Συγκέντρωση του μέτρου

Πρόχειρες Σημειώσεις

Τμήμα Μαθηματικών
Πανεπιστήμιο Αθηνών
Αθήνα – 2016

Περιεχόμενα

1 Ισοπεριμετρικές ανισότητες και συγκέντρωση του μέτρου 1

1.1 Μετρικοί χώροι πιθανότητας . . . 1

1.1α′ Ορισμός και παραδείγματα . . . 1

1.1β′ Το ισοπεριμετρικό πρόβλημα . . . 3

1.2 Κλασικές ισοπεριμετρικές ανισότητες . . . 4

1.2α′ Ανισότητα Brunn-Minkowski και η ισοπεριμετρική ανισότητα . . . 4

1.2β′ Ισοπεριμετρική ανισότητα στην σφαίρα . . . 9

1.2γ′ Ισοπεριμετρική ανισότητα στο χώρο του Gauss . . . 11

1.2δ′ Η ισοπεριμετρική ανισότητα στον E_2^n . . . 12

1.3 Συνάρτηση συγκέντρωσης . . . 13

1.4 Προσεγγιστικές ισοπεριμετρικές ανισότητες . . . 17

1.4α′ Η συνάρτηση συγκέντρωσης της σφαίρας . . . 17

1.4β′ Η συνάρτηση συγκέντρωσης του χώρου του Gauss . . . 19

1.4γ′ Η συνάρτηση συγκέντρωσης του διακριτού κύβου . . . 20

1.5 Εφαρμογή: το θεώρημα του Dvoretzky . . . 26

1.5α′ Το θεώρημα του John και το λήμμα των Dvoretzky και Rogers . . . 28

1.5β′ Απόδειξη του θεωρήματος του Dvoretzky . . . 29

1.6 Ασκήσεις . . . 33

2 Αθροίσματα ανεξάρτητων τυχαίων μεταβλητών 39

2.1 Η ανισότητα του Hoeffding και η ανισότητα του Chernoff . . . 39

2.1α′ Εφαρμογή: βαθμός κορυφών τυχαίου γραφήματος . . . 44

2.2 Υποκανονικές τυχαίες μεταβλητές . . . 45

2.3 Ανισότητα Khintchine-Kahane . . . 51

2.3α′ Η ανισότητα του Khintchine . . . 51

2.3β′ Η ανισότητα του Kahane 54
2.3γ′ Ανισότητα Kahane-Khinthcine για λογαριθμικά κοίλα μέτρα 56
2.4 Υποεκθετικές τυχαίες μεταβλητές και ανισότητες Bernstein 59
2.5 Ασκήσεις 65

3 Η μέθοδος των martingales 69
3.1 Ανισότητα του Azuma 69
3.2 Εφαρμογές στο διακριτό κύβο 74
3.3 Συγκέντρωση του μέτρου στην S_n 76
3.4 Ασκήσεις 81

4 Συναρτησοειδές Laplace και ελαχιστική συνέλιξη 85
4.1 Συναρτησοειδές Laplace 85
4.2 Ελαχιστική συνέλιξη 90
4.3 Η ιδιότητα (τ) 93
4.4 Η ανισότητα του Talagrand για το εκθετικό μέτρο γινόμενο 96
4.5 Η ιδιότητα (τ) στο χώρο του Gauss 100
4.6 Ασκήσεις 102

5 Ανισότητα Poincaré 105
5.1 Ανισότητα Cheeger και ανισότητα Poincaré 105
5.2 Ανισότητα Poincaré και συγκέντρωση του μέτρου 110
5.3 Ημιομάδες Markov 113
5.4 Ημιομάδα Ornstein–Uhlenbeck 119
5.5 Ανισότητα Poincaré και ιδιοτιμές του τελεστή Laplace 122
5.6 Ανισότητα Poincaré στο διακριτό κύβο 124
5.7 Ασκήσεις 126

6 Λογαριθμική ανισότητα Sobolev 129
6.1 Εντροπία και υποκανονικές εκτιμήσεις 129
6.2 Λογαριθμική ανισότητα Sobolev 131
6.3 Λογαριθμική ανισότητα Sobolev σε χώρους γινόμενα 133
6.4 Ημιομάδες Markov και η λογαριθμική ανισότητα Sobolev 135
6.5 Η λογαριθμική ανισότητα Sobolev στο χώρο του Gauss 136
6.6 Λογαριθμική ανισότητα Sobolev και υπερσυσταλτότητα 139
6.7 Λογαριθμική ανισότητα Sobolev στο διακριτό κύβο 141

6.8 Ασκήσεις 150

Κεφάλαιο 1

Ισοπεριμετρικές ανισότητες και συγκέντρωση του μέτρου

1.1 Μετρικοί χώροι πιθανότητας

1.1α΄ Ορισμός και παραδείγματα

Ορισμός 1.1.1 (μετρικός χώρος πιθανότητας)**.** Έστω (X,d) ένας μετρικός χώρος. Αν μ είναι ένα μέτρο πιθανότητας στη σ-άλγεβρα $\mathcal{B}(X)$ των Borel υποσυνόλων του (X,d), τότε η τριάδα (X,d,μ) λέγεται μετρικός χώρος πιθανότητας.

Παραδείγματα 1.1.2. 1. *Η σφαίρα S^{n-1}.* Συμβολίζουμε με $|\cdot|$ την Ευκλείδεια νόρμα στον $\mathbb{R}^n$. Θεωρούμε τη μοναδιαία σφαίρα

$$S^{n-1}=\{x\in\mathbb{R}^n : |x|=1\}$$

στον $\mathbb{R}^n$, εφοδιασμένη με τη γεωδαισιακή μετρική ρ: η απόσταση $\rho(x,y)$ δύο σημείων $x,y\in S^{n-1}$ είναι η κυρτή γωνία xoy στο επίπεδο που ορίζεται από την αρχή των αξόνων o και τα x,y. Η S^{n-1} γίνεται χώρος πιθανότητας με το μοναδικό αναλλοίωτο ως προς ορθογώνιους μετασχηματισμούς μέτρο σ: για κάθε Borel σύνολο $A\subseteq S^{n-1}$ θέτουμε

$$\sigma(A):=\frac{|C(A)|}{|B_2^n|},$$

όπου B_2^n είναι η μοναδιαία Ευκλείδεια μπάλα,

$$C(A):=\{sx : x\in A \text{ και } 0\leqslant s\leqslant 1\},$$

και $|Q|$ είναι το n-διάστατο μέτρο Lebesgue του Q. Η ρ είναι όντως μετρική στην S^{n-1} (άσκηση). Είναι εύκολο να δεί κανείς ότι αν $\rho(x,y)=\theta$ τότε

$$|x-y| = 2\sin\frac{\theta}{2},$$

συνεπώς η γεωδαισιακή και η Ευκλείδεια απόσταση των $x,y\in S^{n-1}$ συγκρίνονται μέσω της

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

2. *Ο χώρος του Gauss.* Θεωρούμε τον $\mathbb{R}^n$ με την Ευκλείδεια μετρική $|\cdot|$ και το μέτρο πιθανότητας γ_n που έχει πυκνότητα τη συνάρτηση

$$g_n(x) = (2\pi)^{-n/2}e^{-|x|^2/2}.$$

Δηλαδή, αν A είναι ένα σύνολο Borel στον $\mathbb{R}^n$, τότε

$$\gamma_n(A) = \frac{1}{(2\pi)^{n/2}}\int_A e^{-|x|^2/2}dx.$$

Το γ_n είναι το n-διάστατο *μέτρο του Gauss* και ο χώρος πιθανότητας $\Gamma_n = (\mathbb{R}^n, |\cdot|, \gamma_n)$ είναι ο n-διάστατος *χώρος του Gauss.*

Το μέτρο του Gauss έχει δύο πολύ σημαντικές ιδιότητες. Από τη μία πλευρά είναι μέτρο γινόμενο, πιο συγκεκριμένα $\gamma_n=\gamma_1\otimes\cdots\otimes\gamma_1$. Από την άλλη πλευρά είναι αναλλοίωτο ως προς ορθογώνιους μετασχηματισμούς. Αν $U\in O(n)$ και A είναι ένα Borel υποσύνολο του $\mathbb{R}^n$, τότε

$$\begin{aligned}\gamma_n(U(A)) &= \frac{1}{(2\pi)^{n/2}}\int_{U(A)} e^{-|x|^2/2}dx = \frac{|\det U|}{(2\pi)^{n/2}}\int_A e^{-|Uy|^2/2}dy \\ &= \frac{1}{(2\pi)^{n/2}}\int_A e^{-|y|^2/2}dy = \gamma_n(A).\end{aligned}$$

3. *Ο διακριτός κύβος.* Θεωρούμε το σύνολο $E_2^n=\{-1,1\}^n$, το οποίο ταυτίζουμε με το σύνολο των κορυφών του κύβου $Q_n=[-1,1]^n$ στον $\mathbb{R}^n$. Στον E_2^n ορίζουμε το ομοιόμορφο μέτρο πιθανότητας μ_n που δίνει μάζα 2^{-n} σε κάθε σημείο. Δηλαδή, αν A είναι ένα υποσύνολο του E_2^n, τότε

$$\mu_n(A) = \frac{|A|}{2^n},$$

όπου με $|A|$ (αλλά και με $\mathrm{card}(A)$) συμβολίζουμε τον πληθάριθμο ενός πεπερασμένου συνόλου.

Ο E_2^n γίνεται μετρικός χώρος με απόσταση την

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

1.1β′ Το ισοπεριμετρικό πρόβλημα

Ορισμός 1.1.3 (t-περιοχή)**.** Έστω (X,d) μετρικός χώρος. Για κάθε μη κενό $A \in \mathcal{B}(X)$ και για κάθε $t > 0$ ορίζουμε την *t-περιοχή* του A ως εξής:

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

Ορισμός 1.1.4 (επιφάνεια κατά Minkowski)**.** *Έστω (X,d) μετρικός χώρος και έστω μ ένα (όχι αναγκαστικά πεπερασμένο) μέτρο στην Borel σ-άλγεβρα $\mathcal{B}(X)$. Η επιφάνεια ενός Borel υποσυνόλου A του X ως προς το μ ορίζεται ως εξής:*

$$\mu^+(A) = \liminf_{t\to 0^+}\frac{\mu(A_t \setminus A)}{t},$$

όπου A_t είναι η t-περιοχή του A. Αν $\mu(A) < \infty$ (το οποίο φυσικά ισχύει αν ο (X,d,μ) είναι μετρικός χώρος πιθανότητας) τότε

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

Σε κάθε μετρικό χώρο πιθανότητας μπορούμε να διατυπώσουμε το *ισοπεριμετρικό πρόβλημα*:

> Για δοσμένο $0 < \alpha < 1$, να βρεθεί το
>
> $$\inf\{\mu^+(A) : A \in \mathcal{B}(X), \mu(A) \geqslant \alpha\}$$
>
> και να βρεθούν (αν υπάρχουν) τα σύνολα A για τα οποία πιάνεται αυτό το infimum.

Μπορούμε επίσης να διατυπώσουμε αντίστοιχο πρόβλημα για το μέτρο των t-περιοχών, σταθεροποιώντας $t > 0$:

> Για δοσμένα $0 < \alpha < 1$ και $t > 0$, να βρεθεί το
>
> $$\inf\{\mu(A_t) : A \in \mathcal{B}(X), \mu(A) \geqslant \alpha\}$$
>
> και να βρεθούν (αν υπάρχουν) τα σύνολα A για τα οποία πιάνεται αυτό το infimum.

Το infimum παίρνεται πάνω από όλα τα $A \in \mathcal{B}(X)$ για τα οποία $\mu(A) \geqslant \alpha$ (και όχι $\mu(A) = \alpha$) για καθαρά τεχνικούς λόγους: στο γενικό πλαίσιο που συζητάμε, μπορεί, για κάποια τιμή του α, να μην υπάρχουν σύνολα $A \in \mathcal{B}(X)$ ώστε $\mu(A) = \alpha$.

Οι λύσεις του δεύτερου προβλήματος μπορεί να είναι διαφορετικές για διαφορετικές τιμές του t. Στα κλασικά όμως παραδείγματα δεν εξαρτώνται από το t και αυτό σημαίνει ότι είναι και λύσεις του πρώτου προβλήματος. Είναι μάλιστα, όπως θα δούμε, πολύ «συμμετρικά υποσύνολα» του X, το οποίο σημαίνει ότι μπορούμε σχετικά εύκολα να υπολογίσουμε το μέτρο της t-περιοχής τους και την επιφάνειά τους.

1.2 Κλασικές ισοπεριμετρικές ανισότητες

Σε αυτή την παράγραφο θα συζητήσουμε το ισοπεριμετρικό πρόβλημα για τα κλασικά παραδείγματα μετρικών χώρων πιθανότητας που ορίστηκαν στην προηγούμενη παράγραφο: την Ευκλείδεια μοναδιαία σφαίρα S^{n-1}, το χώρο του Gauss Γ_n και το διακριτό κύβο E_2^n. Πριν όμως από αυτό, αποδεικνύουμε την ανισότητα Brunn-Minkowski και μέσω αυτής την κλασική ισοπεριμετρική ανισότητα στον $\mathbb{R}^n$.

1.2α΄ Ανισότητα Brunn-Minkowski και η ισοπεριμετρική ανισότητα

Ορισμός 1.2.1 (άθροισμα Minkowski)**.** Έστω A και B μη κενά υποσύνολα του $\mathbb{R}^n$. Ορίζουμε

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

και για κάθε $t \geqslant 0$,

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

Η ανισότητα Brunn-Minkowski συνδέει το άθροισμα Minkowski με τον όγκο $|\cdot|$ στον $\mathbb{R}^n$:

Θεώρημα 1.2.2 (ανισότητα Brunn-Minkowski)**.** *Έστω K και T δύο μη κενά Borel υποσύνολα του $\mathbb{R}^n$. Τότε,*

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

Παρατηρήσεις 1.2.3. Στην περίπτωση που τα K και T είναι κυρτά σώματα (κυρτά συμπαγή σύνολα με μη κενό εσωτερικό), ισότητα στην ανισότητα Brunn-Minkowski μπορεί να ισχύει μόνο αν τα K και T είναι ομοιοθετικά (δηλαδή, αν $K = aT + x$ για κάποιον $a \geqslant 0$ και κάποιο $x \in \mathbb{R}^n$).

Η ανισότητα Brunn-Minkowski εκφράζει με μια έννοια το γεγονός ότι ο όγκος είναι *κοίλη* συνάρτηση ως προς την πρόσθεση κατά Minkowski. Για το λόγο αυτό συχνά γράφεται στην ακόλουθη μορφή: Αν K,T είναι μη κενά Borel υποσύνολα του $\mathbb{R}^n$ και αν $\lambda \in (0,1)$, τότε

$$|\lambda K + (1-\lambda)T|^{1/n} \geqslant \lambda |K|^{1/n} + (1-\lambda)|T|^{1/n}.$$

Χρησιμοποιώντας την τελευταία ανισότητα σε συνδυασμό με την ανισότητα αριθμητικού-γεωμετρικού μέσου, μπορούμε ακόμα να γράψουμε:

$$|\lambda K + (1-\lambda)T| \geqslant |K|^{\lambda}|T|^{1-\lambda}.$$

Η απόδειξη που θα δώσουμε εδώ θα βασιστεί στη *συναρτησιακή ανισότητα* των Prékopa και Leindler, η οποία είναι η γενίκευση της ανισότητας Brunn-Minkowski στο πλαίσιο των μετρήσιμων θετικών συναρτήσεων.

Θεώρημα 1.2.4 (ανισότητα Prékopa-Leindler)**.** *Έστω $f,g,h:\mathbb{R}^n \to \mathbb{R}^+$ μετρήσιμες συναρτήσεις και έστω $\lambda \in (0,1)$. Υποθέτουμε ότι οι f και g είναι ολοκληρώσιμες και ότι, για κάθε $x,y \in \mathbb{R}^n$,*

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

Τότε,

$$\int_{\mathbb{R}^n} h \geqslant \left(\int_{\mathbb{R}^n} f\right)^{\lambda} \left(\int_{\mathbb{R}^n} g\right)^{1-\lambda}.$$

Απόδειξη. Θα δείξουμε την ανισότητα με επαγωγή ως προς τη διάσταση n.

(α) $n=1$: Μπορούμε να υποθέσουμε ότι οι f και g είναι συνεχείς και γνησίως θετικές. Ορίζουμε $x,y:(0,1)\to\mathbb{R}$ μέσω των

$$\int_{-\infty}^{x(t)} f = t\int f \qquad \text{και} \qquad \int_{-\infty}^{y(t)} g = t\int g.$$

Σύμφωνα με τις υποθέσεις μας, οι x,y είναι παραγωγίσιμες και για κάθε $t\in(0,1)$ έχουμε

$$x'(t)f(x(t)) = \int_{\mathbb{R}} f \qquad \text{και} \qquad y'(t)g(y(t)) = \int_{\mathbb{R}} g.$$

Ορίζουμε $z:(0,1)\to\mathbb{R}$ με

$$z(t) = \lambda x(t) + (1-\lambda)y(t).$$

Οι x και y είναι γνησίως αύξουσες. Συνεπώς, η z είναι κι αυτή γνησίως αύξουσα. Από την ανισότητα αριθμητικού-γεωμετρικού μέσου,

$$z'(t) = \lambda x'(t) + (1-\lambda)y'(t) \geqslant (x'(t))^{\lambda}(y'(t))^{1-\lambda}.$$

Μπορούμε λοιπόν να εκτιμήσουμε το ολοκλήρωμα της h κάνοντας την αλλαγή μεταβλητών $s = z(t)$:

$$\begin{aligned}
\int h(s)ds &= \int_0^1 h(z(t))z'(t)dt \\
&\geqslant \int_0^1 h(\lambda x(t) + (1-\lambda)y(t))(x'(t))^{\lambda}(y'(t))^{1-\lambda}dt \\
&\geqslant \int_0^1 [f(x(t))]^{\lambda}[g(y(t))]^{1-\lambda}\left(\frac{\int_{\mathbb{R}} f}{f(x(t))}\right)^{\lambda}\left(\frac{\int_{\mathbb{R}} g}{g(y(t))}\right)^{1-\lambda} dt \\
&= \int_0^1 \left[\left(\int_{\mathbb{R}} f\right)^{\lambda}\left(\int_{\mathbb{R}} g\right)^{1-\lambda}\right] dt \\
&= \left(\int_{\mathbb{R}} f\right)^{\lambda}\left(\int_{\mathbb{R}} g\right)^{1-\lambda}.
\end{aligned}$$

(β) *Επαγωγικό βήμα:* Υποθέτουμε ότι $n \geqslant 2$ και ότι το θεώρημα έχει αποδειχθεί για $k \in \{1, \ldots, n-1\}$. Έστω f, g, h όπως στο θεώρημα. Για κάθε $s \in \mathbb{R}$ ορίζουμε $h_s : \mathbb{R}^{n-1} \to \mathbb{R}^+$ με $h_s(w) = h(w, s)$ και με ανάλογο τρόπο ορίζουμε $f_s, g_s : \mathbb{R}^{n-1} \to \mathbb{R}^+$. Επίσης, ορίζουμε $H : \mathbb{R} \to \mathbb{R}^+$ με $H(s) = \int_{\mathbb{R}^{n-1}} h_s$ και με ανάλογο τρόπο ορίζουμε $F, G : \mathbb{R} \to \mathbb{R}^+$. Από την υπόθεση του θεωρήματος για τις f, g και h έπεται ότι, αν $x, y \in \mathbb{R}^{n-1}$ και $s_0, s_1 \in \mathbb{R}$ τότε

$$h_{\lambda s_1 + (1-\lambda)s_0}(\lambda x + (1-\lambda)y) \geqslant f_{s_1}(x)^{\lambda} g_{s_0}(y)^{1-\lambda},$$

και η επαγωγική υπόθεση μας δίνει

$$\begin{aligned}
H(\lambda s_1 + (1-\lambda)s_0) &:= \int_{\mathbb{R}^{n-1}} h_{\lambda s_1 + (1-\lambda)s_0} \\
&\geqslant \left(\int_{\mathbb{R}^{n-1}} f_{s_1}\right)^{\lambda}\left(\int_{\mathbb{R}^{n-1}} g_{s_0}\right)^{1-\lambda} =: F^{\lambda}(s_1)G^{1-\lambda}(s_0).
\end{aligned}$$

Εφαρμόζοντας τώρα ξανά την επαγωγική υπόθεση για $n = 1$ στις συναρτήσεις F, G και H, παίρνουμε

$$\int_{\mathbb{R}^n} h = \int_{\mathbb{R}} H \geqslant \left(\int_{\mathbb{R}} F\right)^{\lambda}\left(\int_{\mathbb{R}} G\right)^{1-\lambda} = \left(\int_{\mathbb{R}^n} f\right)^{\lambda}\left(\int_{\mathbb{R}^n} g\right)^{1-\lambda},$$

όπου οι ισότητες στην τελευταία γραμμή προκύπτουν άμεσα από το θεώρημα Fubini. □

Χρησιμοποιώντας την ανισότητα Prékopa–Leindler μπορούμε να αποδείξουμε την ανισότητα Brunn–Minkowski:

Απόδειξη του Θεωρήματος 1.2.2. Έστω K, T μη κενά Borel υποσύνολα του $\mathbb{R}^n$, και έστω $\lambda \in (0,1)$. Ορίζουμε $f = \chi_K$, $g = \chi_T$, και $h = \chi_{\lambda K + (1-\lambda)T}$. Εύκολα ελέγχουμε ότι ικανοποιούνται οι υποθέσεις του θεωρήματος 1.2.4. Πράγματι, αν $x \notin K$ ή $y \notin T$ τότε

$$h(\lambda x + (1-\lambda)y) \geqslant 0 = [f(x)]^{\lambda}[g(y)]^{1-\lambda},$$

ενώ αν $x \in K$ και $y \in T$ τότε $\lambda x + (1-\lambda)y \in \lambda K + (1-\lambda)T$, άρα

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

Εφαρμόζοντας την ανισότητα Prékopa-Leindler παίρνουμε

$$|\lambda K + (1-\lambda)T| = \int h \geqslant \left(\int f\right)^{\lambda}\left(\int g\right)^{1-\lambda} = |K|^{\lambda}|T|^{1-\lambda}. \tag{1.2.1}$$

Θεωρούμε τώρα K και T όπως στο Θεώρημα 1.2.2 (με $|K| > 0$ και $|T| > 0$, αλλιώς δεν έχουμε τίποτα να δείξουμε), και ορίζουμε

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

Τα K_1 και T_1 έχουν όγκο 1, οπότε από την (1.2.1) παίρνουμε

$$|\lambda K_1 + (1-\lambda)T_1| \geqslant 1. \tag{1.2.2}$$

Όμως,

$$\lambda K_1 + (1-\lambda)T_1 = \frac{K+T}{|K|^{1/n} + |T|^{1/n}},$$

επομένως η (1.2.2) παίρνει τη μορφή

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

και έπεται το ζητούμενο. □

Χρησιμοποιώντας την ανισότητα Brunn–Minkowski μπορούμε να δώσουμε την απάντηση στο ισοπεριμετρικό πρόβλημα στον $\mathbb{R}^n$:

Ανάμεσα σε όλα τα μη κενά Borel υποσύνολα του $\mathbb{R}^n$ που έχουν δεδομένο όγκο α, η μπάλα όγκου α έχει ελάχιστη επιφάνεια.

Ο ορισμός της επιφάνειας που θα χρησιμοποιήσουμε είναι αυτός του Minkowski, ο οποίος βασίζεται στην έννοια της *t-περιοχής*: θυμηθείτε ότι αν A είναι ένα μη κενό Borel υποσύνολο του $\mathbb{R}^n$ και αν $t > 0$, η t-περιοχή του A είναι το σύνολο $A_t = \{x \in \mathbb{R}^n : d(x, A) < t\}$, όπου $d(x, A) = \inf\{|x - a| : a \in A\}$ είναι η Ευκλείδεια απόσταση του x από το σύνολο A. Παρατηρήστε ότι, για κάθε $t > 0$,

$$A_t = A + tD_n,$$

όπου $D_n = \{x \in \mathbb{R}^n : |x| < 1\}$ είναι η ανοικτή Ευκλείδεια μοναδιαία μπάλα στον $\mathbb{R}^n$. Σύμφωνα με τον ορισμό της επιφάνειας κατά Minkowski, αν A είναι ένα μη κενό Borel υποσύνολο του $\mathbb{R}^n$ με πεπερασμένο όγκο, η *επιφάνεια* $\partial(A)$ του A ορίζεται από την

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

Μπορεί να ελέγξει κανείς ότι αν το A είναι κυρτό σώμα, τότε το $\liminf$ στο δεξιό μέλος είναι όριο.

Η απάντηση στο ισοπεριμετρικό πρόβλημα για τον Ευκλείδειο χώρο δίνεται από το εξής θεώρημα.

Θεώρημα 1.2.5. *Αν A είναι μη κενό Borel υποσύνολο του $\mathbb{R}^n$ με πεπερασμένο όγκο, τότε*

$$\partial(A) \geqslant n|A|^{\frac{n-1}{n}}|B_2^n|^{\frac{1}{n}}.$$

Πράγματι, παρατηρούμε αρχικά ότι για κάθε $t > 0$, ανάμεσα σε όλα τα μη κενά Borel υποσύνολα του Ευκλείδειου χώρου που έχουν δεδομένο όγκο, η μπάλα έχει τη «μικρότερη t-επέκταση».

Πρόταση 1.2.6. *Έστω B μια μπάλα στον $\mathbb{R}^n$. Αν A είναι ένα μη κενό Borel υποσύνολο του $\mathbb{R}^n$ με $|A| = |B|$, τότε $|A_t| \geqslant |B_t|$ για κάθε $t > 0$.*

Απόδειξη. Λόγω του αναλλοίωτου του όγκου ως προς μεταφορές, μπορούμε να υποθέσουμε ότι $B = rB_2^n$ για κάποιον $r > 0$. Από την ανισότητα Brunn-Minkowski παίρνουμε

$$\begin{aligned} |A + tD_n|^{1/n} &\geqslant |A|^{1/n} + |tD_n|^{1/n} = |A|^{1/n} + t|D_n|^{1/n} = r|B_2^n|^{1/n} + t|B_2^n|^{1/n} \\ &= |(r+t)B_2^n|^{1/n} = |(r+t)D_n|^{1/n} = |B + tD_n|^{1/n}, \end{aligned}$$

απ' όπου έπεται το ζητούμενο. □

Τώρα, από τον ορισμό της επιφάνειας έχουμε:

Θεώρημα 1.2.7. *Έστω A μη κενό Borel υποσύνολο του $\mathbb{R}^n$ με πεπερασμένο όγκο και έστω $r > 0$ τέτοιος ώστε $|A| = |rB_2^n|$. Τότε,*

$$\partial(A) \geqslant \partial(rB_2^n).$$

Απόδειξη. Χρησιμοποιώντας την προηγούμενη πρόταση γράφουμε

$$\begin{aligned}\partial(A) &= \liminf_{t\to 0^+} \frac{|A_t| - |A|}{t} = \liminf_{t\to 0^+} \frac{|A_t| - |rB_2^n|}{t} \geqslant \liminf_{t\to 0^+} \frac{|(rB_2^n)_t| - |rB_2^n|}{t} \\ &= \partial(rB_2^n).\end{aligned}$$

□

Απόδειξη του Θεωρήματος 1.2.5. Αρκεί να παρατηρήσουμε ότι αν $|A| = |rB_2^n|$ τότε $r = (|A|/|B_2^n|)^{1/n}$ και ότι

$$\partial(rB_2^n) = \lim_{t\to 0^+} \frac{|(r+t)D_n| - |rB_2^n|}{t} = \lim_{t\to 0^+} \frac{(r+t)^n - r^n}{t} |B_2^n| = nr^{n-1}|B_2^n|.$$

Αντικαθιστώντας το r βλέπουμε ότι $nr^{n-1}|B_2^n| = n|A|^{\frac{n-1}{n}}|B_2^n|^{\frac{1}{n}}$. □

1.2β′ Ισοπεριμετρική ανισότητα στην σφαίρα

Το ισοπεριμετρικό πρόβλημα στη σφαίρα διατυπώνεται ως εξής:

> Δίνονται $\alpha \in (0,1)$ και $t > 0$. Ανάμεσα σε όλα τα Borel υποσύνολα A της σφαίρας για τα οποία $\sigma(A) = \alpha$, να βρεθούν εκείνα για τα οποία ελαχιστοποιείται η επιφάνεια $\sigma(A_t)$ της t-περιοχής του A.

Η απάντηση δίνεται από το ακόλουθο θεώρημα:

Θεώρημα 1.2.8 (ισοπεριμετρική ανισότητα στην σφαίρα)**.** *Έστω $\alpha \in (0,1)$ και έστω*

$$B(x,r) = \{y \in S^{n-1} : \rho(x,y) \leqslant r\}$$

μια μπάλα στην S^{n-1} με ακτίνα $r > 0$ που επιλέγεται ώστε $\sigma(B(x,r)) = \alpha$. Τότε, για κάθε $A \subseteq S^{n-1}$ με $\sigma(A) = \alpha$ και για κάθε $t > 0$ έχουμε

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

Δηλαδή, για οποιοδήποτε δοσμένο μέτρο α και οποιοδήποτε $t > 0$ οι γεωδαισιακές μπάλες μέτρου α δίνουν τη λύση του ισοπεριμετρικού προβλήματος.

Η απόδειξη της ισοπεριμετρικής ανισότητας γίνεται με σφαιρική συμμετρικοποίηση και επαγωγή ως προς την διάσταση. Ας θεωρήσουμε την ειδική περίπτωση $\alpha = 1/2$. Αν $\sigma(A) = 1/2$ και $t > 0$, τότε μπορούμε να εκτιμήσουμε το μέγεθος του A_t χρησιμοποιώντας την ισοπεριμετρική ανισότητα:

$$\sigma(A_t) \geqslant \sigma\left(B\big(x, \frac{\pi}{2} + t\big)\right)$$

για κάθε $t > 0$ και $x \in S^{n-1}$. Εκτιμώντας από κάτω το δεξιό μέλος αυτής της ανισότητας οδηγούμαστε στο εξής:

Θεώρημα 1.2.9. *Εστω $A \subseteq S^{n+1}$ με $\sigma(A) = \frac{1}{2}$ και έστω $t > 0$. Τότε,*

$$\sigma(A_t) \geqslant 1 - \sqrt{\pi/8}\exp(-t^2 n/2).$$

Απόδειξη. Λόγω της σφαιρικής ισοπεριμετρικής ανισότητας, αρκεί να φράξουμε από κάτω το $\sigma(B(x, \frac{\pi}{2} + t))$. Παρατηρήστε ότι

$$\sigma(B(x, \frac{\pi}{2} + t)) = \frac{\int_0^{\frac{\pi}{2}+t} \sin^n \theta d\theta}{\int_0^{\pi} \sin^n \theta d\theta},$$

οπότε θέτοντας $h(t,n) = 1 - \sigma(B(x, \frac{\pi}{2} + t))$, ζητάμε άνω φράγμα γιά την

$$h(t,n) = \frac{\int_{\frac{\pi}{2}+t}^{\pi} \sin^n \theta d\theta}{\int_0^{\pi} \sin^n \theta d\theta} = \frac{\int_t^{\frac{\pi}{2}} \cos^n \phi d\phi}{2I_n},$$

όπου

$$I_n = \int_0^{\pi/2} \cos^n \phi d\phi.$$

Κάνοντας την αλλαγή μεταβλητής $s = \phi\sqrt{n}$ παίρνουμε

$$h(t,n) = \frac{1}{2\sqrt{n}I_n} \int_{t\sqrt{n}}^{\frac{\pi}{2}\sqrt{n}} \cos^n(s/\sqrt{n})ds.$$

Συγκρίνοντας τα αναπτύγματα Taylor των συναρτήσεων $\cos s$ και $\exp(-s^2/2)$ βλέπουμε οτι

$$\cos s \leqslant \exp(-s^2/2)$$

στο $[0, \pi/2]$, επομένως

$$\begin{aligned} h(t,n) &\leqslant \frac{1}{2\sqrt{n}I_n}\int_{t\sqrt{n}}^{\frac{\pi}{2}\sqrt{n}} e^{-\frac{s^2}{2}}ds \\ &= \frac{1}{2\sqrt{n}I_n}\int_0^{(\frac{\pi}{2}-t)\sqrt{n}} \exp(-(s+t\sqrt{n})^2/2)ds \\ &\leqslant \frac{\exp(-t^2n/2)}{2\sqrt{n}I_n}\int_0^{\infty}\exp(-s^2/2)ds \\ &= \frac{\sqrt{\pi/8}}{\sqrt{n}I_n}\exp(-t^2n/2). \end{aligned}$$

Γιά την απόδειξη του θεωρήματος αρκεί λοιπόν να δείξουμε οτι $\sqrt{n}I_n \geqslant 1$ γιά κάθε $n \geqslant 1$. Γιά το σκοπό αυτό παρατηρούμε ότι από την αναδρομική σχέση $(n+2)I_{n+2} = (n+1)I_n$ έπεται ότι

$$\sqrt{n+2}I_{n+2} = \sqrt{n+2}\frac{n+1}{n+2}I_n = \frac{n+1}{\sqrt{n+2}}I_n \geqslant \sqrt{n}I_n,$$

το οποίο σημαίνει ότι αρκεί να ελέγξουμε τις

$$I_1 = \int_0^{\pi/2}\cos\phi\, d\phi = 1 \geqslant 1$$

και

$$\sqrt{2}I_2 = \sqrt{2}\int_0^{\pi/2}\cos^2\phi\, d\phi = \sqrt{2}\frac{\pi}{4} \geqslant 1.$$

Αυτό ολοκληρώνει την απόδειξη. □

Παρατήρηση. Αυτό που έχει σημασία σε σχέση με την εκτίμηση του θεωρήματος 1.2.9 είναι ότι, όσο μικρό $t > 0$ κι αν διαλέξουμε, η ακολουθία $\exp(-t^2n/2)$ τείνει στο 0 καθώς $n \to \infty$ και μάλιστα με πολύ γρήγορο ρυθμό (εκθετικά ως προς n). Συνεπώς, το ποσοστό της σφαίρας που μένει έξω από την t-περιοχή οποιουδήποτε υποσυνόλου A της S^{n+1} με $\sigma(A) = 1/2$ είναι «σχεδόν μηδενικό» αν η διάσταση n είναι αρκετά μεγάλη.

1.2γ΄ Ισοπεριμετρική ανισότητα στο χώρο του Gauss

Η ισοπεριμετρική ανισότητα στο χώρο του Gauss είναι η εξής.

Θεώρημα 1.2.10. *Έστω $\alpha \in (0,1)$, $\theta \in S^{n-1}$, και $H = \{x \in \mathbb{R}^n : \langle x, \theta\rangle \leqslant \lambda\}$ ένας ημίχωρος του $\mathbb{R}^n$ με $\gamma_n(H) = \alpha$. Τότε, γιά κάθε $t > 0$ και για κάθε Borel υποσύνολο A του $\mathbb{R}^n$ με $\gamma_n(A) = \alpha$ ισχύει*

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

Μια απόδειξη του Θεωρήματος 1.2.10 βασίζεται στην «παρατήρηση του Poincaré» και ουσιαστικά ανάγει το πρόβλημα στο ισοπεριμετρικό πρόβλημα για την σφαίρα. Αυτό που μας ενδιαφέρει περισσότερο είναι η παρακάτω ανισότητα, η οποία είναι συνέπεια του Θεωρήματος 1.2.10.

Θεώρημα 1.2.11. *Αν $\gamma_n(A) \geqslant 1/2$, τότε για κάθε $t > 0$*

$$1 - \gamma_n(A_t) \leqslant \frac{1}{2}\exp(-t^2/2).$$

Απόδειξη. Από το Θεώρημα 1.2.10 γνωρίζουμε ότι

$$1 - \gamma_n(A_t) \leqslant 1 - \gamma_n(H_t)$$

όπου H ημίχωρος μέτρου $1/2$. Αφού το γ_n είναι αναλλοίωτο ως προς ορθογώνιους μετασχηματισμούς, μπορούμε να υποθέσουμε ότι $H = \{x \in \mathbb{R}^n : x_1 \leqslant 0\}$, οπότε ολοκληρώνοντας πρώτα ως προς $x_2, \ldots, x_n$ βλέπουμε ότι

$$1 - \gamma_n(H_t) = \frac{1}{\sqrt{2\pi}}\int_t^{\infty} e^{-s^2/2}ds.$$

Παραγωγίζοντας δείχνουμε ότι η συνάρτηση

$$F(x) = e^{x^2/2}\int_x^{\infty} e^{-s^2/2}ds$$

είναι φθίνουσα στο $[0, +\infty)$. Από την $F(t) \leqslant F(0)$ προκύπτει το συμπέρασμα. □

1.2δ′ Η ισοπεριμετρική ανισότητα στον E_2^n

Ως t-περιοχή ενός $A \subseteq E_2^n$ θεωρούμε το σύνολο $A_t = \{x \in E_2^n : d_n(x, A) \leqslant t\}$ (παρατηρήστε ότι ο ορισμός μας είναι λίγο διαφορετικός από το γενικό ορισμό: επιτρέπουμε απόσταση μικρότερη ή ίση του t). Οι τιμές που μπορεί να πάρει η d_n είναι πεπερασμένες το πλήθος: $0, 1/n, 2/n, \ldots, 1$. Επομένως, αυτές είναι οι τιμές του t γιά τις οποίες η t-περιοχή του A παρουσιάζει ενδιαφέρον, με την έννοια ότι το A_t παραμένει αμετάβλητο όταν το t παίρνει τιμές σε ένα διάστημα της μορφής $[k/n, (k+1)/n)$.

Το ισοπεριμετρικό πρόβλημα είναι λοιπόν το εξής. Μας δίνουν ένα φυσικό $m = 1, 2, \ldots, 2^n$ και κάποιο $t = k/n$, $k = 1, \ldots, n$. Γιά ποιό σύνολο A με πλήθος στοιχείων m είναι η k/n-περιοχή του A η μικρότερη δυνατή; Η απάντηση είναι ότι το A θα πρέπει να έχει όσο το δυνατόν «λιγότερα κενά». Αν περιέχει μιά n-άδα $x = (x_1, \ldots, x_n)$, τότε θα πρέπει να περιέχει κατά σειρά προτεραιότητας και τις «γειτονικές» της n-άδες, αυτές δηλαδή που διαφέρουν από την x σε μία συντεταγμένη, δύο συντεταγμένες, κ.ο.κ. (εφόσον το πλήθος των στοιχείων του A επαρκεί). Αυτό, γιατί η παραμικρή επέκταση του A θα τις συμπεριλάβει ούτως ή άλλως. Τα πιό οικονομικά σύνολα είναι οι d_n-μπάλες (οι λεγόμενες Hamming μπάλες του E_2^n). Αποδεικνύεται η ακόλουθη ισοπεριμετρική ανισότητα γιά τον E_2^n.

Θεώρημα 1.2.12. *Έστω $A \subseteq E_2^n$ με $m = \sum_{k=0}^{l} \binom{n}{k}$ στοιχεία. Τότε, γιά κάθε $s = 1, \dots, n-l$, έχουμε*

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

όπου x τυχόν στοιχείο του E_2^n. □

Η ισοπεριμετρική αυτή ανισότητα οδηγεί σε μια προσεγγιστική ισοπεριμετρική ανισότητα για τον E_2^n.

Θεώρημα 1.2.13. *Αν $\mu_n(A) \geqslant 1/2$ και $t > 0$, τότε*

$$\mu_n(A_t^c) \leqslant \frac{1}{2} \exp(-2t^2 n).$$

Το Θεώρημα 1.2.13 ερμηνεύεται ως εξής: για να εκτιμήσουμε το $\mu_n(A_t)$ αρκεί να θέσουμε $l = n/2$ και $s = tn$ στο θεώρημα 1.2.12. Τότε βλέπουμε ότι

$$\mu_n(A_t^c) \leqslant \frac{1}{2^n} \sum_{j=(\frac{1}{2}+t)n}^{n} \binom{n}{j},$$

το οποίο φθίνει εκθετικά στο 0 καθώς $n \to \infty$, γιατί οι «ακραίοι» διωνυμικοί συντελεστές είναι πολύ μικροί σε σύγκριση με τους «μεσαίους» όταν το n είναι μεγάλο.

Η απόδειξη του Θεωρήματος 1.2.12 είναι συνδυαστική και συνήθως γίνεται με επαγωγή ως προς n.

1.3 Συνάρτηση συγκέντρωσης

Ορισμός 1.3.1 (συνάρτηση συγκέντρωσης)**.** *Έστω (X, d, μ) μετρικός χώρος πιθανότητας. Η συνάρτηση συγκέντρωσης του (X, d, μ) ορίζεται στο $(0, \infty)$ από την*

$$\alpha_\mu(t) := \sup\{1 - \mu(A_t) : \mu(A) \geqslant 1/2\}.$$

Η συνάρτηση α_μ είναι προφανώς φθίνουσα και, όπως δείχνει η επόμενη πρόταση, έχουμε πάντα $\lim_{t\to\infty} \alpha_\mu(t) = 0$.

Πρόταση 1.3.2. *Σε κάθε μετρικό χώρο πιθανότητας* (X, d, μ) *ισχύει*

$$\lim_{t\to\infty} \alpha_\mu(t) = 0.$$

Απόδειξη. Σταθεροποιούμε $x \in X$ και $0 < \varepsilon \leqslant 1/2$. Από το γεγονός ότι $X = \bigcup_{n=1}^{\infty} B(x,n)$ έπεται ότι υπάρχει $r \in \mathbb{N}$ ώστε

$$\mu(B(x,r)) > 1 - \varepsilon.$$

Τότε, για κάθε $A \in \mathcal{B}(X)$ με $\mu(A) \geqslant 1/2$ έχουμε $A \cap B(x,r) \neq \emptyset$, απ' όπου έπεται ότι $B(x,r) \subseteq A_{2r}$. [Πράγματι, υπάρχει $a \in A$ ώστε $d(x,a) < r$ και τότε, για κάθε $y \in B(x,r)$ έχουμε $d(y,a) \leqslant d(y,x) + d(x,a) < 2r$, δηλαδή $y \in A_{2r}$]. Τότε, για κάθε $t \geqslant 2r$ έχουμε

$$1 - \mu(A_t) \leqslant 1 - \mu(A_{2r}) \leqslant 1 - \mu(B(x,r)) < \varepsilon,$$

άρα $\alpha_\mu(t) \leqslant \varepsilon$. □

Ορισμός 1.3.3 (συγκέντρωση του μέτρου)**.** Λέμε ότι υπάρχει «συγκέντρωση του μέτρου» στο μετρικό χώρο πιθανότητας (X, d, μ) αν η $\alpha_\mu(t)$ φθίνει γρήγορα καθώς το $t \to \infty$. Πιο συγκεκριμένα:

(α) Λέμε ότι το μ έχει *κανονική συγκέντρωση* στον (X, d) με σταθερές $A, \beta > 0$ αν για κάθε $t > 0$ ισχύει

$$\alpha_\mu(t) \leqslant A e^{-\beta t^2}.$$

Τα αποτελέσματα της προηγούμενης παραγράφου (πιο συγκεκριμένα, οι εκτιμήσεις που προκύπτουν από τις λύσεις των αντίστοιχων ισοπεριμετρικών προβλημάτων) δείχνουν ότι αυτό ισχύει για τη σφαίρα, για το διακριτό κύβο και για το χώρο του Gauss. Σύμφωνα με τον Ορισμό 1.3.1 της συνάρτησης συγκέντρωσης, σε αυτούς τους χώρους έχουμε:

(i) Για τη σφαίρα (S^{n-1}, ρ, σ) ισχύει

$$\alpha_\sigma(t) \leqslant \sqrt{\pi/8}\, \exp(-t^2 n/2).$$

(ii) Για το χώρο του Gauss $(\mathbb{R}^n, |\cdot|, \gamma_n)$ ισχύει

$$\alpha_{\gamma_n}(t) \leqslant \frac{1}{2} \exp(-t^2/2).$$

(iii) Για το διακριτό κύβο (E_2^m, d_n, μ_n) ισχύει

$$\alpha_{\mu_n}(t) \leqslant \frac{1}{2}\exp(-2t^2 n).$$

(α) Λέμε ότι το μ έχει *εκθετική συγκέντρωση* στον (X,d) με σταθερές $A, \beta > 0$ αν για κάθε $t > 0$ ισχύει

$$\alpha_\mu(t) \leqslant Ae^{-\beta t}.$$

Πολλές από τις εφαρμογές της συγκέντρωσης του μέτρου βασίζονται στο εξής θεώρημα.

Θεώρημα 1.3.4. *Έστω (X,d,μ) μετρικός χώρος πιθανότητας. Αν $f : X \to \mathbb{R}$ είναι μια συνάρτηση Lipschitz με σταθερά 1, δηλαδή αν $|f(x) - f(y)| \leqslant d(x,y)$ για κάθε $x, y \in X$, τότε*

$$\mu\left(\{x \in X : |f(x) - \mathrm{med}(f)| \geqslant t\}\right) \leqslant 2\alpha_\mu(t)$$

για κάθε $t > 0$, όπου $\mathrm{med}(f)$ είναι ένας μέσος Lévy της f.

Σημείωση. Μέσος Lévy $\mathrm{med}(f)$ της f είναι ένας αριθμός για τον οποίο

$$\mu(\{f \geqslant \mathrm{med}(f)\}) \geqslant 1/2 \text{ και } \mu(\{f \leqslant \mathrm{med}(f)\}) \geqslant 1/2.$$

Κάθε μετρήσιμη συνάρτηση $f : X \to \mathbb{R}$ έχει τουλάχιστον ένα μέσο Lévy, ο οποίος μπορεί να μην είναι μονοσήμαντα ορισμένος.

Απόδειξη του Θεωρήματος 1.3.4. Θέτουμε $A = \{x : f(x) \geqslant \mathrm{med}(f)\}$ και $B = \{x : f(x) \leqslant \mathrm{med}(f)\}$. Αν $y \in A_t$ τότε υπάρχει $x \in A$ με $d(x,y) < t$, οπότε

$$f(y) = f(y) - f(x) + f(x) \geqslant -d(y,x) + \mathrm{med}(f) > \mathrm{med}(f) - t$$

αφού η f είναι 1-Lipschitz. Ομοίως, αν $y \in B_t$ τότε υπάρχει $x \in B$ με $d(x,y) < t$, οπότε

$$f(y) = f(y) - f(x) + f(x) \leqslant d(y,x) + \mathrm{med}(f) < \mathrm{med}(f) + t.$$

Δηλαδή, αν $y \in A_t \cap B_t$ τότε $|f(x) - \mathrm{med}(f)| < t$. Με άλλα λόγια,

$$\{x \in X : |f(x) - \mathrm{med}(f)| \geqslant t\} \subseteq (A_t \cap B_t)^c = A_t^c \cup B_t^c.$$

Όμως, από τον ορισμό της συνάρτησης συγκέντρωσης έχουμε $\mu(A_t) \geqslant 1 - \alpha_\mu(t)$ και $\mu(B_t) \geqslant 1 - \alpha_\mu(t)$. Έπεται ότι

$$\mu\left(\{|f - \mathrm{med}(f)| \geqslant t\}\right) \leqslant (1 - \mu(A_t)) + (1 - \mu(B_t)) \leqslant 2\alpha_\mu(t).$$

Πόρισμα 1.3.5. *Έστω (X,d,μ) μετρικός χώρος πιθανότητας. Αν $f:X\to\mathbb{R}$ είναι μια συνάρτηση Lipschitz με σταθερά $\|f\|_{\rm Lip}$, δηλαδή αν $|f(x)-f(y)|\leqslant \|f\|_{\rm Lip}d(x,y)$ για κάθε $x,y\in X$, τότε*

$$\mu\left(\{x\in X:|f(x)-{\rm med}(f)|\geqslant t\}\right)\leqslant 2\alpha_\mu(t/\|f\|_{\rm Lip})$$

για κάθε $t>0$, όπου ${\rm med}(f)$ είναι μέσος Lévy της f.

Στην περίπτωση που η συνάρτηση συγκέντρωσης φθίνει πολύ γρήγορα, το Θεώρημα 1.3.4 δείχνει ότι οι 1-Lipschitz συνεχείς συναρτήσεις είναι «σχεδόν σταθερές» σε «σχεδόν ολόκληρο το χώρο». Ισχύει μάλιστα και το αντίστροφο.

Θεώρημα 1.3.6. *Έστω (X,d,μ) μετρικός χώρος πιθανότητας. Αν για κάποιο $\eta>0$ και κάποιο $t>0$ ισχύει*

$$\mu\left(\{x\in X:|f(x)-{\rm med}(f)|\geqslant t\}\right)\leqslant \eta$$

για κάθε 1-Lipschitz συνάρτηση $f:X\to\mathbb{R}$, τότε $\alpha_\mu(t)\leqslant\eta$.

Απόδειξη. Έστω A Borel υποσύνολο του X με $\mu(A)\geqslant 1/2$. Θεωρούμε τη συνάρτηση $f(x)=d(x,A)$. Η f είναι 1-Lipschitz και ${\rm med}(f)=0$ γιατί η f παίρνει μη αρνητικές τιμές και $\mu(\{x:f(x)=0\})\geqslant 1/2$. Από την υπόθεση παίρνουμε

$$\mu(\{x\in X:d(x,A)\geqslant t\})\leqslant\eta,$$

δηλαδή $1-\mu(A_t)\leqslant\eta$. Έπεται ότι $\alpha_\mu(t)\leqslant\eta$. □

Μια άλλη ποσότητα που παίζει το ρόλο του μέσου για την $f:(X,d,\mu)\to\mathbb{R}$ είναι η *μέση τιμή*

$$\mathbb{E}_\mu(f)=\int_X f(x)\,d\mu(x).$$

Δεδομένου ότι μπορούμε να υπολογίσουμε ή έστω να εκτιμήσουμε ευκολότερα τη μέση τιμή από το μέσο Lévy, θα θέλαμε να έχουμε μια εκτίμηση αντίστοιχη με αυτήν του Θεωρήματος 1.4.3 στην οποία να εμφανίζεται η $\mathbb{E}_\mu(f)$ στη θέση του ${\rm med}(f)$. Το επόμενο θεώρημα εξασφαλίζει ότι αυτό ισχύει αν το μ έχει κανονική συγκέντρωση.

Θεώρημα 1.3.7. *Έστω $(X,d\mu)$ μετρικός χώρος πιθανότητας. Υποθέτουμε ότι υπάρχουν $A,\beta>0$ τέτοιες ώστε*

$$\alpha_\mu(t)\leqslant Ae^{-\beta t^2}$$

για κάθε $t>0$. Τότε, για κάθε 1-Lipschitz συνάρτηση $f:(X,d)\to\mathbb{R}$ και για κάθε $t>0$ έχουμε

$$\mu\left(\{x\in X:|f(x)-\mathbb{E}_\mu(f)|\geqslant t\}\right)\leqslant 2A\exp(-\beta t^2/8).$$

Απόδειξη. Θεωρούμε το μέτρο γινόμενο $\mu \otimes \mu$ στον $X \times X$ και παρατηρούμε ότι

$$\begin{aligned}(\mu \otimes \mu)(\{(x,y) : |f(x) - f(y)| \geqslant t\}) &\leqslant (\mu \otimes \mu)(\{(x,y) : |f(x) - \mathrm{med}(f)| \geqslant t/2\}) \\ &+ (\mu \otimes \mu)(\{(x,y) : |f(y) - \mathrm{med}(f)| \geqslant t/2\}) \\ &= \mu(\{x : |f(x) - \mathrm{med}(f)| \geqslant t/2\}) \\ &+ \mu(\{y : |f(y) - \mathrm{med}(f)| \geqslant t/2\}) \\ &\leqslant 2A\exp(-\beta t^2/4)\end{aligned}$$

για κάθε $t > 0$. Τότε, για κάθε $\lambda > 0$ έχουμε

$$\begin{aligned}&\int_X \int_X \big(\exp(\lambda^2(f(x) - f(y))^2)\big)\, d\mu(x)\, d\mu(y) \\ &\quad = 2\int_0^\infty \lambda^2 t e^{\lambda^2 t^2} (\mu \otimes \mu)(\{(x,y) : |f(x) - f(y)| \geqslant t\})\, dt \\ &\quad \leqslant 2A \int_0^\infty 2\lambda^2 t e^{\lambda^2 t^2 - \beta t^2/4} dt\end{aligned}$$

Αν επιλέξουμε $\lambda = \sqrt{\beta/8}$ παίρνουμε

$$\int_X \int_X \exp\Big(\frac{\beta}{8}(f(x) - f(y))^2\Big)\, d\mu(x)\, d\mu(y) \leqslant 2A\int_0^\infty \frac{\beta}{4} t e^{-\beta t^2/8}\, dt = 2A.$$

Η συνάρτηση $t \mapsto \exp(\beta t^2/8)$ είναι κυρτή, οπότε η ανισότητα Jensen μας δίνει

$$\begin{aligned}&\mathbb{E}_\mu \exp\Big(\frac{\beta}{8}(f(x) - \mathbb{E}_\mu(f))^2\Big)\, d\mu(x) \\ &\quad \leqslant \int_X \int_X \exp\Big(\frac{\beta}{8}(f(x) - f(y))^2\Big)\, d\mu(y)\, d\mu(x) \leqslant 2A.\end{aligned}$$

Από την ανισότητα Markov έπεται ότι

$$\begin{aligned}\mu(\{x : |f(x) - \mathbb{E}_\mu(f)| \geqslant t\}) &\leqslant e^{-\beta t^2/8}\mathbb{E}_\mu \exp\Big(\frac{\beta}{8}(f(x) - \mathbb{E}_\mu(f))^2\Big)\, d\mu(x) \\ &\leqslant 2Ae^{-\beta t^2/8}\end{aligned}$$

για κάθε $t > 0$. □

1.4 Προσεγγιστικές ισοπεριμετρικές ανισότητες

1.4α΄ Η συνάρτηση συγκέντρωσης της σφαίρας

Η απόδειξη του Θεωρήματος 1.2.9 βασίζεται πολύ ισχυρά στην σφαιρική ισοπεριμετρική ανισότητα. Για τις περισσότερες όμως εφαρμογές που έχουμε στο νού μας είναι αρκετή μια ανισότητα της μορφής

$$\sigma(A_t) \geqslant 1 - c_1 \exp(-c_2 t^2 n)$$

και όχι η ακριβής λύση του ισοπεριμετρικού προβλήματος. Θα δώσουμε μια απλή απόδειξη εκτίμησης «αυτού του τύπου» χωρίς να περάσουμε μέσα από την ισοπεριμετρική ανισότητα, χρησιμοποιώντας την ανισότητα Brunn-Minkowski.

Λήμμα 1.4.1. *Θεωρούμε το ομοιόμορφο μέτρο πιθανότητας μ στην Ευκλείδεια μοναδιαία μπάλα B_2^n. Δηλαδή, $\mu(A) = |A|/|B_2^n|$ για κάθε Borel σύνολο $A \subseteq B_2^n$. Αν $A, C \subseteq B_2^n$ είναι Borel σύνολα και*

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

τότε

$$\min\{\mu(A), \mu(C)\} \leqslant \exp(-\rho^2 n/8).$$

Απόδειξη. Θεωρούμε το σύνολο $\frac{A+C}{2}$. Από την ανισότητα Brunn–Minkowski παίρνουμε $\left|\frac{A+C}{2}\right| \geqslant \min\{|A|,|C|\}$. Συνεπώς,

$$\mu\left(\frac{A+C}{2}\right) \geqslant \min\{\mu(A), \mu(C)\}.$$

Από την άλλη πλευρά, αν $a \in A$ και $c \in C$, ο κανόνας του παραλληλογράμμου δίνει

$$|a+c|^2 = 2|a|^2 + 2|c|^2 - |a-c|^2 \leqslant 4 - \rho^2,$$

επομένως

$$\frac{A+C}{2} \subseteq \sqrt{1 - \frac{\rho^2}{4}} B_2^n.$$

Συνδυάζοντας τα παραπάνω βλέπουμε ότι

$$\min\{\mu(A), \mu(C)\} \leqslant \mu\left(\frac{A+C}{2}\right) \leqslant \left(1 - \frac{\rho^2}{4}\right)^{n/2} \leqslant \exp(-\rho^2 n/8).$$

□

Απόδειξη του Θεωρήματος 1.2.9: Έστω $A \subseteq S^{n-1}$ με $\sigma(A) = 1/2$ και έστω $t > 0$. Θέτουμε $C = S^{n-1} \setminus A_t$ και θεωρούμε τα υποσύνολα

$$A_1 = \{\rho a : a \in A, 1/2 \leqslant \rho \leqslant 1\} \text{ και } C_1 = \{\rho a : a \in C, 1/2 \leqslant \rho \leqslant 1\}$$

της B_2^n. Εύκολα ελέγχουμε ότι

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

Από το Λήμμα 1.4.1 συμπεραίνουμε ότι

$$|C_1| \leqslant \exp(-d^2n/8)|B_2^n| \leqslant \exp\left(-t^2n/(8\pi^2)\right)|B_2^n|.$$

Όμως, από τον ορισμό του σ έχουμε $|B_2^n|\sigma(C) = |\tilde{C}|$ και $|C_1| = (1-2^{-n})|\tilde{C}|$. Έπεται ότι

$$\sigma(A_t^c) = \sigma(C) \leqslant \frac{1}{1-2^{-n}} \exp\left(-t^2n/(8\pi^2)\right).$$

Δηλαδή,

$$\sigma(A_t) \geqslant 1 - c_1\exp(-c_2t^2n)$$

όπου $c_1 = 2$ και $c_2 = 1/(8\pi^2)$. Αυτή είναι η ανισότητα του Θεωρήματος 1.2.9 αν εξαιρέσουμε τις ακριβείς τιμές των σταθερών c_1 και c_2. □

1.4β′ Η συνάρτηση συγκέντρωσης του χώρου του Gauss

Όπως και στην περίπτωση της σφαίρας, η απόδειξη της προσεγγιστικής ισοπεριμετρικής ανισότητας του Πορίσματος 1.2.11 χρησιμοποιεί ισχυρά την ισοπεριμετρική ανισότητα του Θεωρήματος 1.2.10. Μπορούμε όμως να αποδείξουμε απευθείας την προσεγγιστική ισοπεριμετρική ανισότητα γιά το χώρο του Gauss.

