Τζιατζιος Νικολαος

Η Πιθανοθεωρητικη Μεθοδος

Μεταπτυχιακη Εργασια

Πανεπιστήμιο Αιγαίου, Τμήμα Μαθηματικών
Σάμος 2 Ιουνίου 2005

Εισηγητης: Τσολομύτης Αντώνης

Επιτροπη

Ανούσης Μιχάλης

Τσολομύτης Αντώνης

Φελουζής Ευάγγελος

Στη μνήμη του αγαπημένου μου πατέρα

Περιεχόμενα

Πρόλογος

Ο όρος «πιθανοθεωρητική μέθοδος» χρησιμοποιείται για ένα γενικό σχήμα απόδειξης. Το ζητούμενο είναι να αποδείξουμε την ύπαρξη κάποιας δομής με προκαθορισμένες ιδιότητες. Αντί να κατασκευάσουμε συγκεκριμένο παράδειγμα τέτοιας δομής, κατασκευάζουμε μια οικογένεια υποψηφίων δομών. Την εφοδιάζουμε με ένα μέτρο πιθανότητας και εξετάζοντας την τυπική συμπεριφορά των μελών της οικογένειας αποδεικνύουμε ότι, με θετική πιθανότητα έχουν τις ιδιότητες που ζητάμε.

Η πιθανοθεωρητική μέθοδος εμφανίζεται όλο και συχνότερα σε διάφορους κλάδους των μαθηματικών, χρησιμοποιείται για την απόδειξη προτάσεων που κατ' αρχήν μοιάζουν να μην έχουν καμία σχέση με τη θεωρία πιθανοτήτων. Η μέθοδος χρησιμοποιείται συστηματικά στην συνδυαστική, τη θεωρία αριθμών, την αρμονική ανάλυση και την ασυμπτωτική γεωμετρική ανάλυση.

Τελειώνοντας, θα ήθελα να ευχαριστήσω τα μέλη της εξεταστικής επιτροπής κο. Ανούση Μιχάλη και κο. Φελουζή Ευάγγελο για τον χρόνο που διάθεσαν και τις εύστοχες παρατηρήσεις τους. Τον κο. Γιαννόπουλο Αποστόλη, ο οποίος μου έδωσε την ιδέα για το θέμα της εργασίας, με τις διαλέξεις και τις σημειώσεις του στα πλαίσια του θερινού σχολείου που πραγματοποιήθηκε στο Καρλόβασι Σάμου το καλοκαίρι του 2004. Ιδιαίτερα θα ήθελα να ευχαριστήσω τον εισηγητή αυτής της μεταπτυχιακής εργασίας κο. Τσολομύτη Αντώνη, χωρίς τον οποίο η εκπόνησή της θα καθίστατο αδύνατη.

Ν. Τζιάτζιος, Σάμος 2005.

Κεφάλαιο 1

Η πιθανοθεωρητική μέθοδος

1.1 Στοιχεία θεωρίας πιθανοτήτων

Σε αυτή την παράγραφο θα δούμε μερικά βασικά στοιχεία από τη θεωρία πιθανοτήτων.

Ορισμός 1.1. *Έστω Ω μη κενό σύνολο και $\mathcal{A}$ μια μη κενή οικογένεια υποσυνόλων του Ω. Η $\mathcal{A}$ λέγεται **σ-άλγεβρα** του Ω αν ικανοποιεί τα εξής:*

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

(ii) *Αν* $A \in \mathcal{A}$ *τότε και* $\Omega \setminus A \in \mathcal{A}$.

(iii) *Αν* $A_1, A_2, \dots, A_n, \dots \in \mathcal{A}$ *τότε και* $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

Παρατήρηση: Αν $A_1, A_2, \dots, A_n, \dots \in \mathcal{A}$ από τους κανόνες *De'Morgan* προκύπτει ότι και $\bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$.

Ορισμός 1.2. *Έστω $\mathcal{A}$ σ-άλγεβρα του Ω. Η απεικόνιση $\mathbb{P} : \mathcal{A} \to \mathbb{R}$ λέγεται **μέτρο πιθανότητας** στην $\mathcal{A}$ αν ικανοποιεί τα εξής:*

(i) $\mathbb{P}(\Omega) = 1$.

(ii) $\mathbb{P}(A) \geq 0$, *για κάθε* $A \in \mathcal{A}$.

(iii) *Αν* $A_1, A_2, \dots, A_n, \dots \in \mathcal{A}$ *ξένα μεταξύ τους, τότε*

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

Ορισμός 1.3. *Η τριάδα $(\Omega, \mathcal{A}, \mathbb{P})$ λέγεται **χώρος πιθανότητας**, όταν*

- *το Ω είναι μη κενό σύνολο.*
- *η $\mathcal{A}$ είναι σ-άλγεβρα στο Ω.*
- *το $\mathbb{P}$ είναι μέτρο πιθανότητας στην $\mathcal{A}$.*

Τα στοιχεία της $\mathcal{A}$ λέγονται **ενδεχόμενα** και τα στοιχεία του Ω λέγονται **στοιχειώδη** ή **απλά ενδεχόμενα**. Αν το A είναι ενδεχόμενο, τότε ο αριθμός $\mathbb{P}(A)$ είναι η **πιθανότητα** του A.
Παράδειγμα: Έστω Ω πεπερασμένο σύνολο και $\mathcal{A}$ το δυναμοσύνολο του Ω. Αν $p : \Omega \to [0,1]$ μια συνάρτηση με την ιδιότητα

$$\sum_{\omega \in \Omega} p(\omega) = 1,$$

τότε η απεικόνιση $\mathbb{P} : \mathcal{A} \to [0,1]$ με

$$\mathbb{P}(A) = \sum_{\omega \in A} p(\omega)$$

είναι ένα μέτρο πιθανότητας.

Ένα μέτρο πιθανότητας αυτού του είδους είναι η ομοιόμορφη κατανομή στο Ω. Σε αυτή την περίπτωση η πιθανότητα κάθε ενδεχομένου A είναι

$$\mathbb{P} = \frac{|A|}{|\Omega|},$$

όπου με $|B|$ συμβολίζουμε τον πληθάριθμο ενός πεπερασμένου συνόλου B.

Μια σημαντική ιδιότητα του μέτρου πιθανότητας, που θα χρησιμοποιούμε συχνά στο εξής, είναι η υποαθροιστικότητα, η οποία εκφράζεται από το παρακάτω λήμμα.

Λήμμα 1.1. *Έστω $A_1, A_2, \ldots, A_n \in \mathcal{A}$, τότε*

$$\mathbb{P}\Big(\bigcup_{i=1}^{n} A_i\Big) \leq \sum_{i=1}^{n} \mathbb{P}(A_i).$$

Απόδειξη: Θέτουμε $B_1 = A_1$ και για κάθε $i = 2, \ldots, n$ θέτουμε

$$B_i = A_i \backslash (A_1 \cup A_2 \cup \ldots \cup A_{i-1}).$$

Τότε προφανώς

$$\bigcup_{i=1}^{n} B_i = \bigcup_{i=1}^{n} A_i, \qquad \mathbb{P}(B_i) \leq \mathbb{P}(A_i)$$

και τα ενδεχόμενα $B_1, \ldots, B_n$ είναι ξένα μεταξύ τους. Από την προσθετικότητα του μέτρου έχουμε

$$\mathbb{P}\Big(\bigcup_{i=1}^{n} A_i\Big) = \mathbb{P}\Big(\bigcup_{i=1}^{n} B_i\Big) = \sum_{i=1}^{n} \mathbb{P}(B_i) \leq \sum_{i=1}^{n} \mathbb{P}(A_i).$$

□

Ορισμός 1.4. *Τα ενδεχόμενα A, B λέγονται **ανεξάρτητα** αν*

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

Γενικότερα, τα ενδεχόμενα $A_1, \ldots, A_n$ λέγονται ανεξάρτητα αν για κάθε σύνολο δεικτών $I \subseteq \{1, \ldots, n\}$ έχουμε

$$\mathbb{P}\Big(\bigcap_{i \in I} A_i\Big) = \prod_{i \in I} \mathbb{P}(A_i).$$

Ορισμός 1.5. *Έστω A, B ενδεχόμενα με $\mathbb{P}(B) > 0$. Η **δεσμευμένη πιθανότητα** του A δεδομένου του B ορίζεται ως*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Παρατηρείστε ότι αν A, B ανεξάρτητα τότε $\mathbb{P}(A|B) = \mathbb{P}(A)$.

Ορισμός 1.6. *Έστω $(\Omega, \mathcal{A}, \mathbb{P})$ χώρος πιθανότητας και $X : \Omega \to \mathbb{R}$ μια συνάρτηση. Η X λέγεται **πραγματική τυχαία μεταβλητή** αν για κάθε $a \in \mathbb{R}$ το σύνολο $\{\omega \in \Omega : \ X(\omega) \leq a\}$ είναι στοιχείο της $\mathcal{A}$.*

Ορισμός 1.7. *Έστω X πραγματική τυχαία μεταβλητή. Η **μέση τιμή** της X ορίζεται να είναι η ποσότητα*

$$\mathbb{E}(X) = \int_{\Omega} X(\omega) d\mathbb{P}(\omega).$$

Παρατήρηση: Αν ο χώρος πιθανότητας είναι πεπερασμένος, τότε κάθε συνάρτηση $X : \Omega \to \mathbb{R}$ είναι τυχαία μεταβλητή και έχει μέση τιμή

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} p(\omega) X(\omega).$$

Στο εξής θα χρησιμοποιούμε συχνά την εξής σημαντική ιδιότητα: Αν $\mathbb{E}(X) \geq a$ τότε υπάρχει $\omega \in \Omega$ τέτοιο ώστε $X(\omega) \geq a$. Πράγματι, έστω ότι αυτό δεν ισχύει. Τότε για κάθε $\omega \in \Omega$ έχουμε ότι $X(\omega) < a$ και συνεπώς

$$\int_{\Omega} X(\omega) d\mathbb{P}(\omega) < \int_{\Omega} a d\mathbb{P}(\omega) \Rightarrow \mathbb{E}(X) < a\mathbb{P}(\Omega) = a,$$

πράγμα που είναι άτοπο αφού υποθέσαμε ότι $\mathbb{E}(X) \geq a$. Εντελώς ανάλογα δείχνουμε και ότι αν $\mathbb{E}(X) < a$ τότε υπάρχει $\omega \in \Omega$ τέτοιο ώστε $X(\omega) < a$.

Ορισμός 1.8. *Οι τυχαίες μεταβλητές X και Y λέγονται **ανεξάρτητες** αν για κάθε ζευγάρι A, B υποσυνόλων του $\mathbb{R}$ ισχύει*

$$\mathbb{P}(X \in A \text{ και } Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B).$$

Για να ελέγξουμε την ανεξαρτησία δύο τυχαίων μεταβλητών αρκεί να θεωρήσουμε σύνολα της μορφής $A = (-\infty, a]$ και $B = (-\infty, b]$ όπου $a, b \in \mathbb{R}$. Δηλαδή, αν για κάθε $a, b \in \mathbb{R}$ ισχύει

$$\mathbb{P}(X \leq a \text{ και } Y \leq b) = \mathbb{P}(X \leq a)\mathbb{P}(Y \leq b)$$

τότε οι X και Y είναι ανεξάρτητες.

Λήμμα 1.2. *Έστω X, Y τυχαίες μεταβλητές και $a, b \in \mathbb{R}$. Τότε*

$$\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y).$$

Η απόδειξη είναι άμεση συνέπεια της γραμμικότητας του ολοκληρώματος.

Ορισμός 1.9. *Έστω A ενδεχόμενο. Ορίζουμε τη **δείκτρια τυχαία μεταβλητή** $I_A(\omega)$ του A ως εξής:*

$$I_A(\omega) = \begin{cases} 1, & \text{αν } \omega \in A \\ 0, & \text{αν } \omega \notin A \end{cases}$$

Παρατήρηση: Ισχύει ότι $\mathbb{E}(I_A) = \mathbb{P}(A)$. Πράγματι,

$$\begin{aligned} \mathbb{E}(I_A) &= \int_\Omega I_A(\omega) d\mathbb{P}(\omega) \\ &= \int_A I_A(\omega) d\mathbb{P}(\omega) + \int_{\Omega \setminus A} I_A(\omega) d\mathbb{P}(\omega) \\ &= \int_A 1 d\mathbb{P}(\omega) + \int_{\Omega \setminus A} 0 d\mathbb{P}(\omega) \\ &= \int_A d\mathbb{P}(\omega) \\ &= \mathbb{P}(A). \end{aligned}$$

Λήμμα 1.3. *Αν X, Y ανεξάρτητες τυχαίες μεταβλητές τότε*

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y).$$

Απόδειξη: **(Για πεπερασμένους χώρους)** Έστω V_X και V_Y τα (πεπερασμένα) σύνολα τιμών των X και Y αντίστοιχα. Επειδή X, Y είναι ανεξάρτητες, για κάθε $a \in V_X$ και για κάθε $b \in V_Y$ έχουμε

$$\mathbb{P}(X = a \text{ και } Y = b) = \mathbb{P}(X = a)\mathbb{P}(Y = b).$$

Τώρα,

$$\begin{aligned}
\mathbb{E}(XY) &= \sum_{a \in V_X,\ b \in V_Y} ab\mathbb{P}(X = a \text{ και } Y = b) \\
&= \sum_{a \in V_X,\ b \in V_Y} ab\mathbb{P}(X = a)\mathbb{P}(Y = b) \\
&= \left(\sum_{a \in V_X} a\mathbb{P}(X = a)\right)\left(\sum_{b \in V_Y} b\mathbb{P}(Y = b)\right) \\
&= \mathbb{E}(X)\mathbb{E}(Y).
\end{aligned}$$

□

Η απόδειξη στη γενική περίπτωση είναι ανάλογη, αλλά τεχνικά δυσκολότερη.

1.2 Τριγωνικά Πολυώνυμα

Ορισμός 1.10. *Έστω $1 \le p \le \infty$, η **p-νόρμα** της f ορίζεται ως,*

$$\|f\|_p = \left(\frac{1}{2\pi} \int_0^{2\pi} |f(x)|^p dx \right)^{\frac{1}{p}}.$$

Θεώρημα 1.1. *(Ανισότητα Hölder): Έστω $g, h : [0, 2\pi] \to \mathbb{R}$ ολοκληρώσιμες συναρτήσεις, αν $p > 1$ και q συζυγής εκθέτης του p δηλαδή $\frac{1}{p} + \frac{1}{q} = 1$, τότε ισχύουν τα εξής:*

(i)

$$\frac{1}{2\pi} \int_0^{2\pi} gh \le \left(\frac{1}{2\pi} \int_0^{2\pi} g^p \right)^{\frac{1}{p}} \left(\frac{1}{2\pi} \int_0^{2\pi} h^q \right)^{\frac{1}{q}}.$$

(ii) *Αν $p, q, r \ge 1$ και $r = ap + (1-a)q$ για $a \in (0, 1)$. Τότε,*

$$\frac{1}{2\pi} \int_0^{2\pi} |f|^r = \frac{1}{2\pi} \int_0^{2\pi} \Big(|f|^p\Big)^a \Big(|f|^q\Big)^{1-a} \le \left(\frac{1}{2\pi} \int_0^{2\pi} |f|^p \right)^a \left(\frac{1}{2\pi} \int_0^{2\pi} |f|^q \right)^{1-a}.$$

Θα αποδείξουμε το εξής θεώρημα του *Uchiyama*.

Θεώρημα 1.2. *Υπάρχει σταθερά $c > 0$ ώστε να ισχύουν τα εξής:*
Αν $A = \{n_1 < \cdots < n_N\} \subseteq \mathbb{N}$, τότε υπάρχει $E \subseteq \{1, \ldots, N\}$ τέτοιο ώστε,

$$\Big\| \sum_{j \in E} e^{i n_j x} \Big\|_1 \ge c\sqrt{N}.$$

Απόδειξη: Θεωρούμε τον χώρο $E_2^N = \{-1, 1\}^N$, για κάθε $\varepsilon \in E_2^N$ έχουμε $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_N)$ όπου $\varepsilon_i = \pm 1$, για κάθε $i = 1, \ldots, N$. (*Παρατήρηση:* ο χώρος E_2^N είναι οι κορυφές του κύβου κέντρου 0). Εφοδιάζουμε τον E_2^N με το ομοιόμορφο μέτρο πιθανότητας, δηλαδή αν $A \subseteq E_2^N$ τότε,

$$\mathbb{P}(A) = \frac{|A|}{|E_2^N|} = \frac{|A|}{2^N}$$

όπου $|B|$ ο πληθάριθμος του συνόλου B. Θεωρούμε τη συνάρτηση

$$f_\varepsilon(x) = \sum_{j=1}^N \varepsilon_j e^{i n_j x}, \quad \forall\, \varepsilon \in E_2^N.$$

Η απόδειξη του θεωρήματος θα γίνει σε τρία βήματα.

Βήμα 1:
$$\forall\, \varepsilon \in E_2^N, \quad \|f_\varepsilon\|_2^2 \le \|f_\varepsilon\|_1^{2/3}\, \|f_\varepsilon\|_4^{4/3}.$$

Βήμα 2:
$$\forall\, \varepsilon \in E_2^N, \quad \|f_\varepsilon\|_2^2 = N.$$

Βήμα 3:
$$\text{Υπάρχει} \quad \varepsilon \in E_2^N \quad \text{τέτοιο ώστε} \quad \|f_\varepsilon\|_4^4 \le 2N^2.$$

Αν ισχύουν τα παραπάνω τότε για το τελευταίο ε έχουμε,

$$N = \|f_\varepsilon\|_2^2 \le \|f_\varepsilon\|_1^{2/3}\, \|f_\varepsilon\|_4^{4/3} \le \Big(2N^2\Big)^{1/3} \|f_\varepsilon\|_1^{2/3} \Rightarrow$$

$$N^{1/3} \le 2^{1/3}\, \|f_\varepsilon\|_1^{2/3} \Rightarrow \|f_\varepsilon\|_1 \ge \frac{1}{\sqrt{2}}\sqrt{N}.$$

Έτσι έχουμε βρει

$$f_\varepsilon(x) = \sum_{j=1}^{N} \varepsilon_j e^{in_jx} \quad \text{τέτοιο ώστε} \quad \|f_\varepsilon\|_1 \ge \frac{1}{\sqrt{2}}\sqrt{N}.$$

Θέτουμε, $E_1 = \{j \le N :\ \varepsilon_j = 1\}$ και $E_2 = \{j \le N :\ \varepsilon_j = -1\}$. Τότε,

$$f_\varepsilon(x) = \sum_{j\in E_1} e^{in_jx} - \sum_{j\in E_2} e^{in_jx}.$$

Οπότε,

$$\frac{\sqrt{N}}{\sqrt{2}} \le \Big\|\sum_{j\in E_1} e^{in_jx} - \sum_{j\in E_2} e^{in_jx}\Big\|_1 \le \Big\|\sum_{j\in E_1} e^{in_jx}\Big\|_1 + \Big\|\sum_{j\in E_2} e^{in_jx}\Big\|_1$$

$$\Rightarrow \Big\|\sum_{j\in E_i} e^{in_jx}\Big\|_1 \ge \frac{1}{2\sqrt{2}}\sqrt{N} \qquad \text{για} \quad i = 1 \text{ ή } i = 2.$$

Απόδειξη βήματος 1: Αφού $2 = \frac{2}{3}1 + \frac{1}{3}4$ από την ανισότητα *Hölder* παίρνουμε,

$$\begin{aligned} \|f\|_2^2 &= \frac{1}{2\pi}\int_0^{2\pi} |f(x)|^2 dx = \frac{1}{2\pi}\int_0^{2\pi} |f(x)||f(x)| dx \\ &\le \left(\frac{1}{2\pi}\int_0^{2\pi} |f(x)| dx\right)^{2/3} \left(\frac{1}{2\pi}\int_0^{2\pi} |f(x)|^4 dx\right)^{1/3} \end{aligned}$$

$$\Rightarrow \|f\|_2^2 \leq \|f\|_1^{2/3} \, \|f\|_4^{4/3}.$$

Απόδειξη βήματος 2: Γενικά, ισχύει ότι

$$\Big(\sum_{i=1}^{n} x_i\Big)^2 = \sum_{j,\, i=1}^{n} x_i x_j.$$

Άρα,

$$\begin{aligned}
\|f_\varepsilon\|_2^2 &= \frac{1}{2\pi}\int_0^{2\pi} |f_\varepsilon(x)|^2 dx = \frac{1}{2\pi}\int_0^{2\pi} \Big|\sum_{j=1}^{N} \varepsilon_j e^{in_j x}\Big|^2 dx \\
&= \frac{1}{2\pi}\int_0^{2\pi} \left(\sum_{k=1}^{N} \varepsilon_k e^{in_k x} \sum_{j=1}^{N} \varepsilon_j e^{-in_j x}\right) dx \\
&= \frac{1}{2\pi}\sum_{k,\, j=1}^{N} \varepsilon_k \varepsilon_j \int_0^{2\pi} e^{i(n_k - n_j)x} dx.
\end{aligned}$$

Παρατηρούμε ότι, όταν $k \neq j$ τότε, $\int_0^{2\pi} e^{i(n_k-n_j)x}dx = 0$. Συνεπώς μένουν μόνο οι όροι όταν $k = j$ και άρα και $n_k = n_j$. Οπότε έχουμε

$$\|f_\varepsilon\|_2^2 = \frac{1}{2\pi}\sum_{j=1}^{N} \varepsilon_j^2 \int_0^{2\pi} e^0 dx = \frac{1}{2\pi} 2\pi \sum_{j=1}^{N} 1 = N.$$

Απόδειξη βήματος 3: Με όμοιο τρόπο δείχνουμε ότι,

$$|f_\varepsilon(x)|^4 = \left(\sum_{j,\, k=1}^{N} \varepsilon_j \varepsilon_k e^{i(n_k - n_j)x}\right)^2 = \left(\sum_{s\in\mathbb{Z}} r_\varepsilon(s) e^{isx}\right)^2$$

όπου

$$r_\varepsilon(s) = \sum_{\{(j,k):\, n_j - n_k = s\}} \varepsilon_j \varepsilon_k.$$

Παρατηρούμε ότι $r_\varepsilon(0) = N$ και $r_\varepsilon(-s) = r_\varepsilon(s)$. Τώρα θα δείξουμε ότι

$$\|f_\varepsilon(x)\|_4^4 = \sum_{s\in\mathbb{Z}} \Big(r_\varepsilon(s)\Big)^2.$$

Πράγματι,

$$
\begin{aligned}
\|f_\varepsilon(x)\|_4^4 &= \frac{1}{2\pi}\int_0^{2\pi} |f_\varepsilon(x)|^4 dx \\
&= \frac{1}{2\pi}\int_0^{2\pi} \left(\sum_{j,\,k=1}^{N} \varepsilon_j\varepsilon_k e^{i(n_k-n_j)x}\right)^2 dx \\
&= \frac{1}{2\pi}\int_0^{2\pi} \left(\sum_{s\in\mathbb{Z}} r_\varepsilon(s)e^{isx}\right)^2 dx \\
&= \frac{1}{2\pi}\int_0^{2\pi} \left(\sum_{s,t\in\mathbb{Z}} r_\varepsilon(s)r_\varepsilon(t)e^{i(s+t)x}\right) dx \\
&= \frac{1}{2\pi}\sum_{s,t\in\mathbb{Z}} \left(r_\varepsilon(s)r_\varepsilon(t)\int_0^{2\pi} e^{i(s+t)x}dx\right) \\
&= \sum_{s\in\mathbb{Z}} r_\varepsilon(s)r_\varepsilon(-s) \\
&= \sum_{s\in\mathbb{Z}} \Big(r_\varepsilon(s)\Big)^2.
\end{aligned}
$$

Τώρα θα υπολογίσουμε την μέση τιμή της $\|f_\varepsilon\|_4^4$.

$$
\mathbb{E}(\|f_\varepsilon(x)\|_4^4) = \mathbb{E}\Big(\sum_{s\in\mathbb{Z}}(r_\varepsilon(s))^2\Big) = \sum_{s\in\mathbb{Z}}\mathbb{E}\Big(r_\varepsilon(s)^2\Big).
$$

Αν $s=0$ τότε $\mathbb{E}(r_\varepsilon(s)^2) = \mathbb{E}(r_\varepsilon(0)^2) = \mathbb{E}(N^2) = N^2$. Αν $s\neq 0$ τότε,

$$
\begin{aligned}
\mathbb{E}\Big(r_\varepsilon(s)^2\Big) &= \mathbb{E}\left(\Big(\sum_{n_j-nk=s} \varepsilon_i\varepsilon_k\Big)^2\right) \\
&= \sum_{n_{j_1}-n_{k_1}=s}\ \sum_{n_{j_2}-n_{k_2}=s} \mathbb{E}\Big((\varepsilon_{j_1}\varepsilon_{k_1})(\varepsilon_{j_2}\varepsilon_{k_2})\Big).
\end{aligned}
$$

Όμως $\mathbb{E}\Big((\varepsilon_{j_1}\varepsilon_{k_1})(\varepsilon_{j_2}\varepsilon_{k_2})\Big) \neq 0 \Leftrightarrow j_1 = j_2$ και $k_1 = k_2$. Άρα

$$
\begin{aligned}
\mathbb{E}\Big(r_\varepsilon(s)^2\Big) &= \sum_{n_j-n_k=s} \mathbb{E}\Big((\varepsilon_{j_1}\varepsilon_{k_1})^2\Big) \\
&= \sum_{n_j-n_k=s} 1 \\
&= \Big|\{(j,k):\ n_j-n_k=s\}\Big|.
\end{aligned}
$$

Τελικά σε κάθε περίπτωση έχουμε ότι,

$$\begin{aligned}
\mathbb{E}(\|f_\varepsilon(x)\|_4^4) &= N^2 + \sum_{s\neq 0} |\{(j,k):\ n_j - n_k = s\}| \\
&= N^2 + \big|\{(j,k):\ n_j - n_k \neq 0\}\big| \\
&= N^2 + \big|\{(j,k):\ j \neq k\}\big| \\
&= N^2 + N^2 - N \\
&\leq 2N^2.
\end{aligned}$$

Άρα υπάρχει $\varepsilon \in E_2^N$ τέτοιο ώστε $\|f_\varepsilon(x)\|_4^4 \leq 2N^2$.

