Lecture 9: “Random Walks”

Sotiris Nikoletseas
Professor

CEID - ETY Course
2017 - 2018

Overview

A. The reflection principle

B. Random walks on the line (1D)

C. Random walks on the plane (2D)

D. Random walks on Z^3 (3D)

A. Το αξίωμα της αντανάκλασης (The reflection principle)

Εντυπωσιακό:

- με απλό τρόπο (απλά λογικά επιχειρήματα, απλή συνδυαστική, χωρίς αναλυτικές μεθόδους)
- "βαθιά"αποτελέσματα (που με αναλυτικές μεθόδους αποδεικνύονται με ιδιαίτερη δυσκολία)

A. Το αξίωμα της αντανάκλασης

Παράδειγμα: (The ballot theorem)

Θεώρημα: ψηφοφορία 2 υποψηφίων

- 1ος: n ψήφους
- 2ος: m ψήφους
- n $>$ m $\Rightarrow$ τελικός νικητής ο 1ος

Pr{ στην καταμέτρηση προηγείται συνεχώς ο τελικός νικητής} =
$$\frac{n-m}{n+m}$$

Απόδειξη:

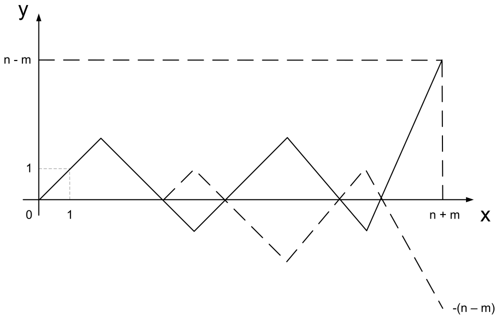
#(Όλων των τρόπων να εξελιχθεί η καταμέτρηση)= $\binom{n+m}{n}$

(απόδειξη με δύο τρόπους)

Απόδειξη α)

α) Παριστάνουμε γραφικά την εξέλιξη της διαφοράς των ψήφων του 1ου από το 2ο: Συσσωρευτικό άθροισμα:

- +1, ψήφος του 1ου
- -1, ψήφος του 2ου



κάθε καταμέτρηση αντιστοιχεί σε ένα μονοπάτι (μία τροχιά) από το (0,0) στο (n+m, n-m)

Απόδειξη β)

Παρατήρηση:
Οι ευνοϊκές τροχιές περνάνε υποχρεωτικά από το (1, 1).

β1 Όλες οι τροχιές από (1, 1) σε (n+m, n-m) είναι

$$\binom{n+m-1}{m} = \binom{n+m-1}{n-1}$$

β2 Δεν είναι όμως όλες αυτές οι τροχιές ευνοϊκές. Υπάρχουν ΠΡΟΒΛΗΜΑΤΙΚΕΣ τροχιές.

- *Ποιές είναι οι προβληματικές τροχιές;*
  Απάντηση: Όσες τέμνουν τον άξονα y = 0.

Πόσες είναι οι προβληματικές τροχιές;

Σύμφωνα με το αξίωμα της αντανάκλασης, είναι τόσες όσες και οι συμμετρικές τους (περί τον άξονα των x).

- Κάθε τέτοια "συμμετρική"τροχιά έχει ως τελικό σημείο το (n+m, m-n)
- Έστω x "άνοδοι"(+1) σε αυτές τις συμμετρικές τροχιές. Έστω y "κάθοδοι"(-1). Η διαφορα "καθόδων"και "ανόδων"ισούται με την απόλυτη μεταβολή της τεταγμένης μεταξύ των σημείων (1, 1) και (n+m, m-n).

