



ΣΤΟΧΑΣΤΙΚΕΣ ΔΙΑΔΙΚΑΣΙΕΣ

Μιχάλης Λουλάκης



Ελληνικά Ακαδημαϊκά Ηλεκτρονικά
Συγγράμματα και Βοηθήματα
www.kallipos.gr

HEALINK
Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
ανάπτυξη στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
Πρόγραμμα για το έθνος
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

ΜΙΧΑΛΗΣ ΛΟΥΛΑΚΗΣ
Αναπληρωτής Καθηγητής
Σχολή Εφαρμοσμένων Μαθηματικών & Φυσικών Επιστημών
Εθνικό Μετσόβιο Πολυτεχνείο

Στοχαστικές Διαδικασίες



Ελληνικά Ακαδημαϊκά Ηλεκτρονικά
Συγγράμματα και Βοηθήματα
www.kallipos.gr

Στοχαστικές Διαδικασίες

Συγγραφή

Μιχάλης Λουλάκης

Κριτικός αναγνώστης

Μιχάλης Κολουντζάκης

Συντελεστές έκδοσης

Γλωσσική Επιμέλεια: Θεόφιλος Τραμπούλης

Τεχνική Επεξεργασία: Μανώλης Γκαραγκούνης-Βλατάκης

Εξώφυλλο: Έλενα Ζακυνθινού

ISBN: 978-960-603-169-4

Copyright © ΣΕΑΒ, 2015

Επικαιροποιημένη έκδοση 2019



Το παρόν έργο αδειοδοτείται υπό τους όρους της άδειας Creative Commons Αναφορά Δημιουργού - Μη Εμπορική Χρήση - Όχι Παράγωγα Έργα 3.0. Για να δείτε ένα αντίγραφο της άδειας αυτής επισκεφτείτε τον ιστότοπο <https://creativecommons.org/licenses/by-nc-nd/3.0/gr/>

ΣΥΝΔΕΣΜΟΣ ΕΛΛΗΝΙΚΩΝ ΑΚΑΔΗΜΑΪΚΩΝ ΒΙΒΛΙΟΘΗΚΩΝ

Εθνικό Μετσόβιο Πολυτεχνείο

Ηρώων Πολυτεχνείου 9, 15780 Ζωγράφου

www.kallipos.gr

Στον Γιάννη και στον Δημήτρη

Περιεχόμενα

1	Εισαγωγή στις Στοχαστικές Διαδικασίες	1
1.1	Εισαγωγή	1
1.2	Κατανομές πεπερασμένης διάστασης	1
1.3	Διαδικασίες σε διακριτό χρόνο και χώρο	5
1.4	Μαρκοβιανές αλυσίδες	6
1.5	Ασκήσεις	11
1.6	Αριθμητικά πειράματα	13
2	Μαρκοβιανές αλυσίδες	14
2.1	Εισαγωγή	14
2.2	Πιθανότητες μετάβασης ανώτερης τάξης	14
2.3	Δομή του χώρου των καταστάσεων	18
2.4	Ισχυρή μαρκοβιανή ιδιότητα	19
2.5	Επαναληπτικότητα - Παροδικότητα	23
2.6	Ασκήσεις	29
2.7	Αριθμητικά πειράματα	31
3	Θεωρία Δυναμικού	33
3.1	Εισαγωγή	33
3.2	Πιθανότητες απορρόφησης	33
3.3	Στατιστικά του χρόνου άφιξης	42
3.4	Ασκήσεις	46
3.5	Αριθμητικά πειράματα	49
4	Martingales	51
4.1	Εισαγωγή	51
4.2	Δεσμευμένη μέση τιμή	51
4.3	Martingales	58
4.4	Martingales και μαρκοβιανές αλυσίδες	63
4.5	Ασκήσεις	67
4.6	Αριθμητικά πειράματα	70
5	Αναλλοίωτες κατανομές	72
5.1	Εισαγωγή	72
5.2	Αναλλοίωτες κατανομές	72
5.3	Η δομή του $\mathcal{I}(P)$	74
5.4	Παραδείγματα	82
5.5	Χρονική αντιστρεψιμότητα και ακριβής ισορροπία	84

5.6	Ασκήσεις	87
5.7	Αριθμητικά πειράματα	89
6	Ασυμπτωτικά Θεωρήματα	92
6.1	Εισαγωγή	92
6.2	Περιοδικότητα	92
6.3	Σύζευξη	95
6.4	Ασυμπτωτική κατανομή	99
6.5	Το εργοδικό θεώρημα	102
6.6	Ασκήσεις	107
6.7	Αριθμητικά πειράματα	109
7	Εφαρμογές των μαρκοβιανών αλυσίδων	112
7.1	Εισαγωγή	112
7.2	Μια κατανομή αξίας δισεκατομμυρίων	112
7.3	Μαρκοβιανές αλυσίδες και ηλεκτρικά κυκλώματα	114
7.4	Ο αλγόριθμος Metropolis-Hastings και το μοντέλο Ising	118
7.5	Προσομοιωμένη απόπτωση (simulated annealing)	121
7.6	Ασκήσεις	124
7.7	Αριθμητικά πειράματα	124
8	Διαδικασίες Poisson	128
8.1	Εισαγωγή	128
8.2	Ορισμός και ιδιότητες	128
8.3	Πρόσθεση και εκλέπτυνση	137
8.4	Ασκήσεις	145
	Βιβλιογραφία	147
	Γλωσσάρι τεχνικών όρων	149

Πίνακας ακρωνυμίων

ΠΣΤ	πρόβλημα συνοριακών τιμών
σ.μ.π.	συνάρτηση μάζας πιθανότητας
σ.π.π.	συνάρτηση πυκνότητας πιθανότητας
τ.μ.	τυχαία μεταβλητή
MCMC	Markov Chain Monte Carlo

Εισαγωγή

Το βιβλίο αυτό απευθύνεται σε φοιτητές που έχουν παρακολουθήσει κάποιο εισαγωγικό μάθημα Θεωρίας Πιθανοτήτων και ενδιαφέρονται να μάθουν περισσότερα Στοχαστικά Μαθηματικά. Δεν προϋποθέτει γνώσεις Θεωρίας Μέτρου, αλλά ελπίζω να προετοιμάσει τον αναγνώστη για να παρακολουθήσει ένα μετροθεωρητικό μάθημα στις Πιθανότητες και να τον προτρέψει να το κάνει.

Προκειμένου να είναι το κείμενο προσιτό σε προπτυχιακούς φοιτητές χωρίς να θυσιαστεί η μαθηματική αυστηρότητα, το βιβλίο αυτό περιλαμβάνει κυρίως στοχαστικές διαδικασίες που εξελίσσονται σε διακριτό χρόνο, μέσα σ' ένα διακριτό σύνολο καταστάσεων. Υπάρχουν δύο τρόποι να παρουσιάσει κανείς αυτό το υλικό. Ο ένας είναι να βασιστεί στη Γραμμική Άλγεβρα, με την οποία οι φοιτητές είναι συνήθως περισσότερο εξοικειωμένοι. Στην ελληνική βιβλιογραφία που γνωρίζω συνήθως προτιμάται αυτή η προσέγγιση. Στο βιβλίο αυτό θέλησα να παρουσιάσω μια όσο είναι δυνατόν πιθανοθεωρητική προσέγγιση. Αυτό ίσως να δυσκολέψει τον αναγνώστη στην πρώτη επαφή με κάποιες καινούργιες έννοιες, αλλά είμαι βέβαιος πως θα τον ανταμείψει όταν σπάσει ο πάγος.

Ένα ιδιαίτερο μέρος αυτού του βιβλίου είναι ένα εικονικό εργαστήριο, που εκτυλίσσεται παράλληλα και γύρω από το υλικό της θεωρίας. Το μοτίβο του εργαστηρίου είναι ότι δίνονται κάποιοι κώδικες στη γλώσσα προγραμματισμού Python, ο αναγνώστης τρέχει αυτούς τους κώδικες, βλέπει τι κάνουν, τους διαβάσει και τους τροποποιεί, ώστε να κάνουν κάτι διαφορετικό. Επομένως για να ασχοληθεί κάποιος με τα αριθμητικά πειράματα δεν χρειάζεται να γνωρίζει ήδη τη συγκεκριμένη γλώσσα προγραμματισμού. Ελπίζω αντίθετα ότι τα πειράματα θα προσφέρουν στον αναγνώστη μια εισαγωγή σ' αυτή τη γλώσσα. Αυτός ήταν ο τρόπος που έμαθα τουλάχιστον εγώ να προγραμματίζω σε Python. Όλοι οι κώδικες που αναφέρονται στον βιβλίο μπορούν να βρεθούν στη διεύθυνση

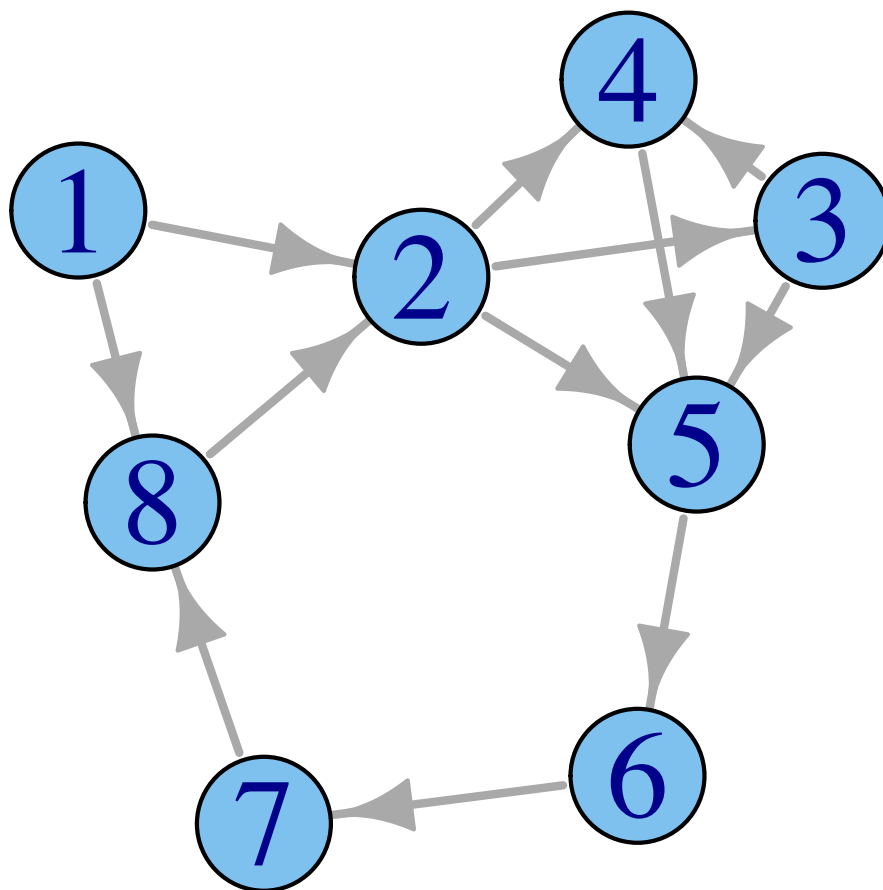
http://www.math.ntua.gr/~loulakis/info/python_codes.html

Τους κώδικες που συνοδεύουν το κείμενο έχει γράψει ο φοιτητής της Σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του ΕΜΠ Μανώλης Γκαραγκούνης-Βλατάκης, η βοήθεια του οποίου στην προετοιμασία αυτού του βιβλίου ήταν ειλικρινά συγκινητική. Τον ευχαριστώ ιδιαίτερα.

Ευχαριστώ τον γλωσσικό επιμελητή Θεόφιλο Τραμπούλη και τον κριτικό αναγνώστη Μιχάλη Κολουντζάκη για τα σχόλιά τους. Ευχαριστώ τους φοιτητές και συναδέλφους που βοήθησαν, συνήθως χωρίς να το ξέρουν, ώστε να πάρει μορφή αυτό το βιβλίο.

Τέλος, ευχαριστώ την Χαρίκλεια για την αγάπη και την ενθάρρυνση, που μου τόσο γενναϊόδωρα μου προσφέρει.

Συνιστώμενες διαδρομές περιήγησης



Κεφάλαιο 1

Εισαγωγή στις Στοχαστικές Διαδικασίες

1.1 Εισαγωγή

Στο εισαγωγικό μάθημα Πιθανοτήτων μαθαίνει κανείς τα βασικά αντικείμενα της θεωρίας των Πιθανοτήτων, τους χώρους πιθανότητας (probability spaces) και τις τυχαίες μεταβλητές (random variables). Με τα αντικείμενα αυτά μπορεί να μελετηθεί η κατάσταση ενός συστήματος ανάλογα με το αποτέλεσμα ενός πειράματος τύχης και γι' αυτόν το λόγο η έννοια του χρόνου συνήθως δεν υπεισέρχεται στα προβλήματα που αναλύονται. Σκοπός αυτών των σημειώσεων είναι να εισαγάγουν τον αναγνώστη στη μοντελοποίηση και στην ανάλυση φαινομένων τα οποία εξελίσσονται στον χρόνο με τρόπο που εμφανίζει κάποια τυχειότητα. Τα μαθηματικά αντικείμενα που μοντελοποιούν τέτοια φαινόμενα είναι οι *στοχαστικές διαδικασίες* (stochastic processes), με την περιγραφή και την κατασκευή των οποίων θα ασχοληθούμε στο πρώτο κεφάλαιο.

1.2 Κατανομές πεπερασμένης διάστασης

Υπάρχουν διάφοροι ισοδύναμοι τρόποι να ορίσει κανείς μια στοχαστική διαδικασία. Σε κάθε περίπτωση χρειαζόμαστε ένα σύνολο \mathbb{X} , που ονομάζεται χώρος καταστάσεων (state space) και στο οποίο ανήκουν οι τιμές που παίρνει η στοχαστική διαδικασία και ένα σύνολο T που είναι συνήθως -χωρίς αυτό να είναι κανόνας- ένα σύνολο χρόνων, π.χ.

$$T = [0, +\infty) \quad \text{ή} \quad T = \mathbb{N}_0 = \{0, 1, 2, \dots\}.$$

Ο απλούστερος τρόπος να ορίσει κανείς μια στοχαστική διαδικασία, ορισμένη στο T και με τιμές στο \mathbb{X} είναι ως μια συλλογή από τυχαίες μεταβλητές $\{X_t\}_{t \in T}$, ορισμένες σε έναν χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbb{P})$, με τιμές στο \mathbb{X} . Σε αυτή την περίπτωση, για κάθε $t \in T$, η τυχαία μεταβλητή $X_t : \Omega \ni \omega \mapsto X_t(\omega) \in \mathbb{X}$ περιγράφει την κατάσταση του συστήματος που μας ενδιαφέρει την χρονική στιγμή t και η συλλογή τους περιγράφει την κατάσταση του συστήματος κάθε χρονική στιγμή στο σύνολο T .

Ένας άλλος τρόπος να δει κανείς μια στοχαστική διαδικασία είναι να θεωρήσει για κάθε $\omega \in \Omega$ την τροχιά της στοχαστικής διαδικασίας, καθώς ο δείκτης t μεταβάλλεται στο T , να δει δηλαδή την $t \mapsto X_t(\omega)$ ως μια συνάρτηση από το T στο \mathbb{X} . Σε αυτή την περίπτωση η στοχαστική διαδικασία είναι μια τυχαία συνάρτηση, δηλαδή μια τυχαία μεταβλητή ορισμένη στο Ω με τιμές στο σύνολο \mathbb{X}^T των συναρτήσεων από το T στο \mathbb{X} .

Παράδειγμα 1 Στριβούμε ένα κέρμα τρεις φορές. Ξεκινώντας από το μηδέν κάνουμε ένα βήμα προς τα δεξιά κάθε φορά που φέρνουμε κεφαλή και ένα βήμα προς τα αριστερά κάθε φορά που φέρνουμε γράμματα. Μπορούμε να περιγράψουμε τη θέση μας ως μια στοχαστική διαδικασία ορισμένη στο σύνολο χρόνων $T = \{0, 1, 2, 3\}$ με τιμές στον χώρο καταστάσεων \mathbb{Z} . Πράγματι, αν για $k = 0, 1, 2, 3$ συμβολίζουμε με X_k τη θέση μας μετά από k στριψίματα, τότε οι $\{X_k\}_{k \in T}$ είναι τυχαίες μεταβλητές με τιμές στον \mathbb{Z} , που μπορούμε να τις ορίσουμε στον δειγματικό χώρο

$$\Omega = \{KKK, KKT, KTK, KTF, GK K, GK T, GFK, GTF\}.$$

Έχουμε λοιπόν $X_0(\omega) = 0$ για κάθε $\omega \in \Omega$ και

$$X_1(\omega) = \begin{cases} +1, & \text{αν } \omega \in \{KKK, KKT, KTK, KTF\} \\ -1, & \text{αν } \omega \in \{GKK, GK T, GFK, GTF\}, \end{cases}$$

$$X_2(\omega) = \begin{cases} +2, & \text{αν } \omega \in \{KKK, KKT\} \\ 0, & \text{αν } \omega \in \{KTK, KTF, GK K, GK T\} \\ -2, & \text{αν } \omega \in \{GFK, GTF\}, \end{cases}$$

$$X_3(\omega) = \begin{cases} +3, & \text{αν } \omega = KKK \\ +1, & \text{αν } \omega \in \{KKT, KTK, GK K\} \\ -1, & \text{αν } \omega \in \{GFK, GK T, KTF\} \\ -3, & \text{αν } \omega = GTF. \end{cases}$$

Αν δούμε την ίδια στοχαστική διαδικασία ως μια τυχαία μεταβλητή με τιμές στον \mathbb{Z}^T έχουμε ότι

$$X.(KKK) = (0, 1, 2, 3), \quad X.(KKT) = (0, 1, 2, 1), \quad X.(KTK) = (0, 1, 0, 1) \quad \text{κ.λπ.}$$

Παρατηρήστε ότι κάθε $\omega \in \Omega$ απεικονίζεται σε μια τροχιά μήκους $|T| = 4$ στο \mathbb{Z} . □

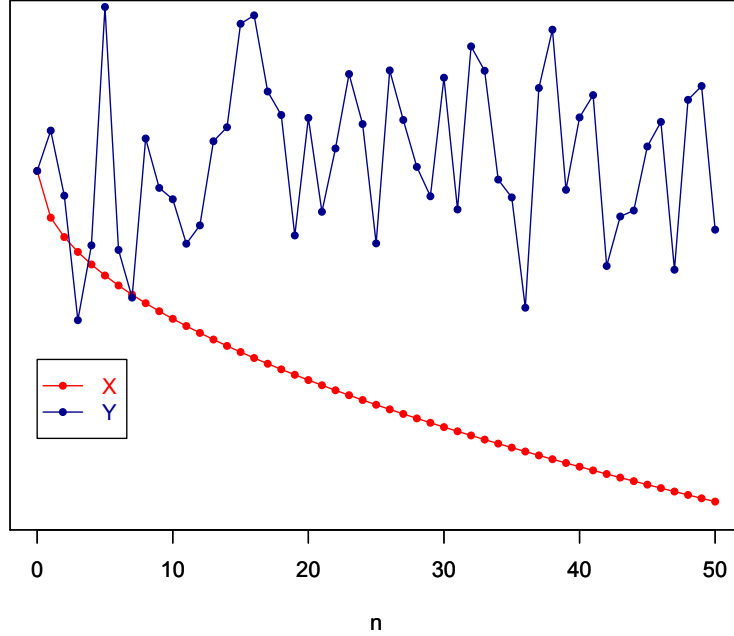
Στο προηγούμενο παράδειγμα κατασκευάσαμε μια στοχαστική διαδικασία περιγράφοντας τον δειγματικό χώρο στον οποίο την έχουμε ορίσει και δίνοντας τον τύπο της, $\omega \mapsto X.(\omega) \in \mathbb{X}^T$. Στην θεωρία των Πιθανοτήτων όμως τα ερωτήματα που μας ενδιαφέρουν δεν αφορούν το πού και πώς έχουμε ορίσει μια τυχαία μεταβλητή, αλλά τις στατιστικές της ιδιότητες, δηλαδή το πόσο πιθανό είναι να βρεθεί η τυχαία μεταβλητή σε διάφορα υποσύνολα του χώρου εντός του οποίου παίρνει τιμές. Αυτή η πληροφορία δίνεται από την κατανομή της τυχαίας μεταβλητής. Γι' αυτό και συνήθως, όταν περιγράφουμε μια τυχαία μεταβλητή, δεν κάνουμε αναφορά στον δειγματικό χώρο Ω στον οποίο είναι ορισμένη και στον τύπο της, αλλά στην κατανομή της. Αυτό θα προσπαθήσουμε να κάνουμε τώρα και για μια στοχαστική διαδικασία. Θα προσπαθήσουμε δηλαδή να την περιγράψουμε μέσω των στατιστικών ιδιοτήτων της.

Θα ήταν άραγε αρκετό να περιγράψουμε την κατανομή της τυχαίας μεταβλητής X_t , για κάθε $t \in T$; Το επόμενο παράδειγμα μας διδάσκει πως η απάντηση είναι αρνητική.

Παράδειγμα 2 Έστω ότι έχουμε ορίσει σε κάποιον χώρο πιθανότητας μια ακολουθία $\{Z_k\}_{k \in \mathbb{N}_0}$ από ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με τυπική κανονική κατανομή. Ορίζουμε τις παρακάτω δύο διαδικασίες διακριτού χρόνου $\{X_n\}_{n \in \mathbb{N}_0}$ και $\{Y_n\}_{n \in \mathbb{N}_0}$.

$$X_n = \sqrt{n}Z_0, \quad \text{για } n \in \mathbb{N}_0 \quad \text{και} \quad Y_0 = 0, \quad Y_n = \sum_{k=1}^n Z_k, \quad \text{για } n \in \mathbb{N}.$$

Εφόσον $Z_0 \sim \mathcal{N}(0, 1)$, έχουμε $X_n = \sqrt{n}Z_0 \sim \mathcal{N}(0, n)$, για κάθε $n \in \mathbb{N}$. Επιπλέον, εφόσον οι $\{Z_k\}_{k \in \mathbb{N}}$ είναι ανεξάρτητες τυπικές κανονικές, τότε και η Y_n θα ακολουθεί κανονική κατανομή για κάθε $n \in \mathbb{N}$, με



Σχήμα 1.1: Τυπικές τροχιές των διαδικασιών $\{X_n\}_n$ και $\{Y_n\}$ του Παραδείγματος 2.

μέση τιμή $\mathbb{E}[Y_n] = n\mathbb{E}[Z_1] = 0$ και διασπορά $V(Y_n) = nV(Z_1) = n$. Έχουμε επομένως ότι $Y_n \sim \mathcal{N}(0, n)$ για κάθε $n \in \mathbb{N}$. Παρατηρήστε ότι κάθε χρονική στιγμή $n \in \mathbb{N}_0$ οι τυχαίες μεταβλητές X_n και Y_n έχουν την ίδια κατανομή. Από την άλλη πλευρά, το Σχήμα 1.1 είναι ενδεικτικό του πόσο διαφέρουν οι τροχιές των δύο στοχαστικών διαδικασιών για μια τυπική πραγματοποίησή τους $\omega \in \Omega$. Παρατηρήστε ακόμα ότι, αν ξέρει κανείς την τιμή $X_1(\omega)$, μπορεί να ανακατασκευάσει όλη την τροχιά $X(\omega)$. Αντίθετα, οι προσauξήσεις της $\{Y_n\}_n$ μετά τη χρονική στιγμή $n = 1$ είναι ανεξάρτητες της Y_1 . \square

Είδαμε λοιπόν ότι δύο στοχαστικές διαδικασίες, που κάθε χρονική στιγμή $t \in T$ έχουν την ίδια κατανομή, μπορεί να έχουν τροχιές με πολύ διαφορετικά ποιοτικά χαρακτηριστικά. Αυτό συμβαίνει γιατί, περιγράφοντας μόνο την κατανομή των X_t , για κάθε $t \in T$, δεν δίνουμε καμία πληροφορία για τη μεταξύ τους εξάρτηση. Η πληροφορία για την αλληλεξάρτηση τυχαίων μεταβλητών βρίσκεται στην από κοινού κατανομή τους (joint distribution). Επομένως, οποιαδήποτε πλήρης περιγραφή των στατιστικών ιδιοτήτων μιας στοχαστικής διαδικασίας θα πρέπει να εμπεριέχει την από κοινού κατανομή των τυχαίων μεταβλητών $X_{t_1}, X_{t_2}, \dots, X_{t_k}$, για κάθε $k \in \mathbb{N}$ και για κάθε διατεταγμένη k -άδα χρόνων $F = (t_1, t_2, \dots, t_k) \subset T^k$. Κάθε τέτοια κατανομή είναι ένα μέτρο πιθανότητας μ_F στον \mathbb{X}^k , που ορίζεται μέσω της

$$\mu_F[A] = \mathbb{P}[(X_{t_1}, \dots, X_{t_n}) \in A] \quad (1.1)$$

για κατάλληλα σύνολα $A \in \mathbb{X}^k$. Θα συμβολίζουμε με $\mathcal{S}(T)$ το σύνολο των πεπερασμένων διατεταγμένων υποσυνόλων του T , δηλαδή

$$\mathcal{S}(T) = \bigcup_{k \in \mathbb{N}} T^k = \bigcup_{k \in \mathbb{N}} \{(t_1, \dots, t_k) : t_i \in T, i = 1, \dots, k\}.$$

Με άλλα λόγια το $\mathcal{S}(T)$ αποτελείται από τα πεπερασμένα διατεταγμένα σύνολα οσωνδήποτε και οποιωνδήποτε χρόνων.

Ορισμός: Ονομάζουμε τη συλλογή $\{\mu_F\}_{F \in \mathcal{S}(T)}$ όλων των κατανομών της (1.1) οικογένεια των κατανομών πεπερασμένης διάστασης της στοχαστικής διαδικασίας $\{X_t\}_{t \in T}$.

Μπορεί να αποδειχθεί ότι η οικογένεια των κατανομών πεπερασμένης διάστασης μιας στοχαστικής διαδικασίας $\{X_t\}_{t \in T}$ προσδιορίζει πλήρως τις στατιστικές της ιδιότητες. Αυτό είναι δύσκολο να γίνει αντιληπτό με μαθηματική αυστηρότητα, χωρίς την ορολογία της Θεωρίας Μέτρου. Για τις ανάγκες αυτού του μαθήματος είναι αρκετό να θυμόμαστε ότι μπορούμε να περιγράψουμε την κατανομή μιας στοχαστικής διαδικασίας προσδιορίζοντας όλες τις κατανομές πεπερασμένης διάστασης. Αν αυτές είναι γνωστές, μπορούμε να απαντήσουμε οποιοδήποτε πιθανοθεωρητικό ερώτημα που αφορά τη διαδικασία, χωρίς να χρειάζεται να γνωρίζουμε σε ποιον δειγματικό χώρο είναι ορισμένη ή ποιος είναι ο τύπος της.

Παράδειγμα 3 Θα χαρακτηρίζουμε μια στοχαστική διαδικασία ως διαδικασία Gauss, αν κάθε κατανομή πεπερασμένης διάστασης μ_F είναι κανονική σε $|F|$ διαστάσεις. Εφόσον κάθε πολυδιάστατη κανονική κατανομή προσδιορίζεται από το διάνυσμα των μέσων τιμών και τον πίνακα συνδιακυμάνσεων, για να περιγράψουμε μια διαδικασία Gauss $\{X_t\}_{t \in T}$, αρκεί να περιγράψουμε τις συναρτήσεις μέσης τιμής και αυτοσυσχέτισης

$$m : T \rightarrow \mathbb{R}, \quad \text{με } m(t) = \mathbb{E}[X_t] \quad \text{και} \quad \rho : T \times T \rightarrow \mathbb{R}, \quad \text{με } \rho(s, t) = \text{Cov}(X_s, X_t).$$

Η τυπική κίνηση Brown είναι μια διαδικασία Gauss, ορισμένη στο $T = [0, +\infty)$ με τιμές στο \mathbb{R} για την οποία

$$m(t) = 0 \quad \text{και} \quad \rho(s, t) = s \wedge t = \min\{s, t\}, \quad \text{για κάθε } s, t \geq 0.$$

Το ότι υπάρχει τέτοια διαδικασία δεν είναι καθόλου αυτονόητο. Μεσολάβησαν είκοσι τρία χρόνια ανάμεσα στην πρώτη ευρετική χρήση αυτής της διαδικασίας από τον Bachelier και στη μαθηματικά αυστηρή κατασκευή της από τον Wiener το 1923.

Για κάθε $t \geq 0$, η κατανομή της τυχαίας μεταβλητής X_t είναι κανονική, αφού η X_t είναι διαδικασία Gauss. Επιπλέον, $\mathbb{E}[X_t] = m(t) = 0$ και $V(X_t) = \text{Cov}(X_t, X_t) = \rho(t, t) = t$, άρα $X_t \sim \mathcal{N}(0, t)$. Ειδικότερα $X_0 = 0$.

Ας δούμε τώρα ποια κατανομή ακολουθεί η προσαύξηση $X_t - X_s$ της κίνησης Brown σ' ένα διάστημα $(s, t]$. Το ζεύγος (X_s, X_t) ακολουθεί κανονική κατανομή σε δύο διαστάσεις, αφού η $\{X_t\}_{t \geq 0}$ είναι διαδικασία Gauss. Η $X_t - X_s$ προκύπτει από τις X_s, X_t μέσω ενός γραμμικού μετασχηματισμού, άρα ακολουθεί κανονική κατανομή. Εύκολα βλέπει κανείς ότι $\mathbb{E}[X_t - X_s] = m(t) - m(s) = 0$, ενώ

$$V(X_t - X_s) = \text{Cov}(X_t - X_s, X_t - X_s) = \text{Cov}(X_t, X_t) - 2\text{Cov}(X_t, X_s) + \text{Cov}(X_s, X_s) = t - 2s + s = t - s.$$

Επομένως $X_t - X_s \sim \mathcal{N}(0, t - s)$.

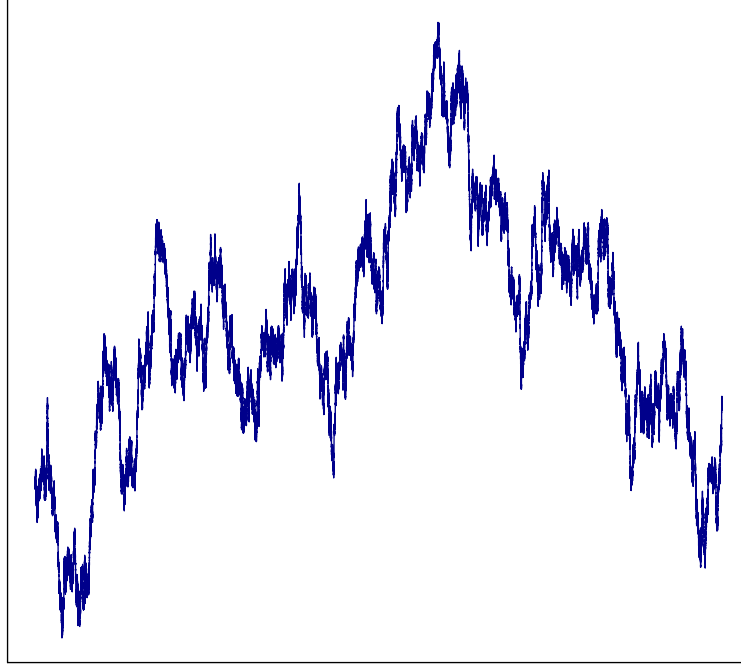
Θεωρήστε τώρα $0 \leq r < s < t$. Θα δείξουμε ότι οι προσαυξήσεις $X_t - X_s$ και $X_s - X_r$ είναι ανεξάρτητες τυχαίες μεταβλητές. Η από κοινού κατανομή των (X_r, X_s, X_t) είναι τριδιάστατη κανονική με διάνυσμα μέσων τιμών $(\mathbb{E}[X_r], \mathbb{E}[X_s], \mathbb{E}[X_t]) = (0, 0, 0)$ και πίνακα συνδιακυμάνσεων

$$\Sigma(r, s, t) = \begin{pmatrix} r & r & r \\ r & s & s \\ r & s & t \end{pmatrix}.$$

Εφόσον οι $(X_t - X_s, X_s - X_r)$ προκύπτουν από τις (X_r, X_s, X_t) μέσω ενός γραμμικού μετασχηματισμού, ακολουθούν κι αυτές κανονική κατανομή (σε δύο διαστάσεις). Αρκεί λοιπόν να δείξουμε ότι είναι ασυσχέτιστες.

$$\text{Cov}(X_t - X_s, X_s - X_r) = \text{Cov}(X_t, X_s) - \text{Cov}(X_t, X_r) - \text{Cov}(X_s, X_s) + \text{Cov}(X_s, X_r) = s - r - s + r = 0.$$

Στο Σχήμα 1.2 φαίνεται μια τυπική τροχιά της κίνησης Brown. □



Σχήμα 1.2: Τυπική τροχιά της κίνησης Brown.

1.3 Διαδικασίες σε διακριτό χρόνο και χώρο

Στο μεγαλύτερο μέρος αυτών των σημειώσεων θα μελετήσουμε στοχαστικές διαδικασίες σε διακριτό χρόνο με τιμές σε έναν αριθμήσιμο χώρο καταστάσεων \mathbb{X} . Σε αυτή την παράγραφο θα δούμε πώς μπορούμε να περιγράψουμε εύκολα τις κατανομές πεπερασμένης διάστασης τέτοιων στοχαστικών διαδικασιών.

Έχοντας επιλέξει $T = \mathbb{N}_0$, προκειμένου να περιγράψουμε όλες τις κατανομές πεπερασμένης διάστασης, αρκεί να περιγράψουμε την από κοινού κατανομή των τυχαίων μεταβλητών (X_0, X_1, \dots, X_n) για κάθε $n \in \mathbb{N}_0$. Πράγματι, για κάθε $F = (n_1, \dots, n_k) \in \mathbb{N}_0^k$, αν θέσουμε $n = \max\{n_1, \dots, n_k\}$, τότε μπορούμε να πάρουμε την μ_F , δηλαδή την από κοινού κατανομή των $(X_{n_1}, \dots, X_{n_k})$, ως μια κατάλληλη περιθώρια κατανομή της από κοινού κατανομής των (X_0, X_1, \dots, X_n) .

Επιπλέον, εφόσον ο \mathbb{X} είναι αριθμήσιμος, ο \mathbb{X}^k είναι και αυτός αριθμήσιμος για κάθε $k \in \mathbb{N}$ και επομένως όλες οι κατανομές πεπερασμένης διάστασης είναι διακριτές κατανομές. Μπορούμε λοιπόν να τις περιγράψουμε μέσω της συνάρτησης μάζας πιθανότητας (σ.μ.π.) τους. Συγκεκριμένα, αν $F_n = (0, 1, \dots, n)$, η σ.μ.π. της κατανομής μ_{F_n} δίνεται, για κάθε $\eta = (x_0, \dots, x_n) \in \mathbb{X}^{n+1}$, από την

$$Q_n(\eta) = \mu_{F_n}[\{\eta\}] = \mathbb{P}[(X_0, \dots, X_n) = \eta] = \mathbb{P}[X_0 = x_0, \dots, X_n = x_n]. \quad (1.2)$$

Επομένως, στην περίπτωση που έχουμε μια στοχαστική διαδικασία σε διακριτό χρόνο $T = \mathbb{N}_0$ με τιμές σ' έναν αριθμήσιμο χώρο καταστάσεων \mathbb{X} , προκειμένου να περιγράψουμε όλες τις κατανομές πεπερασμένης διάστασης της διαδικασίας αρκεί να περιγράψουμε την $Q_n(\eta)$ της σχέσης (1.2), για κάθε $n \in \mathbb{N}_0$ και κάθε $\eta \in \mathbb{X}^{n+1}$. Σ' αυτό το σημείο ας θυμηθούμε τον πολλαπλασιαστικό τύπο για τα μέτρα πιθανότητας. Αν A_0, A_1, \dots, A_k είναι ενδεχόμενα ενός δειγματικού χώρου, τότε

$$\mathbb{P}[A_0 \cap \dots \cap A_n] = \mathbb{P}[A_0] \mathbb{P}[A_1|A_0] \mathbb{P}[A_2|A_0 \cap A_1] \cdots \mathbb{P}[A_n|A_0 \cap \dots \cap A_{n-1}]. \quad (1.3)$$

Αν στην παραπάνω σχέση επιλέξουμε $A_j = \{X_j = x_j\}$, $j = 0, 1, \dots, n$, έχουμε ότι

$$Q_n(\eta) = \mathbb{P}[X_0 = x_0] \mathbb{P}[X_1 = x_1 | X_0 = x_0] \cdots \mathbb{P}[X_n = x_n | X_0 = x_0, \dots, X_{n-1} = x_{n-1}]. \quad (1.4)$$

Ο πρώτος όρος του δεξιού μέλους μπορεί να υπολογιστεί από την κατανομή της τυχαίας μεταβλητής X_0 . Ο γενικός όρος του δεξιού μέλος είναι της μορφής

$$\mathbb{P}[X_k = x_k | X_0 = x_0, \dots, X_{k-1} = x_{k-1}]$$

και μπορεί να υπολογιστεί από τη δεσμευμένη κατανομή της τυχαίας μεταβλητής X_k , δοθέντων των X_0, \dots, X_{k-1} . Επομένως για να περιγράψουμε την $Q_n(\eta)$ της σχέσης (1.2), για κάθε $n \in \mathbb{N}_0$ και κάθε $\eta \in \mathbb{X}^{n+1}$ και άρα για να ορίσουμε ολόκληρη την οικογένεια κατανομών πεπερασμένης διάστασης της στοχαστικής διαδικασίας $\{X_n\}_{n \in \mathbb{N}_0}$, αρκεί να περιγράψουμε

- την κατανομή της X_0 και
- τη δεσμευμένη κατανομή της X_k , δοθέντων των X_0, \dots, X_{k-1} , για κάθε $k \in \mathbb{N}$.

Η πρώτη πληροφορία αφορά την αρχική κατάσταση του συστήματος, ενώ η δεύτερη αφορά τη δυναμική του, δηλαδή τον νόμο που περιγράφει την εξέλιξη του συστήματος στον χρόνο. Παρατηρήστε την αναλογία με τα αιτιοκρατικά δυναμικά συστήματα διακριτού χρόνου, τα οποία μπορούμε να περιγράψουμε δίνοντας την αρχική τους κατάσταση και έναν κανόνα για το πώς προκύπτει η επόμενη κατάσταση από τις προηγούμενες.