Αυτό ολοκληρώνει την απόδειξη του θεωρήματος με σταθερά $c = \sqrt{2}$.

□

Κεφάλαιο 2

Προσεγγιστικές ισοπεριμετρικές ανισότητες

2.1 Συγκέντρωση του μέτρου

Μέχρι στιγμής χρησιμοποιήσαμε την πιθανοθεωρητική μέθοδο για να αποδείξουμε την ύπαρξη κάποιων συνόλων με μια συγκεκριμένη ιδιότητα. Αυτό το πετύχαμε ορίζοντας κατάλληλο χώρο πιθανότητας $(\Omega,\ \mathcal{A},\ \mathbb{P})$ και κατάλληλη τυχαία μεταβλητή $X:\Omega\to\mathbb{R}$ για την οποία ζητούσαμε να υπάρχει $\omega\in\Omega$ τέτοιο ώστε $X(\omega)\leq a$, όπου a κατάλληλος πραγματικός αριθμός. Υπολογίζοντας τη μέση τιμή της X αν καταφέρουμε να δείξουμε ότι $\mathbb{E}(X)\leq a$, τότε εξασφαλίζουμε την ύπαρξη του ω.

Σε πολλά προβλήματα όμως, θέλουμε να αποδείξουμε την ύπαρξη κάποιας δομής που έχει περισσότερες από μία ιδιότητες. Η εφαρμογή της πιθανοθεωρητικής μεθόδου μπορεί να μας οδηγήσει στην εξής κατάσταση: Σε κατάλληλο χώρο πιθανότητας $(\Omega,\ \mathcal{A},\ \mathbb{P})$ ορίζουμε τυχαίες μεταβλητές $X_1,\dots,X_k:\Omega\to\mathbb{R}$ και ζητάμε την ύπαρξη $\omega\in\Omega$ τέτοιο ώστε $X_1(\omega)\leq a_1,\dots,X_k(\omega)\leq a_k$, όπου $a_i\in\mathbb{R}$, για κάθε $i=1,\dots,k$. Αν πετύχουμε να δείξουμε ότι

$$\mathbb{E}(X_1)\leq a_1,\dots,\mathbb{E}(X_k)\leq a_k$$

εξασφαλίζουμε ότι υπάρχουν $\omega_1,\dots,\omega_k\in\Omega$ ώστε

$$X_1(\omega_1)\leq a_1,\dots,X_k(\omega_k)\leq a_k\ .$$

Το ερώτημα που τίθεται είναι αν μπορούμε να βρούμε $\omega=\omega_1=\cdots=\omega_\kappa\in\Omega$.

Η ιδέα για να πετύχουμε την ύπαρξη τέτοιου ω είναι να εκτιμήσουμε την απόκλιση των τυχαίων μεταβλητών X_i από τη μέση τιμή τους $\mathbb{E}(X_i)$, δηλαδή να βρούμε καλά άνω φράγματα για πιθανότητες της μορφής

$$\mathbb{P}(\{\omega\in\Omega:\ |X_i-\mathbb{E}(X_i)|>t_i\}),\quad t_i>0.$$

Αν το καταφέρουμε αυτό, ξεκινώντας από τον υπολογισμό των $\mathbb{E}(X_i)$ για κατάλληλα $t_i > 0$ μπορούμε να δείξουμε ότι

$$\sum_{i=1}^{k} \mathbb{P}(|X_i - \mathbb{E}(X_i)| > t_i) < 1.$$

Συνεπώς από την υποαθροιστικότητα του μέτρου πιθανότητας θα έχουμε ότι

$$\mathbb{P}\Big(\bigcup_{i=1}^{k}(|X_i - \mathbb{E}(X_i)|\ ) > t_i\Big) \leq \sum_{i=1}^{k} \mathbb{P}(|X_i - \mathbb{E}(X_i)| > t_i) < 1.$$

Άρα,

$$\bigcap_{i=1}^{k}\{\omega \in \Omega :\ |X_i - \mathbb{E}(X_i)| \leq t_i\} \neq \emptyset \Longrightarrow$$

$$\mathbb{P}\Big(\bigcap_{i=1}^{k}(|X_i - \mathbb{E}(X_i)|\ ) \leq t_i\Big) > 0,$$

οπότε υπάρχει $\omega \in \Omega$ ώστε $|X_i - \mathbb{E}(X_i)| \leq t_i$, για κάθε $i = 1, \ldots, k$. Αυτό που καταφέραμε να δείξουμε είναι ότι υπάρχει $\omega \in \Omega$ που να ικανοποιεί ταυτόχρονα τις

$$X_1 \leq \mathbb{E}(X_1) + t_1, \ldots, X_k \leq \mathbb{E}(X_k) + t_k\ .$$

Σκοπός μας σε αυτό το κεφάλαιο είναι να κατασκευάσουμε τα κατάλληλα εργαλεία για να μπορούμε να εκτιμήσουμε την $\mathbb{P}(|X - \mathbb{E}(X)| \geq t)$, δηλαδή να βρίσκουμε καλά άνω φράγματα για αυτή τη πιθανότητα. Για το σκοπό αυτό θα ορίσουμε τη διασπορά τυχαίας μεταβλητής και θα δούμε δύο ανισότητες που θα μας βοηθήσουν στην εκτίμηση που θέλουμε να πετύχουμε.

Ορισμός 2.1. *Έστω $(\Omega,\ \mathcal{A},\ \mathbb{P})$ χώρος πιθανότητας και $X : \Omega \to \mathbb{R}$ τυχαία μεταβλητή. Η **διασπορά** της X ορίζεται ως*

$$Var(X) = \mathbb{E}[(X - \mathbb{E}(X))^2] = \mathbb{E}(X^2) - (\mathbb{E}(X))^2\ .$$

Λήμμα 2.1. *(Ανισότητα Markov) Έστω X μια μη αρνητική τυχαία μεταβλητή. Τότε για κάθε $t > 0$,*

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}(X)}{t}.$$

Απόδειξη: Έστω $t > 0$. Τότε

$$\begin{aligned} \mathbb{E}(X) &= \int_{\Omega} X(\omega) d\mathbb{P}(\omega) \\ &\geq \int_{\{\omega\in\Omega: X(\omega)\geq t\}} X(\omega) d\mathbb{P}(\omega) \\ &\geq \int_{\{\omega\in\Omega: X(\omega)\geq t\}} t d\mathbb{P}(\omega) \\ &= t\int_{\{\omega\in\Omega: X(\omega)\geq t\}} d\mathbb{P}(\omega) \\ &= t\mathbb{P}(\{\omega\in\Omega : X(\omega)\geq t\}). \end{aligned}$$

Άρα έπεται το ζητούμενο

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}(X)}{t}.$$

□

Πόρισμα 2.1. *Αν ϕ γνησίως αύξουσα συνάρτηση με θετικές τιμές, τότε για κάθε τυχαία μεταβλητή $X \geq 0$ και για κάθε $t \in \mathbb{R}$ έχουμε*

$$\mathbb{P}(X \geq t) = \mathbb{P}(\phi(X) \geq \phi(t)) \leq \frac{\mathbb{E}(\phi(X))}{\phi(t)}.$$

Θεώρημα 2.1. *(Ανισότητα Chebyshev) Έστω $(\Omega,\ \mathcal{A},\ \mathbb{P})$ χώρος πιθανότητας και $X : \Omega \to \mathbb{R}$ τυχαία μεταβλητή. Τότε για κάθε $t > 0$*

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq t) \leq \frac{Var(X)}{t^2}.$$

Απόδειξη: Έστω $t > 0$, θα εφαρμόσουμε την ανισότητα *Markov* για την τυχαία μεταβλητή $|X - \mathbb{E}(X)|$.

$$\begin{aligned} \mathbb{P}(|X - \mathbb{E}(X)| \geq t) &= \mathbb{P}(|X - \mathbb{E}(X)|^2 \geq t^2) \\ &\leq \frac{\mathbb{E}(|X - \mathbb{E}(X)|^2)}{t^2} \\ &= \frac{Var(X)}{t^2}. \end{aligned}$$

□

Γενικότερα, για $\phi(t) = t^q$, $t \geq 0$, $q > 0$ έχουμε

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq t) \leq \frac{\mathbb{E}(|X - \mathbb{E}(X)|^q)}{t^q}.$$

Στη συνέχεια αυτού του κεφαλαίου θα ορίσουμε τη συνάρτηση συγκέντρωσης σε ένα μετρικό χώρο πιθανότητας και θα δούμε, πώς σε συνδυασμό με τη μέθοδο των προσεγγιστικών ισοπεριμετρικών ανισοτήτων μπορούμε να πετύχουμε άνω φράγματα για την πιθανότητα απόκλισης από τη μέση τιμή.

Ορισμός 2.2. *Έστω (X, d) ένας μετρικός χώρος. Έστω $\mathcal{A}$ η Borel σ-άλγεβρα υποσυνόλων του X. Αν $\mu : \mathcal{A} \to [0,1]$ είναι ένα μέτρο πιθανότητας, τότε η τετράδα $(X, \mathcal{A}, d, \mu)$ λέγεται* ***μετρικός χώρος πιθανότητας****.*

Υπενθύμιση (μετρικού χώρου): Ένα σύνολο X εφοδιασμένο με μια μετρική d λέγεται **μετρικός χώρος**. Μια **μετρική** d στον X είναι μια συνάρτηση $d : X \times X \to \mathbb{R}$ ώστε για κάθε $x, y, z \in X$ να ισχύουν τα εξής:

(i) $d(x,y) = 0 \Leftrightarrow x = y$.

(ii) $d(x,y) = d(y,x)$. (*συμμετρία*)

(iii) $d(x,y) \leq d(x,z) + d(z,y)$. (*τριγωνική ανισότητα*)

Ορισμός 2.3. *Έστω $A \in \mathcal{A}$. Για κάθε $t > 0$ ορίζουμε την* ***t-επέκταση*** *του A να είναι το σύνολο*

$$A_t = \{x \in X : \ d(x, A) \leq t\}$$

όπου $d(x, A) = \inf\{d(x,y) : \ y \in A\}$.

Σε κάθε μετρικό χώρο πιθανότητας το **ισοπεριμετρικό πρόβλημα** διατυπώνεται ως εξής: Έστω $(X, \mathcal{A}, d, \mu)$ μετρικός χώρος πιθανότητας. Για δοσμένα $0 < a < 1$ και $t > 0$ θεωρούμε όλα τα $A \in \mathcal{A}$ με μέτρο $\mu(A) \geq a$ και την t-επέκτασής τους A_t. Να βρεθούν εκείνα τα A για τα οποία ελαχιστοποιείται το $\mu(A_t)$, δηλαδή να βρεθεί το

$$\inf\{\mu(A_t) : \ A \in \mathcal{A}, \ \mu(A) \geq a\}.$$

Αν μπορεί να λυθεί το ισοπεριμετρικό πρόβλημα για κάποιο ζευγάρι (a, t), αν υποθέσουμε ότι $B \in \mathcal{A}$ είναι μια λύση του, τότε για κάθε $A \in \mathcal{A}$ με $\mu(A) \geq a$ έχουμε

$$\mu(A_t) \geq \mu(B_t). \tag{2.1}$$

Η ανισότητα αυτή είναι η **ισοπεριμετρική ανισότητα** για τα δοσμένα a και t.

Σε πολλά ισοπεριμετρικά προβλήματα η λύση τους είναι δύσκολη ή και πολλές φορές αδύνατη. Σε αυτήν την περίπτωση για τον σκοπό μας είναι αρκετή μια ασθενέστερη μορφή της (2.1), δηλαδή μας αρκεί να βρούμε ένα καλό κάτω φράγμα για το $\inf\{\mu(A_t) : \ A \in \mathcal{A}, \ \mu(A) \geq a\}$. Οι ανισότητες που επιτυγχάνουν ένα τέτοιο κάτω φράγμα λέγονται **προσεγγιστικές ισοπεριμετρικές ανισότητες**.

Η σχέση των προσεγγιστικών ισοπεριμετρικών ανισοτήτων με το πρόβλημα της εκτίμησης της πιθανότητας απόκλισης από το μέσο θα φανεί με τον επόμενο ορισμό και το επόμενο θεώρημα.

Ορισμός 2.4. *Έστω $(X, \mathcal{A}, d, \mu)$ μετρικός χώρος πιθανότητας. Η **συνάρτηση συγκέντρωσης** του χώρου X είναι η συνάρτηση*

$$\alpha(X,t) := 1 - \inf\{\mu(A_t) : \ \mu(A) \geq 1/2\}.$$

Ορισμός 2.5. *Για μιά συνάρτηση $f : X \to \mathbb{R}$ ορίζουμε **μέσο Lévy** της f να είναι ένας αριθμός $\mathbf{M_f}$ για τον οποίο ισχύει*

$$\mu(\{x \in X : \ f(x) \geq M_f\}) \geq 1/2$$

και

$$\mu(\{x \in X : \ f(x) \leq M_f\}) \geq 1/2.$$

Παρατήρηση: Αν η f είναι συνεχής τότε ο μέσος Lévy είναι μοναδικός.

Θεώρημα 2.2. *Έστω $(X, \mathcal{A}, d, \mu)$ μετρικός χώρος πιθανότητας. Αν $f : X \to \mathbb{R}$ είναι μια συνάρτηση Lipschitz με σταθερά 1, δηλαδή αν $|f(x) - f(y)| \leq d(x,y)$ για κάθε $x, y \in X$, τότε*

$$\mu(\{x \in X : \ |f(x) - M_f| > t\}) \leq 2\alpha(X,t).$$

Απόδειξη: Θεωρούμε τα σύνολα

$$A = \{x \in X : \ f(x) \geq M_f\}$$

$$B = \{x \in X : \ f(x) \leq M_f\}$$

και τις t-επεκτάσεις τους. Έστω $y \in A_t$ τότε υπάρχει $x \in A$ ώστε $d(x,y) \leq t$. Αφού η f είναι *Lipschitz* με σταθερά 1 έχουμε

$$-d(x,y) \leq f(y) - f(x) \leq d(x,y).$$

Οπότε,

$$f(y) = f(y) - f(x) + f(x) \geq -d(x,y) + M_f \geq M_f - t.$$

Όμοια, αν $y \in B_t$ τότε υπάρχει $x \in B$ τέτοιο ώστε $d(x,y) \leq t$, οπότε

$$f(y) = f(y) - f(x) + f(x) \leq d(x,y) + M_f \leq M_f + t.$$

Δηλαδή, αν $y \in A_t \cap B_t \Rightarrow |f(x) - M_f| \leq t$. Άρα,

$$A_t \cap B_t \subseteq \{x \in X : \ |f(x) - M_f| \leq t\} \Rightarrow$$

$$\{x \in X : |f(x) - M_f| \le t\}^c \subseteq (A_t \cap B_t)^c \Rightarrow$$
$$\{x \in X : |f(x) - M_f| > t\} \subseteq A_t^c \cup B_t^c \Rightarrow$$
$$\mu(\{x \in X : |f(x) - M_f| > t\}) \le \mu(A_t^c \cup B_t^c) \Rightarrow$$
$$\mu(\{x \in X : |f(x) - M_f| > t\}) \le \mu(A_t^c) + \mu(B_t^c) \Rightarrow$$

$$\mu(\{x \in X : |f(x) - M_f| > t\}) \le 1 - \mu(A_t) + 1 - \mu(B_t). \tag{2.2}$$

Όμως από τον ορισμό της συνάρτησης συγκέντρωσης έχουμε

$$1 - \mu(A_t) \le \alpha(X,t) \quad \text{και} \quad 1 - \mu(B_t) \le \alpha(X,t)$$

Αντικαθιστώντας στη (2.2) παίρνουμε το ζητούμενο

$$\mu(\{x \in X : |f(x) - M_f| > t\}) \le \alpha(X,t) + \alpha(X,t) = 2\alpha(X,t).$$

□

Αξίζει να σημειώσουμε ότι ισχύει και το αντίστροφο του θεωρήματος.

Πρόταση 2.1. *Έστω $(X, \mathcal{A}, d, \mu)$ μετρικός χώρος πιθανότητας. Αν για κάποιο $t > 0$ και για κάθε συνάρτηση $f : X \to \mathbb{R}$ με σταθερά Lipschitz 1 έχουμε $\mu(\{x \in X : |f(x) - M_f| > t\}) \le \eta$, τότε $\alpha(X,t) \le \eta$.*

Απόδειξη: Έστω $A \in \mathcal{A}$ με $\mu(A) \ge 1/2$. Θεωρούμε τη συνάρτηση

$$f(x) = d(x,A) = \inf\{d(x,y) : y \in A\}.$$

Η f είναι *Lipschitz* με σταθερά 1. Πράγματι, έστω $x_A, y_A \in A$ και $\varepsilon > 0$, τότε

$$d(x,A) \le d(x,x_A) \le d(x,A) + \varepsilon$$

και

$$d(y,A) \le d(y,y_A) \le d(y,A) + \varepsilon.$$

Έχουμε

$$\begin{aligned} f(x) - f(y) &= d(x,A) - d(y,A) \\ &\le d(x,y_A) - d(y,y_A) + \varepsilon \\ &\le d(x,y) + \varepsilon. \end{aligned}$$

Ομοίως δείχνουμε ότι $f(y) - f(x) \le d(x,y) + \varepsilon$. Άρα $|f(x) - f(y)| \le d(x,y)$. Επίσης ο μέσος *Lévy* της f είναι $M_f = 0$. Πράγματι,

$$\mu(\{x \in X : f(x) \ge 0\}) = \mu(\{x \in X : d(x,A) \ge 0\}) = \mu(X) = 1 \ge 1/2.$$

Θέλουμε να δείξουμε και ότι $\mu(\{x \in X : \ f(x) \leq 0\}) \geq 1/2$. Έχουμε,

$$\mu(\{x \in X : \ f(x) \leq 0\}) = \mu(\{x \in X : \ d(x, A) = 0\}).$$

Όμως αν $x \in A$ τότε $d(x, A) = 0$. Άρα $\{x \in X : \ d(x, A) = 0\} \supseteq A$. Συνεπώς

$$\mu(\{x \in X : \ d(x, A) = 0\}) \geq \mu(A) \geq 1/2.$$

Από την υπόθεση παίρνουμε

$$\mu(\{x \in X : \ |f(x) - 0| > t\}) \leq \eta \Rightarrow$$

$$\mu(\{x \in X : \ d(x, A) > t\}) \leq \eta \Rightarrow$$
$$\mu(A_t^c) \leq \eta \Rightarrow 1 - \mu(A_t) \leq \eta.$$

Έστω $\varepsilon > 0$ τότε,

$$\begin{aligned} \alpha(X, t) - \varepsilon &= 1 - (\inf\{\mu(A_t) : \ \mu(A) \geq 1/2\} + \varepsilon) \\ &\leq 1 - \mu(A_t) \\ &\leq \eta. \end{aligned}$$

Συνεπώς $\alpha(X, t) \leq \eta$.

$\square$

2.2 Η ανισότητα των Prékopa-Leindler

Θεώρημα 2.3. *(Ανισότητα Prékopa-Leindler): Έστω* $h, f, g : \mathbb{R}^n \to \mathbb{R}^+$ *τρεις μετρήσιμες και ολοκληρώσιμες συναρτήσεις και* $a, b \in \mathbb{R}$ *ώστε* $a, b \geq 0$ *και* $a + b = 1$. *Αν*

$$h(ax + by) \geq f(x)^a \, g(y)^b.$$

Τότε,

$$\int_{\mathbb{R}^n} h \geq \left(\int_{\mathbb{R}^n} f \right)^a \left(\int_{\mathbb{R}^n} g \right)^b.$$

Η απόδειξη του θεωρήματος θα γίνει επαγωγικά, θα το αποδείξουμε αρχικά για $n = 1$ και αυτό θα γίνει με τη βοήθεια τριών λημμάτων τα οποία διατυπώνονται και αποδεικνύονται παρακάτω.

Λήμμα 2.2. *Έστω* A, B *μη κενά μετρήσιμα υποσύνολα του* $\mathbb{R}$. *Τότε,*

$$m(A + B) \geq mA + mB,$$

όπου m *το μέτρο Lebesgue και* $A + B = \{a + b : a \in A, \ b \in B\}$.

