' |^2 \, d\mu \right\}^{\frac{1}{2}} = \left\{ \mathbb{E}_\mu [| X - X' |] \right\}^{\frac{1}{2}}$$

δηλαδή η απόσταση της X απο την X' μπορεί να ερμηνευθεί σαν το μέσο τετραγωνικό σφάλμα της X' απο την X. Το ερώτημα λοιπόν που διατυπώσαμε παραπάνω, στην γενική του μορφή, στο πλαίσιο αυτό μπορεί να διατυπωθεί ως

> *Μπορεί να βρεθεί τυχαία μεταβλητή X' μετρήσιμη ως προς την μικρότερη $\sigma-$άλγεβρα $\mathcal{G}$ η οποία να προσεγγίζει την τυχαία μεταβλητή X η οποία ειναι μετρήσιμη ως προς την μεγαλύτερη $\sigma-$άλγεβρα $\mathcal{F}$ με το μικρότερο δυνατό μέσο τετραγωνικό σφάλμα;*

Αυτό όμως δεν είναι τίποτε άλλο απο την γενικότερη διατύπωση του προβλήματος της εκτιμητικής στην στατιστικη και τις πιθανότητες, το οποίο αποτελεί και ένα απο τα θεμελιώδη ερωτήματα του γνωστικού αυτού αντικειμένου.

Θα δείξουμε εδώ ότι η απάντηση στο ερώτημα αυτό ειναι θετική, και όχι μόνο μπορεί να υπάρχει τέτοιο m^* αλλά μπορούμε και να το χαρακτηρίσουμε.

Θεώρημα 6.4.1 *(**Το θεώρημα προβολής**)*
Έστω H ένας χώρος Hilbert και $M \subset H$ ένας υπόχωρος του και x οποιοδήποτε $x \in H$.
Ισχύουν τα ακόλουθα:

1. *Υπάρχει μοναδικό $m^* \in M$ τέτοιο ώστε*

$$\inf_{m \in M} \| x - m \|_H = \| x - m^* \|_H$$

2. *Για το m^* ισχύει ότι $(x - m^*, m)_H = 0$ για κάθε $m \in M$, δηλαδή το $x - m^*$ είναι κάθετο σε κάθε στοιχείο του M. Η συνθήκη αυτή είναι αναγκαία και επαρκής.*

*Το m^* ονομάζεται **προβολή** του x στον υπόχωρο M και συμβολίζεται $m^* = \Pi_M x$.*

Η προβολή Π_M μπορεί να θεωρηθεί σαν μία απεικόνιση $\Pi_M : H \to M$.

Πρόταση 6.4.1 *Η απεικόνιση Π_M έχει τις ακόλουθες ιδιότητες.*

1. $\Pi_M^2 = \Pi_M$
2. *Αν $x \in M$ τότε $\Pi_M x = x$.*
3. *Αν $x \in M^\perp$ τότε $\Pi_M x = 0$.*
4. $(\Pi_M x, y)_H = (x, \Pi_M y)_H$.
5. $\Pi_M(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 \Pi_M x_1 + \lambda_2 \Pi_M x_2$.

6.5 Εφαρμογές του θεωρήματος προβολής

6.5.1 Γραμμική παλινδρόμηση και γραμμικά μοντέλα.

Ας υποθέσουμε ότι έχουμε κάποιο λόγο να πιστεύουμε ότι κάποια τυχαία μεταβλητή Y συνδέεται με κάποιες τυχαίες μεταβλητές $X = (X_1, X_2, \cdots, X_n)$. Θέλουμε να βρούμε ένα γραμμικό μοντέλο το οποίο να συνδέει τις τιμές των τυχαιων μεταβλητών X με τις τιμές της τυχαίας μεταβλητής Y, δηλαδη να βρούμε πραγματικούς αριθμούς $\lambda_1, \lambda_2, \cdots, \lambda_n$ τέτοιους ώστε

$$Y = \lambda_1 X_1 + \lambda_2 X_2 + \cdots + \lambda_n X_n$$

Φυσικά, η σχέση αυτή θα πρέπει να είναι ένας στατιστικός νόμος, και είναι ισως πιο σωστό να ξαναγραφεί σαν

$$Y = \lambda_1 X_1 + \lambda_2 X_2 + \cdots + \lambda_n X_n + \epsilon \tag{6.1}$$