Είδαμε πώς μπορούμε να περιγράψουμε πλήρως τις στατιστικές ιδιότητες μιας στοχαστικής διαδικασίας διακριτού χρόνου που έχει ήδη οριστεί σε κάποιον χώρο πιθανότητας. Αρκεί να ξέρουμε την αρχική της κατανομή και τη δεσμευμένη κατανομή της επόμενης κατάστασης, δοθέντων των προηγούμενων. Σ' αυτό το σημείο αξίζει να αναρωτηθούμε για το αντίστροφο ερώτημα. Είναι πάντα δυνατό να βρούμε έναν κατάλληλο χώρο πιθανότητας και σ' αυτόν να ορίσουμε μια στοχαστική διαδικασία $\{X_n\}_{n \in \mathbb{N}_0}$, ώστε αυτή να έχει μια συγκεκριμένη αρχική κατανομή και έναν συγκεκριμένο νόμο εξέλιξης με την έννοια που περιγράψαμε παραπάνω; Η απάντηση είναι καταφατική και βασίζεται στο Θεώρημα Συνέπειας του Kolmogorov (Θεώρημα 3.5 στο [8]). Η απόδειξη όμως αυτού του ισχυρισμού ξεφεύγει από τους σκοπούς του μαθήματος. Για την ώρα είναι αρκετό να θυμόμαστε ότι το Θεώρημα Συνέπειας μας εξασφαλίζει πως οι στοχαστικές διαδικασίες που θα μελετήσουμε είναι υπαρκτά μαθηματικά αντικείμενα.

1.4 Μαρκοβιανές αλυσίδες

Όταν θέλουμε να μοντελοποιήσουμε πραγματικά συστήματα είναι πολύ συνηθισμένο ο κανόνας που περιγράφει τη δυναμική του συστήματος να εξαρτάται μόνο από την τρέχουσα κατάσταση του συστήματος και όχι από το πώς το σύστημα βρέθηκε εκεί. Τα στοχαστικά συστήματα που έχουν αυτή την ιδιότητα χαρακτηρίζονται ως μαρκοβιανά. Ο αναγνώστης είναι πιθανότατα εξοικειωμένος με τέτοια παραδείγματα μιας και τα περισσότερα επιτραπέζια παιχνίδια αποτελούν μαρκοβιανά συστήματα. Για παράδειγμα, αν μπειτε σ' ένα δωμάτιο όπου δύο φίλοι σας παίζουν τάβλι, για να εκτιμήσετε την πιθανότητα νίκης κάθε παίκτη, αρκεί να δείτε την τρέχουσα διαμόρφωση των πουλιών. Αν ξέρετε αυτή τη διαμόρφωση, η γνώση του τι προηγήθηκε δεν θα προσφέρει τίποτα στην πρόβλεψή σας για την εξέλιξη της παρτίδας.

Ορισμός: Μια στοχαστική διαδικασία $\{X_n\}_{n \in \mathbb{N}_0}$ λέγεται *μαρκοβιανή αλυσίδα* (Markov chain), αν, για κάθε $n \in \mathbb{N}$, η δεσμευμένη κατανομή της X_{n+1} δοθέντων των (X_0, \dots, X_n) , ταυτίζεται με τη δεσμευμένη κατανομή της X_{n+1} με μόνη δοθείσα την X_n . Επομένως, η $\{X_n\}_{n \in \mathbb{N}_0}$ με τιμές σε έναν αριθμησιμο χώρο καταστάσεων \mathbb{X} είναι μαρκοβιανή αλυσίδα, αν, για κάθε $n \in \mathbb{N}$ και κάθε $v_0, \dots, v_{n-1}, x, y \in \mathbb{X}$, έχουμε

$$\mathbb{P}[X_{n+1} = y | X_0 = v_0, \dots, X_{n-1} = v_{n-1}, X_n = x] = \mathbb{P}[X_{n+1} = y | X_n = x]. \quad (1.5)$$

Στις περισσότερες μαρκοβιανές αλυσίδες που παρουσιάζουν ενδιαφέρον, ο κανόνας $\mathbb{P}[X_{n+1} = \cdot | X_n = \cdot]$ που περιγράφει την εξέλιξη της αλυσίδας, δεν εξαρτάται από τη χρονική παράμετρο n . Λέμε τότε ότι η αλυσίδα είναι *χρονικά ομοιογενής* (time homogeneous) και ορίζουμε τις *πιθανότητες μετάβασης* (transition probabilities) της αλυσίδας

$$p : \mathbb{X} \times \mathbb{X} \rightarrow [0, 1], \quad \text{με τύπο } p(x, y) = \mathbb{P}[X_{n+1} = y | X_n = x]. \quad (1.6)$$

Επειδή στις σημειώσεις αυτές θα ασχοληθούμε μόνο με χρονικά ομοιογενείς αλυσίδες, στο εξής θα παραλείπεται η διευκρίνιση για λόγους οικονομίας.

Η συλλογή $P = \{p(x, y)\}_{x, y \in \mathbb{X}}$ ονομάζεται *πίνακας πιθανοτήτων μετάβασης* (transition matrix) της αλυσίδας. Η ορολογία προέρχεται από την περίπτωση που ο χώρος καταστάσεων \mathbb{X} είναι πεπερασμένος, οπότε μπορούμε να περιγράψουμε μία συνάρτηση με πεδίο ορισμού το $\mathbb{X} \times \mathbb{X}$ ως ένα τετραγωνικό πίνακα. Συγκεκριμένα, αν $\mathbb{X} = \{v_1, \dots, v_N\}$, έχουμε

$$P = \begin{pmatrix} p(v_1, v_1) & p(v_1, v_2) & \cdots & p(v_1, v_N) \\ p(v_2, v_1) & p(v_2, v_2) & \cdots & p(v_2, v_N) \\ \vdots & \vdots & \ddots & \vdots \\ p(v_N, v_1) & p(v_N, v_2) & \cdots & p(v_N, v_N) \end{pmatrix}. \quad (1.7)$$

Ας συμβολίζουμε με π_0 την αρχική κατανομή της αλυσίδας, δηλαδή

$$\pi_0 : \mathbb{X} \rightarrow [0, 1], \quad \text{με τύπο } \pi_0(x) = \mathbb{P}[X_0 = x]. \quad (1.8)$$

Ως σ.μ.π. της τυχαίας μεταβλητής X_0 , η π_0 ικανοποιεί τις παρακάτω συνθήκες

$$\pi_0(x) \geq 0, \quad \text{για κάθε } x \in \mathbb{X} \quad \text{και} \quad \sum_{x \in \mathbb{X}} \pi_0(x) = 1. \quad (1.9)$$

Αντίστοιχα, εφόσον η $p(x, \cdot)$ είναι η δεσμευμένη σ.μ.π. της X_{n+1} δεδομένου ότι $X_n = x$, οι πιθανότητες μετάβασης της μαρκοβιανής αλυσίδας $\{X_n\}_{n \in \mathbb{N}_0}$ ικανοποιούν τις συνθήκες

$$p(x, y) \geq 0, \quad \text{για κάθε } x, y \in \mathbb{X} \quad \text{και} \quad \sum_{y \in \mathbb{X}} p(x, y) = 1, \quad \text{για κάθε } x \in \mathbb{X}. \quad (1.10)$$

Ορισμός Θα λέμε ότι ο $P = \{p(x, y)\}_{x, y \in \mathbb{X}}$ είναι *στοχαστικός πίνακας* (stochastic matrix), αν τα στοιχεία του ικανοποιούν τις συνθήκες της (1.10).

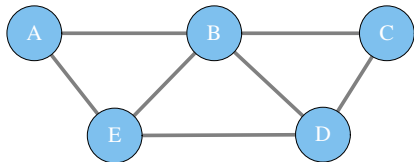
Σύμφωνα με όσα είδαμε στην παράγραφο 1.3, η σ.μ.π. π_0 της (1.8) και ο στοχαστικός πίνακας $P = \{p(x, y)\}_{x, y \in \mathbb{X}}$ της (1.6) καθορίζουν πλήρως τις στατιστικές ιδιότητες της μαρκοβιανής αλυσίδας $\{X_n\}_{n \in \mathbb{N}_0}$. Πράγματι, οι κατανομές πεπερασμένης διάστασης της $\{X_n\}_{n \in \mathbb{N}_0}$ μπορούν να υπολογιστούν με την βοήθεια της (1.4). Για κάθε $n \in \mathbb{N}_0$ και κάθε $(x_0, \dots, x_n) \in \mathbb{X}^{n+1}$ έχουμε

$$\mathbb{P}[X_0 = x_0, \dots, X_n = x_n] = \pi_0(x_0) p(x_0, x_1) p(x_1, x_2) \cdots p(x_{n-1}, x_n). \quad (1.11)$$

Στα επόμενα κεφάλαια θα δούμε ότι για να αναλύσουμε τη συμπεριφορά μιας αλυσίδας αρκεί να γνωρίζουμε αυτές τις ποσότητες. Όπως και στην περίπτωση των πραγματικών τυχαίων μεταβλητών που η κατανομή τους μας είναι γνωστή, ο δειγματικός χώρος Ω στον οποίο έχει οριστεί η αλυσίδα και ο τύπος της απεικόνισης $\omega \mapsto X(\omega)$ δεν θα παίζουν κανένα ρόλο.

Θα δούμε στα παραδείγματα που ακολουθούν ότι, στη μοντελοποίηση πραγματικών συστημάτων από μαρκοβιανές αλυσίδες, συνήθως το πρόβλημα υποδεικνύει την αρχική κατανομή π_0 και τον πίνακα πιθανοτήτων μετάβασης P που θέλουμε να έχει η αλυσίδα η οποία περιγράφει το σύστημα. Το Θεώρημα Συνέπειας

του Kolmogorov μας εξασφαλίζει ότι, για κάθε σ.μ.π. π_0 και για κάθε στοχαστικό πίνακα P , μπορούμε να κατασκευάσουμε μια μαρκοβιανή αλυσίδα $\{X_n\}_{n \in \mathbb{N}_0}$, η οποία έχει την π_0 ως αρχική κατανομή και τον P ως πίνακα πιθανοτήτων μετάβασης. Οι λεπτομέρειες αυτής της κατασκευής είναι αδιάφορες για την ανάλυση του μοντέλου, αφού η π_0 και ο P , όπως είπαμε, καθορίζουν τις στατιστικές ιδιότητες της αλυσίδας.



Παράδειγμα 4 Ένα έντομο κινείται στις κορυφές του διπλανού γράφου. Ξεκινά από την κορυφή A . Σε κάθε βήμα, αν βρίσκεται στην κορυφή x , επιλέγει τυχαία μία από τις κορυφές που συνδέονται με την x μέσω μιας ακμής του γράφου και μεταβαίνει εκεί.

Μπορούμε να περιγράψουμε την κίνηση του εντόμου ως μια μαρκοβιανή αλυσίδα στον χώρο καταστάσεων $\mathbb{X} = \{A, B, C, D, E\}$. Εφόσον η αλυσίδα ξεκινά από την κορυφή A , η αρχική κατανομή της έχει σ.μ.π.

$$\pi_0(x) = \mathbb{P}[X_0 = x] = \delta_A(x) = \begin{cases} 1, & \text{αν } x = A \\ 0, & \text{αν } x \neq A. \end{cases}$$

Μπορούμε να περιγράψουμε την αρχική κατανομή και ως ένα διάνυσμα γραμμής

$$\pi_0 = (\pi_0(A), \pi_0(B), \pi_0(C), \pi_0(D), \pi_0(E)) = (1, 0, 0, 0, 0).$$

Ο πίνακας πιθανοτήτων μετάβασης P της αλυσίδας είναι ο

$$P = \begin{pmatrix} 0 & 1/2 & 0 & 0 & 1/2 \\ 1/4 & 0 & 1/4 & 1/4 & 1/4 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 1/3 & 1/3 & 0 & 1/3 \\ 1/3 & 1/3 & 0 & 1/3 & 0 \end{pmatrix}.$$

Κάθε του γραμμή περιγράφει τις μεταβάσεις από μια κορυφή. Ας δούμε για παράδειγμα πώς προκύπτει η πρώτη γραμμή του P , που περιγράφει τις μεταβάσεις από την κορυφή A . Από την A μπορούμε να πάμε στις κορυφές B ή E με πιθανότητα $1/2$ σε κάθε μία. Άρα η πρώτη γραμμή του P είναι

$$(p(A, A), p(A, B), p(A, C), p(A, D), p(A, E)) = (0, 1/2, 0, 0, 1/2).$$

Παράδειγμα 5 Κάνουμε έναν τυχαίο περίπατο στους ακεραίους ξεκινώντας από το 0 . Σε κάθε μας βήμα στρίβουμε ένα κέρμα που έχει πιθανότητα $p \in (0, 1)$ να φέρει κεφαλή κάθε φορά που το στρίβουμε, ανεξάρτητα από τις άλλες φορές. Αν το αποτέλεσμα είναι κεφαλή, μετατοπιζόμαστε μία θέση δεξιά. Αν το αποτέλεσμα είναι γράμματα, μετατοπιζόμαστε μία θέση αριστερά. Μπορούμε να περιγράψουμε αυτόν τον περίπατο ως μια μαρκοβιανή αλυσίδα στον χώρο καταστάσεων \mathbb{Z} . Εφόσον ξεκινάμε από το 0 , η αρχική κατανομή έχει σ.μ.π.

$$\pi_0(x) = \mathbb{P}[X_0 = x] = \delta_0(x) = \begin{cases} 1, & \text{αν } x = 0 \\ 0, & \text{αν } x \neq 0. \end{cases}$$

Οι πιθανότητες μετάβασης της αλυσίδας, δίνονται από την

$$p(x, y) = \mathbb{P}[X_{n+1} = y | X_n = x] = \begin{cases} p, & \text{αν } y = x + 1 \\ 1 - p, & \text{αν } y = x - 1 \\ 0, & \text{αν } |y - x| \neq 1 \end{cases}, \quad x, y \in \mathbb{Z}.$$

Παράδειγμα 6 Στριβούμε ένα τίμιο κέρμα μέχρι να εμφανιστεί για πρώτη φορά η ακολουθία αποτελεσμάτων ΚΓΚ. Μια πρώτη σκέψη θα ήταν να περιγράψουμε αυτό το πείραμα ως μια μαρκοβιανή αλυσίδα, που η κατάσταση της είναι το αποτέλεσμα των τριών τελευταίων στριψιμάτων. Με αυτή την επιλογή ο χώρος καταστάσεων θα ήταν ο $\mathbb{X}_1 = \{K, \Gamma\}^3$. Η αρχική κατανομή της αλυσίδας θα προκύψει από το αποτέλεσμα των τριών πρώτων στριψιμάτων. Εφόσον το κέρμα είναι τίμιο, θα πρέπει να πάρουμε σαν σ.μ.π. της X_0 την $\pi_0(x) = 1/8$, για κάθε $x \in \mathbb{X}_1$. Είναι εύκολο να περιγράψουμε και τις πιθανότητες μετάβασης της αλυσίδας. Για παράδειγμα, από την κατάσταση ΚΓΓ η αλυσίδα μπορεί να μεταβεί είτε στην ΓΓΚ με πιθανότητα $1/2$ (αν φέρουμε κορώνα) είτε στην ΚΓΓ με πιθανότητα $1/2$ (αν φέρουμε γράμματα). Επειδή το πείραμα τελειώνει όταν η αλυσίδα φτάσει στην ΚΓΚ, μπορούμε να θεωρήσουμε ότι από την ΚΓΚ παραμένουμε στην ίδια κατάσταση με πιθανότητα 1. Μπορείτε να συμπληρώσετε τον πίνακα πιθανοτήτων μετάβασης για εξάσκηση.

Μπορούμε όμως να περιγράψουμε το πείραμα και πιο οικονομικά, θεωρώντας σαν κατάσταση της αλυσίδας που το περιγράφει το αποτέλεσμα των δύο τελευταίων στριψιμάτων και προσθέτοντας μια κατάσταση στην οποία η αλυσίδα θα μεταβαίνει, όταν το πείραμα τελειώσει. Με αυτή την επιλογή ο χώρος καταστάσεων της αλυσίδας είναι ο $\mathbb{X} = \{KK, K\Gamma, \Gamma K, \Gamma\Gamma, T\}$, όπου η κατάσταση T κωδικοποιεί το τέλος του πειράματος. Η αρχική κατανομή της αλυσίδας θα προκύψει τώρα από το αποτέλεσμα των δύο πρώτων στριψιμάτων. Εφόσον το κέρμα είναι τίμιο, η αλυσίδα αρχικά θα βρεθεί σε μία από τις $KK, K\Gamma, \Gamma K$ ή $\Gamma\Gamma$ με πιθανότητα $1/4$ στην κάθε μία. Μπορούμε λοιπόν να περιγράψουμε την αρχική κατανομή της αλυσίδας ως ένα διάνυσμα γραμμής

$$\pi_0 = (\pi_0(KK), \pi_0(K\Gamma), \pi_0(\Gamma K), \pi_0(\Gamma\Gamma), \pi_0(T)) = (1/4, 1/4, 1/4, 1/4, 0).$$

Ο πίνακας πιθανοτήτων μετάβασης της αλυσίδας είναι ο

$$P = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Παράδειγμα 7 Το επιτραπέζιο παιχνίδι φιδάκι ανάμεσα σε δύο παίκτες μπορεί να περιγραφεί ως μια μαρκοβιανή αλυσίδα στον χώρο καταστάσεων

$$\mathbb{X} = \{1, 2, \dots, 100\} \times \{1, 2, \dots, 100\} \times \{A, B\}.$$

Το παιχνίδι βρίσκεται στην κατάσταση (m, n, b) , αν το πόνι του παίκτη A βρίσκεται στο τετράγωνο m , το πόνι του παίκτη B βρίσκεται στο τετράγωνο n και είναι η σειρά του παίκτη b να ρίξει το ζάρι. Η αρχική κατανομή της αλυσίδας, αν ο παίκτης A παίζει πρώτος, είναι η $\pi_0(x) = \delta_{(1,1,A)}(x)$. Θα ήταν κοπιώδες να περιγράψουμε τον πίνακα πιθανοτήτων μετάβασης αυτής της αλυσίδας σαν ένα τετραγωνικό 20.000×20.000 πίνακα, αλλά οι πιθανότητες μετάβασης μπορούν να περιγραφούν σαφώς. Για παράδειγμα $p((1, 1, A), (3, 1, B)) = 1/6$, αφού το παιχνίδι θα μεταβεί από την κατάσταση $(1, 1, A)$ στην κατάσταση $(3, 1, B)$, αν η ζαριά του παίκτη A είναι 2.

Παράδειγμα 8 Ένας μπασκετμπολίστας προπονείται στις ελεύθερες βολές. Κάθε φορά που εκτελεί μια βολή, η πιθανότητα να ευστοχήσει είναι $8/10$, αν έχει ευστοχήσει και στις δύο προηγούμενες προσπάθειές του, $5/10$, αν έχει αστοχήσει και στις δύο προηγούμενες προσπάθειες και $7/10$ σε κάθε άλλη περίπτωση. Ορίζουμε τη στοχαστική διαδικασία $\{X_n\}_{n \in \mathbb{N}}$, όπου $X_n = 1$, αν η n -οστή του προσπάθεια είναι εύστοχη και $X_n = 0$, αν η n -οστή του προσπάθεια είναι άστοχη. Είναι η $\{X_n\}_{n \in \mathbb{N}}$ μαρκοβιανή αλυσίδα;

Διαισθητικά περιμένει κανείς ότι η απάντηση είναι αρνητική. Η τιμή της X_n έχει την πληροφορία μόνο για το αποτέλεσμα της τελευταίας βολής, ενώ η πιθανότητα ευστοχίας στην επόμενη προσπάθεια εξαρτάται

από το αποτέλεσμα των δύο τελευταίων βολών. Για να αποδείξουμε αυστηρά ότι η $\{X_n\}_{n \in \mathbb{N}}$ δεν είναι μαρκοβιανή, αρκεί να παρατηρήσουμε ότι

$$\mathbb{P}[X_{n+1} = 1 | X_{n-1} = X_n = 1] = 8/10 \neq 7/10 = \mathbb{P}[X_{n+1} = 1 | X_{n-1} = 0, X_n = 1].$$

Αν όμως η $\{X_n\}_{n \in \mathbb{N}}$ ήταν μαρκοβιανή, οι δύο πιθανότητες στην παραπάνω σχέση θα ήταν ίσες, αφού και οι δύο θα συνέπιπταν με την $\mathbb{P}[X_{n+1} = 1 | X_n = 1]$.

Το πρόβλημα εδώ είναι ότι ο χώρος καταστάσεων της αλυσίδας $\mathbb{X} = \{0, 1\}$ είναι πολύ μικρός για να κωδικοποιήσει την πληροφορία που χρειαζόμαστε για να περιγράψουμε την κατανομή της επόμενης κατάστασης. Μπορούμε όμως, μεγαλώνοντας τον χώρο καταστάσεων, να περιγράψουμε την προπόνηση του αθλητή ως μια μαρκοβιανή αλυσίδα.

Αν για παράδειγμα ορίσουμε τη στοχαστική διαδικασία $\{Y_n\}_{n \in \mathbb{N}}$, με $Y_n = (X_n, X_{n+1})$, αυτή είναι μια μαρκοβιανή αλυσίδα στον χώρο καταστάσεων $\mathbb{X}' = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$ με πίνακα πιθανοτήτων μετάβασης

$$P = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 3/10 & 7/10 \\ 3/10 & 7/10 & 0 & 0 \\ 0 & 0 & 2/10 & 8/10 \end{pmatrix}. \quad \square$$

Συνηθίζουμε να λέμε ότι για μια μαρκοβιανή αλυσίδα παρελθόν και μέλλον είναι ανεξάρτητα, με δεδομένο το παρόν. Η μαρκοβιανή ιδιότητα (Markov property) είναι η μαθηματική διατύπωση αυτής της έκφρασης.

Θεώρημα 1 (μαρκοβιανή ιδιότητα) Έστω $\{X_n\}_{n \in \mathbb{N}_0}$ μαρκοβιανή αλυσίδα σε ένα αριθμήσιμο χώρο καταστάσεων \mathbb{X} με πίνακα πιθανοτήτων μετάβασης P . Για $k \in \mathbb{N}_0$, ορίζουμε τη στοχαστική διαδικασία $\{Y_n\}_{n \in \mathbb{N}_0}$ με τύπο $Y_n = X_{n+k}$. Δεδομένου ότι $X_k = x \in \mathbb{X}$, η διαδικασία $\{Y_n\}_{n \in \mathbb{N}_0}$ είναι μαρκοβιανή αλυσίδα στον \mathbb{X} , με αρχική κατανομή δ_x , πίνακα πιθανοτήτων μετάβασης P και είναι ανεξάρτητη από τις X_0, X_1, \dots, X_k .

Απόδειξη: Αρκεί να δείξουμε ότι, δεδομένου ότι $X_k = x$, για κάθε $n \in \mathbb{N}$ οι (Y_0, Y_1, \dots, Y_n) είναι ανεξάρτητες από τις X_0, X_1, \dots, X_k και η από κοινού τους κατανομή δίνεται από την (1.11) με αρχική κατανομή δ_x . Αρκεί επομένως να δείξουμε ότι, για κάθε $y = (y_0, \dots, y_n) \in \mathbb{X}^{n+1}$ και κάθε $z = (z_0, \dots, z_k) \in \mathbb{X}^{k+1}$

$$\mathbb{P}[\{(Y_0, \dots, Y_n) = y\} \cap A(z) | X_k = x] = \delta_x(y_0)p(y_0, y_1) \cdots p(y_{n-1}, y_n)\mathbb{P}[A(z) | X_k = x],$$

όπου $A(z) = \{(X_0, \dots, X_k) = z\}$. Πράγματι όμως,

$$\begin{aligned} \mathbb{P}[\{(Y_0, \dots, Y_n) = y\} \cap A(z) | X_k = x] &= \delta_x(y_0)\delta_x(z_k) \frac{\mathbb{P}[X_0 = z_0, \dots, X_k = x, X_{k+1} = y_1, \dots, X_{n+k} = y_n]}{\mathbb{P}[X_k = x]} \\ &= \frac{\delta_x(z_k)\pi_0(z_0)p(z_0, z_1) \cdots p(z_{k-1}, x)}{\mathbb{P}[X_k = x]} \delta_x(y_0)p(x, y_1) \cdots p(y_{n-1}, y_n) \\ &= \mathbb{P}[A(z) | X_k = x] \delta_x(y_0)p(y_0, y_1) \cdots p(y_{n-1}, y_n). \quad \square \end{aligned}$$

Βλέπουμε λοιπόν ότι, αν γνωρίζουμε την παρούσα θέση της αλυσίδας, δεδομένου δηλαδή ότι $X_k = x$, τότε το μέλλον της αλυσίδας, το οποίο εκφράζεται από την αλυσίδα $\{Y_n\}_{n \in \mathbb{N}_0}$, είναι ανεξάρτητο του παρελθόντος της, δηλαδή των τυχαίων μεταβλητών X_0, \dots, X_k και έχει μάλιστα τις ίδιες στατιστικές ιδιότητες όπως μια αλυσίδα που ξεκινά από το x με τις ίδιες πιθανότητες μετάβασης, P . \square

Όταν η αλυσίδα $\{X_n\}_{n \in \mathbb{N}_0}$ ξεκινά από την κατάσταση $x \in \mathbb{X}$, θα χρησιμοποιούμε τον συμβολισμό $\mathbb{P}_x[A]$

για την πιθανότητα ενός ενδεχομένου A που εξαρτάται από την τροχιά της αλυσίδας. Για παράδειγμα, αν $v = (v_0, \dots, v_n) \in \mathbb{X}^{n+1}$, έχουμε

$$\mathbb{P}_x[(X_0, \dots, X_n) = v] = \delta_x(v_0)p(v_0, v_1) \cdots p(v_{n-1}, v_n).$$

Χρησιμοποιώντας τη μαρκοβιανή ιδιότητα για $k = 0$, έχουμε ότι $\mathbb{P}[\cdot | X_0 = x] = \mathbb{P}_x[\cdot]$.

1.5 Ασκήσεις

Άσκηση 1 Θεωρήστε δύο ανεξάρτητες τυχαίες μεταβλητές A, Θ και ορίστε τη στοχαστική διαδικασία $\{X_t\}_{t \geq 0}$ με τύπο

$$X_t = A \sin(\omega t + \Theta),$$

όπου $\omega \in \mathbb{R}$. Αν η A ακολουθεί εκθετική κατανομή με ρυθμό $\lambda = 1$ και η Θ ακολουθεί ομοιόμορφη κατανομή στο $[0, 2\pi]$, υπολογίστε τις $\mathbb{E}[X_t]$ και $\mathbb{E}[X_t X_s]$, για κάθε $s, t \geq 0$.

Άσκηση 2 Αν $\{X_t\}_{t \geq 0}$ είναι η στοχαστική διαδικασία του Παραδείγματος 3, ορίζουμε

$$Y_t = e^{-t} X_{e^{2t}}.$$

Δείξτε ότι η $\{Y_t\}_{t \geq 0}$ είναι διαδικασία Gauss, με $m(t) = \mathbb{E}[Y_t] = 0$ και $\rho(s, t) = \text{Cov}(Y_t, Y_s) = e^{-|t-s|}$.

Άσκηση 3 Θεωρήστε μια διαδικασία Gauss $\{X_t\}_{t \geq 0}$, με $m(t) = \mathbb{E}[X_t] = 0$, για κάθε $t \geq 0$ και

$$\rho(s, t) = \text{Cov}(X_t, X_s) = \frac{1}{2}(t^{2H} + s^{2H} - |t - s|^{2H}), \quad s, t \geq 0,$$

για κάποιο $H \in (0, 1)$.

α) Δείξτε ότι για κάθε $t \geq 0$ έχουμε $X_t \sim \mathcal{N}(0, t^{2H})$.

β) Υπολογίστε για $t \geq 0$ και $h > 0$ τη μέση τιμή και τη διασπορά της προσαύξησης $X_{t+h} - X_t$ και δείξτε ότι η $X_{t+h} - X_t$ ακολουθεί την ίδια κατανομή με την X_h .

γ) Δείξτε ότι για κάθε $t, h > 0$, η $X_{t+h} - X_t$ είναι ανεξάρτητη από την X_t , αν και μόνο αν $H = 1/2$.

δ) Ποια γνωστή μας στοχαστική διαδικασία είναι η $\{X_t\}_{t \geq 0}$ όταν $H = 1/2$;

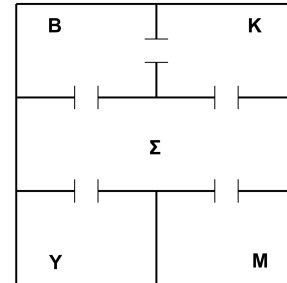
ε) Ορίζουμε για κάθε $t \geq 0$: $Y_t = \alpha^{-H} X_{\alpha t}$, όπου α θετικός πραγματικός αριθμός. Δείξτε ότι $\mathbb{E}[Y_t] = \mathbb{E}[X_t]$, για κάθε $t \geq 0$ και $\text{Cov}(Y_t, Y_s) = \text{Cov}(X_t, X_s)$, για κάθε $s, t \geq 0$. Πώς ερμηνεύετε αυτό το αποτέλεσμα;

Άσκηση 4 α) Αντικαταστήστε τα σύμβολα * με αριθμούς, ώστε ο πίνακας

$$P = \begin{pmatrix} 0 & 3/4 & * & 0 & 0 \\ 3/4 & 0 & 0 & 1/8 & * \\ 1/2 & 1/4 & 1/4 & * & * \\ * & 3/5 & 1/5 & 1/5 & * \\ 0 & 0 & 1/10 & 1/5 & * \end{pmatrix}$$

να είναι πίνακας πιθανοτήτων μετάβασης μιας μαρκοβιανής αλυσίδας $\{X_n\}_{n \in \mathbb{N}_0}$ στον χώρο καταστάσεων $\mathbb{X} = \{1, 2, \dots, 7\}$ με πιθανότητες μετάβασης $p(i, j)$ για κάθε $i, j \in \mathbb{X}$.

Άσκηση 5 Στο διπλανό σχήμα φαίνεται η κάτοψη ενός σπιτιού με πέντε δωμάτια: κουζίνα (K), βιβλιοθήκη (B), σαλόνι (Σ), υπνοδωμάτιο (Υ) και μπάνιο (M), καθώς και οι πόρτες που τα συνδέουν. Ένα έντομο που ζει στο σπίτι κάθε βράδυ διασχίζει τυχαία μία από τις πόρτες του δωματίου στο οποίο βρίσκεται και παραμένει στο δωμάτιο που οδηγεί η πόρτα μέχρι το επόμενο βράδυ. Αρχικά το έντομο βρίσκεται στο μπάνιο. Αν $\{X_n\}_{n \in \mathbb{N}_0}$ είναι η μαρκοβιανή αλυσίδα στον χώρο καταστάσεων $\mathbb{X} = \{K, B, \Sigma, \Upsilon, M\}$ που περιγράφει τη θέση του εντόμου μετά από n βράδυα, βρείτε τον πίνακα πιθανοτήτων μετάβασής της.



Άσκηση 6 Έχουμε πέντε χαρτιά της τράπουλας, τα τέσσερα είναι επτά κούπα και το ένα ρήγας σπαθί. Τα απλώνουμε στη σειρά σε ένα τραπέζι και σε κάθε βήμα επιλέγουμε ένα από τα δύο ακραία χαρτιά (το αριστερότερο με πιθανότητα $2/3$, το δεξιότερο με πιθανότητα $1/3$) και το τοποθετούμε στη μέση. Κατασκευάστε έναν κατάλληλο χώρο καταστάσεων και τις αντίστοιχες πιθανότητες μετάβασης μιας μαρκοβιανής αλυσίδας που θα περιέγραφε την θέση του ρήγα.

Άσκηση 7 Ρίχνουμε ένα ζάρι μέχρι να φέρουμε πέντε συνεχόμενες φορές έξι. Περιγράψτε μια μαρκοβιανή αλυσίδα στον χώρο καταστάσεων $\mathbb{X} = \{0, 1, 2, 3, 4, 5\}$ που θα μοντελοποιούσε αυτό το παιχνίδι. Κάντε το ίδιο για την περίπτωση στην οποία ρίχνουμε το ζάρι μέχρι να εμφανιστεί η ακολουθία 65656.

Άσκηση 8 Επαναλαμβάνουμε ρίψεις ενός ζαριού και για $n \in \mathbb{N}$ συμβολίζουμε με S_n το άθροισμα των n πρώτων ζαριών μας. Αν $X_n = S_n \pmod{5}$, δείξτε ότι η $\{X_n\}_{n \in \mathbb{N}}$ είναι μια μαρκοβιανή αλυσίδα και βρείτε τον πίνακα πιθανοτήτων μετάβασής της.

Άσκηση 9 Αυτή η άσκηση μας διδάσκει πώς να προσομοιώσουμε μια κατανομή στον \mathbb{X} με τη βοήθεια μιας γεννήτριας τυχαίων αριθμών. Έστω $\mathbb{X} = \{v_1, v_2, \dots\}$ ένας αριθμησιμος χώρος καταστάσεων. Για μια σ.μ.π. $\{p_i\}_{i \in \mathbb{N}}$ στον \mathbb{X} ορίζουμε τη συνάρτηση:

$$\Phi_p : (0, 1] \rightarrow \mathbb{X}, \quad \Phi_p(x) = v_k, \quad \text{αν } \sum_{j=1}^{k-1} p_i < x \leq \sum_{j=1}^k p_i.$$

Δείξτε ότι, αν η τυχαία μεταβλητή U έχει ομοιόμορφη κατανομή στο $[0, 1]$, τότε η $\Phi_p(U)$ έχει σ.μ.π. $\{p_i\}_{i \in \mathbb{N}}$.

Άσκηση 10 Θεωρήστε έναν αριθμησιμο χώρο \mathbb{X} και μια συνάρτηση $\Phi : \mathbb{X} \times [0, 1] \rightarrow \mathbb{X}$. Αν $\{U_n\}_{n \in \mathbb{N}}$ είναι μια ακολουθία από ανεξάρτητες, ισόνομες τυχαίες μεταβλητές, με ομοιόμορφη κατανομή στο $[0, 1]$ και ορίσουμε αναδρομικά τη στοχαστική διαδικασία $\{X_n\}_{n \in \mathbb{N}_0}$ ως

$$X_0 = x \in \mathbb{X}, \quad X_n = \Phi(X_{n-1}, U_n) \text{ για } n \in \mathbb{N},$$

τότε η $\{X_n\}_{n \in \mathbb{N}_0}$ είναι μαρκοβιανή αλυσίδα. Ποιες είναι οι πιθανότητες μετάβασης αυτής της αλυσίδας; Με τη βοήθεια και της προηγούμενης άσκησης, εξηγήστε πώς μπορούμε να προσομοιώσουμε μια μαρκοβιανή αλυσίδα με δεδομένες πιθανότητες μετάβασης.

Άσκηση 11 Σ' ένα ράφι της βιβλιοθήκης σας υπάρχουν τρία βιβλία: Algebra, Basic Topology, Calculus, που θα συμβολίζουμε με A, B, C για συντομία. Κάθε πρωί παίρνετε τυχαία ένα βιβλίο από τη θέση του με πιθανότητες p, q, r , αντίστοιχα. Όταν τελειώνετε το διάβασμά σας για την ημέρα, το ξαναβάζετε στο ράφι στην αριστερότερη θέση. Η διάταξη των βιβλίων είναι μια μαρκοβιανή αλυσίδα στον χώρο \mathbb{X} των μεταθέσεων των συμβόλων $\{A, B, C\}$. Ποιος είναι ο πίνακας πιθανοτήτων μετάβασης αυτής της αλυσίδας;

Άσκηση 12 Στο μοντέλο διάχυσης του Ehrenfest, N σωματίδια τοποθετούνται σε ένα δοχείο με δύο διαμερίσματα, A και B. Σε κάθε βήμα επιλέγουμε τυχαία ένα από τα N σωματίδια και του αλλάζουμε διαμέρισμα. Έστω $\{X_n\}_{n \in \mathbb{N}_0}$ η μαρκοβιανή αλυσίδα στον χώρο καταστάσεων $\mathbb{X} = \{0, 1, \dots, N\}$ που περιγράφει το πλήθος των σωματιδίων στο διαμέρισμα A μετά από n βήματα. Ποιες είναι οι πιθανότητες μετάβασης της $\{X_n\}_n$;

Άσκηση 13 Μοντελοποιήστε το παιχνίδι του παραδείγματος 6 ως μια μαρκοβιανή αλυσίδα σε έναν χώρο \mathbb{X} με 4 μόνο καταστάσεις.

Άσκηση 14 Ένας παντοπώλης εφοδιάζεται κάθε πρωί με δύο πακέτα μπισκότα Alfajor. Η ημερήσια ζήτηση για τα μπισκότα αυτά είναι μια τυχαία μεταβλητή που ακολουθεί γεωμετρική κατανομή με παράμετρο $p = 1/2$. Αν χθες βράδυ δεν είχαν μείνει καθόλου μπισκότα Alfajor και $\{X_n\}_{n \in \mathbb{N}}$ είναι το πλήθος των πακέτων που υπάρχει στο παντοπωλείο το βράδυ της n -οστής ημέρας, δείξτε ότι η $\{X_n\}_{n \in \mathbb{N}}$ είναι μαρκοβιανή αλυσίδα και βρείτε τις πιθανότητες μετάβασης της.

Άσκηση 15 Θεωρήστε δύο ακολουθίες $\{X_n, Y_n\}_{0 \leq n \leq N-1}$ από ανεξάρτητα, τυχαία, δεκαδικά ψηφία. Σχηματίστε τους ακεραίους

$$X = X_0 + X_1 \times 10 + \dots + X_{N-1} \times 10^{N-1}, \quad Y = Y_0 + Y_1 \times 10 + \dots + Y_{N-1} \times 10^{N-1}$$

και προσθέστε τους, όπως μάθαμε στο δημοτικό, ξεκινώντας από τα ψηφία των μονάδων X_0, Y_0 , συνεχίζοντας με τα ψηφία των δεκάδων X_1, Y_1 κ.λπ., μεταφέροντας το κρατούμενο, όπου χρειάζεται. Αν $C_n \in \{0, 1\}$ είναι το κρατούμενο της πρόσθεσης που μεταφέρουμε στο n -οστό βήμα, δείξτε ότι η $\{C_n\}_{0 \leq n \leq N-1}$ είναι μια μαρκοβιανή αλυσίδα στον χώρο καταστάσεων $\mathbb{X} = \{0, 1\}$ και βρείτε τον πίνακα πιθανοτήτων μετάβασης της αλυσίδας.

1.6 Αριθμητικά πειράματα

Άσκηση 16 Κατεβάστε και εγκαταστήστε τη γλώσσα Python 2.7.7. Μπορείτε να βρείτε τα σχετικά αρχεία εδώ. Ανάλογα με το λειτουργικό σύστημα του Η/Υ σας, μπορείτε να βρείτε και ολοκληρωμένα πακέτα. Αν ο υπολογιστής σας δεν τρέχει το λειτουργικό σύστημα Linux, πιθανά να βρείτε χρήσιμο να εγκαταστήσετε έναν εξομοιωτή, σύμφωνα με τις οδηγίες που θα βρείτε εδώ.

Άσκηση 17 Κατεβάστε το πρόγραμμα `simple_markov_chain_lib.py` και αποθηκεύστε το στον κατάλογο που θα δουλέψετε. Σ' αυτή την φάση δεν χρειάζεται καν να το ανοίξετε. Το πρόγραμμα αυτό υλοποιεί τον αλγόριθμο της Άσκησης 10. Θα το χρησιμοποιούμε σαν βιβλιοθήκη στα επόμενα προγράμματα που θα φτιάξουμε.