Απόδειξη: Θα διακρίνουμε τις εξής δύο περιπτώσεις. Τα A, B να είναι φραγμένα και τουλάχιστον ένα από τα A, B όχι φραγμένο.
Περίπτωση I: Έστω A, B μη κενά μετρήσιμα και φραγμένα υποσύνολα του $\mathbb{R}$.
Αφού A, B φραγμένα και μη κενά, υπάρχει το $\sup A$ και το $\inf B$.
Έστω $\varepsilon > 0$, τότε υπάρχει $x_A \in A : \sup A - \varepsilon < x_A$. Έχουμε $x_A \in A \Rightarrow 0 \in A - x_A$ (με $A - x_A$ εννοούμε το σύνολο $A - \{x_A\}$).
Έστω $x \in A$ τότε,

$$x - x_A \leq x - (\sup A - \varepsilon) = (x - \sup A) + \varepsilon \leq \varepsilon, \quad \text{αφού } x - \sup A \leq 0.$$

Αυτό που δείξαμε είναι ότι αν $x \in A \Rightarrow x - x_A \leq \varepsilon$, άρα $A - x_A \subseteq (-\infty, \varepsilon]$.
Εργαζόμαστε εντελώς ανάλογα και στο σύνολο B.
Έστω $\varepsilon > 0$, τότε $\exists \, y_B \in B : y_B < \inf B + \varepsilon$. Έχουμε $y_B \in B \Rightarrow 0 \in B - y_B$.
Έστω $y \in B$ τότε,

$$y - y_B \geq y - (\inf B + \varepsilon) = (y - \inf B) - \varepsilon \geq -\varepsilon, \quad \text{αφού } y - \inf B \geq 0.$$

Έτσι δείξαμε ότι αν $y \in B \Rightarrow y - y_B \geq -\varepsilon$. Άρα $B - y_B \subseteq [-\varepsilon, \infty)$. Τώρα το μέτρο *Lebesgue* είναι αναλλοίωτο στις μεταθέσεις, δηλαδή για κάθε $\lambda \in \mathbb{R}$, $mA = m(A + \lambda)$. Έτσι έχουμε ότι

$$m(A + B) = m((A + B) - (x_A + y_B)) = m((A - x_A) + (B - y_B)).$$

Αφού $0 \in A - x_A$ έπεται ότι

$$(A - x_A) + (B - y_B) \supseteq \{0\} + (B - y_B) = B - y_B.$$

Ομοίως επειδή $0 \in B - y_B$ έπεται ότι

$$(A - x_A) + (B - y_B) \supseteq \{0\} + (A - x_A) = A - x_A.$$

Άρα,

$$(A-x_A)+(B-y_B) \supseteq \Big((A-x_A)+\{0\}\Big)\cup\Big((B-y_B)+\{0\}\Big) = (A-x_A)\cup(B-y_B) \Rightarrow$$

$$m((A - x_A) + (B - y_B)) \geq m((A - x_A) \cup (B - y_B))$$

Άρα

$$\begin{aligned} m(A+B) &= m((A - x_A) + (B - y_B)) \\ &\geq m((A - x_A) \cup (B - y_B)) \\ &= m(A - x_A) + m(B - y_B) - m((A - x_A) \cap (B - y_B)) \\ &\geq m(A) + m(B) - 2\varepsilon, \end{aligned}$$

αφού $(A-x_A)\cap(B-y_B) \subseteq [-\varepsilon, \varepsilon]$ οπότε $m((A-x_A)\cap(B-y_B)) \leq m([-\varepsilon, \varepsilon]) = 2\varepsilon$. Αφήνοντας το ε να πάει στο 0 παίρνουμε το ζητούμενο

$$m(A+B) \geq mA + mB.$$

Περίπτωση ΙΙ: Έστω A ή B μη φραγμένα. Εδώ θα διακρίνουμε δύο υποπεριπτώσεις.
Αν $mA = \infty$ ή $mB = \infty$ τότε προφανώς $m(A+B) \geq mA + mB$.
Έστω τώρα ότι $mA < \infty$ και $mB < \infty$. Ορίζουμε

$$I_n = (-(n+1), n] \cup [n, n+1), \quad n \in \mathbb{N}$$

Παρατηρούμε ότι τα I_n είναι ξένα μεταξύ τους για κάθε $n \in \mathbb{N}$ και ότι

$$\bigcup_{n=1}^{\infty} I_n = \mathbb{R}$$

Αφού I_n ξένα, τότε και $I_n \cap A$ ξένα και $I_n \cap B$ ξένα. Έχουμε,

$$A = \bigcup_{n=1}^{\infty} (I_n \cap A).$$

Πράγματι,

$$\bigcup_{n=1}^{\infty}(I_n \cap A) = \Big(\bigcup_{n=1}^{\infty} I_n\Big) \cap A = \mathbb{R} \cap A = A.$$

Ομοίως δείχνουμε και ότι

$$B = \bigcup_{n=1}^{\infty}(I_n \cap B).$$

Από την αθροιστικότητα του μέτρου παίρνουμε

$$\infty > mA = m\Big(\bigcup_{n=1}^{\infty}(I_n \cap A)\Big) = \sum_{n=1}^{\infty} m(I_n \cap A)$$

και

$$\infty > mB = m\Big(\bigcup_{n=1}^{\infty}(I_n \cap B)\Big) = \sum_{n=1}^{\infty} m(I_n \cap B).$$

Άρα οι σειρές συγκλίνουν, οπότε έχουμε ότι

$$\forall\, \varepsilon > 0,\ \exists\, N_1 \in \mathbb{N} :\ \forall\, n \geq N_1$$

$$\Big|\sum_{i=1}^{\infty} m(I_i \cap A) - \sum_{i=1}^{n} m(I_i \cap A)\Big| < \varepsilon \Rightarrow \sum_{i=N_1+1}^{\infty} m(I_i \cap A) < \varepsilon.$$

Ομοίως,

$$\forall\, \varepsilon > 0,\ \exists\, N_2 \in \mathbb{N} :\ \forall\, n \geq N_2 \quad \sum_{i=N_2+1}^{\infty} m(I_i \cap B) < \varepsilon.$$

Για $N = \max\{N_1,\ N_2\}$ παίρνουμε

$$\sum_{i=N+1}^{\infty} m(I_i \cap A) < \varepsilon \quad \text{και} \quad \sum_{i=N+1}^{\infty} m(I_i \cap B) < \varepsilon.$$

Θέτουμε $A_N = A \cap [-N,\ N] \subseteq A$ και $B_N = B \cap [-N,\ N] \subseteq B$. Παρατηρούμε ότι ισχύουν οι εξής σχέσεις:

$$m(A \backslash A_N) = \sum_{i=N+1}^{\infty} m(I_i \cap A) < \varepsilon$$

και

$$m(B \backslash B_N) = \sum_{i=N+1}^{\infty} m(I_i \cap B) < \varepsilon$$

Τα A_N και B_N είναι φραγμένα, άρα από περίπτωση I έχουμε

$$\begin{aligned} m(A_N+B_N) &\geq mA_N+mB_N \\ &= mA-m(A\backslash A_N)+mB-m(B\backslash B_N) \\ &\geq mA+mB-\varepsilon-\varepsilon \\ &= mA+mB-2\varepsilon \end{aligned}$$

Τώρα $A \supseteq A_N$ και $B \supseteq B_N \Rightarrow A+B \supseteq A_N+B_N$. Άρα

$$m(A+B) \geq m(A_N+B_N) \geq mA+mB-2\varepsilon.$$

Τελικά καθώς το ε πηγαίνει στο 0 παίρνουμε το ζητούμενο. Άρα σε κάθε περίπτωση δείξαμε ότι

$$m(A+B) \geq mA+mB.$$

□

Λήμμα 2.3. *Έστω $f:\mathbb{R}\to\mathbb{R}_+$ μετρήσιμη, ολοκληρώσιμη και μη αρνητική συνάρτηση. Τότε,*

$$\int_{\mathbb{R}} f(x)dx = \int_0^\infty m\{x\in\mathbb{R}:\ f(x)\geq y\}dy.$$

Απόδειξη: Το λήμμα αυτό αποδεικνύεται με μια απλή εφαρμογή του θεωρήματος *Fubini*. Ορίζω $A=\{(x,y):\ f(x)\geq y\}$. Τα σημεία του A αποτελούν τα ευθύγραμμα τμήματα μεταξύ των σημείων $(x,0)$ και $(x,f(x))$, για κάθε $x\in\mathbb{R}$. Έστω $\mathcal{X}_A$ η χαρακτηριστική συνάρτηση του A (δηλαδή $\mathcal{X}_A(x)=1$ αν $x\in A$ και $\mathcal{X}_A(x)=0$ διαφορετικά). Από το θεώρημα *Fubibi* έχουμε,

$$\begin{aligned} \infty > \int_{\mathbb{R}} f(x)dx &= \int_{\mathbb{R}^2} \mathcal{X}_A(x,y)dxdy \\ &= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} \mathcal{X}_A(x,y)dx\right)dy \\ &= \int_0^{\infty}\left(\int_{-\infty}^{\infty} \mathcal{X}_A(x,y)dx\right)dy. \end{aligned}$$

Κρατώντας το y σταθερό παίρνουμε:

$$\mathcal{X}_A(x,y)=1 \Leftrightarrow (x,y)\in A \Leftrightarrow f(x)\geq y.$$

Άρα

$$\int_{-\infty}^{\infty} \mathcal{X}_A(x,y)dx = m\{x\in\mathbb{R}:\ f(x)\geq y\}.$$

Αντικαθιστώντας στην προηγούμενη σχέση παίρνουμε το ζητούμενο

$$\int_{\mathbb{R}} f(x)dx = \int_0^{\infty} m\{x \in \mathbb{R} : \ f(x) \geq y\}dy.$$

□

Λήμμα 2.4. *(Ανισότητα αριθμητικού γεωμετρικού μέσου): Έστω $\lambda \in [0,1]$ και $x, y \in \mathbb{R}$ με $x, y \geq 0$. Τότε*

$$\lambda x + (1-\lambda)y \geq x^{\lambda}y^{1-\lambda}.$$

Απόδειξη: Αν $x = 0$ ή $y = 0$ ή $\lambda = 0$ προφανώς ισχύει. Έστω τώρα $x \neq 0$ και $y \neq 0$ και $\lambda \neq 0$. Θέλουμε να δείξουμε ότι:

$$\lambda x + (1-\lambda)y \geq x^{\lambda}y^{1-\lambda} \Leftrightarrow$$

$$\lambda \frac{x}{x^{\lambda}y^{1-\lambda}} + (1-\lambda)\frac{y}{x^{\lambda}y^{1-\lambda}} \geq 1 \Leftrightarrow$$

$$\lambda \Big(\frac{x}{y}\Big)^{1-\lambda} + (1-\lambda)\Big(\frac{y}{x}\Big)^{\lambda} \geq 1.$$

Θέτουμε

$$z = \frac{x}{y} > 0.$$

Άρα αρκεί να δείξουμε ότι:

$$\lambda z^{1-\lambda} + (1-\lambda)\frac{1}{z^{\lambda}} \geq 1, \quad \forall\ z > 0,\ \forall\ \lambda \in [0,1]$$

$$\Leftrightarrow \lambda z + (1-\lambda)z \geq z^{\lambda} \Leftrightarrow \lambda z + (1-\lambda) - z^{\lambda} \geq 0.$$

Ορίζουμε

$$g(z) = \lambda z + (1-\lambda) - z^{\lambda}, \quad z > 0.$$

Η g είναι παραγωγίσιμη, άρα

$$g'(z) = 0 \Leftrightarrow \lambda - \lambda\frac{1}{z^{1-\lambda}} = 0 \Leftrightarrow \lambda\Big(1 - \frac{1}{z^{1-\lambda}}\Big) = 0 \Leftrightarrow z = 1,$$

αφού υποθέσαμε ότι $\lambda \neq 0$. Τώρα παρατηρούμε ότι

$$\forall\ z > 1,\ g'(z) > 0 \quad \text{και} \quad \forall\ z < 1,\ g'(z) < 0.$$

Άρα στο $z = 1$ η g έχει ελάχιστο, επομένως

$$g(z) \geq g(1) = \lambda 1 + (1-\lambda) - 1^{\lambda} = 0 \Rightarrow$$

$$\lambda z + (1-\lambda) - z^{\lambda} \geq 0.$$

□

Είμαστε έτοιμοι τώρα να αποδείξουμε την ανισότητα *Prékopa-Leindler.*

Απόδειξη: Αρχικά θα το δείξουμε για $n = 1$. Έστω $t > 0$, ισχυριζόμαστε ότι

$$\{z \in \mathbb{R} : \ h(z) \geq t\} \supseteq a\{x \in \mathbb{R} : \ f(x) \geq t\} + b\{y \in \mathbb{R} : \ g(y) \geq t\}.$$

Πράγματι, έστω $x \in \{x \in \mathbb{R} : \ f(x) \geq t\}$ και $y \in \{y \in \mathbb{R} : \ g(y) \geq t\}$. Από υπόθεση έχουμε

$$h(ax + by) \geq f(x)^a g(y)^b \geq t^a t^b = t^{a+b} = t^1 = t.$$

Άρα $ax + by \in \{z \in \mathbb{R} : \ h(z) \geq t\}$. Έτσι αποδείχτηκε ο ισχυρισμός. Από το λήμμα 2.2 παίρνουμε:

$$\begin{aligned} m\{z \in \mathbb{R} : \ h(z) \geq t\} &\geq m(a\{x \in \mathbb{R} : \ f(x) \geq t\} + b\{y \in \mathbb{R} : \ g(y) \geq t\}) \\ &\geq am(\{x \in \mathbb{R} : \ f(x) \geq t\}) + bm(\{y \in \mathbb{R} : \ g(y) \geq t\}). \end{aligned}$$

Από την γραμμικότητα του ολοκληρώματος συνεπάγεται ότι

$$\int_0^\infty m\{z \in \mathbb{R} : \ h(z) \geq t\} dt \geq$$

$$a\int_0^\infty m\{x \in \mathbb{R} : \ f(x) \geq t\} dt + b\int_0^\infty m\{y \in \mathbb{R} : \ g(y) \geq t\} dt.$$

Το λήμμα 2.3 και η ανισότητα *αριθμητικού γεωμετρικού μέσου* μας δίνουν το ζητούμενο

$$\int_{\mathbb{R}} h \geq a\int_{\mathbb{R}} f + b\int_{\mathbb{R}} g \geq \left(\int_{\mathbb{R}} f\right)^a \left(\int_{\mathbb{R}} g\right)^b.$$

Επαγωγικό βήμα: Έστω h, f, g όπως στο θεώρημα. Έστω ότι ισχύει για $k = n - 1$, δηλαδή,

$$\int_{\mathbb{R}^{n-1}} h \geq \left(\int_{\mathbb{R}^{n-1}} f\right)^a \left(\int_{\mathbb{R}^{n-1}} g\right)^b.$$

Θέλουμε να το αποδείξουμε για $k = n$. Για κάθε $s \in \mathbb{R}$ ορίζουμε τις συναρτήσεις $h_s, f_s, g_s : \mathbb{R}^{n-1} \to \mathbb{R}_+$ με

$$h_s(w) = h(w, s), \quad f_s(w) = f(w, s), \quad g_s(w) = g(w, s).$$

Για κάθε $x, y \in \mathbb{R}^{n-1}$ και $s_0, s_1 \in \mathbb{R}$ από την υπόθεση έπεται ότι

$$\begin{aligned} h_{as_1+bs_0}(ax+by) &= h(ax+by,\ as_1+bs_0) \\ &= h(a(x,s_1)+b(y,s_0)) \\ &\geq \Big(f(x,s_1)\Big)^a \Big(g(y,s_0)\Big)^b \\ &= \Big(f_{s_1}(x)\Big)^a \Big(g_{s_0}(y)\Big)^b. \end{aligned}$$

Δηλαδή δείξαμε ότι,

$$h_{as_1+bs_0}(ax+by) \geq \Big(f_{s_1}(x)\Big)^a \Big(g_{s_0}(y)\Big)^b.$$

Η επαγωγική υπόθεση μάς δίνει,

$$\begin{aligned} H(as_1+bs_0) &:= \int_{\mathbb{R}^{n-1}} h_{as_1+bs_0} \\ &\geq \left(\int_{\mathbb{R}^{n-1}} f_{s_1}\right)^a \left(\int_{\mathbb{R}^{n-1}} g_{s_0}\right)^b \\ &= F(s_1)^a\, G(s_0)^b. \end{aligned}$$

Εφαρμόζουμε τώρα ξανά την επαγωγική υπόθεση για $n=1$ και για τις συναρτήσεις H, F, G. Έχουμε ότι,

$$\begin{aligned} \int_{\mathbb{R}^n} h &= \int_{\mathbb{R}} \int_{\mathbb{R}^{n-1}} h_{as_0+bs_1} \\ &= \int_{\mathbb{R}} H \\ &\geq \left(\int_{\mathbb{R}} F\right)^a \left(\int_{\mathbb{R}} G\right)^b \\ &= \left(\int_{\mathbb{R}} \int_{\mathbb{R}^{n-1}} f_{s_1}\right)^a \left(\int_{\mathbb{R}} \int_{\mathbb{R}^{n-1}} g_{s_0}\right)^b \\ &= \left(\int_{\mathbb{R}^n} f\right)^a \left(\int_{\mathbb{R}^n} g\right)^b. \end{aligned}$$

$\square$

2.3 Η ανισότητα των Brunn-Minkowski

Ορισμός 2.6. *Έστω A, B μη κενά υποσύνολα του $\mathbb{R}^n$. Ορίζουμε το* ***άθροισμα Minkowski*** *των A και B να είναι το σύνολο*

$$A + B = \{a + b : \ a \in A, \ b \in B\}.$$

Επιπλέον για κάθε $t \geq 0$ ορίζουμε

$$tA = \{ta : \ a \in A\}.$$

Θεώρημα 2.4. *(Ανισότητα Brunn-Minkowski) Έστω K, T δύο μη κενά υποσύνολα του $\mathbb{R}^n$. Τότε,*

$$|K + T|^{1/n} \geq |K|^{1/n} + |T|^{1/n}. \tag{2.3}$$

Παρατήρηση: Η ανισότητα (2.3) εκφράζει το γεγονός ότι ο όγκος είναι **κοίλη** συνάρτηση ως προς το άθροισμα *Minkowski*. Για το λόγο αυτό γράφεται συχνά στη μορφή

$$|\lambda K + (1 - \lambda)T|^{1/n} \geq \lambda |K|^{1/n} + (1 - \lambda)|T|^{1/n}, \tag{2.4}$$

όπου $\lambda \in (0, 1)$. Θα δούμε πρώτα μια άλλη μορφή του θεωρήματος 2.4 η οποία ονομάζεται *ασθενής ανισότητα Brunn-Minkowski.*

Θεώρημα 2.5. *(Ασθενής ανισότητα Brunn-Minkowski) Έστω Έστω K, T δύο μη κενά υποσύνολα του $\mathbb{R}^n$ και $\lambda \in (0, 1)$. Τότε,*

$$|\lambda K + (1 - \lambda)T| \geq |K|^{\lambda} \, |T|^{1-\lambda}. \tag{2.5}$$

Η ανισότητα αυτή παρόλο που είναι ασθενέστερη της (2.3), έχει το πλεονέκτημα ότι είναι ανεξάρτητη της διάστασης n.

Απόδειξη: Η απόδειξη θα γίνει με μια απλή εφαρμογή του θεωρήματος 2.3. Έστω K, T δύο μη κενά υποσύνολα του $\mathbb{R}^n$ και $\lambda \in (0, 1)$. Ορίζουμε τις εξής χαρακτηριστικές συναρτήσεις

$$f = \mathcal{X}_K, \quad g = \mathcal{X}_T, \quad h = \mathcal{X}_{\lambda K + (1-\lambda)T}$$

Θα δείξουμε ότι οι συναρτήσεις αυτές ικανοποιούν τις προϋποθέσεις του θεωρήματος 2.3. Πράγματι, προφανώς οι $h, f, g : \mathbb{R}^n \to \mathbb{R}_+$ είναι μετρήσιμες και ολοκληρώσιμες. Αρκεί να δείξουμε ότι

$$h(\lambda x + (1 - \lambda)y) \geq f(x)^{\lambda} \, g(y)^{1-\lambda}.$$

Αν $x \notin K$ ή $y \notin T$ τότε $f(x) = 0$ ή $g(y) = 0$. Άρα

$$h(\lambda x + (1-\lambda)y) \geq 0 = f(x)^{\lambda}\, g(y)^{1-\lambda}.$$

Αν $x \in K$ και $y \in T$ τότε $\lambda x + (1-\lambda)y \in \lambda K + (1-\lambda)T$, άρα $f(x) = g(y) = h(\lambda x + (1-\lambda)y) = 1$. Επομένως,

$$h(\lambda x + (1-\lambda)y) = 1 = 1^{\lambda}\, 1^{1-\lambda} = f(x)^{\lambda}\, g(y)^{1-\lambda}.$$

Τώρα εφαρμόζουμε το θεώρημα 2.3 για τις συναρτήσεις h, f, g και παίρνουμε,

$$\begin{aligned} |\lambda K + (1-\lambda)T| &= \int_{\mathbb{R}^n} \mathcal{X}_{\lambda K + (1-\lambda)T} \\ &= \int_{\mathbb{R}^n} h \\ &\geq \left(\int_{\mathbb{R}^n} f\right)^{\lambda} \left(\int_{\mathbb{R}^n} g\right)^{1-\lambda} \\ &= \left(\int_{\mathbb{R}^n} \mathcal{X}_K\right)^{\lambda} \left(\int_{\mathbb{R}^n} \mathcal{X}_T\right)^{1-\lambda} \\ &= |K|^{\lambda}\, |T|^{1-\lambda}, \end{aligned}$$

δηλαδή το αποτέλεσμα.

□

Απόδειξη:(της Ανισότητας Brunn-Minkowski) Έστω K, T μη κενά υποσύνολα του $\mathbb{R}^n$ και $\lambda \in (0,1)$. Αν $|K| = |T| = 0$ τότε η (2.3) είναι προφανής. Έστω ότι $|K| \neq 0$ και $|T| \neq 0$. Θα εφαρμόσουμε το θεώρημα 2.5 για τα εξής σύνολα:

$$K_1 = \frac{K}{|K|^{1/n}}, \quad T_1 = \frac{T}{|T|^{1/n}} \quad \text{και για} \quad \lambda = \frac{|K|^{1/n}}{|K|^{1/n} + |T|^{1/n}}.$$

Παρατηρούμε ότι $|K_1| = |T_1| = 1$. Πράγματι,

$$|K_1| = \left|\frac{K}{|K|^{1/n}}\right| = \left(\frac{1}{|K|^{1/n}}\right)^n |K| = \frac{|K|}{|K|} = 1.$$

Ομοίως δείχνουμε και ότι $|T_1| = 1$. Συνεπώς από την (2.5) έχουμε ότι

$$|\lambda K_1 + (1-\lambda)T_1| \geq 1 \tag{2.6}$$

Όμως,

$$\lambda K_1 + (1-\lambda)T_1 =$$

$$\frac{|K|^{1/n}}{|K|^{1/n}+|T|^{1/n}}\,\frac{K}{|K|^{1/n}} + \left(1-\frac{|K|^{1/n}}{|K|^{1/n}+|T|^{1/n}}\right)\frac{T}{|T|^{1/n}} =$$

$$\frac{|K|^{1/n}K}{|K|^{1/n}(|K|^{1/n}+|T|^{1/n})} + \frac{(|K|^{1/n}+|T|^{1/n}-|K|^{1/n})T}{|T|^{1/n}(|K|^{1/n}+|T|^{1/n})} =$$

$$\frac{|K|^{1/n}|T|^{1/n}K + |K|^{1/n}|T|^{1/n}T}{|K|^{1/n}|T|^{1/n}(|K|^{1/n}+|T|^{1/n})} = \frac{K+T}{|K|^{1/n}+|T|^{1/n}}.$$

Αντικαθιστώντας στην (2.6) παίρνουμε

$$\left|\frac{K+T}{|K|^{1/n}+|T|^{1/n}}\right| \geq 1 \Rightarrow$$

$$\left(\frac{1}{|K|^{1/n}+|T|^{1/n}}\right)^n |K+T| \geq 1 \Rightarrow$$

$$|K+T| \geq \left(|K|^{1/n}+|T|^{1/n}\right)^n,$$

Από όπου συνεπάγεται το ζητούμενο:

$$|K+T|^{1/n} \geq |K|^{1/n}+|T|^{1/n}.$$

□

2.4 Ισοπεριμετρική ανισότητα στον Ευκλείδειο χώρο

Ορισμός 2.7. *Έστω A μη κενό υποσύνολο του $\mathbb{R}^n$, η* ***επιφάνεια κατά Minkowski*** $\boldsymbol{\partial(A)}$ *του A ορίζεται ως*

$$\partial(A) := \liminf_{t\to 0^+} \frac{|A_t| - |A|}{t}.$$

Θεώρημα 2.6. *Έστω A μη κενό συμπαγές υποσύνολο του $\mathbb{R}^n$. Τότε*

$$\left(\frac{\partial(A)}{\partial(B)}\right)^{\frac{1}{n-1}} \geq \left(\frac{|A|}{|B|}\right)^{\frac{1}{n}}. \tag{2.7}$$

όπου $B = B(0,1)$ η μοναδιαία Ευκλείδεια μπάλα κέντρου 0 και ακτίνας 1.