Πόσες είναι οι προβληματικές τροχιές; (Συνέχεια)

x + y = n + m - 1 **(1)**

y - x = n - m + 1 **(2)**

(Οι κάθοδοι είναι τόσες παραπάνω όση η απόλυτη μεταβολή της τεταγμένης.) Από **(1)** και **(2)** προκύπτει ότι $\Rightarrow$

- y = n
- x = m - 1
- Άρα προβληματικές είναι όσες έχουν x = m - 1 ανόδους και y = n καθόδους,

$$\text{δηλαδή} \binom{n+m-1}{m-1}$$

Τελικό Συμπέρασμα

Συνολικά

$$P(\text{προηγείται συνεχώς ο 1ος}) = \frac{\binom{n+m-1}{m} - \binom{n+m-1}{m-1}}{\binom{n+m}{n}} =$$

$$= \frac{\frac{(n+m-1)!}{m!\cdot(n-1)!} - \frac{(n+m-1)!}{n!\cdot m!}}{\frac{(n+m)!}{m!\cdot n!}} =$$

$$= \frac{\frac{(n+m-1)!}{m!\cdot n!}\left[n-m\right]}{\frac{(n+m)!}{m!\cdot n!}} = \frac{n-m}{n+m}$$

B. Τυχαίοι Περίπατοι στην ευθεία

Ορισμός

- Έστω στοχαστικά ανεξάρτητες επαναλήψεις πειράματος με δύο ισοπίθανα αποτελέσματα: +1, -1, δηλαδή αν X_k (k $\geq$ 1) το αποτέλεσμα της k $-$οστής επανάληψης:

$$X_k = \begin{cases} +1, \text{ με πιθανότητα } 1/2 \\ -1, \text{ με πιθανότητα } 1/2 \end{cases}$$

- Έστω S_n (n $\geq$ 1) το συσσωρευτικό άθροισμα των n πρώτων αποτελεσμάτων:

$$S_n = X_1 + X_2 + \ldots + X_n$$

- Η ακολουθία $\{S_n\}_{n=1}^{\infty}$ (δηλαδή η $S_1, S_2, \ldots, S_n, \ldots$) καλείται (συμμετρικός) τυχαίος περίπατος (random walk) στην ευθεία.

B. Τυχαίοι Περίπατοι - Τι περιγράφουν (I)

α) Κίνηση σωματιδίου

- Ένα σωματίδιο αρχικά (την στιγμή t=0) βρίσκεται στην αφετηρία (στο σημείο 0 της ευθείας των πραγματικών αριθμών)
- Σε κάθε "βήμα"(διακριτή χρονική στιγμή t = 1, 2, ...) επιλέγει τυχαία και ισοπίθανα να κινηθεί μία μονάδα δεξιά ($X_t = 1$) ή αριστερά ($X_t = -1$).

$\Rightarrow S_n$: η θέση του σωματιδίου τη χρονική στιγμή $t = n$.

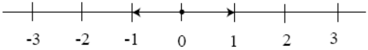
B. Τυχαίοι Περίπατοι - Τι περιγράφουν (I)

Γενίκευση: Τα X_k παίρνουν τιμές στο R^m.

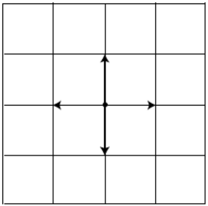
- $m = 2$ (επίπεδο): $X_k = \{$ βορράς, νότος, ανατολή, δύση$\}$ με πιθανότητα 1/4 για κάθε κίνηση. (π.χ. κίνηση ανθρώπου σε μία πόλη ρυμοτομημένη σε πλέγμα τετραγώνων.)

- $m = 3$ (χώρος): $X_k = \{$ αριστερά, δεξιά, εμπρός, πίσω, πάνω, κάτω$\}$ ισοπίθανα. (π.χ. κίνηση ανθρώπου σε ένα κτίριο.)

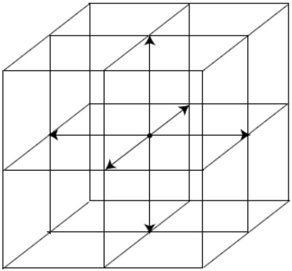
Β. Τυχαίοι Περίπατοι - Οπτικοποίηση κίνησης στο R^m



a. Random walk in one dimension.



b. Random walk in two dimensions.