όπου ϵ είναι ένας όρος σφάλματος. Ο όρος σφάλματος περιέχει τις αποκλίσεις απο τον γραμμικό αυτό νόμο.

Η γραμμική παλινδρόμηση είναι η διαδικασία η οποία μας επιτρέπει να υπολογίζουμε αυτούς τους συντελεστές $\lambda_1^*, \lambda_2^*, \cdots, \lambda_n^*$ για τους οποίους η εκτιμήτρια

$$\hat{Y} = \lambda_1^* X_1 + \lambda_2^* X_2 + \cdots + \lambda_n^* X_n$$

δίνει την καλύτερη δυνατή προσέγγιση στην Y όπου με την καλύτερη δυνατή προσέγγιση εννοούμε την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος $\mathbb{E}_\mu[(Y - \hat{Y})^2]$.

Γνωρίζοντας τις ιδιότητες του χώρου $L^2(\mu)$ και το θεώρημα προβολής μπορούμε να ξαναγράψουμε αυτό το πρόβλημα, της γραμμικής παλινδρόμησης, χρησιμοποιώντας την θεωρία των χώρων Hilbert.

Ας ορίσουμε τον χώρο Hilbert $H = L^2(\Omega, \mathcal{F}, \mu)$. Ας πάρουμε τα στοιχεία $X_1, X_2, \cdots, X_n \in H$ και τον υπόχωρο M του H ο οποίος ορίζεται ως

$$M = \{y \in H \ \mid \ \lambda_1, \cdots, \lambda_n \in \mathbb{R}, \ y = \lambda_1 X_1 + \cdots + \lambda_n X_n\}$$

Με άλλα λόγια, το M είναι ο υπόχωρος του H ο οποίος περιέχει τα στοιχεία που μπορεί να γραφούν σαν γραμμικός συνδυασμός των στοιχείων $X_1, \cdots, X_n$. Το γιατί ο M είναι υπόχωρος αφήνεται σαν άσκηση.

Το πρόβλημα της γραμμικής παλινδρομησης μπορεί τώρα να γραφεί σαν ένα πρόβλημα προβολής. Το γραμμικό μοντέλο που περιγράφει η εξίσωση (6.1) δεν είναι τίποτε άλλο παρά η προβολή της οποιασδήποτε τυχαίας μεταβλητής $Y \in H$ στον υπόχωρο M, δηλαδή στον υπόχωρο των τυχαίων μεταβλητών που μπορεί να γραφούν σαν γραμμικός συνδυασμός των $X_1, \cdots, X_n$. Η ιδιότητα της προβολής μας εξασφαλίζει ότι αν $\hat{Y} = \Pi_M Y$ τότε θα πάρουμε την εκτιμήτρια που μας εξασφαλίζει το ελάχιστο μέσο τετραγωνικό σφάλμα.

Ας δούμε τώρα πως το θεώρημα της προβολής μπορεί να μας βοηθήσει στην εκτίμηση του μοντέλου. Θα εκμεταλλευτούμε τις συνθήκες καθετότητας. Σύμφωνα με αυτές το $m^* = \hat{Y}$ θα είναι η προβολή αν και μόνο αν $< x - m^*, m >= 0$ για κάθε $m \in M$. Έστω ότι $m^* = \lambda_1^* X_1 + \cdots + \lambda_n^* X_n$ ενώ οποιοδήποτε στοιχείο $m \in M$ μπορεί να γραφεί ως $m = \lambda_1 X_1 + \cdots + \lambda_n X_n$ για οποιαδήποτε $\lambda_1, \cdots, \lambda_n \in \mathbb{R}$. Συνεπώς οι συνθήκες καθετότητας μπορεί να γραφούν ως

$$\mathbb{E}_\mu[(X - \sum_{i=1}^n \lambda_i^* X_i)(\sum_{j=1}^n \lambda_j X_j] = 0, \ \forall \lambda_1, \cdots, \lambda_n \in \mathbb{R}$$