Άσκηση 18 Κατεβάστε και τρέξτε το πρόγραμμα `test.py`. Το πρόγραμμα αυτό προσομοιώνει τα πρώτα δέκα βήματα μιας αλυσίδας που κινείται στον χώρο καταστάσεων $\mathbb{X} = \{1, 2, 3, 4, 5\}$ με πίνακα πιθανοτήτων μετάβασης

$$P = \begin{pmatrix} 0 & 1/2 & 1/2 & 0 & 0 \\ 1/3 & 0 & 0 & 2/3 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1/2 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Ξεκινώντας από την κατάσταση 1. Τρέξτε το πρόγραμμα μερικές φορές και στη συνέχεια φτιάξτε ένα πρόγραμμα που προσομοιώνει τα είκοσι πρώτα βήματα της Άσκησης 11.

Κεφάλαιο 2

Μαρκοβιανές αλυσίδες

2.1 Εισαγωγή

Σ' αυτό το κεφάλαιο θα αναλύσουμε τις βασικές ιδιότητες μιας μαρκοβιανής αλυσίδας. Θα δούμε πώς μπορούμε να βρούμε την κατανομή της κατάστασής της οποιαδήποτε χρονική στιγμή και πώς μπορούμε να απλοποιήσουμε την ανάλυση μιας μαρκοβιανής αλυσίδας, διαμερίζοντας τον χώρο καταστάσεων σε κλάσεις επικοινωνίας. Στη συνέχεια θα εισαγάγουμε την έννοια του χρόνου διακοπής και θα δείξουμε την ιδιότητα που έχουν οι μαρκοβιανές αλυσίδες να ανανεώνονται έπειτα από τέτοιους χρόνους, ξεχνώντας το παρελθόν τους. Τέλος, θα δούμε κάποιες βασικές ιδιότητες που έχουν οι τροχιές της αλυσίδας, όπως η επαναληπτικότητα και η παροδικότητα. Μια πολύ καλή αναφορά που περιέχει παρόμοιο υλικό είναι το [7]. Για την προσομοίωση μαρκοβιανών αλυσίδων μπορείτε να δείτε και το [4].

2.2 Πιθανότητες μετάβασης ανώτερης τάξης

Ας θεωρήσουμε μια μαρκοβιανή αλυσίδα $\{X_n\}_{n \in \mathbb{N}_0}$ με τιμές σ' έναν αριθμήσιμο χώρο καταστάσεων \mathbb{X} , που έχει αρχική κατανομή π_0 και πίνακα πιθανοτήτων μετάβασης $P = \{p(x, y)\}_{x, y \in \mathbb{X}}$. Ο στόχος μας σ' αυτή την παράγραφο είναι να μάθουμε να υπολογίζουμε την κατανομή της κατάστασης της αλυσίδας X_n , για κάθε $n \in \mathbb{N}$.

Μπορούμε να αναγάγουμε τον υπολογισμό της κατανομής της X_{n+1} στον υπολογισμό της κατανομής της X_n , χρησιμοποιώντας τον τύπο της ολικής πιθανότητας. Αν για $n \in \mathbb{N}_0$ συμβολίζουμε με π_n τη σ.μ.π. της X_n , τότε για κάθε $y \in \mathbb{X}$ έχουμε

$$\pi_{n+1}(y) = \mathbb{P}[X_{n+1} = y] = \sum_{x \in \mathbb{X}} \mathbb{P}[X_n = x] \mathbb{P}[X_{n+1} = y | X_n = x] = \sum_{x \in \mathbb{X}} \pi_n(x) p(x, y). \quad (2.1)$$

Παρατηρήστε ότι, αν ο \mathbb{X} είναι πεπερασμένος, με $\mathbb{X} = \{v_1, v_2, \dots, v_N\}$, μπορούμε να αναπαριστούμε μια σ.μ.π. π ως ένα διάνυσμα γραμμής

$$\pi = (\pi(v_1), \pi(v_2), \dots, \pi(v_N))$$

και τον πίνακα πιθανοτήτων μετάβασης P ως ένα $N \times N$ τετραγωνικό πίνακα, όπως στην (1.7). Η (2.1) μας δίνει τότε ότι

$$\pi_{n+1} = \pi_n P, \quad \text{για κάθε } n \in \mathbb{N}_0, \quad (2.2)$$

όπου η πράξη στο δεξί μέλος είναι ο συνήθης πολλαπλασμός πινάκων. Ορμώμενοι από αυτή την παρατήρηση, μπορούμε να ορίσουμε, για έναν αριθμήσιμο χώρο καταστάσεων \mathbb{X} , τον πολλαπλασιασμό μιας σ.μ.π. π με

ένα στοχαστικό πίνακα $P = \{p(x, y)\}_{x, y \in \mathbb{X}}$, ως

$$\pi P(y) = \sum_{x \in \mathbb{X}} \pi(x) p(x, y), \quad \text{για κάθε } y \in \mathbb{X}.$$

Το αποτέλεσμα αυτού του πολλαπλασιασμού είναι μια σ.μ.π., αφού $\pi P(y) \geq 0$, για κάθε $y \in \mathbb{X}$, ενώ χρησιμοποιώντας το θεώρημα Fubini-Tonelli για να εναλλάξουμε τα δύο αθροίσματα μη αρνητικών όρων

$$\sum_{y \in \mathbb{X}} \pi P(y) = \sum_{y \in \mathbb{X}} \sum_{x \in \mathbb{X}} \pi(x) p(x, y) = \sum_{x \in \mathbb{X}} \pi(x) \sum_{y \in \mathbb{X}} p(x, y) = \sum_{x \in \mathbb{X}} \pi(x) = 1.$$

Ανάλογα μπορούμε να ορίσουμε τον πολλαπλασιασμό δύο στοχαστικών πινάκων, $P = \{p(x, y)\}_{x, y \in \mathbb{X}}$ και $Q = \{q(x, y)\}_{x, y \in \mathbb{X}}$, ως $PQ = \{r(x, y)\}_{x, y \in \mathbb{X}}$, με

$$r(x, y) = \sum_{z \in \mathbb{X}} p(x, z) q(z, y), \quad \text{για κάθε } y \in \mathbb{X}.$$

Το αποτέλεσμα είναι ένας στοχαστικός πίνακας, αφού $r(x, y) \geq 0$, για κάθε $x, y \in \mathbb{X}$, ενώ από το Θεώρημα Fubini-Tonelli, για κάθε $x \in \mathbb{X}$, έχουμε

$$\sum_{y \in \mathbb{X}} r(x, y) = \sum_{y \in \mathbb{X}} \sum_{z \in \mathbb{X}} p(x, z) q(z, y) = \sum_{z \in \mathbb{X}} p(x, z) \sum_{y \in \mathbb{X}} q(z, y) = \sum_{z \in \mathbb{X}} p(x, z) = 1.$$

Οι πράξεις αυτές γενικεύουν τον συνήθη πολλαπλασιασμό πινάκων και μπορεί εύκολα να ελεγχθεί ότι έχουν την προσεταιριστική ιδιότητα, δηλαδή

$$(\pi P)Q = \pi(PQ) \quad \text{και} \quad (PQ)R = P(QR)$$

για κάθε σ.μ.π. π και οποιουδήποτε στοχαστικούς πίνακες P, Q, R . Ο στοχαστικός πίνακας $I = \{i(x, y)\}_{x, y \in \mathbb{X}}$, με στοιχεία $i(x, y) = \delta_x(y)$ είναι ταυτοτικό στοιχείο αυτού του πολλαπλασιασμού. Για κάθε στοχαστικό πίνακα P , έχουμε $PI = IP = P$.

Ορίζουμε τη δύναμη $P^n = \{p^{(n)}(x, y)\}_{x, y \in \mathbb{X}}$ ενός στοχαστικού πίνακα P για $n \in \mathbb{N}_0$ αναδρομικά ως εξής. $P^0 = I$ και $P^n = P^{n-1}P$, για $n \in \mathbb{N}$.

Με αυτούς τους ορισμούς η (2.2) παραμένει σε ισχύ ακόμα κι αν ο \mathbb{X} είναι άπειρος και εύκολα ελέγχεται επαγωγικά ότι

$$\pi_n = \pi_0 P^n. \quad (2.3)$$

Από τον ορισμό του πολλαπλασιασμού πινάκων,

$$p^{(n)}(x, y) = \sum_{x_1, \dots, x_{n-1} \in \mathbb{X}} \cdots \sum p(x, x_1) \cdots p(x_{n-1}, y), \quad \text{για κάθε } x, y \in \mathbb{X}. \quad (2.4)$$

Με τη βοήθεια της προηγούμενης σχέσης η (2.3) γράφεται και ως

$$\pi_n(y) = \sum_{x_0 \in \mathbb{X}} \pi_0(x_0) p^{(n)}(x_0, y) = \sum_{x_0, \dots, x_{n-1} \in \mathbb{X}} \pi_0(x_0) p(x_0, x_1) \cdots p(x_{n-1}, y), \quad \text{για κάθε } y \in \mathbb{X}.$$

Στην ίδια σχέση θα μπορούσαμε να είχαμε καταλήξει και από την (1.11). Η π_n είναι η σ.μ.π. της X_n , επομένως μπορεί να βρεθεί αθροίζοντας την από κοινού σ.μ.π. των (X_0, X_1, \dots, X_n) ως προς όλες τις τιμές που μπορούν να πάρουν οι X_0, X_1, \dots, X_{n-1} .

Ας θεωρήσουμε τώρα ένα $k \in \mathbb{N}_0$. Από τη μαρκοβιανή ιδιότητα (Θεώρημα 1), δεδομένου ότι $X_k = x$, η

διαδικασία $\{X_{n+k}\}_{n \in \mathbb{N}_0}$ έχει αρχική κατανομή δ_x και πίνακα πιθανοτήτων μετάβασης P . Έτσι, η (2.3) δίνει ότι

$$\mathbb{P}[X_{n+k} = y | X_k = x] = p^{(n)}(x, y), \quad \text{για κάθε } k \in \mathbb{N}_0, x, y \in \mathbb{X}. \quad (2.5)$$

Ορισμός: Αν ο P είναι ο πίνακας πιθανοτήτων μετάβασης μιας μαρκοβιανής αλυσίδας $\{X_n\}_n$, ονομάζουμε τα στοιχεία $p^{(n)}(x, y)$ του P^n *πιθανότητες μετάβασης n -οστής τάξης* της αλυσίδας.

Από την προσεταιριστική ιδιότητα του πολλαπλασιασμού έχουμε ότι $P^{m+n} = P^m P^n$, για κάθε $m, n \in \mathbb{N}_0$. Αν γράψουμε αυτή την ταυτότητα για τα στοιχεία των παραπάνω στοχαστικών πινάκων, παίρνουμε ότι οι πιθανότητες μετάβασης ανώτερης τάξης μιας μαρκοβιανής αλυσίδας ικανοποιούν τις εξισώσεις Chapman-Kolmogorov (Chapman-Kolmogorov equations):

$$p^{(m+n)}(x, y) = \sum_{z \in \mathbb{X}} p^{(m)}(x, z) p^{(n)}(z, y), \quad \text{για κάθε } m, n \in \mathbb{N}_0, x, y \in \mathbb{X}. \quad (2.6)$$

Λαμβάνοντας υπόψιν την (2.5), οι εξισώσεις Chapman-Kolmogorov είναι ο τύπος της ολικής πιθανότητας για το ενδεχόμενο $\{X_{m+n} = y\}$, ανάλογα με την κατάσταση της αλυσίδας την ενδιαμέση χρονική στιγμή m . Πράγματι, αν θυμηθούμε ότι η δεσμευμένη πιθανότητα $\mathbb{P}[\cdot | X_0 = x]$ είναι κι αυτή ένα μέτρο πιθανότητας, τότε ο τύπος ολικής πιθανότητας δίνει

$$\begin{aligned} \mathbb{P}[X_{m+n} = y | X_0 = x] &= \sum_{z \in \mathbb{X}} \mathbb{P}[X_m = z | X_0 = x] \mathbb{P}[X_{m+n} = y | X_m = z, X_0 = x] \\ &= \sum_{z \in \mathbb{X}} \mathbb{P}[X_m = z | X_0 = x] \mathbb{P}[X_{m+n} = y | X_m = z] \end{aligned}$$

από τη μαρκοβιανή ιδιότητα. Σύμφωνα με την (2.5), αυτές είναι ακριβώς οι εξισώσεις Chapman-Kolmogorov.

Παράδειγμα 9 Ας θεωρήσουμε μια μαρκοβιανή αλυσίδα σ' έναν χώρο με δύο καταστάσεις $\mathbb{X} = \{a, b\}$ και πίνακα πιθανοτήτων μετάβασης

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix},$$

όπου $p, q \in (0, 1)$. Επομένως, όταν η αλυσίδα βρίσκεται στην κατάσταση a , μεταβαίνει στην κατάσταση b με πιθανότητα p και παραμένει στην κατάσταση a με πιθανότητα $1-p$, ενώ, όταν η αλυσίδα βρίσκεται στην κατάσταση b , μεταβαίνει στην κατάσταση a με πιθανότητα q και παραμένει στην κατάσταση b με πιθανότητα $1-q$. Δεν κάνουμε κάποια υπόθεση για την αρχική κατανομή της αλυσίδας. Από την (2.3) έχουμε ότι

$$(\pi_n(a), \pi_n(b)) = (\pi_0(a), \pi_0(b)) P^n. \quad (2.7)$$

Προκειμένου να υπολογίσουμε τον P^n για οποιαδήποτε τιμή του n , μπορούμε να διαγωνιοποιήσουμε τον πίνακα P . Η χαρακτηριστική εξίσωση $|\lambda I - P| = 0$ δίνει ότι οι ιδιοτιμές του P είναι οι $\lambda_1 = 1$ και $\lambda_2 = 1 - (p + q)$. Τα αντίστοιχα ιδιοδιανύσματα είναι

$$u_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{και} \quad u_2 = \begin{pmatrix} -p \\ q \end{pmatrix}.$$

Επομένως,

$$P = \begin{pmatrix} 1 & -p \\ 1 & q \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} 1 & -p \\ 1 & q \end{pmatrix}^{-1}$$

και άρα,

$$P^n = \begin{pmatrix} 1 & -p \\ 1 & q \end{pmatrix} \begin{pmatrix} 1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} \begin{pmatrix} 1 & -p \\ 1 & q \end{pmatrix}^{-1} = \frac{1}{p+q} \begin{pmatrix} q + p\lambda_2^n & p - p\lambda_2^n \\ q - q\lambda_2^n & p + q\lambda_2^n \end{pmatrix}.$$

Αντικαθιστώντας στην (2.7), παίρνουμε ότι

$$(\pi_n(a), \pi_n(b)) = \left(\frac{q}{p+q}, \frac{p}{p+q} \right) + \lambda_2^n \frac{p\pi_0(a) - q\pi_0(b)}{p+q} (1, -1).$$

Αξίζει σ' αυτό το σημείο να διερευνήσουμε την απάντηση που βρήκαμε. Προσέξτε ότι, εφόσον $p, q \in (0, 1)$ έχουμε

$$0 < p + q < 2 \Leftrightarrow |\lambda_2| < 1.$$

Επομένως, η εξάρτηση της π_n από την αρχική κατανομή της αλυσίδας φθίνει εκθετικά και σε βάθος χρόνου, καθώς δηλαδή $n \rightarrow \infty$, έχουμε

$$(\mathbb{P}[X_n = a], \mathbb{P}[X_n = b]) \rightarrow \left(\frac{q}{p+q}, \frac{p}{p+q} \right),$$

ανεξάρτητα από την αρχική κατανομή της αλυσίδας. Θα δούμε παρακάτω ότι αυτό το συμπέρασμα είναι τυπικό για μαρκοβιανές αλυσίδες σε πεπερασμένους χώρους καταστάσεων, στις οποίες δεν απαντώνται κάποιες συγκεκριμένες παθολογίες που θα γνωρίσουμε.

Παράδειγμα 10 Ένα ηλεκτρονικό ζάρι είναι πειραγμένο έτσι ώστε, κάθε φορά που το ρίχνουμε να φέρνει την ένδειξη που έφερε στην προηγούμενη ζαριά με πιθανότητα $1/2$ και οποιαδήποτε από τις άλλες ενδείξεις με πιθανότητα $1/10$. Αν στην πρώτη ζαριά μας φέρουμε 6, ποια είναι η πιθανότητα να φέρουμε 6 και στην n -οστή μας ζαριά;

Για $n \in \mathbb{N}$ ορίζουμε την τυχαία μεταβλητή X_n ώστε να είναι ίση με 1, αν η n -οστή μας ζαριά είναι 6 και ίση με 0, αν η n -οστή μας ζαριά δεν είναι 6. Η $\{X_n\}_{n \in \mathbb{N}}$ είναι μια μαρκοβιανή αλυσίδα στον χώρο καταστάσεων $\mathbb{X} = \{0, 1\}$ με αρχική κατανομή $\pi_1 = (\pi_1(0), \pi_1(1)) = (0, 1)$ και πίνακα πιθανοτήτων μετάβασης

$$P = \begin{pmatrix} 9/10 & 1/10 \\ 1/2 & 1/2 \end{pmatrix}.$$

Οι ιδιοτιμές του P είναι $\lambda_1 = 1$ και $\lambda_2 = 2/5$. Μπορούμε να χρησιμοποιήσουμε το αποτέλεσμα του προηγούμενου παραδείγματος για να βρούμε την ζητούμενη πιθανότητα $\mathbb{P}[X_n = 1] = \pi_n(1)$. Εναλλακτικά, από την (2.3), έχουμε ότι

$$\pi_n = \pi_1 P^{n-1} = \pi_1 U \begin{pmatrix} 1 & 0 \\ 0 & (\frac{2}{5})^{n-1} \end{pmatrix} U^{-1}.$$

Παρατηρήστε τώρα ότι στο δεξί μέλος της παραπάνω σχέσης το n εμφανίζεται μόνο στον εκθέτη του $2/5$. Οι πράξεις που θα κάναμε για να υπολογίσουμε το δεξί μέλος είναι γραμμικοί συνδυασμοί των στοιχείων των πινάκων που εμφανίζονται εκεί. Επομένως, το αποτέλεσμα για την $\pi_n(1)$ θα ήταν

$$\pi_n(1) = a + b \left(\frac{2}{5} \right)^n, \quad \text{για κάθε } n \in \mathbb{N},$$

για κάποιες σταθερές $a, b \in \mathbb{R}$. Οι σταθερές αυτές μπορούν να υπολογιστούν εύκολα, αφού $\pi_1(1) = 1$ και $\pi_2(1) = \mathbb{P}[X_2 = 1 | X_1 = 1] = p(1, 1) = 1/2$. Έχουμε λοιπόν τελικά,

$$\pi_n(1) = \mathbb{P}[X_n = 1] = \frac{1}{6} + \frac{1}{3} \left(\frac{2}{5} \right)^{n-2}.$$

2.3 Δομή του χώρου των καταστάσεων

Σε αυτή την παράγραφο θα δούμε πώς μπορούμε να οργανώσουμε τις καταστάσεις του χώρου μιας μαρκοβιανής αλυσίδας έτσι ώστε να διευκολύνουμε την κατανόηση της συμπεριφοράς της. Θα δούμε ότι ο χώρος διαμερίζεται σε κλάσεις επικοινωνίας (communication classes) εντός των οποίων η αλυσίδα είναι πιο εύκολο να αναλυθεί. Ας θεωρήσουμε λοιπόν μια μαρκοβιανή αλυσίδα $\{X_n\}_{n \in \mathbb{N}_0}$ στον χώρο καταστάσεων \mathbb{X} με πίνακα πιθανοτήτων μετάβασης $P = \{p(x, y)\}_{x, y \in \mathbb{X}}$.

Ορισμός: Θα λέμε ότι η κατάσταση $y \in \mathbb{X}$ είναι προσβάσιμη από την κατάσταση $x \in \mathbb{X}$ και θα συμβολίζουμε $x \rightarrow y$, αν υπάρχει $n \geq 0$ τέτοιο ώστε $p^{(n)}(x, y) > 0$.

Ορισμός: Θα λέμε ότι δύο καταστάσεις $x, y \in \mathbb{X}$ επικοινωνούν και θα συμβολίζουμε $x \leftrightarrow y$, αν $x \rightarrow y$ και $y \rightarrow x$.

Πρόταση: Η σχέση \leftrightarrow είναι μια σχέση ισοδυναμίας, έχει δηλαδή τις παρακάτω ιδιότητες:

1. Αυτοπαθής: $x \leftrightarrow x$, για κάθε $x \in \mathbb{X}$.
2. Συμμετρική: $x \leftrightarrow y \Rightarrow y \leftrightarrow x$, για κάθε $x, y \in \mathbb{X}$.
3. Μεταβατική: $(x \leftrightarrow u \text{ και } u \leftrightarrow y) \Rightarrow x \leftrightarrow y$, για κάθε $x, y, u \in \mathbb{X}$.

Απόδειξη: Η (1) ικανοποιείται γιατί $p^{(0)}(x, x) = \delta_x(x) = 1$, για κάθε $x \in \mathbb{X}$, ενώ η (2) είναι φανερή από τον ορισμό της σχέσης. Για να δείξουμε την (3), παρατηρήστε ότι

$$x \rightarrow u \Rightarrow \exists m \geq 0 : p^{(m)}(x, u) > 0 \quad \text{και} \quad u \rightarrow y \Rightarrow \exists n \geq 0 : p^{(n)}(u, y) > 0.$$

Από τις εξισώσεις Chapman-Kolmogorov (2.6) έχουμε τώρα

$$p^{(m+n)}(x, y) = \sum_{z \in \mathbb{X}} p^{(m)}(x, z)p^{(n)}(z, y) \geq p^{(m)}(x, u)p^{(n)}(u, y) > 0.$$

Επομένως $(x \rightarrow u \text{ και } u \rightarrow y) \Rightarrow x \rightarrow y$. Εναλλάσσοντας τον ρόλο των x, y έχουμε το ζητούμενο. \square

Όπως κάθε σχέση ισοδυναμίας, η \leftrightarrow διαμερίζει τον \mathbb{X} σε κλάσεις, που τις ονομάζουμε κλάσεις επικοινωνίας της αλυσίδας.

Ορισμός: Μια κλάση επικοινωνίας \mathcal{C} θα λέγεται ανοιχτή, αν υπάρχουν καταστάσεις $x \in \mathcal{C}$ και $y \notin \mathcal{C}$ τέτοιες ώστε $x \rightarrow y$. Αν μια κλάση δεν είναι ανοιχτή, θα λέγεται κλειστή.

Σύμφωνα με τον παραπάνω ορισμό, ανοιχτές είναι οι κλάσεις από τις οποίες η μαρκοβιανή αλυσίδα μπορεί να φύγει προς άλλες κλάσεις. Μάλιστα, αν η αλυσίδα φύγει από μια ανοιχτή κλάση, δεν μπορεί να ξαναγυρίσει σε αυτήν (Άσκηση 22). Αντίθετα, η αλυσίδα δεν μπορεί να φύγει από μια κλειστή κλάση. Αν η αλυσίδα ξεκινήσει ή βρεθεί, προϊόντος του χρόνου, σε μια κλειστή κλάση, θα παραμείνει σε αυτή την κλάση για πάντα. Αυτό μπορεί να απλοποιήσει την ανάλυση της αλυσίδας, αφού προκειμένου να καταλάβουμε τι θα συμβεί από τότε και μετά, αρκεί να μελετήσουμε την αλυσίδα περιορισμένη στην κλειστή κλάση.

Ορισμός: Μια μαρκοβιανή αλυσίδα λέγεται μη υποβιβάσιμη (irreducible), αν ολόκληρος ο χώρος καταστάσεων είναι μια (κλειστή αναγκαστικά) κλάση.

Παρατηρήστε ότι, αν περιορίσουμε μια μαρκοβιανή αλυσίδα σε μια κλειστή κλάση της \mathcal{C} , αν δηλαδή θεωρήσουμε τη μαρκοβιανή αλυσίδα με χώρο καταστάσεων την \mathcal{C} και τις ίδιες πιθανότητες μετάβασης, αυτή η αλυσίδα είναι μη υποβιβάσιμη.

Για να βρούμε τις κλάσεις επικοινωνίας μιας μαρκοβιανής αλυσίδας, είναι πολλές φορές χρήσιμη η ακόλουθη παρατήρηση.

Παρατήρηση: Αν $x, y \in \mathbb{X}$ και $x \neq y$, τότε $x \rightarrow y$, αν και μόνο αν υπάρχει μονοπάτι

$$x = x_0, x_1, x_2, \dots, x_{n-1}, x_n = y \in \mathbb{X}, \text{ τέτοιο ώστε } p(x_i, x_{i+1}) > 0, \text{ για } i = 0, 1, \dots, n-1.$$

Πράγματι, από την (2.4) έχουμε για κάθε $n \in \mathbb{N}$

$$p^{(n)}(x, y) = \sum_{x_1, \dots, x_{n-1} \in \mathbb{X}} \cdots \sum p(x, x_1) \cdots p(x_{n-1}, y).$$

Επειδή όλοι οι προσθετέοι στο δεξί μέλος είναι μη αρνητικοί, έχουμε $p^{(n)}(x, y) > 0$, αν και μόνο αν κάποιος από αυτούς είναι αυστηρά θετικός. Αυτό όμως είναι προφανώς ισοδύναμο με τον ισχυρισμό της παρατήρησης.

Μπορούμε τώρα να κατασκευάσουμε έναν γράφο με κορυφές τα σημεία του \mathbb{X} , προσθέτοντας την προσανατολισμένη ακμή από την u στην $v \neq u$, όταν $p(u, v) > 0$. Με βάση την παραπάνω παρατήρηση, μια κατάσταση y είναι προσβάσιμη από την x , αν και μόνο αν υπάρχει προσανατολισμένο μονοπάτι του γράφου που οδηγεί από την x στην y .

Παράδειγμα 11 Μια μαρκοβιανή αλυσίδα εξελίσσεται στον χώρο καταστάσεων $\mathbb{X} = \{1, 2, 3, 4, 5, 6, 7, 8\}$ και έχει πίνακα πιθανοτήτων μετάβασης

$$P = \begin{pmatrix} 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 1/4 & 1/4 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2/3 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 0 & 1/3 & 1/3 \\ 0 & 1/6 & 1/6 & 0 & 1/2 & 0 & 1/6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/4 & 3/4 \\ 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 1/2 \end{pmatrix}.$$

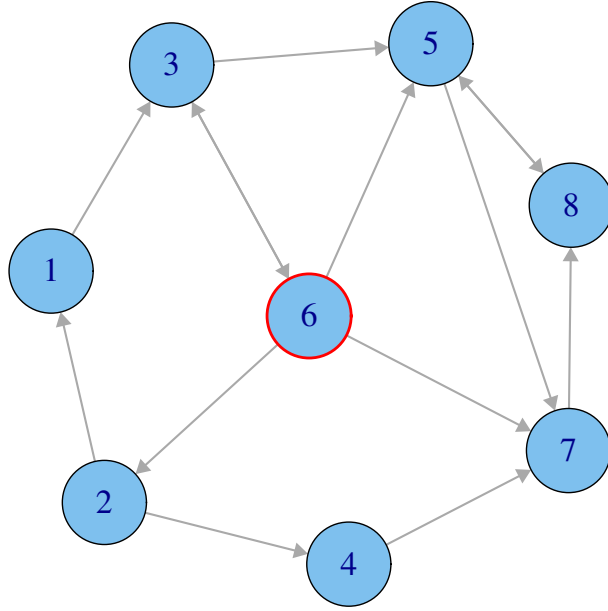
Για να βρούμε τις κλάσεις επικοινωνίας της αλυσίδας, κατασκευάζουμε τον γράφο των δυνατών μεταβάσεων που περιγράψαμε παραπάνω, όπως στο Σχήμα 11. Έχουμε π.χ. $1 \rightarrow 2$, μέσω του μονοπατιού $1, 3, 6, 2$. Επίσης $2 \rightarrow 1$, άμεσα. Επομένως $1 \leftrightarrow 2$. Όμοια, βρίσκουμε ότι η κατάσταση 1 επικοινωνεί με τις 3 και 6, ενώ δεν επικοινωνεί με καμία άλλη κατάσταση. Άρα η κλάση επικοινωνίας που περιέχει την κατάσταση 1 είναι η $\mathcal{C}_1 = \{1, 2, 3, 6\}$. Η κατάσταση 4 δεν επικοινωνεί με καμία άλλη, επομένως είναι το μόνο στοιχείο της κλάσης της $\mathcal{C}_2 = \{4\}$. Τέλος, οι καταστάσεις 5, 7 και 8 επικοινωνούν μεταξύ τους και αποτελούν την κλάση $\mathcal{C}_3 = \{5, 7, 8\}$. Έτσι, στο παράδειγμά μας, η διαμέριση του \mathbb{X} σε κλάσεις επικοινωνίας είναι

$$\mathbb{X} = \mathcal{C}_1 \cup \mathcal{C}_2 \cup \mathcal{C}_3 = \{1, 2, 3, 6\} \cup \{4\} \cup \{5, 7, 8\}.$$

Οι κλάσεις \mathcal{C}_1 και \mathcal{C}_2 είναι ανοιχτές, αφού $6 \rightarrow 7$ και $4 \rightarrow 7$, αντίστοιχα. Αντίθετα, η \mathcal{C}_3 είναι κλειστή. \square

2.4 Ισχυρή μαρκοβιανή ιδιότητα

Είδαμε στο Κεφάλαιο 1 ότι, αν γνωρίζουμε τη θέση $x \in \mathbb{X}$ μιας μαρκοβιανής αλυσίδας κάποια χρονική στιγμή $n \in \mathbb{N}_0$, από τότε και μετά η αλυσίδα θα εξελιχθεί ως μια μαρκοβιανή αλυσίδα που ξεκινά από το x , έχει τις ίδιες πιθανότητες μετάβασης και είναι ανεξάρτητη από ό,τι έχει προηγηθεί. Η ιδιότητα αυτή συνεχίζει να ισχύει, αν αντικαταστήσουμε την καθορισμένη χρονική στιγμή n με οποιαδήποτε στρατηγική διακοπής T που δεν μπορεί να χρησιμοποιήσει πληροφορία από το μέλλον. Ονομάζουμε αυτές τις στρατηγικές χρόνους διακοπής και την ιδιότητα μιας μαρκοβιανής αλυσίδας να ανανεώνεται σε αυτούς τους



Σχήμα 2.1: Ο γράφος των δυνατών μεταβάσεων του Παραδείγματος 11.

χρόνους, ισχυρή μαρκοβιανή ιδιότητα (strong Markov property). Αυτές οι έννοιες είναι το αντικείμενο της παρούσας παραγράφου.

Θεωρήστε μια μαρκοβιανή αλυσίδα $\{X_n\}_{n \in \mathbb{N}_0}$ με τιμές σ' έναν αριθμησιμο χώρο καταστάσεων \mathbb{X} .

Ορισμός: Θα λέμε ότι ένα ενδεχόμενο A ανήκει στην οικογένεια ενδεχομένων \mathcal{F}_n , αν η δείκτρια συνάρτησή του

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

είναι μια συνάρτηση των X_0, X_1, \dots, X_n και μόνο.

Με βάση τον παραπάνω ορισμό, ένα ενδεχόμενο A ανήκει στην κλάση \mathcal{F}_n όταν, προκειμένου να αποφασίσουμε αν συμβαίνει, αρκεί να ξέρουμε τις τιμές των X_0, X_1, \dots, X_n , δηλαδή την τροχιά της διαδικασίας μέχρι τη στιγμή n .

Παράδειγμα 12 Αν $B \subset \mathbb{X}$, το ενδεχόμενο $A_1 = \{X_n \in B\}$ ανήκει στην \mathcal{F}_n . Πράγματι, $\mathbb{1}_{A_1} = \mathbb{1}_B(X_n)$. Επίσης, το ενδεχόμενο η αλυσίδα να επισκεφτεί το B μέχρι και την χρονική στιγμή $n \in \mathbb{N}_0$, είναι στην \mathcal{F}_n . Πράγματι, αν συμβολίσουμε με A_2 το παραπάνω ενδεχόμενο,

$$\mathbb{1}_{A_2} = 1 - \mathbb{1}_{B^c}(X_0) \cdots \mathbb{1}_{B^c}(X_n).$$

Πρόταση: Οι οικογένειες ενδεχομένων $\{\mathcal{F}_n\}_{n \in \mathbb{N}_0}$ έχουν τις παρακάτω ιδιότητες:

1. Για κάθε $n \in \mathbb{N}_0$ έχουμε $\mathcal{F}_n \subset \mathcal{F}_{n+1}$.
2. $\Omega \in \mathcal{F}_n$, για κάθε $n \in \mathbb{N}_0$.
3. Αν $A \in \mathcal{F}_n$, τότε $A^c \in \mathcal{F}_n$.

4. Αν $A, B \in \mathcal{F}_n$, τότε $A \cap B \in \mathcal{F}_n$.

5. Αν $A, B \in \mathcal{F}_n$, τότε $A \cup B \in \mathcal{F}_n$.

Απόδειξη: Η (1) είναι προφανής, αφού μια συνάρτηση των X_0, \dots, X_n είναι και συνάρτηση των X_0, \dots, X_{n+1} , που δεν εξαρτάται από την τελευταία μεταβλητή. Για την (2) παρατηρήστε ότι η $\mathbb{1}_\Omega$ είναι σταθερή και ίση με 1, επομένως είναι (με τετριμμένο τρόπο) συνάρτηση των X_0, \dots, X_n , για κάθε $n \in \mathbb{N}_0$. Η ιδιότητα (3) ισχύει γιατί $\mathbb{1}_{A^c} = 1 - \mathbb{1}_A$, ενώ η (4) γιατί $\mathbb{1}_{A \cap B} = \mathbb{1}_A \mathbb{1}_B$. Τέλος, η (5) προκύπτει από την ταυτότητα $A \cup B = (A^c \cap B^c)^c$ και τις ιδιότητες (3) και (4), που ήδη αποδείξαμε. \square

Ορισμός: Θα λέμε μια τυχαία μεταβλητή $T : \Omega \rightarrow \mathbb{N}_0 \cup \{\infty\}$ *χρόνο διακοπής* (stopping time) της $\{X_n\}_{n \in \mathbb{N}_0}$, αν για κάθε $n \in \mathbb{N}_0$ το ενδεχόμενο $\{T = n\}$ ανήκει στην \mathcal{F}_n .

Μπορούμε να φανταζόμαστε έναν χρόνο διακοπής μιας μαρκοβιανής αλυσίδας ως μια στρατηγική σταματήματος της αλυσίδας που απαγορεύεται να δει το μέλλον. Η στρατηγική αυτή μπορεί να εξαρτάται από την τροχιά της αλυσίδας (αφού ένας χρόνος διακοπής είναι μια τυχαία μεταβλητή), αλλά ο ορισμός επιβάλλει ότι η απόφαση για το αν θα σταματήσουμε τη στιγμή n ή όχι, μπορεί να εξαρτάται μόνο από τις X_0, \dots, X_n και όχι από τις μετέπειτα καταστάσεις της αλυσίδας.

Παράδειγμα 13 Οι σταθεροί χρόνοι είναι χρόνοι διακοπής, δηλαδή ο $T : \Omega \rightarrow \mathbb{N}_0$, με $T(\omega) = N$, για κάθε $\omega \in \Omega$ είναι χρόνος διακοπής. Πράγματι, για κάθε $n \in \mathbb{N}_0$ έχουμε

$$\{T = n\} = \begin{cases} \Omega, & \text{αν } n = N \\ \emptyset, & \text{αν } n \neq N. \end{cases}$$

Σε κάθε περίπτωση, έχουμε $\Omega \in \mathcal{F}_n$ και $\emptyset = \Omega^c \in \mathcal{F}_n$.

Παράδειγμα 14 Αν $A \subset \mathbb{X}$, τότε ο *χρόνος πρώτης άφιξης* (hitting time) στο A ,

$$T_A = \inf\{k \geq 0 : X_k \in A\}$$

είναι χρόνος διακοπής της $\{X_n\}_{n \in \mathbb{N}_0}$. Πράγματι, $\{T_A = 0\} = \{X_0 \in A\} \in \mathcal{F}_0$, ενώ για κάθε $n \in \mathbb{N}$,

$$\{T_A = n\} = \{X_0 \in A^c\} \cap \dots \cap \{X_{n-1} \in A^c\} \cap \{X_n \in A\}.$$

Είδαμε όμως ότι $\{X_n \in A\} \in \mathcal{F}_n$, ενώ για $k = 0, \dots, n-1$ έχουμε $\{X_k \in A^c\} \in \mathcal{F}_k \subset \mathcal{F}_n$. Εφόσον η \mathcal{F}_n είναι κλειστή ως προς τις τομές, έχουμε ότι $\{T_A = n\} \in \mathcal{F}_n$.

Παράδειγμα 15 Αν $A \subset \mathbb{X}$, τότε η τελευταία φορά που η αλυσίδα επισκέπτεται το A δεν είναι χρόνος διακοπής της $\{X_n\}_{n \in \mathbb{N}_0}$. Πράγματι, αν $S = \sup\{k \geq 0 : X_k \in A\}$, τότε

$$\{S = n\} = \{X_n \in A\} \cap \left(\bigcap_{k=n+1}^{\infty} \{X_k \in A^c\} \right).$$

Για να αποφασίσουμε αν το ενδεχόμενο $\{S = n\}$ συνέβη, αν δηλαδή η χρονική στιγμή n είναι η τελευταία φορά που η αλυσίδα επισκέπτεται το A , χρειάζεται να ξέρουμε τις καταστάσεις της αλυσίδας μετά τη χρονική στιγμή n . \square

Το ακόλουθο λήμμα προσφέρει έναν ισοδύναμο χαρακτηρισμό των χρόνων διακοπής που είναι συχνά χρήσιμος.

Λήμμα 1 Η τυχαία μεταβλητή $T : \Omega \rightarrow \mathbb{N}_0 \cup \{\infty\}$ είναι χρόνος διακοπής, αν και μόνο αν, για κάθε $n \in \mathbb{N}_0$, το ενδεχόμενο $\{T \leq n\}$ ανήκει στην \mathcal{F}_n .

Απόδειξη: Έστω ότι ο T είναι χρόνος διακοπής. Για κάθε $n \in \mathbb{N}_0$ έχουμε

$$\{T \leq n\} = \bigcup_{k=0}^n \{T = k\}.$$

Όμως $\{T = k\} \in \mathcal{F}_k \subset \mathcal{F}_n$, για $k = 0, 1, \dots, n$. Εφόσον η \mathcal{F}_n είναι κλειστή ως προς τις πεπερασμένες ενώσεις, έχουμε ότι $\{T \leq n\} \in \mathcal{F}_n$.

Για το αντίστροφο, ας υποθέσουμε ότι $\{T \leq n\} \in \mathcal{F}_n$, για κάθε $n \in \mathbb{N}_0$. Έχουμε $\{T = 0\} = \{T \leq 0\} \in \mathcal{F}_0$, ενώ για $n \in \mathbb{N}$

$$\{T = n\} = \{T \leq n\} \cap \{T \leq n-1\}^c \in \mathcal{F}_n,$$

αφού όλες οι \mathcal{F}_n είναι κλειστές ως προς συμπληρώματα και τομές. Επομένως ο T είναι χρόνος διακοπής.