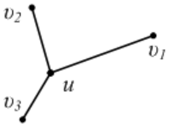
c. Random walk in three dimensions.

B. Τυχαίοι Περίπατοι - Τι περιγράφουν (II)

β) Τη διαφορά αποτελεσμάτων "κεφαλή"από τα αποτελέσματα "γράμματα"κατά τις αλλεπάλληλες ρίψεις ενός τίμιου νομίσματος. ($X_k = +1$ για το αποτέλεσμα "κεφαλή", αλλιώς $X_k = -1$)

γ) Το συσσωρευτικό κεφάλαιο ενός παίκτη ενός τυχερού παιχνιδιού ($X_k = +1$ αν νικάει στον γύρο k , αλλιώς $X_k = -1$)

δ) Τυχαίος περίπατος σε ένα γράφημα. π.χ.



βρισκόμαστε στην κορυφή u και επιλέγουμε τυχαία κάποια από τις γειτονικές της κορυφές v_1, v_2, v_3 κ.ο.κ

B. Τυχαίοι Περίπατοι - Μετρικές

Στο σημερινό μάθημα:

- η πιθανότητα επιστροφής στην αφετηρία σε μια ορισμένη χρονική στιγμή
- η πιθανότητα της πρώτης επίσκεψης στην αφετηρία σε μία χρονική στιγμή
- η πιθανότητα επιστροφής στην αφετηρία (κάποτε)

Άλλες ενδιαφέρουσες μετρικές:

- ο χρόνος ώστε ένας τυχαίος περίπατος σε ένα γράφημα να επισκεφθεί όλες τις κορυφές τουλάχιστον μία φορά την κάθε μια (cover time)
- ο μέσος χρόνος μεταξύ διαδοχικών επισκέψεων σε μία κορυφή

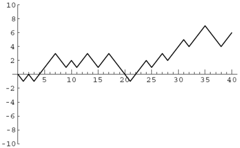
Β. Τυχαίοι Περίπατοι - Στην ευθεία

- Ορισμός:

$$X_k = \begin{cases} +1, \text{ με πιθανότητα } 1/2 \\ -1, \text{ με πιθανότητα } 1/2 \end{cases}$$

$$S_n = \sum_{k=1}^{n} X_k = X_1 + X_2 + \ldots + X_n$$

- Οπτικοποίηση: Τεθλασμένη γραμμή στο επίπεδο, όπου ο οριζόντιος άξονας είναι ο χρόνος (n) και ο κάθετος άξονας είναι το συσσωρευτικό άθροισμα (S_n).



A random walk of length 40.

- Ορισμός: $P_{n,r} = Pr\{S_n = r\}$

Θεώρημα

Η πιθανότητα επιστροφής στο 0 τη χρονική στιγμή 2n είναι:

$$u_{2n} = \frac{\binom{2n}{n}}{2^{2n}}$$

Αποδειξη: Όλα τα δυνατά μονοπάτια είναι 2^{2n} .
Τα ευνοϊκά μονοπάτια περιέχουν n τιμές +1 (ανόδους) (οπότε υποχρεωτικά και n καθόδους) οπότε είναι

$$\binom{2n}{n}$$

Λήμμα

$$u_{2n} \sim \frac{1}{\sqrt{\pi n}}$$

<u>Αποδειξη:</u>

$$u_{2n} = \frac{(2n)!}{(n!)(n!)} 2^{-2n} \sim \frac{\sqrt{2\pi 2n}\left(\frac{2n}{e}\right)^{2n}}{\left(\sqrt{2\pi n}\left(\frac{n}{e}\right)^{n}\right)^{2}} 2^{-2n} =$$

$$= \frac{2\sqrt{\pi n}\frac{2^{2n} n^{2n}}{e^{2n}}}{2\pi n \frac{n^{2n}}{e^{2n}}} 2^{-2n} = \frac{1}{\sqrt{\pi n}}$$

<u>Παρατήρηση:</u> n αυξάνει $\Rightarrow$ u_{2n} μικραίνει, δηλαδή όσο περνάει ο χρόνος η επιστροφή στην αφετηρία γίνεται λιγότερο πιθανή!