Αυτό σημαίνει ότι έχουμε

$$\mathbb{E}_\mu[(X - \sum_{i=1}^n \lambda_i^* X_i) X_j] = 0, \ \ j = 1, \cdots, n$$

Η σχέση αυτή μπορεί να γραφεί και ως

$$\mathbb{E}_\mu[X\, X_j] - \sum_{i=1}^n \lambda_i^* \, \mathbb{E}_\mu[X_i \, X_j] \ \ j = 1, \cdots, n$$

που μπορεί να ερμηνευθούν σαν ένα σύστημα απο n γραμμικές εξισώσεις για n αγνώστους. Η λύση αυτών μας δίνει τα λ_i^* τα οποία μας δίνουν και την εκτιμήτρια.

Στην περίπτωση όπου θεωρούμε ότι οι τυχαίες μεταβλητές έχουν μέσο 0 η εξίσωση αυτή μπορεί να ξαναγραφεί ως

$$\sum_{i=1}^n Cov(X_i X_j) \lambda_i^* = \mathbb{E}_\mu[X\, X_j] =: R_j, \ \ j = 1, \cdots, n$$

που είναι οι κλασσικές εξισώσεις της γραμμικής παλινδρόμησης.

6.5.2 Υπο συνθήκη μέση τιμή

Ορισμός 6.5.1 *Έστω $(\Omega, \mathcal{F}, \mu)$ ένας χώρος πιθανοτήτων και $\mathcal{G} \subset \mathcal{F}$.*
Αν X είναι οποιαδήποτε τυχαία μεταβλητή $X \in L^2(\Omega, \mathcal{F}, \mu)$ λέμε ότι η τυχαία μεταβλητη $\hat{X} \in L^2(\Omega, \mathcal{G}, \mu)$ είναι η υπο συνθήκη μέση τιμή της X δεδομένης της σ-άλγεβρας $\mathcal{G}$ αν ισχύει

$$\mathbb{E}_\mu[X\, Z] = \mathbb{E}_\mu[\hat{X}\, Z], \ \ \forall\ Z \in L^2(\Omega, \mathcal{G}, \mu)$$

Συμβολίζουμε $\hat{X} := \mathbb{E}_\mu[X \mid \mathcal{G}]$.

Μπορούμε να γράψουμε την σχέση αυτή με την μορφή

$$\mathbb{E}_\mu[(X - \hat{X})\, Z] = 0, \ \ \forall\ Z \in L^2(\Omega, \mathcal{G}, \mu) \tag{6.2}$$

Παρατηρούμε ότι αν ορίσουμε σαν H τον χώρο Hilbert $L^2(\Omega, \mathcal{F}, \mu)$ και σαν M τον υπόχωρο του $L^2(\Omega, \mathcal{G}, \mu)$ τότε η συνθήκη (6.2) δεν είναι τίποτε άλλο παρά μία συνθήκη καθετότητας, σύμφωνα με την οποία αν $x = X$ είναι οποιοδήποτε στοιχείο του H, το $\hat{X} = m^*$ είναι αυτό το στοιχείο του M για το οποίο να ισχύει ότι $x - m^*$ είναι κάθετο σε κάθε $m = Z \in M$, δηλαδή

$$< x - m^*, m >= 0 \ \ \forall\ m \in M$$

Όμως απο το θεώρημα της προβολής βλέπουμε ότι το m^* θα είναι η προβολή του x στον υπόχωρο M δηλαδή θα είναι το μοναδικό στοιχείο του M για το οποίο ελαχιστοποιείται η ποσότητα $|| \ x - m \ ||$ για όλα τα πιθανά $m \in M$.
Συνεπώς,