Θεώρημα 1.4.2. *Εστω A μη κενό Borel υποσύνολο του $\mathbb{R}^n$. Τότε,*

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

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

$$1 - \gamma_n(A_t) \leqslant 2\exp(-t^2/4)$$

γιά κάθε $t > 0$.

Απόδειξη. Συμβολίζουμε με $g_n(x)$ την συνάρτηση $(2\pi)^{-n/2}\exp(-|x|^2/2)$, και θεωρούμε τις συναρτήσεις

$$f(x) = e^{d(x,A)^2/4}g_n(x) \ \ , \ \ g(x) = \chi_A(x)g_n(x) \ \ , \ \ m(x) = g_n(x).$$

Γιά κάθε $x \in \mathbb{R}^n$ και $y \in A$ έχουμε

$$\begin{aligned}(2\pi)^n f(x)g(y) &= e^{d(x,A)^2/4} e^{-|x|^2/2} e^{-|y|^2/2} \\ &\leqslant \exp\left(\frac{|x-y|^2}{4} - \frac{|x|^2}{2} - \frac{|y|^2}{2}\right) \\ &= \exp\left(-\frac{|x+y|^2}{4}\right) \\ &= \left(\exp\left(-\frac{1}{2}\left|\frac{x+y}{2}\right|^2\right)\right)^2 \\ &= (2\pi)^n \left(m\left(\frac{x+y}{2}\right)\right)^2,\end{aligned}$$

όπου χρησιμοποιήσαμε τον κανόνα του παραλληλογράμμου και την $d(x,A) \leqslant |x-y|$. Παρατηρώντας ότι $g(y) = 0$ αν $y \notin A$, βλέπουμε ότι οι f, g, m ικανοποιούν τις υποθέσεις της ανισότητας Prékopa-Leindler με $\lambda = 1/2$. Εφαρμόζουμε λοιπόν το Θεώρημα 1.2.4 και έχουμε

$$\begin{aligned}\left(\int_{\mathbb{R}^n} e^{d(x,A)^2/4} g_n(dx)\right)\gamma_n(A) &= \left(\int_{\mathbb{R}^n} f(x)d\gamma_n(x)\right)\left(\int_{\mathbb{R}^n} g(x)d\gamma_n(x)\right) \\ &\leqslant \left(\int_{\mathbb{R}^n} m(x)d\gamma_n(x)\right)^2 = 1.\end{aligned}$$

Αυτό αποδεικνύει τον πρώτο ισχυρισμό του θεωρήματος. Γιά τον δεύτερο, παρατηρούμε ότι αν $\gamma_n(A) = 1/2$ τότε

$$e^{t^2/4}\gamma_n(x : d(x,A) \geqslant t) \leqslant \int_{\mathbb{R}^n} e^{d(x,A)^2/4} d\gamma_n(x) \leqslant \frac{1}{\gamma_n(A)} = 2.$$

Δηλαδή, $\gamma_n(A_t^c) \leqslant 2\exp(-t^2/4)$. □

1.4γ΄ Η συνάρτηση συγκέντρωσης του διακριτού κύβου

Θεωρούμε το σύνολο $E_2^n = \{-1,1\}^n$, το οποίο ταυτίζουμε με το σύνολο των κορυφών του κύβου $Q_n = [-1,1]^n$ στον $\mathbb{R}^n$. Στο E_2^n ορίζουμε το κανονικό μέτρο πιθανότητας μ_n που δίνει μάζα 2^{-n} σε κάθε σημείο. Για κάθε μη κενό $A \subseteq E_2^n$ θέτουμε

$$\phi_A(x) = \inf\{|x-y| : y \in \mathrm{conv}(A)\}.$$

Το βασικό θεώρημα αυτής της παραγράφου είναι το εξής.

Θεώρημα 1.4.3. *Για κάθε* $A \subseteq E_2^n$,

$$\mathbb{E}\exp(\phi_A^2/8) \leqslant \frac{1}{\mu_n(A)}.$$

Απόδειξη. Με επαγωγή ως προς το πλήθος των σημείων του A. Αν $\mathrm{card}(A) = 1$ δηλαδή $A = \{y\}$, τότε

$$\phi_A(x) = \inf\{|x-z| : z \in \mathrm{conv}(A) = \{y\}\} = |x-y|.$$

Άρα,

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

Κάθε $x \in E_2^n$ διαφέρει από το y σε i θέσεις, $i = 0, 1, \ldots, n$. Το πλήθος των $x \in E_2^n$ που διαφέρουν σε i θέσεις από το y είναι $\binom{n}{i}$. Παρατηρούμε ότι $|x-y|^2 = 4i$ όταν το x διαφέρει απο το y σε i θέσεις. Άρα,

$$\begin{aligned}\mathbb{E}e^{|x-y|^2/8} &= \frac{1}{2^n}\sum_{i=0}^{n}\binom{n}{i}e^{i/2} = \frac{1}{2^n}\left(1+e^{1/2}\right)^n \\ &= \left(\frac{1+e^{1/2}}{2}\right)^n \leqslant 2^n = \frac{1}{\mu_n(A)},\end{aligned}$$

αφού $e^{1/2} \leqslant e \leqslant 3$.

Έστω τώρα ότι $\mathrm{card}(A) \geqslant 2$. Εξετάζουμε πρώτα την περίπτωση $n = 1$. Αναγκαστικά έχουμε $A = E_1$, επομένως $\phi_A(x) = 0$ για κάθε $x \in E_1$. Άρα,

$$\mathbb{E}\exp(\phi_A^2/8) = \mathbb{E}e^0 = 1 = 1/\mu_1(A).$$

Για το επαγωγικό βήμα θεωρούμε $A \subseteq E_{n+1}$ με $\mathrm{card}(A) \geqslant 2$. Χωρίς περιορισμό της γενικότητας μπορούμε να υποθέσουμε ότι

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

όπου $A_1, A_{-1} \neq \emptyset$. Μπορούμε επίσης να υποθέσουμε ότι $\mathrm{card}(A_{-1}) \leqslant \mathrm{card}(A_1)$.

Λήμμα 1.4.4. *Για κάθε* $x \in E_2^n$,

$$\phi_A((x,1)) \leqslant \phi_{A_1}(x).$$

Απόδειξη. Αρκεί να δείξουμε ότι

$$\{|x-y| : y \in \mathrm{conv}(A_1)\} \subseteq \{|(x,1)-z| : z \in \mathrm{conv}(A)\}.$$

Έστω $y \in \mathrm{conv}(A_1)$. Τότε, $y = \sum_{i=1}^m t_i x_i$ όπου $t_i \geqslant 0$ με $\sum_{i=1}^m t_i = 1$ και $x_i \in A_1$. Τότε όμως $(x_i, 1) \in A$ και

$$\sum_{i=1}^m t_i(x_i,1) = \Big(\sum_{i=1}^m t_i x_i, \sum_{i=1}^m t_i\Big) = (y,1),$$

δηλαδή $(y,1) \in \mathrm{conv}(A)$. Αφού $|x-y| = |(x,1)-(y,1)|$ και

$$|(x,1)-(y,1)| \in \{|(x,1)-z| : z \in \mathrm{conv}(A)\},$$

έχουμε το ζητούμενο. □

Λήμμα 1.4.5. *Για κάθε $x \in E_2^n$ και για κάθε $0 \leqslant a \leqslant 1$,*

$$\phi_A^2((x,-1)) \leqslant 4a^2 + a\phi_{A_1}^2(x) + (1-a)\phi_{A_{-1}}^2(x).$$

Απόδειξη. Έστω $z_i \in \mathrm{conv}(A_i)$ $(i = 1, -1)$. Τότε, όπως προηγουμένως, $(z_i, i) \in \mathrm{conv}(A)$. Το $\mathrm{conv}(A)$ είναι κυρτό, άρα

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

Έχουμε

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

Αφού τα $z_i \in \mathrm{conv}(A_i)$ ήταν τυχόντα, έπεται ότι

$$\phi_A^2(x,-1) \leqslant a\phi_{A_1}^2(x) + (1-a)\phi_{A_{-1}}^2(x) + 4a^2. \qquad \square$$

Χρησιμοποιώντας τα δύο λήμματα, γράφουμε

$$\begin{aligned}
\mathbb{E}\left(e^{\phi_A^2/8}\right) &= \frac{1}{2^{n+1}} \sum_{x\in E_{n+1}} e^{\phi_A^2(x)/8} \\
&= \frac{1}{2^{n+1}} \sum_{x\in E_n} e^{\phi_A^2((x,1))/8} + \frac{1}{2^{n+1}} \sum_{x\in E_n} e^{\phi_A^2((x,-1))/8} \\
&\leqslant \frac{1}{2^{n+1}} \sum_{x\in E_n} e^{\phi_{A_1}^2(x)/8} + \frac{1}{2^{n+1}} e^{a^2/2} \sum_{x\in E_n} e^{a\phi_{A_1}^2(x)/8+(1-a)\phi_{A_{-1}}^2(x)/8} \\
&\leqslant \frac{1}{2^{n+1}} \sum_{x\in E_n} e^{\phi_{A_1}^2(x)/8} \\
&+ \frac{1}{2^{n+1}} e^{a^2/2} \left(\sum_{x\in E_n} e^{\phi_{A_1}^2(x)/8}\right)^a \left(\sum_{x\in E_n} e^{\phi_{A_{-1}}^2(x)/8}\right)^{1-a} \\
&= \frac{1}{2}\mathbb{E}(e^{\phi_{A_1}^2/8}) + \frac{1}{2} e^{a^2/2} \left(\mathbb{E}(e^{\phi_{A_1}^2/8})\right)^a \left(\mathbb{E}(e^{\phi_{A_{-1}}^2/8})\right)^{1-a}.
\end{aligned}$$

Θέτουμε

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

και

$$u_{-1} = \mathbb{E}\left(e^{\phi_{A_{-1}}^2/8}\right) \quad , \quad v_{-1} = \frac{1}{\mu_n(A_{-1})}.$$

Από την επαγωγική υπόθεση έχουμε $u_1 \leqslant v_1$ και $u_{-1} \leqslant v_{-1}$ (επίσης, η $\mathrm{card}(A_{-1}) \leqslant \mathrm{card}(A_1)$ γράφεται $v_1 \leqslant v_{-1}$). Άρα, η προηγούμενη ανισότητα παίρνει τη μορφή

$$\begin{aligned}
\mathbb{E}e^{\phi_A^2/8} &\leqslant \frac{1}{2}u_1 + \frac{1}{2}e^{a^2/2}(u_1)^a(u_{-1})^{1-a} \\
&\leqslant \frac{1}{2}u_1 + \frac{1}{2}e^{a^2/2}(v_1)^a(v_{-1})^{1-a} \\
&\leqslant \frac{v_1}{2}[1 + e^{a^2/2}(v_1/v_{-1})^{a-1}].
\end{aligned}$$

Η τελευταία ποσότητα ελαχιστοποιείται όταν $a = -\ln(v_1/v_{-1})$. Η τιμή $-\ln(v_1/v_{-1})$ είναι περίπου ίση με $1-v_1/v_{-1}$. Επιλέγουμε $a_0 = 1-v_1/v_{-1}$. Αφού $v_1 \leqslant v_{-1}$ έχουμε $0 \leqslant a_0 \leqslant 1$, οπότε μπορούμε να γράψουμε

$$\mathbb{E}(e^{\phi_A^2/8}) \leqslant \frac{v_1}{2}[1 + e^{a_0^2/2}(1-a_0)^{a_0-1}].$$

Λήμμα 1.4.6. *Για κάθε* $0 \leqslant a \leqslant 1$ *έχουμε*

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

Απόδειξη. Απλές πράξεις δείχνουν ότι η ανισότητα που ζητάμε είναι ισοδύναμη με την

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

Παραγωγίζοντας βλέπουμε ότι $g'' \geqslant 0$ και $g'(0) = 0$. Άρα η g είναι αύξουσα στο $[0,1]$. Αφού $g(0) = 0$, έπεται το ζητούμενο. □

Χρησιμοποιώντας το Λήμμα 1.4.6 έχουμε

$$\begin{aligned}\mathbb{E}(e^{\phi_A^2/8}) &\leqslant \frac{v_1}{2}\frac{4}{2-a_0} = \frac{2v_1}{1+v_1/v_{-1}} \\ &= \frac{2}{1/v_1 + 1/v_{-1}} = \frac{2}{\mu_n(A_1)+\mu_n(A_{-1})} \\ &= \frac{1}{\mu_{n+1}(A)}.\end{aligned}$$

Η τελευταία ισότητα είναι φανερή αφού $\mu_{n+1}(A_i \times \{i\}) = \mu_n(A_i)/2$, $i = \pm 1$. Έτσι ολοκληρώνονται το επαγωγικό βήμα και η απόδειξη του Θεωρήματος 1.4.3. □

Πόρισμα 1.4.7. *Για όλα τα $t > 0$, έχουμε*

$$\mu_n(\phi_A \geqslant t) \leqslant \frac{1}{\mu_n(A)} e^{-t^2/8}.$$

Απόδειξη. Από το Θεώρημα 1.4.3 έχουμε $\mathbb{E}\exp(\phi_A^2/8) \leqslant \frac{1}{\mu_n(A)}$. Άρα,

$$\begin{aligned}e^{t^2/8}\mu_n(\phi_A \geqslant t) &\leqslant \int_{\{\phi_A \geqslant t\}} e^{t^2/8} \leqslant \int_{\{\phi_A \geqslant t\}} e^{\phi_A^2/8} \\ &\leqslant \mathbb{E}(e^{\phi_A^2/8}) \leqslant \frac{1}{\mu_n(A)}.\end{aligned}$$

Έστω A μη κενό υποσύνολο του E_2^n. Η συνάρτηση ϕ_A του Θεωρήματος 1.4.3 και η συνάρτηση απόστασης από το A

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

συγκρίνονται σύμφωνα με το επόμενο λήμμα.

Λήμμα 1.4.8. *Γιά κάθε μη κενό $A \subseteq E_2^n$ έχουμε*

$$2\sqrt{n}\, d_n(x,A) \leqslant \phi_A(x) \quad , \quad x \in E_2^n. \tag{1.4.1}$$

Απόδειξη. Έστω $x \in E_2^n$. Γιά κάθε $y \in A$ ισχύει

$$\langle x-y, x\rangle = \sum_{i=1}^{n} x_i(x_i-y_i) = 2nd_n(x,y) \geqslant 2nd_n(x,A). \tag{1.4.2}$$

Από την (1.4.2) έπεται ότι γιά κάθε $y \in \mathrm{conv}(A)$

$$\sqrt{n}|x-y| \geqslant \langle x-y, x\rangle \geqslant 2nd_n(x,A).$$

Αυτό αποδεικνύει την (1.4.1). □

Συνδυάζοντας τα δύο παραπάνω λήμματα δείχνουμε την προσεγγιστική ισοπεριμετρική ανισότητα γιά τον E_2^n:

Θεώρημα 1.4.9. *Έστω $A \subseteq E_2^n$ με $\mu_n(A) = 1/2$. Τότε, γιά κάθε $t > 0$ έχουμε*

$$\mu_n(A_t) \geqslant 1 - 2\exp(-t^2 n/2).$$

Απόδειξη. Αν $x \notin A_t$, τότε $d_n(x,A) \geqslant t$ και το Λήμμα 1.4.8 δείχνει ότι $\phi_A(x) \geqslant 2t\sqrt{n}$.

Ομως, από το Θεώρημα 1.4.3 έχουμε

$$e^{t^2n/2}\mu_n(x : \phi_A(x) \geqslant 2t\sqrt{n}) \leqslant \int_{E_2^n} \exp(\phi_A^2(x)/8) d\mu_n(x) \leqslant \frac{1}{\mu_n(A)} = 2,$$

το οποίο σημαίνει ότι

$$\mu_n(A_t^c) \leqslant \mu_n(x : \phi_A(x) \geqslant 2t\sqrt{n}) \leqslant 2\exp(-t^2n/2). \qquad \square$$

Το Θεώρημα 1.4.9 έχει σαν συνέπεια την συγκέντρωση των κυρτών Lipschitz συναρτήσεων γύρω από τον μέσο Lévy τους.

Θεώρημα 1.4.10. *Θεωρούμε μια κυρτή Lipschitz συνάρτηση $f : \mathbb{R}^n \to \mathbb{R}$ με σταθερά Lipschitz σ. Έστω M ένας μέσος Lévy της f στον E_2^n. Τότε, για κάθε $t > 0$ έχουμε*

$$\mu_n(\{|f-M| \geqslant t\}) \leqslant 4e^{-t^2/8\sigma^2}.$$

Απόδειξη. Για τον M ισχύουν οι $\mu_n(\{f \geqslant M\}) \geqslant 1/2$ και $\mu_n(\{f \leqslant M\}) \geqslant 1/2$.

Θέτουμε $A = \{f \leqslant M\}$. Αφού η f είναι κυρτή, για κάθε $y \in \mathrm{conv}A$ έχουμε $f(y) \leqslant M$. Αν λοιπόν $f(x) \geqslant M + t$ για κάποιο $x \in E_2^n$, τότε

$$f(x) \geqslant M + t \geqslant f(y) + t$$

για κάθε $y \in \mathrm{conv}(A)$. Άρα, $\sigma|x-y| \geqslant |f(x)-f(y)| \geqslant t$. Αυτό σημαίνει ότι

$$\phi_A(x) \geqslant t/\sigma.$$

Από το Πόρισμα 1.4.7 και από την $\mu_n(A) \geqslant 1/2$ έχουμε

$$\begin{aligned}\mu_n(\{f \geqslant M+t\}) &\leqslant \mu_n(\{\phi_A \geqslant t/\sigma\}) \\ &\leqslant \frac{1}{\mu_n(A)} e^{-t^2/8\sigma^2} \leqslant 2e^{-t^2/8\sigma^2}.\end{aligned}$$

Έστω $t>0$ και $B = \{f \leqslant M-t\}$. Όπως πριν ελέγχουμε ότι

$$f(x) \geqslant M \implies \phi_B(x) \geqslant t/\sigma$$

και με χρήση του Πορίσματος 1.4.7 έχουμε

$$\mu_n(\{f(x) \geqslant M\}) \leqslant \mu_n(\{\phi_B \geqslant t/\sigma\}) \leqslant \frac{1}{\mu_n(B)} e^{-t^2/8\sigma^2}.$$

Όμως $1/2 \leqslant \mu_n(\{f(x) \geqslant M\})$, άρα

$$\mu_n(B) \leqslant 2e^{-t^2/8\sigma^2}.$$

Συνδυάζοντας τα παραπάνω βλέπουμε ότι

$$\begin{aligned}\mu_n(\{|f-M| \geqslant t\}) &= \mu_n(\{f \geqslant M+t\}) + \mu_n(\{f \leqslant M-t\}) \\ &\leqslant 2e^{-t^2/8\sigma^2} + 2e^{-t^2/8\sigma^2} = 4e^{-t^2/8\sigma^2}.\end{aligned}$$

1.5 Εφαρμογή: το θεώρημα του Dvoretzky

Σε αυτή την παράγραφο περιγράφουμε συνοπτικά την απόδειξη του κλασικού θεωρήματος του Dvoretzky. Ιστορικά, αυτή είναι η πρώτη μεγάλη εφαρμογή του «φαινομένου» της συγκέντρωσης του μέτρου στη σφαίρα.

Έστω K συμμετρικό (ως προς την αρχή των αξόνων) κυρτό σώμα στον $\mathbb{R}^n$. Η απεικόνιση $\|x\| = \min\{\lambda \geqslant 0 : x \in \lambda K\}$ είναι νόρμα στον $\mathbb{R}^n$. Ο $\mathbb{R}^n$ εφοδιασμένος με τη νόρμα $\|\cdot\|$ θα συμβολίζεται με X_K. Αντίστροφα, αν $X = (\mathbb{R}^n, \|\cdot\|)$ είναι ένας χώρος με νόρμα, τότε η μοναδιαία του μπάλα $K_X = \{x \in \mathbb{R}^n : \|x\| \leqslant 1\}$ είναι συμμετρικό κυρτό σώμα στον $\mathbb{R}^n$. Η δυϊκή νόρμα $\|\cdot\|_*$ της $\|\cdot\|$ ορίζεται από την

$$\|y\|_* = \max\{|\langle x, y\rangle| : \|x\| \leqslant 1\}.$$

Έστω X και Y δύο χώροι με νόρμα. Ένας γραμμικός τελεστής $T : X \to Y$ λέγεται *ισομορφισμός* αν είναι γραμμικός, ένα προς ένα και επί τελεστής, και οι $T : X \to Y$, $T^{-1} : Y \to X$ είναι φραγμένοι τελεστές. Ο $T : X \to Y$ λέγεται *ισομετρικός ισομορφισμός* αν είναι ισομορφισμός και, επιπλέον, για κάθε $x \in X$ ισχύει $\|T(x)\| = \|x\|$. Δύο χώροι X και Y με νόρμα λέγονται *ισομετρικά ισόμορφοι* αν υπάρχει ισομετρικός ισομορφισμός $T : X \to Y$. Είναι γνωστό ότι αν X είναι ένας n-διάστατος χώρος με νόρμα τότε μπορούμε να ορίσουμε νόρμα $\|\cdot\|'$ στον $\mathbb{R}^n$ έτσι ώστε ο X να είναι ισομετρικά ισόμορφος με τον $(\mathbb{R}^n, \|\cdot\|')$.

Η έννοια της απόστασης Banach-Mazur εμφανίζεται στο βιβλίο του Banach «Théorie des opérations linéaires» (1932). Έστω X και Y δύο χώροι με νόρμα, άπειρης ενδεχομένως διάστασης, και ας υποθέσουμε ότι ο X είναι ισόμορφος με τον Y (γράφουμε $X \sim Y$). Ορίζουμε την *απόσταση Banach-Mazur* των X και Y ως εξής:

$$d(X,Y) := \inf\{\|T\| \cdot \|T^{-1}\| \;\big|\; T : X \to Y \text{ ισομορφισμός}\}.$$

Αν οι X και Y δεν είναι ισόμορφοι ($X \not\sim Y$), θέτουμε $d(X,Y) = +\infty$. Οι βασικές ιδιότητες της απόστασης Banach-Mazur περιγράφονται στην επόμενη Πρόταση.

Πρόταση 1.5.1. *Έστω X, Y, Z χώροι με νόρμα. Τότε,*

(i) $d(X,Y) \geq 1$.

(ii) $d(X,Y) = d(Y,X)$.

(iii) $d(X,Y) \leqslant d(X,Z)\, d(Z,Y)$.

(iv) *Αν οι X και Y είναι αυτοπαθείς, τότε $d(X^*, Y^*) = d(X,Y)$.*

Η γεωμετρική ερμηνεία της απόστασης Banach-Mazur δίνεται από την παρακάτω σχέση:

$$d(X,Y) = \inf\{d > 0 \mid \exists T : X \to Y \;:\; B_Y \subseteq T(B_X) \subseteq d\, B_Y\}.$$

Στην περίπτωση που $\dim X = \dim Y = n$, ένα απλό επιχείρημα συμπάγειας δείχνει ότι υπάρχει ισομορφισμός $T : X \to Y$ ώστε $d(X,Y) = \|T\|\,\|T^{-1}\|$. Ειδικότερα, $d(X,Y) = 1$ αν και μόνο αν ο X είναι ισομετρικά ισόμορφος με τον Y.

Το θεώρημα του Dvoretzky, το οποίο απαντά καταφατικά σε ένα ερώτημα του Grothendieck, διατυπώνεται ως εξής.

Θεώρημα 1.5.2 (Dvoretzky). *Έστω $\varepsilon > 0$ και έστω k φυσικός αριθμός. Υπάρχει $N = N(k,\varepsilon)$ με την εξής ιδιότητα: Αν X είναι χώρος με νόρμα διάστασης $n \geqslant N$, μπορούμε να βρούμε k-διάστατο υπόχωρο F του X με $d(F, \ell_2^k) \leqslant 1 + \varepsilon$.*

Σε γεωμετρική γλώσσα, το θεώρημα του Dvoretzky μας λέει ότι: για κάθε $k \in \mathbb{N}$, κάθε συμμετρικό κυρτό σώμα αρκετά μεγάλης διάστασης έχει κεντρικές τομές διάστασης k που είναι σχεδόν ελλειψοειδή. Η ακριβής εξάρτηση του $N(k,\varepsilon)$ από τα k και ε μελετήθηκε συστηματικά, και το θεώρημα του Dvoretzky πήρε πολύ πιο συγκεκριμένη ποσοτική μορφή.

Θεώρημα 1.5.3. *Έστω X ένας n-διάστατος χώρος με νόρμα και έστω $\varepsilon \in (0,1)$. Υπάρχουν ακέραιος $k \geqslant c\varepsilon^2(\log 1/\varepsilon)^{-1}\log n$ και k-διάστατος υπόχωρος F του X ο οποίος ικανοποιεί την $d(F,\ell_2^k) \leqslant 1+\varepsilon$.*

Δηλαδή, το Θεώρημα 1.5.2 ισχύει με $N(k,\varepsilon) = \exp(c\varepsilon^{-2}|\log\varepsilon|k)$.Η αρχική απόδειξη του Dvoretzky έδινε την εκτίμηση $N(k,\varepsilon) = \exp(c\varepsilon^{-2}k^2\log k)$. Η (βέλτιστη ως προς n) εκτίμηση του Θεωρήματος 1.5.3 αποδείχτηκε από τον Milman (1971). Για την απόδειξη του Θεωρήματος 1.5.2 θα χρειαστούμε τη σφαιρική ιοπεριμετρική ανισότητα και τα θεωρήματα των John και Dvoretzky-Rogers τα οποία συζητάμε στην επόμενη υποπαράγραφο.

1.5α′ Το θεώρημα του John και το λήμμα των Dvoretzky και Rogers

Το θεώρημα του F. John (1948) δίνει ακριβές άνω φράγμα για την απόσταση Banach-Mazur τυχόντος n-διάστατου χώρου με νόρμα από τον ℓ_2^n.

Θεώρημα 1.5.4 (John). *Για κάθε n-διάστατο χώρο με νόρμα X έχουμε $d(X,\ell_2^n) \leqslant \sqrt{n}$. Ακριβέστερα, αν $X = (\mathbb{R}^n, \|\cdot\|)$ και αν η μοναδιαία Ευκλείδεια μπάλα B_2^n είναι το ελλειψοειδές μέγιστου όγκου που περιέχεται στη μοναδιαία μπάλα K_X του X, τότε $B_2^n \subseteq \sqrt{n}K_X$. Ισοδύναμα, για κάθε $x \in \mathbb{R}^n$ έχουμε*

$$\frac{1}{\sqrt{n}}\|x\| \leqslant |x| \leqslant \|x\|.$$

Θα χρειαστούμε ένα ακόμα αποτέλεσμα για τη σχέση ενός συμμετρικού κυρτού σώματος με το ελλειψοειδές μέγιστου όγκου του. Αποτελέσματα αυτού του είδους αποδείχθηκαν από τους Dvoretzky και Rogers (1950). Το συγκεκριμένο μας λέει χοντρικά ότι αν η B_2^n είναι το ελλειψοειδές μέγιστου όγκου της K_X τότε υπάρχουν «πολλές» κάθετες ανά δύο διευθύνσεις στις οποίες οι $|\cdot|$ και $\|\cdot\|$ συγκρίνονται καλά.

Θεώρημα 1.5.5 (Dvoretzky-Rogers). *Έστω K συμμετρικό κυρτό σώμα στον $\mathbb{R}^n$. Υποθέτουμε ότι η Ευκλείδεια μοναδιαία μπάλα B_2^n είναι το ελλειψοειδές μέγιστου όγκου που περιέχεται στο K. Τότε, υπάρχει ορθοκανονική βάση $\{x_1,\dots,x_n\}$ του $\mathbb{R}^n$ ώστε*

$$1 = \|x_i\| \geqslant \frac{1}{4}$$

για κάθε $i = 1,\dots,\lceil\frac{n}{2}\rceil$.

Θεωρούμε τη *μέση τιμή* της $r(x) = \|x\|$ στην S^{n-1}:

$$M = \int_{S^{n-1}} \|x\| \, d\sigma(x).$$

Συνέπεια του Θεωρήματος 1.5.5 είναι το ακόλουθο κάτω φράγμα για την παράμετρο M.

Πρόταση 1.5.6. *Έστω K συμμετρικό κυρτό σώμα στον $\mathbb{R}^n$. Υποθέτουμε ότι η Ευκλείδεια μοναδιαία μπάλα B_2^n είναι το ελλειψοειδές μέγιστου όγκου που περιέχεται στο K. Τότε,*

$$M = \int_{S^{n-1}} \|x\| \, d\sigma(x) \geqslant c_1 \sqrt{\frac{\log n}{n}},$$

όπου $c_1 > 0$ είναι μια απόλυτη σταθερά.

1.5β′ Απόδειξη του θεωρήματος του Dvoretzky

Για την απόδειξη του Θεωρήματος 1.5.3 μπορούμε να υποθέσουμε ότι ο χώρος X είναι ο $\mathbb{R}^n$ εφοδιασμένος με μια νόρμα $\|\cdot\|$ και ότι το ελλειψοειδές μέγιστου όγκου που περιέχεται στη μοναδιαία μπάλα K_X του X είναι Ευκλείδεια μοναδιαία μπάλαB_2^n. Συνεπώς, ισχύουν οι

$$\frac{1}{\sqrt{n}}|x| \leqslant \|x\| \leqslant |x|, \qquad x \in \mathbb{R}^n$$

και

$$M = \int_{S^{n-1}} \|x\| \, d\sigma(x) \geqslant c_1 \sqrt{\frac{\log n}{n}}.$$

Από την πρώτη ανισότητα βλέπουμε ότι η συνάρτηση $r : S^{n-1} \to \mathbb{R}$ με $r(x) = \|x\|$, είναι Lipschitz συνεχής με σταθερά 1. Από το Θεώρημα 1.2.9 και το Θεώρημα 1.3.7 συμπεραίνουμε ότι για κάθε $\eta \in (0,1)$ ισχύει

$$\sigma\left(\{x \in S^{n-1} : | \, \|x\| - M \, | \geqslant \eta M\}\right) \leqslant c_2 \exp(-c_3 \eta^2 M^2 n). \tag{1.5.1}$$

Λήμμα 1.5.7. *Υπάρχει $n_0(\varepsilon) \in \mathbb{N}$ ώστε για κάθε $n \geqslant n_0$ να ισχύει το εξής: αν $m \leqslant \exp(c_4 \eta^2 M^2 n)$ και $y_1, \ldots, y_m \in S^{n-1}$, τότε υπάρχει $U \in O(n)$ ώστε, για κάθε $i = 1, \ldots, m$,*

$$(1-\eta)M \leqslant \|Uy_i\| \leqslant (1+\eta)M.$$

Απόδειξη. Θα χρησιμοποιήσουμε το γεγονός ότι υπάρχει φυσιολογικό μέτρο πιθανότητας ν στην $O(n)$ (το μέτρο Haar) το οποίο έχει την εξής ιδιότητα: αν $x_0 \in S^{n-1}$ και $A \subseteq S^{n-1}$, τότε

$$\sigma(A) = \nu\{U \in O(n) \, : \, Ux_0 \in A\}. \tag{1.5.2}$$

Ορίζουμε το σύνολο

$$A = \{x \in S^{n-1} : (1-\eta)M \leqslant \|x\| \leqslant (1+\eta)M\}.$$

Τότε, από την (1.5.1) έχουμε

$$\sigma(A) \geqslant 1 - c_2 e^{-c_3\eta^2 M^2 n}. \tag{1.5.3}$$

Για κάθε $i = 1, \dots, m$ θέτουμε

$$B_i = \{U \in O(n) \, : \, (1-\eta)M \leqslant \|Uy_i\| \leqslant (1+\eta)M\}.$$

Οι (1.5.2) και (1.5.3) δείχνουν ότι

$$\nu(B_i) > 1 - c_2 e^{-c_3\eta^2 M^2 n}.$$

Αν το n είναι αρκετά μεγάλο και η σταθερά c_4 επιλεγεί κατάλληλα, υποθέτοντας ότι $m \leqslant \exp(c_2\varepsilon^2 n/2)$ έχουμε ότι το $B = \bigcap B_i$ έχει μέτρο

$$\nu(B) \geqslant 1 - \sum_{i=1}^{m} \nu(B_i^c) \geqslant 1 - c_2 m \exp(-c_3\eta^2 M^2 n) > 0.$$

Συνεπώς, $B \neq \emptyset$ και για οποιονδήποτε $U \in B$ παίρνουμε

$$(1-\eta)M \leqslant \|Uy_i\| \leqslant (1+\eta)M$$

για κάθε $i = 1, \dots, m$. □

Λήμμα 1.5.8 (δ-δίκτυο). *Έστω $\delta \in (0,1)$. Υπάρχει $\mathcal{N} \subset S^{k-1}$ με τις εξής ιδιότητες:*

(i) *Για κάθε $y \in S^{k-1}$ υπάρχει $x \in \mathcal{N}$ ώστε $|x-y| < \delta$.*

(ii) $|\mathcal{N}| \leqslant \left(1 + \frac{2}{\delta}\right)^k$.

Απόδειξη. Έστω $\mathcal{N} = \{x_1, \dots, x_m\}$ ένα υποσύνολο της S^{k-1} του οποίου τα σημεία έχουν ανά δύο απόσταση μεγαλύτερη ή ίση του δ και έχει το μέγιστο δυνατό πληθάριθμο. Τέτοιο υποσύνολο υπάρχει λόγω της συμπάγειας της S^{k-1}.

Τότε, το $\mathcal{N}$ ικανοποιεί το (i): αν όχι, υπάρχει $y \in S^{k-1}$ ώστε $\|x_i - y\|_2 \geqslant \delta$ για κάθε $i \in \{1, \dots, m\}$. Όμως τότε, το $\mathcal{N}' = \{x_1, \dots, x_m, y\}$ είναι ένα σύνολο του οποίου τα σημεία ανήκουν στην S^{k-1} και οι αποστάσεις τους ανά δύο είναι μεγαλύτερες ή ίσες του δ. Αυτό είναι άτοπο αφού το $\mathcal{N}'$ έχει περισσότερα στοιχεία από το $\mathcal{N}$.

Για το (ii) θεωρούμε τα σύνολα $x_i + \frac{\delta}{2}B_2^k$, $i \leqslant m$. Αυτά έχουν ανά δύο ξένα εσωτερικά, και

$$x_i + \frac{\delta}{2}B_2^k \subset B_2^k + \frac{\delta}{2}B_2^k$$

για κάθε $i = 1, \ldots, m$. Συνεπώς,

$$\left| \bigcup_{i=1}^{m} \left(x_i + \frac{\delta}{2}B_2^k \right) \right| \leqslant \left| \left(1 + \frac{\delta}{2} \right) B_2^k \right|.$$

Έπεται ότι

$$\sum_{i=1}^{m} \left| \frac{\delta}{2}B_2^k \right| \leqslant \left(1 + \frac{\delta}{2} \right)^k |B_2^k|,$$

δηλαδή,

$$m \left(\frac{\delta}{2} \right)^k |B_2^k| \leqslant \left(1 + \frac{\delta}{2} \right)^k |B_2^k|.$$

Άρα, $m \leqslant (1 + \frac{2}{\delta})^k$. □

Χρησιμοποιώντας τα παραπάνω δείχνουμε το εξής.

Πρόταση 1.5.9. *Έστω $\delta, \eta \in (0,1)$. Αν $(1 + 2/\delta)^k \leqslant \exp(c_4 \eta^2 M^2 n)$, τότε υπάρχουν k-διάστατος υπόχωρος F του $\mathbb{R}^n$ και δ-δίκτυο $\mathcal{N}$ της $S_F : S^{n-1} \cap F$ με την ιδιότητα*

$$(1 - \eta)M \leqslant \|x\| \leqslant (1 + \eta)M$$

για κάθε $x \in \mathcal{N}$.

Απόδειξη. Σταθεροποιούμε υπόχωρο F_0 του $\mathbb{R}^n$ με διάσταση $\dim(F_0) = k$. Από το Λήμμα 1.5.8, υπάρχει δ-δίκτυο $\{y_1, \ldots, y_m\}$ της μοναδιαίας σφαίρας S_{F_0} του F_0, με $m \leqslant (1+2/\delta)^k$.

Αφού $(1 + 2/\delta)^k \leqslant \exp(c_2 \eta^2 n/2)$, το Λήμμα 1.5.7 δείχνει ότι υπάρχει $U \in O(n)$ με την εξής ιδιότητα: για κάθε $1 \leqslant i \leqslant m$,

$$(1 - \eta)M \leqslant \|Uy_i\| \leqslant (1 + \eta)M.$$

Θέτουμε $F := U(F_0)$ και $x_i := Uy_i$ $(i = 1, \ldots, m)$. Αφού ο U είναι ορθογώνιος μετασχηματισμός, το $\{x_1, \ldots, x_m\}$ είναι δ-δίκτυο της S_F, για το οποίο ισχύει

$$(1 - \eta)M \leqslant \|x\| \leqslant (1 + \eta)M.$$

Αυτό αποδεικνύει την πρόταση. □

Χρησιμοποιώντας τώρα το γεγονός ότι η $\|\cdot\|$ είναι νόρμα, θα περάσουμε από το δ-δίκτυο $\mathcal{N}$ της S_F σε ολόκληρη την S_F.

Πρόταση 1.5.10. *Έστω F ένας k-διάστατος υπόχωρος του $(\mathbb{R}^n, \|\cdot\|)$ για τον οποίο υπάρχει δ-δίκτυο $\mathcal{N}$ της S_F με την ιδιότητα*

$$(1-\eta)M \leqslant \|x\| \leqslant (1+\eta)M$$

για κάθε $x \in \mathcal{N}$. Τότε, για κάθε $y \in S_F$ έχουμε

$$\frac{1-\eta-2\delta}{1-\delta}M \leqslant \|y\| \leqslant \frac{1+\eta}{1-\delta}M.$$

Απόδειξη. Έστω $y \in S_F$ με τη μέγιστη δυνατή $\|\cdot\|$-νόρμα: $\|y\| = \ell$. Υπάρχει $x \in \mathcal{N}$ τέτοιο ώστε $|y-x| \leqslant \delta$. Από την τριγωνική ανισότητα έχουμε

$$\ell = \|y\| \leqslant \|x\| + \|y-x\| \leqslant \|x\| + \ell|y-x| \leqslant (1+\eta)M + \ell\delta,$$

δηλαδή

$$\ell \leqslant \frac{1+\eta}{1-\delta}M.$$

Αυτό αποδεικνύει τη δεξιά ανισότητα. Τώρα, για κάθε $z \in S_F$ βρίσκουμε πάλι $x \in \mathcal{N}$ τέτοιο ώστε $|z-x| \leqslant \delta$ και γράφουμε

$$\|z\| \geqslant \|x\| - \|z-x\| \geqslant (1-\eta)M - \frac{1+\eta}{1-\delta}M\delta = \frac{1-\eta-2\delta}{1-\delta}M.$$

Έτσι έχουμε αποδείξει την (1.5.10). □

Απόδειξη του Θεωρήματος 1.5.3. Από την Πρόταση 1.5.3, αν $\eta, \delta \in (0,1)$ και ο $k \in \mathbb{N}$ ικανοποιεί την $(1+2/\delta)^k \leqslant \exp(c_4\eta^2M^2n$, τότε υπάρχει k-διάστατος υπόχωρος F του $\mathbb{R}^n$ ώστε: για κάθε $x \in S_F$,

$$\frac{1-\eta-2\delta}{1-\delta}M \leqslant \|y\| \leqslant \frac{1+\eta}{1-\delta}M.$$

Αν επιλέξουμε $\eta = \delta = \varepsilon/10$ τότε για κάθε $x \in S_F$ έχουμε

$$\frac{1-\varepsilon/3}{1-\varepsilon/10}M \leqslant \|y\| \leqslant \frac{1+\varepsilon/10}{1-\varepsilon/10}M,$$

το οποίο δείχνει ότι

$$d(F, \ell_2^k) \leqslant \frac{1+\varepsilon/10}{1-\varepsilon/3} \leqslant 1+\varepsilon.$$

Μένει να προσδιορίσουμε τη μέγιστη τιμή του k η οποία ικανοποιεί την

$$\left(1+\frac{20}{\varepsilon}\right)^k \leqslant \exp\left(\frac{c_4}{100}\varepsilon^2M^2n\right).$$

Δεδομένου ότι
$$M^2 n \geqslant c_1^2 \log n,$$
βλέπουμε ότι αρκεί το n να είναι αρκετά μεγάλο και να ικανοποιείται η $k \leqslant c(\varepsilon)\log n$, όπου $c(\varepsilon)$ θετική σταθερά που εξαρτάται μόνο από το ε. Μπορούμε λοιπόν να επιλέξουμε k αυτής της τάξης μεγέθους. $\square$

Η γεωμετρική διατύπωση του Θεωρήματος 1.5.3 είναι η εξής: Για κάθε συμμετρικό κυρτό σώμα K στον $\mathbb{R}^n$ και για κάθε $\varepsilon > 0$, υπάρχουν $k \geqslant c\varepsilon^2\big(\log(1/\varepsilon)\big)^{-1}\log n$, υπόχωρος F του $\mathbb{R}^n$ διάστασης k, και ελλειψοειδές $\mathcal{E}$ στον F ώστε
$$\mathcal{E} \subset K \cap F \subset (1+\varepsilon)\mathcal{E}.$$

1.6 Ασκήσεις

1. Θεωρούμε τη μοναδιαία Ευκλείδεια σφαίρα $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$ στον $\mathbb{R}^n$. Ορίζουμε «απόσταση» $\rho(x,y)$ δύο σημείων $x, y \in S^{n-1}$ να είναι η κυρτή γωνία xoy στο επίπεδο που ορίζεται από την αρχή των αξόνων o και τα x, y.

(α) Δείξτε ότι: αν $\rho(x,y) = \theta$ τότε
$$|x-y| = 2\sin\frac{\theta}{2}$$
και συμπεράνατε ότι
$$\frac{2}{\pi}\rho(x,y) \leqslant |x-y| \leqslant \rho(x,y), \qquad x, y \in S^{n-1}.$$
(β) Δείξτε ότι η ρ είναι μετρική στην S^{n-1}.

2. Έστω K ένα συμμετρικό κυρτό σώμα στον $\mathbb{R}^n$ και έστω
$$\|x\| = \min\{t \geqslant 0 : x \in tK\}$$
η νόρμα που επάγεται στον $\mathbb{R}^n$ από το K (ελέγξτε ότι $K = \{x : \|x\| \leqslant 1\}$ δηλαδή το K είναι η μοναδιαία μπάλα του $(\mathbb{R}^n, \|\cdot\|)$). Δείξτε ότι
$$\int_{\mathbb{R}^n} e^{-\|x\|}dx = |K| \cdot \int_0^\infty t^n e^{-t}dt.$$
Γενικότερα, δείξτε ότι για κάθε $p > 0$,
$$\int_{\mathbb{R}^n} e^{-\|x\|^p}dx = |K| \cdot \int_0^\infty pt^{n+p-1}e^{-t^p}dt.$$
Υπόδειξη: Μπορείτε να γράψετε
$$\int_{\mathbb{R}^n} e^{-\|x\|^p}dx = \int_{\mathbb{R}^n}\left(\int_{\|x\|}^\infty pt^{p-1}e^{-t^p}dt\right)dx.$$

3. Η συνάρτηση $\Gamma : (0,+\infty) \to (0,+\infty)$ ορίζεται μέσω της
$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt.$$
Δείξτε ότι:

(α) $\Gamma(1) = 1$.

(β) $\Gamma(x+1) = x\Gamma(x)$ *για κάθε* $x > 0$.

(γ) $\Gamma(n+1) = n!$ *για κάθε* $n = 0, 1, 2, \dots$

(δ) $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

Δείξτε επίσης ότι η συνάρτηση Γ είναι *λογαριθμικά κυρτή*: η $\log\Gamma$ είναι κυρτή συνάρτηση.

4. Για κάθε $p \geqslant 1$, η συνάρτηση
$$\|x\|_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p}$$
είναι νόρμα στον $\mathbb{R}^n$. Ορίζουμε
$$B_p^n = \{x \in \mathbb{R}^n : \|x\|_p \leqslant 1\}.$$
Δείξτε ότι ο όγκος της B_p^n είναι ίσος με
$$|B_p^n| = \frac{\left[2\Gamma\left(\frac{1}{p}+1\right)\right]^n}{\Gamma\left(\frac{n}{p}+1\right)}.$$

5. Έστω B κυρτό σώμα στον $\mathbb{R}^n$. Για κάθε Lebesgue μετρήσιμο υποσύνολο A του $\mathbb{R}^n$, ορίζουμε
$$\mu_B(A) = \frac{|A\cap B|}{|B|}.$$
(α) Έστω $M \subseteq \mathbb{R}^n$ κυρτό και συμμετρικό ως προς το 0. Αποδείξτε ότι για κάθε $t > 1$ ισχύει ο εγκλεισμός
$$\mathbb{R}^n \setminus M \supseteq \frac{2}{t+1}(\mathbb{R}^n \setminus tM) + \frac{t-1}{t+1}M.$$
(β) Υποθέτουμε επιπλέον ότι $\mu_B(M) = a > 0$. Χρησιμοποιώντας το (α) και την ανισότητα Brunn-Minkowski αποδείξτε ότι, για κάθε $t > 1$,
$$1 - \mu_B(tM) \leqslant a\left(\frac{1-a}{a}\right)^{(t+1)/2}.$$

6. Έστω $A \subseteq \mathbb{R}^n$ κλειστό, κυρτό και συμμετρικό ως προς το 0. Αποδείξτε ότι για κάθε $x \in \mathbb{R}^n$ ισχύουν οι ανισότητες
$$e^{-|x|^2/2}\gamma_n(A) \leqslant \gamma_n(A+x) \leqslant \gamma_n(A).$$

7. Έστω $f,g,h:[0,\infty)\to[0,\infty)$ ολοκληρώσιμες συναρτήσεις που ικανοποιούν την

$$h(\sqrt{rs}) \geqslant \sqrt{f(r)}\cdot\sqrt{g(s)}$$

για κάθε $r,s>0$. Αποδείξτε ότι

$$\int_0^\infty h(x)dx \geqslant \left(\int_0^\infty f(x)dx\cdot\int_0^\infty g(x)dx\right)^{1/2}.$$

8. Έστω $f,g,h:[0,\infty)\to[0,\infty)$ ολοκληρώσιμες συναρτήσεις που ικανοποιούν την

$$h\left(\frac{2}{\frac{1}{r}+\frac{1}{s}}\right) \geqslant f(r)^{\frac{s}{r+s}}g(s)^{\frac{r}{r+s}}$$

για κάθε $r,s>0$. Θεωρούμε $p>0$ και θέτουμε

$$\begin{aligned} A &= \left(\int_0^\infty f(r)r^{p-1}\,dr\right)^{1/p}, \\ B &= \left(\int_0^\infty g(r)r^{p-1}\,dr\right)^{1/p}, \\ C &= \left(\int_0^\infty h(r)r^{p-1}\,dr\right)^{1/p}. \end{aligned}$$

Αποδείξτε ότι

$$C \geqslant \frac{2}{\frac{1}{A}+\frac{1}{B}}.$$

9. Έστω (X,d,μ) μετρικός χώρος πιθανότητας και $\varphi:(X,d)\to(Y,\sigma)$ Lipschitz συνάρτηση με σταθερά 1. Θεωρούμε το μέτρο πιθανότητας μ_φ στην $\mathcal{B}(Y)$ που ορίζεται από την

$$\mu_\varphi(A)=\mu(\varphi^{-1}(A)), \qquad A\in\mathcal{B}(Y).$$

Αποδείξτε ότι, για κάθε $t>0$,

$$\alpha_{\mu_\varphi}(t)\leqslant\alpha_\mu(t).$$

10. Έστω (X,d,μ) μετρικός χώρος πιθανότητας και έστω α_μ η συνάρτηση συγκέντρωσης του μ. Υποθέτουμε ότι για κάποιο $\varepsilon\in(0,1)$ και για κάποιο $t>0$ ισχύει $\alpha_\mu(t)<\varepsilon$. Αποδείξτε ότι: αν $A\in\mathcal{B}(X)$ και $\mu(A)\geqslant\varepsilon$, τότε $1-\mu(A_{t+r})\leqslant\alpha_\mu(r)$ για κάθε $r>0$.

11. Έστω (X,d,μ) μετρικός χώρος πιθανότητας και έστω α_μ η συνάρτηση συγκέντρωσης του μ. Αποδείξτε ότι:

(α) Αν $F:(X,d)\to\mathbb{R}$ είναι μια 1-Lipschitz συνεχής συνάρτηση, τότε

$$(\mu\otimes\mu)\left(\{(x,y)\in X\times X:|F(x)-F(y)|\geqslant t\}\right)\leqslant 4\alpha_\mu(t/2)$$

για κάθε $t > 0$.

(β) Αν $A, B \in \mathcal{B}(X)$ και $\mathrm{dist}(A, B) = \delta > 0$, τότε

$$\mu(A)\mu(B) \leqslant 4\alpha_\mu(\delta/2).$$

12. Σταθεροποιούμε έναν (μικρό) θετικό αριθμό κ. Έστω (X, d, μ) μετρικός χώρος πιθανότητας. Ορίζουμε τη *μερική διάμετρο* του (X, d) ως προς το μ ως εξής:

$$PD_\mu(X, d) := \inf\Big\{\mathrm{diam}(A) : A \in \mathcal{B}(X), \mu(A) \geqslant 1 - \kappa\Big\}.$$

Στη συνέχεια, για κάθε 1-Lipschitz συνάρτηση $F : (X, d) \to \mathbb{R}$ θεωρούμε το μέτρο μ_F που ορίζεται στη Borel σ-άλγεβρα του $\mathbb{R}$ μέσω της

$$\mu_F(A) = \mu(F^{-1}(A)), \qquad A \in \mathcal{B}(\mathbb{R})$$

και ορίζουμε την *παρατηρήσιμη διάμετρο* $OD_\mu(X, d)$ θέτοντας

$$OD_\mu(X, d) = \sup\{PD_{\mu_F}(\mathbb{R}) : \text{η } F : (X, d) \to \mathbb{R} \text{ είναι } 1 - \text{Lipschitz}\}.$$

Αποδείξτε ότι

$$OD_\mu(X, d) \leqslant 2\alpha_\mu^{-1}(\kappa/2).$$

13. Έστω (X, d, μ) μετρικός χώρος πιθανότητας. Υποθέτουμε ότι $\alpha_\mu(t) \leqslant Ce^{-ct^2}$ για κάποιες σταθερές $C, c > 0$ και για κάθε $t > 0$. Αποδείξτε ότι (αν η σταθερά $\kappa > 0$ της προηγούμενης άσκησης είναι αρκετά μικρή σε σχέση με τις C, c) τότε

$$OD_\mu(X, d) \leqslant 2\sqrt{\frac{1}{c}\,\ln\frac{2C}{\kappa}}.$$

Εκτιμήστε τις

$$OD_\sigma(S^{n-1}), \qquad OD_{\gamma_n}(\mathbb{R}^n), \qquad OD_{\mu_n}(E_2^n).$$

14. (α) Για κάθε $u \in S^{n-1}$ και για κάθε $\varepsilon \in (0, 1)$ ορίζουμε $C(u, \varepsilon) = \{\theta \in S^{n-1} : \langle u, \theta\rangle \geqslant \varepsilon\}$. Αποδείξτε ότι

$$\sigma(C(u, \varepsilon)) \leqslant \exp(-\varepsilon^2 n/2).$$

(β) Για κάθε $u \in S^{n-1}$ και για κάθε $\varepsilon \in (0, 1)$ ορίζουμε $B(u, \varepsilon) = \{\theta \in S^{n-1} : |\theta - u| \leqslant \varepsilon\}$. Αποδείξτε ότι

$$\sigma(B(u, \varepsilon)) \geqslant \left(\frac{\varepsilon}{\varepsilon + 2}\right)^n \geqslant \left(\frac{\varepsilon}{3}\right)^n.$$

(γ) Έστω $v_1, \ldots, v_m \in \mathbb{R}^n$. Υποθέτουμε ότι το συμμετρικό κυρτό πολύεδρο

$$K = \{y \in \mathbb{R}^n : |\langle y, v_j\rangle| \leqslant 1 \text{ για κάθε } j = 1, \ldots, m\}$$

ικανοποιεί την $B_2^n \subseteq K \subseteq \alpha B_2^n$ για κάποιον $\alpha > 1$. Δείξτε ότι $m \geqslant \exp(n/(2\alpha^2))$.

15. Έστω $(X_i, \|\cdot\|_i)$, $i \leqslant n$, πεπερασμένη ακολουθία χώρων με νόρμα. Για κάθε $i \leqslant n$ θεωρούμε ένα πεπερασμένο υποσύνολο Ω_i του X_i με διάμετρο μικρότερη ή ίση του 1. Έστω P_i μέτρο πιθανότητας στο Ω_i. Θεωρούμε τον χώρο γινόμενο $X^{(n)} = \big(\sum_{i\leqslant n} \oplus X_i\big)_2$ και θέτουμε

$$\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$$

και

$$P = P^n = P_1 \times P_2 \times \cdots \times P_n.$$

(το μέτρο γινόμενο στο Ω). Για κάθε $A \subseteq \Omega$ ορίζουμε

$$\phi_A(t) = d(t, \mathrm{conv}(A)),$$

την απόσταση του t από την κυρτή θήκη $\mathrm{conv}(A)$ του A στον $X^{(n)}$. Δείξτε ότι, για κάθε $A \subseteq \Omega$,

$$\mathbb{E}\left(e^{\phi_A^2}/4\right) \leqslant \frac{1}{P(A)}.$$

Κεφάλαιο 2

Αθροίσματα ανεξάρτητων τυχαίων μεταβλητών

2.1 Η ανισότητα του Hoeffding και η ανισότητα του Chernoff

Σε αυτή την παράγραφο αποδεικνύουμε δύο τυπικές ανισότητες του είδους που θα μας απασχολήσει, οι οποίες αφορούν συμμετρικές τυχαίες μεταβλητές Bernoulli.