Ορισμός:

Ο περίπατος κάνει πρώτη επίσκεψη στο 0 τη χρονική στιγμή 2n αν και μόνο αν:

$$S_1 \neq 0, S_2 \neq 0, \ldots, S_{2n-1} \neq 0, S_{2n} = 0$$

Ορισμός:

f_{2n} = Pr{ πρώτη επίσκεψη στο 0 την χρονική στιγμή 2n }

Λήμμα (σχέση επισκέψεων – πρώτων επισκέψεων

$$u_{2n} = f_2 \cdot u_{2n-2} + f_4 \cdot u_{2n-4} + \ldots + f_{2n} \cdot u_0 \qquad (n \geq 1)$$

Λήμμα (σχέση επισκέψεων – πρώτων επισκέψεων

$$u_{2n} = f_2 \cdot u_{2n-2} + f_4 \cdot u_{2n-4} + \ldots + f_{2n} \cdot u_0 \qquad (n \geq 1)$$

Αποδειξη: u_{2n} = Pr $\{S_{2n} = 0\}$ =

=Pr{(1η επίσκεψη τη χ.σ. 2 ∩ επίσκεψη μετά από 2n-2 βήματα) ∪

∪ (1η επίσκεψη τη χ.σ. 4 ∩ επίσκεψη μετά από 2n-4 βήματα) ∪

⋮

∪ (1η επίσκεψη τη χ.σ. 2n ∩ επίσκεψη μετά από 0 βήματα) }

Και η απόδειξη ολοκληρώνεται αθροίζοντας τις πιθανότητες της ένωσης γεγονότων (είναι διακριτά) και πολλαπλασιάζοντας τις πιθανότητες κάθε τομής (λόγω ανεξαρτησίας, αφού ο περίπατος δεν έχει μνήμη). □

B. Τυχαίοι Περίπατοι

Λήμμα

Η πιθανότητα μη επιστροφής στην αφετηρία μέχρι και τη χρονική στιγμή 2n είναι ίση με την πιθανότητα επιστροφής τη χρονική στιγμή 2n:

$$Pr\{S_1 \neq 0, ..., S_{2n} \neq 0\} = Pr\{S_{2n} = 0\} = u_{2n}$$

Παρατήρηση:$Pr\{S_1 \neq 0, ..., S_{2n} \neq 0\}$ =
$= Pr\{S_1 > 0, ..., S_{2n} > 0\} + Pr\{S_1 < 0, ..., S_{2n} < 0\}$
Αλλά λόγω συμμετρίας οι δύο πιθανότητες (για θετικό, αρνητικό μονοπάτι) είναι ίσες. Άρα αρκεί να δείξουμε το εξής:

Λήμμα

$$Pr\{S_1 > 0, ..., S_{2n} > 0\} = \frac{1}{2} \cdot u_{2n}$$

Λήμμα

$$Pr\{S_1 > 0, ..., S_{2n} > 0\} = \frac{1}{2} \cdot u_{2n}$$

Απόδειξη: Είναι: $Pr\{S_1 > 0, ..., S_{2n} > 0\} =$

$$= \sum_{r=1}^{\infty} Pr\{S_1 > 0, ..., S_{2n-1} > 0, S_{2n} = 2r\}$$

(όπου προφανώς οι όροι r > n είναι μηδενικοί)

Αλλά: $Pr\{S_1 > 0, ..., S_{2n-1} > 0, S_{2n} = 2r\} =$

$$= Pr\{S_1 = 1\} \cdot Pr\{S_2 > 0, ..., S_{2n-1} > 0, S_{2n} = 2r | S_1 = 1\}$$