Πρόταση 6.5.1 *Έστω $(\Omega, \mathcal{F}, \mu)$ ένας χώρος πιθανότητας, $\mathcal{G} \subset \mathcal{F}$ μία σ-ύποάλγεβρα της $\mathcal{F}$ και $X \in L^2(\Omega, \mathcal{F}, \mu)$. Η υπο συνθήκη μέση τιμή $\mathbb{E}_\mu[X \mid \mathcal{G}]$ της X, είναι η προβολή της X στον υπόχωρο $M = L^2(\Omega, \mathcal{G}, \mu)$.*

Η ερμηνεία αυτή πρώτον μας εξασφαλίζει την ύπαρξη της υπο συνθήκη μέσης τιμής, για μία οποιαδήποτε τυχαία μεταβλητή αρκεί να είναι τετραγωνικά ολοκληρώσιμη. Για να το δούμε αυτό δεν έχουμε παρά να εφαρμόσουμε το κομμάτι της ύπαρξης για την προβολή. Αυτό απο μόνο του είναι πολύ σημαντικό.

Η ερμηνεία όμως της υπο συνθήκη μέσης τιμής σαν προβολή κάνει και κάτι παραπάνω. Μπορεί να χαρακτηρίσει της τυχαία μεταβλητή $\hat{X} := \mathbb{E}_\mu[X \mid \mathcal{G}]$ σαν την μοναδική $\mathcal{G}$-μετρήσιμη τυχαία μεταβλητή για την οποία ελαχιστοποιείται η ποσότητα $|| \ X - Y \ ||_{L^2(\mu)}$ επάνω σε όλες τις τυχαίες μεταβλητές Y οι οποίες είναι τετραγωνικά ολοκληρώσιμες και $\mathcal{G}$-μετρήσιμες. Δηλαδή

$$\inf_{Y \in L^2(\Omega, \mathcal{G}, \mu)} \mathbb{E}_\mu[(X - Y)^2] = \mathbb{E}_\mu[(X - \mathbb{E}_\mu[X \mid \mathcal{G}])^2]$$

Η νόρμα αυτή μπορεί να ερμηνευθεί σαν το μέσο τετραγωνικό σφάλμα κατά την προσέγγιση της $\mathcal{F}$-μετρήσιμης τυχαίας μεταβλητής X απο την $\mathcal{G}$-μετρήσιμη τυχαία μεταβλητή Y. Συνεπώς η $\hat{X} := \mathbb{E}[X \mid \mathcal{G}]$ είναι η καλύτερη εκτιμήτρια της τυχαίας μεταβλητής X δεδομένης της πληροφορίας που περιέχεται στην $\mathcal{G}$ (και που βέβαια δεν είναι επαρκής για να μας επιτρέψει την πληρη περιγραφή της) υπό την έννοια ότι ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα της εκτίμησης.

6.6 Οι χώροι $L^p(\mu)$

Θα ασχοληθούμε τώρα με τους χώρους των συναρτήσεων (τυχαίων μεταβλητών) X για τις οποίες το ολοκλήρωμα της $| \ X \ |^p$, $p \geq 1$ είναι πεπερασμένο.