Πόρισμα 1 Αν οι T, S είναι χρόνοι διακοπής, τότε οι $T \wedge S$, με $(T \wedge S)(\omega) = \min\{T(\omega), S(\omega)\}$ και $T \vee S$, με $(T \vee S)(\omega) = \max\{T(\omega), S(\omega)\}$ είναι κι αυτοί χρόνοι διακοπής.

Απόδειξη: Ο ισχυρισμός προκύπτει άμεσα από το Λήμμα 1 και την κλειστότητα της \mathcal{F}_n σε ενώσεις και τομές, αφού

$$\{T \wedge S \leq n\} = \{T \leq n\} \cup \{S \leq n\} \quad \text{και} \quad \{T \vee S \leq n\} = \{T \leq n\} \cap \{S \leq n\}. \quad \square$$

Θα εξετάσουμε τώρα την ισχυρή μαρκοβιανή ιδιότητα (strong Markov property) των μαρκοβιανών αλυσίδων, την ιδιότητά τους δηλαδή να ανανεώνονται έπειτα από κάθε χρόνο διακοπής και να εξελίσσονται ανεξάρτητα από το παρελθόν. Τι σημαίνει όμως το παρελθόν του χρόνου διακοπής; Αυτό δεν είναι σαφές, αφού ένας χρόνος διακοπής είναι μια τυχαία μεταβλητή, επομένως μπορεί να έχει διαφορετικές τιμές σε διαφορετικά σημεία του Ω .

Ορισμός: Αν T είναι ένας χρόνος διακοπής, θα λέμε ότι το ενδεχόμενο A ανήκει στο παρελθόν του T και θα συμβολίζουμε $A \in \mathcal{F}_T$, αν για κάθε $n \in \mathbb{N}_0$ το ενδεχόμενο $A \cap \{T = n\}$ ανήκει στην \mathcal{F}_n .

Σύμφωνα με τον παραπάνω ορισμό, ένα ενδεχόμενο A ανήκει στο παρελθόν του χρόνου διακοπής T , αν στις εκβάσεις $\omega \in \Omega$ που ο T παίρνει την τιμή n , μπορούμε να προσδιορίσουμε αν το A συμβαίνει, γνωρίζοντας μόνο τις καταστάσεις X_0, \dots, X_n .

Παράδειγμα 16 Όλα τα ενδεχόμενα της μορφής $\{T \in B\}$, όπου $B \subset \mathbb{N}_0 \cup \{\infty\}$, ανήκουν στο παρελθόν του χρόνου διακοπής T , αφού

$$\{T \in B\} \cap \{T = n\} = \begin{cases} \{T = n\}, & \text{αν } n \in B \\ \emptyset, & \text{αν } n \notin B \end{cases} \in \mathcal{F}_n.$$

Θεώρημα 2 (ισχυρή μαρκοβιανή ιδιότητα) Έστω $\{X_n\}_{n \in \mathbb{N}_0}$ μαρκοβιανή αλυσίδα στον αριθμησιμο χώρο καταστάσεων \mathbb{X} με πίνακα πιθανοτήτων μετάβασης P . Αν ο T είναι ένας χρόνος διακοπής της $\{X_n\}_{n \in \mathbb{N}_0}$, τότε, με δεδομένο ότι $T < \infty$ και $X_T = x \in \mathbb{X}$, η στοχαστική διαδικασία $\{Y_n\}_{n \in \mathbb{N}_0}$ με $Y_n = X_{T+n}$ είναι μια μαρκοβιανή αλυσίδα που ξεκινά από την x , έχει πίνακα πιθανοτήτων μετάβασης P και είναι ανεξάρτητη από οποιοδήποτε ενδεχόμενο $A \in \mathcal{F}_T$.

Απόδειξη: Χρειάζεται να δείξουμε ότι για κάθε $y = (y_0, \dots, y_n) \in \mathbb{X}^{n+1}$ και κάθε $A \in \mathcal{F}_T$,

$$\mathbb{P}[\{(Y_0, \dots, Y_n) = y\} \cap A \mid T < \infty, X_T = x] = \mathbb{P}_x[(X_0, \dots, X_n) = y] \mathbb{P}[A \mid T < \infty, X_T = x].$$

Έχουμε ότι

$$\begin{aligned} \mathbb{P}[\{(Y_0, \dots, Y_n) = y\} \cap A \cap \{T < \infty, X_T = x\}] &= \sum_{k=0}^{\infty} \mathbb{P}[\{(Y_0, \dots, Y_n) = y\} \cap A \cap \{T = k, X_k = x\}] \\ &= \sum_{k=0}^{\infty} \mathbb{P}[\{(X_k, \dots, X_{k+n}) = y\} \cap A \cap \{T = k\} | X_k = x] \mathbb{P}[X_k = x]. \end{aligned}$$

Ορίζουμε $A_k = A \cap \{T = k\}$. Από τον ορισμό του παρελθόντος του T έχουμε ότι $A_k \in \mathcal{F}_k$ για κάθε $k \in \mathbb{N}_0$. Επομένως, το A_k προσδιορίζεται από τις τιμές των X_0, \dots, X_k και από τη μαρκοβιανή ιδιότητα (Θεώρημα 1) παίρνουμε ότι

$$\mathbb{P}[\{(X_k, \dots, X_{k+n}) = y\} \cap A_k | X_k = x] = \mathbb{P}_x[(X_0, \dots, X_n) = y] \mathbb{P}[A_k | X_k = x].$$

Αντικαθιστώντας την παραπάνω ισότητα στην προηγούμενή της, παίρνουμε ότι

$$\begin{aligned} \mathbb{P}[\{(Y_0, \dots, Y_n) = y\} \cap A \cap \{T < \infty, X_T = x\}] &= \sum_{k=0}^{\infty} \mathbb{P}_x[(X_0, \dots, X_n) = y] \mathbb{P}[A_k \cap \{X_k = x\}] \\ &= \mathbb{P}_x[(X_0, \dots, X_n) = y] \sum_{k=0}^{\infty} \mathbb{P}[A \cap \{T = k, X_T = x\}] \\ &= \mathbb{P}_x[(X_0, \dots, X_n) = y] \mathbb{P}[A \cap \{T < \infty, X_T = x\}]. \end{aligned}$$

Το ζητούμενο τώρα προκύπτει από τον ορισμό της δεσμευμένης πιθανότητας. \square

2.5 Επαναληπτικότητα - Παροδικότητα

Ας θεωρήσουμε μια μαρκοβιανή αλυσίδα $\{X_n\}_{n \in \mathbb{N}_0}$ με τιμές σ' έναν αριθμήσιμο χώρο καταστάσεων \mathbb{X} . Για $x \in \mathbb{X}$ ορίζουμε το συνολικό πλήθος των επισκέψεων της αλυσίδας στην κατάσταση x

$$V(x) = \sum_{k=0}^{\infty} \mathbb{1}\{X_k = x\}.$$

Η $V(x)$ είναι τυχαία μεταβλητή με τιμές στο $\mathbb{N}_0 \cup \{\infty\}$. Στις εκβάσεις $\omega \in \Omega$ για τις οποίες $V(x) < \infty$ η αλυσίδα τελικά εγκαταλείπει την κατάσταση x και δεν επιστρέφει ποτέ σ' αυτήν. Αντίθετα, στις εκβάσεις $\omega \in \Omega$ για τις οποίες $V(x) = \infty$ η αλυσίδα δεν σταματά ποτέ να επιστρέφει στην x .

Ορισμός: Θα λέμε μια κατάσταση $x \in \mathbb{X}$ επαναληπτική (recurrent), αν $\mathbb{P}_x[V(x) < \infty] = 0$. Θα λέμε μια κατάσταση $x \in \mathbb{X}$ παροδική (transient), αν $\mathbb{P}_x[V(x) < \infty] = 1$.

Από τον ορισμό που δώσαμε δεν προκύπτει ότι κάθε κατάσταση $x \in \mathbb{X}$ μπορεί να χαρακτηριστεί ως επαναληπτική ή παροδική. Υπάρχει θεωρητικά η δυνατότητα η $\mathbb{P}_x[V(x) < \infty]$ να είναι μικρότερη του 1 αλλά θετική. Θα δούμε όμως ότι κάτι τέτοιο δεν μπορεί να συμβεί. Η εξήγηση γι' αυτό είναι απλή. Κάθε φορά που η αλυσίδα βρίσκεται στην κατάσταση x έχει την ευκαιρία να δραπετεύσει και να μην ξαναγυρίσει. Από την ισχυρή μαρκοβιανή ιδιότητα, έχει κάθε φορά την ίδια πιθανότητα (έστω p) να τα καταφέρει, ανεξάρτητα από τις προηγούμενες προσπάθειες. Επομένως το $\{V(x) < \infty\}$ είναι στην ουσία το ενδεχόμενο να έχουμε επιτυχία σε πεπερασμένο χρόνο, σε μια σειρά από ανεξάρτητες δοκιμές Bernoulli, με πιθανότητα επιτυχίας p . Κάτι τέτοιο όμως μπορεί να συμβεί είτε με πιθανότητα 1, αν $p > 0$, είτε με πιθανότητα 0, αν $p = 0$. Η απόδειξη του Θεωρήματος 3 που ακολουθεί, επαναλαμβάνει με αυστηρότητα το παραπάνω επιχείρημα.

Ορίζουμε τον χρόνο πρώτης επιστροφής στη x ως $T_x^+ = \inf\{k > 0 : X_k = x\}$. Ορίζουμε επίσης αναδρομικά τους χρόνους της n -οστής επιστροφής της αλυσίδας $\{X_n\}_{n \in \mathbb{N}_0}$ στην κατάσταση x ως εξής.

$$T_0 = 0, \quad T_{n+1} = \inf\{k > T_n : X_k = x\}, \quad \text{για } n \in \mathbb{N}_0. \quad (2.8)$$

Λήμμα 2 Έστω $n \in \mathbb{N}$. Δεδομένου ότι $T_n < \infty$, η τυχαία μεταβλητή $T_{n+1} - T_n$ είναι ανεξάρτητη από οποιοδήποτε ενδεχόμενο στο παρελθόν του T_n και έχει την ίδια κατανομή όπως ο χρόνος πρώτης επιστροφής T_x^+ , όταν η αλυσίδα ξεκινά από το x .

Απόδειξη: Από την Άσκηση 28 ο T_n είναι χρόνος διακοπής, για κάθε $n \in \mathbb{N}$. Από τον ορισμό του T_n αν $T_n < \infty$, έχουμε $X_{T_n} = x$. Από την ισχυρή μαρκοβιανή ιδιότητα, δεδομένου ότι $T_n < \infty$, η στοχαστική διαδικασία $\{Y_k\}_{k \in \mathbb{N}_0}$ με $Y_k = X_{T_n+k}$ είναι μια μαρκοβιανή αλυσίδα που ξεκινά από το x , έχει τις ίδιες πιθανότητες μετάβασης όπως η $\{X_n\}_{n \in \mathbb{N}_0}$ και είναι ανεξάρτητη από ενδεχόμενα στο παρελθόν του T_n . Ειδικότερα, είναι ανεξάρτητη από τον T_n , αφού όπως είδαμε στο Παράδειγμα 16, κάθε ενδεχόμενο της μορφής $\{T_n \in B\}$ ανήκει στο παρελθόν του T_n . Παρατηρήστε τώρα ότι, για κάθε $n \in \mathbb{N}$, έχουμε

$$T_{n+1} = T_n + \inf\{k > 0 : X_{T_n+k} = x\} = T_n + \inf\{k > 0 : Y_k = x\}.$$

Επομένως, δεδομένου ότι $T_n < \infty$, η $T_{n+1} - T_n = \inf\{k > 0 : Y_k = x\}$ είναι ανεξάρτητη από οποιοδήποτε ενδεχόμενο στο παρελθόν του T_n και έχει την ίδια κατανομή όπως ο χρόνος της πρώτης επιστροφής στο x όταν η αλυσίδα ξεκινά από το x , δηλαδή

$$\mathbb{P}[T_{n+1} - T_n = k \mid T_n < \infty] = \mathbb{P}_x[T_x^+ = k], \quad \text{για κάθε } k \in \mathbb{N}. \quad \square$$

Αν μια αλυσίδα $\{X_n\}$ ξεκινά από την επαναληπτική κατάσταση $x \in \mathbb{X}$, τότε με πιθανότητα 1 όλοι οι χρόνοι $\{T_n\}_{n \in \mathbb{N}}$ της (2.8) είναι πεπερασμένοι, αφού, από τον ορισμό της επαναληπτικότητας, η αλυσίδα επισκέπτεται άπειρες φορές τη x . Παίρνουμε λοιπόν το ακόλουθο πόρισμα του Λήμματος 2.

Πόρισμα 2 Αν η αλυσίδα $\{X_n\}_{n \in \mathbb{N}}$ ξεκινά από μια επαναληπτική κατάσταση $x \in \mathbb{X}$, τότε για τους χρόνους επιστροφής στο x , $\{T_n\}_{n \in \mathbb{N}}$, έχουμε ότι

$$T_n = \sum_{k=1}^n (T_k - T_{k-1}) = \sum_{k=1}^n Z_k,$$

όπου η $\{Z_k\}_{k \in \mathbb{N}}$ είναι μια ακολουθία από ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με κατανομή αυτή της T_x^+ .

Θεώρημα 3 Έστω $f(x) = \mathbb{P}_x[T_x^+ < \infty]$. Αν $f(x) = 1$, η κατάσταση x είναι επαναληπτική. Αν $f(x) < 1$, η κατάσταση x είναι παροδική. Ειδικότερα, κάθε κατάσταση είναι είτε επαναληπτική είτε παροδική.

Απόδειξη: Θα δείξουμε πρώτα επαγωγικά ότι για κάθε $n \in \mathbb{N}_0$

$$\mathbb{P}_x[V(x) > n] = f(x)^n. \quad (2.9)$$

Για $n = 0$ και τα δύο μέλη είναι ίσα με 1, αφού, αν $X_0 = x$, τότε $V(x) \geq 1$. Έστω τώρα ότι ο ισχυρισμός μας ισχύει για κάποιο $n \in \mathbb{N}_0$. Τότε,

$$\begin{aligned} \mathbb{P}_x[V(x) > n+1] &= \mathbb{P}_x[T_{n+1} < \infty] \\ &= \mathbb{P}_x[\{T_n < \infty\} \cap \{T_{n+1} - T_n < \infty\}] \\ &= \mathbb{P}_x[T_{n+1} - T_n < \infty \mid T_n < \infty] \mathbb{P}_x[T_n < \infty] \\ &= f(x)f(x)^n, \end{aligned}$$

όπου το τελευταίο βήμα προκύπτει από το Λήμμα 2 και την επαγωγική υπόθεση. Επομένως, η (2.9) αληθεύει για κάθε $n \in \mathbb{N}_0$. Παρατηρήστε τώρα ότι η ακολουθία ενδεχομένων $C_n = \{V(x) > n\}$ είναι φθίνουσα. Επομένως,

$$\begin{aligned} \mathbb{P}_x[V(x) = \infty] &= \mathbb{P}_x[\bigcap_n \{V(x) > n\}] = \lim_{n \rightarrow \infty} \mathbb{P}_x[V(x) > n] = \lim_{n \rightarrow \infty} f(x)^n \\ &= \begin{cases} 1, & \text{αν } f(x) = 1 \\ 0, & \text{αν } f(x) < 1. \end{cases} \quad \square \end{aligned}$$

Εφαρμόζοντας τον τύπο της ολικής πιθανότητας ανάλογα με το πρώτο βήμα της αλυσίδας έχουμε ότι

$$f(x) = \sum_{y \in \mathbb{X}} \mathbb{P}[T_x^+ < \infty | X_0 = x, X_1 = y] p(x, y).$$

Από τη μαρκοβιανή ιδιότητα, δεδομένου ότι $X_1 = y$, η μαρκοβιανή διαδικασία $\{Y_n\}_{n \in \mathbb{N}_0}$, με $Y_n = X_{n+1}$, είναι μια μαρκοβιανή αλυσίδα που ξεκινά από το y , με πιθανότητες μετάβασης $\{p(x, y)\}_{x, y \in \mathbb{X}}$. Επιπλέον,

$$T_x^+ = \inf\{k > 0 : X_k = x\} = 1 + \inf\{k \geq 0 : Y_k = x\}.$$

Έχουμε επομένως ότι $T_x^+ < \infty$, ακριβώς όταν ο χρόνος άφιξης της $\{Y_n\}_{n \in \mathbb{N}_0}$ στο x είναι πεπερασμένος. Έτσι,

$$f(x) = \sum_{y \in \mathbb{X}} p(x, y) \mathbb{P}_y[T_x < \infty]. \quad (2.10)$$

Στο επόμενο κεφάλαιο θα μάθουμε τεχνικές για να υπολογίζουμε τις πιθανότητες $\mathbb{P}_y[T_x < \infty]$ και να βρίσκουμε επομένως αν μια αλυσίδα είναι επαναληπτική ή παροδική. Το επόμενο θεώρημα μας δίνει έναν εναλλακτικό τρόπο, χρησιμοποιώντας τις πιθανότητες μετάβασης ανώτερης τάξης.

Θεώρημα 4 Αν η σειρά $\sum_n p^{(n)}(x, x)$ αποκλίνει, τότε η κατάσταση x είναι επαναληπτική. Αν η σειρά αυτή συγκλίνει, τότε η κατάσταση x είναι παροδική.

Απόδειξη: Από την απόδειξη του Θεωρήματος 3, φαίνεται ότι, αν η κατάσταση x είναι παροδική, το πλήθος των επισκέψεων της αλυσίδας στην κατάσταση x ακολουθεί γεωμετρική κατανομή, με παράμετρο $1 - f(x) > 0$. Πράγματι,

$$\mathbb{P}_x[V(x) = n] = \mathbb{P}_x[V(x) > n - 1] - \mathbb{P}_x[V(x) > n] = f(x)^{n-1} - f(x)^n = f(x)^{n-1}(1 - f(x)).$$

Επομένως, ο αναμενόμενος αριθμός επισκέψεων σε μια παροδική κατάσταση x , όταν η αλυσίδα ξεκινά από την x , είναι

$$\mathbb{E}_x[V(x)] = \frac{1}{1 - f(x)} < +\infty.$$

Επιπλέον, αν η κατάσταση x είναι επαναληπτική, τότε $\mathbb{P}_x[V(x) = \infty] = 1$, άρα $\mathbb{E}_x[V(x)] = \infty$. Άρα, μια κατάσταση είναι επαναληπτική ή παροδική, ανάλογα με το αν η $\mathbb{E}_x[V(x)]$ είναι πεπερασμένη ή άπειρη. Όμως, αν υπολογίσουμε την $\mathbb{E}_x[V(x)]$ από τον ορισμό της V , έχουμε

$$\mathbb{E}_x[V(x)] = \mathbb{E}_x\left[\sum_{n=0}^{\infty} \mathbb{1}\{X_n = x\}\right] = \sum_{n=0}^{\infty} \mathbb{E}_x[\mathbb{1}\{X_n = x\}] = \sum_{n=0}^{\infty} \mathbb{P}_x[X_n = x] = \sum_{n=0}^{\infty} p^{(n)}(x, x),$$

απ' όπου έπεται ο ισχυρισμός του θεωρήματος. □

Λήμμα 3 Η επαναληπτικότητα και η παροδικότητα είναι ιδιότητες κλάσης. Αν δηλαδή η κατάσταση x είναι επαναληπτική/παροδική, τότε κάθε κατάσταση y που επικοινωνεί με την x είναι ομοίως επαναληπτική/παροδική.

Απόδειξη: Ας θεωρήσουμε μια παροδική κατάσταση x και μια κατάσταση $y \in \mathbb{X}$ τέτοια ώστε $x \leftrightarrow y$. Εφόσον $x \rightarrow y$, υπάρχει $j \geq 0$, τέτοιο ώστε $p^{(j)}(x, y) > 0$. Ομοίως, εφόσον $y \rightarrow x$, υπάρχει $k \geq 0$, τέτοιο ώστε $p^{(k)}(y, x) > 0$. Παρατηρήστε τώρα ότι για κάθε $n \geq 0$, οι εξισώσεις Chapman-Kolmogorov δίνουν

$$p^{(j+n+k)}(x, x) = \sum_{u, v \in \mathbb{X}} p^{(j)}(x, u) p^{(n)}(u, v) p^{(k)}(v, x) \geq p^{(j)}(x, y) p^{(n)}(y, y) p^{(k)}(y, x).$$

Αθροίζοντας αυτές τις ανισότητες ως προς n έχουμε ότι

$$\sum_{n=0}^{\infty} p^{(n)}(y, y) \leq \frac{1}{p^{(j)}(x, y) p^{(k)}(y, x)} \sum_{n=j+k}^{\infty} p^{(n)}(x, x).$$

Από το Θεώρημα 4 η σειρά στο δεξί μέλος συγκλίνει, επομένως συγκλίνει και η σειρά στο αριστερό μέλος, άρα και η y είναι παροδική.

Αν μια κατάσταση x είναι επαναληπτική και $x \leftrightarrow y$, τότε η y θα είναι κι αυτή επαναληπτική, αφού, αν ήταν παροδική, τότε η x θα ήταν παροδική, όπως δείξαμε παραπάνω. \square

Με βάση το προηγούμενο λήμμα έχει νόημα να μιλάμε για επαναληπτικές κλάσεις (όταν οι καταστάσεις της κλάσης είναι επαναληπτικές) και για παροδικές κλάσεις (όταν οι καταστάσεις της κλάσης είναι παροδικές). Μπορούμε επίσης να χαρακτηρίσουμε μια μη υποβιβάσιμη μαρκοβιανή αλυσίδα ως επαναληπτική ή παροδική, ανάλογα με το τι είναι η μοναδική της κλάση.

Θεώρημα 5 Κάθε ανοιχτή κλάση είναι παροδική.

Απόδειξη: Αν μια κλάση \mathcal{C} είναι ανοιχτή, υπάρχουν $x \in \mathcal{C}$ και $y \notin \mathcal{C}$, τέτοια ώστε $p(x, y) > 0$. Θα πρέπει τότε να έχουμε $p^{(n)}(y, x) = 0$, για κάθε $n \in \mathbb{N}$, αφού διαφορετικά οι x, y θα ανήκαν στην ίδια κλάση. Έτσι,

$$\{X_0 = x, X_1 = y\} \subset \{X_0 = x, V(x) < +\infty\}$$

και

$$\mathbb{P}_x[V(x) < \infty] \geq \mathbb{P}_x[X_1 = y] = p(x, y) > 0.$$

Επομένως η x δεν είναι επαναληπτική και από το Θεώρημα 3 είναι παροδική. Λόγω του Λήμματος 3, η \mathcal{C} είναι κι αυτή παροδική. \square

Λήμμα 4 Αν μια κλάση \mathcal{C} είναι παροδική, τότε για κάθε $x, y \in \mathcal{C}$ έχουμε $\mathbb{P}_x[V(y) < +\infty] = 1$. Αν μια κλάση \mathcal{C} είναι επαναληπτική, τότε για κάθε $x, y \in \mathcal{C}$ έχουμε $\mathbb{P}_x[V(y) < +\infty] = 0$.

Απόδειξη: Εφόσον $y \rightarrow x$, υπάρχει $n \geq 0$ τέτοιο ώστε $p^{(n)}(y, x) > 0$. Από την παροδικότητα της κατάστασης y , τον τύπο ολικής πιθανότητας ανάλογα με την κατάσταση της αλυσίδας τη στιγμή n και τη μαρκοβιανή ιδιότητα παίρνουμε

$$0 = \mathbb{P}_y[V(y) = \infty] = \sum_{z \in \mathbb{X}} \mathbb{P}_y[V(y) = \infty | X_n = z] p^{(n)}(y, z) \geq \mathbb{P}_x[V(y) = \infty] p^{(n)}(y, x).$$

Έχουμε όμως $p^{(n)}(y, x) > 0$, άρα θα πρέπει $\mathbb{P}_x[V(y) = \infty] = 0 \Leftrightarrow \mathbb{P}_x[V(y) < +\infty] = 1$. Ομοίως, αν η \mathcal{C} είναι επαναληπτική, έχουμε

$$0 = \mathbb{P}_y[V(y) < \infty] = \sum_{z \in \mathbb{X}} \mathbb{P}_y[V(y) < \infty | X_n = z] p^{(n)}(y, z) \geq \mathbb{P}_x[V(y) < \infty] p^{(n)}(y, x).$$

Επομένως, αν επιλέξουμε το n ώστε $p^{(n)}(y, x) > 0$, παίρνουμε ότι $\mathbb{P}_x[V(y) < +\infty] = 0$. \square

Βλέπουμε από το παραπάνω λήμμα ότι μια μη υποβιβάσιμη, επαναληπτική αλυσίδα εξερευνά ολόκληρο τον χώρο καταστάσεων και μάλιστα με πιθανότητα 1 επισκέπτεται κάθε κατάσταση άπειρες φορές.

Θεώρημα 6 Κάθε πεπερασμένη κλειστή κλάση είναι επαναληπτική.

Απόδειξη: Έστω C πεπερασμένη και κλειστή κλάση. Κάθε αλυσίδα που ξεκινάει από κάποιο $x \in C$ παραμένει στη C για πάντα. Επομένως, θα υποθέσουμε χωρίς βλάβη ότι η αλυσίδα είναι μη υποβιβάζσιμη. Από το θεώρημα Fubini-Tonelli έχουμε ότι

$$\sum_{y \in C} V(y) = \sum_{y \in C} \sum_{n=0}^{\infty} \mathbb{1}\{X_n = y\} = \sum_{n=0}^{\infty} \sum_{y \in C} \mathbb{1}\{X_n = y\} = \sum_{n=0}^{\infty} \mathbb{1}_C(X_n) = \sum_{n=0}^{\infty} 1 = \infty.$$

Εφόσον η C είναι πεπερασμένη, θα υπάρχει κάποιο $y \in C$ τέτοιο ώστε $\mathbb{P}_x[V(y) = \infty] > 0$. Από το προηγούμενο λήμμα όμως αυτό σημαίνει ότι η C δεν είναι παροδική. \square

Τα Θεωρήματα 5 και 6 απλοποιούν πολύ τον χαρακτηρισμό των κλάσεων σε πεπερασμένους χώρους καταστάσεων X . Είναι εν γένει εύκολο να βρούμε αν μια κλάση είναι ανοιχτή ή κλειστή και, αν ο X είναι πεπερασμένος, οι ανοιχτές κλάσεις είναι παροδικές και οι κλειστές κλάσεις είναι επαναληπτικές. Όταν ο X είναι άπειρος, οι ανοιχτές κλάσεις είναι παροδικές, μια κλειστή κλάση όμως μπορεί να είναι επαναληπτική ή παροδική, όπως θα δούμε στα παραδείγματα.

Παράδειγμα 17 Στο Παράδειγμα 11 οι κλάσεις C_1 και C_2 είναι παροδικές, αφού είναι ανοιχτές, ενώ η κλάση C_3 είναι επαναληπτική, αφού είναι κλειστή και πεπερασμένη. Επομένως, όποια κι αν είναι η αρχική της κατάσταση, θα ξοδέψει πεπερασμένο χρόνο στις ανοιχτές κλάσεις της και τελικά θα καταλήξει στην C_3 , όπου θα παραμείνει για πάντα.

Παράδειγμα 18 Θεωρούμε έναν απλό τυχαίο περίπατο στους ακεραίους που ξεκινά από το μηδέν, δηλαδή μια μαρκοβιανή αλυσίδα στο \mathbb{Z} με $X_0 = 0$ και πιθανότητες μετάβασης

$$p(x, x+1) = q \in (0, 1), \quad p(x, x-1) = 1 - q, \quad \text{για κάθε } x \in \mathbb{Z}.$$

Είναι φανερό ότι η αλυσίδα αυτή είναι μη υποβιβάζσιμη, επομένως όλος ο χώρος καταστάσεων \mathbb{Z} είναι μια κλειστή κλάση. Μπορούμε να υπολογίσουμε τις $p^{(n)}(0, 0)$, για κάθε $n \in \mathbb{N}$. Η αλυσίδα δεν μπορεί να επιστρέψει στο 0 έπειτα από ένα περιττό αριθμό βημάτων, αφού $X_{2n} = 0 \pmod{2}$ και $X_{2n+1} = 1 \pmod{2}$. Για να επιστρέψει η αλυσίδα στο μηδέν έπειτα από $2n$ βήματα, πρέπει να κάνει ακριβώς n βήματα αριστερά και n βήματα δεξιά. Υπάρχουν $\binom{2n}{n}$ τέτοια μονοπάτια, όσοι και οι τρόποι που μπορούμε να επιλέξουμε τα n βήματα που η αλυσίδα πάει αριστερά ανάμεσα στα $2n$ βήματα, καθένα από τα οποία έχει πιθανότητα $q^n(1-q)^n$. Επομένως,

$$\sum_{n=0}^{\infty} p^{(n)}(0, 0) = \sum_{n=0}^{\infty} p^{(2n)}(0, 0) = \sum_{n=0}^{\infty} \binom{2n}{n} q^n (1-q)^n. \quad (2.11)$$

Από το κριτήριο του λόγου μπορούμε να βρούμε ότι η παραπάνω σειρά συγκλίνει για $q \neq \frac{1}{2}$, επομένως η αλυσίδα του παραδείγματος είναι παροδική σε αυτή την περίπτωση. Αυτό το αποτέλεσμα είναι μάλλον αναμενόμενο. Μπορούμε να γράψουμε την αλυσίδα ως ένα άθροισμα ανεξάρτητων, ισόνομων τυχαίων μεταβλητών $\{Z_i\}_{i \in \mathbb{N}}$

$$X_n = \sum_{i=1}^n Z_i,$$

όπου οι Z_i παίρνουν την τιμή $+1$ με πιθανότητα q και την τιμή -1 με πιθανότητα $1-q$. Εύκολα υπολογίζουμε τη μέση τιμή τους $\mathbb{E}[Z_i] = 2q - 1 \neq 0$. Ο νόμος των μεγάλων αριθμών δίνει τώρα ότι

$$\mathbb{P}\left[\frac{X_n}{n} \rightarrow 2q - 1\right] = 1.$$

Επομένως, αν $q > 1/2$ (αντίστοιχα $q < 1/2$), με πιθανότητα 1 η αλυσίδα θα είναι τελικά μεγαλύτερη του $(q - 1/2)n$ (αντίστοιχα μικρότερη του $(q - 1/2)n$), άρα θα τείνει στο $+\infty$ (αντίστοιχα $-\infty$). Έχουμε δηλαδή

$$\mathbb{P}[X_n \rightarrow +\infty] = 1, \quad \text{αν } q > 1/2 \quad \text{και} \quad \mathbb{P}[X_n \rightarrow -\infty] = 1, \quad \text{αν } q < 1/2.$$

Αν η αλυσίδα έχει μια τάση να κινείται προς τ' αριστερά ή τα δεξιά, τότε, προϊόντος του χρόνου, θα ξεφύγει προς αυτή την κατεύθυνση.

Για να εξετάσουμε την περίπτωση $q = 1/2$, χρειαζόμαστε μια καλή εκτίμηση για την ασυμπτωτική συμπεριφορά του $2^{-2n} \binom{2n}{n}$. Θα μπορούσαμε να πάρουμε μια τέτοια εκτίμηση από τον τύπο του Stirling

$$\lim_{n \rightarrow \infty} \frac{n!}{\left(\frac{n}{e}\right)^n \sqrt{2\pi n}} \rightarrow 1. \quad (2.12)$$

Είναι όμως διασκεδαστικό να αποδείξουμε το ακόλουθο λήμμα, σύμφωνα με το οποίο η σειρά της (2.11) αποκλίνει όταν $q = 1/2$, επομένως ο απλός συμμετρικός ($q = 1/2$) τυχαίος περίπατος στο \mathbb{Z} είναι επαναληπτικός. Σύμφωνα με το Λήμμα 4, αν σταθούμε σε οποιοδήποτε ακέραιο, με πιθανότητα 1, η αλυσίδα δεν θα σταματήσει ποτέ να μας επισκέπτεται. \square

Λήμμα 5 Για κάθε $n \in \mathbb{N}$ έχουμε

$$\frac{1}{\sqrt{\pi(n + \frac{1}{2})}} < \frac{1}{2^{2n}} \binom{2n}{n} < \frac{1}{\sqrt{\pi n}}.$$

Απόδειξη: Θα χρησιμοποιήσουμε κάποιες ιδιότητες της συνάρτησης Γάμμα

$$\Gamma(p) = \int_0^\infty x^{p-1} e^{-x} dx, \quad p > 0.$$

Εύκολα μπορούμε να δούμε ότι $\Gamma(1) = 1$, ενώ με μια ολοκλήρωση κατά μέλη βλέπουμε ότι $\Gamma(p+1) = p\Gamma(p)$, για κάθε $p > 0$. Από τις δύο αυτές ιδιότητες προκύπτει με επαγωγή ότι $\Gamma(n) = (n-1)!$ για κάθε $n \in \mathbb{N}$. Με την αλλαγή μεταβλητής $z = \sqrt{2x}$ βρίσκουμε επίσης ότι $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Τέλος, από την ανισότητα Cauchy-Schwarz, έχουμε

$$\Gamma(p + \frac{1}{2}) = \int_0^\infty \sqrt{x} x^{p-1} e^{-x} dx \leq \left(\int_0^\infty x^p e^{-x} dx \int_0^\infty x^{p-1} e^{-x} dx \right)^{\frac{1}{2}} = \sqrt{\Gamma(p+1)\Gamma(p)} = \sqrt{p}\Gamma(p). \quad (2.13)$$

Έχουμε τώρα όλα τα υλικά που χρειαζόμαστε από τη συνάρτηση Γάμμα και προχωράμε στην εκτίμηση που θέλουμε να κάνουμε.

$$\begin{aligned} \binom{2n}{n} \frac{1}{2^{2n}} &= \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2^n} \frac{2 \cdot 4 \cdot 6 \cdots (2n)}{2^n (n!)^2} = \left(\frac{1}{2}\right) \left(\frac{3}{2}\right) \left(\frac{5}{2}\right) \cdots \left(\frac{2n-1}{2}\right) \frac{2^n n!}{2^n (n!)^2} \\ &= \Gamma\left(\frac{1}{2}\right) \left(\frac{1}{2}\right) \left(\frac{3}{2}\right) \left(\frac{5}{2}\right) \cdots \left(n - \frac{1}{2}\right) \frac{1}{\Gamma\left(\frac{1}{2}\right)n!} = \frac{\Gamma\left(n + \frac{1}{2}\right)}{\sqrt{\pi n!}}. \end{aligned}$$

Από την (2.13) έχουμε τώρα ότι

$$\binom{2n}{n} \frac{1}{2^{2n}} = \frac{\Gamma\left(n + \frac{1}{2}\right)}{\sqrt{\pi n!}} \leq \frac{\sqrt{n}\Gamma(n)}{\sqrt{\pi n}\Gamma(n)} = \frac{1}{\sqrt{\pi n}},$$

και

$$\binom{2n}{n} \frac{1}{2^{2n}} = \frac{\Gamma\left(n + \frac{1}{2}\right)}{\sqrt{\pi}\Gamma\left(n + \frac{1}{2}\right)} \geq \frac{\Gamma\left(n + \frac{1}{2}\right)}{\sqrt{\pi\left(n + \frac{1}{2}\right)\Gamma\left(n + \frac{1}{2}\right)}} = \frac{1}{\sqrt{\pi\left(n + \frac{1}{2}\right)}}. \quad \square$$

2.6 Ασκήσεις

Άσκηση 19 Η $\{X_n\}_{n \in \mathbb{N}_0}$ είναι μια μαρκοβιανή αλυσίδα στον $\mathbb{X} = \{1, 2, 3\}$ με πίνακα πιθανοτήτων μετάβασης

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 2/3 & 1/3 \\ p & 1-p & 0 \end{pmatrix}.$$

Υπολογίστε την $\mathbb{P}[X_n = 1 | X_0 = 1]$ στις περιπτώσεις: α) $p = 1/16$, β) $p = 1/6$, γ) $p = 1/12$.

Άσκηση 20 Η $\{X_n\}_{n \in \mathbb{N}_0}$ είναι μια μαρκοβιανή αλυσίδα στον $\mathbb{X} = \{1, 2, 3\}$ με πίνακα πιθανοτήτων μετάβασης

$$P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/3 & 1/3 & 1/3 \\ p & 2/3-p & 1/3 \end{pmatrix}.$$

Υπολογίστε την $\mathbb{P}[X_n = 1 | X_0 = 1]$, για κάθε $n \in \mathbb{N}$, στις περιπτώσεις: α) $p = 0$, β) $p = 1/6$, γ) $p = 2/3$. Πώς θα μπορούσατε να υπολογίσετε το $\lim_{n \rightarrow \infty} P^n$, χωρίς πολλές πράξεις;

Άσκηση 21 Η $\{X_n\}_{n \in \mathbb{N}_0}$ είναι μια μαρκοβιανή αλυσίδα στον χώρο καταστάσεων $\mathbb{X} = \{1, 2, 3, 4\}$ με πίνακα πιθανοτήτων μετάβασης

$$P = \begin{pmatrix} 0 & 1/6 & 1/12 & 3/4 \\ 1/2 & 0 & 1/4 & 1/4 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 1/3 & 1/6 & 0 \end{pmatrix}.$$

Αν π_n είναι η κατανομή της αλυσίδας μετά από n βήματα, δείξτε ότι, ανεξάρτητα από την αρχική της κατανομή π_0 έχουμε $\pi_n \rightarrow \pi_*$ για κάποια κατανομή π_* που θα προσδιορίσετε. Δείξτε επιπλέον ότι

$$\|\pi_n - \pi_*\| = \sum_{x \in \mathbb{X}} |\pi_n(x) - \pi_*(x)| \leq C(n+1)2^{-n}$$

για κάποια σταθερά $C > 0$.

Άσκηση 22 Δείξτε ότι το ενδεχόμενο μια μαρκοβιανή αλυσίδα να ξαναγυρίσει σε μια ανοιχτή κλάση από την οποία έχει φύγει έχει πιθανότητα 0. Υπόδειξη: Το εν λόγω ενδεχόμενο είναι η αριθμήσιμη ένωση

$$E = \bigcup_{i \in \mathbb{N}_0} \bigcup_{j \in \mathbb{N}} \bigcup_{k \in \mathbb{N}} \bigcup_{x \in \mathcal{C}} \bigcup_{z \notin \mathcal{C}} \bigcup_{y \in \mathcal{C}} \{X_i = x, X_{i+j} = z, X_{i+j+k} = y\}.$$

Αρκεί λοιπόν να δείξουμε ότι $\mathbb{P}[X_i = x, X_{i+j} = z, X_{i+j+k} = y] = 0$, για κάθε $i \in \mathbb{N}_0, j \in \mathbb{N}, k \in \mathbb{N}$ και $x, y \in \mathcal{C}, z \notin \mathcal{C}$.