Ορισμός 2.1.1. Μια τυχαία μεταβλητή X έχει συμμετρική κατανομή Bernoulli αν παίρνει τις τιμές 1 και -1 με πιθανότητα $1/2$. Δηλαδή, αν

$$\mathbb{P}(X=1)=\mathbb{P}(X=-1)=\frac{1}{2}.$$

Λέμε ότι η τυχαία μεταβλητή Z έχει κατανομή Bernoulli αν η $X=2Z-1$ έχει συμμετρική κατανομή Bernoulli. Δηλαδή, αν

$$\mathbb{P}(Z=1)=\mathbb{P}(Z=0)=\frac{1}{2}.$$

Θεώρημα 2.1.2 (ανισότητα Hoeffding)**.** *Έστω $X_1,\ldots,X_N$ ανεξάρτητες συμμετρικές Bernoulli τυχαίες μεταβλητές και έστω $a=(a_1,\ldots,a_N)\in\mathbb{R}^n$. Τότε, για κάθε $t>0$ έχουμε*

$$\mathbb{P}\left(\sum_{i=1}^N a_iX_i\geqslant t\right)\leqslant \exp\Big(-\frac{t^2}{2|a|^2}\Big). \tag{2.1.1}$$

Απόδειξη. Σταθεροποιούμε $\lambda > 0$ και χρησιμοποιώντας το γεγονός ότι η εκθετική συνάρτηση είναι αύξουσα και την ανισότητα Markov γράφουμε

$$\mathbb{P}\left(\sum_{i=1}^N a_iX_i \geqslant t\right) = \mathbb{P}\left(\exp\Big(\lambda\sum_{i=1}^N a_iX_i\Big) \geqslant e^{\lambda t}\right) \tag{2.1.2}$$
$$\leqslant e^{-\lambda t}\mathbb{E}\left[\exp\Big(\lambda\sum_{i=1}^N a_iX_i\Big)\right].$$

Η συνάρτηση (του λ) στο δεξιό μέλος είναι η *ροπογεννήτρια* του αθροίσματος $\sum_{i=1}^N a_iX_i$. Οι τυχαίες μεταβλητές $e^{\lambda a_iX_i}$ είναι ανεξάρτητες διότι οι X_i έχουν υποτεθεί ανεξάρτητες. Συνεπώς,

$$\mathbb{E}\left[\exp\Big(\lambda\sum_{i=1}^N a_iX_i\Big)\right] = \prod_{i=1}^N \mathbb{E}\big(e^{\lambda a_iX_i}\big). \tag{2.1.3}$$

Απευθείας υπολογισμός δείχνει ότι

$$\mathbb{E}\big(e^{\lambda a_iX_i}\big) = \frac{1}{2}e^{\lambda a_i} + \frac{1}{2}e^{-\lambda a_i} = \cosh(\lambda a_i).$$

Συγκρίνοντας τα αναπτύγματα Taylor των $\cosh(z)$ και $\exp(z^2/2)$ βλέπουμε ότι

$$\cosh(z) \leqslant \exp(z^2/2)$$

για κάθε $z \in \mathbb{R}$. Πράγματι,

$$\cosh(z) = \sum_{m=0}^\infty \frac{z^{2m}}{2(2m)!} \leqslant \sum_{m=0}^\infty \frac{z^{2m}}{2^m m!} = \exp(z^2/2).$$

Συνεπώς,

$$\mathbb{E}\big(e^{\lambda a_iX_i}\big) \leqslant \exp(\lambda^2 a_i^2/2).$$

Επιστρέφοντας στην (2.1.3) και κατόπιν στην (2.1.2) βλέπουμε ότι

$$\mathbb{P}\left(\sum_{i=1}^N a_iX_i \geqslant t\right) \leqslant e^{-\lambda t}\prod_{i=1}^N \exp(\lambda^2 a_i^2/2)$$
$$= \exp\left(-\lambda t + \frac{\lambda^2}{2}\sum_{i=1}^N a_i^2\right) = \exp\left(-\lambda t + \frac{\lambda^2|a|^2}{2}\right).$$

Τέλος, επιλέγουμε τη βέλτιστη τιμή του λ: παίρνοντας

$$\lambda_0 = \frac{t}{|a|^2}$$

παίρνουμε την (2.1.1). □

Παρατήρηση 2.1.3. Εφαρμόζοντας την (2.1.1) για το $-a = (-a_1, \ldots, -a_N)$ παίρνουμε

$$\mathbb{P}\left(-\sum_{i=1}^{N} a_i X_i \geqslant t \right) \leqslant \exp\Big(-\frac{t^2}{2|a|^2} \Big). \tag{2.1.4}$$

για κάθε $t > 0$. Άρα,

$$\mathbb{P}\left(\Big| \sum_{i=1}^{N} a_i X_i \Big| \geqslant t \right) \leqslant 2\exp\Big(-\frac{t^2}{2|a|^2} \Big) \tag{2.1.5}$$

για κάθε $t > 0$.

Το ερώτημα αν παρόμοιες ανισότητες ισχύουν για γενικότερες κλάσεις τυχαίων μεταβλητών είναι το κεντρικό θέμα αυτού του κεφαλαίου. Μια καλή άσκηση είναι να δοκιμάσετε τώρα να δείξετε το εξής: Αν $X_1, \ldots, X_N$ είναι ανεξάρτητες τυχαίες μεταβλητές με $\mathbb{P}(a_i \leqslant X_i \leqslant b_i) = 1$ για $i = 1, \ldots, N$, τότε

$$\mathbb{P}\left(\sum_{i=1}^{N} (X_i - \mathbb{E}(X_i)) \geqslant t \right) \leqslant \exp\left(-\frac{2t^2}{\sum_{i=1}^{N} (b_i - a_i)^2} \right)$$

για κάθε $t > 0$.

Θεώρημα 2.1.4 (ανισότητα Chernoff)**.** *Έστω $X_1, \ldots, X_N$ ανεξάρτητες τυχαίες μεταβλητές Bernoulli με παραμέτρους $p_i \in (0,1)$. Δηλαδή, για κάθε $i = 1, \ldots, N$,*

$$\mathbb{P}(X_i = 1) = p_i \quad \textit{και} \quad \mathbb{P}(X_i = 0) = 1 - p_i.$$

Θεωρούμε το άθροισμα $S_N = \sum_{i=1}^{N} X_i$ και θέτουμε $\mu = \mathbb{E}(S_N) = \sum_{i=1}^{N} p_i$. Τότε, για κάθε $t > \mu$ έχουμε

$$\mathbb{P}(S_N \geqslant t) \leqslant e^{-\mu} \left(\frac{e\mu}{t} \right)^t.$$

Ειδικότερα, για κάθε $t \geqslant e^2 \mu$ έχουμε

$$\mathbb{P}(S_N \geqslant t) \leqslant e^{-t}. \tag{2.1.6}$$

Απόδειξη. Όπως και στην απόδειξη της ανισότητας Hoeffding, σταθεροποιούμε $\lambda > 0$ και χρησιμοποιώντας την ανισότητα Markov γράφουμε

$$\mathbb{P}(S_N \geqslant t) \leqslant e^{-\lambda t} \prod_{i=1}^{N} \mathbb{E}\left(\exp(\lambda X_i) \right). \tag{2.1.7}$$

Αφού η X_i παίρνει την τιμή 1 με πιθανότητα p_i και την τιμή 0 με πιθανότητα $1-p_i$, έχουμε

$$\mathbb{E}(\exp(\lambda X_i)) = e^{\lambda}p_i + (1-p_i) = 1 + (e^{\lambda}-1)p_i \leqslant \exp((e^{\lambda}-1)p_i),$$

χρησιμοποιώντας την $1+x \leqslant e^x$ στο τελευταίο βήμα. Αντικαθιστώντας αυτές τις ανισότητες (για $i=1,\dots,N$) στην (2.1.7) παίρνουμε

$$\mathbb{P}(S_N \geqslant t) \leqslant e^{-\lambda t}\exp\left((e^{\lambda}-1)\sum_{i=1}^{N} p_i\right) = e^{-\lambda t}\exp\left((e^{\lambda}-1)\mu\right).$$

Επιλέγουμε $\lambda_0 = \ln(t/\mu)$ (παρατηρήστε ότι $\lambda_0 > 0$ διότι $t/\mu > 1$). Τότε, παίρνουμε

$$\begin{aligned}\mathbb{P}(S_N \geqslant t) &\leqslant e^{-t\ln(t/\mu)}e^{\left(\frac{t}{\mu}-1\right)\mu} = \left(\frac{\mu}{t}\right)^t e^{t-\mu}\\ &= e^{-\mu}\left(\frac{e\mu}{t}\right)^t.\end{aligned}$$

Αν $t \geqslant e^2\mu$ τότε $e\mu/t \leqslant e^{-1}$, απ' όπου έπεται ότι $\mathbb{P}(S_N \geqslant t) \leqslant e^{-\mu}e^{-t} \leqslant e^{-t}$. □

Παρόμοιο επιχείρημα μας δίνει άνω φράγμα για την $\mathbb{P}(S_N \leqslant t)$ στην περίπτωση που $t < \mu$.

Πρόταση 2.1.5. *Έστω $X_1,\dots,X_N$ ανεξάρτητες τυχαίες μεταβλητές Bernoulli με παραμέτρους $p_i \in (0,1)$. Θεωρούμε το άθροισμα $S_N = \sum_{i=1}^{N} X_i$ και θέτουμε $\mu = \mathbb{E}(S_N)$. Τότε, για κάθε $t < \mu$ έχουμε*

$$\mathbb{P}(S_N \leqslant t) \leqslant e^{-\mu}\left(\frac{e\mu}{t}\right)^t.$$

Απόδειξη. Όπως και στην προηγούμενη απόδειξη, γράφουμε

$$\mathbb{P}(S_N \leqslant t) = \mathbb{P}(e^{-\lambda S_N} \geqslant e^{-\lambda t}) \leqslant e^{\lambda t}\prod_{i=1}^{N}\mathbb{E}\left(\exp(-\lambda X_i)\right). \tag{2.1.8}$$

Αφού η X_i παίρνει την τιμή 1 με πιθανότητα p_i και την τιμή 0 με πιθανότητα $1-p_i$, έχουμε

$$\mathbb{E}(\exp(-\lambda X_i)) = e^{-\lambda}p_i + (1-p_i) = 1 + (e^{-\lambda}-1)p_i \leqslant \exp(-(1-e^{-\lambda})p_i),$$

χρησιμοποιώντας την $1+x \leqslant e^x$ στο τελευταίο βήμα. Αντικαθιστώντας αυτές τις ανισότητες (για $i=1,\dots,N$) στην (2.1.8) παίρνουμε

$$\mathbb{P}(S_N \leqslant t) \leqslant e^{\lambda t}\exp\left(-(1-e^{-\lambda})\sum_{i=1}^{N} p_i\right) = e^{\lambda t}\exp\left(-(1-e^{\lambda})\mu\right).$$

Επιλέγουμε $\lambda_0 = \ln(\mu/t)$ (παρατηρήστε ότι $\lambda_0 > 0$ διότι $\mu/t > 1$). Τότε, παίρνουμε

$$\begin{aligned}\mathbb{P}(S_N \geqslant t) &\leqslant e^{t\ln(\mu/t)} e^{-\left(1-\frac{t}{\mu}\right)\mu} = \left(\frac{\mu}{t}\right)^t e^{-(\mu-t)} \\ &= e^{-\mu}\left(\frac{e\mu}{t}\right)^t,\end{aligned}$$

που είναι το ζητούμενο. □

Από το Θεώρημα 2.1.4 και την Πρόταση 2.1.5 προκύπτει το εξής φράγμα για την $\mathbb{P}(|S_N - \mu| \geqslant \delta\mu)$, όπου $\delta \in (0,1)$.

Πρόταση 2.1.6. *Έστω $X_1, \ldots, X_N$ ανεξάρτητες τυχαίες μεταβλητές Bernoulli με παραμέτρους $p_i \in (0,1)$. Θεωρούμε το άθροισμα $S_N = \sum_{i=1}^N X_i$ και θέτουμε $\mu = \mathbb{E}(S_N)$. Τότε, για κάθε $\delta \in (0,1/4)$ έχουμε*

$$\mathbb{P}(|S_N - \mu| \geqslant \delta\mu) \leqslant 2e^{-\delta^2\mu/4}.$$

Απόδειξη. Φράσσουμε χωριστά τις $\mathbb{P}(S_N - \mu \geqslant \delta\mu)$ και $\mathbb{P}(S_N - \mu \leqslant -\delta\mu)$. Για την πρώτη πιθανότητα, από το Θεώρημα 2.1.4 έχουμε

$$\mathbb{P}(S_N - \mu \geqslant \delta\mu) = \mathbb{P}(S_N \geqslant (1+\delta)\mu) \leqslant \frac{e^{\delta\mu}}{(1+\delta)^{(1+\delta)\mu}} = e^{(\delta-(1+\delta)\ln(1+\delta))\mu}.$$

Παραγωγίζοντας βλέπουμε ότι η συνάρτηση $g(\delta) = \ln(1+\delta) - \delta + \frac{\delta^2}{2}$ είναι αύξουσα στο $[0,1)$ και $g(0) = 0$, άρα $\ln(1+\delta) > \delta - \frac{\delta^2}{2}$ για κάθε $\delta \in (0,1)$. Συνεπώς,

$$(1+\delta)\ln(1+\delta) - \delta > (1+\delta)\left(\delta - \frac{\delta^2}{2}\right) - \delta = \frac{\delta^2}{2}(1-\delta) > \frac{\delta^2}{4}$$

αν $0 < \delta < 1/2$. Δηλαδή,

$$\mathbb{P}(S_N - \mu \geqslant \delta\mu) \leqslant e^{-\frac{\delta^2\mu}{4}}, \qquad 0 < \delta < \frac{1}{2}.$$

Για τη δεύτερη πιθανότητα, από την Πρόταση 2.1.5 έχουμε

$$\mathbb{P}(S_N - \mu \leqslant \delta\mu) = \mathbb{P}(S_N \leqslant (1-\delta)\mu) \leqslant \frac{e^{-\delta\mu}}{(1-\delta)^{(1-\delta)\mu}} = e^{(-\delta-(1-\delta)\ln(1-\delta))\mu}.$$

Παραγωγίζοντας βλέπουμε ότι η συνάρτηση $h(\delta) = \ln(1-\delta) + \delta + \frac{2\delta^2}{3}$ είναι αύξουσα στο $[0,1/4)$ και $h(0) = 0$, άρα

$$(1-\delta)\ln(1-\delta) + \delta > (1-\delta)(-\delta - 2\delta^2/3) + \delta = \frac{\delta^2}{3}\left(1 - \frac{2\delta}{3}\right) > \frac{5\delta^2}{18} > \frac{\delta^2}{4}$$

αν $0 < \delta < 1/4$. Δηλαδή,

$$\mathbb{P}(S_N - \mu \geqslant \delta\mu) \leqslant e^{-\frac{\delta^2\mu}{4}}, \qquad 0 < \delta < \frac{1}{4}.$$

Συνδυάζοντας τις δύο ανισότητες έχουμε το συμπέρασμα της πρότασης. □

2.1α′ Εφαρμογή: βαθμός κορυφών τυχαίου γραφήματος

Δίνουμε ένα παράδειγμα εφαρμογής της ανισότητας Chernoff στα τυχαία γραφήματα. Το κλασικό μοντέλο τυχαίου γραφήματος των Erdös-Renyi είναι το τυχαίο γράφημα $G \sim G(n,p)$ που κατασκευάζεται ως εξής: έχουμε n κορυφές και συνδεουμε κάθε ζεύγος κορυφών τυχαία και ανεξάρτητα, με πιθανότητα $p \in (0,1)$.

Ο *βαθμός* μιας κορυφής του γραφήματος είναι το πλήθος των ακμών στις οποίες ανήκει. Ο μέσος βαθμός κάθε κορυφής του G είναι ίσος με

$$d := (n-1)p.$$

Η επόμενη πρόταση δείχνει ότι αν $d \gg \log n$ (δηλαδή, αν $p \gg \log n/n$) τότε ο βαθμός όλων των κορυφών του G είναι πολύ κοντά στο μέσο βαθμό d.

Πρόταση 2.1.7. *Έστω $G(n,p)$ ένα τυχαίο γράφημα με μέσο βαθμό $d \geqslant C \log n$. Τότε, με πιθανότητα μεγαλύτερη από 0.9 όλες οι κορυφές του γραφήματος G έχουν βαθμό μεταξύ $0.9d$ και $1.1d$.*

Απόδειξη. Σταθεροποιούμε μια κορυφή i του γραφήματος (μπορούμε να υποθέσουμε ότι το σύνολο των κορυφών είναι το $[n] = \{1,2,\ldots,n\}$). Ο βαθμός d_i της κορυφής i είναι το άθροισμα n ανεξάρτητων τυχαίων μεταβλητών Bernoulli με παράμετρο p. Από την Πρόταση 2.1.6 έχουμε

$$\mathbb{P}(|d_i - d| \geqslant 0.1d) \leqslant 2e^{-d/400}.$$

Συνεπώς,

$$\mathbb{P}\left(\bigcup_{i=1}^{n}\{|d_i - d| \geqslant 0.1d\}\right) \leqslant 2ne^{-d/400} < \frac{1}{10}$$

αν

$$d > 400\log(20n).$$

Αυτό αποδεικνύει ότι αν $d \geqslant C\log n$ για κάποια, αρκετά μεγάλη, απόλυτη σταθερά $C > 0$ τότε με πιθανότητα μεγαλύτερη από 0.9 έχουμε

$$|d_i - d| < 0.1d$$

για κάθε $i = 1,\ldots,n$. □

2.2 Υποκανονικές τυχαίες μεταβλητές

Ορισμός 2.2.1 (υποκανονική τυχαία μεταβλητή)**.** Μια τυχαία μεταβλητή X λέγεται υποκανονική με σταθερά $\alpha > 0$ αν για κάθε $t > 0$ ισχύει

$$\mathbb{P}(|X| \geqslant t) \leqslant 2e^{-t^2/\alpha^2}.$$

Τυπικό παράδειγμα υποκανονικής τυχαίας μεταβλητής είναι η $g \sim N(0,1)$. Παρατηρήστε ότι:

(i) Για κάθε $t > 0$,
$$\mathbb{P}(|g| \geqslant t) = \frac{2}{\sqrt{2\pi}} \int_t^\infty e^{-s^2/2} ds \leqslant 2e^{-t^2/2}.$$

(ii) Για κάθε $p \geqslant 1$ έχουμε
$$\begin{aligned}\|g\|_p^p &= \frac{2}{\sqrt{2\pi}} \int_0^\infty s^p e^{-s^2/2} ds = \frac{2}{\sqrt{2\pi}} \int_0^\infty (2y)^{\frac{p-1}{2}} e^{-y} dy = \frac{2^{p/2}}{\sqrt{2\pi}} \int_0^\infty y^{\frac{p+1}{2}-1} e^{-y} dy \\ &= \frac{2^{p/2}}{\sqrt{2\pi}} \Gamma\left(\frac{p+1}{2}\right),\end{aligned}$$

και από τον τύπο του Stirling ($\Gamma(x+1) \sim (x/e)^x \sqrt{2\pi x}$ καθώς το $x \to \infty$) παίρνουμε
$$\|g\|_p = \sqrt{2} \left(\frac{\Gamma\left(\frac{p+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)} \right)^{1/p} \leqslant c\sqrt{p},$$

όπου $c > 0$ είναι μια απόλυτη σταθερά.

(iii) Για κάθε $\lambda \in \mathbb{R}$ έχουμε
$$\mathbb{E}(e^{\lambda g}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{\lambda s - \frac{s^2}{2}} ds = e^{\lambda^2/2}.$$

Η επόμενη πρόταση δείχνει ότι οι υποκανονικές τυχαίες μεταβλητές «χαρακτηρίζονται από οποιαδήποτε από τις παραπάνω ιδιότητες». Ακριβέστερα έχουμε τα ακόλουθα.

Πρόταση 2.2.2. *Έστω X μια τυχαία μεταβλητή. Οι παρακάτω ιδιότητες είναι ισοδύναμες (με τις σταθερές α, β, γ να διαφέρουν το πολύ κατά μία απόλυτη σταθερά).*

(i) *Υπάρχει $\alpha > 0$ ώστε: για κάθε $t > 0$,*
$$\mathbb{P}(|X| \geqslant t) \leqslant 2e^{-t^2/\alpha^2}.$$

(ii) *Υπάρχει* $\beta > 0$ *ώστε: για κάθε* $p \geqslant 1$,

$$\|X\|_p \leqslant \beta\sqrt{p}.$$

(iii) *Υπάρχει* $\gamma > 0$ *ώστε*

$$\mathbb{E}(e^{X^2/\gamma^2}) \leqslant 2.$$

Αν επιπλέον $\mathbb{E}(X) = 0$ *τότε οι παραπάνω τρεις ιδιότητες είναι ισοδύναμες με την ακόλουθη:*

(iv) Υπάρχει $\delta > 0$ *ώστε: για κάθε* $\lambda \in \mathbb{R}$,

$$\mathbb{E}(e^{\lambda X}) \leqslant e^{\delta^2\lambda^2}.$$

Απόδειξη. (i) $\Longrightarrow$ (ii): Έστω $p \geqslant 1$. Γράφουμε

$$\begin{aligned}\|X\|_p^p &= \int_0^\infty pt^{p-1}\mathbb{P}(|X| \geqslant t)\,dt \leqslant 2\int_0^\infty pt^{p-1}e^{-t^2/\alpha^2}dt\\ &= \alpha^2\int_0^\infty p(\alpha^2 s)^{\frac{p}{2}-1}e^{-s}ds = \alpha^p p\int_0^\infty s^{\frac{p}{2}-1}e^{-s}ds = \alpha^p p\Gamma\left(\frac{p}{2}\right)\\ &\leqslant (c\alpha\sqrt{p})^p,\end{aligned}$$

όπου $c > 0$ είναι μια απόλυτη σταθερά. Δηλαδή, $\|X\|_p \leqslant \beta\sqrt{p}$, με $\beta = c\alpha$.
(ii) $\Longrightarrow$ (iii): Έστω $t > 0$. Γράφουμε

$$\begin{aligned}\mathbb{E}(e^{t^2X^2}) &= \mathbb{E}\left(1 + \sum_{k=1}^\infty \frac{t^{2k}X^{2k}}{k!}\right) = 1 + \sum_{k=1}^\infty \frac{t^{2k}}{k!}\mathbb{E}(X^{2k})\\ &\leqslant 1 + \sum_{k=1}^\infty \frac{t^{2k}}{k!}\beta^{2k}(2k)^k \leqslant \sum_{k=0}^\infty (2et^2\beta^2)^k = \frac{1}{1-2et^2\beta^2}\\ &\leqslant e^{4et^2\beta^2} \leqslant \sqrt{e} \leqslant 2,\end{aligned}$$

αν $t \leqslant \frac{1}{\sqrt{8e}\beta}$, χρησιμοποιώντας στο τέλος την ανισότητα $\frac{1}{1-x} \leqslant e^{2x}$ για $0 \leqslant x \leqslant 1/2$. Επομένως, αν επιλέξουμε $\gamma = \sqrt{8e}\beta$ έχουμε

$$\mathbb{E}(e^{X^2/\gamma^2}) \leqslant 2.$$

(iii) $\Longrightarrow$ (i): Για κάθε $t > 0$ έχουμε

$$\mathbb{P}(|X| \geqslant t) = \mathbb{P}(e^{X^2/\gamma^2} \geqslant e^{t^2/\gamma^2}) \leqslant 2e^{-t^2/\gamma^2}$$

από την ανισότητα Markov.

Δείχνουμε τώρα ότι αν $\mathbb{E}(X) = 0$ τότε (ii)$\Longrightarrow$(iv). Χρησιμοποιούμε την ανισότητα $e^x \leqslant x + e^{x^2}$, η οποία ισχύει για κάθε $x \in \mathbb{R}$. Μπορούμε λοιπόν να γράψουμε

$$\mathbb{E}(e^{\lambda X}) \leqslant \mathbb{E}(\lambda X + e^{\lambda^2 X^2}) = \mathbb{E}(e^{\lambda^2 X^2}).$$

Επαναλαμβάνοντας τον υπολογισμό που κάναμε παραπάνω, βλέπουμε ότι

$$\begin{aligned}\mathbb{E}(e^{\lambda^2 X^2}) &= \mathbb{E}\left(1 + \sum_{k=1}^{\infty} \frac{\lambda^{2k} X^{2k}}{k!}\right) = 1 + \sum_{k=1}^{\infty} \frac{\lambda^{2k}}{k!}\mathbb{E}(X^{2k}) \\ &\leqslant 1 + \sum_{k=1}^{\infty} \frac{\lambda^{2k}}{k!}\beta^{2k}(2k)^k \leqslant \sum_{k=0}^{\infty}(2e\lambda^2\beta^2)^k = \frac{1}{1-2e\lambda^2\beta^2} \\ &\leqslant e^{4e\beta^2\lambda^2},\end{aligned}$$

αν $|\lambda| \leqslant 1/(\sqrt{8e}\beta)$. Από την άλλη πλευρά, αν $|\lambda| > 1/(\sqrt{8e}\beta)$ τότε θέτοντας $\gamma = \sqrt{4e}\beta$ από την ανισότητα $2\lambda X \leqslant \lambda^2\gamma^2 + \frac{X^2}{\gamma^2}$ παίρνουμε

$$\begin{aligned}\mathbb{E}(e^{\lambda X}) &\leqslant e^{\gamma^2\lambda^2/2}\mathbb{E}(e^{X^2/2\gamma^2}) = \leqslant e^{2e^2\beta^2\lambda^2}\mathbb{E}(e^{X^2/(8e\beta^2)}) \\ &\leqslant e^{2e^2\beta^2\lambda^2}e^{1/2} \leqslant e^{2e^2\beta^2\lambda^2}e^{8e\beta^2\lambda^2} \leqslant e^{5e^2\beta^2\lambda^2}.\end{aligned}$$

Άρα, η (iv) ισχύει με $\delta = \sqrt{5}e\gamma$.

Τέλος αποδεικνύουμε ότι (iv)$\Longrightarrow$(i). Για κάθε $t > 0$ και για κάθε $\lambda > 0$ έχουμε

$$\mathbb{P}(X \geqslant t) = \mathbb{P}(e^{\lambda X} \geqslant e^{\lambda t}) \leqslant e^{-\lambda t}\mathbb{E}(e^{\lambda X}) \leqslant e^{-\lambda t + \delta^2\lambda^2}.$$

Επιλέγοντας $\lambda = \frac{t^2}{2\delta^2}$ βλέπουμε ότι

$$\mathbb{P}(X \geqslant t) \leqslant e^{-\frac{t^2}{4\delta^2}}.$$

Εφαρμόζοντας το ίδιο επιχείρημα για την $-X$ ολοκληρώνουμε την απόδειξη με $\alpha = 2\delta$. $\square$

Ορισμός 2.2.3 (χώροι Orlicz)**.** Έστω $(Q, \mathcal{A}, \mu)$ χώρος πιθανότητας και έστω $\psi : \mathbb{R} \to [0, +\infty)$ μια άρτια κυρτή συνάρτηση που ικανοποιεί τα ακόλουθα: $\psi(0) = 0$ και $\lim_{x\to\infty}\psi(x) = +\infty$. Τότε λέμε ότι η ψ είναι *συνάρτηση Orlicz.*

Ο χώρος Orlicz $L_\psi(\mu)$ που αντιστοιχεί στην ψ αποτελείται από όλες τις $\mathcal{A}$-μετρήσιμες συναρτήσεις f για τις οποίες υπάρχει $\kappa > 0$ τέτοιος ώστε

$$\int_X \psi(f/\kappa)d\mu < \infty.$$

Η νόρμα Orlicz (ως προς την ψ) μιας $f \in L_\psi(\mu)$ ορίζεται τότε ως εξής:

$$\|f\|_{\psi(\mu)} := \inf\left\{\kappa > 0 : \int_X \psi(f/\kappa)d\mu \leqslant 1\right\}.$$

Ισχύει ότι $L_\psi(\mu) \subseteq L_1(\mu)$: αν μια μετρήσιμη συνάρτηση f έχει πεπερασμένη $\psi(\mu)$-νόρμα τότε η f είναι ολοκληρώσιμη ως προς μ. Για να το δούμε αυτό, παρατηρούμε πρώτα ότι, αφού η ψ είναι κυρτή και $\psi(0) = (0)$, η συνάρτηση $t \mapsto \frac{\psi(t)}{t}$ είναι αύξουσα. Έπεται ότι $\psi(t) \geqslant \frac{\psi(t_0)}{t_0} \cdot t$ για κάθε $t > t_0$, όπου t_0 είναι οποιοσδήποτε θετικός αριθμός για τον οποίο $\psi(t_0) > 0$. Τότε, για κάθε $\kappa > \|f\|_{\psi(\mu)}$ μπορούμε να γράψουμε

$$\begin{aligned}\frac{1}{\kappa}\mathbb{E}_\mu(|f|) &= \mathbb{E}_\mu\Big(\frac{|f|}{\kappa} \cdot \mathbf{1}_{\{|f| \leqslant t_0\kappa\}}\Big) + \mathbb{E}_\mu\Big(\frac{|f|}{\kappa} \cdot \mathbf{1}_{\{|f| > t_0\kappa\}}\Big) \\ &\leqslant t_0 + \frac{t_0}{\psi(t_0)}\mathbb{E}_\mu(\psi(|f|/\kappa)) \leqslant t_0 \cdot [1 + (\psi(t_0))^{-1}] < +\infty.\end{aligned}$$

Εφαρμόζοντας την ανισότητα Jensen για την κυρτή συνάρτηση ψ παίρνουμε

$$\psi(\mathbb{E}_\mu(|f|/\kappa)) \leqslant \mathbb{E}_\mu(\psi(|f|/\kappa)) \leqslant 1$$

για κάθε $\kappa > \|f\|_{\psi(\mu)}$. Συνεπώς,

$$\mathbb{E}_\mu(|f|) \leqslant \psi_*^{-1}(1) \cdot \|f\|_{\psi(\mu)},$$

όπου $\psi_*^{-1}(1) = \inf\{s > 0 : \psi(t) > 1 \text{ για κάθε } t \geqslant s\}$.

Η οικογένεια των ψ_α-νορμών, η οποία είναι υποοικογένεια των νορμών Orlicz, παίζει σημαντικό ρόλο σε αυτό το μάθημα.

Ορισμός 2.2.4 (ψ_α-νόρμα)**.** Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας και έστω $f : X \to \mathbb{R}$ μια $\mathcal{A}$-μετρήσιμη συνάρτηση. Για κάθε $\alpha \in [1, 2]$ ορίζουμε την ψ_α-νόρμα της f ως εξής:

$$\|f\|_{\psi_\alpha} := \inf\left\{t > 0 : \int_X \exp\left(\frac{|f|}{t}\right)^\alpha d\mu \leqslant 2\right\},$$

αρκεί το σύνολο στο δεξιό μέλος να είναι μη κενό. Παρατηρήστε ότι η ψ_α-νόρμα είναι ακριβώς η νόρμα Orlicz που αντιστοιχεί στην κυρτή συνάρτηση $t \in \mathbb{R} \to e^{|t|^\alpha} - 1$.

Με βάση τον ορισμό της ψ_2-νόρμας, μπορούμε να επαναδιατυπώσουμε τα συμπεράσματα της Πρότασης 2.2.2 ως εξής:

Πρόταση 2.2.5. *Έστω X μια τυχαία μεταβλητή. Αν $X \in L_{\psi_2}$ τότε:*

(i) *Για κάθε* $t > 0$,
$$\mathbb{P}(|X| \geqslant t) \leqslant 2e^{-c_1 t^2/\|X\|_{\psi_2}^2}$$
όπου $c_1 > 0$ *είναι μια απόλυτη σταθερά.*

(ii) *Για κάθε* $p \geqslant 1$,
$$\|X\|_p \leqslant c_2\sqrt{p}\|X\|_{\psi_2}$$
όπου $c_2 > 0$ *είναι μια απόλυτη σταθερά.*

(iii) *Ισχύει η ανισότητς*
$$\mathbb{E}(e^{X^2/\|X\|_{\psi_2}^2}) \leqslant 2.$$

(iv) *Αν επιπλέον* $\mathbb{E}(X) = 0$ *τότε, για κάθε* $\lambda \in \mathbb{R}$,
$$\mathbb{E}(e^{\lambda X}) \leqslant e^{c_3\|X\|_{\psi_2}^2\lambda^2}$$
όπου $c_3 > 0$ *είναι μια απόλυτη σταθερά.*

Τυπικά παραδείγματα υποκανονικών τυχαίων μεταβλητών είναι τα παρακάτω:

(i) *Τυπικές κανονικές τυχαίες μεταβλητές.* Αν $X \sim N(0,1)$ έχουμε ήδη δει ότι $\|X\|_{\psi_2} \leqslant C$. Γενικότερα, αν $X \sim N(0,\sigma^2)$ τότε $\|X\|_{\psi_2} \leqslant C\sigma$.

(ii) *Συμμετρικές Bernoulli τυχαίες μεταβλητές.* Στην περίπτωση αυτή έχουμε $|X| = 1$, άρα $\|X\|_{\psi_2} = \frac{1}{\sqrt{\ln 2}}$.

(iii) *Φραγμένες τυχαίες μεταβλητές.* Αν $\|X\|_\infty < \infty$, από τον ορισμό της ψ_2-νόρμας είναι φανερό ότι $\|X\|_{\psi_2} \leqslant C\|X\|_\infty$, όπου $C = \frac{1}{\sqrt{\ln 2}}$.

Κλείνουμε αυτή την παράγραφο με μια χρήσιμη εκτίμηση για τη μέση τιμή του maximum πεπερασμένων το πλήθος υποκανονικών (ψ_2) τυχαίων μεταβλητών.

Πρόταση 2.2.6. *Έστω* $X_1, \ldots, X_N$, $N \geqslant 2$, *υποκανονικές τυχαίες μεταβλητές τέτοιες ώστε*
$$\|X_i\|_{\psi_2} \leqslant b$$
για κάθε $i = 1, \ldots, N$. *Τότε,*
$$\mathbb{E}\max_{1\leqslant i\leqslant N}|X_i| \leqslant Cb\sqrt{logN},$$
όπου $C > 0$ *είναι μια απόλυτη σταθερά.*

Απόδειξη. Από τον ορισμό της ψ_2-νόρμας και την ανισότητα Markov, για κάθε $t > 0$ έχουμε

$$\mathbb{P}\left(\max_{1\leqslant i\leqslant N}|X_i|\geqslant t\right)\leqslant\sum_{i=1}^{N}\mathbb{P}\left(|X_i|\geqslant t\right)\leqslant 2Ne^{-t^2/b^2}.$$

Συνεπώς, για κάθε $\alpha > 0$ μπορούμε να γράψουμε

$$\begin{aligned}\mathbb{E}\max_{1\leqslant i\leqslant N}|X_i| &= \int_0^\infty \mathbb{P}\left(\max_{1\leqslant i\leqslant N}|X_i|\geqslant t\right)dt\\ &\leqslant \alpha+\int_\alpha^\infty \mathbb{P}\left(\max_{1\leqslant i\leqslant N}|X_i|\geqslant t\right)dt\\ &\leqslant \alpha+2N\int_\alpha^\infty e^{-t^2/b^2}dt.\end{aligned}$$

Αν επιλέξουμε $\alpha = 2b\sqrt{\log N}$ παίρνουμε

$$\begin{aligned}\int_\alpha^\infty e^{-t^2/b^2}dt &= 2b\sqrt{\log N}\int_1^\infty \exp(-4s^2\log N)ds\leqslant 2b\sqrt{\log N}\int_1^\infty \exp(-4s\log N)ds\\ &\leqslant 2b\sqrt{\log N}\exp(-2\log N)\int_1^\infty e^{-s}ds\\ &\leqslant 2bN^{-2}\sqrt{\log N},\end{aligned}$$

όπου χρησιμοποιήσαμε την

$$\exp(-4s\log N)\leqslant\exp(-2\log N)\cdot e^{-s}$$

που ισχύει για κάθε $s\geqslant 1$. Έπεται ότι

$$\mathbb{E}\max_{1\leqslant i\leqslant N}|X_i|\leqslant Cb\sqrt{\log N}$$

με $C = 4$. □

Η επόμενη πρόταση δείχνει ότι το φράγμα της Πρότασης 2.2.6 είναι βέλτιστο. Αν οι X_i είναι ανεξάρτητες τυπικές κανονικές τυχαίες μεταβλητές τότε ισχύει και η αντίστροφη ανισότητα.

Πρόταση 2.2.7. *Έστω $g_1,\dots,g_N$ ανεξάρτητες τυπικές κανονικές τυχαίες μεταβλητές. Τότε,*

$$\mathbb{E}\max_{1\leqslant j\leqslant N}|g_j|\geqslant c\sqrt{\log N},$$

όπου $c > 0$ είναι μια απόλυτη σταθερά.

Απόδειξη. Παρατηρούμε πρώτα ότι

$$\mathbb{E}\max_{1\leqslant j\leqslant N}|g_j| = \int_{\mathbb{R}^m}\max_{1\leqslant j\leqslant N}|x_j|\,d\gamma_N(x),$$

όπου γ_N είναι το μέτρο του Gauss στον $\mathbb{R}^N$. Χρησιμοποιώντας την ανισότητα

$$\frac{1}{\sqrt{2\pi}}\int_s^\infty e^{-t^2/2}dt \geqslant \frac{1}{\sqrt{2\pi}}\frac{s}{s^2+1}e^{-s^2/2}$$

η οποία ισχύει για κάθε $s>0$, μπορούμε να γράψουμε

$$\begin{aligned}\gamma_N\Big(\Big\{x:\max_{1\leqslant j\leqslant N}|x_j|<s\Big\}\Big) &= (2\pi)^{-N/2}\int_{-s}^s\cdots\int_{-s}^s\exp\Big(-\frac{1}{2}\sum_{j=1}^N x_j^2\Big)\,dx_1\dots dx_N\\ &= \Big(\frac{1}{\sqrt{2\pi}}\int_{-s}^s e^{-t^2/2}dt\Big)^N\\ &\leqslant \Big(1-\frac{2s}{(s^2+1)\sqrt{2\pi}}e^{-s^2/2}\Big)^N.\end{aligned}$$

Αν επιλέξουμε $s=\sqrt{\log N}$, και το N είναι αρκετά μεγάλο, βλέπουμε ότι

$$\gamma_N\Big(\Big\{x:\max_{j\leqslant N}|x_j|\geqslant\sqrt{\log N}\Big\}\Big)\geqslant\frac{1}{2}.$$

Τότε,

$$\mathbb{E}\max_{1\leqslant j\leqslant N}|g_j|\geqslant\sqrt{\log N}\gamma_N\Big(\Big\{x:\max_{j\leqslant N}|x_j|\geqslant\sqrt{\log N}\Big\}\Big)\geqslant\frac{1}{2}\sqrt{\log N}$$

αν το N είναι αρκετά μεγάλο, και έπεται το ζητούμενο. □

2.3 Ανισότητα Khintchine-Kahane

2.3α′ Η ανισότητα του Khintchine

Η κλασική ανισότητα του Khintchine ισχυρίζεται ότι για κάθε $p>0$ υπάρχουν σταθερές $A_p,B_p>0$ με την ακόλουθη ιδιότητα: για κάθε $n\geqslant 1$ και για κάθε n-άδα πραγματικών αριθμών $a_1,\dots,a_n$,

$$A_p\Big(\sum_{i=1}^n a_i^2\Big)^{1/2}\leqslant\Big(\int_{E_2^n}\Big|\sum_{i=1}^n a_i\epsilon_i\Big|^p d\mu_n(\epsilon)\Big)^{1/p}\leqslant B_p\Big(\sum_{i=1}^n a_i^2\Big)^{1/2}. \tag{2.3.1}$$

Δεδομένου ότι

$$\Big(\sum_{i=1}^n a_i^2\Big)^{1/2}=\Big(\int_{E_2^n}\Big|\sum_{i=1}^n a_i\epsilon_i\Big|^2 d\mu_n(\epsilon)\Big)^{1/2}$$

για κάθε $a_1, \dots, a_n$, ένας ισοδύναμος τρόπος διατύπωσης της ανισότητας του Khintchine είναι ο ακόλουθος.

Θεώρημα 2.3.1 (Khintchine). *Για κάθε $p > 0$ υπάρχουν $A_p, B_p > 0$ τέτοιοι ώστε, για κάθε $n \geqslant 1$ και για κάθε $a = (a_1, \dots, a_n) \in \mathbb{R}^n$,*

$$A_p \Big\| \sum_{i=1}^n a_i \epsilon_i \Big\|_{L_2(E_2^n)} \leqslant \Big\| \sum_{i=1}^n a_i \epsilon_i \Big\|_{L_p(E_2^n)} \leqslant B_p \Big\| \sum_{i=1}^n a_i \epsilon_i \Big\|_{L_2(E_2^n)}. \tag{2.3.2}$$

Συμβολίζουμε με A_p^*, B_p^* τις βέλτιστες σταθερές για τις οποίες αληθεύει ο ισχυρισμός του Θεωρήματος 2.3.1. Από την ανισότητα Hölder είναι φανερό ότι $A_p^* = 1$ αν $p \geqslant 2$ και $B_p^* = 1$ αν $0 < p \leqslant 2$. Οι ακριβείς τιμές των A_p^* και B_p^* προσδιορίστηκαν από τον Szarek ($A_1^* = 1/\sqrt{2}$) και τον Haagerup (για κάθε p). Σχετικά με τη συμπεριφορά της σταθεράς B_p^* για μεγάλες τιμές του p είναι σημαντικό το γεγονός ότι $B_p^* \leqslant C\sqrt{p}$, $p \geqslant 1$.

Θα δώσουμε μια απόδειξη αυτής της ανισότητας χρησιμοποιώντας τα αποτελέσματα της προηγούμενης παραγράφου. Ξεκινάμε από την ακόλουθη παρατήρηση.

Πρόταση 2.3.2. *Έστω $X_1, \dots, X_N$ ανεξάρτητες υποκανονικές τυχαίες μεταβλητές με $\mathbb{E}(X_i) = 0$. Τότε, η $S_N := X_1 + \cdots + X_N$ είναι υποκανονική τυχαία μεταβλητή, και*

$$\Big\| \sum_{i=1}^N X_i \Big\|_{\psi_2}^2 \leqslant C \sum_{i=1}^N \|X_i\|_{\psi_2}^2,$$

όπου $C > 0$ είναι μια απόλυτη σταθερά.

Απόδειξη. Για κάθε $\lambda \in \mathbb{R}$ έχουμε (χρησιμοποιώντας την ανεξαρτησία των X_i και το γεγονός ότι είναι υποκανονικές):

$$\begin{aligned} \mathbb{E}\Big[\exp\Big(\sum_{i=1}^N X_i \Big) \Big] &= \prod_{i=1}^N \mathbb{E}(e^{\lambda X_i}) \leqslant \prod_{i=1}^N e^{c_1 \lambda^2 \|X_i\|_{\psi_2}^2} \\ &= e^{\lambda^2 D^2}, \end{aligned}$$

όπου

$$D^2 = c_1 \sum_{i=1}^N \|X_i\|_{\psi_2}^2.$$

Από την ισοδυναμία των (iii) και (iv) στην Πρόταση 2.2.2 έπεται το συμπέρασμα. □

Μπορούμε τώρα, από την Πρόταση 2.2.5, να διατυπώσουμε το συμπέρασμα της Πρότασης 2.3.2 ως εξής:

Θεώρημα 2.3.3 (γενική ανισότητα Hoeffding)**.** *Έστω $X_1,\dots,X_N$ ανεξάρτητες υποκανονικές τυχαίες μεταβλητές με $\mathbb{E}(X_i)=0$. Για κάθε $a=(a_1,\dots,a_N)\in\mathbb{R}^N$ και για κάθε $t>0$ έχουμε*

$$\mathbb{P}\left(\Big|\sum_{i=1}^N a_iX_i\Big|\geqslant t\right)\leqslant 2\exp\left(-\frac{ct^2}{\sum_{i=1}^N a_i^2\|X_i\|_{\psi_2}^2}\right)\leqslant 2\exp\left(-\frac{ct^2}{D^2|a|^2}\right),$$

όπου $D:=\max\limits_{1\leqslant i\leqslant N}\|X_i\|_{\psi_2}$.

Απόδειξη. Απλώς παρατηρούμε ότι

$$\Big\|\sum_{i=1}^N a_iX_i\|_{\psi_2}^2\leqslant C\sum_{i=1}^N\|a_iX_i\|_{\psi_2}^2=C\sum_{i=1}^N a_i^2\|X_i\|_{\psi_2}^2\leqslant CD^2|a|^2.$$

Κατόπιν, χρησιμοποιούμε την Πρόταση 2.2.6. □

Απόδειξη του Θεωρήματος 2.3.1. Η ανισότητα του Khintchine (για $p\geqslant 2$) είναι απλή εφαρμογή των παραπάνω. Θεωρούμε ανεξάρτητες συμμετρικές τυχαίες μεταβλητές Bernoulli $\epsilon_1,\dots,\epsilon_n$. Γνωρίζουμε ότι $\|\epsilon_i\|_{\psi_2}=\frac{1}{\sqrt{\ln 2}}$, άρα $D=\frac{1}{\sqrt{\ln 2}}$. Συνεπώς, για κάθε $n\geqslant 1$, για κάθε $a=(a_1,\dots,a_n)\in\mathbb{R}^n$ και για κάθε $t>0$ έχουμε

$$\mathbb{P}\left(\Big|\sum_{i=1}^n a_i\epsilon_i\Big|\geqslant t\right)\leqslant 2\exp\left(-\frac{ct^2}{|a|^2}\right).$$

Τώρα, είναι εύκολο να εκτιμήσουμε την p-νόρμα της $\sum_{i=1}^n a_i\epsilon_i$ γράφοντας

$$\begin{aligned}\Big\|\sum_{i=1}^n a_i\epsilon_i\Big\|_{L_p(E_2^n)}^p &= \int_0^\infty pt^{p-1}\mathbb{P}\left(\Big|\sum_{i=1}^n a_i\epsilon_i\Big|\geqslant t\right)dt\\ &\leqslant \int_0^\infty pt^{p-1}2\exp\left(-\frac{ct^2}{|a|^2}\right)dt\\ &\leqslant (c_2\sqrt{p}|a|)^p,\end{aligned}$$

απ΄ όπου έπεται η δεξιά ανισότητα στην (2.3.1). Για να χειριστούμε την αριστερή ανισότητα στην περίπτωση που $0<p<2$, χρησιμοποιούμε το εξής τέχνασμα. Για δοθέν $0<p<2$, ο 2 είναι κυρτός συνδυασμός των p και 4. Μπορούμε δηλαδή να βρούμε $\theta\in(0,1)$ τέτοιον ώστε $2=p\theta+4(1-\theta)$: η τιμή του θ είναι $2/(4-p)$. Αν $Y=\sum_{i=1}^n a_k\epsilon_i$ τότε, από την ανισότητα Hölder,

$$\mathbb{E}|Y|^2=\mathbb{E}(|Y|^{p\theta}|Y|^{4(1-\theta)})\leqslant(\mathbb{E}|Y|^p)^\theta\left(\mathbb{E}|Y|^4\right)^{1-\theta}. \tag{2.3.3}$$

Έχουμε δείξει ότι $\|Y\|_{L_4} \leqslant c_0 \|Y\|_{L_2}$. Άρα, η (2.3.3) μας δίνει

$$\mathbb{E}|Y|^2 \leqslant c_0^{4(1-\theta)} \left(\mathbb{E}|Y|^p\right)^\theta \left(\mathbb{E}|Y|^2\right)^{2(1-\theta)}.$$

Παρατηρήστε ότι $1 - 2(1-\theta) = p\theta/2$. Έπεται ότι

$$(\mathbb{E}|Y|^2)^{\frac{p\theta}{2}} \leqslant c_0^{4(1-\theta)} \left(\mathbb{E}|Y|^p\right)^\theta ,$$

απ' όπου έπεται η

$$\|Y\|_{L_2} \leqslant c_p \|Y\|_{L_p}$$

για κάποια σταθερά $c_p > 0$ που εξαρτάται μόνο από το p. □

2.3β′ Η ανισότητα του Kahane

Η ανισότητα Kahane-Khintchine γενικεύει την ανισότητα του Khintchine.

Θεώρημα 2.3.4. *Υπάρχει σταθερά K ώστε για κάθε χώρο με νόρμα X, για κάθε $n \in \mathbb{N}$, για κάθε $x_1, \ldots, x_n \in X$ και για κάθε $p \geqslant 1$,*

$$\left(\mathbb{E}\Big\| \sum_{i\leqslant n} \epsilon_i x_i \Big\|^p\right)^{1/p} \leqslant 2\mathbb{E}\Big\| \sum_{i\leqslant n} \epsilon_i x_i \Big\| + K\sigma\sqrt{p},$$

όπου

$$\sigma^2 = \sup\Big\{ \sum_{i\leqslant n} |x^*(x_i)|^2 : x^* \in X^*, \|x^*\| \leqslant 1\Big\}.$$

Παρατήρηση 2.3.5. Από το Θεώρημα 2.3.4 έπεται ότι, ειδικότερα, υπάρχει σταθερά $K > 0$ ώστε για κάθε χώρο με νόρμα X, για κάθε $n \in \mathbb{N}$, για κάθε $x_1, \ldots, x_n \in X$ και για κάθε $p \geqslant 1$,

$$\Big(\mathbb{E}\Big\| \sum_{i\leqslant n} \epsilon_i x_i \Big\|^p\Big)^{1/p} \leqslant (2 + K\sqrt{p})\mathbb{E}\Big\| \sum_{i\leqslant n} \epsilon_i x_i \Big\|.$$

Για την (2.3.5) παρατηρούμε ότι, από την ανισότητα του Khintchine, αν $\|x^*\| \leqslant 1$ τότε

$$\Big(\sum_{i\leqslant n} |x^*(x_i)|^2\Big)^{1/2} \leqslant \sqrt{2}\,\mathbb{E}\Big| \sum_{i\leqslant n} \epsilon_i x^*(x_i)\Big| \leqslant \sqrt{2}\,\mathbb{E}\Big\| \sum_{i\leqslant n} \epsilon_i x_i \Big\|,$$

το οποίο δείχνει ότι

$$\sigma \leqslant \mathbb{E}\Big\| \sum_{i\leqslant n} \epsilon_i x_i \Big\|.$$

Η απόδειξη του Θεωρήματος 2.3.4 βασίζεται στο εξής πόρισμα του Θεωρήματος 1.4.10.

Πόρισμα 2.3.6. *Έστω X χώρος με νόρμα και $(x_i)_{i\leqslant n}$ ακολουθία διανυσμάτων στον X. Αν M είναι ένας μέσος Lévy της $\|\sum_{i\leqslant n}\epsilon_i x_i\|$ στον E_2^n τότε, για κάθε $t\geqslant 0$,*

$$\mu_n\left(\{|\ \|\sum_{i\leqslant n}\epsilon_i x_i\| - M\ |\geqslant t\}\right)\leqslant 4e^{-t^2/8\sigma^2}.$$

Απόδειξη. Θεωρούμε την $f(u)=\|\sum_{i\leqslant n}u_i x_i\|$. Χρησιμοποιώντας τις ιδιότητες της νόρμας ελέγχουμε εύκολα ότι η f είναι κυρτή συνάρτηση. Έστω $x^*\in X^*$ με $\|x^*\|\leqslant 1$ και $u,v\in\mathbb{R}^n$. Από την ανισότητα Cauchy-Schwarz έχουμε

$$\begin{aligned}
\left|x^*\left(\sum_{i\leqslant n}u_i x_i-\sum_{i\leqslant n}v_i x_i\right)\right| &= \left|\sum_{i\leqslant n}u_i x^*(x_i)-\sum_{i\leqslant n}v_i x^*(x_i)\right| \\
&= \left|\sum_{i\leqslant n}(u_i-v_i)x^*(x_i)\right| \\
&\leqslant \left(\sum_{i\leqslant n}|x^*(x_i)|^2\right)^{1/2}\left(\sum_{i\leqslant n}(u_i-v_i)^2\right)^{1/2} \\
&\leqslant \sigma\|u-v\|_2.
\end{aligned}\tag{2.3.4}$$

Από το Θεώρημα Hahn-Banach συμπεραίνουμε ότι

$$|f(u)-f(v)|\leqslant\left\|\sum_{i\leqslant n}u_i x_i-\sum_{i\leqslant n}v_i x_i\right\|\leqslant\sigma\|u-v\|_2,$$

επομένως η f είναι Lipschitz συνεχής με σταθερά σ. Εφαρμόζοντας το Θεώρημα 1.4.10 για την f έχουμε το ζητούμενο. □

Μπορούμε τώρα να δώσουμε μια απόδειξη της ανισότητας Khintchine-Kahane με βέλτιστη εξάρτηση από το p.

Απόδειξη του θεωρήματος 2.3.4. Θεωρούμε την $f:E_2^n\to\mathbb{R}^+$ με

$$f(\epsilon_1,\dots,\epsilon_n)=\big|\ \|\sum_{i\leqslant n}\epsilon_i x_i\|-M\ \big|.$$

Από το Πόρισμα 2.3.6, κάνοντας την αλλαγή μεταβλητής $x = t^2/8\sigma^2$ έχουμε

$$\begin{aligned}\int_{E_2^n} \Big| \, \|\sum_{i\leqslant n} \epsilon_i x_i\| - M \, \Big|^p d\mu_n(\epsilon) &= p\int_0^\infty t^{p-1}\mu_n(\{\epsilon : |\|\sum_{i\leqslant n} \epsilon_i x_i\| - M| \geqslant t\})dt \\ &\leqslant 4\int_0^\infty pt^{p-1}e^{-t^2/8\sigma^2}dt \\ &= 2^{p+1}p(\sqrt{2}\sigma)^p\int_0^\infty e^{-x}x^{p/2-1}dx \leqslant (K\sigma\sqrt{p})^p.\end{aligned}$$

Άρα,

$$\left(\int_{E_2^n} |\, \|\sum_{i\leqslant n} \epsilon_i x_i\| - M|^p d\mu_n(\epsilon)\right)^{1/p} \leqslant K\sigma\sqrt{p}.$$

Από την τριγωνική ανισότητα,

$$\left(\int_{E_2^n} \|\sum_{i\leqslant n} \epsilon_i x_i\|^p d\mu_n(\epsilon)\right)^{1/p} \leqslant M + K_1\sigma p^{1/2}$$

για κάθε $p \geqslant 1$. Τέλος, παρατηρούμε ότι $M \leqslant 2\mathbb{E}\,\|\sum_{i\leqslant n} \epsilon_i x_i\|$ από την ανισότητα Markov.
□

2.3γ΄ Ανισότητα Kahane-Khinthcine για λογαριθμικά κοίλα μέτρα

Ορισμός 2.3.7. Μια συνάρτηση $f : \mathbb{R}^n \to \mathbb{R}^+$ λέγεται λογαριθμικά κοίλη αν, για κάθε $x, y \in \mathbb{R}^n$ και για κάθε $\lambda \in (0, 1)$,

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

Θεωρούμε την κλάση $\mathcal{M}_n$ όλων των Borel μέτρων πιθανότητας μ στον $\mathbb{R}^n$ που έχουν λογαριθμικά κοίλη πυκνότητα f_μ: η f_μ είναι λογαριθμικά κοίλη, το ολοκλήρωμά της ισούται με 1 και για κάθε σύνολο Borel A ισχύει

$$\mu(A) = \int_A f_\mu(x)dx.$$

Παραδείγματα μέτρων στην $\mathcal{M}_n$. (α) Έστω K κυρτό σώμα όγκου 1 στον $\mathbb{R}^n$. Ορίζουμε ένα μέτρο πιθανότητας μ_K στον $\mathbb{R}^n$, θέτοντας

$$\mu_K(A) = |K \cap A| = \int_A \chi_K(x)dx$$

για κάθε Borel $A \subseteq \mathbb{R}^n$. Χρησιμοποιώντας την κυρτότητα του K ελέγχουμε εύκολα ότι η χ_K είναι λογαριθμικά κοίλη συνάρτηση, άρα $\mu_K \in \mathcal{M}_n$.

(β) Για κάθε $c > 0$, η συνάρτηση $f_c(x) = \exp(-c\|x\|_2^2)$ είναι άρτια και λογαριθμικά κοίλη στον $\mathbb{R}^n$: παρατηρούμε ότι η Ευκλείδεια νόρμα είναι κυρτή συνάρτηση, και η $t \mapsto ct^2$ είναι επίσης κυρτή. Άρα η σύνθεσή τους $c\|x\|_2^2 = -\log f(x)$ είναι μια άρτια, κυρτή συνάρτηση. Αυτό σημαίνει ότι, για κάθε $c > 0$, το μέτρο

$$\gamma_{r,c}(A) = \frac{1}{I(c)} \int_A \exp(-c\|x\|_2^2)dx$$

όπου $I(c) = \int_{\mathbb{R}^r} \exp(-c\|x\|_2^2)dx$, ανήκει στην κλάση $\mathcal{M}_n$. Ειδικότερα, το μέτρο του Gauss $\gamma_n \in \mathcal{M}_n$.

Ορισμός 2.3.8. Ένα Borel μέτρο πιθανότητας μ στον $\mathbb{R}^n$ λέγεται λογαριθμικά κοίλο αν για οποιαδήποτε μη κενά Borel υποσύνολα A, B του $\mathbb{R}^n$ και για κάθε $\lambda \in (0,1)$,

$$\mu(\lambda A + (1-\lambda)B) \geq [\mu(A)]^\lambda [\mu(B)]^{1-\lambda}.$$

Η πρόταση που ακολουθεί δείχνει ότι η κλάση $\mathcal{M}_n$ περιέχεται στην κλάση των λογαριθμικά κοίλων μέτρων.