Το θεώρημα 2.6 εκφράζει την *ισοπεριμετρική ανισότητα στον Ευκλείδειο χώρο*, μας δείχνει με άλλα λόγια ότι, από όλα τα μη κενά συμπαγή υποσύνολα του Ευκλείδειου χώρου που έχουν δεδομένο σταθερό όγκο, η μπάλα έχει τη μικρότερη επιφάνεια. Αντίστροφα, αν κρατήσουμε την επιφάνεια σταθερή τότε η μπάλα έχει το μεγαλύτερο όγκο.

Απόδειξη: Παρατηρούμε ότι $A_t = A + tB$. Πράγματι, έστω $x \in A_t$, επειδή το A είναι συμπαγές έπεται ότι υπάρχει $y \in A$ ώστε

$$|x-y| = d(x,A) \leq t \Rightarrow x - y \in tB \Rightarrow x \in y + tB \subseteq A + tB \Rightarrow x \in A + tB.$$

Άρα

$$A_t \subseteq A + tB. \tag{2.8}$$

Αντίστροφα, έστω $x \in A + tB$, υπάρχουν $x_A \in A$ και $x_B \in B$ ώστε $x = x_A + tx_B \Rightarrow x - x_A = tx_B$. Έχουμε

$$d(x,A) \leq |x - x_A| = t|x_B| \leq t \Rightarrow x \in A_t.$$

Άρα

$$A + tB \subseteq A_t. \tag{2.9}$$

Από τις σχέσεις (2.8) και (2.9) παίρνουμε ότι $A_t = A + tB$. Χρησιμοποιώντας την ανισότητα *Brunn-Minkowski* έχουμε

$$\frac{|A_t| - |A|}{t} = \frac{|A + tB| - |A|}{t} \geq \frac{(|A|^{1/n} + |tB|^{1/n})^n - |A|}{t} \Longrightarrow$$

$$\Longrightarrow \liminf_{t\to 0^+} \frac{|A_t| - |A|}{t} \geq \liminf_{t\to 0^+} \frac{(|A|^{1/n} + t|B|^{1/n})^n - |A|}{t}.$$

Το όριο στο δεξί μέλος μας δίνει απροσδιόριστη μορφή $\frac{0}{0}$. Εφαρμόζοντας *De-l'Hospital* ως προς t παίρνουμε

$$\begin{aligned}
\partial(A) &\geq \liminf_{t\to 0^+} \frac{n(|A|^{1/n} + t|B|^{1/n})^{n-1}\,(0 + |B|^{1/n}) - 0}{1} \\
&\geq \liminf_{t\to 0^+} n(|A|^{1/n} + t|B|^{1/n})^{n-1}\,|B|^{1/n} \\
&\geq n(|A|^{1/n} + 0)^{n-1}\,|B|^{1/n}.
\end{aligned}$$

Τελικά παίρνουμε την εξής ανισότητα

$$\partial(A) \geq n|A|^{\frac{n-1}{n}}|B|^{1/n}\ . \tag{2.10}$$

Τώρα θα υπολογίσουμε την επιφάνεια της μοναδιαίας μπάλας $\partial(B)$.

$$\begin{aligned}
\partial(B) &= \liminf_{t\to 0^+} \frac{|B_t| - |B|}{t} \\
&= \liminf_{t\to 0^+} \frac{|B + tB| - |B|}{t} \\
&= \liminf_{t\to 0^+} \frac{|(t+1)B| - |B|}{t} \\
&= \liminf_{t\to 0^+} \frac{(t+1)^n|B| - |B|}{t} \\
&= \liminf_{t\to 0^+} \frac{n(t+1)^{n-1}|B| - 0}{1} \\
&= \liminf_{t\to 0^+} n(t+1)^{n-1}|B| \\
&= n|B|\ .
\end{aligned}$$

Συνεπώς,

$$\partial(B) = n|B| \Longrightarrow n = \frac{\partial(B)}{|B|}\ .$$

Αντικαθιστώντας στην (2.10) παίρνουμε το ζητούμενο

$$\begin{aligned}
\partial(A) &\geq \frac{\partial(B)}{|B|}\,|A|^{\frac{n-1}{n}}\,|B|^{1/n} \\
\frac{\partial(A)}{\partial(B)} &\geq \frac{|A|^{\frac{n-1}{n}}}{|B|^{\frac{n-1}{n}}} \\
\left(\frac{\partial(A)}{\partial(B)}\right)^{\frac{1}{n-1}} &\geq \left(\frac{|A|}{|B|}\right)^{\frac{1}{n}}\ .
\end{aligned}$$

□

Πόρισμα 2.2. *Έστω A μη κενό συμπαγές υποσύνολο του $\mathbb{R}^n$ και $r > 0$ ώστε $|A| = |rB|$. Τότε*

$$\partial(A) \geq \partial(rB) \ .$$

Απόδειξη: Με όμοιο τρόπο, όπως υπολογίσαμε την επιφάνεια της μπάλας, δείχνουμε ότι $\partial(rB) = nr^{n-1}|B|$. Πράγματι,

$$\begin{aligned}
\partial(rB) &= \liminf_{t\to 0^+} \frac{|(rB)_t| - |rB|}{t} \\
&= \liminf_{t\to 0^+} \frac{|rB + tB| - |rB|}{t} \\
&= \liminf_{t\to 0^+} \frac{(r+t)^n|B| - r^n|B|}{t} \\
&= \liminf_{t\to 0^+} \frac{n(r+t)^{n-1}|B| - 0}{1} \\
&= nr^{n-1}|B| \ .
\end{aligned}$$

Αντικαθιστώντας στην (2.10) και χρησιμοποιώντας την υπόθεση παίρνουμε το ζητούμενο:

$$\begin{aligned}
\partial(A) &\geq n|A|^{\frac{n-1}{n}} \ |B|^{1/n} = \\
&= n|rB|^{\frac{n-1}{n}} \ |B|^{1/n} \\
&= nr^{n-1}|B|^{\frac{n-1}{n}} \ |B|^{1/n} \\
&= nr^{n-1}|B| \\
&= \partial(rB) \ .
\end{aligned}$$

□

Πρόταση 2.2. *Αν $|A| = |B|$ τότε για κάθε $t \geq 0$, $|A_t| \geq |B_t|$, όπου B η μοναδιαία Ευκλείδεια μπάλα.*

Απόδειξη: Το αποτέλεσμα αυτό είναι άμεση συνέπεια της ανισότητας *Brunn-Minkowski.*

$$\begin{aligned}
|A_t|^{1/n} &= |A + tB|^{1/n} \geq |A|^{1/n} + |tB|^{1/n} \\
&= |B|^{1/n} + t|B|^{1/n} = (t+1)|B|^{1/n} \\
&= |(t+1)B|^{1/n} = |B + tB|^{1/n} = |B_t|^{1/n} \ .
\end{aligned}$$

Έτσι δείξαμε ότι $|A_t|^{1/n} \geq |B_t|^{1/n}$ από όπου έπεται το ζητούμενο:

$$|A_t| \geq |B_t| \ .$$

□

2.5 Ισοπεριμετρική ανισότητα στη σφαίρα

Θεωρούμε το μετρικό χώρο πιθανότητας $(\mathbb{S}^{n-1},\ \mathcal{A},\ \rho,\ \sigma)$ όπου,

- $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n :\ \|x\|_2 = 1\}$ η μοναδιαία σφαίρα στον $\mathbb{R}^n$.

- $\mathcal{A}$, η *Borel* σ-άλγεβρα υποσυνόλων της $\mathbb{S}^{n-1}$.

- Εφοδιάζουμε την $\mathbb{S}^{n-1}$ με τη γεωδαισιακή μετρική ρ. Η απόσταση δύο σημείων $x,\ y \in \mathbb{S}^{n-1}$, δηλαδή το $\rho(x,y)$, είναι η κυρτή γωνία $x\hat{O}y$ στο επίπεδο που ορίζεται από την αρχή των αξόνων O και τα $x,\ y$. Αν

$$\rho(x,y) = \theta \text{ τότε } \sin\frac{\theta}{2} = \frac{\|x-y\|_2}{2} \ .$$

Συνεπώς,

$$\rho(x,y) = 2\arcsin\frac{\|x-y\|_2}{2} \ .$$

Η γεωδαισιακή μετρική ρ είναι ισοδύναμη με την Ευκλείδεια μετρική και συνδέονται μέσω της σχέσης

$$\frac{2}{\pi}\rho(x,y) \leq \|x-y\|_2 \leq \rho(x,y) \ .$$

Πράγματι, γενικά για κάθε $x \in [0,\frac{\pi}{2}]$ ισχύει $\frac{2}{\pi} \leq \sin x \leq x$. Επομένως αφού $\theta \in [0,\pi]$ (είναι κυρτή γωνία) έχουμε ότι $\frac{\theta}{2} \in [0,\frac{\pi}{2}]$. Άρα

$$\frac{2}{\pi}\frac{\theta}{2} \leq \sin\frac{\theta}{2} \leq \frac{\theta}{2} \Rightarrow \frac{2}{\pi}\frac{\theta}{2} \leq \frac{\|x-y\|_2}{2} \leq \frac{\theta}{2}$$

$$\Longrightarrow \frac{2}{\pi}\rho(x,y) \leq \|x-y\|_2 \leq \rho(x,y) \ .$$

- Για κάθε σύνολο *Borel* $A \subseteq \mathbb{S}^{n-1}$ ορίζουμε το μέτρο πιθανότητας σ ορίζοντας

$$\sigma(A) := \frac{|\tilde{A}|}{|B|}$$

όπου $B = B(0,1)$ η μοναδιαία Ευκλείδεια μπάλα και

$$\tilde{A} := \{sx :\ x \in A,\ 0 \leq s \leq 1\} \ .$$

Το $\sigma(A)$ είναι το ποσοστό που καταλαμβάνει η επιφάνεια του A, πάνω στην επιφάνεια της σφαίρας.

Διατύπωση ισοπεριμετρικού προβλήματος: Δίνονται $a \in (0,1)$ και $t > 0$. Ανάμεσα σε όλα τα *Borel* σύνολα $A \subseteq \mathbb{S}^{n-1}$ για τα οποία $\sigma(A) = a$, να βρεθούν εκείνα για τα οποία ελαχιστοποιείται η επιφάνεια $\sigma(A_t)$ της t-επέκτασης του A. Η απάντηση δίνεται από το ακόλουθο θεώρημα.

Θεώρημα 2.7. *Έστω $a \in (0,1)$ και $B(x,r) = \{y \in \mathbb{S}^{n-1} : \rho(x,y) \leq r\}$ μια μπάλα στην $\mathbb{S}^{n-1}$ ώστε $\sigma(B(x,r)) = a$. Τότε για κάθε $A \subseteq \mathbb{S}^{n-1}$ με $\sigma(A) = a$ και για κάθε $t > 0$ έχουμε*

$$\sigma(A_t) \geq \sigma(B(x,r)_t) = \sigma(B(x,r+t)) \ .$$

Δηλαδή η μπάλα στην $\mathbb{S}^{n-1}$ είναι η λύση του ισοπεριμετρικού προβλήματος.

Εμείς θα αποδείξουμε την προσεγγιστική ισοπεριμετρική ανισότητα στην ειδική περίπτωση όπου $a = \frac{1}{2}$, η οποία διατυπώνεται μέσω του θεωρήματος 2.8.

Θεώρημα 2.8. *Έστω $A \subseteq \mathbb{S}^{n-1}$ με $\sigma(A) = \frac{1}{2}$ και έστω $t > 0$. Τότε*

$$\sigma(A_t) \geq 1 - \sqrt{\frac{\pi}{8}}\, e^{\frac{-t^2 n}{2}} \ .$$

Παρατήρηση: Όσο μικρό $t > 0$ και αν διαλέξουμε η ακολουθία $e^{\frac{-t^2 n}{2}}$ τείνει στο 0 όταν η διάσταση n πηγαίνει στο άπειρο και μάλιστα με ταχύ ρυθμό (εκθετικά ως προς n). Επομένως το ποσοστό της σφαίρας που μένει έξω από την t-επέκταση οποιουδήποτε συνόλου $A \subseteq \mathbb{S}^{n-1}$ με $\sigma(A) = \frac{1}{2}$ είναι «σχεδόν μηδενικό».

Η απόδειξη του θεωρήματος 2.8 θα γίνει με την βοήθεια του εξής λήμματος.

Λήμμα 2.5. *Θεωρούμε το ομοιόμορφο μέτρο πιθανότητας μ_B στην Ευκλείδεια μοναδιαία μπάλα $B = B(0,1)$, δηλαδή για κάθε σύνολο Borel $A \subseteq B$*

$$\mu_B(A) = \frac{|A|}{|B|} \ .$$

Αν A, C συμπαγή υποσύνολα της B και

$$d(A,C) := \min\{\|a-c\|_2 : \ a \in A, \ c \in C\} = \rho > 0 \ .$$

Τότε

$$\min\{\mu_B(A), \ \mu_B(C)\} \leq e^{\frac{-\rho^2 n}{8}} \ .$$

Απόδειξη: Θεωρούμε το σύνολο $\frac{A+C}{2}$. Από την ανισότητα *Brunn-Minkowski* έχουμε ότι

$$\left|\frac{A+C}{2}\right| \geq \min\{|A|, \ |C|\} \ .$$

Άρα,

$$\left|\frac{A+C}{2}\right| \geq \min\{|A|,\ |C|\} \Rightarrow \frac{\left|\frac{A+C}{2}\right|}{|B|} \geq \min\Big\{\frac{|A|}{|B|},\ \frac{|A|}{|B|}\Big\}\ .$$

Συνεπώς,

$$\mu_B\left(\frac{A+C}{2}\right) \geq \min\{\mu_B(A),\ \mu_B(C)\}\ . \tag{2.11}$$

Έστω τώρα, $a \in A$ και $c \in C$. Επειδή $A \subseteq B$ και $C \subseteq B$ έχουμε $\|a\|_2 \leq 1$ και $\|c\|_2 \leq 1$. Από τον κανόνα του παραλληλογράμμου παίρνουμε

$$\|a+c\|_2^2 + \|a-c\|_2^2 = 2\|a\|_2^2 + 2\|c\|_2^2 \Longrightarrow$$

$$\|a+c\|_2^2 = 2\|a\|_2^2 + 2\|c\|_2^2 - \|a-c\|_2^2 \leq 2+2-\rho^2 = 4-\rho^2 \Longrightarrow$$

$$\Big\|\frac{a+c}{2}\Big\|_2 \leq \Big(1-\frac{\rho^2}{4}\Big)^{1/2}\ .$$

Επομένως,

$$\frac{A+C}{2} \subseteq \Big(1-\frac{\rho^2}{4}\Big)^{1/2} B\ . \tag{2.12}$$

Από τις σχέσεις (2.11) και (2.12) παίρνουμε

$$\min\{\mu_B(A),\ \mu_B(C)\} \leq \mu_B\Big(\frac{A+C}{2}\Big) \leq \mu_B\Big(\Big(1-\frac{\rho^2}{4}\Big)^{1/2} B\Big) = \Big(1-\frac{\rho^2}{4}\Big)^{n/2}$$

και επειδή γενικά ισχύει ότι $1-x \leq e^{-x}$ παίρνουμε το ζητούμενο

$$\min\{\mu_B(A),\ \mu_B(C)\} \leq \Big(e^{-\frac{\rho^2}{4}}\Big)^{n/2} = e^{-\frac{\rho^2 n}{8}}\ .$$

$\square$

Απόδειξη:(του Θεωρήματος 2.8) Έστω $A \subseteq \mathbb{S}^{n-1}$ με $\sigma(A) = \frac{1}{2}$ και έστω $t > 0$. Θέτουμε $C = \mathbb{S}^{n-1} \setminus A_t$ και θεωρούμε τα υποσύνολα της μοναδιαίας σφαίρας

$$A_1 = \{\rho a :\ a \in A,\ \frac{1}{2} \leq \rho \leq 1\} \text{ και } C_1 = \{\rho c :\ c \in C,\ \frac{1}{2} \leq \rho \leq 1\}\ .$$

Παρατηρούμε ότι

$$d(A_1, C_1) \geq \sin\frac{t}{2} \geq \frac{t}{\pi}\ . \tag{2.13}$$

Πράγματι,

Επειδή $\min\{\mu_B(A_1),\ \mu_B(C_1)\} = \mu_B(C_1)$ από το λήμμα 2.5 έχουμε

$$\mu_B(C_1) \le e^{\frac{-\rho^2 n}{8}} \Longrightarrow |C_1| \le e^{\frac{-\rho^2 n}{8}}\ |B|$$

και από τη σχέση (2.13) παίρνουμε ότι

$$|C_1| \le e^{-\frac{t^2 n}{8\pi^2}}\ |B|\ . \tag{2.14}$$

Τώρα υπολογίζουμε και βρίσκουμε ότι

$$|C_1| = (1-2^{-n})|\tilde{C}| \tag{2.15}$$

Πράγματι, αν

$$C_2 = \frac{1}{2}\tilde{C} \Longrightarrow |C_2| = \frac{1}{2^n}|\tilde{C}|\ .$$

Όμως

$$|\tilde{C}| = |C_1| + |C_2| \Longrightarrow |C_1| = (1-2^{-n})|\tilde{C}|\ .$$

Συνδυάζοντας τις τελευταίες σχέσεις παίρνουμε

$$\begin{aligned} \sigma(A_t^c) &= \sigma(\mathcal{S}^{n-1}\backslash A_t) = \sigma(C) = \frac{|\tilde{C}|}{|B|} = \frac{|C_1|}{(1-2^{-n})|B|} \\ &\le \frac{1}{(1-2^{-n})|B|}\ e^{-\frac{t^2 n}{8\pi^2}}\ |B|\ . \end{aligned}$$

Άρα,

$$1-\sigma(A_t) \le \frac{1}{1-2^{-n}}\ e^{-\frac{t^2 n}{8\pi^2}} \Longrightarrow$$

$$\sigma(A_t) \ge 1 - \frac{1}{1-2^{-n}}\ e^{-\frac{t^2 n}{8\pi^2}}.$$

□

2.6 Ισοπεριμετρική ανισότητα στο χώρο του Gauss

Θα μελετήσουμε το μετρικό χώρο πιθανότητας $\Gamma_n = (\mathbb{R}^n,\ \mathcal{A},\ \|\cdot\|_2,\ \gamma_n)$, όπου

- $\mathcal{A}$, η *Borel* σ-άλγεβρα του $\mathbb{R}^n$.
- $\|\cdot\|_2$, η Ευκλείδεια μετρική στον $\mathbb{R}^n$.
- γ_n, μέτρο πιθανότητας που έχει συνάρτηση πυκνότητας την
$$\gamma_n(x) = \frac{1}{(2\pi)^{n/2}}\ e^{-\|x\|_2^2/2}.$$
Δηλαδή αν A μη κενό *Borel* υποσύνολο του $\mathbb{R}^n$, τότε το μέτρο του A είναι
$$\gamma_n(A) = \frac{1}{(2\pi)^{n/2}}\ \int_A e^{-\|x\|_2^2/2} dx.$$

Το γ_n λέγεται **μέτρο του Gauss** και ο μετρικός χώρος πιθανότητας Γ_n είναι ο n-διάστατος **χώρος του Gauss**. Θα δούμε δύο σημαντικές ιδιότητες του μέτρου του *Gauss*.

(i) Είναι μέτρο γινόμενο, δηλαδή $\gamma_n(x_1,\ldots,x_n) = \gamma_n(x_1)\cdots\gamma_n(x_n)$.
Απόδειξη:
$$\begin{aligned}
\gamma_n(x_1,\ldots,x_n) &= \frac{1}{(2\pi)^{n/2}}\ e^{-\|(x_1,\ldots,x_n)\|_2^2/2} \\
&= \frac{1}{(2\pi)^{1/2}}\cdots\frac{1}{(2\pi)^{1/2}}\ e^{-x_1^2/2}\cdots e^{-x_n^2/2} \\
&= \left(\frac{1}{(2\pi)^{1/2}}\ e^{-x_1^2/2}\right)\cdots\left(\frac{1}{(2\pi)^{1/2}}\ e^{-x_n^2/2}\right) \\
&= \gamma_1(x_1)\cdots\gamma_n(x_n).
\end{aligned}$$

(ii) Είναι αναλλοίωτο ως προς τους ορθογώνιους μετασχηματισμούς. Δηλαδή, αν $A \in \mathcal{A}$ και $U \in O(n) = \{U\ n\times n \text{ πίνακες } : UU^t = I\}$ τότε $\gamma_n(U(A)) = \gamma_n(A)$.
Απόδειξη:
$$\gamma_n(U(A)) = \frac{1}{(2\pi)^{n/2}}\int_{U(A)} e^{-\|x\|_2^2/2} dx.$$
Θέτουμε $x = Uy \in U(A)$, άρα $y \in A$ και $dx = |J|dy = |\det U|\ dy$. Συνεπώς,
$$\gamma_n(U(A)) = \int_A e^{-\|Uy\|_2^2/2} |\det U|\ dy.$$

Όμως, $|\det U| = 1$ και $\|Uy\|_2^2 = \|y\|_2^2$. Πράγματι,

$$\begin{aligned}\|Uy\|^2 &= \langle Uy, Uy\rangle \\ &= \left\langle y, U^t U y\right\rangle \\ &= \langle y, y\rangle \\ &= \|y\|^2.\end{aligned}$$

Άρα,

$$\gamma_n(U(A)) = \int_A e^{-\|y\|_2^2/2} dy = \gamma_n(A).$$

Η ισοπεριμετρική ανισότητα στο χώρο του *Gauss*, εκφράζεται μέσω του παρακάτω θεωρήματος.

Θεώρημα 2.9. *Έστω $a \in (0,1)$ και έστω $H = \{x \in \mathbb{R}^n : \langle x, \theta\rangle \leq \lambda\}$ ένας ημίχωρος του $\mathbb{R}^n$, όπου $\theta \in \mathbb{S}^{n-1}$ και $\lambda \in \mathbb{R}$, με $\gamma_n(H) = a$. Τότε για κάθε $t > 0$ και για κάθε υποσύνολο Borel A του $\mathbb{R}^n$ με $\gamma_n(A) = a$ έχουμε*

$$\gamma_n(A_t) \geq \gamma_n(H_t).$$

Δηλαδή ο ημίχωρος H είναι η λύση του ισοπεριμετρικού προβλήματος.