Άσκηση 23 Βρείτε τις κλάσεις επικοινωνίας της μαρκοβιανής αλυσίδας με πίνακα πιθανοτήτων μετάβασης

$$P = \begin{pmatrix} 1/2 & 0 & 0 & 0 & 1/2 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1/4 & 1/4 & 1/4 & 1/4 \\ 1/2 & 0 & 0 & 0 & 1/2 \end{pmatrix}.$$

Ταξινομήστε τις κλάσεις σε ανοιχτές και κλειστές. Ποιες κλάσεις είναι παροδικές και ποιες επαναληπτικές;

Άσκηση 24 Θεωρήστε μια μαρκοβιανή αλυσίδα $\{X_n\}_n$ στο σύνολο καταστάσεων $\mathbb{X} = \{1, 2, \dots, 8\}$ με πίνακα πιθανοτήτων μετάβασης

$$P = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/4 & 3/4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/4 & 1/4 & 0 & 1/8 & 3/8 & 0 & 0 & 0 \\ 0 & 0 & 1/4 & 0 & 3/4 & 0 & 0 & 0 \\ 0 & 0 & 1/5 & 1/5 & 1/5 & 1/5 & 1/5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 1/2 & 0 \end{pmatrix}.$$

Ταξινομήστε τις καταστάσεις σε κλάσεις επικοινωνίας και χαρακτηρίστε τις ως προς την επαναληπτικότητα. Αν $X_0 = 1$, υπολογίστε την $\mathbb{P}[X_n = k]$ για κάθε $n \in \mathbb{N}$ και κάθε $k \in \mathbb{X}$. Υπόδειξη για το τελευταίο ερώτημα: μην επιχειρήσετε να διαγωνιοποιήσετε τον 8×8 πίνακα P . Χρησιμοποιήστε το αποτέλεσμα του προηγούμενου ερωτήματος.

Άσκηση 25 Δείξτε ότι η πρώτη φορά που μια μαρκοβιανή αλυσίδα $\{X_n\}_{n \in \mathbb{N}_0}$ επισκέπτεται την κατάσταση $x \in \mathbb{X}$ μετά τη χρονική στιγμή 10, $T = \inf\{k > 10 : X_k = x\}$, είναι χρόνος διακοπής.

Άσκηση 26 Δείξτε ότι η δεύτερη φορά που μια μαρκοβιανή αλυσίδα επισκέπτεται ένα σύνολο καταστάσεων A είναι χρόνος διακοπής.

Άσκηση 27 Δείξτε ότι ο $T = \inf\{k \geq 5 : X_k = X_2\}$ είναι χρόνος διακοπής για την αλυσίδα $\{X_n\}_{n \in \mathbb{N}_0}$.

Άσκηση 28 Αν ο T είναι χρόνος διακοπής και $A \subset \mathbb{X}$, δείξτε ότι ο χρόνος

$$S = \inf\{k > T : X_k \in A\}$$

είναι κι αυτός χρόνος διακοπής.

Άσκηση 29 Δώστε ένα παράδειγμα μιας αλυσίδας που έχει μόνο ανοιχτές κλάσεις.

Άσκηση 30 Για τη μαρκοβιανή αλυσίδα της Άσκησης 23 υπολογίστε την $\mathbb{P}[T_4^+ < +\infty \mid X_0 = 4]$ και τον αναμενόμενο αριθμό επισκέψεων στην κατάσταση 4, $\mathbb{E}[V(4) \mid X_0 = 4]$. Κάντε το ίδιο για την κατάσταση 2.

Άσκηση 31 Αν η $y \in \mathbb{X}$ είναι παροδική, αποδείξτε ότι για κάθε $x \in \mathbb{X}$ έχουμε

$$\mathbb{E}_x[V(y)] = \frac{\mathbb{P}_x[T_y < +\infty]}{\mathbb{P}_y[T_y^+ = +\infty]}.$$

Άσκηση 32 Θεωρήστε τη μαρκοβιανή αλυσίδα του Παραδείγματος 18. Χρησιμοποιώντας την ισχυρή μαρκοβιανή ιδιότητα κατά τον χρόνο άφιξης στο $x - 1$, $T_{x-1} = \inf\{k \geq 0 : X_k = x - 1\}$, δείξτε ότι, για κάθε $x \in \mathbb{N}$ έχουμε

$$\mathbb{P}_x[T_0 < \infty] = (\mathbb{P}_1[T_0 < \infty])^x$$

Άσκηση 33 Θεωρήστε έναν απλό συμμετρικό τυχαίο περίπατο στο \mathbb{Z} και ορίστε τον χρόνο πρώτης άφιξης στο 0, $T_0 = \inf\{k \geq 0 : X_k = 0\}$. Χρησιμοποιώντας την ισχυρή μαρκοβιανή ιδιότητα κατά τον χρόνο πρώτης άφιξης στο 1, $T_1 = \inf\{k \geq 0 : X_k = 1\}$, δείξτε ότι, για έναν περίπατο που ξεκινά από το 2, έχουμε $T_0 = S_1 + S_2$, όπου οι S_1, S_2 είναι ανεξάρτητες, ισόνομες τυχαίες μεταβλητές, που έχουν την

ίδια κατανομή όπως ο T_0 για έναν περίπατο που ξεκινά από το 1. Χρησιμοποιήστε στη συνέχεια αυτό το αποτέλεσμα, για να δείξετε ότι, αν ορίσουμε $\psi(s) = \mathbb{E}[s^{T_0} | X_0 = 1]$, $s \in (0, 1)$, τότε η $\psi(s)$ λύνει την εξίσωση

$$\psi(s) = \frac{s}{2} + \frac{s}{2}\psi(s)^2.$$

Λύστε την εξίσωση αυτή για να βρείτε την $\psi(s)$ και συμπεράνετε ότι $\mathbb{E}[T_0 | X_0 = 1] = \infty$. Το συμπέρασμα είναι ότι, παρότι ο απλός συμμετρικός τυχαίος περίπατος στο \mathbb{Z} είναι επαναληπτικός, ο αναμενόμενος χρόνος για να φτάσει από το 1 στο 0 είναι άπειρος.

Άσκηση 34 Ένας απλός, συμμετρικός τυχαίος περίπατος στο \mathbb{Z}^2 είναι μια μαρκοβιανή αλυσίδα, που σε κάθε της βήμα μετατοπίζεται είτε βόρεια, είτε νότια, είτε ανατολικά, είτε δυτικά, με πιθανότητα $1/4$, ανεξάρτητα για κάθε βήμα. Επομένως, αν ορίσουμε $e_1 = (1, 0)$ και $e_2 = (0, 1)$, έχουμε

$$S_n = S_0 + \sum_{i=1}^n W_i,$$

όπου οι $\{W_i\}_{i \in \mathbb{N}}$ είναι ανεξάρτητες, ισόνομες τυχαίες μεταβλητές, που παίρνουν τις τιμές $\pm e_1, \pm e_2$ με πιθανότητα $1/4$. Οι πιθανότητες μετάβασης αυτής της αλυσίδας είναι

$$p(x, x + e_1) = p(x, x - e_1) = p(x, x + e_2) = p(x, x - e_2) = \frac{1}{4}, \quad p(x, y) = 0, \text{ διαφορετικά.}$$

Αν X_n, Y_n είναι οι συντεταγμένες του περιπάτου τη χρονική στιγμή $n \in \mathbb{N}_0$, αν δηλαδή $S_n = (X_n, Y_n)$ και ορίσουμε

$$U_n = X_n + Y_n, \quad V_n = X_n - Y_n,$$

δείξτε ότι οι $\{U_n\}_{n \in \mathbb{N}_0}$ και $\{V_n\}_{n \in \mathbb{N}_0}$ είναι ανεξάρτητοι, απλοί, συμμετρικοί τυχαίοι περίπατοι στο \mathbb{Z} . Δείξτε ότι, αν $S_0 = (0, 0)$, τότε

$$\{S_n = (0, 0)\} = \{U_n = 0\} \cap \{V_n = 0\}$$

και χρησιμοποιήστε το Θεώρημα 4 και το Λήμμα 2.5, για να δείξετε ότι ο $\{S_n\}_{n \in \mathbb{N}_0}$ είναι επαναληπτικός.

2.7 Αριθμητικά πειράματα

Η μέθοδος *Monte Carlo* είναι μια υπολογιστική μέθοδος, που βασίζεται στον νόμο των μεγάλων αριθμών. Υπενθυμίζουμε ότι, αν $\{X_n\}_{n \in \mathbb{N}}$ είναι μια ακολουθία από ανεξάρτητες, ισόνομες τυχαίες μεταβλητές, με πεπερασμένη μέση τιμή $\mathbb{E}[X]$, τότε

$$\mathbb{P}\left[\frac{1}{n} \sum_{k=1}^n X_k \rightarrow \mathbb{E}[X]\right] = 1.$$

Προκειμένου να υπολογίσουμε τη μέση τιμή $\mathbb{E}[X]$ μιας τυχαίας μεταβλητής X , μπορούμε λοιπόν να πάρουμε τον μέσο όρο ενός μεγάλου αριθμού ανεξάρτητων δειγμάτων αυτής της μεταβλητής. Με παρόμοιο τρόπο, μπορούμε να προσεγγίσουμε υπολογιστικά την πιθανότητα ενός ενδεχομένου από το κλάσμα των πραγματοποιησέων του σε μια σειρά από ανεξάρτητες προσομοιώσεις. Σ' αυτή την ιδέα θα βασιστούν τα επόμενα αριθμητικά πειράματα.

Άσκηση 35 Κατεβάστε τον κώδικα `ex1.py`. Το πρόγραμμα αυτό εκτιμά με τη μέθοδο Monte Carlo την πιθανότητα της Άσκησης 19 για $p = 1/6$. Τρέξτε το. Προσεγγίζει το αριθμητικό αποτέλεσμα εκείνο που βρήκατε θεωρητικά; Τρέξτε το πρόγραμμα μερικές ακόμα φορές. Είναι η εκτίμηση που δίνει η ίδια κάθε φορά; Ποια είναι η δειγματική διασπορά;

Άσκηση 36 Διαβάστε τώρα τον κώδικα και προσπαθήστε να καταλάβετε πώς λειτουργεί. Αλλάξτε το πλήθος των επαναλήψεων N που κάνουμε από 1.000 σε 100.000 και επαναλάβετε την Άσκηση 35. Ποια είναι τώρα η δειγματική διασπορά; Κάντε το ίδιο και για τις άλλες τιμές του p της Άσκησης 19.

Άσκηση 37 Βρείτε στον κώδικα το σημείο όπου ορίζεται η αρχική κατανομή της αλυσίδας και αλλάξτε τον κώδικα ώστε η αλυσίδα να ξεκινάει από την κατάσταση 3. Είναι το αποτέλεσμα που βρήκατε διαφορετικό από αυτό της προηγούμενης άσκησης;

Άσκηση 38 Τροποίηστε τον κώδικα ώστε να υπολογίζει αριθμητικά την πιθανότητα της Άσκησης 21. Είναι το αποτέλεσμα που βρήκατε αριθμητικά σε συμφωνία με αυτό που βρήκατε θεωρητικά;

Άσκηση 39 Υπολογίστε προσεγγιστικά την πιθανότητα της Άσκησης 42, προσομοιώνοντας $N = 10.000$ μονοπάτια της αλυσίδας και μετρώντας πόσα από αυτά κατέληξαν σε καθεμιά από τις κλειστές κλάσεις.

Άσκηση 40 Υπολογίστε προσεγγιστικά με Monte Carlo την πιθανότητα νίκης του παίκτη που σερβίρει σε ένα γκέιμ τένις, αν η πιθανότητα που έχει να κερδίσει κάθε πόντο είναι $p = 0,6$.

Άσκηση 41 Η ρουτίνα `markov_chain_lib.py` είναι μια εξελιγμένη εκδοχή της βιβλιοθήκης `simple_markov_chain_lib.py` που χρησιμοποιήσαμε στα προηγούμενα πειράματα. Περιλαμβάνει υλοποιημένη μια μέθοδο για την εύρεση των κλάσεων επικοινωνίας και τον χαρακτηρισμό τους ως ανοιχτών ή κλειστών. Χρησιμοποιήστε αυτήν τη ρουτίνα ως βιβλιοθήκη, αντικαθιστώντας την εντολή

```
import simple_markov_chain_lib as lib
```

από την εντολή

```
import markov_chain_lib as lib
```

στον κώδικα `test.py`. Επινοήστε τώρα μια μαρκοβιανή αλυσίδα με μία ανοιχτή και δύο κλειστές κλάσεις και προσθέστε την εντολή

```
print m.communication_classes()
```

Τρέξτε το πρόγραμμα και επιβεβαιώστε ότι η διαμέριση του χώρου καταστάσεων σε κλάσεις γίνεται σωστά.

Κεφάλαιο 3

Θεωρία Δυναμικού

3.1 Εισαγωγή

Είδαμε στο προηγούμενο κεφάλαιο πώς μπορούμε να διαμερίσουμε τον χώρο καταστάσεων μιας μαρκοβιανής αλυσίδας σε κλάσεις επικοινωνίας και πώς μπορούμε να ταξινομήσουμε τις κλάσεις αυτές σε ανοιχτές και κλειστές. Είδαμε επίσης ότι, αν η αλυσίδα ξεκινήσει ή βρεθεί κάποια στιγμή σε μία από τις κλειστές κλάσεις, τότε θα παραμείνει σε αυτήν για πάντα. Αν η αλυσίδα ξεκινά από μια ανοιχτή κλάση και υπάρχουν περισσότερες από μία κλειστές κλάσεις, ποια είναι η πιθανότητα να απορροφηθεί από καθεμία από τις κλειστές κλάσεις; Και πόσο χρόνο θα πάρει μέχρι να συμβεί αυτό; Αυτά είναι ερωτήματα που θα μπορέσουμε να απαντήσουμε με τις τεχνικές που θα μάθουμε στο παρόν κεφάλαιο. Παρόμοιο υλικό μπορείτε να βρείτε και στις αναφορές [2], [7] και στην ελληνική βιβλιογραφία στην [10].

3.2 Πιθανότητες απορρόφησης

Θεωρούμε μια μαρκοβιανή αλυσίδα $\{X_n\}_{n \in \mathbb{N}_0}$ με τιμές σ' έναν χώρο καταστάσεων \mathbb{X} και πιθανότητες μετάβασης $p(\cdot, \cdot)$. Θεωρούμε ακόμα ένα σύνολο $A \subset \mathbb{X}$ και ορίζουμε τον χρόνο πρώτης άφιξης στο A την τυχαία μεταβλητή $T_A : \Omega \rightarrow \mathbb{N}_0 \cup \{+\infty\}$ με

$$T_A(\omega) = \inf\{k \geq 0 : X_k(\omega) \in A\}.$$

Όταν το σύνολο A είναι ένα μονοσύνολο $\{x\}$, θα συμβολίζουμε εναλλακτικά τον χρόνο άφιξης στο A και ως T_x .

Παράδειγμα 19 Σ' ένα γκέιμ τένις δύο παίκτες A, B είναι ισόπαλοι. Ο αγώνας θα τελειώσει, όταν κάποιος από τους δύο παίκτες κερδίσει δύο πόντους παραπάνω από τον αντίπαλό του. Μπορούμε να μοντελοποιήσουμε το παραπάνω παιχνίδι ως μια μαρκοβιανή αλυσίδα στον χώρο καταστάσεων $\mathbb{X} = \{A, \alpha, \iota, \beta, B\}$, όπου η κατάσταση ι αντιστοιχεί σε ισοπαλία, οι καταστάσεις α, β αντιστοιχούν σε πλεονέκτημα του A ή του B αντίστοιχα και οι καταστάσεις A και B αντιστοιχούν στο τέλος του γκέιμ με νίκη του A ή του B αντίστοιχα. Το γκέιμ τελειώνει την πρώτη φορά που η αλυσίδα θα βρεθεί σε μια από τις καταστάσεις A ή B , επομένως η εναπομείνασα διάρκεια της παρτίδας είναι ο χρόνος πρώτης άφιξης στο σύνολο $E = \{A, B\}$. Νικητής στο game είναι ο παίκτης A , αν η αλυσίδα αυτή φτάσει στην κατάσταση A πριν φτάσει στην κατάσταση B , δηλαδή στο ενδεχόμενο $\{T_A < T_B\} = \{\omega \in \Omega : T_A(\omega) < T_B(\omega)\}$. Θυμηθείτε επίσης ότι οι

χρόνοι άφιξης είναι χρόνοι διακοπής και μπορούν να πάρουν και την τιμή $+\infty$. Αν για παράδειγμα στην έκβαση ω ο A κερδίσει τον πρώτο πόντο και στη συνέχεια ο B κερδίσει τους επόμενους τρεις πόντους, έχουμε $T_E(\omega) = T_B(\omega) = 4$, ενώ $T_A(\omega) = +\infty$.

Όπως στο παραπάνω παράδειγμα, μας ενδιαφέρει συχνά το εξής πρόβλημα: αν $A, B \subset \mathbb{X}$ με $A \cap B = \emptyset$ και $X_0 = x \notin A \cup B$, ποια είναι η πιθανότητα $\mathbb{P}[T_A < T_B | X_0 = x] = \mathbb{P}_x[T_A < T_B]$; Παρότι το πρόβλημα που έχουμε να λύσουμε απαιτεί τον υπολογισμό ενός αριθμού, είναι πιο εύκολο να λύσουμε ένα φαινομενικά πολυπλοκότερο πρόβλημα, του οποίου το αρχικό μας πρόβλημα αποτελεί μέρος. Συγκεκριμένα, θα αναζητήσουμε την τιμή αυτής της πιθανότητας ταυτόχρονα για όλα τα $x \in \mathbb{X}$. Θα θεωρήσουμε λοιπόν αυτή την πιθανότητα ως μία συνάρτηση της αρχικής κατάστασης

$$\Phi_{A,B}(x) = \mathbb{P}_x[T_A < T_B]. \quad (3.1)$$

Για λόγους που θα εξηγήσουμε στη συνέχεια η συνάρτηση αυτή καλείται συχνά *συνάρτηση δυναμικού* (potential) από το A στο B . Όπως και στην περίπτωση του χρόνου πρώτης άφιξης, αν το A είναι μονοσύνολο με $A = \{u\}$, θα συμβολίζουμε το δυναμικό $\Phi_{\{u\},B}$ και ως $\Phi_{u,B}$. Αντίστοιχα θα κάνουμε στην περίπτωση που το B είναι μονοσύνολο. Ορίζουμε επίσης τον γεννήτορα L της αλυσίδας ως έναν τελεστή που δρα σε συναρτήσεις $h : \mathbb{X} \rightarrow \mathbb{R}$ επιστρέφοντας μια καινούργια συνάρτηση $Lh : \mathbb{X} \rightarrow \mathbb{R}$ με τύπο

$$Lh(x) = \sum_{y \in \mathbb{X}} p(x,y)(h(y) - h(x)), \quad \text{για κάθε } x \in \mathbb{X}.$$

Ο τελεστής L έχει την παρακάτω πιθανοθεωρητική ερμηνεία:

$$Lh(x) = \mathbb{E}[h(X_{n+1}) - h(X_n) | X_n = x] = \mathbb{E}_x[h(X_1)] - h(x). \quad (3.2)$$

Επομένως, η τιμή $Lh(x)$ είναι η αναμενόμενη μεταβολή που επιφέρει στην τιμή της συνάρτησης h ένα βήμα μιας μαρκοβιανής αλυσίδας που βρίσκεται στο x και έχει πιθανότητες μετάβασης $\{p(x,y)\}_{x,y \in \mathbb{X}}$. Πράγματι,

$$\mathbb{E}[h(X_{n+1}) - h(X_n) | X_n = x] = \sum_{y \in \mathbb{X}} h(y)\mathbb{P}[X_{n+1} = y | X_n = x] - h(x) = \sum_{y \in \mathbb{X}} p(x,y)h(y) - h(x) = Lh(x),$$

όπου στην τελευταία ισότητα χρησιμοποιήσαμε ότι $\sum_{y \in \mathbb{X}} p(x,y) = 1$. Η δεύτερη ισότητα στην (3.2) προκύπτει θέτοντας $n = 0$.

Θεώρημα 7 Αν $A \cap B = \emptyset$, τότε η συνάρτηση δυναμικού $\Phi_{A,B}$ που ορίσαμε στην (3.1) λύνει το πρόβλημα συνοριακών τιμών (ΠΣΤ)

$$\begin{cases} Lh(x) = 0, & \text{αν } x \notin A \cup B \\ h(x) = 1, & \text{αν } x \in A \\ h(x) = 0, & \text{αν } x \in B. \end{cases} \quad (3.3)$$

Απόδειξη: Αν $x \in A$, στο ενδεχόμενο $\{\omega : X_0(\omega) = x\}$ έχουμε $T_A(\omega) = \inf\{k \geq 0 : X_k(\omega) \in A\} = 0$, ενώ $T_B(\omega) > 0$, αφού $A \cap B = \emptyset$. Επομένως,

$$x \in A \Rightarrow \Phi_{A,B}(x) = \mathbb{P}_x[T_A < T_B] = 1.$$

Με εντελώς αντίστοιχο τρόπο έχουμε ότι $x \in B \Rightarrow \Phi_{A,B}(x) = \mathbb{P}_x[T_A < T_B] = 0$.

Έστω τώρα $x \notin A \cup B$. Θα επιχειρήσουμε να υπολογίσουμε την $\mathbb{P}_x[T_A < T_B]$ κάνοντας *ανάλυση πρώτου*

βήματος. Αυτό πολύ απλά σημαίνει ότι θα εφαρμόσουμε τον τύπο της ολικής πιθανότητας για τη διαμέριση του Ω που επάγει το πρώτο βήμα της αλυσίδας,

$$\Omega = \bigcup_{y \in \mathbb{X}} \{X_1 = y\}.$$

Συγκεκριμένα,

$$\begin{aligned} \Phi_{A,B}(x) &= \mathbb{P}_x[T_A < T_B] = \sum_{y \in \mathbb{X}} \mathbb{P}[X_1 = y | X_0 = x] \times \mathbb{P}[T_A < T_B | X_0 = x, X_1 = y] \\ &= \sum_{y \in \mathbb{X}} p(x, y) \mathbb{P}[T_A^+ < T_B^+ | X_0 = x, X_1 = y], \end{aligned}$$

όπου $T_A^+(\omega) = \inf\{k \geq 1 : X_k(\omega) \in A\}$ και αντίστοιχα ορίζεται ο T_B^+ . Προσέξτε ότι η διαφορά του T_A από τον T_A^+ έγκειται στο ότι για τον μεν T_A εξετάζουμε αν η αλυσίδα είναι αρχικά στο A , ενώ για τον T_A^+ όχι. Στην τελευταία ισότητα παραπάνω, αντικαταστήσαμε το ενδεχόμενο $\{T_A < T_B\}$ από το $\{T_A^+ < T_B^+\}$, γιατί για $x \notin A \cup B$ στο ενδεχόμενο $\{X_0 = x\}$ έχουμε $T_A(\omega) = T_A^+(\omega)$ και $T_B(\omega) = T_B^+(\omega)$. Όμως από τη μαρκοβιανή ιδιότητα έχουμε

$$\begin{aligned} \mathbb{P}[T_A^+ < T_B^+ | X_0 = x, X_1 = y] &= \mathbb{P}[T_A^+ < T_B^+ | X_1 = y] \\ &= \mathbb{P}[T_A < T_B | X_0 = y] = \Phi_{A,B}(y), \end{aligned}$$

αφού η $\{Y_n\}_{n \in \mathbb{N}_0}$ με $Y_n = X_{n+1}$ είναι μια μαρκοβιανή αλυσίδα με τις ίδιες πιθανότητες μετάβασης όπως η $\{X_n\}_{n \in \mathbb{N}_0}$, ενώ $T_A^+ = \inf\{k \geq 0 : Y_n \in A\} + 1$ και αντίστοιχα $T_B^+ = \inf\{k \geq 0 : Y_n \in B\} + 1$. Έχουμε λοιπόν ότι για $x \notin A \cup B$

$$\Phi_{A,B}(x) = \sum_{y \in \mathbb{X}} p(x, y) \Phi_{A,B}(y) \Leftrightarrow L\Phi_{A,B}(x) = 0. \quad (3.4)$$

□

Παρατήρηση: Στην περίπτωση των μαρκοβιανών αλυσίδων με διακριτό χώρο καταστάσεων \mathbb{X} , το ΠΣΤ (3.3) δεν είναι παρά ένα σύστημα από τόσες γραμμικές εξισώσεις, όσες και οι καταστάσεις στο $(A \cup B)^c$. Τόσες είναι και οι τιμές της $\Phi_{A,B}$ που ψάχνουμε, αφού για $x \in A$, έχουμε $\Phi_{A,B}(x) = 1$, ενώ για $x \in B$ έχουμε $\Phi_{A,B}(x) = 0$. Όπως θα δούμε στο επόμενο κεφάλαιο, αν ο \mathbb{X} είναι πεπερασμένος και το σύνολο $(A \cup B)$ είναι προσβάσιμο από κάθε σημείο του $(A \cup B)^c$, τότε το ΠΣΤ (3.3) έχει πάντα μοναδική λύση, κάτι που μας επιτρέπει να υπολογίσουμε την $\Phi_{A,B}$, όπως θα δούμε στα επόμενα παραδείγματα.

Παράδειγμα 20 Στο Παράδειγμα 19 ας υποθέσουμε ότι η πιθανότητα νίκης του παίκτη A σε κάθε πόντο είναι p , ενώ η πιθανότητα νίκης του παίκτη B είναι q , ανεξάρτητα από το αποτέλεσμα των άλλων πόντων. Ο πίνακας πιθανοτήτων μετάβασης της αλυσίδας είναι επομένως ο

$$P = \begin{matrix} & A & \alpha & \iota & \beta & B \\ \begin{matrix} A \\ \alpha \\ \iota \\ \beta \\ B \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ p & 0 & q & 0 & 0 \\ 0 & p & 0 & q & 0 \\ 0 & 0 & p & 0 & q \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}.$$

Ας υποθέσουμε ότι αυτή τη στιγμή το γκέιμ είναι ισόπαλο, δηλαδή $X_0 = \iota$. Θα υπολογίσουμε την πιθανότητα να είναι νικητής του γκέιμ ο παίκτης A ως συνάρτηση του p , δηλαδή τη

$$\Phi_{A,B}(\iota) = \mathbb{P}[T_A < T_B | X_0 = \iota].$$

Από το Θεώρημα 7 έχουμε ότι η $\Phi_{A,B}$ λύνει το ΠΣΤ (3.3). Αν $h : \mathbb{X} \rightarrow \mathbb{R}$ είναι μια λύση αυτού του προβλήματος, θα πρέπει $h(A) = 1$, $h(B) = 0$, ενώ γράφοντας την εξίσωση $Lh(x) = 0$ διαδοχικά για $x = \alpha, \iota, \beta$ έχουμε:

$$\begin{aligned} x = \alpha : & \quad h(\alpha) = ph(A) + (1-p)h(\iota) = p + (1-p)h(\iota) \\ x = \iota : & \quad h(\iota) = ph(\alpha) + (1-p)h(\beta) \\ x = \beta : & \quad h(\beta) = ph(\iota) + (1-p)h(B) = ph(\iota). \end{aligned}$$

Λύνοντας τις παραπάνω εξισώσεις βρίσκουμε ότι το σύστημα έχει μοναδική λύση την

$$h(\alpha) = \frac{p(1-p+p^2)}{p^2+(1-p)^2}, \quad h(\iota) = \frac{p^2}{p^2+(1-p)^2}, \quad h(\beta) = \frac{p^3}{p^2+(1-p)^2}$$

και άρα $\Phi_{A,B}(\iota) = \frac{p^2}{p^2+(1-p)^2}$. □

Παράδειγμα 21 Δύο παίκτες A,B στρίβουν ένα τίμιο κέρμα μέχρι να εμφανιστεί η ακολουθία κεφαλή-γράμματα-κεφαλή (ΚΓΚ) ή να έρθουν τρεις διαδοχικές φορές γράμματα (ΓΓΓ). Στην πρώτη περίπτωση νικητής αναδεικνύεται ο A, ενώ στη δεύτερη περίπτωση ο B. Ποια είναι η πιθανότητα νίκης κάθε παίκτη;

Όπως στο Παράδειγμα 6, θα περιγράψουμε την κατάσταση του παιχνιδιού χρησιμοποιώντας το αποτέλεσμα των δύο τελευταίων στριψιμάτων και δύο ακόμα καταστάσεις A, B , που θα αντιστοιχούν σε νίκη του παίκτη A και σε νίκη του παίκτη B. Πιο συγκεκριμένα, μπορούμε να περιγράψουμε το παιχνίδι μας ως μια μαρκοβιανή αλυσίδα στον χώρο καταστάσεων

$$\mathbb{X} = \{KK, K\Gamma, \Gamma K, \Gamma\Gamma, A, B\}.$$

Εάν για παράδειγμα κάποια στιγμή η αλυσίδα βρεθεί στην κατάσταση ΓK , αυτό σημαίνει ότι καμία από τις δύο ακολουθίες τερματισμού δεν έχει εμφανιστεί ακόμη, ενώ τα δύο τελευταία στριψίματα ήταν γράμματα (το προτελευταίο) και κεφαλή (το τελευταίο). Σε αυτή την περίπτωση η επόμενη κατάσταση της αλυσίδας μπορεί να είναι είτε η KK (με πιθανότητα $1/2$), αν φέρουμε κεφαλή στο επόμενο στριψίμο, είτε η $K\Gamma$ (με πιθανότητα $1/2$), αν φέρουμε γράμματα. Αν κάποια στιγμή η αλυσίδα βρεθεί στην κατάσταση $K\Gamma$, τότε η επόμενη κατάστασή της θα είναι είτε η $\Gamma\Gamma$, αν φέρουμε γράμματα, είτε η A , αν φέρουμε κεφαλή, αφού σε αυτή την περίπτωση θα έχει σχηματιστεί η ακολουθία κεφαλή-γράμματα-κεφαλή και το παιχνίδι θα έχει τελειώσει με νικητή τον A. Με παρόμοιους συλλογισμούς μπορούμε να γράψουμε τον πίνακα πιθανοτήτων μετάβασης αυτής της αλυσίδας.

$$P = \begin{matrix} & \begin{matrix} A & KK & K\Gamma & \Gamma K & \Gamma\Gamma & B \end{matrix} \\ \begin{matrix} A \\ KK \\ K\Gamma \\ \Gamma K \\ \Gamma\Gamma \\ B \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}.$$

Η αρχική κατάσταση αυτής της αλυσίδας εξαρτάται από τα αποτελέσματα των δύο πρώτων στριψιμάτων. Επειδή στρίβουμε ένα τίμιο κέρμα, καθεμία από τις καταστάσεις $KK, K\Gamma, \Gamma K, \Gamma\Gamma$ έχει πιθανότητα $1/4$ να είναι η αρχική κατάσταση της αλυσίδας. Μπορούμε λοιπόν από τον τύπο της ολικής πιθανότητας να γράψουμε

$$\begin{aligned} \mathbb{P}[T_A < T_B] &= \sum_{x \in \mathbb{X}} \mathbb{P}[T_A < T_B \mid X_0 = x] \mathbb{P}[X_0 = x] \\ &= \frac{1}{4} \left(\Phi_{A,B}(KK) + \Phi_{A,B}(K\Gamma) + \Phi_{A,B}(\Gamma K) + \Phi_{A,B}(\Gamma\Gamma) \right). \end{aligned} \quad (3.5)$$

Αρκεί λοιπόν να βρούμε τη συνάρτηση δυναμικού $\Phi_{A,B}$, κάτι που μπορούμε να κάνουμε με τη βοήθεια του Θεωρήματος 7. Έστω $h : \mathbb{X} \rightarrow \mathbb{R}$ μία λύση του ΠΣΤ (2). Τότε θα έχουμε $h(A) = 1$, $h(B) = 0$, ενώ η εξίσωση $Lh(x) = 0$ για $x = KK, K\Gamma, \Gamma K, \Gamma\Gamma$ δίνει αντίστοιχα:

$$\begin{aligned} x = KK : \quad h(KK) &= \frac{1}{2}h(KK) + \frac{1}{2}h(K\Gamma) \Leftrightarrow h(KK) = h(K\Gamma) \\ x = K\Gamma : \quad h(K\Gamma) &= \frac{1}{2}h(A) + \frac{1}{2}h(\Gamma\Gamma) \Leftrightarrow 2h(K\Gamma) = 1 + h(\Gamma\Gamma) \\ x = \Gamma K : \quad h(\Gamma K) &= \frac{1}{2}h(KK) + \frac{1}{2}h(K\Gamma) \\ x = \Gamma\Gamma : \quad h(\Gamma\Gamma) &= \frac{1}{2}h(\Gamma K) + \frac{1}{2}h(B) \Leftrightarrow h(\Gamma K) = 2h(\Gamma\Gamma). \end{aligned}$$

Το σύστημα των τεσσάρων αυτών εξισώσεων έχει μοναδική λύση την

$$h(KK) = h(K\Gamma) = h(\Gamma K) = \frac{2}{3} \quad \text{και} \quad h(\Gamma\Gamma) = \frac{1}{3},$$

οπότε από τη σχέση (3.5) λαμβάνουμε ότι $\mathbb{P}[T_A < T_B] = \frac{7}{12}$. Μπορείτε να βρείτε μια διαισθητική εξήγηση, γιατί η πιθανότητα αυτή είναι μεγαλύτερη από $1/2$, γιατί δηλαδή είναι πιο πιθανό να εμφανιστεί πρώτα η ακολουθία ΚΓΚ παρά η ακολουθία ΓΓΓ; \square

Παράδειγμα 22 Παίζετε ένα παιχνίδι στο οποίο η πιθανότητα νίκης σας σε κάθε γύρο είναι $p \in (0, 1)$. Σε κάθε γύρο στοιχηματίζετε ένα κέρμα. Αν κερδίσετε σε έναν γύρο παίρνετε πίσω το κέρμα σας και ένα ακόμα, ενώ αν χάσετε χάνετε το κέρμα σας. Η αρχική σας περιουσία είναι x κέρματα, ενώ αυτή του αντιπάλου σας είναι y κέρματα. Το παιχνίδι συνεχίζεται μέχρι είτε εσείς είτε ο αντίπαλός σας να μείνετε χωρίς καθόλου κέρματα, οπότε νικητής αναδεικνύεται εκείνος που μάζεψε όλα τα $N = x + y$ κέρματα. Ποια είναι η πιθανότητα να νικήσετε;

Η περιουσία σας στο συγκεκριμένο παιχνίδι μπορεί να περιγραφεί από μια μαρκοβιανή αλυσίδα στον χώρο καταστάσεων $\mathbb{X} = \{0, 1, 2, \dots, N\}$. Σε κάθε γύρο είτε θα κερδίσετε (με πιθανότητα p) και η περιουσία σας θα αυξηθεί κατά 1, είτε θα χάσετε (με πιθανότητα $1 - p$) και η περιουσία σας θα μειωθεί κατά 1. Επομένως αν $z \in \{1, 2, \dots, N - 1\}$, οι πιθανότητες μετάβασης αυτής της αλυσίδας είναι

$$p(z, z + 1) = p, \quad p(z, z - 1) = 1 - p, \quad p(z, w) = 0, \quad \text{για } |z - w| \neq 1.$$

Αν ορίσουμε $T_u = \inf\{k \geq 0 : X_k = u\}$ τον χρόνο πρώτης άφιξης στην κατάσταση u , τότε το ενδεχόμενο νίκης σας είναι το $\{T_N < T_0\}$, επομένως ψάχνουμε την πιθανότητα $\mathbb{P}_x[T_N < T_0] = \Phi_{N,0}(x)$. Από την (3.3) έχουμε ότι για $z \in \{1, 2, \dots, N - 1\}$

$$\begin{aligned} 0 &= L\Phi_{N,0}(z) = \sum_{u \in \mathbb{X}} p(z, u) (\Phi_{N,0}(u) - \Phi_{N,0}(z)) \\ &= p(\Phi_{N,0}(z + 1) - \Phi_{N,0}(z)) + (1 - p)(\Phi_{N,0}(z - 1) - \Phi_{N,0}(z)). \end{aligned}$$

Ξαναγράφουμε την παραπάνω σχέση ως

$$\Phi_{N,0}(z + 1) - \Phi_{N,0}(z) = \frac{1 - p}{p} (\Phi_{N,0}(z) - \Phi_{N,0}(z - 1)), \quad z \in \{1, 2, \dots, N - 1\}. \quad (3.6)$$

Στην περίπτωση όπου $p = \frac{1}{2}$ η (3.6) σημαίνει ότι η $\{\Phi_{N,0}(n)\}_{n \in \mathbb{X}}$ είναι μια αριθμητική πρόοδος και άρα

$$\Phi_{N,0}(n) = \Phi_{N,0}(0) + cn = cn, \quad \text{για } n \in \mathbb{X}.$$

Η διαφορά της προόδου c μπορεί να βρεθεί θέτοντας $n = N$ στην παραπάνω σχέση, οπότε

$$1 = cN \Rightarrow c = \frac{1}{N}$$

και άρα

$$\Phi_{N,0}(x) = \frac{x}{N}. \quad (3.7)$$

Επομένως, όταν το παιχνίδι είναι δίκαιο, η πιθανότητα νίκης σας σε αυτό είναι απλά το κλάσμα της αρχικής σας περιουσίας στη συνολική περιουσία.

Αν τώρα έχουμε $p \neq \frac{1}{2}$, η (3.6) δίνει ότι η $\Phi_{N,0}(n+1) - \Phi_{N,0}(n)$ είναι γεωμετρική πρόοδος, άρα

$$\Phi_{N,0}(n+1) - \Phi_{N,0}(n) = \left(\frac{1-p}{p}\right)^n (\Phi_{N,0}(1) - \Phi_{N,0}(0)) = \left(\frac{1-p}{p}\right)^n \Phi_{N,0}(1), \quad \text{για } n = 0, 1, \dots, N-1.$$

Προσθέτοντας κατά μέλη τις παραπάνω σχέσεις για $n = 0, 1, \dots, z-1$, το άθροισμα στο αριστερό μέλος τηλεσκοπεί, ενώ στο δεξί μέλος έχουμε το άθροισμα των z πρώτων όρων μιας γεωμετρικής προόδου με λόγο $(1-p)/p$, οπότε παίρνουμε

$$\Phi_{N,0}(z) = \frac{\left(\frac{1-p}{p}\right)^z - 1}{\frac{1-p}{p} - 1} \Phi_{N,0}(1) = \frac{p}{1-2p} \left(\left(\frac{1-p}{p}\right)^z - 1 \right) \Phi_{N,0}(1), \quad z \in \mathbb{X}. \quad (3.8)$$

Η $\Phi_{N,0}(1)$ μπορεί τώρα να υπολογιστεί από την τελική συνθήκη $\Phi_{N,0}(N) = 1$ επιλέγοντας $z = N$ στην (3.8). Συγκεκριμένα,

$$1 = \Phi_{N,0}(N) = 1 = \frac{p}{1-2p} \left(\left(\frac{1-p}{p}\right)^N - 1 \right) \Phi_{N,0}(1).$$

Αντικαθιστώντας την τιμή που προκύπτει για την $\Phi_{N,0}(1)$ πίσω στην (3.8) βρίσκουμε ότι

$$\Phi_{N,0}(x) = \frac{\left(\frac{1-p}{p}\right)^x - 1}{\left(\frac{1-p}{p}\right)^N - 1}. \quad (3.9)$$

Παρατηρήστε ότι αν $p < 1/2$, αν δηλαδή το παιχνίδι που παίζουμε είναι εις βάρος μας, τότε $\frac{1-p}{p} > 1$ και από τη στοιχειώδη ανισότητα

$$\frac{\alpha^n - 1}{\alpha^N - 1} \leq \alpha^{n-N} \quad \text{για } \alpha \neq 1, n \leq N$$

έχουμε ότι

$$\mathbb{P}_x[T_N < T_0] \leq \left(\frac{1-p}{p}\right)^{x-N}.$$

Η πιθανότητα νίκης πέφτει λοιπόν εκθετικά με την αρχική περιουσία του αντιπάλου σας. Αξίζει τον κόπο να συγκρίνουμε τα αποτελέσματα των (3.7) και (3.9) για διάφορες τιμές του p . Παίρνοντας $N = 100$ και $x = 50$ βρίσκουμε ότι αν $p = 4/10$, η πιθανότητα νίκης μας είναι μικρότερη από 10^{-8} , αν $p = 1/2$, η πιθανότητα νίκης μας είναι ακριβώς $1/2$, ενώ αν $p = 6/10$, η πιθανότητα νίκης μας είναι μεγαλύτερη από $1 - 10^{-8}$.