Πρόταση 2.3.9. *Αν $\mu \in \mathcal{M}_n$, τότε το μ είναι λογαριθμικά κοίλο.*

Απόδειξη. Αφού $\mu \in \mathcal{M}_n$, υπάρχει $f : \mathbb{R}^r \to \mathbb{R}^+$ λογαριθμικά κοίλη, ώστε $\mu(A) = \int_A f(x)dx$. Έστω $\lambda \in (0,1)$ και A, B μη κενά Borel υποσύνολα του $\mathbb{R}^n$. Θέλουμε να δείξουμε ότι

$$\begin{aligned} \mu(\lambda A + (1-\lambda)B) &= \int_{\mathbb{R}^r} \chi_{\lambda A + (1-\lambda)B}(x) f(x) dx \\ &\geqslant \left(\int_{\mathbb{R}^r} \chi_A(x) f(x) dx \right)^\lambda \left(\int_{\mathbb{R}^r} \chi_B(x) f(x) dx \right)^{1-\lambda} \\ &= [\mu(A)]^\lambda [\mu(B)]^{1-\lambda}. \end{aligned}$$

Ορίζουμε $w(x) = \chi_A(x) f(x)$, $g(x) = \chi_B(x) f(x)$ και $h(x) = \chi_{\lambda A + (1-\lambda)B}(x) f(x)$. Εύκολα ελέγχουμε ότι οι g, h, w ικανοποιούν τις υποθέσεις της ανισότητας Prékopa-Leindler, απ' όπου έπεται το ζητούμενο. □

Σημείωση. Ένα θεώρημα του Borell δείχνει ότι κάθε μη εκφυλισμένο λογαριθμικά κοίλο μέτρο πιθανότητας στον $\mathbb{R}^n$ ανήκει στην κλάση $\mathcal{M}_n$.

Θεώρημα 2.3.10. *Έστω μ ένα λογαριθμικά κοίλο μέτρο πιθανότητας στον $\mathbb{R}^n$ με την ιδιότητα $\mu(H) < 1$ για κάθε υπερεπίπεδο H. Τότε, το μ είναι απολύτως συνεχές ως προς το μέτρο Lebesgue και έχει μια λογαριθμικά κοίλη πυκνότητα f, δηλαδή $d\mu(x) = f(x)\,dx$.*

Σκοπός μας εδώ είναι να αποδείξουμε την ανισότητα Kahane-Khintchine για λογαριθμικά κοίλα μέτρα πιθανότητας. Η ακριβής διατύπωση είναι η εξής.

Θεώρημα 2.3.11. *Έστω $\mu \in \mathcal{M}_n$. Αν $g : \mathbb{R}^n \to \mathbb{R}$ είναι μια συνάρτηση που ικανοποιεί την $|g(tx)| = |t||g(x)|$ για κάθε $t \in \mathbb{R}$ και $|g(x+y)| \leqslant |g(x)| + |g(y)|$ για $x, y \in \mathbb{R}^n$, τότε, για κάθε $q > p \geqslant 1$, έχουμε*

$$\left(\int |g|^p \,d\mu\right)^{1/p} \leqslant \left(\int |g|^q \,d\mu\right)^{1/q} \leqslant c\frac{q}{p}\left(\int |g|^p \,d\mu\right)^{1/p},$$

όπου $c > 0$ είναι μια απόλυτη σταθερά.

Η απόδειξη θα βασιστεί στο *λήμμα του Borell.*

Λήμμα 2.3.12. *Έστω $\mu \in \mathcal{M}_n$. Για κάθε συμμετρικό κυρτό σύνολο A στον $\mathbb{R}^n$ με $\mu(A) = \alpha \in (0,1)$ και για κάθε $t > 1$ έχουμε*

$$1 - \mu(tA) \leqslant \alpha\left(\frac{1-\alpha}{\alpha}\right)^{\frac{t+1}{2}}.$$

Απόδειξη. Χρησιμοποιώντας τη συμμετρία και την κυρτότητα του A ελέγχουμε ότι

$$\frac{2}{t+1}\mathbb{R}^n \setminus (tA) + \frac{t-1}{t+1}A \subseteq \mathbb{R}^n \setminus A.$$

για κάθε $t > 1$. Κατόπιν, χρησιμοποιούμε το γεγονός ότι το μ είναι λογαριθμικά κοίλο για να φτάσουμε το συμπέρασμα. □

Απόδειξη του Θεωρήματος 2.3.11. Γράφουμε $\|g\|_p^p := \int |g|^p \,d\mu$. Τότε, το σύνολο

$$A = \{x \in \mathbb{R}^n : |g(x)| \leqslant 3\|g\|_p\}$$

είναι συμμετρικό και κυρτό. Επίσης, για κάθε $t > 0$ έχουμε

$$tA = \{x \in \mathbb{R}^n : |g(x)| \leqslant 3t\|g\|_p\}$$

και $\mu(A) \geqslant 1 - 3^{-p} \geqslant \frac{2}{3}$. Από το λήμμα του Borell βλέπουμε ότι

$$\mu(x : |g(x)| \geqslant 3t\|g\|_p) \leqslant \frac{1}{3}e^{-c_1 p(t-1)}$$

για κάθε $t > 1$, όπου $c_1 = \frac{\ln 2}{2}$. Τώρα, μπορούμε να γράψουμε

$$\begin{aligned}\int |g|^q \, d\mu &= \int_0^\infty q s^{q-1} \mu(x : |g(x)| \geqslant s) \, ds \\ &\leqslant (3\|g\|_p)^q + \frac{1}{3}(3\|g\|_p)^q \int_1^\infty q t^{q-1} e^{-c_1 p(t-1)} \, dt \\ &\leqslant (3\|g\|_p)^q + \frac{e^{c_1 p}}{3}(3\|g\|_p)^q \int_0^\infty q t^{q-1} e^{-c_1 p t} \, dt \\ &\leqslant (3\|g\|_p)^q + \frac{e^{c_1 p}}{3} \left(\frac{3\|g\|_p}{c_1 p} \right)^q \Gamma(q+1).\end{aligned}$$

Από τον τύπο του Stirling και από την $(a+b)^{1/q} \leqslant a^{1/q} + b^{1/q}$ για κάθε $a, b > 0$ και $q \geqslant 1$, έπεται ότι $\|g\|_{L_q(\mu)} \leqslant c\frac{q}{p}\|g\|_{L_p(\mu)}$. □

2.4 Υποεκθετικές τυχαίες μεταβλητές και ανισότητες Bernstein

Ορισμός 2.4.1 (υποεκθετική τυχαία μεταβλητή)**.** Μια τυχαία μεταβλητή X λέγεται υποεκθετική με σταθερά $\alpha > 0$ αν για κάθε $t > 0$ ισχύει

$$\mathbb{P}(|X| \geqslant t) \leqslant 2e^{-t/\alpha}.$$

Η επόμενη πρόταση δίνει χαρακτηρισμούς των υποεκθετικών τυχαίων μεταβλητών. Η απόδειξη είναι ανάλογη με αυτήν της Πρότασης 2.2.2.

Πρόταση 2.4.2. *Έστω X μια τυχαία μεταβλητή. Οι παρακάτω ιδιότητες είναι ισοδύναμες (με τις σταθερές α, β, γ να διαφέρουν το πολύ κατά μία απόλυτη σταθερά).*

(i) *Υπάρχει $\alpha > 0$ ώστε: για κάθε $t > 0$,*

$$\mathbb{P}(|X| \geqslant t) \leqslant 2e^{-t/\alpha}.$$

(ii) *Υπάρχει $\beta > 0$ ώστε: για κάθε $p \geqslant 1$,*

$$\|X\|_p \leqslant \beta p.$$

(iii) *Υπάρχει $\gamma > 0$ ώστε*

$$\mathbb{E}(e^{|X|/\gamma}) \leqslant 2.$$

Με βάση τον ορισμό της ψ_1-νόρμας, μπορούμε να επαναδιατυπώσουμε τα συμπεράσματα της Πρότασης 2.4.2 ως εξής:

Πρόταση 2.4.3. *Έστω X μια τυχαία μεταβλητή. Αν $X \in L_{\psi_1}$ τότε:*

(i) *Για κάθε $t > 0$,*
$$\mathbb{P}(|X| \geqslant t) \leqslant 2e^{-c_1 t/\|X\|_{\psi_1}}$$
όπου $c_1 > 0$ είναι μια απόλυτη σταθερά.

(ii) *Για κάθε $p \geqslant 1$,*
$$\|X\|_p \leqslant c_2 p \|X\|_{\psi_1}$$
όπου $c_2 > 0$ είναι μια απόλυτη σταθερά.

(iii) *Ισχύει η ανισότητα*
$$\mathbb{E}(e^{|X|/\|X\|_{\psi_1}}) \leqslant 2.$$

Τα επόμενα δύο λήμματα δείχνουν ότι το γινόμενο δύο υποκανονικών τυχαίων μεταβλητών είναι υποεκθετική τυχαία μεταβλητή. Ειδικότερα, το τετράγωνο μιας υποκανονικής τυχαίας μεταβλητής είναι υποεκθετική τυχαία μεταβλητή.

Λήμμα 2.4.4. *Μια τυχαία μεταβλητή X είναι υποκανονική αν και μόνο αν η X^2 είναι υποεκθετική. Επιπλέον,*
$$\|X^2\|_{\psi_1} = \|X\|_{\psi_2}^2.$$

Απόδειξη. Αρκεί να παρατηρήσουμε ότι η $\|X^2\|_{\psi_1}$ είναι το infimum των $\alpha > 0$ για τους οποίους $\mathbb{E}\left(\exp(X^2/\alpha)\right) \leqslant 2$ ενώ η $\|X\|_{\psi_2}$ είναι το infimum των $\beta > 0$ για τους οποίους $\mathbb{E}\left(\exp(X^2/\beta^2)\right) \leqslant 2$. Οι δύο ορισμοί συμπίπτουν αν θέσουμε $\alpha = \beta^2$. □

Λήμμα 2.4.5. *Έστω X και Y υποκανονικές τυχαίες μεταβλητές. Τότε, η XY είναι υποεκθετική και*
$$\|XY\|_{\psi_1} \leqslant \|X\|_{\psi_2}\|Y\|_{\psi_2}.$$

Απόδειξη. Θέτουμε $\alpha = \|X\|_{\psi_2}$ και $\beta = \|Y\|_{\psi_2}$. Παρατηρούμε ότι
$$\begin{aligned}
\mathbb{E}\left(e^{\frac{|XY|}{\alpha\beta}}\right) &\leqslant \mathbb{E}\left(e^{\frac{X^2}{2\alpha^2}+\frac{Y^2}{2\beta^2}}\right) \\
&= \mathbb{E}\left(e^{\frac{X^2}{2\alpha^2}}e^{\frac{Y^2}{2\beta^2}}\right) \\
&\leqslant \frac{1}{2}\mathbb{E}\left(e^{\frac{X^2}{\alpha^2}} + e^{\frac{Y^2}{\beta^2}}\right) \\
&= \frac{1}{2}(2+2) = 2,
\end{aligned}$$

χρησιμοποιώντας δύο φορές την $ab \leqslant \frac{1}{2}(a^2+b^2)$. Αυτό αποδεικνύει ότι $\|XY\|_{\psi_1} \leqslant \alpha\beta$. □

Η επόμενη πρόταση δείχνει ότι η ροπογεννήτρια μιας υποεκθετικής τυχαίας μεταβλητής είναι πεπερασμένη σε μια περιοχή του 0.

Πρόταση 2.4.6. *Έστω X μια υποεκθετική τυχαία μεταβλητή με $\mathbb{E}(X)=0$. Αν $|\lambda| \leqslant c_1/\|X\|_{\psi_1}$, όπου $c_1>0$ είναι μια απόλυτη σταθερά, τότε*

$$\mathbb{E}\left(e^{\lambda X}\right) \leqslant e^{c_2\lambda^2\|X\|_{\psi_1}^2},$$

όπου $c_2>0$ είναι μια απόλυτη σταθερά.

Απόδειξη. Χωρίς περιορισμό της γενικότητας υποθέτουμε ότι $\|X\|_{\psi_1}=1$. Γράφουμε

$$\mathbb{E}\left(e^{\lambda X}\right) = \mathbb{E}\left(1+\lambda X+\sum_{k=2}^{\infty}\frac{(\lambda X)^k}{k!}\right) = 1+\sum_{k=2}^{\infty}\frac{\lambda^k\mathbb{E}(X^k)}{k!},$$

χρησιμοποιώντας την υπόθεση ότι $\mathbb{E}(X)=0$. Γνωρίζουμε ότι $\mathbb{E}(X^k) \leqslant (c_3k)^k$, από την Πρόταση 2.4.3. Παίρνοντας υπόψη και την $k! \geqslant (k/e)^k$ βλέπουμε ότι

$$\mathbb{E}\left(e^{\lambda X}\right) \leqslant 1+\sum_{k=2}^{\infty}\frac{(c_3k\lambda)^l}{(k/e)^l} = 1+\sum_{k=2}^{\infty}(c_4\lambda)^k.$$

Αν $|c_4\lambda| < \frac{1}{2}$ τότε η γεωμετρική αυτή σειρά συγκλίνει και έχουμε

$$\mathbb{E}\left(e^{\lambda X}\right) \leqslant 1+2(c_\lambda)^2 \leqslant e^{2(c_4\lambda)^2}.$$

Αυτό αποδεικνύει το ζητούμενο. □

Στο υπόλοιπο αυτής της παραγράφου συζητάμε μερικές ανισότητες τύπου Bernstein για αθροίσματα ανεξάρτητων τυχαίων μεταβλητών $X_1,\ldots,X_N$, αρχίζοντας από την υποεκθετική περίπτωση.

Θεώρημα 2.4.7. *Έστω $X_1,\ldots,X_N$ ανεξάρτητες υποεκθετικές τυχαίες μεταβλητές με μέση τιμή $\mathbb{E}(X_i)=0$. Τότε, για κάθε $t>0$,*

$$\mathbb{P}\Big(\Big\{\Big|\sum_{i=1}^{N}X_i\Big| \geqslant t\Big\}\Big) \leqslant 2\exp\Big(-c\min\Big\{\frac{t^2}{\sum_{i=1}^{N}\|X_i\|_{\psi_1}^2},\frac{t}{\max_{1\leqslant i\leqslant N}\|X_i\|_{\psi_1}}\Big\}\Big),$$

όπου $c>0$ είναι μια απόλυτη σταθερά.

Απόδειξη. Θέτουμε $S_N = X_1 + \cdots + X_N$. Ακολουθώντας τη συνήθη διαδικασία, για κάθε $\lambda > 0$ έχουμε

$$\mathbb{P}(S_N \geqslant t) \leqslant e^{-\lambda t} \prod_{i=1}^{N} \mathbb{E}\left(e^{\lambda X_i}\right). \tag{2.4.1}$$

Αν επιλέξουμε

$$|\lambda| \leqslant \frac{c_1}{\max_{1\leqslant i\leqslant N} \|X_i\|_{\psi_1}}, \tag{2.4.2}$$

από την Πρόταση 2.4.6 έχουμε

$$\mathbb{E}\left(e^{\lambda X_i}\right) \leqslant e^{c_2\lambda^2\|X_i\|_{\psi_1}^2}, \qquad i = 1, \ldots, N.$$

Αντικαθιστώντας στην (2.4.1) παίρνουμε

$$\mathbb{P}(S_N \geqslant t) \leqslant \exp(-\lambda t + C^2\lambda^2\sigma^2),$$

όπου

$$\sigma^2 = \sum_{i=1}^{N} \|X_i\|_{\psi_1}^2.$$

Τώρα, ελαχιστοποιούμε αυτή την ποσότητα ως προς λ, υπό τον περιορισμό (2.4.2). Η βέλτιστη επιλογή είναι

$$\lambda = \min\left\{\frac{t}{2c_2\sigma^2}, \frac{c_1}{\max_{1\leqslant i\leqslant N} \|X_i\|_{\psi_1}}\right\},$$

η οποία μας δίνει

$$\mathbb{P}(S_N \geqslant t) \leqslant \exp\Big(-\min\Big\{\frac{t^2}{4c_2\sum_{i=1}^{N}\|X_i\|_{\psi_1}^2}, \frac{c_1 t}{2\max_{1\leqslant i\leqslant N}\|X_i\|_{\psi_1}}\Big\}\Big).$$

Επαναλαμβάνοντας το ίδιο επιχείρημα με τις $-X_i$ στη θέση των X_i παίρνουμε το ίδιο φράγμα για την $\mathbb{P}(-S_n \geqslant t)$. Συνδυάζοντας τα δύο φράγματα έχουμε το θεώρημα. □

Εφαρμόζοντας το προηγούμενο θεώρημα για τις a_iX_i αντί των X_i έχουμε το εξής.

Θεώρημα 2.4.8. *Έστω $X_1, \ldots, X_N$ ανεξάρτητες υποεκθετικές τυχαίες μεταβλητές με μέση τιμή $\mathbb{E}(X_i) = 0$. Υποθέτουμε ότι, για κάποιον $M > 0$, ισχύει $\|X_i\|_{L_{\psi_1}} \leqslant M$ για κάθε $i = 1, \ldots, N$. Τότε, για κάθε $a = (a_1, \ldots, a_N) \in \mathbb{R}^N$ και για κάθε $t > 0$,*

$$\mathbb{P}\Big(\Big\{\Big|\sum_{i=1}^{N} a_iX_i\Big| \geqslant t\Big\}\Big) \leqslant 2\exp\Big(-c\min\Big\{\frac{t^2}{M^2|a|^2}, \frac{t}{M\|a\|_\infty}\Big\}\Big),$$

όπου $c > 0$ είναι μια απόλυτη σταθερά.

Μπορούμε επίσης να διατυπώσουμε την ανισότητα Bernstein ως ποσοτική μορφή του νόμου των μεγάλων αριθμών.

Θεώρημα 2.4.9. *Έστω $X_1,\dots,X_N$ ανεξάρτητες υποεκθετικές τυχαίες μεταβλητές με μέση τιμή $\mathbb{E}(X_i)=0$. Υποθέτουμε ότι, για κάποιον $M>0$, ισχύει $\|X_i\|_{L_{\psi_1}}\leqslant M$ για κάθε $i=1,\dots,N$. Τότε, για κάθε $t>0$,*

$$\mathbb{P}\Big(\Big\{\Big|\frac{1}{N}\sum_{i=1}^{N}X_i\Big|\geqslant t\Big\}\Big)\leqslant 2\exp\Big(-cN\,\min\Big\{\frac{t^2}{M^2},\frac{t}{M}\Big\}\Big),$$

όπου $c>0$ είναι μια απόλυτη σταθερά.

Δίνουμε δύο ακόμα παραδείγματα ανισοτήτων τύπου Bernstein, με διαφορετικές υποθέσεις για τις X_i.

Θεώρημα 2.4.10. *Έστω $X_1,\dots,X_N$ ανεξάρτητες τυχαίες μεταβλητές με μέση τιμή $\mathbb{E}(X_i)=0$. Υποθέτουμε ότι για κάποιον $M>0$ ισχύει $\|X_i\|_{\infty}\leqslant M$ για κάθε $i=1,\dots,N$. Θέτουμε $\sigma^2=\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}(X_i^2)$. Τότε, για κάθε $t>0$,*

$$\mathbb{P}\Big(\Big\{\sum_{j=1}^{N}X_j\geqslant tN\Big\}\Big)\leqslant \exp\Big(-\frac{\sigma^2N}{M^2}F\Big(\frac{Mt}{\sigma^2}\Big)\Big),$$

όπου $F(u)=(1+u)\log(1+u)-u$, $u>0$.

Απόδειξη. Έστω $t>0$. Από την ανισότητα του Markov και από την ανεξαρτησία των X_i έχουμε

$$\begin{aligned}\mathbb{P}\Big(\sum_{j=1}^{N}X_j\geqslant tN\Big)&=\mathbb{P}\Big(\Big\{\exp\Big(\frac{\lambda}{N}\sum_{j=1}^{N}X_j\Big)\geqslant e^{\lambda t}\Big\}\Big)\\&\leqslant e^{-\lambda t}\mathbb{E}\,\exp\Big(\frac{\lambda}{N}\sum_{j=1}^{N}X_j\Big)\\&=e^{-\lambda t}\prod_{j=1}^{N}\mathbb{E}\,\exp(\lambda X_j/N)\end{aligned}$$

για κάθε $\lambda>0$. Στη συνέχεια παρατηρούμε ότι, αφού $\mathbb{E}(X_j)=0$,

$$\begin{aligned}\mathbb{E}\,\exp(\lambda X_j/N)&=1+\sum_{k=2}^{\infty}\frac{\lambda^k\mathbb{E}(X_j^k)}{N^kk!}\leqslant 1+\mathbb{E}(X_j^2)\sum_{k=2}^{\infty}\frac{\lambda^kM^{k-2}}{N^kk!}\\&=1+\frac{\mathbb{E}(X_j^2)}{M^2}\Big(e^{\frac{\lambda M}{N}}-\frac{\lambda M}{N}-1\Big).\end{aligned}$$

Αφού $e^u \geqslant 1 + u$ για κάθε $u \in \mathbb{R}$, παίρνουμε

$$\prod_{j=1}^{N} \mathbb{E}\exp(\lambda X_j/N) \leqslant \exp\Big(\frac{\sum_{j=1}^{N}\mathbb{E}(X_j)^2}{M^2}\Big(e^{\frac{\lambda M}{N}} - \frac{\lambda M}{N} - 1\Big)\Big).$$

Από τον ορισμό του σ^2 έπεται ότι

$$\mathbb{P}\Big(\sum_{j=1}^{N} X_j \geqslant tN\Big) \leqslant \exp\Big(\frac{\sigma^2 N}{M^2}\Big(e^{\frac{\lambda M}{N}} - \frac{\lambda M}{N} - 1\Big) - \lambda t\Big)$$

για κάθε $\lambda > 0$. Επιλέγουμε το λ έτσι ώστε να ικανοποιείται η $\exp(\lambda M/N) = 1 + tM/\sigma^2$ και έπεται το ζητούμενο. □

Στο επόμενο θεώρημα υποθέτουμε ότι έχουμε κάποιο ομοιόμορφο φράγμα για την L_1 και την L_∞-νόρμα των X_i.

Θεώρημα 2.4.11. *Έστω $X_1, \ldots, X_N$ ανεξάρτητες τυχαίες μεταβλητές με μέση τιμή $\mathbb{E}(X_i) = 0$. Υποθέτουμε ότι οι X_i είναι φραγμένες και ότι $\mathbb{E}|X_i| \leqslant 2$ και $\|X_i\|_\infty \leqslant M$ για κάθε $i = 1, \ldots, N$ και κάποια σταθερά $M > 0$. Τότε, για κάθε $0 < t < 1$,*

$$\mathbb{P}\Big(\Big\{\Big|\sum_{i=1}^{N} X_i\Big| \geqslant tN\Big\}\Big) \leqslant 2\exp(-t^2N/(8M)).$$

Απόδειξη. Χρησιμοποιώντας την $e^x \leqslant 1 + x + x^2$, που ισχύει για $0 \leqslant x \leqslant 1$, μαζί με τις υποθέσεις $\mathbb{E}(X_i) = 0$ και $|X_i| \leqslant M$, βλέπουμε ότι αν $0 < \lambda \leqslant 1/M$ τότε

$$\mathbb{E}\exp(\lambda X_i) \leqslant 1 + \lambda^2\mathbb{E}(X_i^2) \leqslant 1 + \lambda^2\|X_i\|_1\|X_i\|_\infty \leqslant \exp(2\lambda^2 M).$$

Από την ανεξαρτησία των X_i έπεται ότι

$$\mathbb{E}\exp\Big(\sum_{i=1}^{N}\lambda X_i\Big) = \prod_{i=1}^{N}\mathbb{E}\exp(\lambda X_i) \leqslant \exp(2\lambda^2 MN).$$

Τέλος, επιλέγοντας $\lambda = t/(4M)$ παίρνουμε

$$\begin{aligned}\mathbb{P}\Big(\Big\{\sum_{i=1}^{N} X_i > tN\Big\}\Big) &\leqslant \exp(-\lambda tN)\mathbb{E}\exp\Big(\lambda\sum_{i=1}^{N} X_i\Big)\\ &\leqslant \exp(-\lambda tN)\exp(2\lambda^2 MN)\\ &= \exp(2\lambda^2 MN - \lambda tN) = \exp(-t^2N/(8M)).\end{aligned}$$

Αντικαθιστώντας τις X_i με τις $-X_i$ στο προηγούμενο επιχείρημα, παίρνουμε τελικά το ζητούμενο. □

2.5 Ασκήσεις

1. Έστω $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$ η πυκνότητα μιας τυπικής τυχαίας μεταβλητής $X \sim N(0,1)$. Αποδείξτε ότι

$$\phi'(t) + t\phi(t) = 0$$

και χρησιμοποιώντας αυτή την ταυτότητα αποδείξτε ότι

$$\left(\frac{1}{t} - \frac{1}{t^3}\right)\phi(t) \leqslant \mathbb{P}(X \geqslant t) = \frac{1}{\sqrt{2\pi}}\int_t^\infty e^{-s^2/2}ds \leqslant \left(\frac{1}{t} - \frac{1}{t^3} + \frac{1}{t^5}\right)\phi(t)$$

για κάθε $t > 0$.

2. Έστω X μια μη αρνητική τυχαία μεταβλητή και έστω $t > 0$. Αποδείξτε ότι

$$\inf\left\{\frac{\mathbb{E}(X^k)}{t^k} : k = 0,1,2,\dots\right\} \leqslant \inf\left\{\frac{\mathbb{E}(e^{\lambda X})}{e^{\lambda t}} : \lambda > 0\right\}.$$

3. Έστω X τυχαία μεταβλητή με $\mathbb{E}(X) = 0$ και $\mathbb{P}(a \leqslant X \leqslant b) = 1$. Αν $\psi_X(\lambda) = \log(\mathbb{E}(e^{\lambda X}))$, αποδείξτε ότι, για κάθε $\lambda \in \mathbb{R}$,

$$\psi_X''(\lambda) \leqslant \frac{(b-a)^2}{4}$$

και συμπεράνατε ότι

$$\mathbb{E}(e^{\lambda X}) \leqslant \exp\left(\frac{\lambda^2(b-a)^2}{8}\right).$$

4. Έστω X τυχαία μεταβλητή με $\mathbb{E}(X) = 0$ και $\sigma^2 := \mathbb{E}(X^2)$. Υποθέτουμε ότι

$$\mathbb{E}(X^k) \leqslant \frac{1}{2}k!\sigma^2 b^{k-2}$$

για κάθε $k \geqslant 3$, όπου b είναι μια θετική σταθερά. Αποδείξτε ότι

$$\mathbb{E}(e^{\lambda X}) \leqslant e^{\frac{\lambda^2\sigma^2}{2(1-b|\lambda|)}}$$

για κάθε $|\lambda| < \frac{1}{b}$, και συμπεράνατε ότι

$$\mathbb{P}(|X| \geqslant t) \leqslant 2e^{-\frac{t^2}{2(\sigma^2+bt)}}$$

για κάθε $t > 0$.

5. Έστω $X_1, \dots, X_N$ ανεξάρτητες τυπικές κανονικές τυχαίες μεταβλητές ($X_i \sim N(0,1)$). Αποδείξτε ότι:

$$\mathbb{P}\left(\Big|\frac{1}{N}\sum_{i=1}^N X_i^2 - 1\Big| \geqslant t\right) \leqslant 2e^{-t^2N/8}$$

για κάθε $0 < t < 1$.

6. Μια συνάρτηση $\psi : [0,\infty) \to [0,\infty)$ λέγεται συνάρτηση Orlicz αν είναι κυρτή, αύξουσα με $\psi(0) = 0$ και $\lim_{x\to\infty} \psi(x) = +\infty$. Λέμε ότι μια πραγματική τυχαία μεταβλητή X στο χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$ ανήκει στην κλάση $L^{\psi}(\mu)$ αν υπάρχει $\varrho > 0$ τέτοιος ώστε $\mathbb{E}(\psi(|X|/\varrho)) < \infty$. Αποδείξτε ότι η κλάση $L^{\psi}(\mu)$ είναι γραμμικός χώρος και γίνεται χώρος με νόρμα αν θέσουμε

$$\|X\|_{\psi} = \inf\left\{\varrho > 0 : \mathbb{E}(\psi(|X|/\varrho)) \leqslant 1\right\}.$$

7. Θεωρούμε τη συνάρτηση Orlicz $\psi_1(t) = e^t - 1$. Έστω X τυχαία μεταβλητή στο χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$. Αποδείξτε ότι τα ακόλουθα είναι ισοδύναμα:

(α) Υπάρχει $\alpha > 0$ ώστε: $\|X\|_{\psi_1} \leqslant \alpha$.

(β) Υπάρχει $\beta > 0$ ώστε: για κάθε $p \geqslant 1$ ισχύει $\|X\|_p \leqslant \beta p$.

(γ) Υπάρχει $\gamma > 0$ ώστε: για κάθε $t > 0$ ισχύει $\mathbb{P}(|X| \geqslant t) \leqslant 2e^{-t/\gamma}$.

Επιπλέον, $\beta \leqslant c_1\alpha$, $\gamma \leqslant c_2\beta$, $\alpha \leqslant c_3\gamma$, όπου $c_i > 0$ είναι απόλυτες σταθερές.

8. Έστω X υποκανονική τυχαία μεταβλητή. Αποδείξτε ότι η $X - \mathbb{E}(X)$ είναι υποκανονική και

$$\|X - \mathbb{E}(X)\|_{\psi_2} \leqslant c_2\|X\|_{\psi_2},$$

όπου $c_2 > 0$ απόλυτη σταθερά.

Ομοίως, έστω X υποεκθετική τυχαία μεταβλητή. Αποδείξτε ότι η $X - \mathbb{E}(X)$ είναι υποεκθετική και

$$\|X - \mathbb{E}(X)\|_{\psi_1} \leqslant c_1\|X\|_{\psi_1},$$

όπου $c_1 > 0$ απόλυτη σταθερά.

Ποιές είναι οι καλύτερες τιμές που μπορείτε να βρείτε για τις σταθερές c_2 και c_1;

9. Έστω $Q_n = [-1,1]^n$ ο μοναδιαίος κύβος στον $\mathbb{R}^n$ και έστω $\varepsilon_1, \ldots, \varepsilon_n$ ανεξάρτητες συμμετρικές τυχαίες μεταβλητές Bernoulli. Για κάθε $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ και για κάθε $p \geqslant 1$ να συγκρίνετε τις ποσότητες

$$\left(\frac{1}{|Q_n|}\int_{Q_n}\Big|\sum_{i=1}^n a_ix_i\Big|^p dx\right)^{1/p}$$

και

$$\left(\mathbb{E}\Big|\sum_{i=1}^n a_i\varepsilon_i\Big|^p\right)^{1/p} = \left(\int_{E_2^n}\Big|\sum_{i=1}^n a_i\varepsilon_i\Big|^p d\mu_n(\varepsilon)\right)^{1/p}.$$

10. Έστω $p \geqslant 2$ και $\varepsilon_1, \ldots, \varepsilon_n$ ανεξάρτητες συμμετρικές τυχαίες μεταβλητές Bernoulli. Αποδείξτε ότι: για κάθε $a_1, \ldots, a_n \in \mathbb{R}$ ισχύει

$$\left(\mathbb{E}\Big|\sum_{i=1}^n a_i\varepsilon_i\Big|^p\right)^{1/p} \leqslant \sum_{i\leqslant p} a_i^* + c\sqrt{p}\left(\sum_{i>p}(a_i^*)^2\right)^{1/2},$$

όπου $(a_1^*, \ldots, a_n^*)$ είναι η φθίνουσα αναδιάταξη της n-άδας $(|a_1|, \ldots, |a_n|)$.

11. Έστω $\varepsilon_1, \ldots, \varepsilon_n$ ανεξάρτητες συμετρικές τυχαίες μεταβλητές Bernoulli. Για κάθε φραγμένο μη κενό $A \subseteq \mathbb{R}^n$ αποδείξτε ότι

$$\sup_{a\in A} \mathrm{Var}\left(\sum_{i=1}^n a_i\varepsilon_i\right) = \sup_{a\in A}\sum_{i=1}^n a_i^2$$

όπου $a = (a_1, \ldots, a_n)$, και

$$\mathrm{Var}\left(\sup_{a\in A}\sum_{i=1}^n a_i\varepsilon_i\right) \leqslant 4\sup_{a\in A}\sum_{i=1}^n a_i^2.$$

Κεφάλαιο 3

Η μέθοδος των martingales

3.1 Ανισότητα του Azuma

Δίνουμε πρώτα τους βασικούς ορισμούς της δεσμευμένης μέσης τιμής και του martingale σε ένα χώρο πιθανότητας $(\Omega, \mathcal{A}, P)$, και στη συνέχεια αποδεικνύουμε την ανισότητα του Azuma.

Ορισμός 3.1.1 (δεσμευμένη μέση τιμή)**.** Έστω $(\Omega, \mathcal{A}, P)$ ένας χώρος πιθανότητας. Αν $\mathcal{G}$ είναι μια υπο-σ-άλγεβρα της $\mathcal{A}$ και αν $f \in L_1(\Omega, \mathcal{A}, P)$, τότε η συνολοσυνάρτηση

$$\mu(A) = \int_A f dP, \quad A \in \mathcal{G}$$

ορίζει ένα μέτρο στην $\mathcal{G}$, το οποίο είναι απολύτως συνεχές ως προς το $P|_{\mathcal{G}}$. Από το θεώρημα Radon–Nikodym, υπάρχει μοναδική $h \in L_1(\Omega, \mathcal{G}, P)$ με την ιδιότητα

$$\int_A h dP = \int_A f dP$$

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

Παράδειγμα 3.1.2. Τυπικό παράδειγμα, το οποίο είναι βασικά το παράδειγμα που θα μας απασχολήσει, είναι αυτό όπου $\mathcal{G}$ είναι η άλγεβρα που παράγεται από μια διαμέριση $\{A_1, \ldots, A_m\}$ του Ω, με τα $A_i \in \mathcal{A}$ να έχουν θετικό μέτρο $P(A_i) > 0$. Τότε, εύκολα ελέγχουμε ότι αν $f \in L_1(\Omega, \mathcal{A}, P)$ η δεσμευμένη τιμή της f είναι η συνάρτηση h που είναι σταθερή σε κάθε A_i, με τιμή

$$h|_{A_i} = \frac{1}{P(A_i)} \int_{A_i} f\,dP.$$

Βασικές ιδιότητες της δεσμευμένης μέσης τιμής είναι οι εξής:

Λήμμα 3.1.3. (α) *Ο τελεστής $f \mapsto \mathbb{E}(f|\mathcal{G})$ είναι θετικός, γραμμικός και έχει νόρμα 1 σε κάθε L_p, $1 \leqslant p \leqslant \infty$.*

(β) *Αν $\mathcal{G}_1$ είναι μια υπο-σ-άλγεβρα της $\mathcal{G}$, τότε $\mathbb{E}(\mathbb{E}(f|\mathcal{G})|\mathcal{G}_1) = \mathbb{E}(f|\mathcal{G}_1)$.*

(γ) *Αν $g \in L_\infty(\Omega, \mathcal{G}, P)$ τότε $\mathbb{E}(f \cdot g|\mathcal{G}) = g \cdot \mathbb{E}(f|\mathcal{G})$.*

(δ) *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:*

$$\mathbb{E}(f|\mathcal{G}) = \mathbb{E}f = \int f dP.$$

Απόδειξη. (α) Η γραμμικότητα έπεται άμεσα από τον ορισμό της δεσμευμένης μέσης τιμής. Δείχνουμε ότι ο τελεστής $f \mapsto \mathbb{E}(f|\mathcal{G})$ είναι θετικός: Αν $f \in L_1(\Omega, \mathcal{A}, P)$ και $f \geqslant 0$, τότε υπάρχει $h \in L_1(\Omega, \mathcal{G}, P)$ ώστε

$$\int_A h\,dP = \int_A f\,dP \geqslant 0$$

για κάθε $A \in \mathcal{G}$. Αν θεωρήσουμε το $E_n = \{\omega : h(\omega) \leqslant -\frac{1}{n}\}$ έχουμε ότι $E_n \in \mathcal{G}$ και

$$0 \leqslant \int_{E_n} f\,dP = \int_{E_n} h\,dP \leqslant -\frac{1}{n} P(E_n),$$

απ' όπου έπεται ότι $P(E_n) = 0$. Άρα,

$$P(\omega : h(\omega) < 0) = P\left(\bigcup_{n=1}^{\infty} E_n\right) = 0.$$

Από το γεγονός ότι ο τελεστής $T(f) = \mathbb{E}(f|\mathcal{G})$ είναι θετικός και γραμμικός έπεται ότι είναι μονότονος: αν $f, g \in L_1(\Omega, \mathcal{A}, P)$ με $f \leqslant g$, τότε $\mathbb{E}(f|\mathcal{G}) \leqslant \mathbb{E}(g|\mathcal{G})$. Ειδικότερα, έπεται ότι

$$|\mathbb{E}(f|\mathcal{G})| \leqslant \mathbb{E}(|f|\;|\mathcal{G})$$

για κάθε $f \in L_1(\Omega, \mathcal{A}, P)$. Τότε, η δεσμευμένη μέση τιμή $T : L_1 \to L_1$ είναι φραγμένος τελεστής νόρμας 1. Πράγματι·

$$\|\mathbb{E}(f|\mathcal{G})\|_1 = \int |\mathbb{E}(F|\mathcal{G})|\,dP \leqslant \int \mathbb{E}(|f|\;|\mathcal{G})\,dP = \int |f|\,dP = \|f\|_1,$$

όπου στην προτελευταία ισότητα έχουμε χρησιμοποιήσει τον ορισμό της δεσμευμένης μέσης τιμής. Επίσης, είναι $\mathbb{E}(\mathbf{1}|\mathcal{G}) = \mathbf{1}$. Από την ανισότητα Hölder έπεται ότι $L_p \subseteq L_1$ για κάθε $1 \leqslant p \leqslant \infty$. Επομένως, αν $f \in L_\infty$, τότε

$$|\mathbb{E}(f|\mathcal{G})| \leqslant \mathbb{E}(|f||\mathcal{G}) \leqslant \mathbb{E}(\|f\|_\infty|\mathcal{G}) = \|f\|_\infty.$$

Συνεπώς, για κάθε $f \in L_\infty$ έπεται ότι $\mathbb{E}(f|\mathcal{G}) \in L_\infty$ και μάλιστα $\|\mathbb{E}(f|\mathcal{G})\|_\infty \leqslant \|f\|_\infty$. Με άλλα λόγια, η δεσμευμένη μέση τιμή $T : L_\infty \to L_\infty$ είναι καλά ορισμένος τελεστής νόρμας 1. Τέλος, αυτό που μένει να δείξουμε είναι ότι ο $T : L_p \to L_p$ είναι επίσης καλά ορισμένος. Αυτό έπεται από τον ακόλουθο ισχυρισμό:

Ισχυρισμός. Έστω $f \in L_1$ και $\varphi : \mathbb{R} \to \mathbb{R}$ κυρτή ώστε $\mathbb{E}|\varphi(f)| < \infty$. Τότε ισχύει

$$\varphi(\mathbb{E}(f|\mathcal{G})) \leqslant \mathbb{E}(\varphi(f)|\mathcal{G}).$$

Απόδειξη του ισχυρισμού. Είναι γνωστό ότι υπάρχουν ακολουθίες πραγματικών αριθμών (a_n), (b_n) ώστε $\varphi(x) = \sup_n(a_n x + b_n)$ για κάθε $x \in \mathbb{R}$. Τότε, για κάθε $n \in \mathbb{N}$ ισχύει

$$a_n f(x) + b_n \leqslant \varphi(f(x))$$

σχεδόν παντού. Έπεται ότι για κάθε $n \in \mathbb{N}$ υπάρχει $E_n \in \mathcal{G}$ με $P(E_n) = 0$ και

$$a_n \mathbb{E}(f|\mathcal{G}) + b_n \leqslant \mathbb{E}(\varphi(f)|\mathcal{G})$$

για κάθε $x \in \Omega \setminus E_n$. Αν θέσουμε $E = \bigcup_{n=1}^\infty E_n$, τότε $P(E) = 0$ και αν $x \in \Omega \setminus E$ είναι

$$a_n \mathbb{E}(f|\mathcal{G}) + b_n \leqslant \mathbb{E}(\varphi(f)|\mathcal{G})$$

για κάθε $n \in \mathbb{N}$. Συνεπώς, παίρνοντας supremum ως προς n έχουμε ότι

$$\varphi(\mathbb{E}(f|\mathcal{G})) \leqslant \mathbb{E}(\varphi(f)|\mathcal{G})$$

σχεδόν για κάθε $x \in \Omega$. Εφαρμόζοντας αυτήν την ανισότητα για την $\varphi(t) = |t|^p$ έχουμε ότι

$$|\mathbb{E}(f|\mathcal{G})|^p \leqslant \mathbb{E}(|f|^p|\mathcal{G})$$

και ολοκληρώνοντας βλέπουμε ότι $\|\mathbb{E}(f|\mathcal{G})\|_p \leqslant \|f\|_p$ για κάθε $f \in L_p$.

(β) Έστω $g = \mathbb{E}(f|\mathcal{G})$. Τότε για κάθε $A \in \mathcal{G}$ ισχύει $\int_A f\,dP = \int_A g\,dP$. Αν $B \in \mathcal{G}_1 \subseteq \mathcal{G}$ τότε έχουμε $\int_B g\,dP = \int_B f\,dP$. Από τον ορισμό έπεται ότι $\mathbb{E}(g|\mathcal{G}_1) = \mathbb{E}(f|\mathcal{G}_1)$.

(γ) Αρκεί να το δείξουμε για χαρακτηριστικές συναρτήσεις που είναι $\mathcal{G}$-μετρήσιμες. Για $g = \chi_A$ και $A, B \in \mathcal{G}$ έχουμε

$$\int_B \mathbb{E}(fg|\mathcal{G})\, dP = \int_B fg\, dP = \int_{A\cap B} f\, dP = \int_{A\cap B} \mathbb{E}(f|\mathcal{G})\, dP = \int_B g\mathbb{E}(f|\mathcal{G})\, dP.$$

Έτσι, $\mathbb{E}(fg|\mathcal{G}) = g\mathbb{E}(f|\mathcal{G})$, διότι $\chi_A\mathbb{E}(f|\mathcal{G}) \in L_1(\Omega, \mathcal{G}, P)$.

(δ) Άμεσο από τον ορισμό και το παράδειγμα μετά από τον ορισμό. □

Ορισμός 3.1.4 (martingale)**.** Έστω $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}$ μια ακολουθία σ-αλγεβρών. Μια ακολουθία $f_0, f_1, \dots$ συναρτήσεων $f_i \in L_1(\Omega, \mathcal{F}_i, P)$ λέγεται martingale ως προς την $\{\mathcal{F}_i\}$ αν $\mathbb{E}(f_i|\mathcal{F}_{i-1}) = f_{i-1}$ για κάθε $i \geqslant 1$.

Η ανισότητα του Azuma δίνει εκτίμηση της πιθανότητας απόκλισης μιας φραγμένης συνάρτησης από τη μέση τιμή της.

Θεώρημα 3.1.5. *Έστω $f \in L_\infty(\Omega, \mathcal{F}, P)$ και έστω $\{\emptyset, \Omega\} = \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}_n = \mathcal{F}$ μια ακολουθία σ-αλγεβρών. Θέτουμε $d_i = \mathbb{E}(f|\mathcal{F}_i) - \mathbb{E}(f|\mathcal{F}_{i-1})$, $i = 1, \dots, n$. Τότε, για κάθε $t > 0$ έχουμε*

$$P(|f - \mathbb{E}f| \geqslant t) \leqslant 2\exp\left(-t^2/4\sum_{i=1}^n \|d_i\|_\infty^2\right).$$

Απόδειξη. Παρατηρούμε πρώτα ότι η ακολουθία $\{\mathbb{E}(f|\mathcal{F}_i)\}_{i=0}^n$ είναι martingale ως προς $\{\mathcal{F}_i\}_{i=0}^n$. Πράγματι, έχουμε

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

και $\mathbb{E}(f|\mathcal{F}_i) \in L_1(\Omega, \mathcal{F}_i, P)$ από τον ορισμό της δεσμευμένης μέσης τιμής. Επιπλέον, έχουμε $\mathbb{E}(d_i|\mathcal{F}_{i-1}) = 0$ για κάθε $i \geqslant 1$:

$$\begin{aligned} \mathbb{E}(d_i|\mathcal{F}_{i-1}) &= \mathbb{E}\big(\mathbb{E}(f|\mathcal{F}_i) - \mathbb{E}(f|\mathcal{F}_{i-1})|\mathcal{F}_{i-1}\big) \\ &= \mathbb{E}(\mathbb{E}(f|\mathcal{F}_i)|\mathcal{F}_{i-1}) - \mathbb{E}\big(\mathbb{E}(f|\mathcal{F}_{i-1})|\mathcal{F}_{i-1}\big) \\ &= \mathbb{E}(f|\mathcal{F}_{i-1}) - \mathbb{E}(f|\mathcal{F}_{i-1}) = 0. \end{aligned} \tag{3.1.1}$$

Συγκρίνοντας τις δυναμοσειρές των e^x και $e^{x^2/2}$ βλέπουμε ότι $e^x \leqslant x + e^{x^2}$ για κάθε $x \in \mathbb{R}$. Αφού ο τελεστής $f \mapsto \mathbb{E}(f|\mathcal{F})$ είναι θετικός, συμπεραίνουμε οτι για κάθε $\lambda \in \mathbb{R}$

$$\mathbb{E}(e^{\lambda d_k}|\mathcal{F}_{k-1}) \leqslant \mathbb{E}(\lambda d_k|\mathcal{F}_{k-1}) + \mathbb{E}(e^{\lambda^2 d_k^2} \mid \mathcal{F}_{k-1}).$$

Από τη γραμμικότητα της $f \mapsto \mathbb{E}(f|\mathcal{F}_{k-1})$ έχουμε $\mathbb{E}(\lambda d_k|\mathcal{F}_{k-1}) = \lambda\mathbb{E}(d_k|\mathcal{F}_{k-1}) = 0$. Επίσης, έχουμε υποθέσει ότι $f \in L_\infty(\Omega, \mathcal{F}, P)$, άρα κάθε $d_k \in L_\infty(\Omega, \mathcal{F}_k, P)$. Συνεπώς,

$$\mathbb{E}(e^{\lambda d_k}|\mathcal{F}_{k-1}) \leqslant \mathbb{E}(e^{\lambda^2 d_k^2} \mid \mathcal{F}_{k-1}) \leqslant \mathbb{E}(e^{\lambda^2 \|d_k\|_\infty^2}|\mathcal{F}_{k-1}) = e^{\lambda^2\|d_k\|_\infty^2}.$$

Ισχυρισμός. Ισχύει η ανισότητα

$$\mathbb{E}\left(e^{\sum_{i=1}^n \lambda d_i}\right) \leqslant e^{\lambda^2 \sum_{i=1}^n \|d_i\|_\infty^2}.$$

Απόδειξη. Με επαγωγή δείχνουμε ότι $\mathbb{E}\left(e^{\lambda\sum_{j=1}^k d_j}\right) \leqslant e^{\sum_{j=1}^k \lambda^2\|d_j\|^2}$ για κάθε $k \leqslant n$: Για $k = 1$ έχουμε $\mathbb{E}\left(e^{\lambda d_1}\right) = \mathbb{E}\left[\mathbb{E}\left(e^{\lambda d_1}|F_0\right)\right] \leqslant e^{\lambda^2\|d_1\|_\infty^2}$. Υποθέτουμε ότι

$$\mathbb{E}\left(e^{\lambda\sum_{j=1}^k d_j}\right) \leqslant e^{\sum_{j=1}^k \lambda^2\|d_j\|^2}$$

για κάποιον $k < n$. Αφού $e^{\sum_{j=1}^k \lambda d_j} \in L_\infty(\Omega, \mathcal{F}_k, P)$, από το Λήμμα 3.1.3 (γ) παίρνουμε

$$\mathbb{E}(e^{\sum_{j=1}^k \lambda d_j} e^{\lambda d_{k+1}}|\mathcal{F}_k) = e^{\sum_{j=1}^k \lambda d_j}\mathbb{E}(e^{\lambda d_{k+1}}|\mathcal{F}_k).$$

Χρησιμοποιώντας και την επαγωγική υπόθεση έχουμε:

$$\begin{aligned}
\mathbb{E}\left(e^{\lambda\sum_{j=1}^{k+1} d_j}\right) &= \mathbb{E}\left[\mathbb{E}\left(e^{\sum_{j=1}^k \lambda d_j} e^{\lambda d_{k+1}} \mid \mathcal{F}_k\right)\right] \\
&= \mathbb{E}\left[e^{\sum_{j=1}^k \lambda d_j}\mathbb{E}\left(e^{\lambda d_{k+1}} \mid \mathcal{F}_k\right)\right] \\
&\leqslant \mathbb{E}\left[e^{\sum_{j=1}^k \lambda d_j} e^{\lambda^2\|d_{k+1}\|_\infty^2}\right] \\
&= e^{\lambda^2\|d_{k+1}\|_\infty^2}\mathbb{E}\left(e^{\sum_{j=1}^k \lambda d_j}\right) \\
&\leqslant e^{\lambda^2\|d_{k+1}\|_\infty^2} e^{\lambda^2\sum_{j=1}^k \|d_i\|_\infty^2} \\
&= e^{\lambda^2\sum_{j=1}^{k+1}\|d_j\|_\infty^2}.
\end{aligned} \tag{3.1.2}$$

Για κάθε $\lambda > 0$ έχουμε

$$\begin{aligned}
P(f - \mathbb{E}f \geqslant t) = P(\mathbb{E}(f|\mathcal{F}_n) - \mathbb{E}(f|\mathcal{F}_0) \geqslant t) &= P\left(\sum_{j=1}^n d_j \geqslant t\right) \\
&\leqslant \mathbb{E}\, e^{\lambda\sum_{j=1}^n d_j - \lambda t} \leqslant e^{\lambda^2\sum_{j=1}^n \|d_j\|_\infty^2 - \lambda t}.
\end{aligned} \tag{3.1.3}$$

Ελαχιστοποιώντας ως προς λ βλέπουμε ότι

$$P(f - \mathbb{E}f \geqslant t) \leqslant \exp\left(-t^2/4\sum_{j=1}^n \|d_j\|_\infty^2\right).$$

Εφαρμόζοντας το ίδιο επιχείρημα για την $-f$, παίρνουμε

$$P(-f + \mathbb{E}f \geqslant t) \leqslant \exp\left(-t^2/4\sum_{j=1}^{n} \|d_j\|_\infty^2\right).$$

Συνδυάζοντας τις δύο ανισότητες παίρνουμε το ζητούμενο. □.

3.2 Εφαρμογές στο διακριτό κύβο

Θα χρησιμοποιήσουμε πρώτα την ανισότητα του Azuma για να δώσουμε μια απλή απόδειξη της ανισότητας του Khintchine.

Απόδειξη της (2.3.1) **με martingales**. Έστω $a_1, \ldots, a_n \in \mathbb{R}$ με $\sum_{i=1}^{n} a_i^2 = 1$ (μπορούμε να κάνουμε αυτήν την υπόθεση γιατί η ανισότητα Khintchine είναι ομογενής). Για κάθε $j \geqslant 1$ θεωρούμε την άλγεβρα $\mathcal{F}_j$ υποσυνόλων του E_2^n που παράγεται απο τα σύνολα

$$A_{\epsilon_1,\ldots,\epsilon_j} = \{\zeta : \zeta_1 = \epsilon_1, \ldots, \zeta_j = \epsilon_j\}$$

όπου $\epsilon_1, \ldots, \epsilon_j \in \{-1, 1\}$. Παρατηρήστε ότι το πλήθος των ατόμων της $\mathcal{F}_j$ ισούται με 2^j. Θεωρούμε την ακολουθία

$$\{\emptyset, E_2^n\} = \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_n = \mathcal{P}(E_2^n).$$

Τότε, $\mathcal{F}_j \subseteq \mathcal{F}_{j+1}$ για κάθε $j = 0, 1, \ldots, n-1$. Πράγματι, κάθε άτομο της $\mathcal{F}_j$ γράφεται στη μορφή

$$A_{\epsilon_1,\ldots,\epsilon_j} = A_{\epsilon_1,\ldots,\epsilon_j,1} \cup A_{\epsilon_1,\ldots,\epsilon_j,-1},$$

δηλαδή ανήκει στην $\mathcal{F}_{j+1}$. Για κάθε $k \leqslant n$ οι τυχαίες μεταβλητές $\epsilon_1, \ldots, \epsilon_k$ είναι μετρήσιμες ως προς την F_k και η $\left\{\sum_{i=1}^{k} a_i\epsilon_i\right\}_{k=1}^{n}$ είναι martingale ως προς την $\{\mathcal{F}_k\}_{k=1}^{n}$. Πράγματι,

$$\mathbb{E}\left(\sum_{i=1}^{k} a_i\epsilon_i|\mathcal{F}_{k-1}\right) = \sum_{i=1}^{k} a_i\mathbb{E}(\epsilon_i|\mathcal{F}_{k-1}) = \sum_{i=1}^{k-1} a_i\mathbb{E}(\epsilon_i|\mathcal{F}_{k-1}) + a_k\mathbb{E}(\epsilon_k|\mathcal{F}_{k-1}). \tag{3.2.1}$$

Οι ϵ_i, $i = 1, \ldots, k-1$ είναι μετρήσιμες ως προς την $\mathcal{F}_{k-1}$, άρα

$$\mathbb{E}(\epsilon_i|\mathcal{F}_{k-1}) = \epsilon_i, \qquad i = 1, \ldots, k-1. \tag{3.2.2}$$

Επίσης, $\mathbb{E}(\epsilon_k|\mathcal{F}_{k-1}) = 0$. Για το σκοπό αυτό αρκεί να δείξουμε οτι $\int_A \mathbb{E}(\epsilon_k|\mathcal{F}_{k-1})d\mu_n = 0$ για κάθε άτομο της $\mathcal{F}_{k-1}$. Όμως, κάθε άτομο A της $\mathcal{F}_{k-1}$ γράφεται στην μορφή $A = B_1 \cup B_2$, όπου τα B_1, B_2 είναι άτομα της $\mathcal{F}_k$, και

$$\int_A \mathbb{E}(\epsilon_k|\mathcal{F}_{k-1})d\mu_n = \int_A \epsilon_k d\mu_n = \int_{B_1} \epsilon_k d\mu_n + \int_{B_2} \epsilon_k d\mu_n = 0, \tag{3.2.3}$$

αφού σε ένα από τα B_1, B_2 η ϵ_k παίρνει την τιμή 1 και στο άλλο την τιμή -1.

Θέτουμε $f = \sum_{i=1}^n a_i\epsilon_i$. Τότε,

$$\mathbb{E}(f|\mathcal{F}_k) = \sum_{i=1}^{k} a_i\epsilon_i, \tag{3.2.4}$$

και αν θέσουμε $d_k = \mathbb{E}(f|\mathcal{F}_k) - \mathbb{E}(f|\mathcal{F}_{k-1})$ συμπεραίνουμε ότι

$$\|d_k\|_\infty = \left\| \sum_{i=1}^{k} a_i\epsilon_i - \sum_{i=1}^{k-1} a_i\epsilon_i \right\|_\infty = \|a_k\epsilon_k\|_\infty = |a_k|. \tag{3.2.5}$$

Παρατηρήστε επίσης ότι $f \in L_\infty(\mu_n)$ και $\mathbb{E}f = 0$. Από την ανισότητα του Azuma έπεται ότι

$$\mu_n\left(\left|\sum_{i=1}^{n} a_i\epsilon_i\right| \geqslant t\right) \leqslant 2e^{\frac{-t^2}{4\sum_{i=1}^n \|d_i\|_\infty^2}} = 2e^{\frac{-t^2}{4|a|^2}} = 2e^{\frac{-t^2}{4}} \tag{3.2.6}$$

για κάθε $t > 0$.

Μπορούμε λοιπόν, ως συνήθως, να γράψουμε

$$\begin{aligned}\int_{E_2^n} |\sum a_i\epsilon_i|^p d\mu_n &= \int_0^\infty pt^{p-1}\mu_n\left(\left|\sum_{i=1}^{n} a_i\epsilon_i\right| \geqslant t\right) dt \\ &\leqslant \int_0^\infty pt^{p-1}e^{\frac{-t^2}{4}}dt \\ &= 2^p p\int_0^\infty e^{-x}x^{p/2-1}dx \leqslant (C\sqrt{p})^p.\end{aligned}$$

Έπεται ότι η δεξιά ανισότητα της ανισότητας του Khintchine ισχύει με $B_p = O(\sqrt{p})$ για $p \geqslant 2$. □

Το επόμενο θεώρημα δίνει μια ανισότητα απόκλισης για Lipschitz συναρτήσεις στο διακριτό κύβο.