Ορισμός 6.6.1 *Έστω χώρος μέτρου $(\Omega, \mathcal{F}, \mu)$. Θα συμβολίζουμε με $L^p(\mu)$, $p \geq 1$ το σύνολο των συναρτήσεων $X : \Omega \to \mathbb{R}$ οι οποίες είναι μετρήσιμες ως προς την σ-άλγεβρα $\mathcal{F}$ και για τις οποίες ισχύει*

$$\int_\Omega | \ X(\omega) \ |^p, d\mu < \infty$$

Η απεικόνιση $|| \cdot ||_{L^p(\mu)}$: $L^p(\mu) \to \mathbb{R}^+$, που ορίζεται ως $|| \ X \ ||_{L^p(\mu)} := \{\int_\Omega | \ X(\omega) \ |^p, d\mu\}^{\frac{1}{p}}$ ονομάζεται νόρμα του χώρου $L^p(\mu)$.
Ο χώρος αυτός μπορεί να εφοδιαστεί με την μετρική $d(X_1, X_2) = \{\int_\Omega | \ X_1 - X_2 \ |^p, d\mu\}^{\frac{1}{p}} = || \ X_1 - X_2 \ ||_{L^p(\mu)}$.

Σχόλιο 6.6.1 *Για $p = 1$ και $p = 2$ παίρνουμε τους χώρους $L^1(\mu)$ και $L^2(\mu)$ που είδαμε στα προηγούμενα.*

Στην περίπτωση όπου $p = \infty$ παίρνουμε τον χώρο $L^\infty(\mu)$ η νόρμα του οποίου είναι η $|| X ||_{L^\infty} = \sup_{\omega\in\Omega} | X(\omega) |$. Συνεπώς, ο χώρος $L^\infty(\mu)$ περιέχει τις φραγμένες συναρτήσεις (τυχαίες μεταβλητές).

Οι χώροι αυτοι συνδέονται μεταξύ τους με μία σημαντική ανισότητα, την ανισότητα του Hölder .

Πρόταση 6.6.1 (Ανισότητα Hölder)
Ας πάρουμε δύο αριθμούς p, q για τους οποίους ισχύει $\frac{1}{p} + \frac{1}{q} = 1$, $p, q \geq 1$.
Έστω $(\Omega, \mathcal{F}, \mu)$ χώρος μέτρου (πιθανοτήτων) και X, Y μετρήσιμες συναρτήσεις (τυχαίες μεταβλητές) τέτοιες ώστε $X \in L^p(\mu)$ και $Y \in L^q(\mu)$.
Τότε, $XY \in L^1(\mu)$ και μάλιστα

$$|| XY ||_{L^1(\mu)} \leq || X ||_{L^p(\mu)} || Y ||_{L^q(\mu)}$$

Μπορούμε τώρα να ορίσουμε την έννοια της σύγκλισης μιας ακολουθίας στοιχείων του $L^p(\mu)$, $p \geq 1$.

Ορισμός 6.6.2 *Έστω $X^{(n)}$ μία ακολουθία στοιχείων του $L^p(\mu)$. Θα λέμε ότι η ακολουθία $X^{(n)}$ συγκλίνει στο στοιχείο $X \in L^p(\mu)$ αν $|| X^{(n)} - X ||_{L^p(\mu)} \to 0$ δηλαδή αν για κάθε $\epsilon > 0$ μπορεί να βρεθεί $N \in \mathbb{N}$ τέτοιο ώστε για κάθε $n > N$ ισχύει $|| X^{(n)} - X ||_{L^p(\mu)} < \epsilon$. Αυτό μεταφράζεται ισοδύναμα στο $\{\int_\Omega | X^{(n)}(\omega) - X(\omega) |^p \, d\mu(\omega)\}^{\frac{1}{p}} < \epsilon$ για $n > N$.*

Ορισμός 6.6.3 *Η ακολουθία $X^{(n)} \in L^p(\mu)$ ονομάζεται Cauchy αν για κάθε $\epsilon > 0$ υπάρχει $N \in \mathbb{N}$ τέτοιο ώστε $|| X^{(n)} - X^{(m)} ||_{L^p(\mu)} < \epsilon$ για $n, m > N$.*

Θεώρημα 6.6.1 *Οι χώροι $L^p(\mu)$ είναι πλήρεις, δηλαδή για κάθε ακολουθία Cauchy $X^{(n)} \in L^p(\mu)$ υπάρχει $X \in L^p(\mu)$ τέτοια ώστε $X^{(n)} \to X$.*