Πόρισμα 2.3. *Αν $\gamma_n(A) \geq 1/2$, τότε για κάθε $t > 0$*

$$1 - \gamma_n(A_t) \leq \frac{1}{2}\, e^{-t^2/2}.$$

Δηλαδή,

$$\alpha(\Gamma_n, t) \leq \frac{1}{2}\, e^{-t^2/2}.$$

Απόδειξη: Από το θεώρημα 2.9 έχουμε ότι

$$1 - \gamma_n(A_t) \leq 1 - \gamma_n(H_t) \tag{2.16}$$

όπου H ημίχωρος του $\mathbb{R}^n$ με $\gamma_n(H) = 1/2$. Αφού το γ_n είναι αναλλοίωτο ως προς τους ορθογώνιους μετασχηματισμούς, μπορούμε να υποθέσουμε ότι $H = \{x \in \mathbb{R}^n : x_1 \leq 0\}$, οπότε $H_t = \{x \in \mathbb{R}^n : x_1 \leq t\}$. Θα υπολογίσουμε το μέτρο του H_t.

$$\begin{aligned}\gamma_n(H_t) &= \frac{1}{(2\pi)^{n/2}} \int_{H_t} e^{-\|x\|_2^2/2} dx \\ &= \frac{1}{(2\pi)^{n/2}} \int_{-\infty}^{t}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{-x_1^2/2} e^{-x_2^2/2}\cdots e^{-x_n^2/2} dx_1 dx_2\cdots dx_n \\ &= \frac{1}{(2\pi)^{n/2}} \int_{-\infty}^{t} e^{-x_1^2/2}\left(\int_{-\infty}^{\infty} e^{-x^2/2} dx\right)^{n-1} dx_1.\end{aligned}$$

Θέτουμε
$$I = \int_{\mathbb{R}} e^{-x^2/2}dx \Longrightarrow I^2 = \int_{\mathbb{R}^2} e^{-\frac{x^2+y^2}{2}}dxdy.$$
Κάνοντας την αλλαγή μεταβλητής $r^2 = x^2 + y^2$, παίρνουμε
$$I^2 = \int_0^{2\pi}\int_0^{\infty} e^{-r^2/2}rdrd\theta = 2\pi\int_0^{\infty} e^{-r^2/2}rdr = 2\pi.$$
Άρα $I = \sqrt{2\pi}$. Αντικαθιστώντας έχουμε
$$\gamma_n(H_t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-s^2/2}ds.$$
Οπότε,
$$1 - \gamma_n(H_t) = \frac{1}{\sqrt{2\pi}}\int_t^{\infty} e^{-s^2/2}ds. \tag{2.17}$$
Θεωρούμε τη συνάρτηση
$$F(x) = e^{x^2/2}\int_x^{\infty} e^{-s^2/2}ds.$$
Παραγωγίζοντας θα δείξουμε ότι η $F(x)$ είναι φθίνουσα στο $[0,\infty)$. Πράγματι,
$$F'(x) = xe^{x^2/2}\int_x^{\infty} e^{-s^2/2}ds - 1.$$
Τώρα,
$$\int_x^{\infty} e^{-s^2/2}ds = \int_x^{\infty}\frac{1}{s}se^{-s^2/2}ds \leq \frac{1}{x}\int_x^{\infty} e^{-s^2/2}ds$$
διότι $s \geq x \Rightarrow 1/s \leq 1/x$. Άρα
$$xe^{x^2/2}\int_x^{\infty} e^{-s^2/2}ds \leq 1 \Rightarrow F'(x) \leq 0.$$
Επομένως από την ανισότητα $F(t) \geq F(0)$ παίρνουμε ότι
$$e^{t^2/2}\int_t^{\infty} e^{-s^2/2}ds \leq \sqrt{\frac{\pi}{2}}. \tag{2.18}$$
Συνδυάζοντας τις σχέσεις (2.16), (2.17) και (2.18) προκύπτει το ζητούμενο:
$$\begin{aligned} 1-\gamma_n(A_t) &\leq 1-\gamma_n(H_t) \\ &= \frac{1}{\sqrt{2\pi}}\int_t^{\infty} e^{-s^2/2}ds \\ &\leq \frac{1}{\sqrt{2\pi}}\sqrt{\frac{\pi}{2}}\, e^{-t^2/2} \\ &= \frac{1}{2}e^{-t^2/2}. \end{aligned}$$

□

Για την απόδειξη της προσεγγιστικής ισοπεριμετρικής ανισότητας χρησιμοποιήσαμε την ισοπεριμετρική ανισότητα. Τώρα θα αποδείξουμε ένα γενικότερο αποτέλεσμα της προσεγγιστικής ισοπεριμετρικής ανισότητας για το χώρο του *Gauss* χρησιμοποιόντας την ανισότητα *Prékopa-Leindler*.

Θεώρημα 2.10. *Έστω A μη κενό Borel υποσύνολο του $\mathbb{R}^n$. Τότε*

$$\int_{\mathbb{R}^n} e^{d(x,A)^2/4} d\gamma_n(x) \leq \frac{1}{\gamma_n(A)} \tag{2.19}$$

όπου $d(x,A) = \inf\{\|x-y\|_2 : \ y \in A\}$. Επομένως αν $\gamma_n(A) = 1/2$ τότε για κάθε $t > 0$

$$1 - \gamma_n(A_t) \leq 2e^{-t^2/4}. \tag{2.20}$$

Απόδειξη: Θεωρούμε τις εξής συναρτήσεις $m, f, g : \mathbb{R}^n \to \mathbb{R}_+$ με

$$f(x) = e^{d(x,A)^2/4}\gamma_n(x), \quad g(x) = \mathcal{X}_A(x)\gamma_n(x), \quad m(x) = \gamma_n(x).$$

Οι f, g, m είναι προφανώς μετρήσιμες και ολοκληρώσιμες. Αν ικανοποιούνται οι προϋποθέσεις της ανισότητας *Prékopa-Leindler*, τότε

$$\left(\int_{\mathbb{R}^n} m\right)^2 \geq \int_{\mathbb{R}^n} f \int_{\mathbb{R}^n} g.$$

Όμως

$$\left(\int_{\mathbb{R}^n} m\right)^2 = \left(\int_{\mathbb{R}^n} \gamma_n(x)dx\right)^2 = \left(\int_{\mathbb{R}^n} \frac{1}{(2\pi)^{n/2}} e^{-\|x\|_2^2/2} dx\right)^2 = 1.$$

Άρα,

$$\int_{\mathbb{R}^n} f \int_{\mathbb{R}^n} g \leq 1 \Rightarrow$$
$$\int_{\mathbb{R}^n} e^{d(x,A)^2/4}\gamma_n(x)dx \cdot \int_{\mathbb{R}^n} \mathcal{X}_A(x)\gamma_n(x)dx \leq 1 \Rightarrow$$
$$\int_{\mathbb{R}^n} e^{d(x,A)^2/4} d\gamma_n(x) \cdot \int_A \gamma_n(x)dx \leq 1 \Rightarrow$$
$$\int_{\mathbb{R}^n} e^{d(x,A)^2/4} d\gamma_n(x) \cdot \gamma_n(A) \leq 1 \Rightarrow$$

$$\int_{\mathbb{R}^n} e^{d(x,A)^2/4} d\gamma_n(x) \leq \frac{1}{\gamma_n(A)}.$$

Άρα αρκεί να δείξουμε ότι ικανοποιείται η προϋπόθεση της ανισότητας *Prékopa-Leindler*, δηλαδή ότι

$$m\Big(\frac{x+y}{2}\Big)^2 \geq f(x)g(x).$$

Ισοδύναμα ότι

$$\gamma_n\Big(\frac{x+y}{2}\Big)^2 \geq e^{d(x,A)^2/4}\gamma_n(x)\mathcal{X}_A(y)\gamma_n(y) \Leftrightarrow$$

$$\frac{1}{(2\pi)^n}\Big(e^{-\|\frac{x+y}{2}\|_2^2/2}\Big)^2 \geq e^{d(x,A)^2/4}\frac{1}{(2\pi)^{n/2}}e^{-\|x\|_2^2/2}\mathcal{X}_A(y)\frac{1}{(2\pi)^{n/2}}e^{-\|y\|_2^2/2} \Leftrightarrow$$

$$e^{-\|\frac{x+y}{2}\|_2^2} \geq e^{d(x,A)^2/4-\|x\|_2^2/2-\|y\|_2^2/2}\mathcal{X}_A(y).$$

Αν $y \notin A$, τότε το δεξιό μέλος είναι 0 και η ανισότητα ισχύει τετριμμένα. Αν $y \in A$ τότε $\mathcal{X}_A(y) = 1$, άρα αρκεί να δείξουμε ότι

$$-\Big\|\frac{x+y}{2}\Big\|^2 \geq \frac{d(x,A)^2}{4} - \frac{\|x\|_2^2}{2} - \frac{\|y\|_2^2}{2} \Leftrightarrow$$

$$2\|x\|_2^2 + 2\|y\|_2^2 \geq d(x,A)^2 + \|x+y\|^2.$$

Όμως $d(x,A) \leq \|x-y\|_2$ και από τον κανόνα του παραλληλογράμμου έχουμε

$$\|x+y\|^2 + d(x,A)^2 \leq \|x+y\|^2 + \|x-y\|^2 = 2\|x\|_2^2 + 2\|y\|_2^2,$$

δηλαδή το ζητούμενο.

Για να δείξουμε την (2.20) αν $\gamma_n(A) = 1/2$ από την ανισότητα *Markov* έχουμε

$$\begin{aligned}
\int_{\mathbb{R}^n} e^{d(x,A)^2/4} d\gamma_n(x) &\geq \int_{\{x\in\mathbb{R}^n : d(x,A)>t\}} e^{d(x,A)^2/4} d\gamma_n(x)\\
&\geq \int_{\{x\in\mathbb{R}^n : d(x,A)>t\}} e^{t^2/4} d\gamma_n(x)\\
&= e^{t^2/4}\gamma_n(\{x\in\mathbb{R}^n : d(x,A)>t\}).
\end{aligned}$$

Άρα

$$\gamma_n(\{x\in\mathbb{R}^n : d(x,A)>t\}) \leq e^{-t^2/4}\int_{\mathbb{R}^n} e^{d(x,A)^2/4} d\gamma_n(x) \Rightarrow$$

$$\gamma_n(A_t^c) \leq e^{-t^2/4}\frac{1}{\gamma_n(A)} = 2e^{-t^2/4} \Rightarrow$$

$$1-\gamma_n(A_t) \leq 2e^{-t^2/4}.$$

□

Κεφάλαιο 3

Συγκέντρωση του μέτρου σε χώρους γινόμενα

3.1 Η ανισότητα του Talagrand

Ο μετρικός χώρος πιθανότητας που θα μελετήσουμε σε αυτό το κεφάλαιο είναι ο (E_2^n, d_n, μ_n) όπου

- $E_2^n = \{1, -1\}^n$ το σύνολο των κορυφών του κύβου στον $\mathbb{R}^n$. Το τυπικό στοιχείο του E_2^n είναι $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$ με $\varepsilon_i = \pm 1$ για κάθε $i = 1, \dots, n$.

- Εφοδιάζουμε τον E_2^n με το ομοιόμορφο μέτρο πιθανότητας μ_n. Αν A υποσύνολο του E_2^n τότε $\mu_n(A) = |A|/2^n$, όπου $|A|$ ο πληθάριθμος του συνόλου A.

- Έστω $x = (x_1, \dots, x_n)$ και $y = (y_1, \dots, y_n)$ στοιχεία του E_2^n με $x_i, y_i = \pm 1$ για κάθε $i = 1, \dots, n$. Ορίζουμε την μετρική d_n στον E_2^n ως

$$d_n(x, y) = \frac{1}{n}|\{i \leq n : \; x_i \neq y_i\}| = \frac{1}{2n}\sum_{i=1}^{n}|x_i - y_i|.$$

Η d είναι μετρική διότι:

(i) $d_n(x, y) = 0 \Leftrightarrow x = y$. Πράγματι,

$$\begin{aligned} d_n(x, y) = 0 \quad &\Leftrightarrow \quad \frac{1}{n}|\{i \leq n : \; x_i \neq y_i\}| = 0 \\ &\Leftrightarrow \quad |\{i \leq n : \; x_i \neq y_i\}| = 0 \\ &\Leftrightarrow \quad x_i = y_i, \; \forall \; i = 1, \dots, n \\ &\Leftrightarrow \quad x = y. \end{aligned}$$

(ii) Η $d_n(x,y) = d_n(y,x)$ είναι προφανής.

(iii) (Τριγωνική ανισότητα) $d_n(x,y) \leq d_n(x,z) + d_n(z,y)$. Πράγματι,

$$\begin{aligned} d_n(x,y) &= \frac{1}{2n}\sum_{i=1}^{n}|x_i - y_i| \\ &= \frac{1}{2n}\sum_{i=1}^{n}|(x_i - z_i) + (z_i - y_i)| \\ &\leq \frac{1}{2n}\sum_{i=1}^{n}(|x_i - z_i| + |z_i - y_i|) \\ &= \frac{1}{2n}\sum_{i=1}^{n}|x_i - z_i| + \frac{1}{2n}\sum_{i=1}^{n}|z_i - y_i| \\ &= d_n(x,z) + d_n(z,y). \end{aligned}$$

Παρατηρούμε ότι οι τιμές που μπορεί να πάρει η μετρική $d_n(x,y)$ είναι πεπερασμένες το πλήθος: $0, 1/n, 2/n, \ldots, 1$. Η t-επέκταση ενός υποσυνόλου A του E_2^n είναι το σύνολο $A_t = \{x \in E_2^n : \ d_n(x,A) \leq t\}$. Αν το t παίρνει τιμές μέσα από ένα σύνολο της μορφής $[k/n, (k+1)/n]$ η t-επέκτασή του παραμένει αμετάβλητη, επομένως οι τιμές του t που παρουσιάζουν ενδιαφέρον είναι της μορφής $t = k/n$ για $k = 1, \ldots, n$.

Το ισοπεριμετρικό πρόβλημα στο μετρικό χώρο πιθανότητας (E_2^n, d_n, μ_n) διατυπώνεται ως εξής: Δίνονται ένας φυσικός $m \in \{1, 2, \ldots, 2^n\}$ και $t = k/n$ για $k = 1, \ldots, n$. Από όλα τα $A \subseteq E_2^n$ με $|A| = m$ να βρεθούν εκείνα για τα οποία ελαχιστοποιείται το $\mu_n(A_t)$. Δηλαδή, για ποιά $A \subseteq E_2^n$ η k/n-επέκταση $A_{k/n}$ έχει το μικρότερο πλήθος στοιχείων. Η απάντηση στο ισοπεριμετρικό πρόβλημα είναι η παρακάτω ισοπεριμετρική ανισότητα.

Θεώρημα 3.1. ***(Ισοπεριμετρική ανισότητα)*** *Έστω $A \subseteq E_2^n$ με $|A| = m = \sum_{k=0}^{l}\binom{n}{k}$. Τότε για κάθε $s = 1, \ldots, n-1$ έχουμε*

$$\mu_n(A_{s/n}) \geq \frac{1}{2^n}\sum_{k=0}^{l+s}\binom{n}{k} = \mu_n(B(x, l/n)_{s/n}) = \mu_n(B(x, (l+s)/n)).$$

Δηλαδή η μπάλα $B(x, s/n) = \{y \in E_2^n : \ d_n(x,y) \leq s/n\}$ είναι η λύση του ισοπεριμετρικού προβλήματος. Εμείς θα αποδείξουμε την εξής προσεγγιστική ισοπεριμετρική ανισότητα:

Θεώρημα 3.2. *Έστω $A \subseteq E_2^n$. Αν $\mu_n(A) \geq 1/2$ και $t > 0$, τότε*

$$\mu_n(A_t^c) \leq 2e^{-t^2 n/2}.$$

Δηλαδή για τη συνάρτηση συγκέντρωσης του E_2^n έχουμε $\alpha(E_2^n, t) \leq 2e^{-t^2 n/2}$. Η απόδειξη του θεωρήματος βασίζεται στην ανισότητα του Talagrand.

Ορισμός 3.1. ***Κυρτή θήκη*** *ενός συνόλου $A \subseteq \mathbb{R}^n$ είναι το μικρότερο κυρτό σύνολο που περιέχει το A. Συμβολικά*

$$conv(A) = \Big\{ \sum_{i=1}^m t_i y_i : \ y_i \in A, \ t_i \in [0,1], \ \sum_{i=1}^m t_i = 1, \ m \in \mathbb{N} \Big\}.$$

Θεώρημα 3.3. ***(Ανισότητα Talagrand)*** *Έστω A μη κενό υποσύνολο του E_2^n. Για κάθε $x \in E_2^n$ ορίζουμε τη συνάρτηση*

$$\Phi_A(x) := \min\{\|x-y\|_2 : \ y \in conv(A)\}.$$

Τότε

$$\mathbb{E}(e^{\Phi_A^2/8}) = \int_{E_2^n} e^{\Phi_A^2/8} d\mu_n(x) \leq \frac{1}{\mu_n(A)}.$$

Πριν δώσουμε την απόδειξη του θεωρήματος 3.3 θα δούμε πως μέσω αυτού αποδεικνύεται η προσεγγιστική ισοπεριμετρική ανισότητα. Η συνάρτηση Φ_A και η συνάρτηση απόστασης

$$d_n(x, A) = \min\Big\{ \frac{1}{2n} \sum_{i=1}^n |x_i - y_i| : \ y \in A \Big\}$$

συγκρίνονται με το παρακάτω λήμμα.

Λήμμα 3.1. *Για κάθε μη κενό υποσύνολο A του E_2^n και για κάθε $x \in E_2^n$ έχουμε*

$$2\sqrt{n} d_n(x, A) \leq \Phi_A(x).$$

Απόδειξη: Έστω $x \in E_2^n$. Για κάθε $y \in A$ έχουμε

$$\langle x - y, x \rangle = \sum_{i=1}^n x_i(x_i - y_i).$$

Όμως διακρίνοντας περιπτώσεις για τις τιμές των $x_i, y_i = \pm 1$ παρατηρούμε ότι $x_i(x_i - y_i) = |x_i - y_i|$. Άρα

$$\langle x - y, x \rangle = \sum_{i=1}^n |x_i - y_i| = 2n d_n(x, y) \geq 2n d_n(x, A).$$

Τώρα από την ανισότητα Cauchy-Schwarz για κάθε $y \in conv(A)$ έχουμε

$$\langle x - y, x \rangle \leq \|x - y\|_2 \|x\|_2 = \|x - y\|_2 \sqrt{n}.$$

Άρα

$$2nd_n(x,A) \leq \langle x-y, x\rangle \leq \sqrt{n}\|x-y\|_2 \Rightarrow$$
$$2\sqrt{n}d_n(x,A) \leq \langle x-y, x\rangle \leq \|x-y\|_2$$

και από τον ορισμό της Φ_A έπεται το ζητούμενο.

□

Απόδειξη:(Θεωρήματος 3.2) Έστω $A \subseteq E_2^n$ με $\mu_n(A) \geq 1/2$. Αν

$$x \notin A_t \Rightarrow x \in A_t^c \Rightarrow d_n(x,A) > t.$$

Από το λήμμα 3.1 έχουμε ότι

$$\Phi_A(x) \geq 2\sqrt{n}d_n(x,A) \geq 2t\sqrt{n}.$$

Άρα

$$A_t^c \subseteq \{x \in E_2^n : \ \Phi_A(x) \geq 2t\sqrt{n}\} \Rightarrow$$
$$\mu_n(A_t^c) \leq \mu_n(\{x \in E_2^n : \ \Phi_A(x) \geq 2t\sqrt{n}\}).$$

Από την ανισότητα Markov παίρνουμε

$$\begin{aligned}
\mathbb{E}(e^{\Phi_A^2/8}) &= \int_{E_2^n} e^{\Phi_A^2/8} d\mu_n(x) \\
&\geq \int_{\{x\in E_2^n:\ \Phi_A(x)\geq 2t\sqrt{n}\}} e^{\Phi_A^2/8} d\mu_n(x) \\
&\geq \int_{\{x\in E_2^n:\ \Phi_A(x)\geq 2t\sqrt{n}\}} e^{4t^2n/8} d\mu_n(x) \\
&= e^{t^2n/2}\mu_n(\{x \in E_2^n : \ \Phi_A(x) \geq 2t\sqrt{n}\}).
\end{aligned}$$

Άρα

$$\mu_n(A_t^c) \leq \mu_n(\{x \in E_2^n : \ \Phi_A(x) \geq 2t\sqrt{n}\}) \leq e^{-t^2n/2}\int_{E_2^n} e^{\Phi_A^2(x)/8} d\mu_n(x).$$

Τώρα από το θεώρημα 3.3 παίρνουμε

$$\begin{aligned}
\mu_n(A_t^c) &\leq e^{-t^2n/2}\int_{E_2^n} e^{\Phi_A^2(x)/8} d\mu_n(x) \\
&\leq e^{-t^2n/2}\frac{1}{\mu_n(A)} \\
&\leq e^{-t^2n/2}\frac{1}{1/2} \\
&\leq 2e^{-t^2n/2},
\end{aligned}$$

δηλαδή το ζητούμενο.

□

Απόδειξη: (Θεωρήματος 3.3) Η απόδειξη θα γίνει επαγωγικά ως προς το πλήθος των σημείων του συνόλου A. Αν $|A| = 1$ δηλαδή $A = \{y\}$, τότε για κάθε $x \in E_2^n$, $\Phi_A(x) = \|x - y\|_2$. Άρα

$$\mathbb{E}(e^{\Phi_A^2(x)/8}) = \mathbb{E}(e^{\|x-y\|_2^2/8}) = \frac{1}{2^n} \sum_{x \in E_2^n} e^{\|x-y\|_2^2/8}.$$

Παρατηρούμε ότι αν το x διαφέρει από το y σε i το πλήθος συντεταγμένες ($i = 1, \ldots, n$), τότε $\|x - y\|_2^2 = 4i$. Πράγματι, για τις συντεταγμένες που είναι ίδιες έχουμε $|x_i - y_i| = 0$. Για $i = 1$, δηλαδή αν το x διαφέρει από το y σε μία συντεταγμένη, έστω την i-οστή, έχουμε

$$\begin{aligned} \|x - y\|_2^2 &= (x_1 - y_1)^2 + \ldots + (x_i - y_i)^2 + \ldots + (x_n - y_n)^2 \\ &= (x_i - y_i)^2 = 2^2 = 4 = 4i. \end{aligned}$$

Έστω ότι ισχύει για i το πλήθος συντεταγμένες, δηλαδή

$$\|x - y\|_2^2 = (x_1 - y_1)^2 + \ldots + (x_i - y_i)^2 = 4i.$$

Θα δείξουμε ότι ισχύει για $i + 1$. Πράγματι,

$$\begin{aligned} \|x - y\|_2^2 &= (x_1 - y_1)^2 + \ldots + (x_i - y_i)^2 + (x_{i+1} - y_{i+1})^2 \\ &= 4i + (x_{i+1} - y_{i+1})^2 = 4i + 4 = 4(i + 1). \end{aligned}$$

Τώρα το πλήθος των $x \in E_2^n$ που διαφέρουν από το y σε i το πλήθος συντεταγμένες είναι $\binom{n}{i}$, για τα υπόλοιπα όταν $x_i = y_i$ έχουμε ότι $\|x - y\|_2 = 0$. Άρα,

$$\begin{aligned} \mathbb{E}(e^{\|x-y\|_2^2/8}) &= \frac{1}{2^n} \sum_{i=0}^{n} \binom{n}{i} e^{4i/8} + 0 \\ &= \frac{1}{2^n} \sum_{i=0}^{n} \binom{n}{i} (e^{1/2})^i 1^{n-i} \\ &= \frac{1}{2^n} (1 + e^{1/2})^n = \Big(\frac{1 + e^{1/2}}{2}\Big)^n \\ &\le \Big(\frac{1 + e}{2}\Big)^n \le \Big(\frac{1 + 3}{2}\Big)^n = 2^n = \frac{1}{\mu_n(A)}. \end{aligned}$$

Η τελευταία ισότητα είναι σωστή αφού, $|A| = 1$ οπότε $\mu_n(A) = 1/2^n$.