Θεώρημα 3.2.1. *Έστω $f : E_2^n \to \mathbb{R}$ συνάρτηση Lipschitz με σταθερά 1. Τότε,*

$$\mu_n(|f - \mathbb{E}f| \geqslant t) \leqslant 2e^{-t^2n/4}$$

για κάθε $t > 0$.

Απόδειξη. Έστω $\mathcal{F}_j$ η άλγεβρα που παράγεται απο τα σύνολα

$$A_{\varepsilon_1,\dots,\varepsilon_j} = \{\zeta : \zeta_1 = \varepsilon_1, \dots, \zeta_j = \varepsilon_j\}$$

όπου $\varepsilon_1, \dots, \varepsilon_j \in \{-1, 1\}$. Όπως στην προηγούμενη απόδειξη, θεωρούμε την ακολουθία

$$\{\emptyset, E_2^n\} = \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_n = 2^{E_2^n}.$$

Τότε, $\mathcal{F}_j \subseteq \mathcal{F}_{j+1}$ για κάθε $j = 0, 1, \dots, n-1$.

Έστω $f : E_2^n \to \mathbb{R}$ συνάρτηση Lipschitz με σταθερά 1 και έστω $(f_j)_{j=0}^n$ το martingale $f_j = \mathbb{E}(f|\mathcal{F}_j)$ που επάγεται απο την f.

Η βασική παρατήρηση είναι ότι για κάθε άτομο $A = A_{\varepsilon_1, \dots, \varepsilon_j}$ της $\mathcal{F}_j$, υπάρχει προφανής απεικόνιση $\phi : B \to C$ ώστε $d(\zeta, \phi(\zeta)) \leqslant \frac{2}{n}$ για κάθε $\zeta \in B$, όπου $B = A_{\varepsilon_1, \dots, \varepsilon_j, 1}$ και $C = A_{\varepsilon_1, \dots, \varepsilon_j, -1}$. Θέτουμε $\phi(\zeta) = \zeta'$, αλλάζοντας την $(j+1)$-στή συντεταγμένη του ζ από 1 σε -1.

Έτσι βλέπουμε ότι

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

Συνεπώς,

$$|f_{j+1}|_B - f_{j+1}|_C| \leqslant \frac{1}{n}$$

και από την

$$f_j|_A = \frac{1}{2} f_{j+1}|_B + \frac{1}{2} f_{j+1}|_C,$$

παίρνουμε

$$|f_j|_A - f_{j+1}|_B| \leqslant \frac{1}{n} \quad \text{και} \quad |f_j|_A - f_{j+1}|_C| \leqslant \frac{1}{n}.$$

Αν λοιπόν $d_j = f_j - f_{j-1}$, έχουμε $\|d_j\|_\infty \leqslant \frac{1}{n}$ για κάθε $j = 1, \dots, n$. Από την ανισότητα του Azuma συμπεραίνουμε ότι

$$\mu_n(|f - \mathbb{E}f| \geqslant t) \leqslant 2e^{-t^2 n/4}$$

για κάθε $t > 0$. □

3.3 Συγκέντρωση του μέτρου στην S_n

Θεωρούμε την οικογένεια S_n των μεταθέσεων του συνόλου $[n] := \{1, 2, \dots, n\}$ με μετρική την $d(\sigma, \tau) = \frac{1}{n}|\{i : \sigma(i) \neq \tau(i)\}|$ και με το ομοιόμορφο μέτρο P που δίνει μάζα $\frac{1}{n!}$ σε κάθε μετάθεση.

Θεώρημα 3.3.1 (Maurey). *Έστω $f : S_n \to \mathbb{R}$ συνάρτηση Lipschitz με σταθερά 1. Τότε,*

$$P(|f - \mathbb{E}f| \geqslant t) \leqslant 2e^{-t^2 n/16}$$

για κάθε $t > 0$.

Απόδειξη. Έστω $\mathcal{F}_j$ η άλγεβρα που παράγεται απο τα σύνολα

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

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

$$\{\emptyset, S_n\} = \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_{n-1} = \mathcal{P}(S_n).$$

Τότε, $\mathcal{F}_j \subseteq \mathcal{F}_{j+1}$ για κάθε $j = 0, 1, \dots, n-2$. Πράγματι, κάθε άτομο της $\mathcal{F}_j$ γράφεται στην μορφή

$$A_{i_1,i_2,\dots,i_j} = \bigcup_{k\in[n]\setminus\{i_1\dots,i_j\}} A_{i_1,i_2,\dots,i_j,k},$$

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

Έστω $f : S_n \to \mathbb{R}$ συνάρτηση Lipschitz με σταθερά 1 και έστω $(f_j)_{j=0}^n$ το martingale $f_j = \mathbb{E}(f|\mathcal{F}_j)$ που επάγεται απο την f.

Λήμμα 3.3.2. *Για κάθε άτομο $A = A_{i_1,i_2,\dots,i_j}$ της $\mathcal{F}_j$ και κάθε ζευγάρι ατόμων $B = A_{i_1,i_2,\dots,i_j,r}$ και $C = A_{i_1,\dots,i_j,s}$ της $\mathcal{F}_{j+1}$ που περιέχονται στο A, μπορούμε να βρούμε μια $1-1$ και επί απεικόνιση $\phi : B \to C$ ώστε $d(b, \phi(b)) \leqslant \frac{2}{n}$ για κάθε $b \in B$.*

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

Τότε, $\phi(\sigma)(i) = \sigma(i)$ για $i \neq j+1$ και $i \neq \sigma^{-1}(s)$. Αν $i = j+1$ τότε $\phi(\sigma)(j+1) = \pi \circ \sigma(j+1) = \pi(r) = s$ και αν $i = \sigma^{-1}(s)$ τότε $\phi(\sigma)(i) = \pi(s) = r$. Άρα,

$$d(\sigma, \phi(\sigma)) \leqslant \frac{2}{n}.$$

Η ϕ είναι εξ ορισμού 1-1 και αφού $|B| = |C|$ έπεται ότι η ϕ είναι επί. □

Σταθεροποιούμε A, B, C όπως στο Λήμμα 3.3.2. Αφού τα B, C είναι άτομα της $\mathcal{F}_{j+1}$, η f_{j+1} είναι σταθερή στα B, C. Έχουμε

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

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

$$f_{j+1}|B \equiv \frac{1}{P(B)}\frac{1}{n!}\sum_{\sigma\in B} f(\sigma) = \frac{1}{|B|}\sum_{\sigma\in B} f(\sigma).$$

Όμοια δείχνουμε οτι $f_{j+1}|C = \frac{1}{|C|}\sum_{\sigma\in C} f(\sigma)$. Γράφουμε

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

όπου ϕ η συνάρτηση του Λήμματος 3.3.2. Αφού η f είναι συνάρτηση Lipschitz με σταθερά 1,

$$\begin{aligned}|f_{j+1}|B - f_{j+1}|C| &\leqslant \frac{1}{|B|}\sum_{\sigma\in B}|f(\sigma) - f(\phi(\sigma))| \leqslant \frac{1}{|B|}\sum_{\sigma\in B} d(\sigma,\phi(\sigma)) \\ &\leqslant \frac{1}{|B|}\sum_{\sigma\in B}\frac{2}{n} = \frac{2}{n}.\end{aligned} \tag{3.3.1}$$

Έπεται ότι $|f_{j+1}|B - f_j|A| \leqslant \frac{2}{n}$. Πράγματι, έχουμε

$$\begin{aligned}f_j|A &= \frac{1}{|A|}\sum_{\sigma\in A} f(\sigma) = \sum_{s\notin\{i_1,\ldots,i_j\}}\frac{1}{|A|}\sum_{\sigma\in A_{i_1,\ldots,i_j,s}} f(\sigma) \\ &= \sum_{C\subseteq A}\frac{1}{(n-j)|C|}\sum_{\sigma\in C} f(\sigma) = \frac{1}{n-j}\sum_{C\subseteq A} f_{j+1}|C.\end{aligned} \tag{3.3.2}$$

Συνεπώς,

(3.3.3)

$$\begin{aligned}|f_{j+1}|B - f_j|A| &= \left|f_{j+1}|B - \frac{1}{n-j}\sum_{C\subseteq A} f_{j+1}|C\right| = \left|\sum_{C\subseteq A}\frac{1}{n-j}(f_{j+1}|B - f_{j+1}|C)\right| \\ &\leqslant \sum_{C\subseteq A}\frac{1}{n-j}|f_{j+1}|B - f_{j+1}|C| \leqslant \sum_{C\subseteq A}\frac{1}{n-j}\frac{2}{n} = \frac{2}{n}.\end{aligned}$$

Έχουμε $\Omega = \bigcup B_i$ όπου $B_i \in \mathcal{F}_{j+1}$. Θα δείξουμε οτι $|d_{j+1}|B_i| \leqslant \frac{2}{n}$, όπου οι d_j ορίζονται όπως στην ανισότητα του Azuma. Πράγματι,

$$|d_{j+1}|B_i| = |f_{j+1}|B_i - f_j|B_i| = |f_{j+1}|B_i - f_j|A_i| \leqslant \frac{2}{n}$$

όπου A_i το άτομο της F_j που περιέχει το B_i. Άρα,

$$\|d_{j+1}\|_\infty \leqslant \frac{2}{n}.$$

Η f προφανώς ανήκει στον $L_\infty(S_n, \mathcal{F}_n, P)$, άρα η προηγούμενη ανισότητα και η ανισότητα του Azuma δίνουν

$$P(|f - \mathbb{E}f| \geqslant t) \leqslant 2e^{-t^2 n/16}$$

για κάθε $t > 0$. □

Από το Θεώρημα 3.3.1 προκύπτει ότι ο μετρικός χώρος πιθανότητας (S_n, d, P) έχει κανονική συνάρτηση συγκέντρωσης. Αυτό προκύπτει από την επόμενη γενική πρόταση, την οποία θα χρειαστούμε και στη συνέχεια.

Πρόταση 3.3.3. *Έστω (X, d, μ) ένας μετρικός χώρος πιθανότητας. Υποθέτουμε ότι για κάποια συνάρτηση $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$, για κάθε φραγμένη 1-Lipschitz συνάρτηση $f : (X, d) \to \mathbb{R}$ και για κάθε $t > 0$ ισχύει*

$$\mu\left(\{f \geqslant \mathbb{E}_\mu(f) + t\}\right) \leqslant \alpha(t).$$

Τότε, για κάθε Borel σύνολο $A \subseteq X$ με $\mu(A) > 0$ και για κάθε $t > 0$ ισχύει

$$1 - \mu(A_t) \leqslant \alpha(\mu(A)t).$$

Ειδικότερα,

$$\alpha_\mu(t) \leqslant \alpha(t/2), \qquad t > 0.$$

Απόδειξη. Σταθεροποιούμε $A \in \mathcal{B}(X)$ με $\mu(A) > 0$ και $t > 0$. Θεωρούμε τη συνάρτηση $f(x) = \min\{d(x, A), t\}$. Παρατηρήστε ότι $\|f\|_{\mathrm{Lip}} \leqslant 1$ και

$$\int f\, d\mu \leqslant (1 - \mu(A))t.$$

Από την υπόθεση έχουμε

$$\begin{aligned} 1 - \mu(A_t) = \mu(\{f \geqslant t\}) &\leqslant \mu(\{f \geqslant \int f\, d\mu + \mu(A)t\}) \\ &\leqslant \alpha(\mu(A)t). \end{aligned} \tag{3.3.4}$$

Ειδικότερα, αν $\mu(A) \geqslant 1/2$ τότε από την (3.3) έχουμε $\int f\, d\mu \leqslant t/2$ και, επαναλαμβάνοντας το προηγούμενο επιχείρημα, βλέπουμε ότι

$$\begin{aligned} 1 - \mu(A_t) = \mu(\{f \geqslant t\}) &\leqslant \mu\left(\left\{f \geqslant \int f\, d\mu + t/2\right\}\right) \\ &\leqslant \alpha(t/2), \end{aligned} \tag{3.3.5}$$

απ' όπου έπεται ότι $\alpha_\mu(t) \leqslant \alpha(t/2)$. □

Θεώρημα 3.3.4 (Maurey). *Η συνάρτηση συγκέντρωσης του* (S_n, d, P) *ικανοποιεί την*

$$\alpha_P(t) \leqslant 2e^{-t^2 n/64}$$

για κάθε $t > 0$.

Απόδειξη. Από το Θεώρημα 3.3.1 βλέπουμε ότι για κάθε φραγμένη 1-Lipschitz συνάρτηση $f : (X, d) \to \mathbb{R}$ και για κάθε $t > 0$ ισχύει

$$P\left(\left\{f \geqslant \int f d\mu + t\right\}\right) \leqslant P\left(\left\{|f - \int f d\mu| \geqslant t\right\}\right) \leqslant e^{-t^2 n/16}.$$

Το συμπέρασμα έπεται από την Πρόταση 3.3.3. □

Δίνουμε ακόμα μία εφαρμογή της μεθόδου στο εξής πλαίσιο. Υποθέτουμε ότι για κάθε $i = 1, \ldots, n$, ο $(X_i, \mathcal{A}_i, \mu_i)$ είναι χώρος πιθανότητας και θέτουμε $X = X_1 \times \cdots \times X_n$. Στον X θεωρούμε το μέτρο γινόμενο $\mu = \mu_1 \otimes \cdots \otimes \mu_n$: τα μετρήσιμα υποσύνολα του X είναι αριθμήσιμες ενώσεις συνόλων της μορφής $A = A_1 \times \cdots \times A_n$, όπου $A_i \in \mathcal{A}_i$, και $\mu(A) = \mu_1(A_1) \cdots \mu_n(A_n)$.

Στον X θεωρούμε την εξής μετρική: αν $x = (x_1, \ldots, x_n)$ και $y = (y_1, \ldots, y_n) \in X$, ορίζουμε

$$d(x, y) = |\{i \leqslant n : x_i \neq y_i\}|.$$

Θεώρημα 3.3.5. *Έστω* $f : X \to \mathbb{R}$ *ολοκληρώσιμη συνάρτηση και έστω* $a_i > 0$ *ώστε* $|f(x) - f(y)| \leqslant a_i$ *αν τα* x *και* y *διαφέρουν μόνο στην* i*-οστή συντεταγμένη. Τότε,*

$$\mu(|f - \mathbb{E}f| \geqslant t) \leqslant 2e^{-t^2/4A}$$

για κάθε $t > 0$*, όπου* $A = a_1^2 + \cdots + a_n^2$.

Απόδειξη. Για κάθε $j = 0, 1, \ldots, n$ θεωρούμε τη σ-άλγεβρα $\mathcal{F}_j$ που αποτελείται από όλες τις αριθμήσιμες ενώσεις συνόλων της μορφής $A_1 \times \cdots \times A_j \times X_{j+1} \times \cdots \times X_n$, όπου $A_i \in \mathcal{A}_i$, $i = 1, \ldots, j$. Θέτουμε $f_j = \mathbb{E}(f \mid \mathcal{F}_j)$. Παρατηρήστε ότι $f_0 = \mathbb{E}f$, $f_n = f$ και, για κάθε $1 \leqslant j \leqslant n - 1$,

$$f_j(x_1, \ldots, x_n) = \int_{X_{j+1} \times \cdots \times X_n} f(x_1, \ldots, x_j, y_{j+1}, \ldots, y_n)\, d\mu_{j+1}(y_{j+1}) \cdots d\mu_n(y_n).$$

Από την υπόθεση για την f ελέγχουμε εύκολα ότι αν $d_j = f_j - f_{j-1}$ τότε $\|d_j\|_\infty \leqslant a_j$.

Πράγματι, για κάθε $x \in X$ έχουμε

$$
\begin{aligned}
(3.3.6) \qquad & |f_j(x) - f_{j-1}(x)| \\
= \quad & \left| \int_X f(x_1, \ldots, x_j, y_{j+1}, \ldots, y_n) d\mu(y) - \int_X f(x_1, \ldots, x_{j-1}, y_j, \ldots, y_n) d\mu(y) \right| \\
\leqslant \quad & \int_X |f(x_1, \ldots, x_j, y_{j+1}, \ldots, y_n) - f(x_1, \ldots, x_{j-1}, y_j, \ldots, y_n)| d\mu(y) \\
\leqslant \quad & \int_X a_j d\mu(y) = a_j.
\end{aligned}
$$

Το συμπέρασμα έπεται άμεσα από την ανισότητα του Azuma. □

3.4 Ασκήσεις

1. Έστω $Y_1, \ldots, Y_n$ ανεξάρτητες τυχαίες μεταβλητές στον χώρο πιθανότητας $(X, \mathcal{A}, P)$. Υποθέτουμε ότι για κάθε $i = 1, \ldots, n$ υπάρχουν $a_i, b_i \in \mathbb{R}$ ώστε $a_i \leqslant Y_i \leqslant b_i$. Αν $S_n = Y_1 + \cdots + Y_n$, δείξτε ότι

$$P(\{S_n \geqslant \mathbb{E}(S_n) + t\}) \leqslant e^{-t^2/(2D^2)}$$

για κάθε $t > 0$, όπου $D^2 = \sum_{i=1}^n (b_i - a_i)^2$.

2. Θεωρούμε n δοχεία, τα οποία συμβολίζουμε με $1, \ldots, n$ και τοποθετούμε m βώλους στα δοχεία επιλέγοντας τυχαία και ανεξάρτητα, με πιθανότητα $\frac{1}{n}$, ένα δοχείο για κάθε βώλο. Αν X είναι η τυχαία μεταβλητή που δίνει το πλήθος των δοχείων που μένουν κενά, τότε

$$\mathbb{E}(X) = n\left(1 - \frac{1}{n}\right)^m.$$

Αποδείξτε ότι, για κάθε $t > 0$,

$$\mathbb{P}(|X - \mathbb{E}(X)| \geqslant t\sqrt{m}) \leqslant 2e^{-t^2/2}.$$

3. Έστω X χώρος με νόρμα και έστω $x_1, \ldots, x_n \in X$ με $\|x_i\| \leqslant 1$ για κάθε $i = 1, \ldots, n$. Θεωρούμε την τυχαία μεταβλητή

$$X = \|\varepsilon_1 x_1 + \cdots + \varepsilon_n x_n\|,$$

όπου $\varepsilon_1, \ldots, \varepsilon_n$ είναι συμμετρικές τυχαίες μεταβλητές Bernoulli. Αποδείξτε ότι, για κάθε $t > 0$,

$$\mathbb{P}(|X - \mathbb{E}(X)| \geqslant t\sqrt{n}) \leqslant 2e^{-t^2/2}.$$

4. Έστω $Y_1,\dots,Y_n$ ανεξάρτητες τυχαίες μεταβλητές στον χώρο πιθανότητας $(X,\mathcal{A},P)$ με τιμές σε έναν χώρο με νόρμα $(X,\|\cdot\|)$. Υποθέτουμε ότι για κάθε $i=1,\dots,n$ υπάρχει $M_i\in\mathbb{R}$ ώστε $\|Y_i\|\leqslant M_i$. Αν $S_n=Y_1+\cdots+Y_n$, δείξτε ότι

$$P(\{\|S_n\|\geqslant\mathbb{E}(\|S_n\|)+t\})\leqslant e^{-t^2/(2D^2)}$$

για κάθε $t>0$, όπου $D^2=\sum_{i=1}^n M_i^2$.

5. Έστω (X_i,μ_i) χώροι πιθανότητας $(i=1,\dots,n)$ και έστω $\mu=\mu_1\otimes\cdots\otimes\mu_n$ το μέτρο γινόμενο στον $X=X_1\times\cdots\times X_n$. Έστω $c_1,\dots,c_n>0$ και $F:X\to\mathbb{R}$ συνάρτηση που ικανοποιεί την

$$|F(x)-F(y)|\leqslant\sum_{i=1}^n c_i\mathbf{1}_{x_i\neq y_i}$$

για κάθε $x,y\in X$. Δείξτε ότι

$$\mu(\{F\geqslant\mathbb{E}_\mu(F)+t\})\leqslant e^{-t^2/(2D^2)}$$

για κάθε $t>0$, όπου $D^2=\sum_{i=1}^n c_i^2$.

6. Λέμε ότι ένας μετρικός χώρος (X,d) έχει μήκος ℓ αν ο ℓ είναι ο μικρότερος θετικός αριθμός με την εξής ιδιότητα: μπορούμε να βρούμε μια αύξουσα ακολουθία $\{X\}=X^0,X^1,\dots,X^n=\{\{x\}:x\in X\}$ διαμερίσεων του X (αύξουσα σημαίνει ότι η X^i είναι εκλέπτυνση της X^{i-1}) και θετικούς πραγματικούς αριθμούς $a_1,\dots,a_n$ με $\sum_{i=1}^n a_i^2=\ell^2$ ώστε, αν $X^i=\{A_j^i\}_{1\leqslant j\leqslant m_i}$, τότε για κάθε $i=1,\dots,n$, και για κάθε $p=1,\dots,m_{i-1}$ και j,k με $A_j^i,A_k^i\subset A_p^{i-1}$, υπάρχει μια 1-1 και επί απεικόνιση $\phi:A_j^i\to A_k^i$ ώστε $d(x,\phi(x))\leqslant a_i$ για κάθε $x\in A_j^i$.

(α) Δείξτε ότι το μήκος ℓ του (X,d) είναι μικρότερο ή ίσο της διαμέτρου $\mathrm{diam}(X)$ του X.

(β) Έστω (X,d,μ) μετρικός χώρος πιθανότητας με μήκος ℓ. Δείξτε ότι για κάθε 1-Lipschitz συνεχή συνάρτηση $F:X\to\mathbb{R}$ και για κάθε $t>0$ ισχύει

$$\mu(\{F\geqslant\mathbb{E}_\mu(f)+t\})\leqslant e^{-t^2/2\ell^2}.$$

Ειδικότερα, για κάθε $t>0$,

$$\alpha_\mu(t)\leqslant e^{-t^2/8\ell^2}.$$

7. Έστω $F:\mathbb{R}^n\to\mathbb{R}$ συνάρτηση Lipschitz με $\|F\|_{\mathrm{Lip}}\leqslant\alpha$. Υποθέτουμε επίσης ότι

$$|F(x)-F(y)|\leqslant b\|x-y\|_1$$

για κάθε $x,y\in\mathbb{R}^n$. Δείξτε ότι, για κάθε $t>0$,

$$\xi_n(\{F\geqslant M+t\})\leqslant C\exp\left(-\frac{1}{C}\min\left(\frac{t}{b},\frac{t^2}{a^2}\right)\right),$$

όπου ξ_n είναι το συμμετρικό εκθετικό μέτρο γινόμενο στον $\mathbb{R}^n$, $C > 0$ είναι μια απόλυτη σταθερά και M είναι είτε ένας μέσος Lévy της F ή η μέση τιμή $\mathbb{E}(f)$ της f.

8. Για κάθε n-άδα σημείων $x_1, \ldots, x_n \in [0,1]^2$ ορίζουμε

$$G(x_1, \ldots, x_n) = \min_{\sigma \in S_n} \sum_{i=1}^{n} |x_{\sigma(i)} - x_{\sigma(i+1)}|.$$

Θεωρούμε ανεξάρτητες τυχαίες μεταβλητές $X_1, \ldots, X_n$ ομοιόμορφα κατανεμημένες στο $[0,1]^2$ και την τυχαία μεταβλητή $Y := G(X_1, \ldots, X_n)$. Αποδείξτε ότι

$$\mathbb{P}(|Y - \mathbb{E}(Y)| \geqslant t) \leqslant 2 \exp\Big(- \frac{t^2}{c \ln n} \Big),$$

για κάθε $t > 0$, όπου $c > 0$ είναι μια απόλυτη σταθερά.

Κεφάλαιο 4

Συναρτησοειδές Laplace και ελαχιστική συνέλιξη

4.1 Συναρτησοειδές Laplace

Το συναρτησοειδές Laplace ενός μέτρου πιθανότητας μ στον μετρικό χώρο (X,d) μας δίνει μία ακόμα μέθοδο για να αποδείξουμε φράγματα για τη συνάρτηση συγκέντρωσης α_μ μέσω εκθετικών ανισοτήτων απόκλισης των συναρτήσεων Lipschitz $f:(X,d)\to\mathbb{R}$ από τη μέση τιμή τους.

Ορισμός 4.1.1 (συναρτησοειδές Laplace)**.** Έστω (X,d,μ) ένας μετρικός χώρος πιθανότητας. Για κάθε $\lambda\geqslant 0$ ορίζουμε

$$E_\mu(\lambda)=\sup\left\{\int e^{\lambda F}d\mu \mid F:(X,d)\to\mathbb{R},\, 1-\text{Lipschitz συνάρτηση με }\int Fd\mu=0\right\}.$$

Η συνάρτηση $\lambda\mapsto E_\mu(\lambda)$ ονομάζεται *συναρτησοειδές Laplace* του μ.

Η επόμενη Πρόταση δείχνει την σχέση του συναρτησοειδούς Laplace με τη συνάρτηση συγκέντρωσης.

Πρόταση 4.1.2. *Έστω (X,d,μ) ένας μετρικός χώρος πιθανότητας. Για κάθε $t>0$ ισχύει*

$$\alpha_\mu(t)\leqslant\inf_{\lambda\geqslant 0}\left\{e^{-\lambda t/2}E_\mu(\lambda)\right\}. \tag{4.1.1}$$

Απόδειξη. Έστω $f : (X,d) \to \mathbb{R}$ μια 1-Lipschitz συνάρτηση. Θεωρούμε την $F := f - \int f\,d\mu$. Τότε, η F είναι 1-Lipschitz και $\int F\,d\mu = 0$. Από τον ορισμό του συναρτησοειδούς Laplace και από την ανισότητα του Markov, για κάθε $t > 0$ έχουμε

$$e^{\lambda t}\mu(\{F \geqslant t\}) \leqslant \int e^{\lambda F} d\mu \leqslant E_\mu(\lambda),$$

δηλαδή

$$\mu\left(\left\{f - \int f d\mu \geqslant t\right\}\right) \leqslant \alpha(t) := \inf_{\lambda \geqslant 0}\left\{e^{-\frac{\lambda t}{2}} E_\mu(\lambda)\right\}. \tag{4.1.2}$$

Εφαρμόζοντας την Πρόταση 3.3.3 έχουμε το συμπέρασμα. □

Πόρισμα 4.1.3. *Έστω (X,d,μ) ένας μετρικός χώρος πιθανότητας. Υποθέτουμε ότι υπάρχει $c > 0$ ώστε*

$$E_\mu(\lambda) \leqslant e^{\lambda^2/2c}$$

για κάθε $\lambda \geqslant 0$. Τότε, το μ έχει κανονική συνάρτηση συγκέντρωσης: για κάθε $t > 0$ ισχύει

$$\alpha_\mu(t) \leqslant e^{-ct^2/4}.$$

Απόδειξη. Έστω $t > 0$. Ελαχιστοποιούμε την συνάρτηση

$$g(\lambda) = -\frac{\lambda t}{2} + \frac{\lambda^2}{2c}$$

ως προς $\lambda \geqslant 0$ και το συμπέρασμα έπεται από την (4.1.1). □

Πόρισμα 4.1.4. *Έστω (X,d,μ) ένας μετρικός χώρος πιθανότητας. Υποθέτουμε ότι υπάρχει $\lambda_0 > 0$ ώστε*

$$E_\mu(\lambda_0) < +\infty.$$

Τότε, το μ έχει εκθετική συγκέντρωση: για κάθε $t > 0$ ισχύει

$$\alpha_\mu(t) \leqslant E_\mu(\lambda_0)e^{-\lambda_0 t/2}.$$

Απόδειξη. Από την (4.1.2), για κάθε 1-Lipschitz συνάρτηση $f : (X,d) \to \mathbb{R}$ και για κάθε $t > 0$ έχουμε

$$\mu\left(\left\{f - \int f d\mu \geqslant t\right\}\right) \leqslant \alpha_1(t) := E_\mu(\lambda_0)e^{-\frac{\lambda_0 t}{2}}.$$

Το συμπέρασμα έπεται από την Πρόταση 3.3.3. □

Οι επόμενες δύο προτάσεις περιγράφουν δύο πολύ χρήσιμες ιδιότητες του συναρτησοειδούς Laplace: συμπεριφέρεται καλά ως πρός Lipschitz συναρτήσεις μεταξύ μετρικών χώρων και ως προς γινόμενα μετρικών χώρων πιθανότητας αν αυτά εφοδιαστούν με την ℓ_1-μετρική.

Πρόταση 4.1.5. *Έστω (X,d,μ) μετρικός χώρος πιθανότητας και έστω $\varphi : (X,d) \to (Y,\sigma)$ μια Lipschitz συνεχής συνάρτηση με $\|\varphi\|_{\mathrm{Lip}} \leqslant 1$. Θεωρούμε το μέτρο εικόνα $\nu = \varphi(\mu)$ το οποίο ορίζεται μέσω της*

$$\nu(A) = \mu(\varphi^{-1}(A)), \qquad A \in \mathcal{B}(Y).$$

Ισοδύναμα, για κάθε συνεχή συνάρτηση $F : (Y,\sigma) \to \mathbb{R}$ ισχύει

$$\int_X (F \circ \varphi)(x)\, d\mu(x) = \int_Y F(y)\, d\nu(y). \tag{4.1.3}$$

Τότε,

$$E_\nu(\lambda) \leqslant E_\mu(\lambda) \tag{4.1.4}$$

για κάθε $\lambda \geqslant 0$.

Απόδειξη. Έστω $F : (Y,\sigma) \to \mathbb{R}$ μια 1-Lipschitz συνεχής συνάρτηση με $\int F\, d\nu = 0$. Παρατηρήστε ότι η $F \circ \varphi : (X,d) \to \mathbb{R}$ είναι 1-Lipschitz και, από την (4.1.3),

$$\int_X (F \circ \varphi)(x)\, d\mu(x) = \int_Y F(y)\, d\nu(y) = 0.$$

Συνεπώς,

$$\int_Y e^{\lambda F(y)}\, d\nu(y) = \int_X e^{\lambda (F\circ\varphi)(x)}\, d\mu(x) \leqslant E_\mu(\lambda).$$

Παίρνοντας supremum ως προς F καταλήγουμε στην (4.1.4). □

Πρόταση 4.1.6. *Έστω (X,d,μ) και (Y,σ,ν) δύο μετρικοί χώροι πιθανότητας. Θεωρούμε τον $X \times Y$ με μετρική την*

$$\tau((x_1,y_1),(x_2,y_2)) = d(x_1,x_2) + \sigma(y_1,y_2)$$

και το μέτρο γινόμενο $\mu \otimes \nu$. Τότε, για κάθε $\lambda \geqslant 0$,

$$E_{\mu\otimes\nu}(\lambda) \leqslant E_\mu(\lambda) E_\nu(\lambda).$$

Απόδειξη. Έστω $F : (X \times Y, \tau) \to \mathbb{R}$ μια 1-Lipschitz συνεχής συνάρτηση με $\int F\, d(\mu\otimes\nu) = 0$. Για κάθε $y \in Y$ θεωρούμε τη συνάρτηση $F^y : (X,d) \to \mathbb{R}$ με

$$F^y(x) = F(x,y).$$

Επίσης, ορίζουμε $G : (Y, \sigma) \to \mathbb{R}$ με

$$G(y) = \int_X F^y(x)\, d\mu(x).$$

Παρατηρήστε ότι η G και κάθε F^y είναι 1-Lipschitz συναρτήσεις και ότι

$$\int_Y G(y)\, d\nu(y) = \int_{X\times Y} F\, d(\mu\otimes\nu) = 0.$$

Χρησιμοποιώντας την

$$\int_X e^{\lambda(F^y(x) - \int F^y d\mu)} d\mu(x) \leqslant E_\mu(\lambda),$$

γράφουμε

$$\begin{aligned}
\int e^{\lambda F} d(\mu\otimes\nu) &= \int_Y e^{\lambda G(y)} \left(\int_X e^{\lambda(F^y(x) - \int F^y d\mu)} d\mu(x) \right) d\nu(y) \\
&\leqslant \int_Y e^{\lambda G(y)} E_\mu(\lambda) d\nu(y) \\
&\leqslant E_\mu(\lambda) \int_Y e^{\lambda G(y)} d\nu(y) \\
&\leqslant E_\mu(\lambda) E_\nu(\lambda).
\end{aligned}$$

Παίρνοντας supremum ως προς F συμπεραίνουμε ότι $E_{\mu\otimes\nu}(\lambda) \leqslant E_\mu(\lambda)E_\nu(\lambda)$ για κάθε $\lambda \geqslant 0$. □

Μια περίπτωση στην οποία μπορούμε να εφαρμόσουμε το Πόρισμα 4.1.3 είναι όταν ο μετρικός χώρος (X, d) έχει πεπερασμένη, και κυρίως όταν έχει «σχετικά μικρή», διάμετρο.

Θεώρημα 4.1.7. *Έστω (X, d) μετρικός χώρος με $\mathrm{diam}(X, d) = D < +\infty$. Τότε, για κάθε μέτρο πιθανότητας μ στην $\mathcal{B}(X)$ και για κάθε $\lambda \geqslant 0$ ισχύει*

$$E_\mu(\lambda) \leqslant e^{D^2\lambda^2/2}.$$

Απόδειξη. Έστω $F : (X, d) \to \mathbb{R}$ μια 1-Lipschitz συνεχής συνάρτηση με $\int F\, d\mu = 0$. Χρησιμοποιώντας την τελευταία ιδιότητα και την ανισότητα Jensen, για κάθε $\lambda \geqslant 0$ γράφουμε

$$\begin{aligned}
\int e^{\lambda F} d\mu &= \int e^{\lambda F(x)} \cdot e^{-\lambda \int F(y) d\mu(y)} d\mu(x) \\
&= \int e^{\lambda (F(x) - \int F(y) d\mu(y))} d\mu(x) \\
&\leqslant \iint e^{\lambda (F(x) - F(y))} d\mu(x) d\mu(y) \\
&= \sum_{k=0}^{\infty} \frac{\lambda^k \iint (F(x) - F(y))^k d\mu(x) d\mu(y)}{k!} \\
&= \sum_{k=0}^{\infty} \frac{\lambda^{2k} \iint (F(x) - F(y))^{2k} d\mu(x) d\mu(y)}{(2k)!} \\
&\leqslant \sum_{k=0}^{\infty} \frac{(D\lambda)^{2k}}{(2k)!} \\
&\leqslant e^{D^2 \lambda^2 / 2},
\end{aligned} \tag{4.1.5}$$

όπου, στα τελευταία βήματα, χρησιμοποιήσαμε το γεγονός ότι

$$\iint (F(x) - F(y))^k d\mu(x) d\mu(y) = 0$$

αν ο k είναι περιττός, το γεγονός ότι $|F(x) - F(y)| \leqslant d(x,y) \leqslant D$ για κάθε $x, y \in X$ (διότι η F είναι 1-Lipschitz) και το γεγονός ότι

$$\sum_{k=0}^{\infty} \frac{u^{2k}}{(2k)!} \leqslant e^{u^2/2}$$

για κάθε $u \in \mathbb{R}$. □

Πόρισμα 4.1.8. *Έστω (X_i, d_i, μ_i), $1 \leqslant i \leqslant n$, μετρικοί χώροι πιθανότητας με $D_i = \mathrm{diam}(X_i) < \infty$. Θεωρούμε το μέτρο γινόμενο $\mu = \mu_1 \otimes \cdots \otimes \mu_n$ στον $X = X_1 \times \cdots \times X_n$, τον οποίο θεωρούμε εφοδιασμένο με την ℓ_1-μετρική $d = d_1 + \cdots + d_n$. Τότε,*

$$E_\mu(\lambda) \leqslant e^{D^2 \lambda^2 / 2}$$

για κάθε $\lambda \geqslant 0$, όπου $D^2 = D_1^2 + \cdots + D_n^2$. Έπεται ότι, για κάθε 1-Lipschitz συνεχή συνάρτηση $F : (X, d) \to \mathbb{R}$ και για κάθε $t > 0$,

$$\mu\left(\left\{F \geqslant \int F \, d\mu + t\right\}\right) \leqslant e^{-t^2/2D^2}$$

Συνεπώς,

$$\alpha_\mu(t) \leqslant 2e^{-t^2/8D^2}.$$

Απόδειξη. Από την Πρόταση 4.1.6 (με επαγωγή) παίρνουμε

$$E_\mu(\lambda) \leqslant \prod_{i=1}^{n} E_{\mu_i}(\lambda) = \prod_{i=1}^{n} e^{D_i^2\lambda^2/2}$$

για κάθε $\lambda \geqslant 0$, χρησιμοποιώντας την εκτίμηση του Θεωρήματος 4.1.4. Δηλαδή,

$$E_\mu(\lambda) \leqslant e^{\lambda^2 \sum_{i=1}^{n} D_i^2/2} = e^{\lambda^2 D^2/2},$$

από τον ορισμό του D. Οι υπόλοιποι ισχυρισμοί έπονται από τον ορισμό του συναρτησοειδούς Laplace και το Πόρισμα 4.1.3. □

Άμεση εφαρμογή είναι μια δεύτερη απόδειξη της εκτίμησης για τη συνάρτηση συγκέντρωσης του διακριτού κύβου (Θεώρημα 1.4.9).

Θεώρημα 4.1.9. *Θεωρούμε το διακριτό κύβο (E_2^n, d, μ), όπου $\mu(A) = |A|/2^n$ και $d(x,y) = \frac{1}{2n}\sum_{i=1}^{n}|x_i - y_i|$. Τότε, γιά κάθε $t > 0$ έχουμε*

$$\alpha_\mu(t) \leqslant 2\exp(-t^2 n/8).$$

Απόδειξη. Εφαρμόζουμε το Πόρισμα 4.1.8 με $X_i = \{-1, 1\}$ και $d_i(x,y) = \frac{1}{2n}|x-y|$, $1 \leqslant i \leqslant n$. Παρατηρήστε ότι $D_i = 1/n$ για κάθε $i \leqslant n$, άρα

$$D^2 = \sum_{i=1}^{n} D_i^2 = n \cdot \frac{1}{n^2} = \frac{1}{n}.$$

Έτσι, έχουμε

$$\alpha_\mu(t) \leqslant 2e^{-t^2 n/8}$$

από την (4.1.8). □

4.2 Ελαχιστική συνέλιξη

Ορισμός 4.2.1 (ελαχιστική συνέλιξη)**.** Έστω $(X, \mathcal{A}, \mu)$ ένας χώρος πιθανότητας και έστω $\phi : X \times X \to \mathbb{R}^+$. Ονομάζουμε την ϕ *συνάρτηση κόστους* και συνήθως απαιτούμε να ικανοποιεί την $\phi(x,x) = 0$ για κάθε $x \in X$. Τυπικό παράδειγμα είναι η $\phi(x,y) = c|x-y|^2$ στον $\mathbb{R}^n \times \mathbb{R}^n$.

Για κάθε μετρήσιμη συνάρτηση $f : X \to \mathbb{R}$, η *ελαχιστική συνέλιξη* της f με τη συνάρτηση κόστους ϕ είναι η συνάρτηση

$$Q_\phi f(x) = \inf_{y \in X} \big\{ f(y) + \phi(x,y) \big\}, \qquad x \in X.$$

Παρατηρήστε ότι, αν ικανοποιείται η $\phi(x,x) = 0$, τότε

$$Q_\phi f(x) \leqslant f(x), \qquad x \in X.$$

Παρατηρήστε επίσης ότι, για κάθε φραγμένη μετρήσιμη συνάρτηση $f : X \to \mathbb{R}$ ισχύει

$$\int e^f d\mu \cdot \int e^{-f} d\mu \geqslant 1$$

από την ανισότητα Hölder.

Λέμε ότι το μέτρο πιθανότητας μ *ικανοποιεί την ανισότητα ελαχιστικής συνέλιξης ως προς τη συνάρτηση κόστους* ϕ αν για κάθε φραγμένη μετρήσιμη συνάρτηση $f : X \to \mathbb{R}$ ισχύει

$$\int e^{Q_\phi f} d\mu \cdot \int e^{-f} d\mu \leqslant 1.$$

Η επόμενη πρόταση δείχνει την σχέση της ελαχιστικής συνέλιξης με το πρόβλημα της εκτίμησης της συνάρτησης συγκέντρωσης.

Πρόταση 4.2.2. *Έστω $(X, \mathcal{A}, \mu)$ χώρος πιθανότητας. Υποθέτουμε ότι το μ ικανοποιεί την ανισότητα ελαχιστικής συνέλιξης ως προς κάποια συνάρτηση κόστους ϕ. Τότε, για κάθε μετρήσιμο $A \subseteq X$ και για κάθε $t > 0$ έχουμε*

$$1 - \mu\left(\Big\{ \inf_{y \in A} \phi(x,y) < t \Big\}\right) \leqslant \frac{1}{\mu(A)} e^{-t}.$$

Απόδειξη. Για κάθε $n \geqslant t$ θεωρούμε τη συνάρτηση

$$f_{A,n}(x) = \begin{cases} 0, & \text{αν } x \in A \\ n, & \text{αλλιώς.} \end{cases}$$

Παρατηρήστε ότι: $Q_\phi f_{A,n}(x) \geqslant t$ αν και μόνο αν $\inf\limits_{y \in A} \phi(x,y) \geqslant t$. Πράγματι, για κάθε $y \in A$ έχουμε

$$\phi(x,y) = f_{A,n}(y) + \phi(x,y) \geqslant Q_\phi f_{A,n}(x),$$

άρα

$$\inf_{y \in A} \phi(x,y) \geqslant Q_\phi f_{A,n}(x).$$

Από την άλλη πλευρά, αν $y \notin A$ τότε

$$f_{A,n}(y) + \phi(x,y) \geqslant f_{A,n}(x,y) = n \geqslant t.$$

Έπεται ότι

$$\{Q_\phi f_{A,n} \geqslant t\} = \Big\{ \inf_{y \in A} \phi(x,y) \geqslant t \Big\}.$$

Έτσι, χρησιμοποιώντας την ανισότητα του Markov και την ιδιότητα ελαχιστικής συνέλιξης για το ζεύγος (μ, ϕ) γράφουμε

$$\begin{aligned} 1 - \mu\left(\Big\{ \inf_{y \in A} \phi(x,y) < t \Big\} \right) &= \mu\left(\{Q_\phi f_{A,n} \geqslant t\}\right) \\ &\leqslant e^{-t} \int e^{Q_\phi f_{A,n}} d\mu \\ &\leqslant \left(\int e^{-f_{A,n}} d\mu \right)^{-1} e^{-t}. \end{aligned} \tag{4.2.1}$$

Τέλος, παρατηρούμε ότι

$$\int e^{-f_{A,n}} d\mu = \mu(A) + e^{-n}(1 - \mu(A)) \to \mu(A)$$

καθώς το $n \to \infty$. Από την (4.2.1) έχουμε το συμπέρασμα. □

Μια τυπική εφαρμογή αυτού του αποτελέσματος στο πλαίσιο των μετρικών χώρων πιθανότητας είναι η εξής.

Θεώρημα 4.2.3. *Έστω (X,d,μ) μετρικός χώρος πιθανότητας. Θεωρούμε τη συνάρτηση κόστους $\phi(x,y) = \frac{c}{2} d(x,y)^2$ για κάποια σταθερά $c > 0$. Αν το μ ικανοποιεί την ανισότητα ελαχιστικής συνέλιξης ως προς ϕ τότε*

$$E_\mu(\lambda) \leqslant e^{\frac{\lambda^2}{2c}}$$

και

$$\alpha_\mu(t) \leqslant e^{-ct^2/8}$$

Απόδειξη. Έστω $F : (X,d) \to \mathbb{R}$ μια Lipschitz συνεχής συνάρτηση. Για κάθε $x \in X$ έχουμε

$$\begin{aligned} Q_\phi F(x) &= F(x) + \inf_{y \in X} \Big\{ F(y) - F(x) + \frac{c}{2} d(x,y)^2 \Big\} \\ &\geqslant F(x) + \inf_{y \in X} \Big\{ -\|F\|_{\mathrm{Lip}} d(x,y) + \frac{c}{2} d(x,y)^2 \Big\} \\ &\geqslant F(x) - \frac{1}{2c} \|F\|_{\mathrm{Lip}}^2. \end{aligned} \tag{4.2.2}$$

Από την ανισότητα ελαχιστικής συνέλιξης και από την ανισότητα του Jensen, για κάθε $\lambda \geqslant 0$ έχουμε

$$e^{-\frac{\lambda^2}{2c}\|F\|_{\mathrm{Lip}}^2}\int e^{\lambda F}d\mu \leqslant \int e^{Q_\phi(\lambda F)}d\mu \leqslant \left(\int e^{-\lambda F}d\mu\right)^{-1} \leqslant \exp\left(\int \lambda F\, d\mu\right). \tag{4.2.3}$$

Δηλαδή,

$$\int e^{\lambda F}d\mu \leqslant e^{\int \lambda F\, d\mu + \frac{\lambda^2}{2c}\|F\|_{\mathrm{Lip}}^2}.$$

Αν επιπλέον υποθέσουμε ότι $\int F\, d\mu = 0$ και $\|F\|_{\mathrm{Lip}} \leqslant 1$, παίρνουμε

$$\int e^{\lambda F}d\mu \leqslant e^{\frac{\lambda^2}{2c}}.$$

Έπεται ότι

$$E_\mu(\lambda) \leqslant e^{\frac{\lambda^2}{2c}},$$

και από το Πόρισμα 4.1.3 προκύπτει το άνω φράγμα της (4.2.3) για τη συνάρτηση συγκέντρωσης του μ. □

4.3 Η ιδιότητα (τ)

Ορισμός 4.3.1 (ιδιότητα (τ))**.** Έστω f και φ μετρήσιμες συναρτήσεις ορισμένες στον $\mathbb{R}^n$. Με $f\Box\varphi$ συμβολίζουμε την *ελαχιστική συνέλιξη* των f και φ,

$$(f\Box\varphi)(x) = \inf\{f(y) + \varphi(x-y) :\ y \in \mathbb{R}^n\}.$$

Αν μ είναι ένα μέτρο πιθανότητας στον $\mathbb{R}^n$ και φ μία θετική μετρήσιμη συνάρτηση στον $\mathbb{R}^n$, λέμε ότι το ζεύγος (μ,φ) έχει την ιδιότητα (τ) αν για κάθε φραγμένη μετρήσιμη συνάρτηση f στον $\mathbb{R}^n$ ισχύει

$$\left(\int e^{f\Box\varphi}d\mu\right)\left(\int e^{-f}d\mu\right) \leqslant 1.$$

Παρατηρήστε ότι το «(μ,φ) έχει την ιδιότητα (τ)» αν και μόνο αν «το μ ικανοποιεί την ανισότητα ελαχιστικής συνέλιξης ως προς την συνάρτηση κόστους $\phi(x,y) = \varphi(x-y)$».

Τα επόμενα λήμματα περιγράφουν βασικές και χρήσιμες ιδιότητες της ιδιότητας (τ).

Λήμμα 4.3.2. *Αν το (μ_i,φ_i) έχει την ιδιότητα (τ) στον $\mathbb{R}^{n_i}$ για $i = 1,2$, τότε το $(\mu_1\otimes\mu_2,\varphi)$ έχει την ιδιότητα (τ) στον $\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$, όπου $\varphi(x_1,x_2) = \varphi_1(x_1) + \varphi_2(x_2)$.*

Απόδειξη. Έστω $f : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}$ φραγμένη μετρήσιμη συνάρτηση. Ορίζουμε

$$\psi(y) = -\log\left(\int_{\mathbb{R}^{n_1}} e^{-f(x,y)} d\mu_1(x)\right)$$

και $f^y(x) := f(x,y)$. Τότε, χτησιμοποιώντας την ιδιότητα (τ) για το ζεύγος (μ_1, φ_1) βλέπουμε ότι, για κάθε $y, y_1 \in \mathbb{R}^{n_2}$,

$$\int_{\mathbb{R}^{n_1}} e^{f\Box\varphi(x,y)} d\mu_1(x) \leqslant \int_{\mathbb{R}^{n_1}} e^{f^{y_1}\Box\varphi_1(x)+\varphi_2(y-y_1)} d\mu_1(x) \leqslant e^{\psi(y_1)+\varphi_2(y-y_1)}.$$

Έπεται ότι

$$\int_{\mathbb{R}^{n_1}} e^{f\Box\varphi(x,y)} d\mu_1(x) \leqslant e^{\psi\Box\varphi_2(y)}.$$

Κατόπιν, εφαρμόζοντας την ιδιότητα (τ) για το ζεύγος (μ_2, φ_2), βλέπουμε ότι

$$\begin{aligned}\int e^{f\Box\varphi} d(\mu_1 \otimes \mu_2) &\leqslant \int_{\mathbb{R}^{n_2}} e^{\psi\Box\varphi_2(y)} d\mu_2(y) \leqslant \left(\int_{\mathbb{R}^{n_2}} e^{-\psi(y)} d\mu_2(y)\right)^{-1} \\ &\leqslant \left(\int e^{-f} d\mu_1 \otimes d\mu_2\right)^{-1}.\end{aligned}$$

Έπεται ότι το $(\mu_1 \otimes \mu_2, \varphi)$ έχει την ιδιότητα (τ). □

Έστω μ_1 και μ_2 μέτρα πιθανότητας στον $\mathbb{R}^n$. Η συνέλιξη $\mu_1 * \mu_2$ των δύο μέτρων ορίζεται μέσω της

$$\int_{\mathbb{R}^n} h \, d(\mu_1 * \mu_2) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} h(x+y) d\mu_1(x) d\mu_2(y).$$

Στην περίπτωση που τα δύο μέτρα είναι απολύτως συνεχή ως προς το μέτρο Lebesgue με πυκνότητες f_1 και f_2, το $\mu_1 * \mu_2$ είναι το μέτρο που έχει πυκνότητα την $f_1 * f_2$.

Λήμμα 4.3.3. *Αν το (μ_i, φ_i) έχει την ιδιότητα (τ) στον $\mathbb{R}^n$ για $i = 1, 2$, τότε το ζευγάρι $(\mu_1 * \mu_2, \varphi_1 \Box \varphi_2)$ έχει την ιδιότητα (τ) στον $\mathbb{R}^n$.*

Απόδειξη. Η απόδειξη είναι ανάλογη με αυτήν του Λήμματος 4.3.2 και την περιγράφουμε εν συντομία. Έστω g μια φραγμένη μετρήσιμη συνάρτηση στον $\mathbb{R}^n$. Για κάθε $x \in \mathbb{R}^n$ ορίζουμε $g_x : \mathbb{R}^n \to \mathbb{R}$ με $g_x(y) := g(x+y)$. Τότε, για κάθε x έχουμε ότι η g_x είναι μια φραγμένη μετρήσιμη συνάρτηση στον $\mathbb{R}^n$. Ορίζουμε $h : \mathbb{R}^n \to \mathbb{R}$ θέτοντας

$$e^{-h(x)} = \int e^{-g_x} \, d\mu_2.$$

Χρησιμοποιώντας την υπόθεση ότι το (μ_2, φ_2) έχει την ιδιότητα (τ) παίρνουμε:

$$\int e^{\varphi_2 \Box g_x}\, d\mu_2 \leqslant e^{h(x)},$$

και αυτό με τη σειρά του δείχνει ότι

$$(h\Box\varphi_1)(x) \geqslant \log \int \exp([g\Box(\varphi_1\Box\varphi_2)](x+y))\, d\mu_2(y).$$

Συνδυάζοντας αυτή την ανισότητα με το γεγονός ότι το (μ_1, φ_1) έχει την ιδιότητα (τ) παίρνουμε το ζητούμενο. □

Λήμμα 4.3.4. *Έστω ότι το (μ_1, φ_1) έχει την ιδιότητα (τ) στον $\mathbb{R}^{n_1}$. Έστω $\varphi_2 : \mathbb{R}^{n_2} \to \mathbb{R}$ θετική μετρήσιμη συνάρτηση και $g : \mathbb{R}^{n_1} \to \mathbb{R}^{n_2}$ συνάρτηση που ικανοποιεί την $\varphi_2(g(x) - g(y)) \leqslant \varphi_1(x-y)$ για κάθε $x, y \in \mathbb{R}^{n_1}$. Έστω μ_2 το μέτρο πιθανότητας $g(\mu_1)$ στον $\mathbb{R}^{n_2}$, δηλαδή $\mu_2(A) = \mu_1(g^{-1}(A))$. Τότε, το (μ_2, φ_2) έχει την ιδιότητα (τ) στον $\mathbb{R}^{n_2}$.*

Απόδειξη. Ελέγχουμε πρώτα ότι

$$[(f\circ\psi)\Box\varphi_1] \geqslant [(f\Box\varphi_2)\circ\psi]$$

για κάθε φραγμένη μετρήσιμη συνάρτηση $f : \mathbb{R}^{n_2} \to \mathbb{R}$. Από την $\mu_2 = \psi(\mu_1)$ έπεται ότι

$$\begin{aligned}
\int_{\mathbb{R}^{n_2}} e^{(f\Box\varphi_2)(y)} d\mu_2(y) &= \int_{\mathbb{R}^{n_1}} e^{((f\Box\varphi_2)\circ\psi)(x)} d\mu_1(x) \\
&\leqslant \int_{\mathbb{R}^{n_1}} e^{((f\circ\psi)\Box\varphi_1)(x)} d\mu_1(x) \\
&\leqslant \left(\int_{\mathbb{R}^{n_1}} e^{-(f\circ\psi)(x)} d\mu_1(x)\right)^{-1} \\
&= \left(\int_{\mathbb{R}^{n_2}} e^{-f(y)} d\mu_2(y)\right)^{-1}.
\end{aligned}$$

Αυτό αποδεικνύει ότι το (μ_2, φ_2) έχει την ιδιότητα (τ). □

Στο πλαίσιο που συζητάμε, η Πρόταση 4.2.2 παίρνει την εξής μορφή.

Πρόταση 4.3.5. *Έστω ότι το (μ, φ) έχει την ιδιότητα (τ) στον $\mathbb{R}^n$. Για κάθε μετρήσιμο $A \subseteq \mathbb{R}^n$ και για κάθε $t > 0$ έχουμε*

$$\mu\left(x \notin A + \{\varphi < t\}\right) \leqslant (\mu(A))^{-1} e^{-t}.$$

Απόδειξη. Όπως στην απόδειξη της Πρότασης 4.2.2, για κάθε $n \geqslant t$ θεωρούμε τη συνάρτηση

$$f_{A,n}(x) = \begin{cases} 0 & \text{αν } x \in A \\ n & \text{αλλιώς.} \end{cases}$$

και δείχνουμε ότι αν $x \notin A + \{\varphi < t\}$, τότε $(f_{A,n} \square \varphi)(x) \geqslant t$. Πράγματι,

$$(f_{A,n} \square \varphi)(x) = \inf_z \{f_{A,n}(z) + \varphi(x-z)\}. \tag{4.3.1}$$

Αν $z \in A$, τότε $f_{A,n}(z) = 0$ και αφού $x \notin A + \{\varphi < t\}$ έχουμε $\varphi(x-z) \geqslant t$. Άρα,

$$f_{A,n}(z) + \varphi(x-z) \geqslant 0 + t = t.$$

Αν πάλι $z \notin A$, τότε $f_{A,n}(z) = n$ και $\varphi(x-z) \geqslant 0$, άρα

$$f_{A,n}(z) + \varphi(x-z) \geqslant n + 0 \geqslant t.$$

Αυτό αποδεικνύει την (4.3.1). Από την ιδιότητα (τ) έχουμε

$$\begin{aligned} \int_{\mathbb{R}^n} e^{f_{A,n} \square \varphi} d\mu &\leqslant \left(\int_{\mathbb{R}^n} e^{-f_{A,n}} d\mu \right)^{-1} \\ &= \left(\int_A e^{-f_{A,n}} d\mu + \int_{\mathbb{R}^n \setminus A} e^{-f_{A,n}} d\mu \right)^{-1} \\ &= \big(\mu(A) + e^{-n}(1-\mu(A))\big)^{-1} \\ &\leqslant 1/\mu(A). \end{aligned} \tag{4.3.2}$$

Από την ανισότητα του Markov,

$$e^t \mu(x \notin A + \{\varphi < t\}) \leqslant e^t \mu(x : (f_A \square \varphi)(x) \geqslant t) \leqslant \int e^{f_{A,n} \square \varphi} d\mu \leqslant (\mu(A))^{-1}.$$

Άρα,

$$\mu\left(x \notin A + \{\varphi < t\}\right) \leqslant (\mu(A))^{-1} e^{-t}.$$

4.4 Η ανισότητα του Talagrand για το εκθετικό μέτρο γινόμενο

Σε αυτή την παράγραφο χρησιμοποιούμε την ιδιότητα (τ) για να αποδείξουμε μια προσεγγιστική ισοπεριμετρική ανισότητα για το εκθετικό μέτρο γινόμενο ξ_n στον $\mathbb{R}^n$: η πυκνότητα του ξ_n είναι η συνάρτηση $\frac{1}{2^n} e^{-\|x\|_1}$.