6.7 Ορισμένα σχόλια σχετικά με την σύγκλιση

Είδαμε την έννοια της σύγκλισης στον χώρο $L^p(\mu)$. Μία άλλη φυσιολογική έννοια σύγκλισης είναι η λεγόμενη σημειακή σύγκλιση, δηλαδή αν έχουμε την ακολουθία συναρτήσεων $X^{(n)}(\omega)$ να υπάρχει $X(\omega)$ τέτοιο ώστε $X^{(n)}(\omega) \to X(\omega)$, σχεδόν παντού. Με άλλα λόγια $\mu(\{\omega : X^{(n)}(\omega) \not\to X(\omega)\}) = 0$. Η έννοια της σύγκλισης αυτής, είναι πολύ χρήσιμη στις πιθανότητες, και ονομάζεται σχεδόν βέβαιη σύγκλιση ή σύγκλιση με πιθανότητα 1.

Είναι ισοδύναμες αυτές οι δύο έννοιες της σύγκλισης; Δηλαδή, αν υποθέσουμε ότι $X^{(n)} \to X$ σ.β. μπορούμε να εξασφαλίσουμε και ότι $X^{(n)} \to X$ στον L^p, $p \geq 1$; Η απάντηση είναι αρνητική!

6.8 Αποδείξεις ορισμένων αποτελεσμάτων

6.8.1 Απόδειξη του Θεωρήματος 6.4.1

Θα ξεκινήσουμε δείχνοντας την ιδιότητα 2 υποθέτωντας αρχικά την ύπαρξη του m^*.
Ας υποθέσουμε το αντίθετο, δηλαδή ότι αν m^* ελαχιστοποιεί το $|| x - m ||_H$ υπάρχει κάποιο $m \in M$ το οποίο να μην είναι ορθογώνιο στο $x - m^*$, δηλαδή ότι υπάρχει κάποιο $m \in M$ όπου $(x - m^*, m)_H = \delta \neq 0$ για κάποιο $\delta \in \mathbb{R}$.
Χωρίς βλάβη της γενικότητας, μπορούμε να θεωρήσωουμε ότι $|| m ||_H = 1$ (αλλιώς απλά εργάζομαι με το $\frac{m}{||m||_H}$.
Ορίζουμε

$$m_1 = m^* + \delta\, m$$

για το οποίο ισχύει ότι $m_1 \in M$ (απο τον ορισμό του υπόχωρου).
Ας υπολογίσουμε την απόσταση $|| x - m_1 ||_H$. Χρησιμοποιώντας τις ιδιότητες του εσωτερικού γινομένου

$$\begin{aligned}
|| x - m_1 ||_H^2 &= || x - m^* - \delta\, m ||_H^2 \\
&= || x - m^* ||_H^2 - 2(x - m^*, \delta\, m)_H + | \delta |^2 \\
&= || x - m^* ||_H^2 - 2 | \delta |^2 + | \delta |^2 \\
&= || x - m^* ||_H^2 - | \delta |^2 < || x - m^* ||_H^2
\end{aligned}$$

Αυτό όμως σημαίνει ότι υπάρχει κάποιο άλλο στοιχείο του M, το m_1 το οποίο έχει μικρότερη απόσταση απο το x απο το m^*. Αυτό είναι άτοπο εφόσον έχω δεχθεί ότι το m^* είναι αυτό το στοιχείο του M το οποίο έχει την μικρότερη απόσταση απο το x. Συνεπώς, δεν είναι δυνατόν να ισχύει $(x-m^*,m)_H \neq 0$ για οποιοδήποτε $m \in M$. Αντίστροφα, έστω ότι το m^* είναι τέτοιο ώστε $(x-m^*,m)_H = 0$ για κάθε $m \in M$. Απο τις ιδιότητες του εσωτερικού γινομένου έχουμε ότι για οποιοδήποτε $m \in M$ ισχύει

$$|| x-m ||_H^2 = || x-m^*+m^*-m ||_H^2 = || x-m^* ||_H^2 + || m^*-m ||_H^2 \tag{6.3}$$