Παρατήρηση: Αν στο Θεώρημα 1 πάρουμε την ειδική περίπτωση $B = \emptyset$, τότε η συνθήκη $A \cap B = \emptyset$ ικανοποιείται αυτόματα. Επιπλέον,

$$T_B(\omega) = \inf\{k \geq 0 : X_k(\omega) \in \emptyset\} = \inf(\emptyset) = +\infty$$

και άρα

$$\Phi_{A,B}(x) = \mathbb{P}_x[T_A < T_B] = \mathbb{P}_x[T_A < +\infty].$$

Έχουμε λοιπόν το ακόλουθο πόρισμα που μας δίνει τη δυνατότητα να υπολογίζουμε την πιθανότητα η αλυσίδα μας να φτάσει σε κάποιο σύνολο στόχο A σε πεπερασμένο χρόνο.

Πόρισμα 3 Αν $A \subset \mathbb{X}$ και ορίσουμε $\Phi_A(x) = \mathbb{P}_x[T_A < +\infty]$, τότε η Φ_A ικανοποιεί το ΠΣΤ

$$\begin{cases} Lh(x) = 0, & \text{αν } x \notin A \\ h(x) = 1, & \text{αν } x \in A. \end{cases} \quad (3.10)$$

Ας θυμηθούμε σ' αυτό το σημείο τη σχέση (2.10). Το να μπορούμε να υπολογίζουμε την πιθανότητα $\mathbb{P}_y[T_x < \infty]$ μας επιτρέπει να αποφανθούμε αν μια κατάσταση είναι επαναληπτική ή παροδική. Παρότι σε πεπερασμένους χώρους καταστάσεων το ερώτημα αυτό είναι εύκολο να απαντηθεί, όπως είδαμε στο Παράδειγμα 18, τα πράγματα περιπλέκονται. Ας δούμε τώρα πώς μπορούμε να ξαναβρούμε το αποτέλεσμα εκείνου του παραδείγματος, χρησιμοποιώντας τις μεθόδους αυτού του κεφαλαίου.

Παράδειγμα 23 Θεωρήστε έναν απλό συμμετρικό τυχαίο περίπατο $\{X_n\}_{n \in \mathbb{N}_0}$ στο \mathbb{Z} , δηλαδή οι πιθανότητες μετάβασης ικανοποιούν τις σχέσεις

$$p(x, x+1) = p(x, x-1) = \frac{1}{2} \quad \forall x \in \mathbb{Z}.$$

Αν $X_0 = k \in \mathbb{N}$, ποια είναι η πιθανότητα ο περίπατος να φτάσει κάποια στιγμή στο μηδέν;

Με τον συμβολισμό που αναπτύξαμε ψάχνουμε να βρούμε την $\Phi_0(k) = \mathbb{P}[T_0 < +\infty | X_0 = k]$. Με βάση το Πόρισμα 3, η συνάρτηση Φ_0 θα ικανοποιεί για $x \neq 0$ την

$$\begin{aligned} L\Phi_0(x) = 0 &\Leftrightarrow \frac{1}{2}(\Phi_0(x+1) - \Phi_0(x)) + \frac{1}{2}(\Phi_0(x-1) - \Phi_0(x)) = 0 \\ &\Leftrightarrow \Phi_0(x+1) - \Phi_0(x) = \Phi_0(x) - \Phi_0(x-1). \end{aligned}$$

Η παραπάνω σχέση σημαίνει ότι η διαφορά οποιωνδήποτε διαδοχικών όρων της ακολουθίας $\{\alpha_n\}_{n \in \mathbb{N}_0}$ με $\alpha_n = \Phi_0(n)$ είναι σταθερή και άρα η ακολουθία αυτή είναι αριθμητική πρόοδος. Μαζί με τη συνθήκη στο μηδέν $\Phi_0(0) = 1$ αυτό σημαίνει ότι υπάρχει μια σταθερά $c \in \mathbb{R}$ τέτοια ώστε

$$\Phi_0(n) = \mathbb{P}[T_0 < +\infty | X_0 = n] = 1 + cn, \quad \text{για κάθε } n \in \mathbb{N}_0.$$

Θα πρέπει όμως $0 \leq \Phi_0(n) \leq 1$ για κάθε $n \in \mathbb{N}$. Επομένως, θα πρέπει $c = 0$, που δίνει τελικά

$$\Phi_0(k) = 1.$$

Απ' όποιο $k \in \mathbb{N}$ κι αν ξεκινήσει ο συμμετρικός τυχαίος περίπατος, θα καταφέρει να φτάσει στο 0 με πιθανότητα 1. Φυσικά, λόγω συμμετρίας, το ίδιο θα ισχύει αν ξεκινήσει από κάποιο αρνητικό ακέραιο. Η (2.10) δίνει τώρα ότι $f(0) = 1$ και το Θεώρημα 3 ότι το μηδέν είναι επαναληπτική κατάσταση. Ξαναβρίσκουμε λοιπόν ότι ο απλός συμμετρικός τυχαίος περίπατος είναι επαναληπτικός.

Ένας άλλος τρόπος να δούμε το προηγούμενο αποτέλεσμα είναι ο εξής. Αν παίζουμε ένα τίμιο παιχνίδι σε ένα καζίνο (το οποίο υποθέτουμε ότι έχει άπειρο κεφάλαιο) ποντάροντας μία μάρκα κάθε φορά, όποια κι αν είναι η αρχική μας περιουσία, θα χρεωκοπήσουμε κάποια στιγμή με πιθανότητα 1. \square

Παράδειγμα 24 Αν $X_0 = k \in \mathbb{N}$ όπως στο παραπάνω παράδειγμα, αλλά ο περίπατος έχει επιπλέον μια τάση προς τα αριστερά, δηλαδή

$$p(x, x+1) = p < \frac{1}{2}, \quad p(x, x-1) = 1 - p > \frac{1}{2}, \quad \forall x \in \mathbb{Z},$$

περιμένουμε διαισθητικά ότι και πάλι θα φτάσει στο μηδέν με πιθανότητα 1. Πράγματι, κατ' αναλογία με το Παράδειγμα 22, η συνθήκη $L\Phi_0(x) = 0$ για $x > 0$ γράφεται ως

$$\Phi_0(x+1) - \Phi_0(x) = \frac{1-p}{p}(\Phi_0(x) - \Phi_0(x-1))$$

και δίνει ότι για κάθε $x \geq 0$ ισχύει

$$\Phi_0(x) = \Phi_0(0) + c \left[\left(\frac{1-p}{p} \right)^x - 1 \right] = 1 + c \left[\left(\frac{1-p}{p} \right)^x - 1 \right], \quad (3.11)$$

για κάποια σταθερά $c \in \mathbb{R}$. Η σταθερά c δεν μπορεί να είναι αυστηρά θετική, αφού στην περίπτωση αυτή θα είχαμε $\Phi_0(x) > 1$ για $x \in \mathbb{N}$. Δεν μπορούμε όμως να έχουμε ούτε $c < 0$ αφού

$$p \in \left(0, \frac{1}{2}\right) \Rightarrow \frac{1-p}{p} > 1 \Rightarrow \left(\frac{1-p}{p}\right)^x \rightarrow +\infty \text{ καθώς } x \rightarrow \infty$$

και άρα η $\Phi_0(x)$ θα έπαιρνε αρνητικές τιμές για κατάλληλα μεγάλα x . Επομένως $c = 0$ και άρα $\Phi_0(k) = 1$ για κάθε $k \in \mathbb{N}$. Γνωρίζοντας από το προηγούμενο κεφάλαιο ότι ο μη συμμετρικός τυχαίος περίπατος είναι παροδικός, περιμένουμε φυσικά ότι αυτή η πιθανότητα είναι μικρότερη του 1, αν η αλυσίδα μας ξεκινήσει από κάποιο αρνητικό ακέραιο k . Θα εξετάσουμε αυτή την περίπτωση στο Παράδειγμα 25.

Στα παραδείγματα που είδαμε ως τώρα τα ΠΣΤ (3.3) και (3.10) είχαν μοναδική λύση και προσδιόριζαν επομένως τη συνάρτηση δυναμικού. Στην περίπτωση που τα παραπάνω ΠΣΤ έχουν περισσότερες από μία λύσεις, θα θέλαμε να έχουμε ένα κριτήριο που να μας επιτρέπει να επιλέγουμε εκείνη που μας ενδιαφέρει, δηλαδή τη συνάρτηση δυναμικού. Είναι πολλές φορές εύκολο να απορρίψουμε κάποιες λύσεις γιατί π.χ. παίρνουν αρνητικές τιμές, ενώ η συνάρτηση δυναμικού είναι εξ ορισμού μη αρνητική. Κι αυτό όμως δεν είναι πάντα αρκετό, γιατί τα ΠΣΤ ενδέχεται να έχουν περισσότερες από μία μη αρνητικές λύσεις. Το ακόλουθο θεώρημα είναι χρήσιμο σε τέτοιες περιπτώσεις.

Θεώρημα 8 Ας είναι $A, B \subset \mathbb{X}$, με $A \cap B = \emptyset$. Αν η $w : \mathbb{X} \rightarrow [0, \infty)$ είναι μια μη αρνητική λύση του ΠΣΤ (3.3), τότε

$$\Phi_{A,B}(x) = \mathbb{P}_x[T_A < T_B] \leq w(x), \quad \text{για κάθε } x \in \mathbb{X}.$$

Απόδειξη: Ο ισχυρισμός ισχύει κατά τετριμμένο τρόπο, αν $x \in A$, οπότε $\Phi_{A,B}(x) = w(x) = 1$ ή $x \in B$, οπότε $\Phi_{A,B}(x) = w(x) = 0$. Αν πάλι $x \notin A \cup B$, τότε από την (3.3) έχουμε

$$0 = Lw(x) \Leftrightarrow w(x) = \sum_{y \in \mathbb{X}} p(x, y)w(y). \quad (3.12)$$

Χωρίζοντας τους προσθετέους στο παραπάνω άθροισμα ανάλογα με το αν η y είναι στα A, B ή στο $(A \cup B)^c$ και χρησιμοποιώντας τις συνοριακές συνθήκες για την w στα A, B παίρνουμε

$$\begin{aligned} w(x) &= \sum_{y \in A} p(x, y)w(y) + \sum_{y \in B} p(x, y)w(y) + \sum_{y \notin A \cup B} p(x, y)w(y) \\ &= \sum_{y \in A} p(x, y) + 0 + \sum_{y \notin A \cup B} p(x, y)w(y) \\ &= \mathbb{P}_x[1 = T_A < T_B] + \sum_{y \notin A \cup B} p(x, y)w(y). \end{aligned}$$

Εφόσον το τελευταίο άθροισμα εκτείνεται σε όρους $y \notin A \cup B$ μπορούμε να γράψουμε την $w(y)$ με τη βοήθεια της (3.12) ώστε να πάρουμε

$$\begin{aligned}
w(x) &= \mathbb{P}_x[1 = T_A < T_B] + \sum_{\substack{y \notin A \cup B \\ z \in \mathbb{X}}} p(x, y)p(y, z)w(z) \\
&= \mathbb{P}_x[1 = T_A < T_B] + \sum_{\substack{y \notin A \cup B \\ z \in A}} p(x, y)p(y, z) + \sum_{y, z \notin A \cup B} p(x, y)p(y, z)w(z) \\
&= \mathbb{P}_x[1 = T_A < T_B] + \mathbb{P}_x[2 = T_A < T_B] + \sum_{y, z \notin A \cup B} p(x, y)p(y, z)w(z) \\
&= \mathbb{P}_x[T_A \leq 2, T_A < T_B] + \sum_{y, z \notin A \cup B} p(x, y)p(y, z)w(z).
\end{aligned}$$

Επαναλαμβάνοντας το παραπάνω επιχείρημα n φορές, λαμβάνουμε

$$\begin{aligned}
w(x) &= \mathbb{P}_x[T_A \leq n, T_A < T_B] + \sum_{x_1, \dots, x_n \notin A \cup B} p(x, x_1)p(x_1, x_2) \cdots p(x_{n-1}, x_n)w(x_n) \\
&\geq \mathbb{P}_x[T_A \leq n, T_A < T_B],
\end{aligned}$$

αφού η w παίρνει μη αρνητικές τιμές. Η παραπάνω ανισότητα ισχύει για κάθε $n \in \mathbb{N}$, ενώ τα ενδεχόμενα $R_n = \{T_A \leq n\} \cap \{T_A < T_B\}$ σχηματίζουν μια αύξουσα ακολουθία ενδεχομένων. Παίρνοντας το όριο καθώς $n \rightarrow \infty$ έχουμε

$$\begin{aligned}
w(x) &\geq \lim_{n \rightarrow \infty} \mathbb{P}_x[T_A \leq n, T_A < T_B] \\
&= \mathbb{P}_x\left[\bigcup_n \{T_A \leq n\} \cap \{T_A < T_B\}\right] \\
&= \mathbb{P}_x[T_A < T_B] = \Phi_{A,B}(x).
\end{aligned}$$

Επομένως, σε κάθε περίπτωση το δυναμικό $\Phi_{A,B}(x)$ είναι μικρότερο από οποιαδήποτε μη αρνητική λύση της (3.3). \square

Στην ειδική περίπτωση όπου $B = \emptyset$ παίρνουμε το ακόλουθο πόρισμα.

Πόρισμα 4 Αν $A \subset \mathbb{X}$ και $w : \mathbb{X} \rightarrow [0, \infty)$ είναι μια μη αρνητική λύση του ΠΣΤ (3.10), τότε

$$\Phi_A(x) = \mathbb{P}_x[T_A < +\infty] \leq w(x), \quad \text{για κάθε } x \in \mathbb{X}.$$

Παράδειγμα 25 Θεωρήστε έναν τυχαίο περίπατο στους μη αρνητικούς ακεραίους με $p(x, x+1) = p > \frac{1}{2}$, $p(x, x-1) = 1-p$ για κάθε $x \in \mathbb{N}$. Αν τον ξεκινήσουμε από το $n \in \mathbb{N}$, ποια είναι η πιθανότητα να φτάσει κάποτε στο μηδέν;

Έχουμε ήδη δει στο Παράδειγμα 24 ότι, αν $p \leq \frac{1}{2}$ τότε $\Phi_0(x) = \mathbb{P}_x[T_0 < +\infty] = 1$, για κάθε $x \in \mathbb{N}$. Εδώ θα δούμε πώς συμπεριφέρεται αυτή η πιθανότητα, αν $p > 1/2$. Όπως και στο παράδειγμα (22) η συνθήκη $Lw(x) = 0$ για $x > 0$ γράφεται ως

$$w(x+1) - w(x) = \frac{1-p}{p}(w(x) - w(x-1))$$

και δίνει ότι για κάθε $x \geq 0$ έχουμε

$$w(x) = w(0) + c \left[\left(\frac{1-p}{p} \right)^x - 1 \right] = 1 + c \left[\left(\frac{1-p}{p} \right)^x - 1 \right], \quad (3.13)$$

για κάποια σταθερά $c \in \mathbb{R}$. Σε αντίθεση με το Παράδειγμα 22 όπου προσδιορίσαμε τη σταθερά c από τη συνοριακή συνθήκη στο δεξί άκρο του διαστήματος που κινείται η αλυσίδα, εδώ το ΠΣΤ (3.10) έχει άπειρες λύσεις, αφού ικανοποιείται από την w της εξίσωσης (3.13) για κάθε $c \in \mathbb{R}$. Γνωρίζουμε όμως ότι η συνάρτηση δυναμικού είναι μία από αυτές τις λύσεις και με τη βοήθεια του Πορίσματος 4 μπορούμε να προσδιορίσουμε σε ποια τιμή του c αντιστοιχεί η λύση Φ_0 που μας ενδιαφέρει. Παρατηρήστε ότι

$$p \in \left(\frac{1}{2}, 1 \right) \Rightarrow \frac{1-p}{p} < 1 \Rightarrow \left(\frac{1-p}{p} \right)^x \rightarrow 0, \quad \text{καθώς } x \rightarrow \infty.$$

Συμπεραίνουμε ότι θα πρέπει $c \leq 1$, ώστε η λύση w στην (3.13) να είναι θετική. Επιπλέον, εφόσον $\left(\frac{1-p}{p} \right)^x < 1$ για κάθε $x > 0$, η μικρότερη μη αρνητική λύση της (3.13) αντιστοιχεί σε $c = 1$ και άρα

$$\Phi_0(x) = 1 + 1 \left[\left(\frac{1-p}{p} \right)^x - 1 \right] = \left(\frac{1-p}{p} \right)^x, \quad \text{για κάθε } x \in \mathbb{N}.$$

Αξίζει να σημειώσουμε ότι για $c \in [0, 1]$ έχουμε $w(x) \in [0, 1]$, για κάθε $x \in \mathbb{N}$, επομένως δεν θα ήταν δυνατόν να επιλέξουμε την παράμετρο c που δίνει τη συνάρτηση δυναμικού με επιχειρήματα ανάλογα με αυτά που χρησιμοποιήσαμε για $p \leq \frac{1}{2}$. \square

3.3 Στατιστικά του χρόνου άφιξης

Ας υποθέσουμε ότι το σύνολο $A \subset \mathbb{X}$ είναι τέτοιο ώστε $\mathbb{P}_x[T_A < +\infty] = 1$. Μπορούμε να πούμε κάτι παραπάνω για την κατανομή του χρόνου άφιξης στο A ; Ποια είναι η αναμενόμενη τιμή του; Η κατανομή του; Σε αυτή την παράγραφο θα δούμε ότι μπορούμε να απαντήσουμε σε τέτοια ερωτήματα με μεθόδους ανάλογες με αυτές που αναπτύξαμε στην προηγούμενη παράγραφο.

Θεώρημα 9 Ας είναι $A \subset \mathbb{X}$ και $T_A = \inf\{k \geq 0 : X_k \in A\}$ ο χρόνος πρώτης άφιξης στο A . Ορίζουμε $G_A : \mathbb{X} \rightarrow [0, +\infty]$ με τύπο $G_A(x) = \mathbb{E}[T_A | X_0 = x] = \mathbb{E}_x[T_A]$. Αν το ΠΣΤ

$$\begin{cases} Lg(x) = -1, & \text{αν } x \notin A \\ g(x) = 0, & \text{αν } x \in A \end{cases} \quad (3.14)$$

έχει κάποια μη αρνητική και πεπερασμένη λύση $g : \mathbb{X} \rightarrow [0, \infty)$, τότε η G_A επίσης λύνει το (3.14) και μάλιστα

$$G_A(x) \leq g(x), \quad \text{για κάθε } x \in \mathbb{X}.$$

Απόδειξη: Ας θεωρήσουμε μια $g : \mathbb{X} \rightarrow [0, \infty)$ που λύνει το (3.14).

Αν $x \in A$, τότε $T_A(\omega) = 0$ για κάθε ω στο ενδεχόμενο $\{X_0 = x\}$. Επομένως, $G_A(x) = 0 = g(x)$.

Έστω τώρα $x \notin A$. Τότε μπορούμε να γράψουμε την $Lg(x) = -1$ ως

$$g(x) = \sum_{y \in \mathbb{X}} p(x, y)g(y) + 1 = \sum_{y \notin A} p(x, y)g(y) + 1, \quad (3.15)$$

όπου η τελευταία ισότητα προκύπτει από τη συνοριακή συνθήκη για την g στο A . Επαναλαμβάνοντας το προηγούμενο επιχείρημα έχουμε

$$\begin{aligned}
g(x) &= 1 + \sum_{y \notin A} p(x, y) \left(\sum_{z \notin A} p(y, z)g(z) + 1 \right) \\
&= 1 + \sum_{y \notin A} p(x, y) + \sum_{y, z \notin A} p(x, y)p(y, z)g(z) \\
&= 1 + \sum_{y \notin A} p(x, y) + \sum_{y, z \notin A} p(x, y)p(y, z) \left(\sum_{u \notin A} p(z, u)g(u) + 1 \right) \\
&= 1 + \sum_{y \notin A} p(x, y) + \sum_{y, z \notin A} p(x, y)p(y, z) + \sum_{y, z, u \notin A} p(x, y)p(y, z)p(z, u)g(u) \\
&= \mathbb{P}_x[T_A > 0] + \mathbb{P}_x[T_A > 1] + \mathbb{P}_x[T_A > 2] + \sum_{y, z, u \notin A} p(x, y)p(y, z)p(z, u)g(u).
\end{aligned}$$

Έπειτα από $n \in \mathbb{N}$ επαναλήψεις καταλήγουμε στην

$$\begin{aligned}
g(x) &= \sum_{k=0}^{n-1} \mathbb{P}_x[T_A > k] + \sum_{x_1, \dots, x_n \notin A} p(x, x_1)p(x_1, x_2) \cdots p(x_{n-1}, x_n)g(x_n) \\
&\geq \sum_{k=0}^{n-1} \mathbb{P}_x[T_A > k],
\end{aligned}$$

αφού η g παίρνει μη αρνητικές τιμές. Παίρνοντας το όριο $n \rightarrow \infty$ έχουμε

$$g(x) \geq \sum_{k=0}^{\infty} \mathbb{P}_x[T_A > k] = \mathbb{E}_x[T_A] = G_A(x).$$

Αποδείξαμε ως τώρα ότι αν το (3.14) έχει μία μη αρνητική και πεπερασμένη λύση g , τότε $G_A(x) \leq g(x) < +\infty$ για κάθε $x \in \mathbb{X}$. Μένει να αποδείξουμε ότι η G_A λύνει και αυτή το ΠΣΤ (3.14). Με ανάλυση πρώτου βήματος έχουμε διαδοχικά

$$\begin{aligned}
G_A(x) &= \mathbb{E}_x[T_A] = \sum_{y \in \mathbb{X}} \mathbb{P}[X_1 = y | X_0 = x] \mathbb{E}[T_A | X_0 = x, X_1 = y] \\
&= \sum_{y \in \mathbb{X}} p(x, y) \mathbb{E}[T_A^+ | X_0 = x, X_1 = y],
\end{aligned} \tag{3.16}$$

όπου $T_A^+ = \inf\{k > 0 : X_k \in A\}$. Από τη μαρκοβιανή ιδιότητα, με δεδομένο ότι $\{X_1 = y\}$, η αλυσίδα $\{Y_n\}_{n \in \mathbb{N}}$ με $Y_n = X_{n+1}$ για $n \in \mathbb{N}_0$ είναι μια αλυσίδα με τις ίδιες πιθανότητες μετάβασης όπως η $\{X_n\}_{n \in \mathbb{N}_0}$ και $Y_0 = y$, ενώ

$$T_A^+ = \inf\{k > 0 : X_k \in A\} = \inf\{k \geq 0 : Y_k \in A\} + 1.$$

Επομένως, η (3.16) γίνεται

$$\begin{aligned}
G_A(x) &= \sum_{y \in \mathbb{X}} p(x, y) \mathbb{E}[T_A^+ | X_1 = y] = \sum_{y \in \mathbb{X}} p(x, y) \mathbb{E}[T_A + 1 | Y_0 = y] \\
&= \sum_{y \in \mathbb{X}} p(x, y) (G_A(y) + 1) = 1 + \sum_{y \in \mathbb{X}} p(x, y) G_A(y).
\end{aligned} \tag{3.17}$$

Αναδιατάσσοντας τους όρους της (3.17), μπορούμε να την ξαναγράψουμε ως $LG_A(x) = -1$. \square

Ας δούμε μερικά παραδείγματα εφαρμογής του θεωρήματος 9.

Παράδειγμα 26 Σε συνέχεια του Παραδείγματος 20, αν κάποια στιγμή ένα γκέιμ τένις είναι ισόπαλο, πόσοι κατά μέση τιμή πόντοι θα παιχτούν ακόμα μέχρι να τελειώσει;

Αν ορίσουμε $E = \{A, B\}$ το σύνολο των καταστάσεων στις οποίες τελειώνει το γκέιμ, τότε, με τον συμβολισμό του Παραδείγματος 20 θέλουμε να υπολογίσουμε την $G_E(\iota) = \mathbb{E}[T_E | X_0 = \iota]$. Θεωρούμε το ΠΣΤ

$$\begin{cases} Lg(x) = -1, & \text{αν } x \notin E \\ g(x) = 0, & \text{αν } x \in E. \end{cases}$$

Γράφοντας τις εξισώσεις $Lg(x) = -1$ διαδοχικά για $x \in \{\alpha, \iota, \beta\}$ έχουμε:

$$\text{για } x = \alpha: g(\alpha) = pg(A) + (1-p)g(\iota) + 1 = (1-p)g(\iota) + 1$$

$$\text{για } x = \iota: g(\iota) = pg(\alpha) + (1-p)g(\beta) + 1$$

$$\text{για } x = \beta: g(\beta) = pg(\iota) + (1-p)g(B) + 1 = pg(\iota) + 1$$

Το σύστημα αυτό έχει μοναδική λύση

$$g(\iota) = \frac{2}{p^2 + (1-p)^2}, \quad g(\alpha) = (1-p)g(\iota) + 1, \quad g(\beta) = pg(\iota) + 1$$

που είναι μη αρνητική. Επομένως η G_E θα ικανοποιεί και αυτή το παραπάνω ΠΣΤ και λόγω μοναδικότητας θα ταυτίζεται με αυτήν τη λύση. Δηλαδή,

$$G_E(\iota) = \mathbb{E}[T_E | X_0 = \iota] = \frac{2}{p^2 + (1-p)^2}.$$

Παράδειγμα 27 Ποιος είναι ο αναμενόμενος αριθμός ζαριών που πρέπει να ρίξουμε με ένα τίμιο ζάρι μέχρι να εμφανιστούν τέσσερα διαδοχικά εξάρια; Μέχρι να εμφανιστεί για πρώτη φορά η ακολουθία 6,5,4,3;

Πριν προχωρήσουμε σε υπολογισμούς, αφιερώστε ένα λεπτό για να αναλογιστείτε σε ποια περίπτωση είναι μεγαλύτερος ο αναμενόμενος αριθμός ζαριών μέχρι την εμφάνιση της ακολουθίας στόχου. Αν δεν σας είναι φανερό, ξασκαφεθείτε την ερώτηση αφού έχετε μελετήσει τη λύση.

Μπορούμε να περιγράψουμε το πρώτο πείραμα με μια μαρκοβιανή αλυσίδα στο σύνολο $\mathbb{X} = \mathbb{N}_0$. Διαισθητικά, η κατάσταση της αλυσίδας θα δείχνει το τρέχον σερί μας από εξάρια. Πιο συγκεκριμένα, αν $\{\zeta_n\}_{n \in \mathbb{N}}$ είναι τα αποτελέσματα των ζαριών μας μπορούμε να ορίσουμε $X_0 = 0$, ενώ για $n \in \mathbb{N}$ να θέσουμε

$$X_n = k \Leftrightarrow \zeta_{n-k} \neq 6 \quad \text{και} \quad \zeta_j = 6, \quad \text{για κάθε } j \in \{n-k+1, \dots, n\}.$$

Διαπιστώνει κανείς ότι η $\{X_n\}_n$ είναι μια μαρκοβιανή αλυσίδα με πιθανότητες μετάβασης

$$p(k, k+1) = \frac{1}{6} \quad \text{και} \quad p(k, 0) = \frac{5}{6}, \quad \text{για κάθε } k \in \mathbb{N}_0.$$

Πράγματι, αν έχουμε ήδη σχηματίσει ένα σερί από k εξάρια και αν στην επόμενη ζαριά μας φέρουμε 6, τότε το σερί μας θα διευρυνθεί σε $k+1$, ενώ αν φέρουμε οτιδήποτε άλλο, το τρέχον σερί μας θα πέσει στο μηδέν. Με τον συμβολισμό που αναπτύξαμε, αν ορίσουμε

$$T_4 = \inf\{n \geq 0 : X_n = 4\}$$

τότε η απάντηση στο ερώτημά μας είναι η $\mathbb{E}[T_4 | X_0 = 0]$. Θα υπολογίσουμε αυτή την αναμενόμενη τιμή λύνοντας το ΠΣΤ

$$\begin{cases} Lg(x) = -1, & \text{αν } x \in \{0, 1, 2, 3\} \\ g(x) = 0, & \text{αν } x = 4. \end{cases}$$

Γράφουμε τις εξισώσεις που προκύπτουν διαδοχικά για $x = 0, 1, 2, 3$.

$$x = 0: g(0) = \frac{1}{6}g(1) + \frac{5}{6}g(0) + 1$$

$$x = 1: g(1) = \frac{1}{6}g(2) + \frac{5}{6}g(0) + 1$$

$$x = 2: g(2) = \frac{1}{6}g(3) + \frac{5}{6}g(0) + 1$$

$$x = 3: g(3) = \frac{1}{6}g(4) + \frac{5}{6}g(0) + 1 = \frac{5}{6}g(0) + 1.$$

Το παραπάνω σύστημα έχει μοναδική λύση $g(0) = 1.554$, $g(1) = 1.548$, $g(2) = 1.512$, $g(3) = 1.296$, επομένως από το Θεώρημα 9 θα έχουμε

$$\mathbb{E}[T_4 | X_0 = 0] = 1.554.$$

Σε ό,τι αφορά στο δεύτερο πείραμα, μπορούμε πάλι να το περιγράψουμε με τη βοήθεια μιας μαρκοβιανής αλυσίδας στον χώρο $\mathbb{X} = \{0, 1, 2, 3, 4\}$ που θα μας δείχνει πόσα από τα ψηφία της ακολουθίας-στόχου έχουμε σχηματίσει. Την πρώτη φορά που θα εμφανιστεί η ακολουθία ζαριών 6,5,4,3 η αλυσίδα αυτή θα φτάσει για πρώτη φορά στην κατάσταση 4 και θα παραμείνει στην κατάσταση αυτή εφεξής. Δεν είναι δύσκολο να βρει κανείς τον πίνακα πιθανοτήτων μετάβασης της αλυσίδας

$$P = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{pmatrix} 5/6 & 1/6 & 0 & 0 & 0 \\ 2/3 & 1/6 & 1/6 & 0 & 0 \\ 2/3 & 1/6 & 0 & 1/6 & 0 \\ 2/3 & 1/6 & 0 & 0 & 1/6 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}.$$

Για παράδειγμα, αν κάποια στιγμή η αλυσίδα είναι στην κατάσταση 2, αυτό σημαίνει ότι η προτελευταία ζαριά ήταν 6, ενώ η τελευταία 5. Αν στην επόμενη ζαριά φέρουμε 4 (πιθανότητα 1/6) η επόμενη κατάσταση της αλυσίδας θα είναι η 3, αφού θα έχουμε σχηματίσει 3 από τα 4 ψηφία της ακολουθίας στόχου. Αν φέρουμε 6 (πιθανότητα 1/6) τότε το σερί μας θα έχει σπάσει, αλλά θα έχουμε ήδη έτοιμο το πρώτο ψηφίο (6) της ακολουθίας στόχου, οπότε επόμενη κατάσταση της αλυσίδας θα είναι η 1. Αν τέλος φέρουμε 1, 2, 3, ή 5 (πιθανότητα 2/3), η επόμενη κατάσταση της αλυσίδας θα είναι το 0. Με ανάλογα επιχειρήματα μπορεί κανείς να γεμίσει και τις υπόλοιπες γραμμές του πίνακα πιθανοτήτων μετάβασης. Όπως και πριν, η $\mathbb{E}[T_4 | X_0 = 0]$ μπορεί να βρεθεί λύνοντας το ΠΣΤ

$$\begin{cases} Lg(x) = -1, & \text{αν } x \in \{0, 1, 2, 3\} \\ g(x) = 0, & \text{αν } x = 4. \end{cases}$$

Γράφουμε τις εξισώσεις που προκύπτουν διαδοχικά για $x = 0, 1, 2, 3$.

$$x = 0: g(0) = \frac{1}{6}g(1) + \frac{5}{6}g(0) + 1$$

$$x = 1: g(1) = \frac{1}{6}g(2) + \frac{1}{6}g(1) + \frac{2}{3}g(0) + 1$$

$$x = 2: g(2) = \frac{1}{6}g(3) + \frac{1}{6}g(1) + \frac{2}{3}g(0) + 1$$

$$x = 3: g(3) = \frac{1}{6}g(4) + \frac{1}{6}g(1) + \frac{2}{3}g(0) + 1 = \frac{1}{6}g(1) + \frac{2}{3}g(0) + 1.$$

Το παραπάνω σύστημα έχει μοναδική λύση $g(0) = 1.296$, $g(1) = 1.290$, $g(2) = 1.260$, $g(3) = 1.080$, επομένως από το Θεώρημα 9 θα έχουμε $\mathbb{E}[T_4 | X_0 = 0] = 1.296$. \square

Με παρόμοιο τρόπο, κάνοντας δηλαδή ανάλυση πρώτου βήματος, δεν είναι δύσκολο να βρούμε και τη γεννήτρια συνάρτηση του χρόνου πρώτης άφιξης $\Psi_A : (0, 1) \rightarrow [0, 1]$ με

$$\Psi_A(z; x) = \mathbb{E}_x[z^{T_A}],$$

η οποία χαρακτηρίζει πλήρως την κατανομή του χρόνου T_A .

Θεώρημα 10 Ας είναι $A \subset \mathbb{X}$ και $T_A = \inf\{k \geq 0 : X_k \in A\}$ ο χρόνος πρώτης άφιξης στο A . Για κάθε $z \in (0, 1)$ η συνάρτηση $\psi(x) = \mathbb{E}_x[z^{T_A}]$ ικανοποιεί το ΠΣΤ

$$\begin{cases} Lu(x) = \frac{1-z}{z}u(x), & \text{αν } x \notin A \\ u(x) = 1, & \text{αν } x \in A. \end{cases} \quad (3.18)$$

Επιπλέον, αν $u : \mathbb{X} \rightarrow [0, \infty)$ μια μη αρνητική λύση του (3.18), τότε $u(x) \geq \Psi_A(z; x)$, για κάθε $x \in \mathbb{X}$.

Απόδειξη: Αν $x \in A$, τότε $T_A(\omega) = 0$ στο ενδεχόμενο $\{\omega : X_0(\omega) = x\}$ και άρα $\psi(x) = \mathbb{E}_x[z^0] = 1$.

Αν $x \notin A$ τότε με ανάλυση πρώτου βήματος έχουμε

$$\begin{aligned} \psi(x) &= \mathbb{E}_x[z^{T_A}] = \sum_{y \in \mathbb{X}} \mathbb{P}[X_1 = y | X_0 = x] \mathbb{E}[z^{T_A} | X_0 = x, X_1 = y] \\ &= \sum_{y \in \mathbb{X}} p(x, y) \mathbb{E}[z^{T_A^+} | X_0 = x, X_1 = y] \\ &= \sum_{y \in \mathbb{X}} p(x, y) \mathbb{E}[z^{T_A^+} | X_1 = y] = \sum_{y \in \mathbb{X}} p(x, y) \mathbb{E}[z^{T_A+1} | Y_0 = y] \\ &= z \sum_{y \in \mathbb{X}} p(x, y) \psi(y), \end{aligned}$$

αφού από τη μαρκοβιανή ιδιότητα η $\{Y_n\}_{n \in \mathbb{N}_0}$ με $Y_n = X_{n+1}$ έχει τις ίδιες πιθανότητες μετάβασης όπως η $\{X_n\}_n$ και $T_A^+ = \inf\{k > 0 : X_k \in A\} = \inf\{k \geq 0 : Y_k \in A\} + 1$. Επομένως

$$\left(\frac{1}{z} - 1\right)\psi(x) = \sum_{y \in \mathbb{X}} p(x, y)\psi(y) - \psi(x) = L\psi(x).$$

Η απόδειξη του δεύτερου ισχυρισμού είναι αντίστοιχη με αυτήν του Θεωρήματος 7 και αφήνεται για εξάσκηση.

3.4 Ασκήσεις

Άσκηση 42 Θεωρήστε τη μαρκοβιανή αλυσίδα της Άσκησης 24. Υπολογίστε την πιθανότητα η αλυσίδα να καταλήξει σε καθεμία από τις κλειστές κλάσεις της, όταν ξεκινάει από τις καταστάσεις 3, 4, 5. Ποια είναι αυτή η πιθανότητα, αν αρχικά η αλυσίδα επιλέγει τυχαία μία από τις τρεις αυτές καταστάσεις και ξεκινά από εκεί; Στην τελευταία περίπτωση, ποια είναι η πιθανότητα να φτάσει η αλυσίδα στην κατάσταση 1 πριν φτάσει στην κατάσταση 8; Υπόδειξη για το τελευταίο ερώτημα: δεν χρειάζεται να λύσετε ένα ΠΣΤ από την αρχή. Χρησιμοποιήστε τις ιδιότητες των επαναληπτικών κλάσεων.

Άσκηση 43 Έχετε €1 και θέλετε να συμπληρώσετε γρήγορα ένα ποσό €10. Για το σκοπό αυτό παίζετε ένα παιχνίδι με τους εξής κανόνες. Σε κάθε γύρο η πιθανότητα νίκης σας είναι $0 < p < 1$, ανεξάρτητα από τα αποτελέσματα των προηγούμενων γύρων. Πριν από κάθε γύρο, επιλέγετε το ποσό που στοιχηματίζετε. Αν κερδίσετε σας επιστρέφεται το διπλάσιο του στοιχήματός σας, αν όχι, χάνετε το ποσό που ποντάρετε σ' αυτόν τον γύρο. Έχετε αποφασίσει να ποντάρετε όσα χρήματα έχετε, αν αυτά είναι λιγότερα από €5, διαφορετικά όσα χρειάζεστε για να φτάσετε τα €10.

- α) Ποια είναι η πιθανότητα να φτάσετε ποτέ τα €10 με αυτήν τη στρατηγική;
 β) Ποια είναι η πιθανότητα να φτάσετε ποτέ τα €10, αν σε κάθε γύρο ποντάρετε €1;
 γ) Ποιος είναι ο αναμενόμενος χρόνος μέχρι να χάσετε τα χρήματά σας ή να φτάσετε τα €10 σε καθεμία από τις παραπάνω περιπτώσεις;

Άσκηση 44 Δίνεται ο πίνακας πιθανοτήτων μετάβασης \mathbf{P} μιας μαρκοβιανής αλυσίδας στον $\mathbb{X} = \{s_1, \dots, s_5\}$, $p \in (0, 1)$.