Για κάθε μετρήσιμο $A \subseteq \mathbb{R}^n$ και κάθε $t > 0$,

$$\xi_n(x \notin A + 6\sqrt{t}B_2^n + 9tB_1^n) \leqslant \frac{1}{\xi_n(A)} e^{-t}.$$

Η απόδειξη που παρουσιάζουμε οφείλεται στον Maurey.

Αρχικά, ορίζουμε μια συνάρτηση $W : \mathbb{R} \to \mathbb{R}$ με

$$W(t) = \begin{cases} t^2/18 & \text{αν } |t| \leqslant 2 \\ 2(|t|-1)/9 & \text{αλλιώς.} \end{cases}$$

Η W είναι άρτια, κυρτή, συνεχώς παραγωγίσιμη συνάρτηση. Θεωρούμε επίσης το μέτρο πιθανότητας μ_e στο $\mathbb{R}$, με πυκνότητα την $\chi_{(0,+\infty)}(x)e^{-x}$.

Πρόταση 4.4.1. *Το ζεύγος (μ_e, w) έχει την ιδιότητα (τ).*

Απόδειξη. Έστω f φραγμένη συνεχής συνάρτηση στο $(0,+\infty)$. Γράφουμε ψ για την $f\Box w$ και θέτουμε

$$I_0 = \int_0^{+\infty} e^{-f(x)-x}dx \text{ και } I_1 = \int_0^{+\infty} e^{\psi(y)-y}dy.$$

Για κάθε $t \in (0,1)$ ορίζουμε $x(t)$ και $y(t)$ από τις σχέσεις

$$\int_0^{x(t)} e^{-f(x)-x}dx = tI_0 \text{ και } \int_0^{y(t)} e^{\psi(y)-y}dy = tI_1.$$

Από τις παραπάνω σχέσεις φαίνεται ότι οι $x(t)$, $y(t)$ είναι παραγωγίσιμες, με

$$x'(t) = I_0 e^{f(x(t))+x(t)} \text{ και } y'(t) = I_1 e^{-\psi(y(t))+y(t)}.$$

Έχουμε

$$\psi(y(t)) = \inf_{y\in\mathbb{R}}\{f(y) + w(y(t)-y)\} \leqslant f(x(t)) + w(y(t)-x(t)).$$

Άρα,

$$y'(t) \geqslant I_1 e^{-f(x(t))-w(y(t)-x(t))+y(t)}.$$

Θέτουμε

$$z(t) = \frac{x(t)+y(t)}{2} - W(y(t)-x(t)),$$

οπότε

$$\begin{aligned} z'(t) &= \frac{x'(t)+y'(t)}{2} - W'(y(t)-x(t))(y'(t)-x'(t)) \\ &= \left(\frac{1}{2} + W'(y(t)-x(t))\right)x'(t) + \left(\frac{1}{2} - W'(y(t)-x(t))\right)y'(t). \end{aligned} \tag{4.4.1}$$

Εύκολα ελέγχουμε ότι $|W'| \leqslant 1/2$ στο $\mathbb{R}$, άρα η $z(t)$ είναι αύξουσα.

Γράφουμε x, y αντί των $x(t), y(t)$. Χρησιμοποιώντας την ανισότητα

$$\frac{1}{2}\left(ua + \frac{v}{a}\right) \geqslant \sqrt{uv}, \quad u, v, a > 0$$

με $a = \exp(f(x))$, παίρνουμε

$$\begin{aligned}
z'(t) &\geqslant (1 - 2W'(y-x))I_0 e^x \frac{e^{f(x)}}{2} + (1 + 2W'(y-x))I_1 e^{-W(y-x)+y}\frac{e^{-f(x)}}{2} \\
&\geqslant \sqrt{1 - 4(W'(y-x))^2}\sqrt{I_0 I_1} e^{(x+y)/2 - W(y-x)/2} \\
&= \sqrt{1 - 4(W'(y-x))^2}\sqrt{I_0 I_1} e^{(x+y)/2 - W(y-x)} e^{W(y-x)/2} \\
&= \sqrt{1 - 4(W'(y-x))^2}\sqrt{I_0 I_1} e^{z(t)} e^{W(y-x)/2}.
\end{aligned} \tag{4.4.2}$$

Ισχυρισμός 4.4.2. Για κάθε s,

$$\left(1 - 4(W'(s))^2\right) e^{W(s)} \geqslant 1.$$

Απόδειξη του ισχυρισμού: Αφού η W είναι άρτια, αρκεί να αποδείξουμε την ανισότητα για $s \geqslant 0$. Στο $[2, +\infty)$ έχουμε $W'(s) = 2/9$ και η W είναι αύξουσα. Αν λοιπόν η ανισότητα ισχύει για $s = 2$, τότε θα ισχύει για κάθε $s \geqslant 2$. Ζητάμε

$$\left(1 - 4(2/9)^2\right) e^{2/9} \geqslant 1$$

ή, ισοδύναμα, $e^{2/9} \geqslant 81/65$. Η τελευταία ανισότητα ισχύει γιατί

$$e^{2/9} \geqslant 1 + \frac{2}{9} + \frac{1}{2}\left(\frac{2}{9}\right)^2 = \frac{101}{81} > \frac{81}{65}.$$

Για $s \in [0, 2]$ έχουμε $W'(s) = s/9$, οπότε ζητάμε την $e^{-s^2/18} \leqslant 1 - 4s^2/81$. Αρκεί λοιπόν να δείξουμε ότι η συνάρτηση

$$f(u) = 1 - \frac{4u}{81} - e^{-u/18}$$

παίρνει μη αρνητικές τιμές στο $[0, 4]$. Παραγωγίζοντας βλέπουμε ότι η f είναι κοίλη, άρα αρκεί να εξετάσουμε τις τιμές $f(0)$ και $f(4)$. Όμως, $f(0) = 0$ και η $f(4) \geqslant 0$ είναι ισοδύναμη με την $e^{2/9} \geqslant 81/65$ η οποία, όπως είδαμε, ισχύει. $\square$

Από τον ισχυρισμό και την προηγούμενη ανισότητα συμπεραίνουμε ότι

$$z'(t) \geqslant \sqrt{I_0 I_1} e^{z(t)},$$

άρα
$$\left(-e^{-z(t)}\right)' \geqslant \sqrt{I_0 I_1}.$$
Ολοκληρώνοντας στο $[0,1]$ και χρησιμοποιώντας το γεγονός ότι $z(0)=0$, παίρνουμε
$$1 \geqslant e^{-z(0)} - e^{-z(1)} = \int_0^1 \left(-e^{-z(t)}\right)' dt \geqslant \sqrt{I_0 I_1}.$$
Δηλαδή,
$$\left(\int_0^\infty e^{f\square W} d\mu_e\right)\left(\int_0^\infty e^{-f} d\mu_e\right) = I_0 I_1 \leqslant 1.$$
Αφού η f ήταν τυχούσα, το (μ_e, W) έχει την ιδιότητα (τ). □

Θεωρούμε τώρα τη συμμετρική εικόνα μ_e' του μ_e στο $(-\infty,0)$, με πυκνότητα την $\chi_{(-\infty,0)}(x)e^x$. Λόγω συμμετρίας, το (μ_e', W) έχει την ιδιότητα (τ). Αν ξ είναι το εκθετικό μέτρο πιθανότητας στο $\mathbb{R}$ με πυκνότητα την $\frac{1}{2}e^{-|x|}$, εύκολα ελέγχουμε οτι
$$\xi = \mu_e * \mu_e'.$$
Από το Λήμμα 4.3.3, το ζευγάρι $(\xi, W\square W)$ έχει την ιδιότητα (τ). Παίρνοντας υπ' όψιν τον ορισμό της W, βλέπουμε οτι η $U := W\square W$ δίνεται από την
$$U(t) = \begin{cases} t^2/36 & \text{αν } |t| \leqslant 4 \\ 2(|t|-2)/9 & \text{αλλιώς.} \end{cases}$$
Θεωρούμε τώρα το μέτρο γινόμενο $\xi_n = \xi \otimes \cdots \otimes \xi$ (n φορές) στον $\mathbb{R}^n$. Αν ορίσουμε τη συνάρτηση $U_n : \mathbb{R}^n \to \mathbb{R}$ με
$$U_n(x_1,\ldots,x_n) = \sum_{i=1}^n U(x_i),$$
το Λήμμα 4.3.2 μας δίνει το εξής αποτέλεσμα.

Θεώρημα 4.4.3. *Το ζευγάρι (ξ_n, U_n) έχει την ιδιότητα (τ) στον $\mathbb{R}^n$.* □

Από το Θεώρημα 4.4.3 και από την Πρόταση 4.4.1 έπεται ότι για κάθε μετρήσιμο $A \subseteq \mathbb{R}^n$ και κάθε $t>0$,
$$\xi_n(x \notin A + \{U_n < t\}) \leqslant \frac{1}{\xi_n(A)} e^{-t}.$$
Από τον ορισμό της U_n και την (4.4) προκύπτει η ανισότητα του Talagrand για το ξ_n:

Θεώρημα 4.4.4. *Για κάθε μετρήσιμο $A \subseteq \mathbb{R}^n$ και κάθε $t > 0$,*

$$\xi_n(x \notin A + 6\sqrt{t}B_2^n + 9tB_1^n) \leqslant \frac{1}{\xi_n(A)} e^{-t}.$$

Απόδειξη. Αρκεί να δείξουμε οτι

$$\{U_n < t\} \subseteq 6\sqrt{t}B_2^n + 9tB_1^n.$$

Έστω $x \in \mathbb{R}^n$ με $U_n(x) < t$. Ορίζουμε y και z στον $\mathbb{R}^n$ ως εξής: $y_i = x_i$ αν $|x_i| \leqslant 4$ και $y_i = 0$ αλλιώς, $z_i = x_i$ αν $|x_i| > 4$ και $z_i = 0$ αλλιώς. Προφανώς,

$$x = y + z.$$

Παρατηρούμε οτι

$$|y|^2 = \sum_{\{i:|x_i|\leqslant 4\}} x_i^2 = 36 \sum_{\{i:|x_i|\leqslant 4\}} U(x_i) \leqslant U_n(x) < 36t,$$

άρα $y \in 6\sqrt{t}B_2^n$. Επίσης, αν $|x_i| > 4$, τότε

$$U(x_i) = \frac{2}{9}(|x_i| - 2) \geqslant \frac{2}{9}\left(|x_i| - \frac{|x_i|}{2}\right) = \frac{|x_i|}{9},$$

άρα

$$\|z\|_1 = \sum_{\{i:|x_i|>4\}} |x_i| \leqslant 9 \sum_{\{i:|x_i|>4\}} U(x_i) \leqslant 9U_n(x) < 9t,$$

δηλαδή $z \in 9tB_1^n$. □

4.5 Η ιδιότητα (τ) στο χώρο του Gauss

Θεωρούμε το μέτρο του Gauss γ στο $\mathbb{R}$, με πυκνότητα την $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ και το n-διάστατο μέτρο του Gauss $\gamma_n = \gamma \otimes \cdots \otimes \gamma$ στον $\mathbb{R}^n$.

Θεώρημα 4.5.1. *Το ζευγάρι $(\gamma_n, |x|^2/4)$ έχει την ιδιότητα (τ).*

Απόδειξη. Από το Λήμμα 4.3.2, αρκεί να δείξουμε οτι το $(\gamma, x^2/4)$ έχει την ιδιότητα (τ) στο $\mathbb{R}$. Μπορεί κανείς να δώσει απόδειξη αυτού του ισχυρισμού παρόμοια με αυτήν της Πρότασης 4.4.1. Θα δώσουμε όμως απευθείας απόδειξη στον $\mathbb{R}^n$ χρησιμοποιώντας την ανισότητα Prékopa-Leindler.

Έστω f φραγμένη μετρήσιμη συνάρτηση στον $\mathbb{R}^n$. Ορίζουμε $w(y) = |y|^2/4$ και $\psi = f\Box w$. Αν

$$u(x) = f(x) + \frac{|x|^2}{2} \quad , \quad g(y) = -\psi(y) + \frac{|y|^2}{2} \quad \text{και} \quad h(z) = \frac{|z|^2}{2},$$

τότε εύκολα ελέγχουμε οτι

$$h\left(\frac{x+y}{2}\right) \leqslant \frac{u(x)+g(y)}{2}.$$

Άρα,

$$\left(\int_{\mathbb{R}^n} e^{-u(x)}dx\right)\left(\int_{\mathbb{R}^n} e^{-g(x)}dx\right) \leqslant \left(\int_{\mathbb{R}^n} e^{-h(x)}dx\right)^2.$$

Δηλαδή,

$$\left(\int_{\mathbb{R}^n} e^{-f}d\gamma_n\right)\left(\int_{\mathbb{R}^n} e^{f\Box w}d\gamma_n\right) \leqslant 1$$

που είναι το ζητούμενο. □

Σαν εφαρμογή του Θεωρήματος 4.5.1 παίρνουμε μια ανισότητα του Pisier για τη συγκέντρωση των Lipschitz συναρτήσεων ως προς το μέτρο γ_n.

Θεώρημα 4.5.2. *Έστω $f : \mathbb{R}^n \to \mathbb{R}$ Lipschitz συνάρτηση με σταθερά 1. Για κάθε $t > 0$ ισχύει*

$$\iint \left(\exp\left(\frac{t(f(x)-f(y))}{\sqrt{2}}\right)\right) d\gamma_n(x)d\gamma_n(y) \leqslant e^{t^2/2}.$$

Απόδειξη. Θεωρούμε την $w(y) = |y|^2/4$ και ορίζουμε $\psi = (tf/\sqrt{2})\Box w$. Έστω $x \in \mathbb{R}^n$ και $y \in \mathbb{R}^n$ τέτοιο ώστε

$$\psi(x) = \frac{tf(y)}{\sqrt{2}} + \frac{|x-y|^2}{4}.$$

Τότε, αφού $\|f\|_{\mathrm{Lip}} \leqslant 1$,

$$\begin{aligned} \psi(x) &\geqslant \frac{tf(x)}{\sqrt{2}} - \frac{t}{\sqrt{2}}|x-y| + \frac{|x-y|^2}{4} \geqslant \frac{tf(x)}{\sqrt{2}} + \min_{s\in\mathbb{R}}\left\{\frac{s^2}{4} - \frac{ts}{\sqrt{2}}\right\} \\ &= \frac{tf(x)}{\sqrt{2}} - \frac{t^2}{2}. \end{aligned} \tag{4.5.1}$$

Από το Θεώρημα 4.5.1,

$$\left(\int e^{\psi}d\gamma_n\right)\left(\int e^{-tf/\sqrt{2}}d\gamma_n\right) \leqslant 1.$$

Χρησιμοποιώντας την προηγούμενη ανισότητα παίρνουμε

$$\int e^{tf/\sqrt{2}} d\gamma_n \cdot \int e^{-tf/\sqrt{2}} d\gamma_n \leqslant e^{t^2/2},$$

δηλαδή

$$\iint \left(\exp\left(\frac{t(f(x)-f(y))}{\sqrt{2}}\right)\right) d\gamma_n(x) d\gamma_n(y) \leqslant e^{t^2/2}.$$

Πόρισμα 4.5.3. *Αν η $f:\mathbb{R}^n \to \mathbb{R}$ είναι συνάρτηση Lipschitz με $\|f\|_{\mathrm{Lip}} \leqslant 1$, τότε*

$$\gamma_n \left(x : \left| f(x) - \int_{\mathbb{R}^n} f d\gamma_n \right| > s \right) \leqslant 2e^{-s^2/4}$$

για κάθε $s>0$.

Απόδειξη. Έστω $t>0$. Από το Θεώρημα 4.5.2 και την ανισότητα του Jensen,

$$\int \left(\exp\left(\frac{t}{\sqrt{2}} (f(x) - \mathbb{E}(f))\right)\right) d\gamma_n(x) \leqslant e^{t^2/2}.$$

Άρα, για κάθε $s>0$,

$$\gamma_n \left(x : f(x) - \mathbb{E}f > s \right) \leqslant \exp\left(\frac{t^2}{2} - \frac{ts}{\sqrt{2}} \right).$$

Ελαχιστοποιώντας ως προς t και ακολουθώντας την ίδια διαδικασία για την $-f$ παίρνουμε το ζητούμενο. □

4.6 Ασκήσεις

1. Έστω $F:\mathbb{R}^n \to \mathbb{R}$ συνάρτηση Lipschitz με $\|F\|_{\mathrm{Lip}} \leqslant \alpha$. Υποθέτουμε επίσης ότι

$$|F(x) - F(y)| \leqslant b\|x-y\|_1$$

για κάθε $x,y \in \mathbb{R}^n$. Δείξτε ότι, για κάθε $t>0$,

$$\xi_n(\{F \geqslant M + t\}) \leqslant C \exp\left(-\frac{1}{C} \min\left(\frac{t}{b}, \frac{t^2}{a^2} \right)\right),$$

όπου $C>0$ είναι μια απόλυτη σταθερά και M είναι είτε ένας μέσος Lévy της F ή η μέση τιμή $\mathbb{E}(f)$ της f.

2. Έστω X ένας χώρος με νόρμα, εφοδιασμένος με ένα Borel μέτρο πιθανότητας και έστω $\phi : X \to \mathbb{R}^+$ κυρτή συνάρτηση κόστους. Η ελαχιστική συνέλιξη μιας μετρήσιμης $f:X\to\mathbb{R}$ ορίζεται ως εξής:

$$Q_\phi f(x) = \inf_{y\in X} \left[f(y) + \phi(x-y) \right].$$

Λέμε ότι το μ ικανοποιεί την ανισότητα κυρτής ελαχιστικής συνέλιξης ως προς την ϕ αν για κάθε φραγμένη μετρήσιμη κυρτή συνάρτηση $f : X \to \mathbb{R}$ ισχύει

$$\int e^{Q_\phi f} d\mu \cdot \int e^{-f} d\mu \leqslant 1.$$

(α) Στο προηγούμενο πλαίσιο, υποθέτουμε ότι το μ έχει φορέα κάποιο σύνολο A διαμέτρου $\mathrm{diam}(A) \leqslant 1$. Δείξτε ότι το μ ικανοποιεί την ανισότητα κυρτής ελαχιστικής συνέλιξης ως προς την $\phi(x) = \frac{\|x\|^2}{4}$.

(β) Έστω $X_1, \ldots, X_n$ χώροι με νόρμα. Υποθέτουμε ότι, για κάθε $i = 1, \ldots, n$ έχουμε ένα μέτρο πιθανότητας μ_i στον X_i το οποίο έχει φορέα κάποιο σύνολο A_i διαμέτρου $\mathrm{diam}(A_i) \leqslant 1$. Δείξτε ότι το $\mu = \mu_1 \otimes \cdots \otimes \mu_n$ στον $X = X_1 \times \cdots \times X_n$ ικανοποιεί την ανισότητα κυρτής ελαχιστικής συνέλιξης ως προς την $\phi(x) = \frac{1}{4} \sum_{i=1}^n \|x_i\|^2$.

3. Έστω μ ένα μέτρο πιθανότητας στον $\mathbb{R}^n$ με πυκνότητα e^{-V}, όπου $V : \mathbb{R}^n \to \mathbb{R}$ κυρτή συνάρτηση. Υποθέτουμε ότι υπάρχει $w : \mathbb{R}^n \to \mathbb{R}^+$ ώστε

$$V(x) + V(y) - 2V\left(\frac{x+y}{2}\right) \geqslant w(x-y)$$

για κάθε $x, y \in \mathbb{R}^n$. Δείξτε ότι: για κάθε $f : \mathbb{R}^n \to (-\infty, \infty]$ με $\int e^{-f} d\mu \in (0, \infty)$, ισχύει

$$\int e^{f \square w} d\mu \cdot \int e^{-f} d\mu \leqslant 1.$$

4. Έστω μ ένα μέτρο πιθανότητας στον $\mathbb{R}^n$ με πυκνότητα e^{-V}, όπου $V : \mathbb{R}^n \to \mathbb{R}$ κυρτή συνάρτηση. Υποθέτουμε ότι υπάρχουν $c > 0$, $p \geqslant 1$ και μια νόρμα $\|\cdot\|$ στον $\mathbb{R}^n$ ώστε

$$V(x) + V(y) - 2V\left(\frac{x+y}{2}\right) \geqslant \frac{c}{2p}\|x-y\|^p$$

για κάθε $x, y \in \mathbb{R}^n$. Δείξτε ότι: για κάθε Borel σύνολο $A \subseteq \mathbb{R}^n$ ισχύει

$$\int e^{\frac{c}{2p} d(x,A)^p} d\mu(x) \leqslant \frac{1}{\mu(A)},$$

όπου $d(x, A) = \inf\{\|x-y\| : y \in A\}$. Συνεπώς, αν $\mu(A) \geqslant \frac{1}{2}$ τότε, για κάθε $t > 0$,

$$\mu(\{x : d(x, A) \geqslant t\}) \leqslant 2 \exp\left(-\frac{c}{2p} t^p\right).$$

5. Έστω μ ένα συμμετρικό Borel μέτρο πιθανότητας στον $\mathbb{R}^n$. Ορίζουμε

$$\Lambda(u) = \log\left(\int_{\mathbb{R}^n} e^{\langle u, x\rangle} d\mu(x)\right)$$

και
$$\Lambda^*(v) = \sup\{\langle u, v\rangle - \Lambda(u) : u \in \mathbb{R}^n\}.$$
Έστω $\varphi : \mathbb{R}^n \to \mathbb{R}^+$ κυρτή συνάρτηση κόστους τέτοια ώστε το ζεύγος (μ, φ) να έχει την ιδιότητα (τ). Αποδείξτε ότι
$$\varphi(v) \leqslant 2\Lambda^*(v/2) \leqslant \Lambda^*(v).$$

6. Έστω A, B συμμετρικά κυρτά σώματα στον $\mathbb{R}^n$ και p_A, p_B τα συναρτησοειδή Minkowski που αντιστοιχούν σε αυτά. Περιγράψτε το
$$\{x \in \mathbb{R}^n : (p_A \Box p_B)(x) < 1\}.$$
[*Υπενθύμιση:* $p_C(x) = \inf\{\lambda > 0 : x \in \lambda C\}$.

Κεφάλαιο 5

Ανισότητα Poincaré

5.1 Ανισότητα Cheeger και ανισότητα Poincaré

Υπενθυμίζουμε ότι αν μ είναι ένα Borel μέτρο πιθανότητας στον $\mathbb{R}^n$ τότε η επιφάνεια κατά Minkowski ενός Borel υποσυνόλου A του $\mathbb{R}^n$ ως προς το μ ορίζεται από την

$$\mu^+(A) = \liminf_{t\to 0^+} \frac{\mu(A_t) - \mu(A)}{t}, \tag{5.1.1}$$

όπου $A_t = A + tB_2^n = \{x : d(x,A) < t\}$ είναι η t-επέκταση του A ως προς την Ευκλείδεια μετρική.

Ορισμός 5.1.1. Έστω μ ένα Borel μέτρο πιθανότητας $\mathbb{R}^n$. Λέμε ότι το μ ικανοποιεί την *ανισότητα Cheeger* με σταθερά $\beta \geqslant 0$ αν

$$\mu^+(A) \geqslant \beta \min\{\mu(A), 1 - \mu(A)\} \tag{5.1.2}$$

για κάθε Borel υποσύνολο A του $\mathbb{R}^n$. Η *σταθερά Cheeger* $\mathrm{Is}(\mu)$ του μ είναι η μεγαλύτερη σταθερά $\beta \geqslant 0$ για την οποία η (5.1.2) ισχύει για κάθε A.

Το επόμενο θεώρημα δίνει μια συναρτησιακή περιγραφή της σταθεράς Cheeger. Η σύνδεση με τις συναρτησιακές ανισότητες δίνεται μέσω της *co-area formula.*

Θεώρημα 5.1.2. *Για κάθε μ-ολοκληρώσιμη συνάρτηση Lipschitz $f : \mathbb{R}^n \to \mathbb{R}$ ισχύει*

$$\int_{\mathbb{R}^n} |\nabla f(x)|\, d\mu(x) \geqslant \int_0^\infty \mu^+(\{x : |f(x)| > t\})\, dt. \tag{5.1.3}$$

Απόδειξη. Για κάθε $t > 0$ ορίζουμε $f_t : \mathbb{R}^n \to \mathbb{R}$ με

$$f_t(x) = \sup_{|x-y|<t} |f(y)|.$$

Η f_t είναι μετρήσιμη και για κάθε $x \in \mathbb{R}^n$

$$\limsup_{t\to 0} \frac{f_t(x) - |f(x)|}{t} = \limsup_{y\to x} \frac{|f(y)| - |f(x)|}{|y-x|} \leqslant |\nabla f(x)|.$$

Χρησιμοποιώντας το λήμμα του Fatou γράφουμε

$$\begin{aligned}
\int_{\mathbb{R}^n} |\nabla f(x)|\, d\mu(x) &\geqslant \int_{\mathbb{R}^n} \limsup_{t\to 0} \frac{f_t(x) - |f(x)|}{t}\, d\mu(x) \\
&\geqslant \limsup_{t\to 0} \int_{\mathbb{R}^n} \frac{f_t(x) - |f(x)|}{t}\, d\mu(x) \\
&\leqslant \liminf_{t\to 0} \int_{\mathbb{R}^n} \frac{f_t(x) - |f)x)|}{t}\, d\mu(x) \\
&= \liminf_{t\to 0} \frac{1}{t} \int_0^\infty [\mu(\{x : f_t(x) > s\}) - \mu(\{x : |f(x)| > s\})]\, ds.
\end{aligned}$$

Για σταθερό $s > 0$ ορίζουμε $A(s) = \{x : |f(x)| > s\}$. Παρατηρήστε ότι

$$A(s)_t = A(s) + tB_2^n = \{x : f_t(x) > s\}.$$

Άρα, χρησιμοποιώντας και πάλι το λήμμα του Fatou, έχουμε

$$\begin{aligned}
\int_{\mathbb{R}^n} |\nabla f(x)|\, d\mu(x) &\geqslant \liminf_{t\to 0} \int_0^\infty [\mu(A(s)_t) - \mu(A(s))]\, ds \\
&\geqslant \int_0^\infty \liminf_{t\to 0} \frac{\mu(A(s)_t) - \mu(A(s))}{t}\, ds \\
&= \int_0^\infty \mu^+(A(s))\, ds,
\end{aligned}$$

δηλαδή έχουμε δείξει την (5.1.3). □

Θεώρημα 5.1.3. *Έστω μ ένα Borel μέτρο πιθανότητας στον $\mathbb{R}^n$.*

(α) *Αν το μ ικανοποιεί την (5.1.2) με σταθερά β τότε, για κάθε ολοκληρώσιμη τοπικά Lipschitz συνάρτηση $f : \mathbb{R}^n \to \mathbb{R}$,*

$$\|f - \mathbb{E}_\mu(f)\|_1 \leqslant \frac{1}{c} \|\,|\nabla f|\,\|_1 \tag{5.1.4}$$

για κάποια $c \geqslant \beta/2$.

(β) *Αντίστροφα, αν το μ ικανοποιεί την* (5.1.4) *για κάποια $c > 0$ και όλες τις f όπως παραπάνω, τότε* $\mathrm{Is}(\mu) \geqslant c$.

Απόδειξη. (α) Έστω $f : \mathbb{R}^n \to \mathbb{R}$ μια ολοκληρώσιμη τοπικά Lipschitz συνάρτηση. Χωρίς περιορισμό της γενικότητας μπορούμε να υποθέσουμε ότι η f είναι κάτω φραγμένη, δηλαδή υπάρχει $\delta \in \mathbb{R}$ ώστε η $f_1 := f + \delta$ να είναι θετική συνάρτηση. Παρατηρήστε ότι η f_1 είναι ολοκληρώσιμη και τοπικά Lipschitz, με $\| |\nabla f_1| \|_1 = \| |\nabla f| \|_1$ και $\|f_1 - \mathbb{E}_\mu(f_1)\|_1 = \|f - \mathbb{E}_\mu(f)\|_1$.

Για κάθε $t \geqslant 0$ ορίζουμε $A_t = \{x : f_1(x) > t\}$. Χρησιμοποιώντας την υπόθεση και την (5.1.3) γράφουμε

$$\int_{\mathbb{R}^n} |\nabla f_1(x)|\, d\mu(x) \geqslant \int_0^\infty \mu^+(\{x : f_1(x) > t\})\, dt \geqslant \beta \int_0^\infty \min\{\mu(A_t), 1 - \mu(A_t)\}\, dt.$$

Αφού

$$\|\mathbf{1}_B - \mathbb{E}_\mu(\mathbf{1}_B)\|_1 = 2\mu(B)(1 - \mu(B)) \leqslant 2\min\{\mu(B), 1 - \mu(B)\}$$

για κάθε Borel υποσύνολο B του $\mathbb{R}^n$, μπορούμε να συνεχίσουμε ως εξής: για κάθε $g \in L_\infty(\mu)$ με $\|g\|_\infty \leqslant 1$ έχουμε

$$\begin{aligned}
\int_{\mathbb{R}^n} |\nabla f_1(x)|\, d\mu(x) &\geqslant \frac{\beta}{2} \int_0^\infty \|\mathbf{1}_{A_t} - \mathbb{E}_\mu(\mathbf{1}_{A_t})\|_1\, dt \\
&\geqslant \frac{\beta}{2} \int_0^\infty \int_{\mathbb{R}^n} (\mathbf{1}_{A_t}(x) - \mathbb{E}_\mu(\mathbf{1}_{A_t}) g(x)\, d\mu(x)\, dt \\
&= \frac{\beta}{2} \int_0^\infty \int_{\mathbb{R}^n} \mathbf{1}_{A_t}(x)(g(x) - \mathbb{E}_\mu(g))\, d\mu(x)\, dt \\
&= \frac{\beta}{2} \int_{\mathbb{R}^n} (g(x) - \mathbb{E}_\mu(g)) f_1(x)\, d\mu(x) \\
&= \frac{\beta}{2} \int_{\mathbb{R}^n} (f_1(x) - \mathbb{E}_\mu(f_1)) g(x)\, d\mu(x).
\end{aligned}$$

Παίρνοντας το supremum ως προς g συμπεραίνουμε ότι

$$\int_{\mathbb{R}^n} |\nabla f_1(x)|\, d\mu(x) \geqslant \frac{\beta}{2} \|f_1 - \mathbb{E}_\mu(f_1)\|_1,$$

και ο πρώτος ισχυρισμός του θεωρήματος έχει αποδειχτεί.

Για την απόδειξη του (β), έστω B ένα Borel υποσύνολο του $\mathbb{R}^n$. Μπορούμε να υποθέσουμε ότι $\mu(B) = \mu(\overline{B})$, όπου $\overline{B}$ είναι η κλειστή θήκη του B, αλλιώς $\mu^+(B) = \infty$. Για κάθε $0 < \varepsilon < 1$ θεωρούμε τη συνάρτηση

$$f_\varepsilon(x) = \max\Big\{0, 1 - \frac{d(x, B_{\varepsilon^2})}{\varepsilon(1 - \varepsilon)}\Big\}.$$

Παρατηρήστε ότι $0 \leqslant f_\varepsilon \leqslant 1$, $f_\varepsilon \equiv 1$ στο B_{ε^2} και $f_\varepsilon(x) = 0$ αν $d(x,B) > \varepsilon$. Επίσης, έχουμε

$$|f_\varepsilon(x) - f_\varepsilon(y)| \leqslant \frac{1}{\varepsilon(1-\varepsilon)} |d(x,B_{\varepsilon^2}) - d(y,B_{\varepsilon^2})| \leqslant \frac{|x-y|}{\varepsilon(1-\varepsilon)},$$

δηλαδή η f_ε είναι τοπικά Lipschitz με $\| |\nabla f_\varepsilon| \|_\infty \leqslant \frac{1}{\varepsilon(1-\varepsilon)}$. Αφού $\nabla f_\varepsilon(x) = 0$ για κάθε $x \in B \cup \{x : d(x,B) > \varepsilon\}$ παίρνουμε

$$\| |\nabla f_\varepsilon| \|_1 \geqslant \int_{\{x:d(x,B)\leqslant\varepsilon\}\setminus B} |\nabla f_\varepsilon(x)|\, d\mu(x) \leqslant \frac{\mu(B_{\varepsilon+\varepsilon^2}) - \mu(B)}{\varepsilon(1-\varepsilon)}.$$

Από την (5.1.4) έπεται ότι

$$\|f_\varepsilon - \mathbb{E}_\mu(f_\varepsilon)\|_1 \leqslant \frac{1}{c} \frac{\mu(B_{\varepsilon+\varepsilon^2}) - \mu(B)}{\varepsilon(1-\varepsilon)}.$$

Παρατηρήστε ότι $f_\varepsilon \to \mathbf{1}_{\overline{B}}$ καθώς το $\varepsilon \to 0^+$. Αφήνοντας το $\varepsilon \to 0^+$, από την προηγούμενη ανισότητα παίρνουμε

$$\min\{\mu(B), 1-\mu(B)\} \leqslant 2\mu(B)(1-\mu(B)) \leqslant \liminf_{\varepsilon\to 0^+} \|f_\varepsilon - \mathbb{E}_\mu(f_\varepsilon)\|_1 \leqslant \frac{1}{c}\mu^+(B).$$

Αυτό αποδεικνύει ότι το μ ικνοποιεί την ανισότητα Cheeger με σταθερά $\beta \geqslant c$. $\square$

Στη συνέχεια εισάγουμε την ανισότητα Poincaré και ορίζουμε τη σταθερά Poincaré του μ.

Ορισμός 5.1.4. Έστω μ ένα Borel μέτρο πιθανότητας στον $\mathbb{R}^n$. Λέμε ότι το μ ικανοποιεί την *ανισότητα Poincaré* με σταθερά $\beta > 0$ αν

$$\beta \mathrm{Var}_\mu(f) \leqslant \int |\nabla f|^2 \, d\mu, \tag{5.1.5}$$

για όλες τις ομαλές συναρτήσεις f στον $\mathbb{R}^n$, όπου $\mathrm{Var}_\mu(g) = \mathbb{E}_\mu(g^2) - (\mathbb{E}_\mu(g))^2$ είναι η διασπορά της g ως προς το μ. Η *σταθερά Poincaré* $\mathrm{Poin}(\mu)$ του μ είναι η μεγαλύτερη σταθερά $\beta > 0$ για την οποία ικανοποιείται η (5.1.5).

Το επόμενο θεώρημα δείχνει ότι η σταθερά Poincaré του μ φράσσεται από κάτω από τη σταθερά Cheeger του μ. Με άλλα λόγια, η ανισότητα Poincaré είναι συνέπεια της ισοπεριμετρικής ανισότητας (5.1.2).

Θεώρημα 5.1.5 (Maz'ya, Cheeger)**.** *Έστω μ ένα Borel μέτρο πιθανότητας στον $\mathbb{R}^n$ που ικανοποιεί την ανισότητα Cheeger. Τότε,*

$$\sqrt{\mathrm{Poin}(\mu)} \geqslant \frac{\mathrm{Is}(\mu)}{2}. \tag{5.1.6}$$

Απόδειξη. Θέτουμε $\beta = \mathrm{Is}(\mu)$. Από το Θεώρημα 5.1.2 και τον ορισμό της σταθεράς Cheeger, για κάθε θετική ολοκληρώσιμη τοπικά Lipschitz συνάρτηση g έχουμε

$$\beta \int_0^\infty \min\{\mu(\{g \geqslant s\}), 1 - \mu(\{g \geqslant s\})\}\, ds \leqslant \int_0^\infty \mu^+(\{g \geqslant s\})\, ds \leqslant \int_{\mathbb{R}^n} |\nabla g|\, d\mu. \tag{5.1.7}$$

Έστω f μια ολοκληρώσιμη τοπικά Lipschitz συνάρτηση και έστω $m = \mathrm{med}(f)$. Τότε, έχουμε $\mu(\{f \geqslant m\}) \geqslant \frac{1}{2}$ και $\mu(\{f \leqslant m\}) \geqslant \frac{1}{2}$. Θέτουμε $f^+ = \max\{f - m, 0\}$ και $f^- = -\min\{f - m, 0\}$. Τότε, $f - m = f^+ - f^-$ και από τον ορισμό του m έχουμε

$$\mu(\{(f^+)^2 \geqslant s\}) \leqslant \frac{1}{2} \quad \text{και} \quad \mu(\{(f^-)^2 \geqslant s\}) \leqslant \frac{1}{2}$$

για κάθε $s > 0$. Χρησιμοποιώντας την (5.1.7) με $g = (f^+)^2$ και $g = (f^-)^2$ και εφαρμόζοντας ολοκλήρωση κατά μέρη βλέπουμε ότι

$$\begin{aligned}\beta \int_{\mathbb{R}^n} |f - m|^2 d\mu &= \beta \int_{\mathbb{R}^n} (f^+)^2 d\mu + \beta \int_{\mathbb{R}^n} (f^-)^2 d\mu \\ &= \beta \int_0^\infty \mu(\{(f^+)^2 \geqslant s\})\, ds + \beta \int_0^\infty \mu(\{(f^-)^2 \geqslant s\})\, ds \\ &\leqslant \int_{\mathbb{R}^n} |\nabla((f^+)^2)|\, d\mu + \int_{\mathbb{R}^n} |\nabla((f^-)^2)|\, d\mu \\ &= \int_{\mathbb{R}^n} (|\nabla((f^+)^2)| + |\nabla((f^-)^2)|)\, d\mu.\end{aligned}$$

Παρατηρήστε ότι

$$|\nabla((f^+)^2)| + |\nabla((f^-)^2)| \leqslant 2|f - m|\, |\nabla f|.$$

Συνεπώς, εφαρμόζοντας την ανισότητα Cauchy-Schwarz βλέπουμε ότι

$$\beta \int_{\mathbb{R}^n} |f - m|^2 d\mu \leqslant 2 \left(\int_{\mathbb{R}^n} |f - m|^2 d\mu \right)^{1/2} \left(\int_{\mathbb{R}^n} |\nabla f|^2\, d\mu \right)^{1/2},$$

άρα

$$\frac{\beta^2}{4} \int_{\mathbb{R}^n} |f - m|^2 d\mu \leqslant \int_{\mathbb{R}^n} |\nabla f|^2\, d\mu.$$

Αφού

$$\int_{\mathbb{R}^n} |f - \mathbb{E}_\mu(f)|^2 d\mu = \min_{\alpha \in \mathbb{R}} \int_{\mathbb{R}^n} |f - \alpha|^2 d\mu \leqslant \int_{\mathbb{R}^n} |f - m|^2 d\mu,$$

παίρνουμε το θεώρημα. $\square$

5.2 Ανισότητα Poincaré και συγκέντρωση του μέτρου

Έστω (X,d,μ) μετρικός χώρος πιθανότητας. Μια συνάρτηση $f:(X,d)\to\mathbb{R}$ λέγεται τοπικά Lipschitz αν για κάθε $x\in X$ υπάρχει περιοχή U_x του x ώστε η $f\mid_{U_x}$ να είναι Lipschitz. Για κάθε τοπικά Lipschitz συνάρτηση f ορίζουμε

$$|\nabla f|(x)=\limsup_{y\to x}\frac{|f(x)-f(y)|}{d(x,y)}. \tag{5.2.1}$$

Ορισμός 5.2.1. Λέμε ότι το μ ικανοποιεί την ανισότητα Poincaré με σταθερά β αν

$$\mathrm{Var}_\mu(f)\leqslant\frac{1}{\beta}\int|\nabla f|^2d\mu \tag{5.2.2}$$

για κάθε τοπικά Lipschitz συνάρτηση $f:X\to\mathbb{R}$, όπου

$$\mathrm{Var}_\mu(f)=\mathbb{E}\left(f-\mathbb{E}(f)\right)^2=\mathbb{E}(f^2)-(\mathbb{E}(f))^2. \tag{5.2.3}$$

Σκοπός μας σε αυτήν την παράγραφο είναι να δείξουμε ότι κάθε μέτρο μ που ικανοποιεί την ανισότητα Poincaré με σταθερά β έχει εκθετική συγκέντρωση. Παρουσιάζουμε δύο επιχειρήματα: το πρώτο χρησιμοποιεί την έννοια του *συντελεστή επέκτασης* του μ, ενώ το δεύτερο χρησιμοποιεί το συναρτησοειδές Laplace.

Ορισμός 5.2.2 (συντελεστής επέκτασης)**.** Έστω (X,d,μ) μετρικός χώρος πιθανότητας. Ο συντελεστής επέκτασης του μ ορίζεται για κάθε $\varepsilon>0$ ως εξής:

$$\mathrm{Exp}_\mu(\varepsilon)=\sup\{s\geqslant 1:\mu(B_\varepsilon)\geqslant s\mu(B)\ \text{ για κάθε } B\subseteq X \text{ με } \mu(B_\varepsilon)\leqslant 1/2\}. \tag{5.2.4}$$

.

Πρόταση 5.2.3. *Αν για κάποιο $\varepsilon>0$ έχουμε $\mathrm{Exp}_\mu(\varepsilon)\geqslant s>1$, τότε ο (X,d,μ) έχει εκθετική συγκέντρωση:*

$$\alpha_\mu(t)\leqslant\frac{s}{2}\exp\left(-\frac{\ln s}{\varepsilon}t\right).$$

Απόδειξη. Θεωρούμε $A\subseteq X$ με $\mu(A)\geqslant\frac12$ και θέτουμε $B=A_t^c$. Υπάρχει $k\geqslant 0$ ώστε $k\varepsilon\leqslant t<(k+1)\varepsilon$. Τότε, για κάθε $j\leqslant k$ έχουμε $B_{j\varepsilon}\subseteq A^c$, δηλαδή $\mu(B_{j\varepsilon})\leqslant\frac12$.

Από τον ορισμό του συντελεστή επέκτασης και την υπόθεση ότι $\mathrm{Exp}_\mu(\varepsilon)\geqslant s$, έχουμε

$$\begin{aligned}\mu(A_t^c)&\leqslant\mu(A_{k\varepsilon}^c)\leqslant\frac1s\mu(A_{(k-1)\varepsilon}^c)\leqslant\frac{1}{s^2}\mu(A_{(k-2)\varepsilon}^c)\\&\leqslant\cdots\leqslant\frac{1}{s^k}\mu(A^c)\leqslant\frac12 s^{-k}.\end{aligned}$$

Από την $t < (k+1)\varepsilon$ έπεται ότι

$$\mu(A_t^c) \leqslant \frac{s}{2}e^{-(k+1)\ln s} \leqslant \frac{s}{2}e^{-\frac{\ln s}{\varepsilon}t}.$$

Αυτό αποδεικνύει το ζητούμενο. □

Θεώρημα 5.2.4. *Εστω (X,d,μ) μετρικός χώρος πιθανότητας. Αν το μ ικανοποιεί την ανισότητα Poincaré με σταθερά β, τότε*

$$\alpha_\mu(t) \leqslant \exp\left(-\frac{\sqrt{\beta}t}{3}\right).$$

Απόδειξη. Έστω $A \subseteq X$ με $\mu(A) \geqslant 1/2$. Θέτουμε $B = A_t^c$ και ορίζουμε $a = \mu(A)$, $b = \mu(B)$. Παρατηρήστε ότι $\mathrm{dist}(A,B) \geqslant t$.

Ορίζουμε $f : X \to \mathbb{R}$ με

$$f(x) = \frac{1}{a} - \frac{1}{t}\left(\frac{1}{a}+\frac{1}{b}\right)\min\{t, d(x,A)\}.$$

Τότε, $f(x) = 1/a$ στο A, $f(x) = -1/b$ στο B και

$$|\nabla f|(x) \leqslant \frac{1}{t}\left(\frac{1}{a}+\frac{1}{b}\right), \qquad x \notin A\cup B,$$

ενώ $|\nabla f|(x) = 0$ στο $A\cup B$. Συνεπώς,

$$\int |\nabla f|^2 d\mu \leqslant \frac{1}{t^2}\left(\frac{1}{a}+\frac{1}{b}\right)^2 (1-a-b).$$

Από την άλλη πλευρά, αν $m = \mathbb{E}_\mu(f)$ έχουμε

$$\begin{aligned}\mathrm{Var}_\mu(f) &\geqslant \int_A (f-m)^2 d\mu + \int_B (f-m)^2 d\mu\\ &\geqslant a\left(\frac{1}{a}-m\right)^2 + b\left(-\frac{1}{b}-m\right)^2\\ &\geqslant \frac{1}{a}+\frac{1}{b}.\end{aligned}$$

Από την ανισότητα Poincaré έχουμε

$$\left(\frac{1}{a}+\frac{1}{b}\right) \leqslant \frac{1}{\beta t^2}\left(\frac{1}{a}+\frac{1}{b}\right)^2 (1-a-b),$$

απ' όπου παίρνουμε

$$\beta t^2 \leqslant \frac{a+b}{ab}(1-a-b) \leqslant \frac{1-a-b}{ab} = \frac{1-a}{ab} - \frac{1}{a}.$$

Λύνοντας ως προς b παίρνουμε

$$b \leqslant \frac{1-a}{a} \cdot \frac{1}{\frac{1}{a}+\beta t^2} = \frac{1-a}{1+\beta a t^2} \leqslant \frac{1-a}{1+\frac{\beta t^2}{2}}$$

ιότι $a \geqslant \frac{1}{2}$. Επιλέγουμε $\varepsilon = \sqrt{2/\beta}$. Τότε,

$$\mu(B) \leqslant \frac{1}{2}\mu(B_t).$$

Από το γεγονός ότι το A ήταν τυχόν, βλέπουμε ότι $\mathrm{Exp}_\mu(\sqrt{2/\beta}) \geqslant 2$. Τότε, η Πρόταση 5.2.3 μας δίνει

$$\alpha_\mu(t) \leqslant \exp\left(-\frac{\ln 2\beta}{\sqrt{2}}t\right) \leqslant \exp(-\sqrt{\beta}t/\sqrt{3}).$$

Δεύτερη απόδειξη. Θα δείξουμε ότι αν $F : X \to \mathbb{R}$ είναι Lipschitz συνεχής συνάρτηση με $\|F\|_{\mathrm{Lip}} \leqslant 1$ τότε

$$\mu\left(F \geqslant \mathbb{E}_\mu(F)+t\right) \leqslant 3\exp(-t\sqrt{\beta}/2).$$

Το συμπέρασμα προκύπτει από την Πρόταση 3.3.3. Ορίζουμε

$$\Phi(\lambda) = \int e^{\lambda F} d\mu.$$

Εφαρμόζουμε την ανισότητα Poincaré για την $f = e^{\lambda F/2}$, οπότε

$$\nabla(f) = \frac{\lambda}{2}e^{\lambda F/2}\nabla(F).$$

Από την ανισότητα Poincaré,

$$\int e^{\lambda F}\, d\mu - \left(\int e^{\lambda F/2} d\mu\right)^2 \leqslant \frac{\lambda^2}{4\beta}\int e^{\lambda F}|\nabla F|^2 d\mu.$$

Όμως $|\nabla F| \leqslant 1$ σχεδόν παντού, άρα

$$\int e^{\lambda F}\|\nabla F\|^2 d\mu \leqslant \int e^{\lambda F} d\mu.$$

Συνεπώς,

$$\Phi(\lambda) - \Phi^2(\lambda/2) \leqslant \frac{\lambda^2}{4\beta}\Phi(\lambda).$$

Αν $\lambda < 2\sqrt{\beta}$ συμπεραίνουμε ότι

$$\Phi(\lambda) \leqslant \frac{1}{1-\frac{\lambda^2}{4\beta}}\Phi^2(\lambda/2).$$

Επαγωγικά παίρνουμε

$$\Phi(\lambda) \leqslant \prod_{k=0}^{n-1}\left(\frac{1}{1-\frac{\lambda^2}{4^k\beta}}\right)^{2^k}\Phi^{2^n}(\lambda/2^n).$$

Χρησιμοποιώντας το γεγονός ότι

$$\Phi(\lambda) = 1 + \lambda\mathbb{E}_\mu(F) + O(\lambda^2) = 1 + O(\lambda^2)$$

και αφήνοντας το $n \to \infty$ παίρνουμε $\Phi^{2^n}(\lambda/2^n) \to 1$ και

$$\Phi(\lambda) \leqslant \prod_{k=0}^{\infty}\left(\frac{1}{1-\frac{\lambda^2}{4^k\beta}}\right)^{2^k}.$$

Επιλέγοντας τώρα $\lambda = \frac{\sqrt{\beta}}{2}$ έχουμε τελικά

$$\Phi\left(\frac{\sqrt{\beta}}{2}\right) \leqslant 3.$$

Από την ανισότητα του Markov

$$\mu(F \geqslant t) = \mu(e^{F\sqrt{\beta}/2} \geqslant e^{t\sqrt{\beta}/2}) \leqslant 3\exp(-t\sqrt{\beta}/2)$$

έπεται το συμπέρασμα. □

5.3 Ημιομάδες Markov

Ορισμός 5.3.1. Έστω $(X, \mathcal{A}, \mu)$ ένας χώρος πιθανότητας. Μια οικογένεια $\{P_t\}_{t\geqslant 0}$ γραμμικών τελεστών που ορίζονται στην κλάση των φραγμένων μετρήσιμων συναρτήσεων $f : X \to \mathbb{R}$ λέγεται *ημιομάδα Markov* αν ικανοποιεί τα εξής:

(α) Για κάθε φραγμένη μετρήσιμη f και για κάθε $t \geqslant 0$, η P_tf είναι φραγμένη μετρήσιμη συνάρτηση.

(β) Ο P_0 είναι ο ταυτοτικός τελεστής και $P_{t+s} = P_t \circ P_s$ για κάθε $t, s \geqslant 0$.

(γ) $P_t(\mathbf{1}) = \mathbf{1}$ για κάθε $t \geqslant 0$, όπου $\mathbf{1}$ είναι η σταθερή συνάρτηση με τιμή 1, και αν $f \geqslant 0$ τότε $P_t f \geqslant 0$ για κάθε $t \geqslant 0$.

Το μέτρο πιθανότητας μ λέγεται *αναλλοίωτο* για την $\{P_t\}_{t\geqslant 0}$ αν

$$\int P_t f \, d\mu = \int f \, d\mu \tag{5.3.1}$$

για κάθε $t \geqslant 0$ και κάθε φραγμένη μετρήσιμη $f : X \to \mathbb{R}$.

Χρησιμοποιώντας την ιδιότητα (γ) βλέπουμε ότι για κάθε κυρτή συνάρτηση $\phi : \mathbb{R} \to \mathbb{R}$ ισχύει

$$P_t(\phi(f)) \geqslant \phi(P_t(f))$$

για κάθε f και t. Ειδικότερα, $P_t(|f|^p) \geqslant |P_t f|^p$ για κάθε $1 \leqslant p \leqslant \infty$. Υποθέτοντας ότι το μ είναι αναλλοίωτο βλέπουμε ότι

$$\|P_t f\|_p \leqslant \|f\|_p \tag{5.3.2}$$

για κάθε φραγμένη μετρήσιμη f, άρα μπορούμε να επεκτείνουμε τον P_t στον $L_p(\mu)$ έτσι ώστε η (5.3.2) να εξακολουθεί να ισχύει.

Στον ορισμό της ημιομάδας Markov συμπεριλαμβάνουμε μια ιδιότητα συνέχειας: για κάθε $f \in L_2(\mu)$ απαιτούμε να ισχύει

$$\lim_{t\to 0} \|P_t f - f\|_2 \to 0. \tag{5.3.3}$$

Παρατήρηση 5.3.2. Μια πρώτη απλή συνέπεια αυτών των ιδιοτήτων είναι ότι η συνάρτηση $t \mapsto \mathrm{Var}_\mu(P_t f)$ είναι φθίνουσα. Για να το δούμε, για τυχόντες $t > s \geqslant 0$ γράφουμε

$$\begin{aligned}\mathrm{Var}_\mu(P_t f) &= \|P_t f - \mathbb{E}_\mu(f)\|_2^2 = \|P_t(f - \mathbb{E}_\mu(f))\|_2^2 = \|P_{t-s}(P_s(f - \mathbb{E}_\mu(f)))\|_2^2 \\ &\leqslant \|P_s(f - \mathbb{E}_\mu(f))\|_2^2 = \|P_s f - \mathbb{E}_\mu(f)\|_2^2 = \mathrm{Var}_\mu(P_s f).\end{aligned}$$

Ορισμός 5.3.3. Ο γεννήτορας $\mathcal{L}$ της $\{P_t\}_{t\geqslant 0}$ ορίζεται από την

$$\mathcal{L}(f) = \lim_{t\to 0^+} \frac{P_t f - f}{t},$$

για όλες τις $f \in L^2(\mu)$ για τις οποίες το όριο αυτό υπάρχει στον $L^2(\mu)$. Η κλάση $D(\mathcal{L})$ όλων αυτών των συναρτήσεων είναι το *πεδίο ορισμού* του $\mathcal{L}$, και ο $\mathcal{L} : D(\mathcal{L}) \to L^2(\mu)$ είναι γραμμικός τελεστής. Αποδεικνύεται ότι ο $D(\mathcal{L})$ είναι πυκνός υπόχωρος του $L^2(\mu)$.

Παρατηρήστε ότι αν $f \in D(\mathcal{L})$ τότε

$$\frac{d}{dt}(P_t f) = P_t \mathcal{L}(f) = \mathcal{L} P_t(f). \tag{5.3.4}$$

Για να το δούμε, πρώτα γράφουμε

$$\frac{d}{dt}(P_t f) = \lim_{h\to 0^+} \frac{P_{t+h}f - P_t f}{h} = \lim_{h\to 0^+} P_t\Big(\frac{P_h f - f}{h}\Big) = P_t\mathcal{L}(f),$$

και στη συνέχεια υπολογίζουμε την παράγωγο ως εξής:

$$\frac{d}{dt}(P_t f) = \lim_{h\to 0^+} \frac{P_{t+h}f - P_t f}{h} = \lim_{h\to 0^+} \frac{P_h(P_t f) - P_t f}{h} = \mathcal{L}P_t(f).$$

Αντίστροφα, ο γεννήτορας $\mathcal{L}$ και το πεδίο ορισμού $D(\mathcal{L})$ προσδιορίζουν την ημιομάδα με την έννοια ότι υπάρχει μοναδική ημιομάδα $\{P_t\}_{t\geqslant 0}$ τελεστών στον $L_2(\mu)$ που ικανοποιεί την (5.3.4). Μπορούμε να εκφράσουμε τυπικά τη σχέση ανάμεσα στο γεννήτορα και την ημιομάδα μέσω της σχέσης $P_t = e^{t\mathcal{L}}$. Αυτή η ισότητα γίνεται ακριβής αν ερμηνεύσουμε τον $e^{t\mathcal{L}}$ σαν δυναμοσειρά στην περίπτωση που $D(\mathcal{L}) = L_2(\mu)$.

Ορισμός 5.3.4. Η ημιομάδα $\{P_t\}_{t\geqslant 0}$ λέγεται *συμμετρική* ή *αντιστρέψιμη* ως προς το μ αν

$$\langle f, P_t g\rangle_\mu = \langle P_t f, g\rangle_\mu$$

για κάθε $f, g \in L^2(\mu)$ και $t \geqslant 0$, όπου

$$\langle f, g\rangle_\mu := \int_X fg\, d\mu.$$

Η αντιστρεψιμότητα είναι λοιπόν ισοδύναμη με το γεγονός ότι κάθε P_t είναι αυτοσυζυγής τελεστής. Παίρνοντας υπόψη την $P_t = e^{t\mathcal{L}}$ βλέπουμε ότι αυτό είναι ισοδύναμο με το να είναι ο $\mathcal{L}$ αυτοσυζυγής.

Λέμε ότι η ημιομάδα $\{P_t\}_{t\geqslant 0}$ είναι *εργοδική* αν

$$\lim_{t\to\infty} \|P_t f - \mathbb{E}_\mu(f)\|_2 = 0$$

για κάθε $f \in L^2(\mu)$.

Ορισμός 5.3.5. Έστω $\{P_t\}_{t\geqslant 0}$ μια ημιομάδα Markov με γεννήτορα $\mathcal{L}$ και αναλλοίωτο μέτρο μ. Η *μορφή Dirichlet* ή *ενέργεια* της ημιομάδας ορίζεται από την

$$\mathcal{E}(f, g) = -\langle f, \mathcal{L}g\rangle.$$

Η μορφή Dirichlet $\mathcal{E}(f, f)$ παίζει το ρόλο της μέσης τιμής του τετραγώνου της κλίσης της f στην ακόλουθη αφηρημένη ανισότητα Poincaré.