όπου χρησιμοποιήσαμε την καθετότητα του $x-m^*$ με το $m \in M$.
Η σχέση (6.3) όμως μας δείχνει ότι για κάθε $m \in M$ ισχύει

$$|| x-m^* ||_H < || x-m ||_H$$

συνεπώς το m^* είναι το στοιχείο αυτό του υποχώρου M για το οποίο ελαχιστοποιείται η απόσταση απο το $x \in H$.

Θα αποδείξουμε τώρα το 1.
Έστω ότι $x \notin M$. Ας ορίσουμε

$$\delta := \inf_{m \in M} || x-m ||_H$$

η οποία υπάρχει εφόσον η ποσότητα $|| x-m ||_H$ είναι φραγμένη απο τα κάτω απο το 0.
Θα δείξουμε ότι το μέγιστο κάτω φράγμα αυτό επιτυγχάνεται, δηλαδή ότι υπάρχει $m^* \in M$ τέτοιο ώστε $|| x-m^* ||_H = \delta$.
Θα προσεγγίσουμε το στοιχείο αυτό με μία ακολουθία $m^{(n)} \in M$ τέτοια ώστε $|| x-m^{(n)} ||_H \to \delta$. Προσπαθήστε να σκεφτείτε γιατί είμαστε σίγουροι ότι μία τέτοια ακολουθία υπάρχει.
Απο τον κανόνα του παραλληλογράμμου

$$|| x_1+x_2 ||_H^2 + || x_1-x_2 ||_H^2 = 2 || x_1 ||_H^2 + 2 || x_2 ||_H^2$$

έχουμε ότι

$$|| (m^{(j)}-x)+(x-m^{(i)}) ||_H^2 + || (m^{(j)}-x)-(x-m^{(i)}) ||_H^2 = 2 || m^{(j)}-x ||_H^2 + 2 || x-m^{(i)} ||_H^2$$

Η σχέση αυτή μπορεί να ξαναγραφεί ως

$$|| m^{(i)}-m^{(j)} ||_H^2 = 2 || m^{(j)}-x ||_H^2 + 2 || x-m^{(i)} ||_H^2 - 4 \left|\left| x - \frac{m^{(i)}+m^{(j)}}{2} \right|\right|_H \tag{6.4}$$

Επειδή $\frac{m^{(i)}+m^{(j)}}{2} \in M$ απο τον ορισμό του δ έχουμε ότι

$$\left|\left| x - \frac{m^{(i)}+m^{(j)}}{2} \right|\right|_H \geq \delta$$

Απο την παρατήρηση αυτή, και το γεγονός ότι $|| m^{(j)}-x ||_H \to \delta$ η σχέση (6.4) μας εξασφαλίζει ότι

$$|| m^{(i)}-m^{(j)} ||_H^2 \to 0, \quad i,j \to \infty$$

δηλαδή η $m^{(i)}$ είναι ακολουθία Cauchy . Η ακολουθία αυτή συγκλινει στον H και επειδή ο M είναι υπόχωρος, αν ονομάσουμε m^* το όριο της ακολουθίας είμαστε σίγουροι ότι $m^* \in M$.
Θα δείξουμε ότι για το m^* αυτό ισχύει ότι $|| x-m^* ||_H = \delta$. Αυτό συμβαίνει γιατί $|| \cdot ||_H \colon H \to \mathbb{R}_+$ είναι συνεχής συνάρτηση.