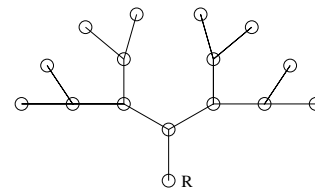
$$\mathbf{P} = \begin{matrix} & s_1 & s_2 & s_3 & s_4 & s_5 \\ \begin{matrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ p & 0 & 1-p & 0 & 0 \\ 0 & 1-p & 0 & p & 0 \\ 0 & 0 & p & 0 & 1-p \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}.$$

α) Αν $T_i = \inf\{k \geq 0 : X_k = s_i\}$ είναι ο χρόνος πρώτης άφιξης στην κατάσταση s_i , υπολογίστε την πιθανότητα

$$\mathbb{P}[T_1 < T_5 \mid X_0 = s_3].$$

β) Αν έχετε ένα κέρμα που φέρνει κορώνα με πιθανότητα $p \neq 1/2$, μπορείτε με τη βοήθεια της παραπάνω αλυσίδας να φτιάξετε έναν αλγόριθμο που μιμείται το στρίψιμο ενός τίμιου κέρματος; Συγκεκριμένα θέλουμε ο αλγόριθμος να τερματίζει με πιθανότητα 1 σε πεπερασμένο χρόνο και με πιθανότητα $1/2$ σε καθεμία από τις δύο δυνατές τελικές καταστάσεις.

Άσκηση 45 Ένα διωνυμικό δέντρο με ρίζα είναι ένας άπειρος γράφος, χωρίς κλειστά μονοπάτια, με μία διακεκριμένη κορυφή R (τη ρίζα) από την οποία διέρχεται μία ακμή, ενώ από κάθε άλλη κορυφή του διέρχονται τρεις ακμές όπως στο διπλανό σχήμα. Ένας τυχαίος περίπατος σ' αυτόν τον γράφο είναι μια μαρκοβιανή αλυσίδα στο σύνολο V των κορυφών του γράφου, που επιλέγει σε κάθε βήμα τυχαία μία από τις κορυφές με τις οποίες συνδέεται με ακμή και μετακινείται εκεί. Αν ξεκινήσουμε από την κορυφή που συνδέεται με τη ρίζα, ποια είναι η πιθανότητα να φτάσουμε κάποια στιγμή στη ρίζα; Είναι αυτός ο περίπατος παροδικός ή επαναληπτικός;



Άσκηση 46 Δείξτε ότι η πιθανότητα να κερδίσει ένα γκέιμ τένις ο παίκτης που σεβρίρει δίνεται, ως συνάρτηση της πιθανότητας p να κερδίσει κάθε πόντο, από τον τύπο

$$f(p) = p^4(10p^2 - 24p + 15) + \frac{20p^5(1-p)^3}{p^2 + (1-p)^2}.$$

Επιβεβαιώστε ότι $f(p) + f(1-p) = 1$. Υπόδειξη: Μετά από έξι πόντους, είτε το γκέιμ θα έχει τελειώσει είτε θα είναι ισόπαλο. Χρησιμοποιήστε τον τύπο ολικής πιθανότητας, ανάλογα με το τι έχει συμβεί μέχρι τότε.

Άσκηση 47 Ποιος είναι ο αναμενόμενος αριθμός των φορών που πρέπει να στρίψουμε ένα τίμιο νόμισμα, μέχρι να εμφανιστεί μια σειρά από N ίδια αποτελέσματα; Ποια είναι η απάντηση, αν η πιθανότητα να φέρουμε γράμματα σε κάθε στρίψιμο είναι $p \neq \frac{1}{2}$;

Άσκηση 48 Κάθε φορά που επισκέπτεστε ένα εστιατόριο επιλέγετε ένα από τα N πιάτα του μενού τυχαία. Ποιος είναι ο αναμενόμενος αριθμός των επισκέψεων που θα χρειαστείτε μέχρι να δοκιμάσετε όλα τα πιάτα;

Άσκηση 49 Ποιος είναι ο αναμενόμενος αριθμός ζαριών που πρέπει να ρίξουμε μέχρι το άθροισμα των ενδείξεων να είναι πολλαπλάσιο του 7;

Άσκηση 50 Δείξτε ότι ο αναμενόμενος αριθμός ζαριών μέχρι να φέρουμε n διαδοχικά θάρια με ένα τίμιο ζάρι είναι $t_n = 1,2 \times (6^n - 1)$.

Άσκηση 51 Απαντήστε τα παρακάτω ερωτήματα για τη μαρκοβιανή αλυσίδα της Άσκησης 6.

α) Αν αρχικά ο ρήγας βρίσκεται στο κέντρο, ποια είναι η πιθανότητα να βρίσκεται στο κέντρο μετά από 4 βήματα;

β) Αν αρχικά ο ρήγας είναι αριστερά, ποιος είναι ο αναμενόμενος αριθμός κινήσεων μέχρι να βρεθεί για πρώτη φορά δεξιά;

Άσκηση 52 Θεωρούμε μια μαρκοβιανή αλυσίδα $\{X_n\}$ στον $\mathbb{X} = \mathbb{N}$ με πιθανότητες μετάβασης

$$p_{k,k-1} = \frac{k-1}{2k} = \frac{1}{2} - \frac{1}{2k}, \quad p_{k,k+1} = \frac{k+1}{2k} = \frac{1}{2} + \frac{1}{2k}, \quad \text{για } k \in \mathbb{N}.$$

Η X_n έχει μια τάση να πηγαίνει δεξιά αλλά η τάση αυτή εξασθενεί όσο απομακρινόμαστε από το 1, οπότε συμπεριφέρεται σχεδόν όπως ένας απλός, συμμετρικός τυχαίος περίπατος. Για τον απλό, συμμετρικό τυχαίο περίπατο γνωρίζουμε ότι απ' όπου κι αν ξεκινήσει θα φτάσει στο 1 με πιθανότητα 1. Αν $T_1 = \inf\{n \geq 0 : X_n = 1\}$ υπολογίστε την πιθανότητα $\mathbb{P}_k[T_1 < +\infty]$.

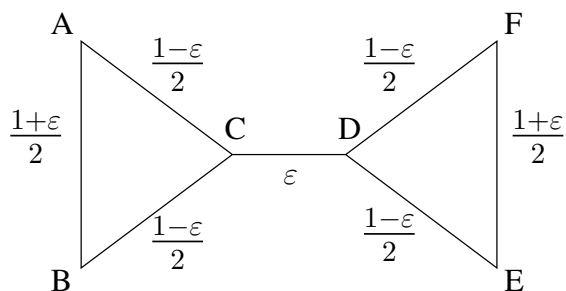
Άσκηση 53 Μια αράχνη κινείται τυχαία στον ιστό της, που αποτελείται από N ομόκεντρα εξάγωνα και τις ακτίνες τους. Πόσο χρόνο κατά μέσο όρο θα της πάρει για να φτάσει στο κέντρο του ιστού;

Άσκηση 54 Ένα έντομο κινείται στους ακεραίους. Σε κάθε βήμα του μετατοπίζεται προς την ίδια κατεύθυνση όπως στο προηγούμενο βήμα με πιθανότητα $1 - \epsilon \in (0, 1)$, ενώ με πιθανότητα ϵ μετατοπίζεται προς την αντίθετη κατεύθυνση. Μπορούμε να περιγράψουμε την κίνηση του εντόμου ως μια μαρκοβιανή αλυσίδα στον χώρο $\mathbb{X} = \mathbb{Z} \times \{+, -\}$. Για παράδειγμα, η κατάσταση $(3, +)$ σημαίνει ότι το έντομο βρίσκεται στο 3 με κατεύθυνση προς τα δεξιά. Η αρχική κατάσταση του εντόμου είναι η $(0, +)$. Υπολογίστε τον αναμενόμενο χρόνο μέχρι το έντομο να φτάσει είτε στο $N \in \mathbb{N}$ είτε στο $-N$.

Άσκηση 55 Μια μαρκοβιανή αλυσίδα κινείται ανάμεσα σε έξι καταστάσεις. Οι δυνατές μεταβάσεις εικονίζονται σαν ακμές στο διπλανό σχήμα. Οι πιθανότητες μετάβασης είναι συμμετρικές, δηλ. $p(x, y) = p(y, x)$ για κάθε $x, y \in \{A, B, C, D, E, F\}$ και δίνονται και αυτές στο σχήμα. Π.χ. $p(C, D) = p(D, C) = \epsilon$, με $0 < \epsilon < 1$.

α) Ορίζουμε $T_\epsilon = \inf\{m \geq 0 : X_m \in \{D, E, F\}\}$ τον χρόνο εισόδου στο $\{D, E, F\}$. Υπολογίστε για $x \in \{A, B, C\}$ την $\mathbb{E}[T_\epsilon | X_0 = x]$.

β) Υπολογίστε την $\psi(x) = \mathbb{E}[e^{-sT_\epsilon} | X_0 = x]$ για $x = A, B, C$ και βρείτε ποιο είναι το όριό της, καθώς $\epsilon \rightarrow 0$. Συμπεράνετε ότι η τυχαία μεταβλητή ϵT_ϵ συγκλίνει κατά κατανομή σε μια εκθετική τυχαία μεταβλητή, με ρυθμό $\lambda = 3$.



Άσκηση 56 Στην άσκηση αυτή θα υπολογίσουμε τις πιθανότητες μετάβασης μιας αλυσίδας, δεσμευμένης να φτάσει σε ένα σύνολο $A \subset \mathbb{X}$ πριν φτάσει στο B . Συγκεκριμένα, ορίζουμε για $x \in \mathbb{X} \setminus (A \cup B)$

$$\tilde{p}(x, y) = \mathbb{P}[X_{n+1} = y \mid X_n = x, n < T_A < T_B].$$

Δείξτε ότι

$$\tilde{p}(x, y) = p(x, y) \frac{\Phi_{A,B}(y)}{\Phi_{A,B}(x)}.$$

Σαν εφαρμογή λύστε το ακόλουθο πρόβλημα. Η Αλίκη και ο Βασίλης είχαν από δέκα κέρματα και έπαιξαν κορώνα/γράμματα, ποντάροντας από ένα κέρμα σε κάθε γύρο, μέχρι ένας από τους δύο να συγκεντρώσει και τα είκοσι κέρματα. Με δεδομένο ότι κέρδισε η Αλίκη, ποια είναι η πιθανότητα να μεσολάβησε κάποια στιγμή που η Αλίκη είχε μείνει με ένα κέρμα;

3.5 Αριθμητικά πειράματα

Για τις παρακάτω ασκήσεις μπορείτε να χρησιμοποιήσετε τη βασική ρουτίνα προσομοίωσης πεπερασμένων μαρκοβιανών αλυσίδων που είδαμε στο Κεφάλαιο 1. Χρησιμοποιήστε τη ρουτίνα αυτή ως βιβλιοθήκη, όπως κάναμε στον κώδικα `ex1.py`, ώστε να γράψετε κώδικες σε `python` που θα λύνουν αριθμητικά τα ακόλουθα προβλήματα.

Άσκηση 57 Υπολογίστε προσεγγιστικά την πιθανότητα της Άσκησης 49 με δύο τρόπους: α) άμεσα και β) θεωρώντας τη μαρκοβιανή αλυσίδα του υπολοίπου της ακεραίας διαίρεσης του αθροίσματος των ζαριών σας με το 7.

Άσκηση 58 Σε αυτή την άσκηση θα δούμε πώς μπορούμε να κάνουμε γραφικές παραστάσεις με την Python. Κατεβάστε το πρόγραμμα `example_plot.py` και αποθηκεύστε το στον κατάλογο που θα δουλέψετε. Τρέξτε το πρόγραμμα. Το πρόγραμμα τυπώνει τη γραφική παράσταση της συνάρτησης $x \mapsto 32x^3$ στο διάστημα $(0,6)$ σε κανονική και λογαριθμική (με βάση το 2) κλίμακα.

α) Γιατί σε λογαριθμική κλίμακα βλέπουμε μια ευθεία;

β) Εκτιμήστε γραφικά την κλίση της ευθείας και το σημείο που τέμνει τον άξονα y' .

Για τα επόμενα δύο ερωτήματα ίσως σας βοηθήσει να έχετε ανοίξει τον κώδικα με έναν επεξεργαστή κειμένου σε ένα τερματικό και να εκτελείτε το πρόγραμμα από ένα άλλο τερματικό.

γ) Αλλάξτε τη συνάρτηση σε $x \mapsto 8x^3$ και ξανατρέξτε το πρόγραμμα. Πώς αλλάζει το διάγραμμα σε λογαριθμική κλίμακα;

δ) Αλλάξτε τη συνάρτηση σε $x \mapsto 8x^2$ και ξανατρέξτε το πρόγραμμα. Πώς αλλάζει το διάγραμμα σε λογαριθμική κλίμακα;

Κλείστε τώρα το παράθυρο γραφικών και αλλάξτε το πρόγραμμα ώστε να ξαναπάρτε τη γραφική παράσταση της $x \mapsto 32x^3$. Αφαιρέστε το σύμβολο `#` από τις δύο γραμμές στο τέλος του προγράμματος

```
a,b = np.polyfit(newx,newy,1)
```

```
print "The fitted line is y=%.2f*x+%.2f " % (a,b)
```

ώστε αυτές να μετατραπούν από σχόλια σε μέρος του εκτελέσιμου προγράμματος. Η πρώτη υπολογίζει τους συντελεστές της ευθείας $y = ax + b$ που προσαρμόζεται καλύτερα στα σημεία (`newx`, `newy`). Η παράμετρος 1 δηλώνει ότι θέλουμε να προσαρμόσουμε ένα πολυώνυμο βαθμού 1, δηλαδή μια ευθεία. Η δεύτερη γραμμή απλά τυπώνει το αποτέλεσμα (με δύο δεκαδικά ψηφία) για τον χρήστη. Ξανατρέξτε το πρόγραμμα.

ε) Συμφωνεί το αποτέλεσμα για τα a, b με αυτό που βρήκατε στο ερώτημα (β);

Άσκηση 59 Παραστήστε σε ένα κοινό γράφημα τα αποτελέσματα της Άσκησης 43 ως συνάρτηση του $p \in (0, 1)$. Τι παρατηρείτε;

Άσκηση 60 Τροποποιήστε τον κώδικα της Άσκησης 40 ώστε να υπολογίζει την πιθανότητα νίκης του παίκτη που σερβίρει για $p \in \{0.01, 0.02, \dots, 0.98, 0.99\}$. Στη συνέχεια, με τη βοήθεια της Άσκησης 58 φτιάξτε σ' ένα κοινό διάγραμμα τις γραφικές παραστάσεις αυτής της πιθανότητας α) όπως την υπολογίστε αριθμητικά και β) από τον αναλυτικό τύπο της που δίνεται στην Άσκηση 46. Είναι τα δύο αποτελέσματα σε καλή συμφωνία;

Άσκηση 61 Ας υποθέσουμε ότι $p = 1/2$ στο σενάριο της Άσκησης 43. Έστω ότι θέλουμε να υπολογίσουμε με μεθόδους Monte Carlo τον αναμενόμενο αριθμό των γύρων που παίχτηκαν, δεδομένου ότι καταφέρατε να φτάσετε τα €10. Υπάρχουν δύο τρόποι να το πετύχετε αυτό. α) Να προσομοιώσετε ένα μεγάλο αριθμό παρτίδων, π.χ. $N = 10.000$, να μετρήσετε το πλήθος m από αυτές κατά τις οποίες φτάσατε πραγματικά τα €10, καθώς και τον αριθμό των γύρων T_1, \dots, T_m που παίχτηκαν σε αυτές τις παρτίδες. Η εκτιμήτρια της πιθανότητας που ψάχνετε είναι τότε η

$$\frac{T_1 + T_2 + \dots + T_m}{m}.$$

β) Να χρησιμοποιήσετε την ιδέα της Άσκησης 56, προσομοιώνοντας $N = 10.000$ φορές την αλυσίδα δεσμευμένη να φτάσει στο 10.

Γράψτε μια ρουτίνα που εκτιμά τον αναμενόμενο χρόνο 50 φορές με καθέναν από τους δύο τρόπους. Σε ποια περίπτωση τα αποτελέσματα που παίρνετε διαφέρουν μεταξύ τους λιγότερο; Μπορείτε να εξηγήσετε γιατί; Κάντε τώρα το ίδιο για $p = 1/4$. Τι παρατηρείτε;

Κεφάλαιο 4

Martingales

4.1 Εισαγωγή

Στο κεφάλαιο αυτό θα εισαγάγουμε την έννοια της δεσμευμένης μέσης τιμής για διακριτές τυχαίες μεταβλητές και θα δούμε πότε χαρακτηρίζουμε μια στοχαστική διαδικασία διακριτού χρόνου ως martingale. Μια βασική ιδιότητα που έχει μια martingale είναι ότι η αναμενόμενη τιμή της μένει σταθερή στον χρόνο ακόμα και όταν αυτή υπολογίζεται για έναν φραγμένο χρόνο διακοπής. Όπως θα δούμε, αυτή η ιδιότητα είναι ένα πολύ ισχυρό υπολογιστικό εργαλείο, ιδιαίτερα όταν η στοχαστική διαδικασία συμβαίνει να είναι ταυτόχρονα μαρκοβιανή. Παρόμοιο υλικό μπορείτε να βρείτε στις εξαιρετικές αναφορές [2], [8] και [9].

4.2 Δεσμευμένη μέση τιμή

Στην παράγραφο αυτή θα ορίσουμε τη δεσμευμένη μέση τιμή ως προς μια διακριτή τυχαία μεταβλητή. Ας υποθέσουμε λοιπόν ότι οι X, Y είναι τυχαίες μεταβλητές, ορισμένες στον ίδιο χώρο πιθανότητας, που παίρνουν τιμές στα αριθμησιμα σύνολα \mathbb{X}, \mathbb{Y} αντίστοιχα. Ας συμβολίζουμε με p_{XY} την από κοινού σ.μ.π. των X και Y , δηλαδή

$$p_{XY}(x, y) = \mathbb{P}[X = x, Y = y], \quad x \in \mathbb{X}, y \in \mathbb{Y}.$$

Οι περιθώριες σ.μ.π. των X, Y μπορούν εύκολα να υπολογιστούν από την από κοινού σ.μ.π. ως εξής.

$$p_X(x) = \mathbb{P}[X = x] = \sum_{y \in \mathbb{Y}} \mathbb{P}[X = x, Y = y], \quad x \in \mathbb{X}$$

και

$$p_Y(y) = \mathbb{P}[Y = y] = \sum_{x \in \mathbb{X}} \mathbb{P}[X = x, Y = y], \quad y \in \mathbb{Y}.$$

Υποθέτουμε χωρίς βλάβη ότι $p_Y(y) > 0$ για κάθε $y \in \mathbb{Y}$, αφού, αν $p_Y(y) = 0$, μπορούμε να αφαιρέσουμε ένα τέτοιο y από το \mathbb{Y} . Η δεσμευμένη πιθανότητα του ενδεχομένου $\{X = x\}$ δεδομένου του ενδεχομένου $\{Y = y\}$ είναι

$$\mathbb{P}[X = x | Y = y] = \frac{\mathbb{P}[X = x, Y = y]}{\mathbb{P}[Y = y]}.$$

Η δεσμευμένη πιθανότητα του ενδεχομένου $\{X = x\}$ ως προς την τυχαία μεταβλητή Y είναι μια τυχαία μεταβλητή, που είναι συνάρτηση της Y και η τιμή της, όταν $Y = y$, είναι $\mathbb{P}[X = x | Y = y]$. Έχει σημασία να έχουμε κατά νου ότι η δεσμευμένη πιθανότητα ως προς μια τυχαία μεταβλητή είναι εν γένει μια τυχαία

μεταβλητή και όχι ένας αριθμός. Η τιμή της αλλάζει ανάμεσα στα σημεία του δειγματικού χώρου Ω . Αν όμως θεωρήσουμε τη διαμέριση του Ω

$$\Omega = \bigcup_{y \in \mathbb{Y}} \{\omega \in \Omega : Y(\omega) = y\} = \bigcup_{y \in \mathbb{Y}} \{Y = y\},$$

τότε, σε καθένα από τα ενδεχόμενα $\{Y = y\}$, η τιμή της $\mathbb{P}[X = x | Y]$ παραμένει σταθερή και ίση προς $\mathbb{P}[X = x | Y = y]$. Έχουμε λοιπόν ότι

$$\mathbb{P}[X = x | Y] = p(x|Y), \quad \text{όπου } p(x|y) = \mathbb{P}[X = x | Y = y] = \frac{\mathbb{P}[X = x, Y = y]}{\mathbb{P}[Y = y]}, \quad y \in \mathbb{Y}.$$

Παρατηρήστε ότι, για κάθε $y \in \mathbb{Y}$, η $p(\cdot | y)$ είναι μια σ.μ.π. στον \mathbb{X} . Επομένως, μπορούμε να φανταζόμαστε την $\mathbb{P}[X = \cdot | Y]$ ως μια τυχαία σ.μ.π.

Ορισμός: Έστω X μια πραγματική, διακριτή, τυχαία μεταβλητή, με $\mathbb{E}[|X|] < \infty$. Ορίζουμε τη δεσμευμένη μέση τιμή (conditional expectation) της X ως προς τη διακριτή τυχαία μεταβλητή Y , ως την αναμενόμενη τιμή της X , υπολογισμένη με βάση την τυχαία σ.μ.π. $\mathbb{P}[X = \cdot | Y]$, δηλαδή

$$\mathbb{E}[X | Y] = \sum_{x \in \mathbb{X}} x \mathbb{P}[X = x | Y]. \quad (4.1)$$

Η συνθήκη $\mathbb{E}[|X|] < \infty$ εξασφαλίζει ότι το δεξί μέλος είναι πεπερασμένο. Παρατηρήστε ότι η $\mathbb{E}[X | Y]$, ως γραμμικός συνδυασμός των $\mathbb{P}[X = x | Y]$, είναι κι αυτή μια πραγματική τυχαία μεταβλητή, η οποία είναι συνάρτηση της Y . Αυτή η συνάρτηση, $g(Y) = \mathbb{E}[X | Y]$, έχει τη σταθερή τιμή $g(y) = \sum_{x \in \mathbb{X}} x \mathbb{P}[X = x | Y = y]$ σε καθένα από τα ενδεχόμενα $\{Y = y\}$.

Παράδειγμα 28 Ρίχνετε ένα ζάρι και στη συνέχεια στρίβετε τόσα κέρματα όσα η ένδειξη του ζαριού, $Y \in \{1, 2, 3, 4, 5, 6\}$. Αν $X \in \{0, 1, 2, 3, 4, 5, 6\}$ είναι το πλήθος των κεφαλών που θα φέρετε, τότε η δεσμευμένη κατανομή της X , δεδομένου ότι $Y = y$, είναι διωνυμική $\text{bin}(y, \frac{1}{2})$, δηλαδή

$$\mathbb{P}[X = x | Y = y] = \binom{y}{x} \frac{1}{2^y}.$$

Η δεσμευμένη μέση τιμή της X ως προς την Y είναι μια τυχαία μεταβλητή, που στο ενδεχόμενο $\{Y = y\}$ έχει την τιμή

$$\sum_{x=0}^y x \binom{y}{x} \frac{1}{2^y} = \frac{y}{2}.$$

Επομένως, $\mathbb{E}[X | Y] = \frac{Y}{2}$.

Θεώρημα 11 Η $g(Y) = \mathbb{E}[X | Y]$ της σχέσης (4.1) είναι η μοναδική συνάρτηση της Y για την οποία

$$\mathbb{E}[Xh(Y)] = \mathbb{E}[g(Y)h(Y)], \quad \text{για κάθε συνάρτηση } h : \mathbb{Y} \rightarrow [0, 1]. \quad (4.2)$$

Απόδειξη: Θα δείξουμε πρώτα ότι η $g(Y) = \mathbb{E}[X | Y]$ της (4.1) έχει την ιδιότητα (4.2). Πράγματι, από τον ορισμό της δεσμευμένης πιθανότητας και το Θεώρημα Fubini-Tonelli έχουμε

$$\begin{aligned} \mathbb{E}[Xh(Y)] &= \sum_{x \in \mathbb{X}} \sum_{y \in \mathbb{Y}} xh(y) \mathbb{P}[X = x, Y = y] \\ &= \sum_{x \in \mathbb{X}} \sum_{y \in \mathbb{Y}} xh(y) \mathbb{P}[X = x | Y = y] \mathbb{P}[Y = y] \\ &= \sum_{y \in \mathbb{Y}} h(y) \left(\sum_{x \in \mathbb{X}} x \mathbb{P}[X = x | Y = y] \right) \mathbb{P}[Y = y] \\ &= \sum_{y \in \mathbb{Y}} h(y)g(y) \mathbb{P}[Y = y] = \mathbb{E}[g(Y)h(Y)]. \end{aligned}$$

Θα δείξουμε τώρα ότι η g είναι η μοναδική συνάρτηση της Y που έχει την ιδιότητα (4.2). Έστω $\phi(Y)$ μια συνάρτηση της Y για την οποία

$$\mathbb{E}[Xh(Y)] = \mathbb{E}[\phi(Y)h(Y)]$$

για οποιαδήποτε συνάρτηση $h : \mathbb{Y} \rightarrow [0, 1]$. Επιλέγοντας

$$h_y(Y) = \begin{cases} 1, & \text{αν } Y = y \\ 0, & \text{αν } Y \neq y \end{cases}$$

παίρνουμε

$$\sum_{x \in \mathbb{X}} x \mathbb{P}[X = x, Y = y] = \phi(y) \mathbb{P}[Y = y].$$

Επομένως, για κάθε $y \in \mathbb{Y}$, έχουμε

$$\phi(y) = \sum_{x \in \mathbb{X}} x \frac{\mathbb{P}[X = x, Y = y]}{\mathbb{P}[Y = y]} = g(y). \quad \square$$

Το παραπάνω θεώρημα μας δίνει έναν εναλλακτικό ορισμό της δεσμευμένης μέσης τιμής ο οποίος έχει δύο πλεονεκτήματα. Αφενός, δεν κάνει αναφορά στο είδος των τυχαίων μεταβλητών X, Y , ενώ ο αρχικός ορισμός υποθέτει ότι οι X, Y έχουν διακριτή κατανομή. Αφετέρου, απλοποιεί συχνά τις αποδείξεις ισχυρισμών που αφορούν στη δεσμευμένη μέση τιμή, όπως θα δούμε και στο ακόλουθο θεώρημα.

Θεώρημα 12 Θεωρούμε Y, Z διακριτές τυχαίες μεταβλητές. Η δεσμευμένη μέση τιμή έχει τις παρακάτω ιδιότητες.

1. $\mathbb{E}[c_1X_1 + c_2X_2 | Y] = c_1\mathbb{E}[X_1 | Y] + c_2\mathbb{E}[X_2 | Y]$, για κάθε $c_1, c_2 \in \mathbb{R}$.
2. $\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$.
3. Αν οι X, Y είναι ανεξάρτητες, τότε η $\mathbb{E}[X|Y]$ είναι σταθερή και έχει την τιμή $\mathbb{E}[X]$. Ειδικότερα, $\mathbb{E}[c|Y] = c$, για κάθε $c \in \mathbb{R}$, ενώ αν η Y είναι σταθερή, τότε $\mathbb{E}[X|Y] = \mathbb{E}[X]$.
4. Αν $X_1 \leq X_2$, τότε $\mathbb{E}[X_1 | Y] \leq \mathbb{E}[X_2 | Y]$.
5. $|\mathbb{E}[X | Y]| \leq \mathbb{E}[|X| | Y]$.
6. Αν η $f : \mathbb{Y} \rightarrow \mathbb{R}$ είναι τέτοια ώστε $\mathbb{E}[|Xf(Y)|] < +\infty$, τότε $\mathbb{E}[Xf(Y) | Y] = f(Y)\mathbb{E}[X | Y]$. Ειδικότερα, αν $\mathbb{E}[|f(Y)|] < +\infty$, τότε $\mathbb{E}[f(Y) | Y] = f(Y)$.

$$7. \mathbb{E}[\mathbb{E}[X | Y, Z] | Y] = \mathbb{E}[X | Y].$$

Απόδειξη: Σε όλη την απόδειξη η h είναι μια συνάρτηση $h : \mathbb{Y} \rightarrow [0, 1]$. Για την (1), αν $g_1(Y) = \mathbb{E}[X_1 | Y]$ και $g_2(Y) = \mathbb{E}[X_2 | Y]$, από τη γραμμικότητα της μέσης τιμής έχουμε ότι

$$\begin{aligned} \mathbb{E}[(c_1 X_1 + c_2 X_2)h(Y)] &= c_1 \mathbb{E}[X_1 h(Y)] + c_2 \mathbb{E}[X_2 h(Y)] \\ &= c_1 \mathbb{E}[g_1(Y)h(Y)] + c_2 \mathbb{E}[g_2(Y)h(Y)] = \mathbb{E}[(c_1 g_1(Y) + c_2 g_2(Y))h(Y)]. \end{aligned}$$

Επομένως, από το Θεώρημα 11,

$$\mathbb{E}[c_1 X_1 + c_2 X_2 | Y] = c_1 g_1(Y) + c_2 g_2(Y) = c_1 \mathbb{E}[X_1 | Y] + c_2 \mathbb{E}[X_2 | Y].$$

Για την (2), αρκεί να επιλέξουμε $h(Y) = 1$ στο Θεώρημα 11.

Για την (3), έχουμε

$$\mathbb{E}[Xh(Y)] = \mathbb{E}[X] \mathbb{E}[h(Y)] = \mathbb{E}[\mathbb{E}[X] h(Y)].$$

Η πρώτη ισότητα ισχύει λόγω της ανεξαρτησίας των X, Y και η δεύτερη γιατί η $\mathbb{E}[X]$ είναι μια σταθερά. Από το Θεώρημα 11 έχουμε λοιπόν ότι $\mathbb{E}[X|Y] = \mathbb{E}[X]$. Υπενθυμίζουμε ότι, αν μια τυχαία μεταβλητή είναι σταθερή, τότε είναι ανεξάρτητη από κάθε άλλη τυχαία μεταβλητή. Επομένως, αν η Y είναι σταθερή, τότε $\mathbb{E}[X|Y] = \mathbb{E}[X]$, για οποιαδήποτε τυχαία μεταβλητή X . Για τον ίδιο λόγο, έχουμε ότι $\mathbb{E}[c|Y] = \mathbb{E}[c] = c$, για οποιαδήποτε τυχαία μεταβλητή Y και $c \in \mathbb{R}$.

Για την (4), από τη γραμμικότητα της δεσμευμένης μέσης τιμής (ιδιότητα 1), αρκεί να δείξουμε ότι

$$X \geq 0 \Rightarrow \mathbb{E}[X | Y] \geq 0.$$

Αν τώρα $g(Y) = \mathbb{E}[X | Y]$ και $h(Y) = \mathbb{1}\{g(Y) < 0\} \geq 0$, τότε, από το Θεώρημα 11

$$0 \leq \mathbb{E}[Xh(Y)] = \mathbb{E}[g(Y)h(Y)] = \mathbb{E}[g(Y)\mathbb{1}\{g(Y) < 0\}] \leq 0.$$

Προκειμένου όμως να έχουμε ισότητα στην τελευταία ανισότητα, θα πρέπει $\mathbb{P}[g(Y) < 0] = 0$.

Για την (5), παρατηρήστε ότι, αν ορίσουμε το θετικό και το αρνητικό μέρος της X ως $X^+ = \max\{X, 0\} \geq 0$ και $X^- = \max\{-X, 0\} \geq 0$ αντίστοιχα, τότε $X = X^+ - X^-$ και $|X| = X^+ + X^-$. Επομένως, από την τριγωνική ανισότητα, έχουμε

$$|\mathbb{E}[X | Y]| = |\mathbb{E}[X^+ | Y] - \mathbb{E}[X^- | Y]| \leq \mathbb{E}[X^+ | Y] + \mathbb{E}[X^- | Y] = \mathbb{E}[|X| | Y].$$

Στην ανισότητα παραπάνω χρησιμοποιήσαμε ότι οι $\mathbb{E}[X^+ | Y]$ και $\mathbb{E}[X^- | Y]$ είναι μη αρνητικές τυχαίες μεταβλητές από την ιδιότητα 4.

Για την (6), θα δείξουμε πρώτα τον ισχυρισμό όταν η f παίρνει τιμές στο $[0, 1]$. Σε αυτή την περίπτωση, εφόσον η $f \cdot h$ είναι κι αυτή μια συνάρτηση από το \mathbb{Y} στο $[0, 1]$, το Θεώρημα 11 δίνει ότι για κάθε $h : \mathbb{Y} \rightarrow [0, 1]$ έχουμε

$$\mathbb{E}[Xf(Y)h(Y)] = \mathbb{E}[X(f \cdot h)(Y)] = \mathbb{E}[g(Y)(f \cdot h)(Y)] = \mathbb{E}[g(Y)f(Y)h(Y)],$$

όπου $g(Y) = \mathbb{E}[X | Y]$. Επομένως, $\mathbb{E}[Xf(Y) | Y] = f(Y)\mathbb{E}[X | Y]$.

Στη συνέχεια θα δείξουμε τον ισχυρισμό όταν οι X και $f(Y)$ είναι μη αρνητικές τυχαίες μεταβλητές. Σε αυτήν την περίπτωση, για κάθε $h : \mathbb{Y} \rightarrow [0, 1]$ έχουμε

$$\mathbb{E}[Xf(Y)h(Y)] = \mathbb{E}[Xh(Y) \int_0^\infty \mathbb{1}\{f(Y) > t\} dt] = \int_0^\infty \mathbb{E}[X \mathbb{1}\{f(Y) > t\} h(Y)] dt,$$

όπου στη δεύτερη ισότητα χρησιμοποιήσαμε το Θεώρημα Fubini-Tonelli. Εφόσον η $\mathbb{1}\{f(Y) > t\}$ είναι μια συνάρτηση της Y με τιμές στο $[0,1]$, η τελευταία ισότητα γίνεται και

$$\mathbb{E}[Xf(Y)h(Y)] = \int_0^\infty \mathbb{E}[g(Y)\mathbb{1}\{f(Y) > t\}h(Y)] dt = \mathbb{E}[g(Y)f(Y)h(Y)],$$

όπου στην τελευταία ισότητα χρησιμοποιήσαμε την ιδιότητα 4 για να συμπεράνουμε ότι $g(Y) = \mathbb{E}[X | Y] \geq 0$ ώστε να χρησιμοποιήσουμε ξανά το Θεώρημα Fubini-Tonelli. Επομένως, από το Θεώρημα 11 η ιδιότητα 6 ικανοποιείται και στην περίπτωση που οι X και $f(Y)$ είναι μη αρνητικές τυχαίες μεταβλητές. Για να αποδείξουμε τώρα την ιδιότητα 6 στη γενική περίπτωση αρκεί να χρησιμοποιήσουμε τη γραμμικότητα (ιδιότητα 1) και τις διασπάσεις

$$X = X^+ - X^- \quad \text{και} \quad f(Y) = f^+(Y) - f^-(Y).$$

Πράγματι, έχουμε ότι

$$\mathbb{E}[|X^\pm f^\pm(Y)|] \leq \mathbb{E}[|X||f(Y)|] = \mathbb{E}[|Xf(Y)|] < +\infty,$$

οπότε

$$\begin{aligned} \mathbb{E}[Xf(Y) | Y] &= \mathbb{E}[(X^+ - X^-)(f^+(Y) - f^-(Y)) | Y] \\ &= \mathbb{E}[X^+f^+(Y) | Y] - \mathbb{E}[X^+f^-(Y) | Y] - \mathbb{E}[X^-f^+(Y) | Y] + \mathbb{E}[X^-f^-(Y) | Y] \\ &= f^+(Y)\mathbb{E}[X^+ | Y] - f^-(Y)\mathbb{E}[X^+ | Y] - f^+(Y)\mathbb{E}[X^- | Y] + f^-(Y)\mathbb{E}[X^- | Y] \\ &= (f^+(Y) - f^-(Y))(\mathbb{E}[X^+ | Y] - \mathbb{E}[X^- | Y]) = f(Y)\mathbb{E}[X | Y]. \end{aligned}$$

Το διαισθητικό συμπέρασμα της ιδιότητας 6 είναι πως όταν δεσμεύουμε ως προς μια τυχαία μεταβλητή Y , οι συναρτήσεις αυτής της μεταβλητής συμπεριφέρονται ως σταθερές. Τέλος, για την (7), αν ορίσουμε $G(Y, Z) = \mathbb{E}[X | Y, Z]$, έχουμε αρχικά από τις ιδιότητες 5 και 2, ότι

$$\mathbb{E}[|G(Y, Z)|] \leq \mathbb{E}[\mathbb{E}[|X| | Y, Z]] = \mathbb{E}[|X|] < \infty.$$

Η δεσμευμένη μέση τιμή $\mathbb{E}[G(Y, Z) | Y]$ είναι λοιπόν καλά ορισμένη. Από το Θεώρημα 11, αρχικά για την $G(Y, Z)$ και στη συνέχεια για τη X , έχουμε ότι

$$\mathbb{E}[G(Y, Z)h(Y)] = \mathbb{E}[Xh(Y)] = \mathbb{E}[g(Y)h(Y)].$$

Επομένως, $\mathbb{E}[G(Y, Z) | Y] = g(Y) = \mathbb{E}[X | Y]$. □

Παράδειγμα 29 Πενήντα φοιτητές από το ΕΜΠ, εβδομήντα φοιτητές από το ΕΚΠΑ και 30 φοιτητές από το ΟΠΑ παίρνουν μέρος σ' ένα διαγώνισμα. Αν επιλέξουμε τυχαία έναν από τους φοιτητές, μπορούμε να θεωρήσουμε ως δειγματικό χώρο του πειράματος τύχης το σύνολο των φοιτητών και τότε το πανεπιστήμιο προέλευσής τους Y είναι μια τυχαία μεταβλητή ορισμένη σε αυτόν τον χώρο, ενώ ο βαθμός τους στο διαγώνισμα X είναι μια άλλη τυχαία μεταβλητή. Η $\mathbb{E}[X|Y]$ είναι μια τυχαία μεταβλητή, που δίνει σε όλους τους φοιτητές του Πανεπιστημίου Y τον μέσο όρο $M(Y)$ των φοιτητών του Y . Τότε η ιδιότητα 2 του Θεωρήματος 12 σημαίνει ότι ο μέσος όρος M των βαθμών όλων των φοιτητών δίνεται από τη

$$M = \frac{50}{150}M(EMΠ) + \frac{70}{150}M(EKΠΑ) + \frac{30}{150}M(OPA),$$

δηλαδή ο μέσος όρος των βαθμών όλων των φοιτητών μπορεί να υπολογιστεί και ως ένας ζυγισμένος μέσος των μέσων βαθμών κατά πανεπιστήμιο, με βάρη τις πιθανότητες ο τυχαία επιλεγμένος φοιτητής να προέρχεται από κάθε πανεπιστήμιο. □

Στην περίπτωση που $\mathbb{E}[X^2] < +\infty$, η δεσμευμένη μέση τιμή έχει μια πολύ όμορφη γεωμετρική ερμηνεία. Πριν φτάσουμε όμως σε αυτήν χρειάζεται να δούμε κάποιες χρήσιμες ιδιότητες.