Έστω: $P = Pr\{S_2 > 0, ..., S_{2n-1} > 0, S_{2n} = 2r | S_1 = 1\}$

B. Τυχαίοι Περίπατοι

$$P = Pr\{S_2 > 0, ..., S_{2n-1} > 0, S_{2n} = 2r | S_1 = 1\}$$

- Η πιθανότητα P προκύπτει από τον αριθμό των μονοπατιών από το (1,1) στο (2n, 2r) που είναι “θετικά”(δεν ακουμπούν τον άξονα y =0).
- Τα μονοπάτια αυτά είναι όσα όλα τα μονοπάτια από το (1,1) στο (2n,2r) μείον το πλήθος των μονοπατιών που τέμνουν τον άξονα y =0.
- Από το αξίωμα της αντανάκλασης, τα μονοπάτια που τέμνουν τον άξονα y =0 είναι όσα και τα συμμετρικά τους περί αυτόν τον άξονα. Αυτά τα συμμετρικά μονοπάτια ξεκινούν από το (1,1) και καταλήγουν στο (2n, -2r) δηλαδή η τεταγμένη των ακραίων σημείων τους διαφέρει κατά απόλυτη τιμή κατά 2r+1.

B. Τυχαίοι Περίπατοι

- Επομένως (διαιρώντας όλους τους αριθμούς των μονοπατιών με τον συνολικό αριθμό 2^{2n-1} των μονοπατιών από το (1,1) στο (2n,2r)) είναι:

$$P = P_{2n-1,2r-1} - P_{2n-1,2r+1}$$

- Άρα $Pr\{S_1 > 0, ..., S_{2n} > 0\} =$

$$= \frac{1}{2}\sum_{r=1}^{n}(P_{2n-1,2r-1} - P_{2n-1,2r+1}) =$$

$$= \frac{1}{2}P_{2n-1,1}$$ (οι άλλοι όροι απαλείφονται)

B. Τυχαίοι Περίπατοι

Αλλά

$$u_{2n} = Pr\{S_{2n} = 0\} =$$

$$= Pr\{S_{2n-1} = +1 \cap X_{2n} = -1\} + Pr\{S_{2n-1} = -1 \cap X_{2n} = +1\}$$

Λόγω συμμετρίας

$$Pr\{S_{2n-1} = 1\} = Pr\{S_{2n-1} = -1\} \Rightarrow$$

$$\Rightarrow u_{2n} = Pr\{S_{2n-1} = 1\}\tfrac{1}{2} \cdot 2 = P_{2n-1,1} \qquad \square$$

Λήμμα

$$f_{2n} = u_{2n-2} - u_{2n}$$

Απόδειξη:

$$u_{2n} = Pr\{S_1 \neq 0, ..., S_{2n-1} \neq 0, S_{2n} \neq 0\}$$

$$u_{2n-2} = Pr\{S_1 \neq 0, ..., S_{2n-2} \neq 0\}$$

Αλλά προφανώς: $Pr\{S_1 \neq 0, ..., S_{2n-2} \neq 0\} =$

$$= Pr\{S_1 \neq 0, ..., S_{2n-2} \neq 0, S_{2n} \neq 0\} +$$

$$+ Pr\{S_1 \neq 0, ..., S_{2n-2} \neq 0, S_{2n} = 0\}$$

$$\Rightarrow u_{2n-2} = u_{2n} + f_{2n} \quad \square$$

Λήμμα

$$f_{2n} = \frac{1}{2n-1} \cdot u_{2n}$$

Απόδειξη: $f_{2n} = u_{2n-2} - u_{2n} =$

$$= \frac{\binom{2n-2}{n-1}}{2^{2n-2}} - \frac{\binom{2n}{n}}{2^{2n}} = \frac{\frac{(2n-2)!}{(n-1)!(n-1)!}}{2^{2n-2}} - u_{2n} =$$

$$= \frac{\binom{2n}{n}\frac{n \cdot n}{(2n-1)2n}}{2^{2n}2^{-2}} - u_{2n} = u_{2n}\left[\frac{4n}{2(2n-1)} - 1\right] =$$

$$= u_{2n} \cdot \frac{1}{2n-1} \qquad \square$$

Recurrence and transience

(we will deeply explore these notions in the next Lecture on Markov Chains)