Έστω τώρα ότι $|A| \geq 2$. Στην περίπτωση που $n = 1$ αναγκαστικά $A = E_2^1 = \{1, -1\}$. Επομένως για κάθε $x \in E_2^1 \Rightarrow x \in A \Rightarrow \Phi_A(x) = 0$. Άρα,

$$\mathbb{E}(e^{\Phi_A^2(x)/8}) = \mathbb{E}(e^0) = 1 = \frac{1}{\mu_1(A)},$$

αφού $\mu_1(A) = |A|/2^1 = 2/2 = 1$. Δηλαδή το ζητούμενο ισχύει σαν ισότητα.

Για το επαγωγικό βήμα θεωρούμε $A \subseteq E_2^{n+1}$ με $|A| \geq 2$. Χωρίς περιορισμό της γενικότητας υποθέτουμε ότι

$$A = (A_1 \times \{1\}) \cup (A_{-1} \times \{-1\})$$

όπου

$$A_1 = \{x \in E_2^n : \ (x, 1) \in A\} \text{ και } A_{-1} = \{x \in E_2^n : \ (x, -1) \in A\}.$$

Επίσης μπορούμε να υποθέσουμε ότι $A_1, \ A_{-1}$ μη κενά και ότι $|A_{-1}| \leq |A_1|$. Θα χρειαστούμε δύο λήμματα.

Λήμμα 3.2. *Για κάθε $x \in E_2^n$, ισχύει*

$$\Phi_A((x, 1)) \leq \Phi_{A_1}(x).$$

Απόδειξη: Έστω $y \in conv(A_1)$ τέτοιο ώστε $\Phi_{A_1}(x) = \|x - y\|_2$. Έχουμε ότι $y = \sum_{i=1}^n t_i x_i$ όπου $t_i \geq 0$ με $\sum_{i=1}^n t_i = 1$ και $x_i \in A_1$. Τότε $(x_i, 1) \in A$ και

$$\sum_{i=1}^n t_i (x_i, 1) = \left(\sum_{i=1}^n t_i x_i, \ \sum_{i=1}^n t_i \right) = (y, 1).$$

Δηλαδή δείξαμε ότι $(y, 1) \in conv(A)$. Αφού $\|x - y\|_2 = \|(x, 1) - (y, 1)\|_2$ παίρνουμε ότι

$$\begin{aligned} \Phi_A((x, 1)) &= \min\{\|(x, 1) - (y, 1)\|_2 : \ (y, 1) \in conv(A), \ A \subseteq E_2^{n+1}\} \\ &\leq \|(x, 1) - (y, 1)\|_2 = \|x - y\|_2 = \Phi_{A_1}(x), \end{aligned}$$

δηλαδή το ζητούμενο.

□

Λήμμα 3.3. *Για κάθε $x \in E_2^n$ και για κάθε $0 \leq a \leq 1$, ισχύει*

$$\Phi_A^2((x, -1)) \leq 4a^2 + a\Phi_{A_1}^2(x) + (1 - a)\Phi_{A_{-1}}^2(x).$$

Απόδειξη: Έστω $z_i \in conv(A_i)$ με $i = \pm 1$. Όπως και προηγουμένως δείχνουμε ότι $(z_i, i) \in conv(A)$. Αφού το $conv(A)$ είναι κυρτό ισχύει

$$z := a(z_1, 1) + (1-a)(z_{-1}, -1) = (az_1 + (1-a)z_{-1}, 2a-1) \in conv(A).$$

Έχουμε ότι,

$$\begin{aligned}
\|(x,-1) - z\|_2^2 &= \|(x - az_1 - (1-a)z_{-1}, -2a)\|_2^2 \\
&= \|(x - az_1 - (1-a)z_{-1}, 0)\|_2^2 + \|(0, -2a)\|_2^2 \\
&= \|x - ax - az_1 + ax - (1-a)z_{-1}\|_2^2 + 4a^2 \\
&= \|a(x - z_1) + (1-a)(x - z_{-1})\|_2^2 + 4a^2 \\
&\leq (a\|x - z_1\|_2 + (1-a)\|x - z_{-1}\|_2)^2 + 4a^2 \\
&\leq a\|x - z_1\|_2^2 + (1-a)\|x - z_{-1}\|_2^2 + 4a^2.
\end{aligned}$$

Η τελευταία ανισότητα ισχύει διότι, θέτοντας $1 - a = b$ $(a + b = 1)$ και $\|x - z_1\|_2 = x_0,\ \|x - z_{-1}\|_2 = y_0$ αρκεί να δείξουμε ότι

$$\begin{aligned}
& (ax_0 + by_0)^2 \leq ax_0^2 + by_0^2 \\
\Leftrightarrow\ & a^2x_0^2 + b^2y_0^2 + 2abx_0y_0 \leq ax_0^2 + by_0^2 \\
\Leftrightarrow\ & 2abx_0y_0 \leq ax_0^2(1-a) + by_0^2(1-b) \\
\Leftrightarrow\ & 2a(1-a)x_0y_0 \leq a(1-a)x_0^2 + a(1-a)y_0^2 \\
\Leftrightarrow\ & 2a(1-a)x_0y_0 \leq a(1-a)(x_0^2 + y_0^2).
\end{aligned}$$

Αν $a = 0$ ισχύει τετριμμένα, διαφορετικά αρκεί να δείξουμε ότι

$$2x_0y_0 \leq x_0^2 + y_0^2 \Leftrightarrow (x_0 - y_0)^2 \geq 0$$

που ισχύει. Τώρα που τα $z_i \in conv(A_i)$ είναι τυχαία έχουμε

$$\begin{aligned}
\Phi_A^2((x,-1)) &\leq \|(x,-1) - z\|_2^2 \\
&\leq a\|x - z_1\|_2^2 + (1-a)\|x - z_{-1}\|_2^2 + 4a^2 \\
&\leq a\Phi_{A_1}(x) + (1-a)\Phi_{A_{-1}}(x) + 4a^2,
\end{aligned}$$

δηλαδή το ζητούμενο.

□

Επιστρέφουμε τώρα την απόδειξη του θεωρήματος 3.3. Από τα λήμματα 3.2 και 3.3 παίρνουμε ότι

$$\begin{aligned}
\mathbb{E}(e^{\Phi_A^2(x)/8}) &= \frac{1}{2^{n+1}} \sum_{x\in E_2^{n+1}} e^{\Phi_A^2(x)/8} \\
&= \frac{1}{2^{n+1}} \sum_{x\in E_2^n} e^{\Phi_A^2((x,1))/8} + \frac{1}{2^{n+1}} \sum_{x\in E_2^n} e^{\Phi_A^2((x,-1))/8} \\
&\leq \frac{1}{2^{n+1}} \sum_{x\in E_2^n} e^{\Phi_{A_1}^2(x)/8} + \frac{1}{2^{n+1}} \sum_{x\in E_2^n} e^{(a\Phi_{A_1}^2(x)+(1-a)\Phi_{A_{-1}}^2(x)+4a^2)/8} \\
&= \frac{1}{2^{n+1}} \sum_{x\in E_2^n} e^{\Phi_{A_1}^2(x)/8} + \frac{1}{2^{n+1}} e^{a^2/2} \sum_{x\in E_2^n} e^{(a\Phi_{A_1}^2(x)+(1-a)\Phi_{A_{-1}}^2(x))/8}.
\end{aligned}$$

Από την ανισότητα Hölder παίρνουμε ότι

$$\begin{aligned}
\mathbb{E}(e^{\Phi_A^2(x)/8}) &\leq \frac{1}{2^{n+1}} \sum_{x\in E_2^n} e^{\Phi_{A_1}^2(x)/8} \\
&\quad + \frac{1}{2^{n+1}} e^{a^2/2} \Bigg(\sum_{x\in E_2^n} e^{\Phi_{A_1}^2(x)/8} \Bigg)^a \Bigg(\sum_{x\in E_2^n} e^{\Phi_{A_{-1}}^2(x)/8} \Bigg)^{1-a} \\
&= \frac{1}{2}\mathbb{E}(e^{\Phi_{A_1}^2(x)/8}) + \frac{1}{2} e^{a^2/2} \Big(\mathbb{E}(e^{\Phi_{A_1}^2(x)/8})\Big)^a \Big(\mathbb{E}(e^{\Phi_{A_{-1}}^2(x)/8})\Big)^{1-a}.
\end{aligned}$$

Θέτουμε

$$u_1 = \mathbb{E}(e^{\Phi_{A_1}^2(x)/8}), \quad v_1 = \frac{1}{\mu_n(A_1)}$$

και

$$u_{-1} = \mathbb{E}(e^{\Phi_{A_{-1}}^2(x)/8}), \quad v_{-1} = \frac{1}{\mu_n(A_{-1})}$$

Από την επαγωγική υπόθεση έχουμε $u_1 \leq v_1$ και $u_{-1} \leq v_{-1}$. Επίσης επειδή υποθέσαμε ότι $|A_{-1}| \leq |A_1|$ ισχύει $v_1 \leq v_{-1}$. Αντικαθιστώντας στην προηγούμενη ανισότητα παίρνουμε

$$\begin{aligned}
\mathbb{E}(e^{\Phi_A^2(x)/8}) &\leq \frac{1}{2}u_1 + \frac{1}{2}e^{a^2/2}(u_1)^a(u_{-1})^{1-a} \\
&\leq \frac{1}{2}v_1 + \frac{1}{2}e^{a^2/2}(v_1)^a(v_{-1})^{1-a} \\
&= \frac{1}{2}v_1\Big(1 + e^{a^2/2}\Big(\frac{v_1}{v_{-1}}\Big)^{a-1}\Big).
\end{aligned}$$

Θέτουμε

$$g(a) = \frac{1}{2}v_1\Big(1 + e^{a^2/2}\Big(\frac{v_1}{v_{-1}}\Big)^{a-1}\Big).$$

Θα ελαχιστοποιήσουμε τη συνάρτηση $g(a)$. Παραγωγίζοντας παίρνουμε

$$g'(a) = \frac{1}{2} v_1 e^{a^2/2} \Big(\frac{v_1}{v_{-1}} \Big)^{a-1} \Big(a + \ln \frac{v_1}{v_{-1}} \Big).$$

Οπότε

$$g'(a) = 0 \Leftrightarrow a = -\ln \frac{v_1}{v_{-1}}.$$

Εύκολα ελέγχει κανείς ότι στο παραπάνω a η g έχει ελάχιστο. Επιλέγουμε

$$a_0 = 1 - \frac{v_1}{v_{-1}} \quad (\text{τιμή κοντά στην ελάχιστη}).$$

Επειδή $v_1 \leq v_{-1}$ έχουμε ότι $0 \leq a_0 \leq 1$. Αντικαθιστώντας το a_0 παίρνουμε

$$\mathbb{E}(e^{\Phi_A^2(x)/8}) \leq \frac{1}{2} v_1 \Big(1 + e^{a_0^2/2} (1-a_0)^{a_0-1} \Big). \tag{3.1}$$

Θα χρειαστούμε και το ακόλουθο λήμμα.

Λήμμα 3.4. *Για κάθε* $0 \leq a \leq 1$ *έχουμε*

$$1 + e^{a^2/2}(1-a)^{a-1} \leq \frac{4}{2-a}. \tag{3.2}$$

Απόδειξη: Η (3.2) με απλές πράξεις γράφεται ισοδύναμα

$$g(a) = \ln(2+a) - \ln(2-a) - a^2/2 - (a-1)\ln(1-a) \geq 0.$$

Παραγωγίζοντας βλέπουμε ότι $g'' > 0$ και $g'(0) = 0$. Άρα η g αύξουσα στο $[0,1]$ και επειδή $g(0) = 0$ έπεται το ζητούμενο.

□

Τώρα από το λήμμα 3.4 και τη σχέση (3.1) παίρνουμε

$$\begin{aligned} \mathbb{E}(e^{\Phi_A^2(x)/8}) &\leq \frac{v_1}{2} \frac{4}{2-a_0} = \frac{2v_1}{1 + v_1/v_{-1}} \\ &= \frac{2}{1/v_1 + 1/v_{-1}} = \frac{2}{\mu_n(A_1) + \mu_n(A_{-1})}. \end{aligned}$$

Όμως,

$$\mu_{n+1}(A_i \times \{i\}) = \frac{|A_i \times \{i\}|}{2^{n+1}} = \frac{1}{2} \frac{|A_i|}{2^n} = \frac{1}{2} \mu_n(A_i).$$

Επομένως,

$$\begin{aligned}
\mathbb{E}(e^{\Phi_A^2(x)/8}) &\leq \frac{2}{2\mu_{n+1}(A_1 \times \{1\}) + 2\mu_{n+1}(A_{-1} \times \{-1\})} \\
&\leq \frac{1}{\mu_{n+1}((A_1 \times \{1\}) \cup (A_{-1} \times \{-1\}))} \\
&= \frac{1}{\mu_{n+1}(A)}.
\end{aligned}$$

Αυτό ολοκληρώνει το επαγωγικό βήμα και την απόδειξη του θεωρήματος 3.3.

□

Κεφάλαιο 4

Αθροίσματα ανεξάρτητων τυχαίων μεταβλητών

Σε αυτό το κεφάλαιο θα ασχοληθούμε με ανισότητες για αθροίσματα ανεξάρτητων τυχαίων μεταβλητών. Σκοπός μας είναι να πετύχουμε άνω φράγματα για την πιθανότητα απόκλισης από την μέση τιμή. Δηλαδή, ανισότητες για την πιθανότητα

$$\mathbb{P}(|S_n - \mathbb{E}(S_n)| \geq t) \tag{4.1}$$

όπου $S_n = X_1 + \ldots + X_n$ και οι X_i είναι ανεξάρτητες τυχαίες μεταβλητές για κάθε $i = 1, \ldots, n$.

Παρατήρηση: Αν $X_1, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές τότε

$$\mathbb{E}(X_1 \cdots X_n) = \mathbb{E}(X_1) \cdots \mathbb{E}(X_n)$$

και

$$Var(X_1 + \ldots + X_n) = Var(X_1) + \ldots + Var(X_n).$$

Από την ανεξαρτησία και την ανισότητα *Chebyshev* μπορούμε να πάρουμε ένα πρώτο άνω φράγμα για την (4.1)

$$\mathbb{P}(|S_n - \mathbb{E}(S_n)| \geq t) \leq \frac{Var(S_n)}{t^2} = \frac{\sum_{i=1}^{n} Var(X_i)}{t^2}.$$

Αν θέσουμε

$$\sigma^2 := \frac{1}{n} \sum_{i=1}^{n} Var(X_i)$$

(το σ^2 είναι η διασπορά της τυχαίας μεταβλητής S_n) παίρνουμε

$$\mathbb{P}\left(\left| \sum_{i=1}^{n} X_i - \mathbb{E}\left(\sum_{i=1}^{n} X_i \right) \right| \geq t \right) \leq \frac{n\sigma^2}{t^2}$$

και κάνοντας την αλλαγή μεταβλητής θέτοντας tn όπου t, προκύπτει ότι

$$\mathbb{P}\left(\frac{1}{n}\left|\sum_{i=1}^{n}(X_i - \mathbb{E}X_i)\right| \geq t\right) \leq \frac{\sigma^2}{nt^2}. \tag{4.2}$$

Το ερώτημα που τίθεται είναι αν μπορούμε να περιμένουμε καλύτερα άνω φράγματα από αυτό της (4.2). Μια απάντηση δίνεται από το **κεντρικό οριακό θεώρημα** το οποίο υπό κάποιες ισχυρές προϋποθέσεις μας δίνει ότι

$$\mathbb{P}\left(\frac{\sqrt{n}}{\sigma}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}X_i)\right) \geq y\right) \leq \frac{1}{\sqrt{2\pi}}\frac{1}{y}e^{-y^2/2}.$$

Για να συγκρίνουμε με την (4.2) κάνουμε την αλλαγή μεταβλητής θέτοντας $t\sqrt{n}/y$ όπου t και παίρνουμε ότι

$$\mathbb{P}\left(\frac{1}{n}\left|\sum_{i=1}^{n}(X_i - \mathbb{E}X_i)\right| \geq t\right) \leq e^{-nt^2/2\sigma^2}$$

το οποίο είναι καλύτερο άνω φράγμα από αυτό της ανισότητας (4.2).

Στη συνέχεια αυτού του κεφαλαίου θα δούμε ανισότητες που δίνουν καλύτερα άνω φράγματα από αυτό της ανισότητας *Chebyshev*.

4.1 Η ανισότητα του Hoeffding

Λήμμα 4.1. *(Ανισότητα Chernoff) Έστω $s > 0$, τότε για κάθε τυχαία μεταβλητή X και για κάθε $t > 0$ έχουμε*

$$\mathbb{P}(X \geq t) \leq e^{-st}\mathbb{E}(e^{sX}).$$

Απόδειξη: Η απόδειξη είναι άμεση συνέπεια της ανισότητας *Markov*. Αφού $s > 0$ έχουμε

$$X \geq t \Leftrightarrow sX \geq st \Leftrightarrow e^{sX} \geq e^{st}.$$

Άρα,

$$\mathbb{P}(X \geq t) = \mathbb{P}(e^{sX} \geq e^{st}) \leq \frac{\mathbb{E}(e^{sX})}{e^{st}} = e^{-st}\mathbb{E}(e^{sX}).$$

□

Τώρα, χρησιμοποιώντας την ανεξαρτησία των X_i και εφαρμόζοντας την ανισότητα *Chernoff* για την τυχαία μεταβλητή $S_n - \mathbb{E}(S_n)$ παίρνουμε

$$\begin{aligned}
\mathbb{P}(S_n - \mathbb{E}S_n \geq t) &\leq e^{-st}\mathbb{E}\big(e^{s(S_n-\mathbb{E}S_n)}\big) \\
&= e^{-st}\mathbb{E}\big(e^{s\sum_{i=1}^n (X_i-\mathbb{E}X_i)}\big) \\
&= e^{-st}\mathbb{E}\big(e^{s(X_1-\mathbb{E}X_1)}\cdots e^{s(X_n-\mathbb{E}X_n)}\big) \\
&= e^{-st}\mathbb{E}\big(e^{s(X_1-\mathbb{E}X_1)}\big)\cdots\mathbb{E}\big(e^{s(X_n-\mathbb{E}X_n)}\big) \\
&= e^{-st}\prod_{i=1}^n \mathbb{E}\big(e^{s(X_i-\mathbb{E}X_i)}\big).
\end{aligned}$$

Δηλαδή, δείξαμε ότι

$$\mathbb{P}(S_n - \mathbb{E}S_n \geq t) \leq e^{-st}\prod_{i=1}^n \mathbb{E}\big(e^{s(X_i-\mathbb{E}X_i)}\big). \tag{4.3}$$

Αυτό που καταφέραμε είναι να ανάγουμε το πρόβλημα στο να βρεθούν άνω φράγματα για τις ροπογεννήτριες των τυχαίων μεταβλητών $X_i - \mathbb{E}X_i$, δηλαδή για την ποσότητα $\mathbb{E}(e^{s(X_i-\mathbb{E}X_i)})$.

Λήμμα 4.2. *Έστω X τυχαία μεταβλητή με $a \leq X \leq b$ και $\mathbb{E}X = 0$. Τότε για κάθε $s > 0$,*

$$\mathbb{E}(e^{sX}) \leq e^{s^2(b-a)^2/8}.$$

Απόδειξη: Χρησιμοποιώντας το γεγονός ότι η εκθετική συνάρτηση είναι κυρτή, για κάθε $a \leq x \leq b$ έχουμε

$$x = \frac{b-x}{b-a}a + \frac{x-a}{b-a}b \Leftrightarrow e^{sx} = e^{\frac{b-x}{b-a}sa + \frac{x-a}{b-a}sb}.$$

Από την κυρτότητα της εκθετικής έχουμε

$$e^{sx} \leq \frac{b-x}{b-a}e^{sa} + \frac{x-a}{b-a}e^{sb}.$$

Άρα

$$\begin{aligned}
\mathbb{E}(e^{sX}) &\leq \mathbb{E}\Big(\frac{b-x}{b-a}e^{sa} + \frac{x-a}{b-a}e^{sb}\Big) \\
&= \frac{b-\mathbb{E}X}{b-a}e^{sa} + \frac{\mathbb{E}X-a}{b-a}e^{sb} \\
&= \frac{b}{b-a}e^{sa} - \frac{a}{b-a}e^{sb}.
\end{aligned}$$

Παρατηρούμε ότι αφού $a \leq X \leq b$ και $\mathbb{E}X = 0$ τότε $a \leq 0$ και $b \geq 0$. Θέτουμε

$$\rho = -\frac{a}{b-a} \in [0,1], \quad 1-\rho = \frac{b}{b-a}.$$

Άρα,

$$\begin{aligned}
\mathbb{E}(e^{sX}) &\leq (1-\rho)e^{-\rho s(b-a)} + \rho e^{(1-\rho)s(b-a)} \\
&= (1-\rho+\rho e^{s(b-a)})e^{-\rho s(b-a)}.
\end{aligned}$$

Τώρα θέτουμε εκ νέου $u = s(b-a)$. Άρα,

$$\begin{aligned}
\mathbb{E}(e^{sX}) &\leq (1-\rho+\rho e^{u})e^{-\rho u} \\
&= e^{\ln(1-\rho+\rho e^{u})-\rho u} \\
&= e^{g(u)}
\end{aligned}$$

όπου $g(u) = \ln(1-\rho+\rho e^u) - \rho u$. Τώρα αρκεί να δείξουμε ότι $g(u) \leq u^2/8$, γιατί τότε

$$\mathbb{E}(e^{sX}) \leq e^{u^2/8} = e^{s^2(b-a)^2/8}.$$

Από το θεώρημα του *Taylor* υπάρχει $\theta \in [0,u]$ τέτοιο ώστε

$$g(u) = g(0) + g'(0)u + \frac{u^2}{2}g''(\theta).$$

Υπολογίζοντας βρίσκουμε ότι $g(0) = g'(0) = 0$, άρα

$$g(u) = \frac{u^2}{2}g''(\theta).$$

Επομένως αρκεί να δείξουμε ότι για κάθε $u \geq 0$,

$$g''(u) = \frac{\rho(1-\rho)e^{-u}}{((1-\rho)e^{-u}+\rho)^2} \leq \frac{1}{4}. \tag{4.4}$$

Θέτουμε $\delta = (1-\rho)e^{-u}$, οπότε η (4.4) γίνεται ισοδύναμα

$$\frac{\rho\delta}{(\rho+\delta)^2} \leq \frac{1}{4} \Leftrightarrow 4\rho\delta \leq (\rho+\delta)^2 \Leftrightarrow (\rho-\delta)^2 \geq 0$$

που ισχύει.