6.8.2 Απόδειξη του θεωρήματος 6.6.1

Έστω $X^{(n)}$ μία οποιαδήποτε ακολουθία Cauchy στον $L^p(\mu)$. Αρκεί να δείξουμε ότι υπάρχει μία υποακολουθία $X^{(n_k)}$ και κάποιο (μοναδικό) $X \in L^p(\mu)$ τέτοιο ώστε $X^{(n_k)} \to X$.
Ας δούμε πως μπορούμε να επιλέξουμε την ακολουθία αυτή. Επιλέγουμε ένα n_1 τέτοιο ώστε $|| X^{(n_1)} - X^{(n)} ||_{L^p(\mu) \leq \frac{1}{2}}$ για κάθε $n \geq n_1$. Αυτό είναι εφικτό εφόσον η ακολουθία $X^{(n)}$ είναι ακολουθία Cauchy . Επιλέγουμε μετα ένα

n_2 τέτοιο ώστε $\| X^{(n_2)} - X^{(n)} \|_{L^p(\mu) \leq \frac{1}{4}}$ για κάθε $n \geq n_2$. Με παρόμοιο τρόπο επιλέγουμε ένα n_k τέτοιο ώστε $\| X^{(n_k)} - X^{(n_{k+1})} \|_{L^p(\mu) \leq \frac{1}{2^k}}$, $k = 1, 2, \cdots$. Ορίζουμε τώρα την μονότονη ακολουθία συναρτήσεων (τυχαίων μεταβλητών)

$$Y^{(m)}(\omega) := \mid X^{(n_1)}(\omega) \mid + \sum_{k=1}^{m} \mid X^{(n_k)}(\omega) - X^{(n_{k+1})}(\omega) \mid$$

Χρησιμοποιούμε τώρα την τριγωνική ανισοτητα για την νόρμα και παίρνουμε

$$\| Y^{(m)} \|_{L^p(\mu)} \leq \| X^{(n_1)} \|_{L^p}\ (\mu) + \sum_{k=1}^{m} \frac{1}{2^k} \leq \| X^{(n_1)} \|_{L^p}\ (\mu) + 1$$

Η ακολουθία συναρτήσεων $Y^{(m)}$ είναι μονότονη, συνεπώς και οι $\mid Y^{(m)} \mid^p$, $p \geq 1$ θα είναι επίσης μία μονότονη ακολουθία και απο το θεώρημα μονότονης σύγκλισης εξασφαλίζεται ότι θα υπάρχει μία συνάρτηση $Y \in L^p(\mu)$ τέτοια ώστε $Y^{(m)} \to Y$.

Θα δείξουμε τώρα ότι και η υποακολουθία $X^{(n_k)}$ συγκλίνει σε κάποιο $X \in L^p(\mu)$. Πραγματικά, ας γράψουμε την $X^{n_{k+1}}$ σαν

$$X^{(n_{k+1})} = X^{(n_1)} + (X^{(n_2)} - X^{(n_1)}) + \cdots + (X^{(n_{k+1})} - X^{n_k})$$

Παρατηρούμε ότι η ακολουθία αυτή συγκλίνει απόλυτα σημειακά (απο τα προηγούμενα) και συνεπώς θα συγκλίνει σε κάποιο $X(\omega)$ σημειακά. Θα πρέπει να δείξουμε ότι η σύγκλιση είναι στον $L^p(\mu)$ και ότι $X \in L^p(\mu)$.

Είναι προφανές ότι $\mid X^{(n_k)} \mid \leq Y$ σημειακά και επειδή $Y \in L^p(\mu)$ απο το θεώρημα κυριαρχημένης σύγκλισης έχουμε ότι $X \in L^p(\mu)$. Επίσης απο κυριαρχημένη σύγκλιση $\| X^{(n_k)} - X \|_{L^p(\mu) \to 0}$ (εφόσον $\mid X^{(n_k)} - X \mid \leq Y + \mid X \mid \in L^p(\mu)$). Συνεπώς, η υποακολουθία $X^{(n_k)}$ συγκλίνει στο X στον L^p.