- Let $\{X_n\}_{n\geq 0}$ the position (state) of a random walk on the line at time n.

- We say that state i is recurrent iff

$$\Pr\{X_n = i \text{ for infinitely many } n\} = 1$$

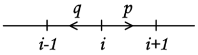
- The state i is transient iff

$$\Pr\{X_n = i \text{ for infinitely many } n\} = 0$$

- In other words:
  - the walk keeps coming back to a recurrent state for ever.
  - it eventually leaves for ever a transient state and never comes back to it.

Is the random walk on the line transient or recurrent?

- The random walk on the line:



where $0 < p = 1 - q < 1$

- Suppose we start at 0. Clearly we can not return to 0 after an odd number of steps, i.e. $P_{00}^{(2n+1)} = 0$ for all n. Note that:

$$P_{00}^{(2n)} = \binom{2n}{n} p^n q^n \text{ (return to 0 after } 2n \text{ steps)}$$

- Using Stirling's approximation $\left(n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n\right)$ we set:

$$P_{00}^{(2n)} \sim \frac{(4pq)^n}{\sqrt{2\pi}\sqrt{\frac{n}{2}}}$$

- In the symmetric case ($p = q = \frac{1}{2}$ so $4pq = 1$) we get:

$$P_{00}^{(2n)} \geq \frac{1}{2\sqrt{2\pi n}}$$

so $\displaystyle\sum_{n=0}^{\infty} P_{00}^{(2n)} = \frac{1}{2\sqrt{2\pi}} \cdot \sum_{n=0}^{\infty} \frac{1}{\sqrt{n}} = \infty$ (since the series diverges)

- Thus, ***the random walk is recurrent in the symmetric case.***

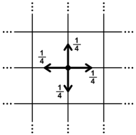
The non-symmetric case

- Let us now study the non-symmetric case ($p \neq q$ so $4pq = r < 1$)

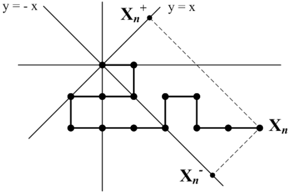
- Then
$$\sum_{n=0}^{\infty} P_{00}^{(2n)} \leq \frac{1}{\sqrt{2\pi}} \cdot \sum_{n=0}^{\infty} r^n < \infty$$
so the random walk is transient and it eventually leaves 0 and never gets back to it.

C. Random walk on the plane(I)

- The diagram of the walk is



- We start at $(0,0)$. Let us call the walk X_n and X_n^+, X_n^- be the vertical projections of the X_n on the diagonal lines $y = x$ and $y = -x$.



- Clearly $X_n = 0$ iff $X_n^+ = X_n^- = 0$

Random walk on the plane(II)

- But X_n^+ and X_n^- are independent symmetric random values on $2^{-\frac{1}{2}}Z$ so for X_n it is:

$$P_{00}^{(2n)} = \left[\binom{2n}{n}\left(\frac{1}{2}\right)^{2n}\right]^2 \sim \frac{1}{\pi n}$$

So $\displaystyle\sum_{n=0}^{\infty} P_{00}^{(2n)} = \infty$ and the walk is recurrent.

D. Random walk on the Z^3 (3D)

- It can be shown that the walk on Z^3 (even the symmetric case when all 6 next positions are equi-probable) is transient.
- This applies to more dimensions too (e.g. the symmetric random walk on Z^4 is transient etc.).