$\square$

Θεώρημα 4.1. *(Ανισότητα Hoeffding) Έστω $X_1, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές με $\mathbb{P}(X_i \in [a_i, b_i]) = 1$. Τότε για κάθε $t > 0$ ισχύουν οι ανισότητες*

$$\mathbb{P}(S_n - \mathbb{E}S_n \geq t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right) \tag{4.5}$$

και

$$\mathbb{P}(S_n - \mathbb{E}S_n \leq -t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right). \tag{4.6}$$

Απόδειξη: Χωρίς περιορισμό της γενικότητας μπορούμε να υποθέσουμε ότι $\mathbb{E}X_i = 0$ για κάθε $i = 1, \ldots, n$. Από την ανισότητα (4.3) και το λήμμα 4.2 έχουμε

$$\begin{aligned}
\mathbb{P}(S_n - \mathbb{E}S_n \geq t) &\leq e^{-st} \prod_{i=1}^n \mathbb{E}\left(e^{s(X_i - \mathbb{E}X_i)}\right) \\
&\leq e^{-st} \prod_{i=1}^n e^{s^2(b_i-a_i)^2/8} \\
&= e^{-st} e^{s^2/8 \sum_{i=1}^n (b_i-a_i)^2} \\
&= e^{-st+s^2/8 \sum_{i=1}^n (b_i-a_i)^2} \\
&= e^{-g(s)}
\end{aligned}$$

όπου

$$g(s) = st - \frac{s^2}{8} \sum_{i=1}^n (b_i - a_i)^2.$$

Ελαχιστοποιώντας την $g(s)$ (επιλέγοντας $s = 4t / \sum_{i=1}^n (b_i - a_i)^2 > 0$) παίρνουμε την (4.5). Εφαρμόζοντας την ανισότητα (4.5) για τις τυχαίες μεταβλητές $-X_1, \ldots, -X_n$ προκύπτει η (4.6) και αυτό ολοκληρώνει την απόδειξη του θεωρήματος.

$\square$

4.2 Η ανισότητα του Bernstein

Θεώρημα 4.2. *(Ανισότητα Bernstein) Έστω $X_1, \dots X_n$ ανεξάρτητες τυχαίες μεταβλητές με $\mathbb{E}X_i = 0$ για κάθε $i = 1, \dots, n$. Επιπλέον έστω ότι οι X_i είναι φραγμένες, δηλαδή υπάρχει σταθερά $M > 0$ τέτοια ώστε $\mathbb{P}(|X_i| \leq M) = 1$. Αν*

$$\sigma^2 := \frac{1}{n}\sum_{i=1}^n Var(X_i),$$

τότε για κάθε $t > 0$

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n X_i > t\right) \leq \exp\Big(-\frac{nt^2}{2\sigma^2 + 2Mt/3}\Big). \tag{4.7}$$

Αν εξαιρέσουμε τον όρο $2Mt/3$ η ανισότητα του *Bernstein* μας δίνει την ίδια εκτίμηση με αυτή του κεντρικού οριακού θεωρήματος.

Λήμμα 4.3. *(Ανισότητα Bennet) Έστω $X_1, \dots X_n$ ανεξάρτητες τυχαίες μεταβλητές με $\mathbb{E}X_i = 0$ για κάθε $i = 1, \dots, n$ και $\mathbb{P}(|X_i| \leq M) = 1$ για κάποια σταθερά $M > 0$. Αν*

$$\sigma^2 := \frac{1}{n}\sum_{i=1}^n Var(X_i),$$

τότε για κάθε $t > 0$

$$\mathbb{P}\left(\sum_{i=1}^n X_i > t\right) \leq \exp\left(-\frac{n\sigma^2}{M^2} h\left(\frac{Mt}{n\sigma^2}\right)\right) \tag{4.8}$$

όπου $h(u) = (1+u)\ln(1+u) - u$ για $u \geq 0$.

Απόδειξη: Έστω $s > 0$, από την ανισότητα *Chernoff* έχουμε ότι

$$\mathbb{P}(S_n \geq t) \leq e^{-st}\prod_{i=1}^n \mathbb{E}(e^{sX_i}). \tag{4.9}$$

Θα αναζητήσουμε άνω φράγμα για την $\mathbb{E}(e^{sX_i})$. Θέτουμε

$$\sigma_i^2 := Var(X_i) = \mathbb{E}(X_i^2)$$

και

$$F_i := \sum_{k=2}^\infty \frac{s^{k-2}\mathbb{E}(X_i^k)}{k!\sigma_i^2}.$$

Αναπτύσσουμε την εκθετική ως δυναμοσειρά με κέντρο το μηδέν και παίρνουμε

$$e^{sx} = \sum_{k=0}^{\infty} \frac{s^k x^k}{k!} = 1 + sx + \sum_{k=2}^{\infty} \frac{s^k x^k}{k!}.$$

Άρα,

$$\begin{aligned} \mathbb{E}(e^{sX_i}) &= \mathbb{E}\left(1 + sX_i + \sum_{k=2}^{\infty} \frac{s^k X_i^k}{k!}\right) \\ &= 1 + s\mathbb{E}(X_i) + \mathbb{E}\left(\sum_{k=2}^{\infty} \frac{s^k X_i^k}{k!}\right) \\ &= 1 + \sum_{k=2}^{\infty} \frac{s^k \mathbb{E}(X_i^k)}{k!} \\ &= 1 + s^2\sigma_i^2 \sum_{k=2}^{\infty} \frac{s^{k-2}\mathbb{E}(X_i^k)}{k!\sigma_i^2} \\ &= 1 + s^2\sigma_i^2 F_i. \end{aligned}$$

Για κάθε y ισχύει ότι $1 + y \leq e^y$, συνεπώς

$$\mathbb{E}(e^{sX_i}) = 1 + s^2\sigma_i^2 F_i \leq e^{s^2\sigma_i^2 F_i}. \tag{4.10}$$

Τώρα, για κάθε $k \geq 2$ έχουμε

$$\begin{aligned} \mathbb{E}(X_i^k) &\leq \mathbb{E}(|X_i|^k) = \mathbb{E}(X_i^2|X_i|^{k-2}) \\ &\leq \mathbb{E}(X_i^2 M^{k-2}) = M^{k-2}\sigma_i^2. \end{aligned}$$

Άρα,

$$\begin{aligned} F_i &= \sum_{k=2}^{\infty} \frac{s^{k-2}\mathbb{E}(X_i^k)}{k!\sigma_i^2} \\ &\leq \sum_{k=2}^{\infty} \frac{s^{k-2}M^{k-2}\sigma_i^2}{k!\sigma_i^2} \\ &= \frac{1}{M^2 s^2} \sum_{k=2}^{\infty} \frac{s^{k-2}s^2 M^{k-2}M^2}{k!} \\ &= \frac{1}{(Ms)^2} \sum_{k=2}^{\infty} \frac{(Ms)^k}{k!}. \end{aligned}$$

Από το ανάπτυγμα *Taylor* της εκθετικής έχουμε

$$e^{Ms} = 1 + Ms + \sum_{k=2}^{\infty} \frac{(Ms)^k}{k!} \Rightarrow \sum_{k=2}^{\infty} \frac{(Ms)^k}{k!} = e^{Ms} - 1 - Ms.$$

Συνεπώς,

$$F_i \leq \frac{e^{Ms} - 1 - Ms}{(Ms)^2}. \tag{4.11}$$

Συνδυάζοντας με την (4.10) παίρνουμε

$$\mathbb{E}(e^{sX_i}) \leq e^{\sigma_i^2 \frac{e^{Ms} - 1 - Ms}{M^2}}. \tag{4.12}$$

Σε συνδυασμό με την (4.9) και θέτοντας $\sum_{i=1}^{n} \sigma_i^2 = n\sigma^2$ παίρνουμε

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{n} X_i \geq t\right) &\leq e^{-st} \prod_{i=1}^{n} e^{\sigma_i^2 \frac{e^{Ms} - 1 - Ms}{M^2}} \\
&= e^{-st} e^{\sum_{i=1}^{n} (\sigma_i^2) \frac{e^{Ms} - 1 - Ms}{M^2}} \\
&= e^{\frac{n\sigma^2 (e^{Ms} - 1 - Ms)}{M^2} - st} \\
&= e^{g(s)}.
\end{aligned}$$

Το άνω φράγμα ελαχιστοποιείται αν επιλέξουμε

$$s := \frac{1}{M} \ln\left(1 + \frac{tM}{n\sigma^2}\right) > 0.$$

Για αυτή την τιμή του s με απλές πράξεις προκύπτει η (4.8).

□

Απόδειξη:(θεωρήματος 4.2) Αν υποθέσουμε ότι

$$h(u) \geq \frac{u^2}{2 + 2u/3}$$

για κάθε $u \geq 0$, τότε συνδυάζοντας με την (4.8) παίρνουμε

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{n} X_i > t\right) &\leq \exp\left(-\frac{n\sigma^2}{M^2} \frac{M^2 t^2}{n^2 \sigma^4 (2 + 2Mt/3n\sigma^2)}\right) \\
&= \exp\left(-\frac{t^2}{n\sigma^2 (2 + 2Mt/3n\sigma^2)}\right) \\
&= \exp\left(-\frac{t^2}{2n\sigma^2 + 2Mt/3}\right).
\end{aligned}$$

Πραγματοποιώντας την αλλαγή μεταβλητής θέτοντας tn όπου t, παίρνουμε

$$\begin{aligned}\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} X_i > t\right) &\leq \exp\left(-\frac{t^2n^2}{2n\sigma^2 + 2Mtn/3}\right)\\ &= \exp\left(-\frac{nt^2}{2\sigma^2 + 2Mt/3}\right).\end{aligned}$$

Για να ολοκληρώσουμε την απόδειξη του θεωρήματος 4.2 μένει να δείξουμε ότι για κάθε $u \geq 0$

$$h(u) = (1+u)\ln(1+u) - u \geq \frac{u^2}{2 + 2u/3}. \tag{4.13}$$

Πράγματι, η (4.13) γράφεται ισοδύναμα

$$\ln(1+u) \geq \frac{5u^2 + 6u}{2u^2 + 8u + 6}.$$

Θεωρώ την συνάρτηση

$$f(u) = \ln(1+u) - \frac{5u^2 + 6u}{2u^2 + 8u + 6}.$$

Εύκολα υπολογίζει κανείς ότι η παράγωγος της f έχει μία πραγματική ρίζα στο $u = 0$. Άρα το $f(0) = 0$ είναι είτε ελάχιστο είτε μέγιστο. Μέγιστο δεν είναι διότι διαφορετικά θα έπρεπε

$$\ln(1+u) \leq \frac{5u^2 + 6u}{2u^2 + 8u + 6}$$

άτοπο αφού το δεξιό μέλος είναι άνω φραγμένο (έχει στο άπειρο όριο 5/2 και ο παρανομαστής δεν μηδενίζεται για $u \geq 0$). Άρα είναι ελάχιστο, οπότε

$$f(u) \geq f(0) \Leftrightarrow \ln(1+u) \geq \frac{5u^2 + 6u}{2u^2 + 8u + 6} \Leftrightarrow h(u) \geq \frac{u^2}{2 + 2u/3}.$$

□

Κεφάλαιο 5

Η μέθοδος των martingales

5.1 Εισαγωγή

Σε αυτό το κεφάλαιο θα ορίσουμε τη δεσμευμένη μέση τιμή και σε συνδυασμό με τη μέθοδο των martingales θα δούμε έναν ακόμα τρόπο εκτίμησης της πιθανότητας απόκλισης μιας φραγμένης τυχαίας μεταβλητής από τη μέση τιμή της.

Θα ξεκινήσουμε με κάποιες έννοιες από τη θεωρία μέτρου. Έστω $(\Omega, \mathcal{A})$ ένας μετρήσιμος χώρος, όπου $\mathcal{A}$ σ-άλγεβρα των υποσυνόλων του Ω. Τα στοιχεία της $\mathcal{A}$ λέγονται **μετρήσιμα** υποσύνολα του Ω. Ένα μέτρο ν λέγεται **απόλυτα συνεχές** ως προς το μέτρο μ αν για κάθε $A \in \mathcal{A}$ με $\mu(A) = 0 \Rightarrow \nu(A) = 0$.

Θεώρημα 5.1. *(Radon-Nikodym) Έστω $(\Omega, \mathcal{A}, \mu)$ μετρήσιμος χώρος με μ σ-πεπερασμένο. Έστω ν μέτρο απόλυτα συνεχές ως προς το μ. Τότε υπάρχει μοναδική μετρήσιμη μη αρνητική συνάρτηση f, τέτοια ώστε για κάθε $A \in \mathcal{A}$ να ισχύει*

$$\nu(A) = \int_A f d\mu. \tag{5.1}$$

Στη συνέχεια θα ορίσουμε τους $\mathbf{L^p}$ **χώρους** $1 \leq p \leq \infty$, πάνω σε ένα χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$.

$$L^p(\Omega, \mathcal{A}, \mathbb{P}) = \left\{ f : \Omega \to \mathbb{R} : \ f \text{ μετρήσιμη}, \ \int_\Omega |f|^p d\mathbb{P} < \infty \right\} \Big/ \sim$$

όπου η σχέση ισοδυναμίας $\sim$ ορίζεται ως εξής: $f \sim g \Leftrightarrow f - g = 0$ σχεδόν παντού. Τα στοιχεία του $L^p(\Omega, \mathcal{A}, \mathbb{P})$ παρόλο που είναι κλάσεις ισοδυναμίας συνεχίζουμε να τα γράφουμε ως συναρτήσεις, επιλέγοντας έναν αντιπρόσωπο της κάθε κλάσης. Αυτό δεν δημιουργεί πρόβλημα στην θεωρία αφού οι διαφορές σε σύνολα μέτρου μηδέν δεν επηρεάζουν το ολοκλήρωμα, με το οποίο ορίζεται η νόρμα του χώρου ως εξής: η **p-νόρμα** της f είναι η ποσότητα

$$\|f\|_p = \left(\int_\Omega |f|^p d\mathbb{P} \right)^{1/p}.$$

Ομοίως ο $L^\infty(\Omega, \mathcal{A}, \mathbb{P})$ είναι ο

$$L^\infty(\Omega, \mathcal{A}, \mathbb{P}) = \{ f : \Omega \to \mathbb{R} : \ f \text{ μετρήσιμη}, \operatorname{ess\,sup} f < \infty \}$$

όπου

$$\begin{aligned} \operatorname{ess\,sup} f = \|f\|_\infty &= \inf\{M > 0 : \mathbb{P}(t : \ f(t) \geq M) = 0\} \\ &= \inf\{M > 0 : \mathbb{P}(t : \ f(t) < M) = 1\} \end{aligned}$$

δηλαδή είναι ο χώρος των σχεδόν παντού φραγμένων συναρτήσεων.

Έστω τώρα $f \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ μη αρνητική μετρήσιμη συνάρτηση και $\mathcal{G}$ μια σ-υποάλγεβρα της $\mathcal{A}$, τότε η συναλοσυνάρτηση

$$\mu(A) = \int_A f d\mathbb{P}$$

ορίζει ένα μέτρο στην $\mathcal{G}$ για κάθε $A \in \mathcal{G}$ το οποίο είναι απόλυτα συνεχές ως προς το $\mathbb{P}|_\mathcal{G}$. Από το θεώρημα *Radon-Nikodym* υπάρχει μοναδική συνάρτηση $h \in L^1(\Omega, \mathcal{G}, \mathbb{P})$ με την ιδιότητα

$$\int_A h d\mathbb{P} = \int_A f d\mathbb{P} \tag{5.2}$$

για κάθε $A \in \mathcal{G}$. Η h ονομάζεται **δεσμευμένη μέση τιμή** της f ως προς την $\mathcal{G}$ και συμβολίζεται με $\mathbb{E}(f|\ \mathcal{G})$.

Παράδειγμα: Έστω $\Omega = [0,1]$, $\mathcal{A}$ τα *Lebesgue* μετρήσιμα υποσύνολα του $[0,1]$ και $\mathbb{P}$ το μέτρο *Lebesgue*. Ο χώρος $(\Omega, \mathcal{A}, \mathbb{P})$ είναι ένας χώρος πιθανότητας. Έστω $f \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ μη αρνητική συνάρτηση. θέτουμε

$$A_i = \left[\frac{i}{10}, \frac{i+1}{10} \right]$$

για $i = 0, \ldots, 9$ και σχηματίζουμε την σ-άλγεβρα που παράγεται από τα A_i

$$\mathcal{G} = \left\{ \bigcup_{i \in I \subseteq \{0,\ldots,9\}} A_i, \ \left(\bigcup_{i \in I} A_i \right)^c, \ \emptyset, \ [0,1] \right\}.$$

Η $\mathcal{G}$ είναι μία σ-υποάλγεβρα της $\mathcal{A}$. Από το θεώρημα *Radon-Nikodym* υπάρχει μοναδική συνάρτηση $h \in L^1(\Omega, \mathcal{G}, \mathbb{P})$ με την ιδιότητα

$$\int_A h d\mathbb{P} = \int_A f d\mathbb{P}$$

για κάθε $A \in \mathcal{G}$. Άρα επειδή $A_0 \in \mathcal{G}$ πρέπει να ισχύει και

$$\int_{A_0} h d\mathbb{P} = \int_{A_0} f d\mathbb{P}.$$

Επομένως το εμβαδόν κάτω από το γράφημα της f στο A_0 πρέπει να είναι ίσο με το εμβαδόν που περικλείεται από το γράφημα της h και το A_0. Επιπλέον η μετρησιμότητα της h ως προς την $\mathcal{G}$ επιβάλει στην h να είναι σταθερή συνάρτηση, διότι διαφορετικά θα μπορούσαμε να βρούμε ένα σύνολο (a, b) τέτοιο ώστε $h^{-1}((a, b)) \notin \mathcal{G}$. Το ίδιο ισχύει και για όλα τα A_i. Η $h = \mathbb{E}(f|\, \mathcal{G})$ είναι η δεσμευμένη μέση τιμή της f ως προς την $\mathcal{G}$. Είναι φανερό ότι

$$\mathbb{E}(f) = \mathbb{E}(\mathbb{E}(f|\, \mathcal{G})) = \mathbb{E}(h)$$

διότι

$$\mathbb{E}(h) = \int_0^1 h = \sum_{i=0}^{9} \int_{A_i} h = \sum_{i=0}^{9} \int_{A_i} f = \int_0^1 f = \mathbb{E}(f).$$

Μερικές βασικές ιδιότητες της δεσμευμένης μέσης τιμής είναι οι εξής:

(i) Ο τελεστής $f \mapsto \mathbb{E}(f|\, \mathcal{G})$ είναι θετικός, γραμμικός και έχει νόρμα 1 σε κάθε L^p χώρο, $1 \leq p \leq \infty$.

(ii) Αν $\mathcal{G}_1$ μια σ-υποάλγεβρα της $\mathcal{G}$, τότε

$$\mathbb{E}(\mathbb{E}(f|\, \mathcal{G})|\, \mathcal{G}_1) = \mathbb{E}(f|\, \mathcal{G}_1). \tag{5.3}$$

(iii) Αν $g \in L^\infty(\Omega, \mathcal{G}, \mathbb{P})$ τότε

$$\mathbb{E}(f \cdot g|\, \mathcal{G}) = g \cdot \mathbb{E}(f|\, \mathcal{G}). \tag{5.4}$$

(iv) Αν $\mathcal{G} = \{\emptyset, \Omega\}$ είναι η τετριμμένη σ-άλγεβρα, τότε η $\mathbb{E}(f|\, \mathcal{G})$ είναι σταθερή και ισούται με τη μέση τιμή της f

$$\mathbb{E}(f|\, \mathcal{G}) = \mathbb{E}(f) = \int_\Omega f d\mathbb{P}. \tag{5.5}$$

Ορισμός 5.1. *Εστω $\mathcal{A}_0 \subseteq \mathcal{A}_1 \subseteq \cdots \subseteq \mathcal{A}$ μια αύξουσα ακολουθία σ-αλγεβρών. Μια ακολουθία συναρτήσεων $f_0, f_1, \ldots$ με $f_i \in L^1(\Omega, \mathcal{A}_i, \mathbb{P})$ λέγεται **martingale** ως προς την $\{A_i\}$, αν $\mathbb{E}(f_i|\, \mathcal{A}_{i-1}) = f_{i-1}$ για κάθε $i \geq 1$.*

5.2 Η ανισότητα του Azuma

Θεώρημα 5.2. *(Ανισότητα Azuma) Έστω* $f \in L^\infty(\Omega, \mathcal{A}, \mathbb{P})$ *και έστω* $\{\emptyset, \Omega\} = \mathcal{A}_0 \subseteq \mathcal{A}_1 \subseteq \cdots \subseteq \mathcal{A}_n = \mathcal{A}$ *μια αύξουσα ακολουθία σ-αλγεβρών. Για* $i = 1, \dots, n$ *θέτουμε*

$$d_i = \mathbb{E}(f|\mathcal{A}_i) - \mathbb{E}(f|\mathcal{A}_{i-1}).$$

Τότε για κάθε $t > 0$ *έχουμε*

$$\mathbb{P}(|f - \mathbb{E}f| \geq t) \leq e^{-t^2/4\sum_{i=1}^n \|d_i\|_\infty^2}. \tag{5.6}$$

Απόδειξη: Από την ιδιότητα (5.3) έχουμε ότι

$$\mathbb{E}(\mathbb{E}(f|\mathcal{A}_i)|\mathcal{A}_{i-1}) = \mathbb{E}(f|\mathcal{A}_{i-1})$$

και από τον ορισμό της δεσμευμένης μέσης τιμής $\mathbb{E}(f|\mathcal{A}_i) \in L^1(\Omega, \mathcal{A}_i, \mathbb{P})$. Άρα η ακολουθία $\{\mathbb{E}(f|\mathcal{A}_i)\}_{i=0}^n$ είναι martingale. Ισχύει ότι $\mathbb{E}(d_i|\mathcal{A}_{i-1}) = 0$. Πράγματι,

$$\begin{aligned}
\mathbb{E}(d_i|\mathcal{A}_{i-1}) &= \mathbb{E}[(\mathbb{E}(f|\mathcal{A}_i) - \mathbb{E}(f|\mathcal{A}_{i-1}))|\mathcal{A}_{i-1}] \\
&= \mathbb{E}(\mathbb{E}(f|\mathcal{A}_i)|\mathcal{A}_{i-1}) - \mathbb{E}(\mathbb{E}(f|\mathcal{A}_{i-1})|\mathcal{A}_{i-1}) \\
&\overset{(5.3)}{=} \mathbb{E}(f|\mathcal{A}_{i-1}) - \mathbb{E}(f|\mathcal{A}_{i-1}) = 0.
\end{aligned}$$

Ισχύει η ανισότητα

$$e^x \leq x + e^{x^2}. \tag{5.7}$$

Πράγματι, αν $x \geq 1$ είναι φανερή διότι

$$x \leq x^2 \Rightarrow e^x \leq e^{x^2} \leq e^{x^2} + x.$$

Έστω τώρα $x < 1$. Με τη βοήθεια του αναπτύγματος *Taylor* η (5.7) γράφεται

$$\sum_{n=0}^\infty \frac{x^n}{n!} \leq x + \sum_{n=0}^\infty \frac{x^{2n}}{n!}$$

η οποία είναι ισοδύναμη με την

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots \leq 1 + \frac{x^2}{1!} + \frac{x^4}{2!} + \frac{x^6}{3!} + \cdots \Leftrightarrow$$

$$\sum_{n=1}^\infty \frac{x^{2n}}{(2n)!} + \sum_{n=1}^\infty \frac{x^{2n+1}}{(2n+1)!} \leq \sum_{n=1}^\infty \frac{x^{2n}}{n!} \tag{5.8}$$

Για $x \leq 0$, έχουμε ότι

$$\sum_{n=1}^{\infty} \frac{x^{2n+1}}{(2n+1)!} \leq 0$$

και επειδή

$$\frac{x^{2n}}{(2n)!} \leq \frac{x^{2n}}{n!}$$

η (5.7) είναι σωστή. Μένει η περίπτωση $0 < x < 1$. Η (5.8) γράφεται ισοδύναμα

$$\sum_{n=1}^{\infty} \frac{x^{2n+1}}{(2n+1)!} \leq \sum_{n=1}^{\infty} x^{2n}\Big(\frac{1}{n!} - \frac{1}{(2n)!}\Big).$$

Αρκεί να δείξουμε ότι

$$\frac{1}{(2n+1)!} \leq \frac{1}{n!} - \frac{1}{(2n)!}$$

για κάθε $n \geq 1$. Ισοδύναμα ότι

$$1 \leq (2n+1)\Big[\frac{(2n)!}{n!} - 1\Big].$$

Το οποίο είναι προφανές αφού επαγωγικά προκύπτει άμεσα ότι

$$\frac{(2n)!}{n!} - 1 \geq 1.$$

Τώρα εφαρμόζουμε την (5.7) για $x = \lambda d_i$ και παίρνουμε

$$e^{\lambda d_i} \leq \lambda d_i + e^{\lambda^2 d_i^2}$$

για κάθε $\lambda \in \mathbb{R}$. Αφού ο τελεστής $f \mapsto \mathbb{E}(f|\mathcal{A})$ είναι γραμμικός έχουμε

$$\mathbb{E}(e^{\lambda d_i}|\mathcal{A}_{i-1}) \leq \mathbb{E}(\lambda d_i|\mathcal{A}_{i-1}) + \mathbb{E}(e^{\lambda^2 d_i^2}|\mathcal{A}_{i-1}).$$

Όμως

$$\mathbb{E}(\lambda d_i|\mathcal{A}_{i-1}) = \lambda \mathbb{E}(d_i|\mathcal{A}_{i-1}) = 0$$

και

$$\mathbb{E}(e^{\lambda^2 d_i^2}|\mathcal{A}_{i-1}) \leq \mathbb{E}(e^{\lambda^2 \|d_i\|_\infty^2}|\mathcal{A}_{i-1}) = e^{\lambda^2 \|d_i\|_\infty^2}.$$

Άρα,

$$\mathbb{E}(e^{\lambda d_i}|\mathcal{A}_{i-1}) \leq e^{\lambda^2 \|d_i\|_\infty^2}. \tag{5.9}$$

Παρατηρούμε ότι $\mathbb{E}(f|\mathcal{A}_i) \in L^\infty(\Omega, \mathcal{A}_i, \mathbb{P})$. Πράγματι, αφού $f \in L^\infty(\Omega, \mathcal{A}, \mathbb{P})$ υπάρχει $M > 0$ τέτοιο ώστε

$$-M \leq f \leq M \Rightarrow -M \leq \mathbb{E}(f|\mathcal{A}_i) \leq M.$$

Ομοίως και $\mathbb{E}(f|\mathcal{A}_{i-1}) \in L^\infty(\Omega, \mathcal{A}_{i-1}, \mathbb{P})$, άρα $d_i \in L^\infty(\Omega, \mathcal{A}_i, \mathbb{P})$.