Θεώρημα 5.3.6 (ανισότητα Poincaré). *Έστω $\{P_t\}_{t\geqslant 0}$ μια αντιστρέψιμη ημιομάδα Markov με αναλλοίωτο μέτρο μ και έστω $\beta > 0$. Τα ακόλουθα είναι ισοδύναμα:*

(α) *Ισχύει η ανισότητα Poincaré: για κάθε f,*

$$\mathrm{Var}_\mu(f) \leqslant \frac{1}{\beta}\,\mathcal{E}(f,f).$$

(β) *Για κάθε f και $t \geqslant 0$*

$$\|P_t f - \mathbb{E}_\mu(f)\|_2 \leqslant e^{-\beta t}\|f - \mathbb{E}_\mu(f)\|_2.$$

(γ) *Για κάθε f και $t \geqslant 0$*

$$\mathcal{E}(P_t f, P_t f) \leqslant e^{-2\beta t}\mathcal{E}(f,f).$$

(δ) *Για κάθε f υπάρχει σταθερά $c_f > 0$ τέτοια ώστε, για κάθε $t \geqslant 0$,*

$$\|P_t f - \mathbb{E}_\mu(f)\|_2 \leqslant c_f \cdot e^{-\beta t}.$$

(ε) *Για κάθε f υπάρχει σταθερά $c_f > 0$ τέτοια ώστε, για κάθε $t \geqslant 0$,*

$$\mathcal{E}(P_t f, P_t f) \leqslant c_f \cdot e^{-2\beta t}.$$

Το επόμενο λήμμα μας δίνει τη σύνδεση ανάμεσα στην $\mathrm{Var}_\mu(f)$ και την $\mathcal{E}(f,f)$.

Λήμμα 5.3.7. *Έστω $\{P_t\}_{t\geqslant 0}$ μια αντιστρέψιμη ημιομάδα Markov με αναλλοίωτο μέτρο μ. Για κάθε f και $t \geqslant 0$ έχουμε*

$$\frac{d}{dt}\big(\mathrm{Var}_\mu(P_t f)\big) = -2\mathcal{E}(P_t f, P_t f).$$

Απόδειξη. Χρησιμοποιώντας το γεγονός ότι το μ είναι αναλλοίωτο γράφουμε

$$\begin{aligned}\frac{d}{dt}\big(\mathrm{Var}_\mu(P_t f)\big) &= \frac{d}{dt}\big(\mathbb{E}_\mu((P_t f)^2) - (\mathbb{E}_\mu(P_t f))^2\big) = \frac{d}{dt}\big(\mathbb{E}_\mu((P_t f)^2) - (\mathbb{E}_\mu(f))^2\big)\\ &= \frac{d}{dt}\big(\mathbb{E}_\mu((P_t f)^2)\big) = \mathbb{E}\Big(2P_t f \cdot \frac{d}{dt}(P_t f)\Big)\\ &= \mathbb{E}\left(2P_t f \cdot \mathcal{L}P_t f\right),\end{aligned}$$

και το λήμμα έπεται από τον ορισμό της μορφής Dirichlet. □

Το Λήμμα 5.3.7 έχει δύο άμεσες συνέπειες. Πρώτα απ΄ όλα, από την Παρατήρηση 5.3.2 γνωρίζουμε ότι η απεικόνιση $t \mapsto \mathrm{Var}_\mu(P_t f)$ είναι φθίνουσα. Τότε, από το Λήμμα 5.3.7 βλέπουμε αμέσως ότι η ενέργεια είναι μη αρνητική.

Πρόταση 5.3.8. *Για κάθε f ισχύει $\mathcal{E}(f,f) \geqslant 0$.*

Έπειτα, μπορούμε να αποδείξουμε μια ολοκληρωτική αναπαράσταση της $\mathrm{Var}_\mu(f)$ συναρτήσει της ενέργειας. Αν υποθέσουμε ότι η ημιομάδα είναι εργοδική, έχουμε $P_t f \to \mathbb{E}_\mu(f)$ καθώς το $t \to \infty$, άρα

$$\mathrm{Var}_\mu(P_t f) \to \mathrm{Var}_\mu(\mathbb{E}_\mu(f)) = 0.$$

Μπορούμε λοιπόν να γράψουμε

$$\mathrm{Var}_\mu(f) = \mathrm{Var}_\mu(P_0 f) - \lim_{t\to\infty} \mathrm{Var}_\mu(P_t f) = -\int_0^\infty \frac{d}{dt}\big(\mathrm{Var}_\mu(P_t f)\big)\, dt.$$

Αυτό αποδεικνύει την επόμενη πρόταση.

Πρόταση 5.3.9. *Έστω $\{P_t\}_{t\geqslant 0}$ μια εργοδική ημιομάδα Markov. Για κάθε f έχουμε*

$$\mathrm{Var}_\mu(f) = 2\int_0^\infty \mathcal{E}(P_t f, P_t f)\, dt.$$

Το επόμενο λήμμα δείχνει ότι αν υποθέσουμε αντιστρεψιμότητα για την ημιομάδα τότε έχουμε ισχυρότερη πληροφορία για τη συμπεριφορά της διασποράς συναρτήσει του χρόνου.

Λήμμα 5.3.10. *Έστω $\{P_t\}_{t\geqslant 0}$ μια αντιστρέψιμη ημιομάδα Markov. Τότε, οι συναρτήσεις $t \mapsto \ln \|P_t f\|_2^2$ και $t \mapsto \ln \mathcal{E}(P_t f, P_t f)$ είναι κυρτές.*

Απόδειξη. Χρησιμοποιώντας το γεγονός ότι ο $\mathcal{L}$ είναι αυτοσυζυγής, γράφουμε

$$\begin{aligned}
\frac{d}{dt}\mathcal{E}(P_t f, P_t f) &= -\frac{d}{dt}\langle P_t f, \mathcal{L}P_t f\rangle_\mu \\
&= -\langle \mathcal{L}P_t f, \mathcal{L}P_t f\rangle_\mu - \langle P_t f, \mathcal{L}^2 P_t f\rangle_\mu \\
&= -2\|\mathcal{L}P_t f\|_2^2.
\end{aligned}$$

Εισάγοντας αυτή την ταυτότητα στον επόμενο υπολογισμό παίρνουμε

$$\begin{aligned}
\frac{d^2}{dt^2}\ln \|P_t f\|_2^2 &= \frac{4\|\mathcal{L}P_t f\|_2^2}{\|P_t f\|_2^2} - \frac{4\mathcal{E}(P_t f, P_t f)^2}{\|P_t f\|_2^4} \\
&= \frac{4}{\|P_t f\|_2^4}\Big(\|\mathcal{L}P_t f\|_2^2 \|P_t f\|_2^2 - \langle P_t f, \mathcal{L}P_t f\rangle_\mu^2\Big),
\end{aligned}$$

και αυτή είναι μια μη αρνητική ποσότητα από την ανισότητα Cauchy-Schwarz. Έπεται ότι η $t \mapsto \ln \|P_t f\|_2^2$ είναι κυρτή συνάρτηση.

Για τον δεύτερο ισχυρισμό του λήμματος παρατηρούμε αρχικά ότι η μορφή Dirichlet ικανοποιεί την ανισότητα Cauchy-Schwarz $\mathcal{E}(f,g)^2 \leqslant \mathcal{E}(f,f)\mathcal{E}(g,g)$. Αυτό προκύπτει από το γεγονός ότι, από την Πρόταση 5.3.8, ισχύει $\mathcal{E}(f+tg, f+tg) \geqslant 0$ για κάθε $t \in \mathbb{R}$. Κατόπιν, υπολογίζουμε την δεύτερη παράγωγο της $\ln \mathcal{E}(P_t f, P_t f)$ όπως κάναμε και στην απόδειξη του πρώτου ισχυρισμού του λήμματος. □

Μπορούμε τώρα να αποδείξουμε το Θεώρημα 5.3.6. Παρατηρήστε ότι η ισοδυναμία των (α) και (γ), αλλά και οι συνεπαγωγές (β) $\Longrightarrow$ (δ), (γ) $\Longrightarrow$ (α) και (γ) $\Longleftarrow$ (δ) δεν απαιτούν την αντιστρεψιμότητα της ημιομάδας.

Απόδειξη του Θεωρήματος 5.3.6. (α) $\Longrightarrow$ (β). Η υπόθεση και το Λήμμα 5.3.7 δείχνουν ότι

$$\frac{d}{dt}\big(\mathrm{Var}_\mu(P_t f)\big) \leqslant -2\beta \mathrm{Var}_\mu(f).$$

Έπεται ότι

$$\|P_t f - \mathbb{E}_\mu(f)\|_2^2 = \mathrm{Var}_\mu(P_t f) \leqslant e^{-2\beta t}\mathrm{Var}_\mu(f) = e^{-2\beta t}\|f - \mathbb{E}_\mu(f)\|_2^2.$$

(β) $\Longrightarrow$ (γ). Από το Λήμμα 5.3.10 η συνάρτηση

$$t \mapsto \frac{d}{dt}\ln \|P_t f\|_2^2 = -\frac{2\mathcal{E}(P_t f, P_t f)}{\|P_t f\|_2^2}$$

είναι αύξουσα. Ειδικότερα, για κάθε $t \geqslant 0$ έχουμε

$$-\frac{2\mathcal{E}(P_t f, P_t f)}{\|P_t f\|_2^2} \geqslant -\frac{2\mathcal{E}(f,f)}{\|f\|_2^2},$$

το οποίο μας δίνει

$$\mathcal{E}(P_t f, P_t f) \leqslant \frac{\|P_t f\|_2^2}{\|f\|_2^2}\,\mathcal{E}(f,f).$$

Τώρα, η (γ) είναι άμεση συνέπεια της υπόθεσης (β).

Είναι φανερό ότι (β) $\Longrightarrow$ (δ) και (γ) $\Longrightarrow$ (δ). Μπορούμε να δείξουμε τις αντίστροφες συνεπαγωγές εφαρμόζοντας το επόμενο λήμμα στις κυρτές συναρτήσεις $t \mapsto \ln \|P_t f\|_2^2$ και $t \mapsto \ln \mathcal{E}(P_t f, P_t f)$.

Λήμμα 5.3.11. *Έστω $g : [0,\infty) \to \mathbb{R}$ μια κυρτή συνάρτηση που ικανοποιεί την $g(t) \leqslant M - \delta t$ για κάποιους M, δ και για κάθε $t \geqslant 0$. Τότε, $g(t) \leqslant g(0) - \delta t$ για κάθε $t \geqslant 0$.*

Απόδειξη. Αρκεί να δείξουμε ότι $g'(t) \leqslant -\delta$ για κάθε $t \geqslant 0$. Ας υποθέσουμε ότι για κάποιον $s \geqslant 0$ ισχύει $g'(s) = -r > -\delta$. Αφού η g' είναι αύξουσα πρέπει να ισχύει $g'(t) \geqslant -r$ για κάθε $t \geqslant s$, άρα $g(t) \geqslant g(s) - r(t-s)$ για κάθε $t \geqslant s$. Τότε,

$$M - rs - g(s) \geqslant (\delta - r)t$$

και αφού $\delta - r > 0$ καταλήγουμε σε άτοπο αφήνοντας το $t \to \infty$. □

Μένει να δείξουμε τη συνεπαγωγή (γ) $\Longrightarrow$ (α). Αυτή προκύπτει από την ολοκληρωτική αναπαράσταση της διασποράς που μας δίνει η Πρόταση 5.3.9. Χρησιμοποιώντας και την (γ) βλέπουμε αμέσως ότι

$$\mathrm{Var}_\mu(f) \leqslant 2\mathcal{E}(f,f)\int_0^\infty e^{-2\beta t}dt = \frac{1}{\beta}\mathcal{E}(f,f).$$

Αυτό ολοκληρώνει την απόδειξη του θεωρήματος. □

Υπάρχουν πολλά σημαντικά παραδείγματα μέτρων πιθανότητας μ για τα οποία θα θέλαμε να αποδείξουμε την ανισότητα Poincaré. Με βάση το Θεώρημα 5.3.6 θα μπορούσαμε να προσπαθήσουμε να ορίσουμε κατάλληλη ημιομάδα Markov στον $L_2(\mu)$ για την οποία το μ να είναι αναλλοίωτο μέτρο. Αυτό ενδέχεται να γίνεται με περισσότερους από έναν τρόπους, με αποτέλεσμα διαφορετικές μορφές Dirichlet άρα και διαφορετικές ανισότητες Poincaré με διαφορετικές έννοιες κλίσης.

Στην επόμενη παράγραφο θα εφαρμόσουμε αυτή τη μέθοδο στο τυπικό n-διάστατο μέτρο του Gauss γ_n.

5.4 Ημιομάδα Ornstein–Uhlenbeck

Ορισμός 5.4.1. Η ημιομάδα Ornstein–Uhlenbeck ορίζεται στον $L^p(\gamma_n)$ ως εξής. Για κάθε $f \in L^p(\gamma_n)$ και για κάθε $t \geqslant 0$ ορίζουμε

$$(T_t f)(x) = \int_{\mathbb{R}^n} f(e^{-t}x + \sqrt{1-e^{-2t}}y)\, d\gamma_n(y).$$

Η $T_t f$ είναι καλά ορισμένη: παρατηρούμε πρώτα ότι το γ_n είναι η εικόνα του $\gamma_n \otimes \gamma_n$ μέσω της

$$(x,y) \mapsto e^{-t}x + \sqrt{1-e^{-2t}}y.$$

Συνεπώς, αν $f \in L^1(\gamma_n)$ έχουμε

$$\int_{\mathbb{R}^n} f(z)d\gamma_n(z) = \iint_{\mathbb{R}^n\times\mathbb{R}^n} f(e^{-t}x + \sqrt{1-e^{-2t}}y)\, d\gamma_n(y)d\gamma_n(x).$$

Χρησιμοποιώντας το θεώρημα Fubini βλέπουμε ότι: αφού

$$\int_{\mathbb{R}^n} |f(x)|^p d\gamma_n(x) = \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}\left|f\left(e^{-t}x + \sqrt{1-e^{-2t}}y\right)\right|^p d\gamma_n(x)\right) d\gamma_n(y),$$

η συνάρτηση $T_t f$ ανήκει στον $L^p(\gamma_n)$ και, από την ανισότητα του Hölder,

$$\|T_t f\|_{L^p(\gamma_n)} \leqslant \|f\|_{L^p(\gamma_n)}.$$

Δουλεύουμε στον χώρο $W^{2,1}(\gamma_n)$, ο οποίος είναι η πλήρωση του $C_0(\mathbb{R}^n)$ ως προς τη νόρμα Sobolev
$$\|f\|_{2,1} = \left(\int_{\mathbb{R}^n} |f(x)|^2 d\gamma_n(x) + \int_{\mathbb{R}^n} |\nabla f(x)|^2 d\gamma_n(x)\right)^{1/2}.$$
Βασικές ιδιότητες, οι οποίες ελέγχονται άμεσα από τον ορισμό, είναι οι εξής:

(i) Για κάθε f,
$$T_0 f = f, \qquad T_{t+s} f = T_t(T_s f), \qquad T_\infty f = \lim_{t\to\infty} T_t f \equiv \int f\, d\gamma_n.$$

(ii) Για κάθε $t \geqslant 0$,
$$\int_{\mathbb{R}^n} T_t f\, d\gamma_n = \int_{\mathbb{R}^n} f\, d\gamma_n.$$

(iii) Για κάθε $t \geqslant 0$,
$$[\,T_t(fg)\,]^2 \leqslant T_t(f^2) \cdot T_t(g^2).$$

(iv) Αν $f \leqslant g$ τότε $T_t f \leqslant T_t g$.

(v) Για κάθε f, g και $a, b \in \mathbb{R}$, $T_t(af + bg) = aT_t(f) + bT_t(g)$. Επίσης, $T_t(1) \equiv 1$.

(vi) Αν ορίσουμε $T_t(g_1, \ldots, g_n) = (T_t(g_1), \ldots, T_t(g_n))$ τότε
$$\nabla T_t(f) = e^{-t} T_t(\nabla f).$$

Ορισμός 5.4.2. Ο γεννήτορας L της ημιομάδας ορίζεται στον $W^{2,2}(\gamma_n)$ από την
$$(Lf)(x) = \lim_{t\to 0} \frac{T_t f(x) - f(x)}{t}$$
και ικανοποιεί τις
$$\frac{d}{dt}(T_t f) = LT_t f = T_t L f$$
και
$$(Lf)(x) = \Delta f(x) - \langle x, \nabla f(x)\rangle.$$
Για την απόδειξη γράφουμε
$$\begin{aligned} L(T_t f) &= \frac{d}{dt}\left(\int_{\mathbb{R}^n} f(e^{-t}x + \sqrt{1-e^{-2t}}y)\, d\gamma_n(y)\right) \\ &= -\int_{\mathbb{R}^n} \langle \nabla f(e^{-t}x + \sqrt{1-e^{-2t}}y), e^{-t}x\rangle\, d\gamma_n(y) \\ &\quad + \int_{\mathbb{R}^n} \left\langle \nabla f(e^{-t}x + \sqrt{1-e^{-2t}}y), \frac{e^{-2t}}{\sqrt{1-e^{-2t}}}y \right\rangle d\gamma_n(y), \end{aligned}$$

και παίρνουμε $t \to 0^+$. Ο πρώτος όρος τείνει στο

$$-\int_{\mathbb{R}^n} \langle \nabla f(x), x\rangle \, d\gamma_n(y) = -\langle \nabla f(x), x\rangle,$$

ενώ ο δεύτερος γράφεται στη μορφή

$$-\int_{\mathbb{R}^n} \langle \nabla g_t(y), \nabla h(y)\rangle \, dy,$$

όπου

$$h(y) = \frac{1}{(2\pi)^{n/2}} e^{-\|y\|_2^2/2} \quad \text{και} \quad g_t(y) = f(e^{-t}x + \sqrt{1-e^{-2t}}y),$$

άρα ισούται με

$$\int_{\mathbb{R}^n} \Delta g_t(y) \, \frac{1}{(2\pi)^{n/2}} e^{-\|y\|_2^2/2} \, dy \to \int_{\mathbb{R}^n} \Delta f(x) \, d\gamma_n(y) = \Delta f(x)$$

καθώς το $t \to 0^+$. Έπεται ότι

$$(Lf)(x) = \lim_{t\to 0} \big(L(T_t f)\big)(x) = \Delta f(x) - \langle \nabla f(x), x\rangle.$$

Μια άλλη ιδιότητα του L, η οποία προκύπτει από την προηγούμενη με εφαρμογή του τύπου του Green, είναι η εξής: για κάθε $f \in W^{2,2}(\gamma_n)$ και $g \in W^{2,1}(\gamma_n)$,

$$\int_{\mathbb{R}^n} Lf \cdot g \, d\gamma_n = -\int_{\mathbb{R}^n} \langle \nabla f, \nabla g\rangle \, d\gamma_n.$$

Θεώρημα 5.4.3 (ανισότητα Poincaré)**.** *Για κάθε $f \in W^{2,1}(\gamma_n)$ ισχύει η ανισότητα*

$$\int_{\mathbb{R}^n} f^2 \, d\gamma_n - \left(\int_{\mathbb{R}^n} f \, d\gamma_n\right)^2 \leqslant \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n.$$

Απόδειξη. Γράφουμε

$$\begin{aligned}
\int_{\mathbb{R}^n} f^2 \, d\gamma_n - \left(\int_{\mathbb{R}^n} f \, d\gamma_n\right)^2 &= \int_{\mathbb{R}^n} (T_0 f)^2 \, d\gamma_n - (T_\infty f)^2 \\
&= \int_{\mathbb{R}^n} (T_0 f)^2 \, d\gamma_n - \int_{\mathbb{R}^n} (T_\infty f)^2 \, d\gamma_n \\
&= -\int_0^\infty \frac{d}{dt} \left(\int_{\mathbb{R}^n} (T_t f)^2\right) dt \\
&= -\int_0^\infty \int_{\mathbb{R}^n} 2T_t f \cdot LT_t f \, d\gamma_n \, dt \\
&= 2\int_0^\infty \int_{\mathbb{R}^n} \langle \nabla T_t f, \nabla T_t f\rangle \, d\gamma_n \, dt \\
&= 2\int_0^\infty \int_{\mathbb{R}^n} |\nabla T_t f|^2 d\gamma_n \, dt.
\end{aligned}$$

Τώρα,

$$
\begin{aligned}
2\int_0^\infty \int_{\mathbb{R}^n} |\nabla T_t f|^2 d\gamma_n\, dt &= 2\int_0^\infty e^{-2t} \int_{\mathbb{R}^n} |T_t(\nabla f)|^2 d\gamma_n\, dt \\
&= 2\int_0^\infty e^{-2t} \int_{\mathbb{R}^n} \sum_{i=1}^n \big(T_t(\partial_{x_i} f)\big)^2 \\
&\leqslant 2\int_0^\infty e^{-2t} \int_{\mathbb{R}^n} \sum_{i=1}^n T_t(1^2) T_t((\partial_{x_i} f)^2) \\
&= 2\int_0^\infty e^{-2t} \int_{\mathbb{R}^n} T_t(|\nabla f|^2)\, d\gamma_n\, dt \\
&= 2\int_0^\infty e^{-2t} \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n\, dt \\
&= 2\int_0^\infty e^{-2t} dt \cdot \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n \\
&= \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n.
\end{aligned}
$$

Έχουμε έτσι ευθεία απόδειξη της ανισότητας του Poincaré μέσω της ημιόμάδας Ornstein-Uhlenbeck. □

5.5 Ανισότητα Poincaré και ιδιοτιμές του τελεστή Laplace

Στο πλαίσιο μιας συμπαγούς πολλαπλότητας Riemann (X, g, μ), όπου g είναι η γεωδαισιακή μετρική και μ είναι ο κανονικοποιημένος όγκος, η ανισότητα Poincaré συνδέεται στενά με τις ιδιοτιμές του τελεστή Laplace-Beltrami

$$\Delta(f) = \mathrm{div}(\nabla f).$$

Αποδεικνύεται ότι οι ιδιοτιμές του $-\Delta$ είναι μη αρνητικές και σχηματίζουν ένα διακριτό σύνολο, οπότε μπορούμε να τις διατάξουμε ως εξής: $0 < \lambda_1 \leqslant \lambda_2 \leqslant \cdots$ (υπολογίζοντας και την πολλαπλότητα των θετικών ιδιοτιμών). Αποδεικνύεται επίσης ότι η μικρότερη θετική ιδιοτιμή λ_1 έχει πολλαπλότητα 1.

Θεώρημα 5.5.1. *Έστω (X, g, μ) μια συμπαγής πολλαπλότητα Riemann. Τότε, το μ ικανοποιεί την ανισότητα Poincaré με σταθερά $\beta = \lambda_1$.*

Απόδειξη. Θεωρούμε των χώρο των C^2-συναρτήσεων $f : X \to \mathbb{R}$ σαν υπόχωρο του $L^2(X)$ με εσωτερικό γινόμενο το $\langle f, g\rangle = \int fg\, d\mu$. Από τον τύπο του Green έχουμε

$$\int (\Delta f) g\, d\mu = -\int \langle \nabla f, \nabla g\rangle\, d\mu$$

για κάθε $f, g \in C^2(X)$. Ο τελεστής $-\Delta$ είναι θετικός και αυτοσυζυγής, και υπάρχει ορθοκανονική βάση $\{f_j\}$ του $L^2(X)$ η οποία αποτελείται από ιδιοσυναρτήσεις που αντιστοιχούν στις ιδιοτιμές λ_j.

Για κάθε $f \in L^2(X)$ έχουμε

$$f = \sum_{j=1}^{\infty} \langle f, f_j \rangle f_j$$

και

$$\|f\|_2^2 = \sum_{j=1}^{\infty} \langle f, f_j \rangle^2$$

από την ταυτότητα του Parseval. Η *ενέργεια*

$$\mathcal{E}(f,g) = -\int \langle \nabla f, \nabla g \rangle \, d\mu$$

είναι διγραμμική και συμμετρική. Συνεπώς, για κάθε $n \geqslant 1$ έχουμε

$$\begin{aligned} 0 &\leqslant \mathcal{E}\left(f - \sum_{j=1}^{n} \langle f, f_j \rangle f_j, f - \sum_{j=1}^{n} \langle f, f_j \rangle f_j \right) \\ &= \mathcal{E}(f,f) - 2\sum_{j=1}^{n} \langle f, f_j \rangle \mathcal{E}(f, f_j) + \sum_{j,k=1}^{n} \langle f, f_j \rangle \langle f, f_k \rangle \mathcal{E}(f_j, f_k) \\ &= \mathcal{E}(f,f) - 2\sum_{j=1}^{n} \langle f, f_j \rangle \langle f, \Delta f_j \rangle + \sum_{j,k=1}^{n} \langle f, f_j \rangle \langle f, f_k \rangle \langle f_j, \Delta f_k \rangle \\ &= \mathcal{E}(f,f) - 2\sum_{j=1}^{n} \langle f, f_j \rangle^2 \lambda_j + \sum_{j=1}^{n} \langle f, f_j \rangle^2 \lambda_j. \end{aligned}$$

Με άλλα λόγια,

$$\sum_{j=1}^{n} \langle f, f_j \rangle^2 \lambda_j \leqslant \mathcal{E}(f,f)$$

για κάθε n, απ΄ όπου έπεται ότι

$$\lambda_1 \|f\|_2^2 = \lambda_1 \sum_{j=1}^{\infty} \langle f, f_j \rangle^2 \leqslant \sum_{j=1}^{\infty} \lambda_j \langle f, f_j \rangle^2 \leqslant \mathcal{E}(f,f).$$

Παρατηρώντας ότι αν αντικαταστήσουμε την f με την $f - \mathbb{E}_\mu(f)$ τότε η ενέργεια $\mathcal{E}(f,f)$ δεν μεταβάλλεται, συμπεραίνουμε ότι

$$\mathrm{Var}_\mu(f) \leqslant \frac{1}{\lambda_1} \mathcal{E}(f,f) = \frac{1}{\lambda_1} \int |\nabla f|^2 d\mu.$$

Δηλαδή, ισχύει η ανισότητα Poincaré με σταθερά $\beta = \lambda_1$. □

5.6 Ανισότητα Poincaré στο διακριτό κύβο

Θεωρούμε το διακριτό κύβο E_2^n με την μετρική $d_n(\varepsilon,\zeta)=\frac{1}{2n}\sum_{j=1}^n|\varepsilon_i-\zeta_i|$ και το ομοιόμορφο μέτρο πιθανότητας μ_n. Οι συναρτήσεις Rademacher r_i, $i=1,\dots,n$, ορίζονται από τις

$$r_i(\varepsilon)=\varepsilon_i.$$

Οι συναρτήσεις Walsh ορίζονται ως εξής: για κάθε $\emptyset\neq A\subseteq\{1,\dots,n\}$ θέτουμε

$$w_A(\varepsilon)=\prod_{i\in A}r_i(\varepsilon),$$

και στην περίπτωση $A=\emptyset$ θέτουμε $w_\emptyset\equiv 1$. Παρατηρήστε ότι $w_{\{i\}}=r_i$.

Οι συναρτήσεις Walsh σχηματίζουν ορθοκανονική βάση στον $L^2(E_2^n)$ (παρατηρήστε ότι είναι ορθοκανονικό σύνολο με πληθάριθμο $2^n=\dim(L^2(E_2^n))$). Συνεπώς, κάθε $f:E_2^n\to\mathbb{R}$ γράφεται στην μορφή

$$f=\sum_A\widehat{f}_Aw_A$$

όπου

$$\widehat{f}=\langle f,w_A\rangle=\int_{E_2^n}f(\varepsilon)w_A(\varepsilon)\,d\mu_n(\varepsilon),$$

και ισχύει η ταυτότητα Parseval

$$\|f\|_2^2=\langle f,f\rangle=\sum_A\widehat{f}_A^2.$$

Για κάθε $\varepsilon=(\varepsilon_1,\dots,\varepsilon_n)\in E_2^n$, οι «γείτονες» του ε είναι εκείνα τα $\zeta=(\zeta_1,\dots,\zeta_n)\in E_2^n$ για τα οποία

$$|\{i\leqslant n:\varepsilon_i\neq\zeta_i\}|=1.$$

Αν τα ε,ζ είναι γείτονες, γράφουμε $\varepsilon\sim\zeta$. Το gradient της f είναι το διάνυσμα

$$(\nabla f)(\varepsilon)=\frac{1}{2}\left(f(\zeta)-f(\varepsilon)\right)_{\zeta\sim\varepsilon}.$$

Θεωρούμε την *διακριτή Λαπλασιανή* $L(f)$ της f, η οποία ορίζεται από την

$$L(f)(\varepsilon)=\frac{1}{2}\sum_{\zeta\sim\varepsilon}[f(\zeta)-f(\varepsilon)].$$

Παρατηρούμε ότι: αν $\zeta\sim\varepsilon$ και $\zeta_i\neq\varepsilon_i$ τότε $w_A(\zeta)=w_A(\varepsilon)$ αν $i\notin A$ και $w_A(\zeta)=-w_A(\varepsilon)$ αν $i\in A$. Έπεται ότι

$$L(w_A)=-|A|w_A,$$

δηλαδή οι συναρτήσεις Walsh είναι ιδιοσυναρτήσεις της διακριτής Λαπλασιανής. Αντικαθιστώντας, παίρνουμε

$$-L(f) = \sum_{i=1}^{m} \widehat{f}_i \varepsilon_i + \sum_{|A|\geqslant 2} |A| \widehat{f}_A w_A.$$

Θεωρούμε επίσης την *ενέργεια* $\mathcal{E}(f)$ της f, η οποία ορίζεται από την

$$\mathcal{E}(f) = -\langle f, L(f)\rangle.$$

Θεώρημα 5.6.1. *Για κάθε $f : E_2^n \to \mathbb{R}$ ισχύει*

$$\mathrm{Var}_{\mu_n}(f) \leqslant \mathcal{E}(f) = \int |\nabla f|^2 d\mu_n.$$

Απόδειξη. Μπορούμε να υποθέσουμε ότι $\mathbb{E}_{\mu_n}(f) = 0$. Θεωρώντας το ανάπτυγμα της f ως προς τις συναρτήσεις Walsh βλέπουμε ότι

$$\mathcal{E}(f) = \sum_{i=1}^{n} \widehat{f}_i^2 + \sum_{|A|\geqslant 2} |A| \widehat{f}_A^2.$$

Επίσης,

$$\begin{aligned}
\mathcal{E}(f) &= -\frac{1}{2}\sum_{\varepsilon}\sum_{\zeta\sim\varepsilon}(f(\zeta)-f(\varepsilon))f(\varepsilon) \\
&= -\frac{1}{4}\left(\sum_{\varepsilon}\sum_{\zeta\sim\varepsilon}(f(\zeta)-f(\varepsilon))f(\varepsilon) - \sum_{\zeta}\sum_{\zeta\sim\varepsilon}(f(\zeta)-f(\varepsilon))f(\zeta)\right) \\
&= \frac{1}{4}\sum_{\varepsilon}\sum_{\zeta\sim\varepsilon}(f(\zeta)-f(\varepsilon))^2 \\
&= \int |\nabla f|^2 d\mu_n.
\end{aligned}$$

Από την ταυτότητα του Parseval,

$$\begin{aligned}
\int f^2 d\mu_n = \sum_{i=1}^{n} \widehat{f}_i^2 + \sum_{|A|\geqslant 2} \widehat{f}_A^2 &\leqslant \sum_{i=1}^{n} \widehat{f}_i^2 + \sum_{|A|\geqslant 2} |A| \widehat{f}_A^2 \\
&= \mathcal{E}(f) = \int |\nabla f|^2 d\mu_n.
\end{aligned}$$

Κλείνουμε αυτή την παράγραφο με ένα θεώρημα των Latala και Oleszkiewicz το οποίο δίνει την καλύτερη σταθερά στην ανισότητα του Kahane για τη σύγκριση της $L^1(X)$ και της $L^2(X)$ νόρμας αθροισμάτων Rademacher.

Θεώρημα 5.6.2. *Θέτουμε $S_n = \sum_{i=1}^n \varepsilon_i x_i$, όπου $\varepsilon_1, \ldots, \varepsilon_n$ είναι ανεξάρτητες τυχαίες μεταβλητές Rademacher και $x_1, \ldots, x_n$ διανύσματα σε ένα χώρο X με νόρμα. Τότε,*

$$\|S_n\|_{L^2(X)} \leqslant \sqrt{2}\,\|S_n\|_{L^1(X)}.$$

Απόδειξη. Θεωρούμε τον E_2^n σαν υποσύνολο του $\mathbb{R}^n$ και ορίζουμε

$$F(t) = \|t_1x_1 + \cdots + t_nx_n\|.$$

Αν $f = F|_{E_2^n}$ τότε

$$f(\varepsilon) = \|S_n(\varepsilon)\|, \quad \langle f, f\rangle = \|S_n\|^2_{L^2(X)}, \quad \mathbb{E}(f) = \|S_n\|_{L^1(X)}.$$

Η f είναι άρτια συνάρτηση, άρα $\widehat{f_i} = 0$ για κάθε $i = 1, \ldots, n$. Επίσης, η F είναι κυρτή και θετικά ομογενής, άρα

$$\frac{1}{n}\sum_{\zeta\sim\varepsilon} f(\zeta) \geqslant F\left(\frac{1}{n}\sum_{\zeta\sim\varepsilon}\zeta\right) = F\left(\frac{(n-2)\varepsilon}{n}\right) = \frac{n-2}{n}f(\varepsilon).$$

Έπεται ότι

$$-L(f)(\varepsilon) \leqslant \frac{1}{2}(nf(\varepsilon) - (n-2)f(\varepsilon)) = f(\varepsilon).$$

Τότε,

$$\mathcal{E}(f) = \langle f, -L(f)\rangle \leqslant \|f\|_2^2,$$

άρα

$$2\|f\|_2^2 \leqslant \|f\|_2^2 + 2(\mathbb{E}(f))^2.$$

Αυτό αποδεικνύει την

$$\|S_n\|_{L^2(X)} \leqslant \sqrt{2}\,\|S_n\|_{L^1(X)}$$

που είναι ο ισχυρισμός του θεωρήματος. □

5.7 Ασκήσεις

1. Έστω μ μέτρο πιθανότητας στον $\mathbb{R}^n$ το οποίο ικανοποιεί την ανισότητα Poincaré με σταθερά $\beta > 0$. Αν $F : \mathbb{R}^n \to \mathbb{R}$ είναι φραγμένη συνάρτηση και αν $d\nu = \frac{1}{\mathbb{E}_\mu(e^F)}e^F d\mu$, δείξτε ότι το ν ικανοποιεί την ανισότητα Poincaré με σταθερά $\beta e^{-4\|F\|_\infty}$.

2. Έστω ν το εκθετικό μέτρο πιθανότητας στο $\mathbb{R}$ με πυκνότητα $\frac{1}{2}e^{-|x|}$, $x \in \mathbb{R}$. Έστω $f : \mathbb{R} \to \mathbb{R}$ συνεχώς παραγωγίσιμη συνάρτηση. Δείξτε ότι

$$\int f d\nu = f(0) + \int \mathrm{sign}(x) f'(x) d\nu(x)$$

και συμπεράνατε ότι
$$\mathrm{Var}_\nu(f) \leqslant 4\int [f'(x)]^2 d\nu(x).$$

3. Έστω $f:\mathbb{R}^n\to\mathbb{R}$ μια 1-Lipschitz συνάρτηση. Για κάθε $\varepsilon>0$ ορίζουμε
$$f_\varepsilon(x) = \frac{1}{|B(x,\varepsilon)|}\int_{B(x,\varepsilon)} f(y)dy.$$
Δείξτε ότι η f_ε είναι διαφορίσιμη και $\|\nabla f_\varepsilon(x)\|_2 \leqslant 1$ για κάθε $x\in\mathbb{R}^n$. Δείξτε επίσης ότι
$$\|f-f_\varepsilon\|_\infty \leqslant \frac{\varepsilon n}{n+1} \leqslant \varepsilon,$$
δηλαδή $f_\varepsilon \to f$ ομοιόμορφα καθώς το $\varepsilon\to 0^+$.

4. Με τις υποθέσεις της προηγούμενης άσκησης δείξτε ότι
$$\|\nabla f_\varepsilon(x) - \nabla f_\varepsilon(y)\|_2 \leqslant \frac{c\sqrt{n}}{\varepsilon}\|x-y\|_2$$
για κάθε $x,y\in\mathbb{R}^n$, όπου $c>0$ είναι μια απόλυτη σταθερά.

5. Έστω $a,b\in\mathbb{R}$ και έστω $F:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^n$ η συνάρτηση που ορίζεται από την $F(x,y) = ax+by$. Δείξτε ότι $F(\gamma_n\otimes\gamma_n)=\gamma_n$, όπου
$$[F(\gamma_n\otimes\gamma_n)](A) = (\gamma_n\otimes\gamma_n)(\{(x,y): F(x,y)\in A\}).$$
Υπόδειξη. Ο $U:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^n\times\mathbb{R}^n$ με $U(x,y)=(ax+by, bx-ay)$ είναι ορθογώνιος.

6. Έστω $\{P_t\}_{t\geqslant 0}$ ημιομάδα Markov με γεννήτορα L και αναλλοίωτο μέτρο μ. Αποδείξτε ότι:

(α) $\mathbb{E}_\mu(Lf)=0$ για κάθε $f\in D(L)$.

(β) Αν $\phi:\mathbb{R}\to\mathbb{R}$ είναι κυρτή συνάρτηση και $f,\phi(f)\in L^2(\mu)$ τότε $P_t\phi(f)\geqslant \phi(P_tf)$ για κάθε $t\geqslant 0$.

(γ) Αν $\phi:\mathbb{R}\to\mathbb{R}$ είναι κυρτή συνάρτηση και $f,\phi(f)\in D(L)$ τότε $L\phi(f)\geqslant \phi'(f)Lf$.

7. Έστω $\{P_t\}_{t\geqslant 0}$ συμμετρική, εργοδική ημιομάδα Markov με αναλλοίωτο μέτρο μ. Ορίζουμε $\mathrm{Cov}_\mu(f,g) = \langle f-\mathbb{E}_\mu(f), g-\mathbb{E}_\mu g\rangle_\mu$. Αποδείξτε ότι:
$$\mathrm{Cov}_\mu(f,g) = 2\int_0^\infty \mathcal{E}(P_tf,P_tg)\,dt = \int_0^\infty \mathcal{E}(f,P_tg)\,dt.$$

8. Έστω $\{P_t\}_{t\geqslant 0}$ ημιομάδα Markov με γεννήτορα L και αναλλοίωτο μέτρο μ. Ορίζουμε
$$\Gamma(f,g) = \frac{1}{2}\,[L(fg) - f\,Lg - g\,Lf].$$
Αποδείξτε ότι:

(α) $\mathcal{E}(f,f) = \int \Gamma(f,f)\,d\mu$ και αν η $\{P_t\}_{t\geqslant 0}$ είναι συμμετρική τότε, επιπλέον,

$$\mathcal{E}(f,g) = \int \Gamma(f,g)\,d\mu.$$

(β) $\Gamma(f,f)\geqslant 0$. [*Υπόδειξη:* Χρησιμοποιήστε την $P_t(f^2)\geqslant (P_tf)^2$ και τον ορισμό του L.]

(γ) $\Gamma(f,g)^2 \leqslant \Gamma(f,f)\Gamma(g,g)$.

9. Έστω $\{P_t\}_{t\geqslant 0}$ ημιομάδα Markov με γεννήτορα L και αναλλοίωτο μέτρο μ. Αποδείξτε ότι:

$$P_t(f^2) - (P_tf)^2 = 2\int_0^t P_{t-s}\Gamma(P_sf, P_sf)\,ds.$$

10. Έστω μ ένα Borel μέτρο πιθανότητας στο $\mathbb{R}$ με πυκνότητα f ως προς το μέτρο Lebesgue, και έστω $m\in\mathbb{R}$ τέτοιος ώστε $\mu([m,+\infty))\geqslant \frac{1}{2}$ και $\mu((-\infty,m])\geqslant\frac{1}{2}$. Ορίζουμε

$$\alpha_+(x) = \int_m^x \frac{1}{f(t)}\,dt, \qquad x\geqslant m$$

και

$$\alpha_-(x) = \int_x^m \frac{1}{f(t)}\,dt, \qquad x\leqslant m.$$

Αποδείξτε ότι: αν

$$b_+ = \int_m^\infty \alpha_+(x)f(x)dx < \infty \quad \text{και} \quad b_- = \int_{-\infty}^m \alpha_-(x)f(x)dx < \infty,$$

τότε το μ ικανοποιεί ανισότητα Poincaré με σταθερά $\beta = 1/\min\{b_+,b_-\}$.

11. Έστω $(X,\mathcal{A},\mu)$ χώρος μέτρου και έστω $A,B\in\mathcal{A}$ με $\mu(A\cap B)>0$ και $\mu(A\cup B)<\infty$. Αποδείξτε ότι: αν η $f:X\to\mathbb{R}$ ικανοποιεί τις

$$\int_{A\cup B} f^2 d\mu < \infty \quad \text{και} \quad \int_{A\cup B} f\,d\mu = 0,$$

τότε

$$\frac{1}{\mu(A)}\left(\int_A f\,d\mu\right)^2 + \frac{1}{\mu(B)}\left(\int_B f\,d\mu\right)^2 \leqslant \left(1-\frac{\mu(A\cap B)}{\mu(A\cup B)}\right)\int_{A\cup B} f^2 d\mu.$$

12. Έστω $\{P_t\}_{t\geqslant 0}$ ημιομάδα Markov με γεννήτορα L και αναλλοίωτο μέτρο μ. Λέμε ότι το μ ικανοποιεί την ανισότητα Sobolev με εκθέτη $p>2$ και σταθερές $\alpha\in\mathbb{R}$, $\gamma>0$, αν για κάθε $f\in D(\mathcal{E})$

$$\|f\|_p^2 \leqslant \alpha\|f\|_2^2 + \gamma\mathcal{E}(f,f).$$

Αποδείξτε ότι: αν το μ ικανοποιεί την ανισότητα Sobolev με εκθέτη $p>2$ και σταθερές $\alpha=0$, $\gamma>0$, τότε το μ ικανοποιεί την ανισότητα Poincaré με σταθερά $\beta = \frac{p-2}{\gamma}$.

Κεφάλαιο 6

Λογαριθμική ανισότητα Sobolev

6.1 Εντροπία και υποκανονικές εκτιμήσεις

Ορισμός 6.1.1 (εντροπία)**.** Έστω $(X, \mathcal{A}, \mu)$ ένας χώρος πιθανότητας. Για κάθε μη αρνητική μετρήσιμη συνάρτηση $f : X \to \mathbb{R}$, η *εντροπία* της f ως προς το μ είναι η ποσότητα

$$\mathrm{Ent}_\mu(f) = \int_X f \ln f \, d\mu - \int_X f \, d\mu \, \ln \int_X f \, d\mu$$

αν $\int f \ln(1+f) \, d\mu < +\infty$, και $+\infty$ αλλιώς. Παρατηρήστε ότι $\mathrm{Ent}_\mu(f) \geqslant 0$ αν εφαρμόσουμε την ανισότητα Jensen για την κυρτή συνάρτηση $g(x) = x \ln x$, και ότι η εντροπία είναι ομογενής βαθμού 1.

Παρατήρηση 6.1.2. Αν υποθέσουμε ότι $f = \frac{d\nu}{d\mu}$ για κάποιο μέτρο πιθανότητας ν, τότε έχουμε

$$\mathrm{Ent}_\mu(f) = \int_X \ln f \, d\nu.$$

Η ποσότητα αυτή ονομάζεται *σχετική εντροπία* του ν ως προς το μ και συχνά συμβολίζεται με $H(\nu|\mu)$. Παρατηρήστε επίσης ότι αν $\int_X f \, d\mu = 1$ τότε

$$\lim p \to 1^+ \frac{p}{p-1} \ln \|f\|_p = \int_X f \ln f \, d\mu.$$

Μια πολύ χρήσιμη περιγραφή της εντροπίας δίνεται από το επόμενο λήμμα.

Λήμμα 6.1.3. *Έστω $f \geqslant 0$ μετρήσιμη συνάρτηση ορισμένη σε ένα χώρο πιθανότητας $(X, \mathcal{A}, \mu)$. Τότε,*

$$\mathrm{Ent}_\mu(f) = \sup\Big\{\int fg\,d\mu : \ \int e^g\,d\mu \leqslant 1\Big\}.$$

Απόδειξη. Η σχέση είναι ομογενής, άρα μπορούμε να υποθέσουμε ότι $\int f\,d\mu = 1$. Από την ανισότητα του Young έχουμε

$$uv \leqslant u\ln u - u + e^v, \ u \geqslant 0, \ v \in \mathbb{R}$$

άρα, αν g είναι μια μετρήσιμη συνάρτηση στον X με $\int e^g\,d\mu \leqslant 1$, τότε

$$\int fg\,d\mu \leqslant \int f\ln f\,d\mu - \int f\,d\mu + \int e^g\,d\mu \leqslant \int f\ln f\,d\mu.$$

Παίρνοντας το supremum βλέπουμε ότι

$$\sup\Big\{\int fg\,d\mu : \ \int e^g\,d\mu \leqslant 1\Big\} \leqslant \int f\ln f\,d\mu = \mathrm{Ent}_\mu(f),$$

χρησιμοποιώντας και την υπόθεση ότι $\int f\,d\mu = 1$. Τέλος, αφού $\int e^{\ln f}\,d\mu = \int f\,d\mu = 1$, έχουμε

$$\sup\Big\{\int fg\,d\mu : \ \int e^g\,d\mu \leqslant 1\Big\} \geqslant \int f\ln f\,d\mu,$$

το οποίο μας δίνει την ισότητα. □

Το επόμενο λήμμα, γνωστό και ως *επιχείρημα του Herbst*, συνδέει την εντροπία με τις υποκανονικές εκτιμήσεις.

Λήμμα 6.1.4 (Herbst)**.** *Έστω $(X, \mathcal{A}, \mu)$ ένας χώρος πιθανότητας και έστω $f : X \to \mathbb{R}$ μετρήσιμη συνάρτηση. Αν*

$$\mathrm{Ent}_\mu(e^{\lambda f}) \leqslant \frac{\lambda^2\sigma^2}{2}\mathbb{E}_\mu(e^{\lambda f})$$

για κάθε $\lambda \geqslant 0$ τότε η f είναι σ^2-υποκανονική.

Απόδειξη. Αρκεί να ελέγξουμε ότι η λογαριθμική ροπογεννήτρια ψ της f ικανοποιεί την

$$\psi(\lambda) := \ln\mathbb{E}_\mu(e^{\lambda(f-\mathbb{E}_\mu(f))}) \leqslant \frac{\lambda^2\sigma^2}{2}$$

για κάθε $\lambda \geqslant 0$. Γράφουμε $\psi(\lambda) = \ln\mathbb{E}_\mu(e^{\lambda f}) - \lambda\mathbb{E}_\mu(f)$ και παραγωγίζοντας παίρνουμε

$$\frac{d}{d\lambda}\Big(\frac{\psi(\lambda)}{\lambda}\Big) = \frac{1}{\lambda}\frac{\mathbb{E}_\mu(fe^{\lambda f})}{\mathbb{E}_\mu(e^f)} - \frac{1}{\lambda^2}\ln\mathbb{E}_\mu(e^{\lambda f}) = \frac{1}{\lambda^2}\frac{\mathrm{Ent}_\mu(e^{\lambda f})}{\mathbb{E}_\mu(e^{\lambda f})}.$$

Κατόπιν, παρατηρούμε ότι $\lim_{\lambda\to 0^+}\frac{\psi(\lambda)}{\lambda}=0$ και χρησιμοποιώντας την υπόθεση γράφουμε

$$\begin{aligned}\frac{\psi(\lambda)}{\lambda} &= \int_0^\lambda \frac{1}{s^2}\frac{\mathrm{Ent}_\mu(e^{sf})}{\mathbb{E}_\mu(e^{sf})}\,ds \leqslant \int_0^\lambda \frac{1}{s^2}\frac{s^2\sigma^2}{2}\,ds\\ &= \frac{\lambda\sigma^2}{2}.\end{aligned}$$

Αυτό αποδεικνύει ότι $\psi(\lambda)\leqslant \lambda^2\sigma^2/2$. □

6.2 Λογαριθμική ανισότητα Sobolev

Ορισμός 6.2.1. Έστω (X,d,μ) μετρικός χώρος πιθανότητας. Λέμε ότι το μ ικανοποιεί τη λογαριθμική ανισότητα Sobolev με σταθερά $\beta>0$ αν για κάθε τοπικά Lipschitz συνάρτηση $f:X\to\mathbb{R}$ ισχύει

$$\mathrm{Ent}_\mu(f^2)\leqslant \frac{2}{\beta}\int |\nabla f|^2\,d\mu. \tag{6.2.1}$$

Η ανισότητα είναι ομογενής, μπορούμε λοιπόν χωρίς περιορισμό της γενικότητας να υποθέσουμε ότι $\int_X f^2d\mu=1$. Τότε μπορούμε να γράψουμε $d\nu=f^2d\mu=gd\mu$ και η ανισότητα παίρνει τη μορφή

$$H(\nu|\mu)\leqslant \frac{1}{2\beta}\int \frac{|\nabla g|^2}{g}d\mu. \tag{6.2.2}$$

Η ποσότητα $\int \frac{|\nabla g|^2}{g}d\mu =: I_\mu(\nu)$ ονομάζεται *πληροφορία Fisher* του ν ως προς το μ.

Η επόμενη πρόταση δείχνει ότι η λογαριθμική ανισότητα Sobolev είναι ισχυρότερη από την ανισότητα Poincaré. Μάλιστα, όπως φαίνεται από την απόδειξη, η δεύτερη ανισότητα είναι μια γραμμικοποίηση της πρώτης.

Πρόταση 6.2.2. *Έστω (X,d,μ) μετρικός χώρος πιθανότητας. Αν το μ ικανοποιεί τη λογαριθμική ανισότητα Sobolev με σταθερά β, τότε το μ ικανοποιεί την ανισότητα Poincaré με σταθερά β.*

Απόδειξη. Έστω $f:X\to\mathbb{R}$ τοπικά Lipschitz συνάρτηση. Θέτουμε $g=f-\mathbb{E}_\mu(f)$. Τότε, $|\nabla g|=|\nabla f|$ και αρκεί να αποδείξουμε ότι αν $\mathbb{E}_\mu(g)=0$ τότε

$$\int g^2d\mu\leqslant \beta\int |\nabla g|^2d\mu.$$

Χωρίς περιορισμό της γενικότητας μπορούμε επίσης να υποθέσουμε ότι $\mathbb{E}_\mu(g^2)=1$. Εφαρμόζουμε τη λογαριθμική ανισότητα Sobolev για τη συνάρτηση $1+\varepsilon g$ και παίρνουμε το όριο καθώς το $\varepsilon\to 0^+$. Ξεκινώντας από την

$$\begin{aligned}2\int(1+\varepsilon g)^2\ln(1+\varepsilon g)\,d\mu-\int(1+\varepsilon g)^2d\mu\,\cdot\,\ln\Big(\int(1+\varepsilon g)^2d\mu\Big)\\ \leqslant\frac{2}{\beta}\varepsilon^2\int|\nabla g|^2d\mu,\end{aligned}$$

και κάνοντας απλούς υπολογισμούς, παίρνουμε

$$\begin{aligned}2\int(1+2\varepsilon g+\varepsilon^2g^2)^2\big(\varepsilon g-\frac{\varepsilon^2g^2}{2}\big)\,d\mu-(1+\varepsilon^2)\ln(1+\varepsilon^2)\\ \leqslant\frac{2}{\beta}\varepsilon^2\int|\nabla g|^2d\mu+O(\varepsilon^3).\end{aligned}$$

Παίρνοντας υπόψη μας τις $\mathbb{E}_\mu(g)=0$ και $\mathbb{E}_\mu(g^2)=1$ στο αριστερό μέλος, βλέπουμε ότι

$$3\varepsilon^2-(1+\varepsilon^2)\ln(1+\varepsilon^2)\leqslant\frac{2}{\beta}\varepsilon^2\int|\nabla g|^2d\mu+O(\varepsilon^3).$$

Από την $(1+\varepsilon^2)\ln(1+\varepsilon^2)=\varepsilon^2+O(\varepsilon^3)$ έπεται ότι

$$1\leqslant\frac{1}{\beta}\int|\nabla g|^2d\mu+O(\varepsilon)$$

και αφήνοντας το $\varepsilon\to 0^+$ παίρνουμε

$$\int g^2d\mu=1\leqslant\frac{1}{\beta}\int|\nabla g|^2d\mu.$$

Αυτό αποδεικνύει ότι το μ ικανοποιεί την ανισότητα Poincaré με σταθερά β. □

Το επόμενο θεώρημα, γνωστό ως επιχείρημα του Herbst, δείχνει ότι η λογαριθμική ανισότητα Sobolev έχει ως συνέπεια κανονική συγκέντρωση σε ένα μετρικό χώρο πιθανότητας. Υπενθυμίζουμε ότι το μ έχει *κανονική συγκέντρωση* στον (X,d) αν υπάρχουν σταθερές $C,c>0$ τέτιες ώστε, για κάθε $t>0$,

$$\alpha_\mu(t)\leqslant Ce^{-ct^2}.$$

Θεώρημα 6.2.3 (Herbst). *Έστω (X,d,μ) ένας μετρικός χώρος πιθανότητας που ικανοποιεί τη λογαριθμική ανισότητα Sobolev* (6.2.1) *με σταθερά β. Τότε, για κάθε $t>0$ έχουμε*

$$\alpha_\mu(t)\leqslant\exp(-\beta t^2/8)$$

Απόδειξη. Έστω $g : X \to \mathbb{R}$ Lipschitz συνεχής συνάρτηση με $\|g\|_{\mathrm{Lip}} \leqslant 1$ και $\mathbb{E}_\mu(g) = 0$. Ορίζουμε $f = e^{\lambda g/2}$. Παρατηρήστε ότι

$$\nabla f = \frac{\nabla(f^2)}{2f} = \frac{\lambda e^{\lambda g}\nabla g}{2e^{\lambda g/2}} = \frac{\lambda}{2}e^{\lambda g/2}\nabla g.$$

Από τη λογαριθμική ανισότητα Sobolev έχουμε

$$\mathrm{Ent}_\mu(f^2) \leqslant \frac{2}{\beta}\int |\nabla f|^2\, d\mu = \frac{\lambda^2}{2\beta}\int_X e^{\lambda g}|\nabla g|^2\, d\mu.$$

Αφού $|\nabla g| \leqslant 1$, βλέπουμε ότι

$$\mathrm{Ent}_\mu(f^2) \leqslant \frac{\lambda^2}{2\beta}\mathbb{E}_\mu(e^{\lambda g}).$$

Τότε, το Λήμμα 6.1.4 δείχνει ότι η g είναι $(1/\beta)$-υποκανονική. Έχουμε λοιπόν δείξει ότι για κάθε 1-Lipschitz συνάρτηση g και για κάθε $t > 0$ ισχύει

$$\mu(\{g - \mathbb{E}_\mu(g) \geqslant t\}) \leqslant \alpha(t) := \exp(-\beta t^2/2),$$

και γνωρίζουμε ότι αυτό έχει ως συνέπεια την $\alpha_\mu(t) \leqslant \alpha(t/2) = \exp(-\beta t^2/8)$ για κάθε $t > 0$. □

6.3 Λογαριθμική ανισότητα Sobolev σε χώρους γινόμενα

Η λογαριθμική ανισότητα Sobolev συμπεριφέρεται καλά ως προς γινόμενα μέτρων. Έστω $(X_i, \mathcal{A}_i, \mu_i)$, $i = 1, \dots, n$ χώροι πιθανότητας και $X = \prod_{i=1}^n X_i$ ο χώρος γινόμενο, εφοδιασμένος με το μέτρο γινόμενο $P = \mu_1 \otimes \dots \otimes \mu_n$. Αν f είναι συνάρτηση ορισμένη στον X, συμβολίζουμε με f_i την συνάρτηση που ορίζεται, για σταθερά $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$, στον X_i ως εξής:

$$f_i(x_i) = f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n).$$

Πρόταση 6.3.1. *Για κάθε μη αρνητική συνάρτηση f ορισμένη στον χώρο γινόμενο X ισχύει*

$$\mathrm{Ent}_P(f) \leqslant \sum_{i=1}^n \int \mathrm{Ent}_{\mu_i}(f_i)\, dP.$$

Απόδειξη. Με βάση το Λήμμα 6.1.3, αρκεί να δείξουμε ότι αν η $g : X \to \mathbb{R}$ ικανοποιεί την $\int e^g\, dP \leqslant 1$, τότε

$$\int fg\, dP \leqslant \sum_{i=1}^n \int \mathrm{Ent}_{\mu_i}(f_i)\, dP.$$

Για κάθε $i = 1, \dots, n$ θέτουμε

$$g^i(x_1, \dots, x_n) = \ln\left(\frac{\int e^g \, d\mu_1(x_1) \dots d\mu_{i-1}(x_{i-1})}{\int e^g \, d\mu_1(x_1) \dots d\mu_i(x_i)}\right).$$

Τότε $g \leqslant \sum_{i=1}^n g^i$ και $\int e^{(g^i)_i} \, d\mu_i = 1$. Πράγματι,

$$\begin{aligned}\sum_{i=1}^n g^i &= \ln\left(\frac{e^g}{\int e^g \, d\mu_1} \frac{\int e^g \, d\mu_1}{\int e^g \, d\mu_1 d\mu_2} \cdots \frac{\int e^g \, d\mu_1 \dots d\mu_{n-1}}{\int e^g \, d\mu_1 \dots d\mu_n}\right) \\ &= \ln\left(\frac{e^g}{\int e^g \, dP}\right),\end{aligned}$$

οπότε αφού $\int e^g \, dP \leqslant 1$ έπεται ότι

$$g \leqslant \ln\left(\frac{e^g}{\int e^g \, dP}\right).$$

Επίσης,

$$\begin{aligned}\int e^{(g^i)_i} \, d\mu_i &= \int \frac{\int e^g \, d\mu_1(x_1) \dots d\mu_{i-1}(x_{i-1})}{\int e^g \, d\mu_1(x_1) \dots d\mu_i(x_i)} \, d\mu_i(x_i) \\ &= \frac{\int e^g \, d\mu_1 \dots d\mu_i}{\int e^g \, d\mu_1 \dots d\mu_i} = 1.\end{aligned}$$

Άρα,

$$\begin{aligned}\int fg \, dP &\leqslant \sum_{i=1}^n \int fg^i \, dP = \sum_{i=1}^n \int \left(\int f_i (g^i)_i \, d\mu_i\right) \, dP \\ &\leqslant \sum_{i=1}^n \int \mathrm{Ent}_{\mu_i}(f_i) \, dP,\end{aligned}$$

οπού η τελευταία ανισότητα προκύπτει από το γεγονός ότι $\int (g^i)_i \, d\mu_i = 1$. □

Πόρισμα 6.3.2. *Έστω (X_i, d_i, μ_i) μετρικοί χώροι πιθανότητας. Υποθέτουμε ότι, για κάθε $i = 1, \dots, n$, υπάρχει σταθερά β_i ώστε για κάθε τοπικά Lipschitz συνάρτηση f στον X_i να ισχύει*

$$\mathrm{Ent}_{\mu_i}(f^2) \leqslant \frac{2}{\beta_i} \int |\nabla_i f|^2 \, d\mu_i,$$

όπου $|\nabla_i f|$ το μέτρο του γενικευμένου gradient στον X_i. Τότε για κάθε τοπικά Lipschitz συνάρτηση f στον $X = \prod_{i=1}^n X_i$ ισχύει

$$\mathrm{Ent}_P(f^2) \leqslant \frac{2}{\min_{1 \leqslant i \leqslant n} \beta_i} \int |\nabla f|^2 \, dP,$$

όπου

$$|\nabla f|^2 = \sum_{i=1}^{n} |\nabla_i f|^2$$

Απόδειξη. Από την προηγούμενη πρόταση έχουμε ότι

$$\begin{aligned}\mathrm{Ent}_P(f^2) &\leqslant \sum_{i=1}^{n} \int \mathrm{Ent}_{\mu_i}(f^2)\, dP \leqslant \sum_{i=1}^{n} \int \left(\frac{2}{\beta_i} \int |\nabla_i f|^2 \, d\mu_i \right) \, dP \\ &\leqslant \frac{2}{\in_{1\leqslant i\leqslant n} \beta_i} \int \left(\int \sum_{i=1}^{n} |\nabla_i f|^2 \, d\mu_i \right) \, dP \\ &= \frac{2}{\min_{1\leqslant i\leqslant n} \beta_i} \int |\nabla f|^2 \, dP.\end{aligned}$$

6.4 Ημιομάδες Markov και η λογαριθμική ανισότητα Sobolev

Σε αυτή την παράγραφο συζητάμε τη λογαριθμική ανισότητα Sobolev στο αφηρημένο πλαίσιο μιας ημιομάδας Markov $\{P_t\}_{t\geqslant 0}$ με αναλλοίωτο μέτρο μ. Το επόμενο θεώρημα, που είναι το ανάλογο του Θεωρήματος 5.3.6, δείχνει ότι το μ ικανοποιεί τη λογαριθμική ανισότητα Sobolev αν και μόνο αν έχουμε εκθετική σύγκλιση της ημιομάδας προς το αναλλοίωτο μέτρο της με την έννοια της εντροπίας.