Λήμμα 5.1. *Ισχύει η ανισότητα*

$$\mathbb{E}(e^{\sum_{i=1}^{n}\lambda d_i}) \leq e^{\lambda^2\sum_{i=1}^{n}\|d_i\|_\infty^2}. \tag{5.10}$$

Απόδειξη: Επαγωγικά ως προς n. Για $n=1$ έχουμε

$$\mathbb{E}(e^{\lambda d_i}) \overset{(5.5)}{=} \mathbb{E}(e^{\lambda d_i}|\mathcal{A}_0) \overset{(5.9)}{\leq} e^{\lambda^2\|d_i\|_\infty^2}.$$

Έστω ότι ισχύει για $n=k$, δηλαδή

$$\mathbb{E}(e^{\sum_{i=1}^{k}\lambda d_i}) \leq e^{\lambda^2\sum_{i=1}^{k}\|d_i\|_\infty^2}. \tag{5.11}$$

Θα δείξουμε ότι ισχύει για $n=k+1$.

$$\begin{aligned}
\mathbb{E}(e^{\sum_{i=1}^{k+1}\lambda d_i}) &\overset{(5.5)}{=} \mathbb{E}(e^{\sum_{i=1}^{k+1}\lambda d_i}|\mathcal{A}_0)\\
&\overset{(5.3)}{=} \mathbb{E}(\mathbb{E}(e^{\sum_{i=1}^{k+1}\lambda d_i}|\mathcal{A}_k)|\mathcal{A}_0))\\
&\overset{(5.5)}{=} \mathbb{E}(\mathbb{E}(e^{\sum_{i=1}^{k+1}\lambda d_i}|\mathcal{A}_k))\\
&= \mathbb{E}(\mathbb{E}(e^{\sum_{i=1}^{k}\lambda d_i}e^{\lambda d_{k+1}}|\mathcal{A}_k))\\
&\overset{(5.4)}{=} \mathbb{E}(e^{\sum_{i=1}^{k}\lambda d_i}\mathbb{E}(e^{\lambda d_{k+1}}|\mathcal{A}_k))\\
&\overset{(5.9)}{\leq} \mathbb{E}\big(e^{\sum_{i=1}^{k}\lambda d_i}e^{\lambda^2\|d_{k+1}\|_\infty^2}\big)\\
&= e^{\lambda^2\|d_{k+1}\|_\infty^2}\mathbb{E}\big(e^{\sum_{i=1}^{k}\lambda d_i}\big)\\
&\overset{(5.11)}{\leq} e^{\lambda^2\|d_{k+1}\|_\infty^2}e^{\sum_{i=1}^{k}\lambda^2\|d_i\|_\infty^2}\\
&= e^{\sum_{i=1}^{k+1}\lambda^2\|d_i\|_\infty^2}.
\end{aligned}$$

□

Συνεχίζουμε την απόδειξη του θεωρήματος 5.2. Έστω $t>0$ και $\lambda>0$. Έχουμε

$$\begin{aligned}
\mathbb{P}(f-\mathbb{E}f\geq t) &= \mathbb{P}(\mathbb{E}(f|\mathcal{A}_n)-\mathbb{E}(f|\mathcal{A}_0)\geq t)\\
&= \mathbb{P}\Big(\sum_{i=1}^{n}d_i\geq t\Big)\\
&= \mathbb{P}\Big(\sum_{i=1}^{n}\lambda d_i\geq \lambda t\Big)\\
&\leq e^{-\lambda t}\mathbb{E}\big(e^{\sum_{i=1}^{n}\lambda d_i}\big)\\
&\overset{(5.10)}{\leq} e^{\lambda^2\sum_{i=1}^{n}\|d_i\|_\infty^2-\lambda t}.
\end{aligned}$$

Ελαχιστοποιώντας την τελευταία ως προς λ, το

$$\lambda = \frac{t}{2\sum_{i=1}^{n}\|d_i\|_\infty^2}$$

δίνει

$$\mathbb{P}(f - \mathbb{E}f \geq t) \leq e^{-t^2/4\sum_{i=1}^{n}\|d_i\|_\infty^2}. \tag{5.12}$$

Τώρα, αν εφαρμόσουμε την (5.12) για τη συνάρτηση $-f$ παίρνουμε

$$\mathbb{P}(f - \mathbb{E}f \leq -t) \leq e^{-t^2/4\sum_{i=1}^{n}\|d_i\|_\infty^2}. \tag{5.13}$$

Οπότε, συνδυάζοντας τις σχέσεις (5.12) και (5.13) προκύπτει η ζητούμενη ανισότητα

$$\mathbb{P}(|f - \mathbb{E}f| \geq t) \leq e^{-t^2/4\sum_{i=1}^{n}\|d_i\|_\infty^2}.$$

□

5.3 Συγκέντρωση του μέτρου στο χώρο των μεταθέσεων

Σε αυτήν την παράγραφο θα δούμε μια εφαρμογή της ανισότητας του *Azuma* για να εκτιμήσουμε τη συνάρτηση συγκέντρωσης $\alpha(S_n, t)$, της οικογένειας S_n των μεταθέσεων του $\{1, 2, \ldots, n\}$. Η S_n γίνεται μετρικός χώρος πιθανότητας αν την εφοδιάσουμε με την μετρική $d(\sigma, \tau) = \frac{1}{n}|\{i : \ \sigma(i) \neq \tau(i)\}|$ και με το ομοιόμορφο μέτρο πιθανότητας $\mathbb{P}$ που δίνει μάζα $\frac{1}{n!}$ σε κάθε μετάθεση.

Θεώρημα 5.3. *Για κάθε $t > 0$,*

$$\alpha(S_n, t) \leq 2\exp(-t^2 n/16).$$

Απόδειξη: Έστω $\mathcal{A}_j$ η άλγεβρα που παράγεται από τα

$$A_{i_1,\ldots,i_j} = \{\sigma : \ \sigma(1) = i_1, \ldots, \sigma(j) = i_j\}$$

όπου $i_1, \ldots, i_j$ διακεκριμένα στοιχεία του $\{1, \ldots, n\}$. Θεωρούμε την ακολουθία

$$\{\emptyset, S_n\} = \mathcal{A}_0 \subseteq \mathcal{A}_1 \subseteq \ldots \subseteq \mathcal{A}_n = 2^{S_n}$$

όπου 2^{S_n} το δυναμοσύνολο του S_n. Τότε $\mathcal{A}_j \subseteq \mathcal{A}_{j+1}$ για κάθε $j = 0, \ldots, n-1$. Πράγματι, κάθε άτομο της $\mathcal{A}_j$ είναι της μορφής

$$A_{i_1,\ldots,i_j} = \bigcup_{k \notin \{i_1,\ldots,i_j\}} \{\sigma : \ \sigma(1) = i_1, \ldots, \sigma(j) = i_j, \sigma(j+1) = k\},$$

δηλαδή ανήκει στην $\mathcal{A}_{j+1}$.

Έστω $f : S_n \to \mathbb{R}$ μια συνάρτηση *Lipschitz* με σταθερά 1 και έστω $\{f_j\}_{j=0}^n$ το martingale $f_j = \mathbb{E}(f|\mathcal{A}_j)$ που παράγεται από την f. Η ακολουθία $\{f_j\}_{j=0}^n$ είναι πράγματι martingale διότι, προφανώς η $\mathbb{E}(f|\mathcal{A}_j)$ ανήκει στον $L^1(S_n, \mathcal{A}_j, \mathbb{P})$ και επειδή $\mathcal{A}_{j-1} \subseteq \mathcal{A}_j$, έχουμε

$$\mathbb{E}(f_j|\mathcal{A}_{j-1}) = \mathbb{E}(\mathbb{E}(f|\mathcal{A}_j)|\mathcal{A}_{j-1}) = \mathbb{E}(f|\mathcal{A}_{j-1}) = f_{j-1}.$$

Λήμμα 5.2. *Για κάθε άτομο $A = A_{i_1,\ldots,i_j}$ της $\mathcal{A}_j$ και κάθε ζευγάρι ατόμων $B = A_{i_1,\ldots,i_j,r}$ και $C = A_{i_1,\ldots,i_j,s}$ της $\mathcal{A}_{j+1}$ που περιέχονται στο A, μπορούμε να βρούμε μια ένα προς ένα και επί απεικόνιση $\phi : B \to C$ τέτοια ώστε*

$$d(b, \phi(b)) \leq \frac{2}{n},$$

για κάθε $b \in B$.

Απόδειξη: Έστω π η μετάθεση που αντιμεταθέτει τα r και s και αφήνει αμετάβλητα τα υπόλοιπα στοιχεία του $\{1,\ldots,n\}$. Ορίζουμε $\phi : B \to C$ με $\phi(\sigma) = \pi \circ \sigma$. Τότε $\phi(\sigma)(i) = \pi(\sigma(i)) = \sigma(i)$ αν $\sigma(i) \neq r, s$. Επομένως

$$d(b, \phi(b)) = \frac{1}{n}|\{i : \ b(i) \neq \phi(b)(i)\}| \leq \frac{2}{n}.$$

Η ϕ είναι ένα προς ένα και επειδή $|B| = |C|$ είναι και επί.

□

Έστω A, B, C όπως στο λήμμα 5.2. Αφού τα B, C είναι άτομα της $\mathcal{A}_{j+1}$, η f_{j+1} είναι σταθερή στα B, C. Από τον ορισμό της δεσμευμένης μέσης τιμής έχουμε

$$\int_B \mathbb{E}(f|\mathcal{A}_{j+1})d\mathbb{P} = \int_B f d\mathbb{P} = \frac{1}{n!}\sum_{\sigma\in B} f(\sigma).$$

Όμως η $f_{j+1} = \mathbb{E}(f|\mathcal{A}_{j+1})$ είναι σταθερή στο B, άρα

$$\begin{aligned}
\int_B f_{j+1}|_B d\mathbb{P} &= \frac{1}{n!}\sum_{\sigma\in B} f(\sigma) \Rightarrow \\
f_{j+1}|_B \mathbb{P}(B) &= \frac{1}{n!}\sum_{\sigma\in B} f(\sigma) \Rightarrow \\
f_{j+1}|_B \frac{|B|}{n!} &= \frac{1}{n!}\sum_{\sigma\in B} f(\sigma) \Rightarrow \\
f_{j+1}|_B &= \frac{1}{|B|}\sum_{\sigma\in B} f(\sigma).
\end{aligned}$$

Ομοίως δείχνουμε και ότι

$$f_{j+1}|_C = \frac{1}{|C|}\sum_{\sigma\in C} f(\sigma).$$

Γράφουμε,

$$f_{j+1}|_C = \frac{1}{|C|}\sum_{\sigma\in C} f(\sigma) = \frac{1}{|\phi(B)|}\sum_{\sigma\in B} f(\phi(\sigma)) = \frac{1}{|B|}\sum_{\sigma\in B} f(\phi(\sigma)).$$

Από το λήμμα 5.2 και επειδή η f είναι συνάρτηση *Lipschitz* με σταθερά 1, παίρνουμε

$$\begin{aligned}\left|f_{j+1}|_B - f_{j+1}|_C\right| &\leq \frac{1}{|B|}\sum_{\sigma\in B}|f(\sigma)-f(\phi(\sigma))|\\ &\leq \frac{1}{|B|}\sum_{\sigma\in B} d(\sigma,\phi(\sigma))\\ &\leq \frac{1}{|B|}\sum_{\sigma\in B}\frac{2}{n}=\frac{2}{n}.\end{aligned}$$

Τώρα, θα δείξουμε ότι $\left|f_{j+1}|_B - f_j|_A\right| \leq \frac{2}{n}$. Πράγματι, έχουμε

$$A = \bigcup_{s\notin\{i_1,\ldots,i_j\}} A_{i_1,\ldots,i_j,s}.$$

Άρα,

$$|A| = \left|\bigcup_{s\notin\{i_1,\ldots,i_j\}} A_{i_1,\ldots,i_j,s}\right| = \sum_{s\notin\{i_1,\ldots,i_j\}} |A_{i_1,\ldots,i_j,s}| = (n-j)|C|.$$

Επομένως,

$$\begin{aligned}f_j|_A &= \frac{1}{|A|}\sum_{\sigma\in A} f(\sigma)\\ &= \frac{1}{(n-j)|C|}\sum_{\substack{C\subseteq A\\ C\in\mathcal{A}_{j+1}}}\sum_{\sigma\in C} f(\sigma)\\ &= \frac{1}{n-j}\sum_{\substack{C\subseteq A\\ C\in\mathcal{A}_{j+1}}} f_{j+1}|_C.\end{aligned}$$

Επειδή η f_{j+1} είναι σταθερή στο B έχουμε

$$\sum_{\substack{C\subseteq A\\ C\in\mathcal{A}_{j+1}}} f_{j+1}|_B = (n-j)f_{j+1}|_B.$$

Άρα,

$$\begin{aligned}
\left|f_{j+1}|_B - f_j|_A\right| &= \left|\frac{1}{n-j}\sum_{\substack{C\subseteq A\\ C\in\mathcal{A}_{j+1}}} f_{j+1}|_B - \frac{1}{n-j}\sum_{\substack{C\subseteq A\\ C\in\mathcal{A}_{j+1}}} f_{j+1}|_C\right| \\
&= \left|\sum_{\substack{C\subseteq A\\ C\in\mathcal{A}_{j+1}}} \frac{1}{n-j}(f_{j+1}|_B - f_{j+1}|_C)\right| \\
&\leq \sum_{\substack{C\subseteq A\\ C\in\mathcal{A}_{j+1}}} \frac{1}{n-j}|f_{j+1}|_B - f_{j+1}|_C| \\
&\leq \sum_{\substack{C\subseteq A\\ C\in\mathcal{A}_{j+1}}} \frac{1}{n-j}\frac{2}{n} = \frac{1}{n-j}(n-j)\frac{2}{n} = \frac{2}{n}.
\end{aligned}$$

Έστω $d_j = \mathbb{E}(f|\mathcal{A}_j) - \mathbb{E}(f|\mathcal{A}_{j-1}) = f_j - f_{j-1}$, όπως στην ανισότητα του *Azuma*. Θα δείξουμε ότι $|d_{j+1}|_B| \leq \frac{2}{n}$. Πράγματι, η f_j είναι σταθερή στο A επειδή το A είναι άτομο της $\mathcal{A}_j$ και $B \subseteq A$ άρα $f_j|_B = f_j|_A$. Έχουμε,

$$\left|d_{j+1}|_B\right| = \left|f_{j+1}|_B - f_j|_B\right| = \left|f_{j+1}|_B - f_j|_A\right| \leq \frac{2}{n}.$$

Επομένως και

$$\|d_{j+1}\|_\infty \leq \frac{2}{n}. \tag{5.14}$$

Η f προφανώς ανήκει στον $L^\infty(S_n, \mathcal{A}_n, \mathbb{P})$ άρα από την (5.14) και την ανισότητα του *Azuma* προκύπτει ότι

$$\begin{aligned}
\mathbb{P}(|f - \mathbb{E}f| \geq t) &\leq 2e^{-t^2/4\sum_{i=1}^n \|d_j\|_\infty^2} \\
&\leq 2e^{-t^2/4\sum_{i=1}^n \frac{4}{n^2}} \\
&= 2e^{-t^2 n/16}.
\end{aligned}$$

Έχουμε λοιπόν αποδείξει την εξής πρόταση.

Πρόταση 5.1. *Έστω $f : S_n \to \mathbb{R}$ συνάρτηση Lipschitz με σταθερά 1. Τότε, για κάθε $t > 0$,*

$$\mathbb{P}(|f - \mathbb{E}f| \geq t) \leq 2e^{-t^2 n/16}.$$

Η απόδειξη του θεωρήματος 5.3 ολοκληρώνεται ως εξής: Έστω A υποσύνολο του S_n με $\mathbb{P}(A) \geq 1/2$. Τότε, αφού το A είναι κλειστό έχουμε ότι, $\{\sigma : \ d(\sigma, A) = 0\} = A$ άρα και

$$\mathbb{P}(\sigma : \ d(\sigma, A) = 0) = \mathbb{P}(A) \geq \frac{1}{2}. \tag{5.15}$$

Η συνάρτηση $\sigma \mapsto d(\sigma, A)$ είναι *Lipschitz* με σταθερά 1. Επιλέγουμε $t > 0$ έτσι ώστε $2e^{-t^2 n/16} = \frac{1}{2} - \delta$, όπου $0 < \delta < 1/2$, δηλαδή

$$t = 4\sqrt{\frac{\ln(4/1 - 2\delta)}{n}}.$$

Από την πρόταση 5.1 έχουμε

$$\mathbb{P}(|d(\cdot, A) - \mathbb{E}(d(\cdot, A))| \geq t) \leq 2e^{-t^2 n/16} \Rightarrow$$

$$1 - \mathbb{P}(|d(\cdot, A) - \mathbb{E}(d(\cdot, A))| < t) \leq 2e^{-t^2 n/16} \Rightarrow$$

$$\mathbb{P}\left(|d(\cdot, A) - \mathbb{E}(d(\cdot, A))| < 4\sqrt{\frac{\ln(4/1 - 2\delta)}{n}}\right) \geq \frac{1}{2} + \delta > \frac{1}{2} \Rightarrow \tag{5.16}$$

$$\mathbb{P}\left(d(\cdot, A) - \mathbb{E}(d(\cdot, A)) < 4\sqrt{\frac{\ln(4/1 - 2\delta)}{n}}\right) \geq \frac{1}{2} + \delta \Rightarrow$$

$$\mathbb{P}\left(d(\cdot, A) < 4\sqrt{\frac{\ln(4/1 - 2\delta)}{n}} + \mathbb{E}(d(\cdot, A))\right) \geq \frac{1}{2} + \delta \Rightarrow$$

$$\mathbb{P}\left(d(\cdot, A) \leq 4\sqrt{\frac{\ln(4/1 - 2\delta)}{n}} + \mathbb{E}(d(\cdot, A))\right) \leq 1 - \left(\frac{1}{2} + \delta\right) = 2e^{-t^2 n/16}. \tag{5.17}$$

Από τη (5.16) προκύπτει ότι

$$\mathbb{P}\left(\mathbb{E}(d(\cdot, A)) - d(\sigma, A) < 4\sqrt{\frac{\ln(4/1 - 2\delta)}{n}}\right) \geq \frac{1}{2} + \delta. \tag{5.18}$$

Από την (5.15) και την (5.18) βλέπουμε ότι υπάρχει $\sigma \in A$ με

$$\mathbb{E}(d(\cdot, A)) \leq 4\sqrt{\frac{\ln(4/1 - 2\delta)}{n}}. \tag{5.19}$$

Αντικαθιστώντας στην (5.17) παίρνουμε ότι

$$\mathbb{P}\left(d(\sigma, A) \geq 8\sqrt{\frac{\ln(4/1-2\delta)}{n}}\right) \leq 2e^{-t^2n/16}.$$

Τώρα, αν $t > 8\sqrt{\frac{\ln(4/1-2\delta)}{n}}$ τότε

$$\mathbb{P}(d(\sigma, A) \geq t) \leq 2e^{-t^2n/16}.$$

Αν $t \leq 8\sqrt{\frac{\ln(4/1-2\delta)}{n}}$ τότε

$$\mathbb{P}(d(\sigma, A) \geq t) \leq \mathbb{P}(S_n \backslash A) \leq \frac{1}{2} = 2e^{-t^2n/16} + \delta.$$

Δηλαδή, για κάθε $t > 0$ έχουμε ότι

$$\mathbb{P}(A_t^c) \leq 2e^{-t^2n/16}.$$

Έπεται ότι

$$\alpha(S_n, t) \leq 2e^{-t^2n/16}.$$

$\square$

Βιβλιογραφία

[1] M. Capinski and E. Kopp, Measure, Integral and Probability, Springer – Verlag, 1999.

[2] Α. Γιαννόπουλος, Συγκέντρωση του Μέτρου, Σημειώσεις Μαθήματος Θερινού Σχολείου, Καρλόβασι Σάμου, 2004.

[3] M. Ledoux, The concentration of Measure Phenomenon, Mathematical Surveys and Monographs, Vol. 89, AMS 2001.

[4] V.D. Milman and G. Schechtman, Asympotic Theory of Finite Dimensional Normed Spaces. LNM 1200 (1986), Springer – Verlag.

[5] Τ. Παπαϊωάννου, Εισαγωγή στις Πιθανότητες, Εκδόσεις Σταμούλη, 2000.

[6] G. Pisier, The Volume of Convex Bodies and Banach Space Geometry, Cambridge tracts in Mathematics, Vol. 94, Cambidge Univ. Press.