Θεώρημα 6.4.1. *Έστω $\{P_t\}_{t\geqslant 0}$ μια ημιομάδα Markov με αναλλοίωτο μέτρο μ και έστω $\beta > 0$. Τα εξής είναι ισοδύναμα:*

(α) *Ισχύει η λογαριθμική ανισότητα Sobolev με σταθερά β: για κάθε $f \geqslant 0$,*

$$\mathrm{Ent}_\mu(f) \leqslant \frac{1}{\beta}\, \mathcal{E}(\ln f, f).$$

(β) *Για κάθε $f \geqslant 0$ και $t \geqslant 0$*

$$\mathrm{Ent}_\mu(P_t f) \leqslant e^{-\beta t} \mathrm{Ent}_\mu(f).$$

Επιπλέον, αν $\mathrm{Ent}_\mu(P_t f) \to 0$ καθώς το $t \to \infty$ και

$$\mathcal{E}(\ln P_t f, P_t f) \leq^{-\beta t} \mathcal{E}(\ln f, f)$$

για κάθε $f \geqslant 0$ και $t \geqslant 0$, τότε ισχύουν οι παραπάνω ισοδύναμες προτάσεις.

Απόδειξη. Αρχικά παραγωγίζουμε την $\mathrm{Ent}_\mu(P_tf)$. Παρατηρήστε ότι

$$\mathbb{E}_\mu(\mathcal{L}P_tf) = \frac{d}{dt}\big(\mathbb{E}_\mu(P_tf)\big) = \frac{d}{dt}\big(\mathbb{E}_\mu(f)\big) = 0.$$

Συνεπώς,

$$\frac{d}{dt}\mathrm{Ent}_\mu(P_tf) = \mathbb{E}_\mu(\mathcal{L}P_tf\cdot \ln P_tf) + \mathbb{E}_\mu(\mathcal{L}P_tf) = -\mathcal{E}(\ln P_tf, P_tf).$$

Υποθέτοντας ότι ισχύει η (α) παρατηρούμε ότι

$$\frac{d}{dt}\mathrm{Ent}_\mu(P_tf) = -\mathcal{E}(\ln P_tf, P_tf) \leqslant -\beta\mathrm{Ent}_\mu(P_tf),$$

απ' όπου έπεται η (β). Αντίστροφα, υποθέτοντας την (β) γράφουμε

$$\mathcal{E}(\ln f, f) = \lim_{t\to 0^+}\frac{\mathrm{Ent}_\mu(f) - \mathrm{Ent}_\mu(P_tf)}{t} \geqslant \lim_{t\to 0^+}\frac{1-e^{-\beta t}}{t}\,\mathrm{Ent}_\mu(f) = \beta\mathrm{Ent}_\mu(f).$$

Τέλος, αν υποθέσουμε την «εργοδική συνθήκη» $\lim_{t\to\infty}\mathrm{Ent}_\mu(P_tf) = 0$ και την (γ) μπορούμε να γράψουμε

$$\mathrm{Ent}_\mu(f) = \int_0^\infty \mathcal{E}(\ln P_tf, P_tf)\,dt \leqslant \mathcal{E}(\ln f, f)\int_0^\infty e^{-\beta t}dt = \frac{1}{\beta}\mathcal{E}(f,f),$$

άρα οι (α) και (β) ισχύουν και οι δύο. □

6.5 Η λογαριθμική ανισότητα Sobolev στο χώρο του Gauss

Μπορούμε να χρησιμοποιήσουμε το Θεώρημα 6.4.1 για να αποδείξουμε τη λογαριθμική ανισότητα Sobolev στο χώρο του Gauss. Θα δώσουμε όμως μια απευθείας απόδειξη, πάλι χρησιμοποιώντας την ημιομάδα Ornstein-Uhlenbeck.

Θεώρημα 6.5.1 (λογαριθμική ανισότητα Sobolev)**.** *Για κάθε $f \in W^{2,1}(\gamma_n)$ ισχύει η ανισότητα*

$$\int_{\mathbb{R}^n} f^2\,\ln|f|\,d\gamma_n - \frac{1}{2}\int_{\mathbb{R}^n} f^2\,d\gamma_n\cdot\ln\left(\int_{\mathbb{R}^n} f^2\,d\gamma_n\right) \leqslant \int_{\mathbb{R}^n}|\nabla f|^2 d\gamma_n,$$

δηλαδή

$$\int |f|^2\ln|f|\,d\gamma_n \leqslant \int|\nabla f|^2 d\gamma_n + \|f\|_2^2\ln\|f\|_2,$$

με τη σύμβαση $f^2\ln|f| = 0$ αν $f=0$.

Απόδειξη. Θα υποθέσουμε ότι $f \in C_b^\infty(\mathbb{R}^n)$ και $f \geqslant c > 0$. Θέτουμε $\phi = f^2$, οπότε $\nabla f = \frac{\nabla\phi}{2\sqrt{\phi}}$ και η ανισότητα παίρνει τη μορφή

$$\int_{\mathbb{R}^n} \phi \ln\phi \, d\gamma_n - \int_{\mathbb{R}^n} \phi \, d\gamma_n \cdot \ln\left(\int_{\mathbb{R}^n} \phi \, d\gamma_n\right) \leqslant \frac{1}{2}\int_{\mathbb{R}^n} \frac{|\nabla\phi|^2}{\phi} \, d\gamma_n.$$

Παρατηρούμε ότι το αριστερό μέλος είναι ίσο με

$$\int_{\mathbb{R}^n} T_0\phi \cdot \ln T_0\phi - \int_{\mathbb{R}^n} T_\infty\phi \cdot \ln T_\infty\phi,$$

πορούμε λοιπόν να το γράψουμε στη μορφή

$$-\int_0^\infty \left(\frac{d}{dt}\left[\int_{\mathbb{R}^n} T_t\phi \cdot \ln T_t\phi \, d\gamma_n\right]\right) dt.$$

Όμως,

$$\frac{d}{dt}\left[\int_{\mathbb{R}^n} T_t\phi \cdot \ln T_t\phi \, d\gamma_n\right] = \int_{\mathbb{R}^n} \left\{LT_t\phi \cdot \ln T_t\phi + \frac{d}{dt}T_t\phi\right\}$$

και

$$\begin{aligned}\int_0^\infty \int_{\mathbb{R}^n} \frac{d}{dt}T_t\phi &= \int_{\mathbb{R}^n} T_\infty\phi \, d\gamma_n - \int_{\mathbb{R}^n} T_0\phi \, d\gamma_n \\ &= \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} \phi \, d\gamma_n\right) d\gamma_n - \int_{\mathbb{R}^n} \phi \, d\gamma_n = 0.\end{aligned}$$

Συνεπώς, το αριστερό μέλος της $(*)$ είναι ίσο με

$$\begin{aligned}-\int_0^\infty \int_{\mathbb{R}^n} LT_tf \cdot \ln T_t\phi \, d\gamma_n &= \int_0^\infty \int_{\mathbb{R}^n} \langle \nabla T_t\phi, \nabla \ln T_t\phi\rangle \, d\gamma_n \, dt \\ &= \int_0^\infty \int_{\mathbb{R}^n} \frac{|\nabla T_t\phi|^2}{T_t\phi} \, d\gamma_n \, dt.\end{aligned}$$

Χρησιμοποιώντας την

$$|T_t\partial_{x_i}\phi|^2 = \left|T_t\left(\sqrt{\phi}\cdot\frac{\partial_{x_i}\phi}{\sqrt{\phi}}\right)\right|^2 \leqslant T_t\phi \cdot T_t\left(\frac{(\partial_{x_i}\phi)^2}{\phi}\right),$$

συμπεραίνουμε ότι

$$\begin{aligned}|\nabla T_t\phi|^2 = e^{-2t}|T_t(\nabla\phi)|^2 &= e^{-2t}\sum_{i=1}^n |T_t\partial_{x_i}\phi|^2 \\ &\leqslant e^{-2t} T_t\phi \cdot \sum_{i=1}^n T_t\left(\frac{(\partial_{x_i}\phi)^2}{\phi}\right).\end{aligned}$$

Έπεται ότι

$$
\begin{aligned}
\int_0^\infty \int_{\mathbb{R}^n} \frac{|\nabla T_t\phi|^2}{T_t\phi}\, d\gamma_n\, dt &\leqslant \int_0^\infty e^{-2t} \int_{\mathbb{R}^n} \sum_{i=1}^n T_t\left(\frac{(\partial_{x_i}\phi)^2}{\phi}\right) d\gamma_n\, dt \\
&= \int_0^\infty e^{-2t} \int_{\mathbb{R}^n} T_t\left(\frac{|\nabla\phi|^2}{\phi}\right) d\gamma_n\, dt \\
&= \int_0^\infty e^{-2t} \int_{\mathbb{R}^n} \frac{|\nabla\phi|^2}{\phi}\, d\gamma_n\, dt \\
&= \int_0^\infty e^{-2t}\, dt \cdot \int_{\mathbb{R}^n} \frac{|\nabla\phi|^2}{\phi}\, d\gamma_n \\
&= \frac{1}{2} \int_{\mathbb{R}^n} \frac{|\nabla\phi|^2}{\phi}\, d\gamma_n,
\end{aligned}
$$

και η απόδειξη είναι πλήρης. □

Σημείωση. Μπορούμε επίσης να δώσουμε μια απόδειξη του φράγματος

$$E_{\gamma_n}(\lambda) \leqslant \exp(\lambda^2/2)$$

για το συναρτησοειδές Laplace του γ_n, χρησιμοποιώντας απευθείας την ημιομάδα Ornstein-Uhlenbeck: ορίζουμε

$$G(t) = \int_{\mathbb{R}^n} e^{\lambda T_t F} d\gamma_n.$$

Τότε,

$$G(0) = \int_{\mathbb{R}^n} e^{\lambda F} d\gamma_n \quad \text{και} \quad G(\infty) = \int_{\mathbb{R}^n} e^{\lambda \int F\, d\gamma_n} d\gamma_n = 1.$$

Συνεπώς,

$$
\begin{aligned}
G(t) &= 1 - \int_t^\infty G'(s)\, ds \\
&= 1 - \int_t^\infty \int_{\mathbb{R}^n} \lambda e^{\lambda T_s F} L T_s(F)\, d\gamma_n\, ds \\
&= 1 + \lambda \int_t^\infty \int_{\mathbb{R}^n} \langle \nabla e^{\lambda T_s F}, \nabla T_s F\rangle\, d\gamma_n\, ds \\
&= 1 + \lambda^2 \int_t^\infty \int_{\mathbb{R}^n} e^{\lambda T_s F} \|\nabla T_s F\|_2^2\, d\gamma_n\, ds \\
&= 1 + \lambda^2 \int_t^\infty \int_{\mathbb{R}^n} e^{\lambda T_s F} e^{-2s} \|T_s(\nabla F)\|_2^2 d\gamma_n\, ds.
\end{aligned}
$$

Όμως, από την $\|\nabla F\|_2 \leqslant 1$ έχουμε $\|T_s(\nabla F)\|_2^2 \leqslant 1$. Άρα,

$$G(t) \leqslant 1 + \lambda^2 \int_t^\infty e^{-2s} G(s)\, ds.$$

Ορίζουμε

$$H(t) := \ln\left(1 + \lambda^2 \int_t^\infty e^{-2s} G(s)\, ds\right).$$

Τότε,

$$H'(t) = -\frac{\lambda^2 e^{-2t} G(t)}{1 + \lambda^2 \int_t^\infty e^{-2s} G(s)\, ds} \geqslant -\lambda^2 e^{-2t},$$

άρα

$$H(t) - H(0) \geqslant -\lambda^2 \int_0^\infty e^{-2t} dt = -\frac{\lambda^2}{2}.$$

Όμως, $\lim\limits_{t\to\infty} H(t) = 0$. Συνεπώς, $H(0) \leqslant \lambda^2/2$. Έπεται ότι

$$\int_{\mathbb{R}^n} e^{\lambda F} d\gamma_n = G(0) \leqslant e^{H(0)} \leqslant \exp(\lambda^2/2).$$

6.6 Λογαριθμική ανισότητα Sobolev και υπερσυσταλτότητα

Έστω $\{P_t\}_{t\geqslant 0}$ μια ημιομάδα Markov με γεννητορα L και αναλλοίωτο μέτρο μ και έστω $\beta > 0$. Η λογαριθμική ανισότητα Sobolev συνδέεται στενά με την υπερσυσταλτότητα: με λίγα λόγια αυτό σημαίνει ότι ο e^{-tL} είναι συστολή από τον L_p σε κάποιον L_q με $q > p$, αν το t είναι αρκετά μεγάλο. Η ακριβής διατύπωση δίνεται στο επόμενο θεώρημα, το οποίο στην περίπτωση του μέτρου του Gauss γ_n οφείλεται στον Gross.

Θεώρημα 6.6.1. *Έστω $\{P_t\}_{t\geqslant 0}$ μια ημιομάδα Markov στον (X,d) με αναλλοίωτο μέτρο μ και γεννήτορα L ώστε*

$$\int \langle \nabla f, \nabla g\rangle\, d\mu = -\int Lf \cdot g\, d\mu, \qquad f, g \in L_2(\mu).$$

Έστω $\beta > 0$. Τα εξής είναι ισοδύναμα:

(α) *Ισχύει η λογαριθμική ανισότητα Sobolev*

$$\mathrm{Ent}_\mu(g^2) \leqslant \frac{2}{\beta} \int |\nabla g|^2 d\mu$$

για κάθε ομαλή συνάρτηση g.

(β) *Για κάθε* $p>1$, *για κάθε θετική* $f\in L_p(\mu)$ *και για κάθε* $q,t>0$ *που ικανοποιούν την* $\frac{q-1}{p-1}\leqslant \exp(2\beta t)$ *ισχύει*

$$\|P_tf\|_q\leqslant \|f\|_p.$$

Απόδειξη. Υποθέτουμε πρώτα ότι ισχύει η λογαριθμική ανισότητα Sobolev με σταθερά β. Η πρώτη παρατήρηση που θα χρησιμοποιήσουμε είναι ότι, για κάθε θετική f,

$$\frac{d}{dq}\left(\int_X f^q\,d\mu\right)=\int_X f^q\ln f\,d\mu=\frac{1}{q}\left(\mathrm{Ent}_\mu(f^q)+\int_X f^q\,d\mu\cdot\ln\left(\int_X f^q\,d\mu\right)\right).$$

Από την άλλη πλευρά, για σταθερό $q>1$, με ολοκλήρωση κατά μέρη παίρνουμε

$$\begin{aligned}\frac{d}{dt}\left(\int_X (P_tf)^q\,d\mu\right)&=q\int_X (P_tf)^{q-1}LP_tf\,d\mu\\&=-q(q-1)\int_X (P_tf)^{q-2}\Gamma(P_tf)\,d\mu,\end{aligned}$$

όπου

$$\Gamma(f,g)=\frac{1}{2}\,[L(fg)-f\,Lg-g\,Lf],$$

άρα

$$\Gamma(g)=\frac{1}{2}L(g^2)-g\,Lg.$$

Τέλος, εφαρμόζοντας τη λογαριθμική ανισότητα Sobolev για την $f^{q/2}$ παίρνουμε

$$\mathrm{Ent}_\mu(f^q)\leqslant\frac{q^2}{2\beta}\int_X f^{q-2}\Gamma(f)\,d\mu. \tag{6.6.1}$$

Για $t\geqslant 0$ και $q>1$ ορίζουμε

$$\Lambda(t,q)=\int_X (P_tf)^q\,d\mu.$$

Χρησιμοποιώντας τις προηγούμενες ταυτότητες και την (6.6.1) βλέπουμε ότι

$$\frac{d}{dq}\Lambda\leqslant-\frac{1}{2\beta(q-1)}\,\frac{d}{dt}\Lambda+\frac{1}{q}\Lambda\ln\Lambda.$$

Αν λοιπόν ορίσουμε

$$H(t)=\frac{1}{q(t)}\ln\Lambda(t,q(t)),$$

όπου $q(t)=1+(p-1)e^{2\beta t}$, παραγωγίζοντας βλέπουμε ότι $H'\leqslant 0$. Άρα, η H είναι φθίνουσα. Ειδικότερα, η $H(t)\leqslant H(0)$ παίρνει τη μορφή

$$\Lambda(t,q(t))^{1/q(t)}\leqslant\Lambda(0,q(0))^{1/q(0)},$$

δηλαδή

$$\left(\int_X (P_t f)^q(t)\,d\mu\right)^{1/q(t)} \leqslant \left(\int_X f^p\,d\mu\right)^{1/p}.$$

Αντίστροφα, αν υποθέσουμε ότι ισχύει η $\|P_t f\|_{q(t)} \leqslant \|f\|_p$ για κάποιο $p > 1$, παραγωγίζοντας στο σημείο $t = 0$ παίρνουμε την (6.6.1) με $q = p$, η οποία είναι ισοδύναμη με τη λογαριθμική ανισότητα Sobolev. □

6.7 Λογαριθμική ανισότητα Sobolev στο διακριτό κύβο

Σκοπός μας είναι να δείξουμε τη λογαριθμική ανισότητα Sobolev για τον E_2^n. Αν $f : E_2^n \to \mathbb{R}^+$ τότε η εντροπία $\mathrm{Ent}(f)$ της f ορίζεται ως συνήθως από την

$$\mathrm{Ent}(f) = \mathbb{E}\,(f \ln f) - \|f\|_1 \ln \|f\|_1.$$

Παρατηρήστε ότι αν $\|f\|_1 = 1$ τότε $\mathrm{Ent}(f) = \mathbb{E}\,(f \ln f)$. Γενικά, αν ν είναι ένα μέτρο πιθανότητας στον E_2^n τότε η εντροπία του ν είναι η ποσότητα

$$\mathrm{Ent}(\nu) = -\sum_{\varepsilon \in E_2^n} \nu(\{\varepsilon\}) \ln_2 \nu(\{\varepsilon\}).$$

Αν θεωρήσουμε το ομοιόμορφο μέτρο πιθανότητας μ_n στον E_2^n έχουμε $\mathrm{Ent}(\mu_n) = n$ και

$$\mathrm{Ent}(\nu) \leqslant \mathrm{Ent}(\mu_n)$$

για κάθε άλλο μέτρο πιθανότητας στον E_2^n. Κάθε $f : E_2^n \to \mathbb{R}^+$ με $\|f\|_1 = 1$ επάγει ένα μέτρο πιθανότητας ν_f στον E_2^n μέσω της $\nu_f(\{\varepsilon\}) = f(\varepsilon)/2^n$. Τότε, η εντροπία $\mathrm{ent}(f)$ του ν_f με την έννοια της θεωρίας πληροφορίας ισούται με

$$\mathrm{ent}(\nu_f) = -\sum_{\varepsilon \in E_2^n} \frac{f(\varepsilon)}{2^n} \ln_2\left(\frac{f(\varepsilon)}{2^n}\right) = n - \frac{\mathrm{Ent}(f)}{\ln 2}.$$

Δηλαδή, η $\mathrm{Ent}(f)$ είναι μια *σχετική εντροπία* που μετράει πόσο απέχει η εντροπία του ν_f από τη μέγιστη δυνατή εντροπία n.

Θυμίζουμε τους ορισμούς της διακριτής Λαπλασιανής και της ενέργειας. Για κάθε $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n) \in E_2^n$, οι «γείτονες» του ε είναι εκείνα τα $\zeta = (\zeta_1, \ldots, \zeta_n) \in E_2^n$ για τα οποία

$$|\{i \leqslant n : \varepsilon_i \neq \zeta_i\}| = 1.$$

Αν τα ε, ζ είναι γείτονες, γράφουμε $\varepsilon \sim \zeta$.

Θεωρούμε τη *διακριτή Λαπλασιανή* $L(f)$ της f, η οποία ορίζεται από την

$$L(f)(\varepsilon) = \frac{1}{2}\sum_{\zeta\sim\varepsilon}[f(\zeta) - f(\varepsilon)].$$

Θεωρούμε επίσης την *ενέργεια* $\mathcal{E}(f)$ της f, η οποία ορίζεται από την

$$\mathcal{E}(f) = -\langle f, L(f)\rangle.$$

Παρατηρούμε ότι: αν $\zeta \sim \varepsilon$ και $\zeta_i \neq \varepsilon_i$ τότε $w_A(\zeta) = w_A(\varepsilon)$ αν $i \notin A$ και $w_A(\zeta) = -w_A(\varepsilon)$ αν $i \in A$. Έπεται ότι

$$L(w_A) = -|A| w_A,$$

δηλαδή οι συναρτήσεις Walsh είναι ιδιοσυναρτήσεις της διακριτής Λαπλασιανής.

Θεώρημα 6.7.1 (λογαριθμική ανισότητα Sobolev)**.** *Για κάθε $f : E_2^n \to \mathbb{R}$ ισχύει η ανισότητα*

$$\mathrm{Ent}(f^2) \leqslant 2E(f).$$

Για την απόδειξη θα χρειαστούμε μια «ανισότητα υπερσυσταλτότητας» που βασίζεται στην ανισότητα της Bonami.

Θεώρημα 6.7.2. *Για κάθε $x, y \in \mathbb{R}$ και $1 < p < \infty$ ορίζουμε*

$$F_p(x,y) = \left(\frac{1}{2}|x + r_p y|^p + \frac{1}{2}|x - r_p y|^p\right)^{1/p},$$

όπου

$$r_p := \frac{1}{\sqrt{p-1}}.$$

Τότε, η $F_p(x,y)$ είναι φθίνουσα συνάρτηση του p στο $(1,\infty)$.

Απόδειξη. Έστω $x, y \in \mathbb{R}$ και $1 < p < q < \infty$. Η ανισότητα $F_q(x,y) \leqslant F_p(x,y)$ είναι ομογενής, οπότε μπορούμε να υποθέσουμε ότι $x = 1$. Διακρίνουμε τρεις περιπτώσεις:

(α) Υποθέτουμε πρώτα ότι $1 < p < q \leqslant 2$ και ότι $0 \leqslant |r_p y| \leqslant 1$. Χρησιμοποιώντας το διωνυμικό θεώρημα βλέπουμε ότι

$$\frac{1}{2}|1 + r_q y|^q + \frac{1}{2}|1 - r_q y|^q = 1 + \sum_{k=1}^{\infty}\binom{q}{2k}\left(\frac{y^2}{q-1}\right)^k.$$

Χρησιμοποιώντας την ανισότητα $(1+x)^\alpha \leqslant 1+\alpha x$ (η οποία ισχύει για $0<\alpha\leqslant 1$ και $x\geqslant 0$) με $\alpha = p/q$, παίρνουμε

$$\left(1+\sum_{k=1}^{\infty}\binom{q}{2k}\left(\frac{y^2}{q-1}\right)^k\right)^{p/q} \leqslant 1+\frac{p}{q}\sum_{k=1}^{\infty}\binom{q}{2k}\left(\frac{y^2}{q-1}\right)^k.$$

Έπεται ότι

$$\begin{aligned} F_q(1,y) &= \left(1+\sum_{k=1}^{\infty}\binom{q}{2k}\left(\frac{y^2}{q-1}\right)^k\right)^{1/q} \\ &\leqslant \left(1+\frac{p}{q}\sum_{k=1}^{\infty}\binom{q}{2k}\left(\frac{y^2}{q-1}\right)^k\right)^{1/p}. \end{aligned}$$

Παρατηρούμε ότι

$$\begin{aligned} \frac{p}{q}\binom{q}{2k}\left(\frac{1}{q-1}\right)^k &= \frac{p}{q}\,\frac{q(q-1)\cdots(q-2k+1)}{(2k)!\,(q-1)^k} \\ &= \frac{p(q-2)\cdots(q-2k+1)}{(2k)!\,(q-1)^{k-1}} \\ &= \frac{p(2-q)\cdots(2k-1-q)}{(2k)!\,(q-1)^{k-1}} \\ &\leqslant \frac{p(2-p)\cdots(2k-1-p)}{(2k)!\,(p-1)^{k-1}} \\ &= \binom{p}{2k}\left(\frac{1}{p-1}\right)^k. \end{aligned}$$

[Παρατηρήστε ότι $(q-2)\cdots(q-2k+1) = (2-q)\cdots(2k-1-q)$ γιατί το πλήθος των όρων στο γινόμενο είναι άρτιο, και ότι $\frac{(2-q)\cdots(2k-1-q)}{(q-1)^{k-1}} \leqslant \frac{(2-p)\cdots(2k-1-p)}{(p-1)^{k-1}}$ γιατί $p<q$.] Επιστρέφοντας στην εκτίμηση για την $F_q(x,y)$ παίρνουμε

$$\begin{aligned} F_q(1,y) &\leqslant \left(1+\frac{p}{q}\sum_{k=1}^{\infty}\binom{q}{2k}\left(\frac{y^2}{q-1}\right)^k\right)^{1/p} \\ &\leqslant \left(1+\sum_{k=1}^{\infty}\binom{p}{2k}\left(\frac{y^2}{p-1}\right)^k\right)^{1/p} \\ &= F_p(1,y). \end{aligned}$$

(β) Υποθέτουμε τώρα ότι $1 < p < q \leqslant 2$ και ότι $|r_p y| \geqslant 1$. Θέτουμε $\lambda = r_q/r_p$ και $\mu = 1/|r_p y|$. Γράφουμε

$$\begin{aligned} F_q(1,y) &= \left(\frac{1}{2}|1+\lambda r_p y|^q + \frac{1}{2}|1-\lambda r_p y|^q\right)^{1/q} \\ &= \frac{1}{\mu}\left(\frac{1}{2}|\lambda+\mu|^q + \frac{1}{2}|\lambda-\mu|^q\right)^{1/q}. \end{aligned}$$

Παρατηρήστε ότι $0 < \lambda, \mu \leqslant 1$. Άρα, $|\lambda-\mu| \leqslant 1-\lambda\mu$ και $\lambda+\mu \leqslant 1+\lambda\mu$. Μπορούμε λοιπόν να γράψουμε

$$F_q(1,y) \leqslant \frac{1}{\mu}\left(\frac{1}{2}|1+\lambda\mu|^q + \frac{1}{2}|1-\lambda\mu|^q\right)^{1/q}.$$

Έχουμε $\lambda\mu = r_q z$, όπου $z = \frac{1}{r_p|r_p y|}$ και $|r_p z| = \frac{1}{|r_p y|} \leqslant 1$. Από το (α) συμπεραίνουμε ότι

$$\begin{aligned} \frac{1}{\mu}\left(\frac{1}{2}|1+\lambda\mu|^q + \frac{1}{2}|1-\lambda\mu|^q\right)^{1/q} &= \frac{1}{\mu}\left(\frac{1}{2}|1+r_q z|^q + \frac{1}{2}|1-r_q z|^q\right)^{1/q} \\ &\leqslant \frac{1}{\mu}\left(\frac{1}{2}|1+r_p z|^p + \frac{1}{2}|1-r_p z|^p\right)^{1/p} \\ &= \frac{1}{\mu}\left(\frac{1}{2}|1+\mu|^p + \frac{1}{2}|1-\mu|^p\right)^{1/p} \\ &= \left(\frac{1}{2}\left|1+\frac{1}{\mu}\right|^p + \frac{1}{2}\left|1-\frac{1}{\mu}\right|^p\right)^{1/p} \\ &= \left(\frac{1}{2}\,|1+r_p y|^p + \frac{1}{2}\,|1-r_p y|^p\right)^{1/p} \\ &= F_p(1,y). \end{aligned}$$

Δηλαδή,

$$F_q(1,y) \leqslant F_p(1,y).$$

(γ) Τέλος, υποθέτουμε ότι $2 \leqslant p < q < \infty$. Θα χρησιμοποιήσουμε δυϊσμό. Θεωρούμε τους συζυγείς εκθέτες p' και q' των p και q. Παρατηρήστε ότι $1 < q' < p' \leqslant 2$. Αν $\lambda = r_{p'}/r_{q'} = \frac{\sqrt{q'-1}}{\sqrt{p'-1}}$ και αν $\kappa(1,1) = \kappa(-1,-1) = 1+\lambda$ και $\kappa(1,-1) = \kappa(-1,1) = 1-\lambda$, τότε τα (α) και (β) δείχνουν ότι για τον τελεστή $T : L^{q'}(E_2^1) \to L^{p'}(E_2^1)$, $1 < p' < q' \leqslant 2$, που ορίζεται μέσω της

$$T(f)(\epsilon) = \int_{E_2^1} \kappa(\epsilon,\zeta) f(\zeta)\, d\zeta,$$

ισχύει η ανισότητα

$$\|T(f)\|_{p'} \leqslant \|f\|_{q'}.$$

Η κ είναι συμμετρική στον $E_2^1 \times E_2^1$, άρα $T^* = T$, όπου $T^* : L^p(E_2^1) \to L^q(E_2^1)$ είναι ο συζυγής τελεστής του T. Από την $\|T^*\| = \|T\|$ έπεται ότι

$$\|T(f)\|_q \leqslant \|f\|_p$$

για κάθε $f : E_2^1 \to \mathbb{R}$. Παρατηρώντας ότι

$$\frac{q'-1}{p'-1} = \frac{p-1}{q-1}$$

έχουμε το ζητούμενο. □

Η ανισότητα του Θεωρήματος 6.7.2 επεκτείνεται εύκολα στην περίπτωση που τα x, y είναι διανύσματα σε έναν χώρο με νόρμα.

Θεώρημα 6.7.3. *Έστω $(X, \|\cdot\|)$ ένας χώρος με νόρμα. Για κάθε $x, y \in X$ και $1 < p < \infty$ ορίζουμε*

$$F_p(x,y) = \left(\frac{1}{2}\|x + r_p y\|^p + \frac{1}{2}\|x - r_p y\|^p\right)^{1/p},$$

όπου

$$r_p := \frac{1}{\sqrt{p-1}}.$$

Τότε, η $F_p(x,y)$ είναι φθίνουσα συνάρτηση του p στο $(1,\infty)$.

Θα χρειαστούμε το εξής λήμμα.

Λήμμα 6.7.4. *Έστω $(X, \|\cdot\|)$ ένας χώρος με νόρμα. Αν $x, z \in X$ και $-1 \leqslant \lambda < 1$, τότε*

$$\|x + \lambda z\| \leqslant \frac{1}{2}\big(\|x+z\| + \|x-z\|\big) + \frac{\lambda}{2}\big(\|x+z\| - \|x-z\|\big).$$

Απόδειξη. Γράφουμε

$$x + \lambda z = \left(\frac{1+\lambda}{2}\right)(x+z) + \left(\frac{1-\lambda}{2}\right)(x-z),$$

οπότε

$$\begin{aligned}\|x + \lambda z\| &\leqslant \left(\frac{1+\lambda}{2}\right)\|x+z\| + \left(\frac{1-\lambda}{2}\right)\|x-z\| \\ &= \frac{1}{2}\big(\|x+z\| + \|x-z\|\big) + \frac{\lambda}{2}\big(\|x+z\| - \|x-z\|\big).\end{aligned}$$

□

Απόδειξη του Θεωρήματος 6.7.3. Έστω $1 < p < q < \infty$. Θέτουμε $u = x + r_p y$, $v = x - r_p y$ και $\lambda = r_q / r_p$. Παρατηρήστε ότι $0 < \lambda < 1$, οπότε το λήμμα μας δίνει

$$\begin{aligned}\|x + \lambda r_p y\| &\leqslant \frac{1}{2}\big(\|x + r_p y\| + \|x - r_p y\|\big) + \frac{\lambda}{2}\big(\|x + r_p y\| - \|x - r_p y\|\big)\\ &= \frac{1}{2}\big(\|u\| + \|v\|\big) + \frac{\lambda}{2}\big(\|u\| - \|v\|\big)\end{aligned}$$

και, εντελώς ανάλογα,

$$\|x - \lambda r_p y\| \leqslant \frac{1}{2}\big(\|u\| + \|v\|\big) - \frac{\lambda}{2}\big(\|u\| - \|v\|\big).$$

Μπορούμε λοιπόν να γράψουμε

$$\left(\frac{1}{2}\|x + r_q y\|^q + \frac{1}{2}\|x - r_q y\|^q\right)^{1/q} = \left(\frac{1}{2}\|x + \lambda r_p y\|^q + \frac{1}{2}\|x - \lambda r_p y\|^q\right)^{1/q}$$

$$\begin{aligned}&\leqslant \left(\frac{1}{2}\left(\frac{1}{2}\big(\|u\| + \|v\|\big) + \frac{\lambda}{2}\big(\|u\| - \|v\|\big)\right)^q + \frac{1}{2}\left(\frac{1}{2}\big(\|u\| + \|v\|\big) - \frac{\lambda}{2}\big(\|u\| - \|v\|\big)\right)^q\right)^{1/q}\\ &\leqslant \left(\frac{1}{2}\left(\frac{1}{2}\big(\|u\| + \|v\|\big) + \frac{1}{2}\big(\|u\| - \|v\|\big)\right)^p + \frac{1}{2}\left(\frac{1}{2}\big(\|u\| + \|v\|\big) - \frac{1}{2}\big(\|u\| - \|v\|\big)\right)^p\right)^{1/p}\\ &= \left(\frac{1}{2}\|u\|^p + \frac{1}{2}\|v\|^p\right)^{1/p}\\ &= \left(\frac{1}{2}\|x + r_p y\|^p + \frac{1}{2}\|x - r_p y\|^p\right)^{1/p},\end{aligned}$$

όπου, για την δεύτερη ανισότητα, χρησιμοποιήσαμε το Θεώρημα 6.7.2 με $x = \frac{\|u\|+\|v\|}{2}$ και $y = \frac{1}{2r_p}\big(\|u\| - \|v\|\big)$. □

Θεώρημα 6.7.5. *Έστω $1 < p < q < \infty$ και έστω $\{x_A : A \subseteq \{1, \ldots, N\}\}$ μια οικογένεια διανυσμάτων σε ένα χώρο με νόρμα $(X, \|\cdot\|)$. Τότε,*

$$\left\|\sum_A r_q^{|A|} w_A x_A\right\|_{L^q(X)} \leqslant \left\|\sum_A r_p^{|A|} w_A x_A\right\|_{L^p(X)},$$

όπου w_A είναι οι συναρτήσεις Walsh.

Απόδειξη. Με επαγωγή ως προς N. Για $N = 1$ το ζητούμενο είναι ακριβώς η ανισότητα που αποδείξαμε στο Θεώρημα 6.7.2 (θεωρούμε $x_\emptyset = x$, $x_{\{1\}} = y$ και $w_{\{1\}}(-1) = 1$, $w_{\{1\}}(1) = 1$.)

Υποθέτουμε ότι το αποτέλεσμα ισχύει για $k = N-1$ και γράφουμε $E_2^N = E_2^{N-1} \times E_2^1$, $\mu_N = \mu_{N-1} \times \mu_{\{N\}}$. Γράφουμε $\mathcal{P}(N-1)$ για το σύνολο των υποσυνόλων του $\{1, \dots, N-1\}$ και $\mathcal{P}(N)$ για το σύνολο των υποσυνόλων του $\{1, \dots, N\}$. Για κάθε $B \in \mathcal{P}(N-1)$ ορίζουμε $B^+ = B \cup \{N\}$. Με αυτό τον συμβολισμό, $\mathcal{P}(N) = \mathcal{P}(N-1) \cup \{B^+ : B \in \mathcal{P}(N-1)\}$. Γράφουμε

$$\sum_A r_p^{|A|} w_A(\omega) x_A = u_p(\varepsilon) + \varepsilon_N(\eta) r_p v_p(\varepsilon),$$

όπου $\varepsilon \in E_2^{N-1}$, $\eta \in E_2^{\{N\}} = \{-1, +1\}$ και $\omega = (\varepsilon, \eta) \in E_2^N$. Ορίζουμε

$$u_p = \sum_{B \in \mathcal{P}(N-1)} r_p^{|B|} w_B x_B \quad \text{και} \quad v_p = \sum_{B \in \mathcal{P}(N-1)} r_p^{|B|} w_B x_{B^+}.$$

Εντελώς ανάλογα,

$$\sum_A r_q^{|A|} w_A x_A = u_q + \varepsilon_N r_q v_q,$$

όπου

$$u_q = \sum_{B \in \mathcal{P}(N-1)} r_q^{|B|} w_B x_B \quad \text{και} \quad v_q = \sum_{B \in \mathcal{P}(N-1)} r_q^{|B|} w_B x_{B^+}.$$

Επομένως αρκεί να δείξουμε ότι

$$\|u_q + r_q \varepsilon_N v_q\|_{L^q} \leqslant \|u_p + r_p \varepsilon_N v_p\|_{L^p}.$$

Από την επαγωγική μας υπόθεση ισχύει

$$\begin{aligned}
\|u_q + r_q \varepsilon_N v_q\|_{L^q} &= \big(\mathbb{E}_{N-1}(\|u_q + r_q \varepsilon_N v_q\|^q)\big)^{1/q} \\
&= \left(\mathbb{E}_{N-1} \| \sum_{B \in \mathcal{P}(N-1)} r_q^{|B|} w_B (x_B + \varepsilon_N r_q x_{B^+}) \|^q \right)^{1/q} \\
&\leqslant \left(\mathbb{E}_{N-1} \| \sum_{B \in \mathcal{P}(N-1)} r_p^{|B|} w_B (x_B + \varepsilon_N r_q x_{B^+}) \|^p \right)^{1/p} \\
&= \big(\mathbb{E}_{N-1}\|u_q + r_q \varepsilon_N v_p\|^p\big)^{1/p}.
\end{aligned}$$

Σε αυτό το σημείο θα χρησιμοποιήσουμε την παρακάτω ανισότητα: Αν $(\Omega_1, \mathcal{A}_1, \mu_1)$ και

$(\Omega_2, \mathcal{A}_2, \mu_2)$ είναι χώροι μέτρου, $f : \Omega_1 \times \Omega_2 \to [0, +\infty]$, και $0 < p \leqslant q < +\infty$, τότε

$$\left(\int_{\Omega_1}\left(\int_{\Omega_2} f(x,y)^p \, d\mu(y)\right)^{q/p} d\mu_1(x)\right)^{1/q} \leqslant \left(\int_{\Omega_2}\left(\int_{\Omega_1} f(x,y)^q \, d\mu_1(x)\right)^{p/q} d\mu_2(y)\right)^{1/p}.$$

Άρα αν $\varepsilon \in E_2^{N-1}, \eta \in \{-1, +1\}$, τότε

$$\begin{aligned}
\|u_q(\varepsilon) + \varepsilon_N(\eta) r_q v_q(\varepsilon)\|_{L^q} &= \left(\mathbb{E}_{\{N\}}(\mathbb{E}_{N-1}\|u_q(\varepsilon) + \varepsilon_N(\eta) r_q v_q(\varepsilon)\|^q)\right)^{1/q} \\
&\leqslant \left(\mathbb{E}_{\{\mathbb{N}\}}(\mathbb{E}_{N-1}\|u_p(\varepsilon) + \varepsilon_N(\eta) r_q v_p(\varepsilon)\|^p)^{q/p}\right)^{1/q} \\
&\leqslant \left(\mathbb{E}_{N-1}(\mathbb{E}_{\{N\}}\|u_p(\varepsilon) + \varepsilon_N(\eta) r_q v_p(\varepsilon)\|^q)^{p/q}\right)^{1/p} \\
&\leqslant \left(\mathbb{E}_{N-1}\,\mathbb{E}_{\{N\}}\|u_p(\varepsilon) + \varepsilon_N(\eta) r_p v_p(\varepsilon)\|^p\right)^{1/p} \\
&= \left(\mathbb{E}_N\|u_p(\varepsilon) + \varepsilon_N(\eta) r_p v_p(\varepsilon)\|^p\right)^{1/p} \\
&= \left\| \sum_{A \in \mathcal{P}(N)} r_p^{|A|} w_A x_A \right\|_{L^p}
\end{aligned}$$

όπου η τελευταία ανισότητα προέκυψε από την περίπτωση $k = 1$. □

Ορισμός 6.7.6. Γράφουμε $C(E_2^n)$ για το χώρο όλων των συναρτήσεων $f : E_2^n \to \mathbb{R}$. Για κάθε $t \geqslant 0$ θεωρούμε τον τελεστή $P_t : C(E_2^n) \to C(E_2^n)$ με

$$P_t(f) = (e^{tL})(f) = \sum_{j=0}^{\infty} \frac{t^j L^j(f)}{j!},$$

όπου $L^j = L \circ \cdots \circ L$ (j φορές). Παρατηρούμε ότι, για κάθε $A \subseteq \{1, \ldots, n\}$,

$$L^j(w_A) = (-1)^j |A|^j w_A$$

(αυτό προκύπτει επαγωγικά από την $L(w_A) = -|A| w_A$). Συνεπώς,

$$P_t(w_A) = \sum_{j=0}^{\infty} \frac{(-1)^j t^j |A|^j w_A}{j!} = (e^{-t|A|})(w_A).$$

Τώρα, άμεση εφαρμογή του θεωρήματος 6.7.5 μας δίνει το εξής:

Θεώρημα 6.7.7 (υπερσυσταλτότητα στο διακριτό κύβο)**.** *Έστω* $1 < p < \infty$ *και* $t \geqslant 0$. *Θέτουμε*

$$q(t) = 1 + (p-1)e^{2t}.$$

Τότε, για κάθε $f : E_2^n \to \mathbb{R}$,

$$\|P_t(f)\|_{q(t)} \leqslant \|f\|_p.$$

Απόδειξη. Παρατηρούμε ότι αν $f = \sum_A \hat{f}_a w_A$, τότε

$$\begin{aligned} P_t f = \sum_A \hat{f}_A e^{-t|A|} w_A &= \sum_A r_p^{|A|} e^{-t|A|} w_A \frac{\hat{f}_A}{r_p^{|A|}} \\ &= \sum_A r_{q(t)}^{|A|} e^{-t|A|} w_A \frac{\hat{f}_A}{r_p^{|A|}}, \end{aligned}$$

αφού $r_{q(t)} = e^{-t|A|} r_p^{|A|}$. Άρα από την γενικευμένη ανισότητα της Bonami (θεώρημα 6.7.7) έχουμε ότι

$$\begin{aligned} \|P_t f\|_{q(t)} &= \left\| \sum_A r^{q(t)} w_A \frac{\hat{f}_A}{r_p^{|A|}} \right\|_{q(t)} \\ &\leqslant \left\| \sum_A r_p^{|A|} w_A \frac{\hat{f}_A}{r_p^{|A|}} \right\|_p = \|f\|_p. \end{aligned}$$

□

Απόδειξη της λογαριθμικής ανισότητας Sobolev στον διακριτό κύβο. Παίρνουμε $p = 2$ και, για κάθε $t \geqslant 0$, θεωρούμε τον $q(t) = 1 + e^{2t}$. Από την $P_t(w_A) = (e^{-t|A|})(w_A)$ βλέπουμε ότι

$$\frac{dP_t(w_A)}{dt} = -|A|(e^{-t|A|})(w_A) = (L \circ P_t)(w_A),$$

άρα

$$\frac{dP_t(f)}{dt} = (L \circ P_t)(f)$$

για κάθε $f : E_2^n \to \mathbb{R}$. Ας υποθέσουμε ότι $\|f\|_2 = 1$. Από το Θεώρημα 6.7.5 έχουμε

$$\|P_t(f)\|_{q(t)} \leqslant 1,$$

δηλαδή

$$\frac{d}{dt}\left(\mathbb{E}\left[(P_t(f))^{q(t)}\right]\right) \leqslant 0$$

στο σημείο $t=0$. Όμως,

$$\begin{aligned}\frac{d}{dt}\,(P_t(f))^{q(t)} &= [P_t(f)]^{q(t)}\,\frac{d}{dt}\Big(\ln[P_t(f)]^{q(t)}\Big)\\ &= [P_t(f)]^{q(t)}\,\frac{d}{dt}\,(q(t)\ln(P_t(f)))\\ &= [P_t(f)]^{q(t)}\,q'(t)\ln(P_t(f)) + [P_t(f)]^{q(t)-1}q(t)(L\circ P_t)(f)\\ &= 2e^{2t}[P_t(f)]^{q(t)}\,\ln(P_t(f)) + (1+e^{2t})[P_t(f)]^{q(t)-1}(L\circ P_t)(f).\end{aligned}$$

Παίρνοντας μέση τιμή στην παραπάνω ανισότητα και θέτοντας $t=0$ βλέπουμε ότι

$$\mathrm{Ent}(f^2) - 2E(f) = \mathbb{E}\,(f^2\ln(f^2)) + 2\mathbb{E}\,(fL(f)) \leqslant 0.$$

Αυτό αποδεικνύει το θεώρημα. □

6.8 Ασκήσεις

1. Έστω L ο γεννήτορας της ημιομάδας Ornstein-Uhlenbeck στον $\mathbb{R}^n$. Αν f είναι μια λεία συνάρτηση στον $\mathbb{R}^n$, δείξτε ότι

$$\frac{1}{2}L(|\nabla f|^2) - \langle \nabla f, \nabla(Lf)\rangle \geqslant |\nabla f|^2.$$

Για κάθε $t\geqslant 0$ θέτουμε $\alpha(t) = \mathrm{Ent}_{\gamma_n}(T_tf)$. Σταθεροποιούμε $t\geqslant 0$ και θέτουμε $F=\ln T_tf$. Δείξτε ότι

$$\alpha''(t) = 2\int\langle\nabla F, \nabla L(\ln F)\rangle d\gamma_n - \int L(|\nabla\ln F|^2)F\,d\gamma_n.$$

Δείξτε ότι $\alpha''(t)\leqslant -2\alpha'(t)$ και συμπεράνατε την λογαριθμική ανισότητα Sobolev.

2. Έστω ν το εκθετικό μέτρο πιθανότητας στο $\mathbb{R}$ με πυκνότητα $\frac{1}{2}e^{-|x|}$, $x\in\mathbb{R}$. Δείξτε ότι: για κάθε $0<\rho<1$ και για κάθε Lipschitz συνεχή συνάρτηση $f:\mathbb{R}\to\mathbb{R}$ με $|f'|\leqslant\rho<1$ σχεδόν παντού, ισχύει

$$\mathrm{Ent}_\nu(e^f) \leqslant \frac{2}{1-\rho}\int (f')^2 e^f\,d\nu.$$

3. Έστω $f:(X,\mathcal{A},\mu)\to\mathbb{R}$ μετρήσιμη συνάρτηση τέτοια ώστε $\int_X f^2\ln(1+f^2)\,d\mu<\infty$. Αποδείξτε ότι, για κάθε $a\in\mathbb{R}$,

$$\mathrm{Ent}_\mu((f+a)^2) \leqslant \mathrm{Ent}_\mu(f^2) + 2\int_X f^2\,d\mu.$$

[*Υπόδειξη:* Θεωρήστε τη συνάρτηση $\psi(r) = \mathrm{Ent}_\mu((rf+a)^2) - \mathrm{Ent}_\mu((rf)^2) - 2\int_X (rf)^2 d\mu$.]

4. Έστω $\{P_t\}_{t\geqslant 0}$ ημιομάδα Markov με γεννήτορα L και αναλλοίωτο μέτρο μ. Αποδείξτε ότι: για κάθε $f \in D(\mathcal{E})$ και για κάθε $t \geqslant 0$,

$$\int f^2 \ln(f^2)\, d\mu \leqslant 2t\mathcal{E}(f,f) + \int f^2 \ln(P_t f)^2\, d\mu.$$

5. Έστω $\{P_t\}_{t\geqslant 0}$ ημιομάδα Markov με γεννήτορα L και αναλλοίωτο μέτρο μ. Αποδείξτε ότι: αν το μ ικανοποιεί τη λογαριθμική ανισότητα Sobolev με σταθερά β τότε, για κάθε $p \geqslant 2$ και για κάθε $f \in L^p(\mu)$,

$$\|f\|_p^2 \leqslant \|f\|_2^2 + \frac{p-2}{\beta}\left(\int \Gamma(f)^{p/2} d\mu\right)^{2/p}.$$

6. Έστω $\{P_t\}_{t\geqslant 0}$ ημιομάδα Markov με γεννήτορα L και αναλλοίωτο μέτρο μ. Αποδείξτε ότι: αν το μ ικανοποιεί τη λογαριθμική ανισότητα Sobolev με σταθερά β τότε, για κάθε $1 \leqslant p < 2$ και για κάθε $f \in D(\mathcal{E})$,

$$\|f\|_2^2 \ln\left(\frac{\|f\|_2^2}{\|f\|_p^2}\right) \leqslant \frac{2-p}{p\beta}\, \mathcal{E}(f,f).$$

7. Έστω $\{P_t\}_{t\geqslant 0}$ ημιομάδα Markov με γεννήτορα L και αναλλοίωτο μέτρο μ. Λέμε ότι το μ ικανοποιεί την ανισότητα Sobolev με εκθέτη $p > 2$ και σταθερές $\alpha \in \mathbb{R}$, $\gamma > 0$, αν για κάθε $f \in D(\mathcal{E})$

$$\|f\|_p^2 \leqslant \alpha \|f\|_2^2 + \gamma \mathcal{E}(f,f).$$

Αποδείξτε ότι αν το μ ικανοποιεί την ανισότητα Sobolev με εκθέτη n και σταθερές $\alpha \geqslant 0$, $\gamma > 0$, τότε για κάθε $f \in D(\mathcal{E})$ με $\int f^2 d\mu = 1$ ισχύει

$$\mathrm{Ent}_\mu(f^2) \leqslant \frac{n}{2} \ln(\alpha + \gamma \mathcal{E}(f,f)).$$

Επίσης,

$$\|f\|_2^{n+2} \leqslant \left(\alpha \|f\|_2^2 + \gamma \mathcal{E}(f,f)\right)^{n/2} \|f\|_1^2.$$

8. Έστω $f : \mathbb{R}^n \to \mathbb{R}$ Lebesgue μετρήσιμη συνάρτηση. Υποθέτουμε ότι $\int f^2 d\lambda(x) = 1$ και $f \in D(\mathcal{E})$, όπου $\mathcal{E}(f,f) = \int_{\mathbb{R}^n} |\nabla f|^2 d\lambda(x)$ και λ είναι το μέτρο Lebesgue στον $\mathbb{R}^n$. Αποδείξτε ότι

$$\mathrm{Ent}_\lambda(f^2) \leqslant \frac{n}{2} \ln\left(\frac{2}{n\pi e}\mathcal{E}(f,f)\right).$$

9. Έστω $g : \mathbb{R}^n \to \mathbb{R}$ ομαλή, γνήσια θετική συνάρτηση με $\int g\, d\gamma_n = 1$ και τέτοια ώστε οι Δg και $g\Delta(\ln g)$ να είναι γ_n-ολοκληρώσιμες. Αποδείξτε ότι

$$\mathrm{Ent}_{\gamma_n}(g) \leqslant \frac{1}{2}\int \Delta g\, d\gamma_n + \frac{n}{2} \ln\left(1 - \frac{1}{n}\int g\Delta(\ln g)\, d\gamma_n\right).$$

10. Αποδείξτε ότι: για κάθε $a, b \in \mathbb{R}$,

$$\frac{a^2}{2}\ln a^2 + \frac{b^2}{2}\ln b^2 - \frac{a^2+b^2}{2}\ln\frac{a^2+b^2}{2} \leqslant \frac{1}{2}(a-b)^2.$$

11. Ξεκινώντας από την προηγούμενη ανισότητα αποδείξτε τη λογαριθμική ανισότητα Sobolev για τον διακριτό κύβο.

12. Αποδείξτε ότι: για κάθε $f : \{-1,1\}^n \to \mathbb{R}$ και για κάθε $p \in (0,1)$,

$$\mathrm{Ent}_{\mu_{n,p}}(f^2) \leqslant c_p \mathcal{E}(f,f),$$

όπου

$$c_p := \frac{1}{1-2p}\ln\frac{1-p}{p}$$

και $\mu_{n,p} = \mu_{1,p} \otimes \cdots \otimes \mu_{1,p}$, όπου $\mu_{1,p}(\{1\}) = p$ και $\mu_{1,p}(\{-1\}) = 1-p$. Παρατηρήστε ότι $\lim\limits_{p\to\frac{1}{2}} c_p = 2$.