Ορισμός: Θα λέμε ότι η τυχαία μεταβλητή X είναι *τετραγωνικά ολοκληρώσιμη*, αν $\mathbb{E}[X^2] < +\infty$. Όταν αυτό συμβαίνει, θα γράφουμε $X \in L^2(\mathbb{P})$.

Το σύνολο $L^2(\mathbb{P})$ των τετραγωνικά ολοκληρώσιμων τυχαίων μεταβλητών είναι ένας γραμμικός διανυσματικός χώρος πάνω από το \mathbb{R} . Συγκεκριμένα,

$$X_1, X_2 \in L^2(\mathbb{P}) \implies c_1 X_1 + c_2 X_2 \in L^2(\mathbb{P}), \quad \text{για κάθε } c_1, c_2 \in \mathbb{R}. \quad (4.3)$$

Αυτό προκύπτει εύκολα από τη στοιχειώδη ανισότητα $(a+b)^2 \leq 2a^2 + 2b^2$, για κάθε $a, b \in \mathbb{R}$. Από την ανισότητα Cauchy-Schwarz, αν $X, Y \in L^2(\mathbb{P})$, τότε

$$\mathbb{E}[|XY|]^2 \leq \mathbb{E}[X^2]\mathbb{E}[Y^2] < +\infty. \quad (4.4)$$

Στην ειδική περίπτωση που η Y είναι σταθερή και ίση με 1 έχουμε ότι $\mathbb{E}[|X|]^2 \leq \mathbb{E}[X^2] < +\infty$. Επομένως, αν $X \in L^2(\mathbb{P})$, η $\mathbb{E}[X | Y]$ είναι καλά ορισμένη.

Θεώρημα 13 Αν $X \in L^2(\mathbb{P})$, τότε

$$\mathbb{E}[X | Y]^2 \leq \mathbb{E}[X^2 | Y]. \quad (4.5)$$

Επιπλέον, $\mathbb{E}[X | Y] \in L^2(\mathbb{P})$ και

$$\mathbb{E}[(\mathbb{E}[X | Y])^2] \leq \mathbb{E}[X^2]. \quad (4.6)$$

Απόδειξη: Για την απόδειξη της (4.5) παρατηρήστε ότι εφόσον $(X - \lambda)^2 \geq 0$ για κάθε $\lambda \in \mathbb{R}$, από τις ιδιότητες 4, 1 και 3 του Θεωρήματος 12 έχουμε ότι

$$\mathbb{E}[(X - \lambda)^2 | Y] \geq 0, \quad \forall \lambda \in \mathbb{R} \Leftrightarrow \lambda^2 - 2\lambda\mathbb{E}[X | Y] + \mathbb{E}[X^2 | Y] \geq 0, \quad \forall \lambda \in \mathbb{R}.$$

Προκειμένου αυτό το τριώνυμο ως προς λ να έχει σταθερό πρόσημο για όλα τα $\lambda \in \mathbb{R}$ θα πρέπει η διακρίνουσά του να είναι μικρότερη ή ίση με το μηδέν, απ' όπου έπεται η (4.5). Ο δεύτερος ισχυρισμός προκύπτει παίρνοντας τη μέση τιμή στα δύο μέλη της (4.5) και χρησιμοποιώντας την ιδιότητα 2 του Θεωρήματος 12 στο δεξί μέλος. \square

Η απεικόνιση από τον $L^2(\mathbb{P}) \times L^2(\mathbb{P}) \mapsto \mathbb{R}$, με $(X, Y) \mapsto \mathbb{E}[XY]$ είναι συμμετρική, γραμμική ως προς τις δύο μεταβλητές ενώ $\mathbb{E}[X^2] \geq 0$, με ισότητα αν και μόνο αν $\mathbb{P}[X = 0] = 1$. Έχει όπως λέμε τις ιδιότητες ενός εσωτερικού γινομένου. Έχοντας εφοδιάσει τον $L^2(\mathbb{P})$ με ένα εσωτερικό γινόμενο μπορούμε να μιλήσουμε για την απόσταση και τη γωνία ανάμεσα σε δύο στοιχεία του, ακριβώς όπως κάνει κανείς και στον \mathbb{R}^n .

Από την (4.6) έχουμε ότι αν $X \in L^2(\mathbb{P})$, τότε $\mathbb{E}[X | Y] \in L^2(\mathbb{P})$. Επιπλέον, από την (4.4) και τις ιδιότητες 2 και 6 του Θεωρήματος 12 έχουμε ότι

$$\mathbb{E}[Xf(Y)] = \mathbb{E}[\mathbb{E}[Xf(Y) | Y]] = \mathbb{E}[\mathbb{E}[X | Y]f(Y)], \quad \text{για κάθε } f(Y) \in L^2(\mathbb{P}).$$

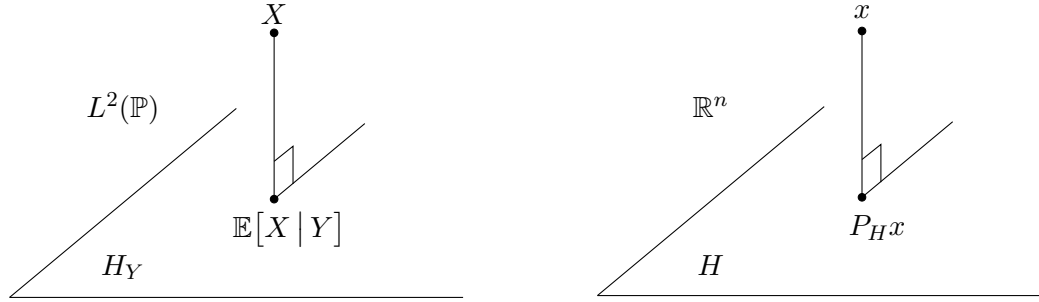
Επομένως, η δεσμευμένη μέση τιμή $\mathbb{E}[X | Y]$ μιας τυχαίας μεταβλητής X έχει το ίδιο εσωτερικό γινόμενο όπως η X με κάθε $f(Y) \in L^2(\mathbb{P})$. Ισοδύναμα, η τυχαία μεταβλητή $X - \mathbb{E}[X | Y]$ είναι ορθογώνια σε όλες τις τετραγωνικά ολοκληρώσιμες συναρτήσεις της Y .

$$\mathbb{E}[(X - \mathbb{E}[X | Y])f(Y)] = 0, \quad \text{για κάθε } f(Y) \in L^2(\mathbb{P}). \quad (4.7)$$

Στην περίπτωση του συνήθους εσωτερικού γινομένου στον \mathbb{R}^n , αν $P_H x$ είναι η ορθογώνια προβολή του διανύσματος x στον υπόχωρο H , έχουμε ότι

$$\langle x, y \rangle = \langle P_H x, y \rangle \iff \langle x - P_H x, y \rangle = 0, \quad \text{για κάθε } y \in H.$$

Μάλιστα η παραπάνω ιδιότητα χαρακτηρίζει την $P_H x$. Με βάση αυτή την παρατήρηση θα μπορούσαμε να ερμηνεύσουμε γεωμετρικά την $\mathbb{E}[X | Y]$ ως την ορθογώνια προβολή της $X \in L^2(\mathbb{P})$ στον υπόχωρο H_Y του $L^2(\mathbb{P})$ που αποτελείται από τις τετραγωνικά ολοκληρώσιμες συναρτήσεις της τυχαίας μεταβλητής Y , όπως φαίνεται στο παρακάτω σχήμα.



Με αυτή την ερμηνεία η ιδιότητα 7 του Θεωρήματος 12 αντιστοιχεί στο ότι αν ο H' είναι υπόχωρος του H , τότε $P_{H'}(P_H x) = P_{H'} x$, δηλαδή η προβολή του $P_H x$ σε έναν υπόχωρο $H' \subset H$ είναι ίδια με την απευθείας προβολή του x στον H' .

Θεώρημα 14 Αν $X \in L^2(\mathbb{P})$, τότε

$$\mathbb{E}[(X - \mathbb{E}[X | Y])^2] = \min_{f: \mathbb{Y} \rightarrow \mathbb{R}} \mathbb{E}[(X - f(Y))^2]. \quad (4.8)$$

Επομένως, αν θέλουμε να προσεγγίσουμε τη X από κάποια συνάρτηση της τυχαίας μεταβλητής Y , εκείνη που ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα της προσέγγισης είναι η $\mathbb{E}[X | Y]$.

Απόδειξη: Ας παρατηρήσουμε πρώτα ότι για την αναζήτηση της λύσης του προβλήματος ελαχιστοποίησης στο δεξί μέλος της (4.8) αρκεί να περιοριστούμε σε συναρτήσεις $f: \mathbb{Y} \rightarrow \mathbb{R}$ για τις οποίες $f(Y) \in L^2(\mathbb{P})$. Πράγματι, αν $\mathbb{E}[f^2(Y)] = +\infty$, τότε θα έχουμε και $\mathbb{E}[(f(Y) - X)^2] = +\infty$ λόγω της (4.3).

Αν τώρα $g(Y) = \mathbb{E}[X | Y]$, από την (4.6) έχουμε ότι $g(Y) \in L^2(\mathbb{P})$. Επιπλέον,

$$\begin{aligned} \mathbb{E}[(X - f(Y))^2] &= \mathbb{E}[(X - g(Y) + g(Y) - f(Y))^2] \\ &= \mathbb{E}[(X - g(Y))^2] + \mathbb{E}[(g(Y) - f(Y))^2] + 2\mathbb{E}[(X - g(Y))(g(Y) - f(Y))]. \end{aligned}$$

Εφόσον $g(Y), f(Y) \in L^2(\mathbb{P})$ λόγω της (4.3) θα έχουμε ότι $g(Y) - f(Y) \in L^2(\mathbb{P})$. Λόγω της (4.7), ο τελευταίος όρος στο δεξί μέλος της παραπάνω σχέσης είναι ίσος με μηδέν και τελικά

$$\mathbb{E}[(X - f(Y))^2] = \mathbb{E}[(X - g(Y))^2] + \mathbb{E}[(g(Y) - f(Y))^2], \quad \text{για κάθε } f(Y) \in L^2(\mathbb{P}). \quad (4.9)$$

Η παραπάνω σχέση αντιστοιχεί στο γνωστό μας Πυθαγόρειο Θεώρημα στον \mathbb{R}^n και μας δίνει ότι

$$\mathbb{E}[(X - f(Y))^2] \geq \mathbb{E}[(X - g(Y))^2], \quad \text{για κάθε } f(Y) \in L^2(\mathbb{P})$$

με ισότητα ακριβώς όταν $\mathbb{P}[f(Y) = g(Y)] = 1$. □

4.3 Martingales

Σε αυτή την παράγραφο θα δούμε πότε χαρακτηρίζουμε μια στοχαστική διαδικασία διακριτού χρόνου ως martingale και κάποιες βασικές ιδιότητες αυτών των διαδικασιών.

Θυμηθείτε από το Κεφάλαιο 2 ότι, αν έχουμε μια ακολουθία από διακριτές τυχαίες μεταβλητές, όπως π.χ. στην περίπτωση μιας στοχαστικής διαδικασίας διακριτού χρόνου $\{X_n\}_{n \in \mathbb{N}_0}$ και το ενδεχόμενο A εξαρτάται μόνο από τις τιμές των X_0, \dots, X_n , γράφουμε $A \in \mathcal{F}_n$. Για κάθε $n \in \mathbb{N}_0$, η $Y_n = (X_0, \dots, X_n)$ είναι μια τυχαία μεταβλητή με τιμές στον διακριτό χώρο \mathbb{X}^{n+1} . Θα συμβολίζουμε τη δεσμευμένη μέση τιμή μιας τυχαίας μεταβλητής Z ως προς την Y_n με $\mathbb{E}[Z | \mathcal{F}_n]$. Επομένως, η

$$\mathbb{E}[Z | \mathcal{F}_n] = \mathbb{E}[Z | X_0, \dots, X_n]$$

είναι μια τυχαία μεταβλητή που εξαρτάται μόνο από τις X_0, \dots, X_n .

Ορισμός: Αν μια τυχαία μεταβλητή X είναι συνάρτηση των X_0, \dots, X_n , θα λέμε ότι η X είναι \mathcal{F}_n -μετρήσιμη.

Ορισμός: Αν, για μια στοχαστική διαδικασία $\{M_n\}_{n \in \mathbb{N}_0}$ η τυχαία μεταβλητή M_n είναι \mathcal{F}_n -μετρήσιμη για κάθε $n \in \mathbb{N}_0$, θα χαρακτηρίζουμε την $\{M_n\}_{n \in \mathbb{N}_0}$ προσαρμοσμένη (adapted) στην $\{\mathcal{F}_n\}_{n \in \mathbb{N}_0}$.

Ορισμός: Θα χαρακτηρίζουμε μια στοχαστική διαδικασία $\{M_n\}_{n \in \mathbb{N}_0}$ με τιμές στο \mathbb{R} ως \mathcal{F}_n -martingale, αν για κάθε $n \in \mathbb{N}_0$ έχουμε

1. $\mathbb{E}[|M_n|] < \infty$,
2. $\mathbb{E}[M_{n+1} | \mathcal{F}_n] = M_n$.

Η πρώτη συνθήκη του ορισμού υπάρχει απλώς για να είναι καλά ορισμένη η δεσμευμένη μέση τιμή που εμφανίζεται στη δεύτερη. Η ουσία του ορισμού βρίσκεται στη δεύτερη συνθήκη. Αυτό που χαρακτηρίζει μια martingale είναι ότι, με δεδομένο οτιδήποτε έχει συμβεί μέχρι τη χρονική στιγμή n , η αναμενόμενη τιμή της μετά το επόμενο βήμα είναι ίδια με τη σημερινή, όπως ακριβώς συμβαίνει, σε κάθε γύρο ενός δίκαιου παιχνιδιού, με την παρουσία κάποιου που στοιχηματίζει σ' αυτό. Παρατηρήστε επίσης ότι, από την ιδιότητα 2 του ορισμού, μια διαδικασία που είναι \mathcal{F}_n -martingale, είναι πάντα προσαρμοσμένη στην $\{\mathcal{F}_n\}_{n \in \mathbb{N}_0}$.

Παράδειγμα 30 Ένας απλός, συμμετρικός τυχαίος περίπατος $\{S_n\}_{n \in \mathbb{N}_0}$ στο \mathbb{Z} είναι martingale. Πράγματι, αν $S_0 = x \in \mathbb{Z}$, έχουμε ότι για κάθε $n \in \mathbb{N}$

$$S_n = S_{n-1} + X_n = x + \sum_{k=1}^n X_k,$$

όπου η $\{X_n\}_{n \in \mathbb{N}}$ είναι μια ακολουθία ανεξάρτητων, ισόνομων τυχαίων μεταβλητών, που παίρνουν τις τιμές $+1$ ή -1 , με πιθανότητα $\frac{1}{2}$ καθεμιά. Έχουμε λοιπόν ότι $|S_n| \leq |x| + n < \infty$, ενώ

$$\mathbb{E}[S_{n+1} | \mathcal{F}_n] = \mathbb{E}[S_n + X_{n+1} | \mathcal{F}_n] = \mathbb{E}[S_n | \mathcal{F}_n] + \mathbb{E}[X_{n+1} | \mathcal{F}_n] = S_n + \mathbb{E}[X_{n+1}] = S_n.$$

Η δεύτερη ισότητα προκύπτει από την ιδιότητα 1 του Θεωρήματος 12, ενώ η τρίτη ισότητα από τις ιδιότητες 3 και 6 του ίδιου θεωρήματος.

Παράδειγμα 31 Αν $\{S_n\}_{n \in \mathbb{N}_0}$ είναι ο απλός, συμμετρικός τυχαίος περίπατος του προηγούμενου παραδείγματος, τότε η διαδικασία $\{M_n\}_{n \in \mathbb{N}_0}$ με $M_n = S_n^2 - n$ είναι martingale. Πράγματι, εφόσον $|S_n| \leq |x| + n$, έχουμε ότι $\mathbb{E}[M_n] \leq (|x| + n)^2 < \infty$. Από τις ιδιότητες 1, 3 και 6 του Θεωρήματος 12, έχουμε τώρα ότι

$$\begin{aligned} \mathbb{E}[S_{n+1}^2 | \mathcal{F}_n] &= \mathbb{E}[S_n^2 + 2S_n X_{n+1} + X_{n+1}^2 | \mathcal{F}_n] = S_n^2 + 2S_n \mathbb{E}[X_{n+1} | \mathcal{F}_n] + \mathbb{E}[1 | \mathcal{F}_n] \\ &= S_n^2 + 2S_n \mathbb{E}[X_{n+1}] + 1 = S_n^2 + 1. \end{aligned}$$

Αφαιρώντας $n + 1$ από τα δύο μέλη, παίρνουμε ότι $\mathbb{E}[M_{n+1} | \mathcal{F}_n] = M_n$.

Παράδειγμα 32 Έστω $z \in (0, 1)$. Για τον απλό, συμμετρικό, τυχαίο, περίπατο των προηγούμενων παραδειγμάτων θα εξετάσουμε αν, για κάποια σταθερά $\alpha = \alpha(z) > 0$, η διαδικασία $\{M_n\}_{n \in \mathbb{N}_0}$ με

$$M_n = \alpha^{S_n} z^n$$

είναι martingale. Η πρώτη συνθήκη του ορισμού ελέγχεται εύκολα όπως στα προηγούμενα παραδείγματα. Τώρα, από τις ιδιότητες 6 και 3 του Θεωρήματος 12

$$\begin{aligned} \mathbb{E}[M_{n+1} | \mathcal{F}_n] &= \mathbb{E}[\alpha^{S_n} \alpha^{X_{n+1}} z^{n+1} | \mathcal{F}_n] = z^{n+1} \alpha^{S_n} \mathbb{E}[\alpha^{X_{n+1}} | \mathcal{F}_n] \\ &= z M_n \mathbb{E}[\alpha^{X_{n+1}}] = \frac{1}{2} z M_n \left(\alpha + \frac{1}{\alpha} \right). \end{aligned}$$

Επομένως, η $\{M_n\}_{n \in \mathbb{N}_0}$ είναι martingale, αν και μόνο αν

$$\alpha + \frac{1}{\alpha} = \frac{2}{z} \Leftrightarrow \alpha = \frac{1 \pm \sqrt{1 - z^2}}{z}.$$

Παράδειγμα 33 (martingale του Doob). Αν η Z είναι μια τυχαία μεταβλητή με $\mathbb{E}[|Z|] < \infty$, τότε η $M_n = \mathbb{E}[Z | \mathcal{F}_n]$ είναι martingale. Πράγματι, από τις ιδιότητες 5 και 2 του Θεωρήματος 12, έχουμε

$$\mathbb{E}[|M_n|] = \mathbb{E}[|\mathbb{E}[Z | \mathcal{F}_n]|] \leq \mathbb{E}[\mathbb{E}[|Z| | \mathcal{F}_n]] = \mathbb{E}[|Z|] < \infty.$$

Επιπλέον, από την ιδιότητα 7 του Θεωρήματος 12, για κάθε $n \in \mathbb{N}_0$ έχουμε

$$\mathbb{E}[M_{n+1} | \mathcal{F}_n] = \mathbb{E}[\mathbb{E}[Z | \mathcal{F}_{n+1}] | \mathcal{F}_n] = \mathbb{E}[Z | \mathcal{F}_n] = M_n.$$

Παράδειγμα 34 (μετασχηματισμός martingale). Έστω $\{M_n\}_{n \in \mathbb{N}_0}$ μια \mathcal{F}_n -martingale ως προς κάποια διαδικασία $\{X_n\}_{n \in \mathbb{N}_0}$. Θεωρούμε μια στοχαστική διαδικασία $\{\phi_n\}_{n \in \mathbb{N}_0}$ προσαρμοσμένη στην $\{\mathcal{F}_n\}_{n \in \mathbb{N}_0}$, τέτοια ώστε, για κάθε $n \in \mathbb{N}_0$ η τυχαία μεταβλητή ϕ_n είναι φραγμένη από μια σταθερά C_n . Ορίζουμε $(\phi \cdot M)_0 = 0$ και για $n \in \mathbb{N}$

$$(\phi \cdot M)_n = \sum_{k=0}^{n-1} \phi_k (M_{k+1} - M_k).$$

Η $\{(\phi \cdot M)_n\}_{n \in \mathbb{N}_0}$ είναι μια martingale. Η ιδιότητα 1 δεν είναι δύσκολο να ελεγχθεί, ενώ

$$\mathbb{E}[(\phi \cdot M)_{n+1} | \mathcal{F}_n] = \mathbb{E}[(\phi \cdot M)_n + \phi_n (M_{n+1} - M_n) | \mathcal{F}_n] = (\phi \cdot M)_n + \phi_n \mathbb{E}[M_{n+1} - M_n | \mathcal{F}_n] = (\phi \cdot M)_n,$$

όπου στη δεύτερη ισότητα χρησιμοποιήσαμε ότι η ϕ_n είναι \mathcal{F}_n -μετρήσιμη ώστε, από την ιδιότητα 6 του Θεωρήματος 12, να τη βγάλουμε εκτός της δεσμευμένης μέσης τιμής.

Έστω ότι παίζουμε ένα τίμιο παιχνίδι όπου πριν από κάθε γύρο ποντάρουμε κάποια νομίσματα. Σε κάθε γύρο κερδίζουμε ή χάνουμε με πιθανότητα $1/2$. Αν κερδίσουμε σε κάποιον γύρο, μας επιστρέφεται το στοιχημά μας στο διπλάσιο, διαφορετικά χάνουμε όσα νομίσματα στοιχηματίσαμε. Αν πριν από κάθε γύρο ποντάρουμε ένα νόμισμα, η περιουσία μας θα είναι ένας τυχαίος περίπατος στους ακεραίους

$$M_n = x + \sum_{k=1}^n X_k.$$

Στην παραπάνω σχέση, x είναι το αρχικό πλήθος των νομισμάτων μας, ενώ οι τυχαίες μεταβλητές X_k παριστάνουν το κέρδος μας κατά τον γύρο k και έχουν μέση τιμή 0. Αν τώρα πριν από τον γύρο $k + 1$ ποντάρουμε ϕ_k νομίσματα, τότε η μεταβολή της περιουσίας μας θα είναι $\phi_k X_{k+1} = \phi_k (M_{k+1} - M_k)$ και

η περιουσία μας μετά τον n -οστό γύρο θα είναι $x + (\phi \cdot M)_n$. Μπορούμε να επιτρέψουμε στο ποντάρισμά μας να εξαρτάται απ' ό,τι έχει συμβεί μέχρι τη στιγμή που ποντάρουμε, αλλά όχι απ' ό,τι θα συμβεί στο μέλλον. Επομένως, για κάθε $k \in \mathbb{N}_0$, η ϕ_k είναι μια τυχαία μεταβλητή, που εξαρτάται μόνο από τις X_1, X_2, \dots, X_k , είναι δηλαδή \mathcal{F}_k -μετρήσιμη. Το προηγούμενο παράδειγμα μας διδάσκει ότι, ακόμα κι αν ποντάρουμε χρησιμοποιώντας τη σοφία που έχουμε κάθε φορά σωρεύσει, η περιουσία μας θα είναι και πάλι μια martingale. Το ακόλουθο θεώρημα μας διδάσκει ότι δεν μπορούμε να αυξήσουμε την αναμενόμενη περιουσία μας ποντάροντας έξυπνα.

Θεώρημα 15 Αν η διαδικασία $\{M_n\}_{n \in \mathbb{N}_0}$ είναι \mathcal{F}_n -martingale, τότε $\mathbb{E}[M_{n+k} | \mathcal{F}_k] = M_k$, για κάθε $k \in \mathbb{N}_0$ και $n \in \mathbb{N}$. Ειδικότερα, $\mathbb{E}[M_n] = \mathbb{E}[M_0]$, για κάθε $n \in \mathbb{N}$.

Απόδειξη: Έστω $k \in \mathbb{N}_0$. Θα αποδείξουμε τον ισχυρισμό με επαγωγή επί του n . Για $n = 1$ ο ισχυρισμός συμπίπτει με την ιδιότητα 2 του ορισμού μιας martingale. Έστω τώρα ότι $\mathbb{E}[M_{n+k} | \mathcal{F}_k] = M_k$ για κάποιο $n \in \mathbb{N}$. Από την ιδιότητα 7 του Θεωρήματος 12, έχουμε

$$\mathbb{E}[M_{n+k+1} | \mathcal{F}_k] = \mathbb{E}\left[\mathbb{E}[M_{n+k+1} | \mathcal{F}_{n+k}] \mid \mathcal{F}_k\right] = \mathbb{E}[M_{n+k} | \mathcal{F}_k] = M_k,$$

όπου η τελευταία ισότητα προκύπτει από την επαγωγική υπόθεση. Ειδικότερα, για $k = 0$ έχουμε $\mathbb{E}[M_n | \mathcal{F}_0] = M_0$. Το ζητούμενο προκύπτει τώρα από την ιδιότητα 2 του Θεωρήματος 12, παίρνοντας την αναμενόμενη τιμή των δύο μελών. \square

Σύμφωνα με το προηγούμενο θεώρημα, η αναμενόμενη τιμή μιας martingale είναι η ίδια σε όποια χρονική στιγμή $n \in \mathbb{N}_0$ κι αν την υπολογίσουμε. Ένα κεντρικό αποτέλεσμα στη μελέτη των martingale είναι ότι η παραπάνω ιδιότητα παραμένει σε ισχύ ακόμα κι αν αυτή η χρονική στιγμή είναι ένας φραγμένος χρόνος διακοπής.

Θεώρημα 16 [επιλεκτικής διακοπής (optional stopping)] Αν η διαδικασία $\{M_n\}_{n \in \mathbb{N}_0}$ είναι \mathcal{F}_n -martingale και ο τυχαίος χρόνος T είναι φραγμένος χρόνος διακοπής της $\{X_n\}_{n \in \mathbb{N}_0}$, τότε

$$\mathbb{E}[M_T] = \mathbb{E}[M_0].$$

Απόδειξη: Έστω N ένα άνω φράγμα του χρόνου T . Έχουμε τότε ότι $1 = \sum_{k=0}^N \mathbb{1}\{T = k\}$ και

$$\mathbb{E}[M_T] = \sum_{k=0}^N \mathbb{E}[M_T \mathbb{1}\{T = k\}] = \sum_{k=0}^N \mathbb{E}[M_k \mathbb{1}\{T = k\}]. \quad (4.10)$$

Από το Θεώρημα 15 έχουμε ότι $\mathbb{E}[M_N | \mathcal{F}_k] = M_k$, για $k = 0, 1, \dots, N-1$. Εφόσον ο T είναι χρόνος διακοπής, η $\mathbb{1}\{T = k\}$ εξαρτάται μόνο από τις X_0, \dots, X_k και παίρνει τιμές στο διάστημα $[0, 1]$. Επομένως, από το Θεώρημα 11 έχουμε ότι

$$\mathbb{E}[M_N \mathbb{1}\{T = k\}] = \mathbb{E}[M_k \mathbb{1}\{T = k\}].$$

Αντικαθιστώντας τώρα στην (4.10) έχουμε ότι

$$\mathbb{E}[M_T] = \sum_{k=0}^N \mathbb{E}[M_N \mathbb{1}\{T = k\}] = \mathbb{E}[M_N] = \mathbb{E}[M_0],$$

όπου η τελευταία ισότητα προκύπτει πάλι από το Θεώρημα 15. \square

Παρατήρηση 1: Το προηγούμενο θεώρημα μας διδάσκει ότι δεν μπορούμε να αυξήσουμε την αναμενόμενη περιουσία μας σ' ένα τίμιο παιχνίδι ακολουθώντας μια στρατηγική εξόδου που μπορεί να εξαρτάται

από ό,τι έχει συμβεί στο παρελθόν, αλλά οφείλει να έχει ολοκληρωθεί μέχρι κάποια συγκεκριμένη χρονική στιγμή N .

Παρατήρηση 2: Η υπόθεση ότι ο χρόνος διακοπής T είναι φραγμένος είναι ουσιαστική. Δεν μπορεί εν γένει να αντικατασταθεί από την ασθενέστερη συνθήκη $\mathbb{P}[T < \infty] = 1$, όπως φαίνεται στο παράδειγμα που ακολουθεί.

Παράδειγμα 35 Θεωρούμε έναν απλό, συμμετρικό, τυχαίο περίπατο $\{S_n\}_{n \in \mathbb{N}_0}$ στο \mathbb{Z} , που ξεκινά από το μηδέν. Έστω $T_1 = \inf\{k \geq 0 : X_k = 1\}$ ο χρόνος πρώτης άφιξης του περιπάτου στο 1. Όπως είδαμε στο Παράδειγμα 14, ο T_1 είναι χρόνος διακοπής, ενώ, από το Παράδειγμα 18, ο $\{S_n\}_{n \in \mathbb{N}_0}$ είναι επαναληπτικός. Το Λήμμα 4 δίνει λοιπόν ότι $\mathbb{P}_0[T_1 < \infty] = 1$. Είναι φανερό όμως ότι στο ενδεχόμενο $\{T_1 < \infty\}$ έχουμε $S_{T_1} = 1$, επομένως $\mathbb{E}[S_{T_1}] = 1$, ενώ $\mathbb{E}[S_0] = 0$.

Παρατήρηση 3: Στην πράξη είναι σπάνιο να μπορεί να εξασφαλίσει κανείς ότι ένας ενδιαφέρων χρόνος διακοπής T είναι φραγμένος. Σύμφωνα όμως με το Παράδειγμα 13, ο σταθερός χρόνος $N \in \mathbb{N}$ είναι χρόνος διακοπής, ενώ, αξιοποιώντας το Πόρισμα 1, ο χρόνος $T \wedge N = \min\{T, N\}$ είναι χρόνος διακοπής και είναι προφανώς φραγμένος από το N . Εφαρμόζουμε λοιπόν συνήθως το Θεώρημα επιλεκτικής διακοπής για τον $T \wedge N$ και προσπαθούμε να περάσουμε το αποτέλεσμα που παίρνουμε στον αρχικό χρόνο διακοπής T , στέλνοντας το N στο άπειρο. Θα δούμε πώς ακριβώς δουλεύει αυτή η διαδικασία στα παραδείγματα που ακολουθούν.

Παράδειγμα 36 Ένας απλός, συμμετρικός τυχαίος περίπατος $\{S_n\}_{n \in \mathbb{N}_0}$ ξεκινάει από το $x \in \mathbb{Z}$. Αν $a, b \in \mathbb{Z}$, με $a < x < b$ και T_a, T_b είναι οι χρόνοι πρώτης άφιξης στα a, b αντίστοιχα,

$$\mathbb{P}_x[T_b < T_a] = \frac{x - a}{b - a}.$$

Είδαμε στο Παράδειγμα 30 ότι η $\{S_n\}_{n \in \mathbb{N}_0}$ είναι martingale. Είδαμε ακόμα στο παράδειγμα 14 ότι οι T_a, T_b αλλά και ο χρόνος

$$T = T_{a,b} = T_a \wedge T_b = \inf\{k \geq 0 : S_k \in \{a, b\}\}$$

είναι χρόνοι διακοπής. Θα εφαρμόσουμε το Θεώρημα επιλεκτικής διακοπής για τη martingale $\{S_n\}_{n \in \mathbb{N}_0}$ και τον φραγμένο χρόνο διακοπής $T \wedge N$. Έχουμε λοιπόν

$$x = \mathbb{E}_x[S_0] = \mathbb{E}_x[S_{T \wedge N}] = \mathbb{E}_x[S_T \mathbb{1}\{T < N\}] + \mathbb{E}_x[S_N \mathbb{1}\{T \geq N\}].$$

Παρατηρήστε τώρα ότι $S_T = a$ στο ενδεχόμενο $\{T_a < T_b \wedge N\}$, ενώ $S_T = b$ στο ενδεχόμενο $\{T_b < T_a \wedge N\}$. Αν αφαιρέσουμε a από τα δύο μέλη της παραπάνω σχέσης, έχουμε επομένως ότι

$$x - a = (b - a) \mathbb{P}_x[T_b < T_a \wedge N] + \mathbb{E}_x[(S_N - a) \mathbb{1}\{T \geq N\}]. \quad (4.11)$$

Παρατηρήστε τώρα ότι η ακολουθία ενδεχομένων $B_N = \{T_b < T_a \wedge N\}$ είναι αύξουσα και $\cup_N B_N = \{T_b < T_a\}$. Επομένως, αν στείλουμε το N στο άπειρο, ο πρώτος προσθετέος στο δεξί μέλος της (4.11) τείνει στο $(b - a) \mathbb{P}_x[T_b < T_a]$. Σε ό,τι αφορά στον δεύτερο προσθετέο, παρατηρήστε ότι στο ενδεχόμενο $C_N = \{T \geq N\}$ ο περίπατος μέχρι τη χρονική στιγμή N δεν έχει βγει από το σύνολο $\{a, a+1, \dots, b-1, b\}$. Επομένως,

$$\left| \mathbb{E}_x[(S_N - a) \mathbb{1}\{T \geq N\}] \right| \leq (b - a) \mathbb{P}_x[T \geq N] \rightarrow (b - a) \mathbb{P}_x[T = \infty],$$

αφού η ακολουθία ενδεχομένων C_N είναι φθίνουσα, με $\cap_N C_N = \{T = \infty\}$. Όπως έχουμε δει όμως ο απλός, συμμετρικός τυχαίος περίπατος είναι επαναληπτικός, επομένως $\mathbb{P}_x[T = \infty] = 0$. Καταλήγουμε λοιπόν στην απόδειξη του ισχυρισμού μας παίρνοντας $N \rightarrow \infty$ στην (4.11).

Παράδειγμα 37 Στο προηγούμενο παράδειγμα, ο αναμενόμενος χρόνος άφιξης του απλού συμμετρικού τυχαίου περιπάτου στα άκρα $\{a, b\}$ είναι

$$\mathbb{E}_x[T_{a,b}] = (x - a)(b - x).$$

Θα χρησιμοποιήσουμε τη martingale του Παραδείγματος 31 και τον φραγμένο χρόνο διακοπής $T_{a,b} \wedge N$. Έχουμε για κάθε $N \in \mathbb{N}$ ότι

$$x^2 = \mathbb{E}_x[M_0] = \mathbb{E}_x[M_{T_{a,b} \wedge N}] = \mathbb{E}_x[S_{T_{a,b} \wedge N}^2] - \mathbb{E}_x[T_{a,b} \wedge N]. \quad (4.12)$$

Όπως στο προηγούμενο παράδειγμα,

$$\lim_{N \rightarrow \infty} \mathbb{E}_x[S_{T_{a,b} \wedge N}^2] = a^2 \mathbb{P}_x[T_a < T_b] + b^2 \mathbb{P}_x[T_b < T_a] = x(a + b) - ab.$$

Προκειμένου να υπολογίσουμε το όριο του $\mathbb{E}_x[T_{a,b} \wedge N]$, μπορούμε είτε να επικαλεστούμε το Θεώρημα μονότονης σύγκλισης της Θεωρίας Μέτρου (δείτε το Πρόσχημα 1.6 στο [8]), είτε να χρησιμοποιήσουμε το ακόλουθο χρήσιμο αποτέλεσμα από τις Πιθανότητες. Αν X είναι μια τυχαία μεταβλητή που παίρνει με πιθανότητα 1 τιμές στους μη αρνητικούς ακεραίους, όπως π.χ. η $T_{a,b}$ και η $T_{a,b} \wedge N$ στο παράδειγμά μας, τότε

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} \mathbb{P}[X > k].$$

Έχουμε λοιπόν ότι

$$\mathbb{E}_x[T_{a,b} \wedge N] = \sum_{k=0}^{\infty} \mathbb{P}_x[T_{a,b} \wedge N > k] = \sum_{k=0}^{N-1} \mathbb{P}_x[T_{a,b} > k].$$

Επομένως,

$$\lim_{N \rightarrow \infty} \mathbb{E}_x[T_{a,b} \wedge N] = \lim_{N \rightarrow \infty} \sum_{k=0}^{N-1} \mathbb{P}_x[T_{a,b} > k] = \sum_{k=0}^{\infty} \mathbb{P}_x[T_{a,b} > k] = \mathbb{E}_x[T_{a,b}].$$

Καταλήγουμε τώρα στην απόδειξη του ισχυρισμού μας παίρνοντας το όριο $N \rightarrow \infty$ στην (4.12).

Στην ειδική περίπτωση όπου $-a = b = R$ και $x = 0$, αν συμβολίσουμε με $T_R = \inf\{k \geq 0 : |S_k| \geq R\}$ τον χρόνο άφιξης του περιπάτου στο $\{-R, R\}$, έχουμε

$$\mathbb{E}_0[T_R] = R^2.$$

Βλέπουμε λοιπόν ότι ο αναμενόμενος χρόνος μέχρι ο απλός, συμμετρικός τυχαίος περίπατος να διανύσει απόσταση R , είναι R^2 . Αυτό είναι ένα χαρακτηριστικό των διαχυτικών διαδικασιών, που δεν έχουν την τάση να μετακινούνται προς κάποια κατεύθυνση και η μετακίνησή τους συμβαίνει μόνο ως αποτέλεσμα τυχαίων διακυμάνσεων.

Παράδειγμα 38 Μπορούμε, χρησιμοποιώντας το Θεώρημα επιλεκτικής διακοπής, να υπολογίσουμε ακόμα και την κατανομή του χρόνου T_R στο προηγούμενο παράδειγμα. Συγκεκριμένα, η γεννήτρια συνάρτηση του χρόνου T_R δίνεται από την

$$\mathbb{E}_x[z^{T_R}] = \frac{\alpha^x(z) + \alpha^{-x}(z)}{\alpha^R(z) + \alpha^{-R}(z)}, \quad z \in (0, 1),$$

όπου $\alpha(z) = \frac{1 - \sqrt{1 - z^2}}{z}$.

Θυμηθείτε από το Παράδειγμα 32 ότι, για $z \in (0, 1)$, η διαδικασία $\alpha^{S_n} z^n$ είναι martingale, αν

$$\alpha = \frac{1 \pm \sqrt{1 - z^2}}{z}.$$

Παρατηρήστε ότι οι παραπάνω τιμές του α είναι η μία αντίστροφη της άλλης. Εύκολα επίσης ελέγχεται ότι το άθροισμα δύο martingale είναι ομοίως μια martingale. Επομένως, η διαδικασία $\{M_n\}_{n \in \mathbb{N}_0}$, με

$$M_n = (\alpha^{S_n} + \alpha^{-S_n})z^n,$$

είναι martingale. Εφαρμόζοντας το Θεώρημα επιλεκτικής διακοπής γι' αυτήν τη martingale και τον φραγμένο χρόνο διακοπής $T_R \wedge N$, έχουμε

$$\alpha^x + \alpha^{-x} = \mathbb{E}_x[M_0] = \mathbb{E}_x[M_{T_R \wedge N}] = \mathbb{E}_x[M_{T_R} \mathbb{1}\{T_R \leq N\}] + \mathbb{E}_x[M_N \mathbb{1}\{T_R > N\}]. \quad (4.13)$$

Παρατηρήστε τώρα ότι, στο ενδεχόμενο $T_R > N$, ο περίπατος βρίσκεται μέσα στο σύνολο $I_R = \{-R, \dots, R\}$ τη χρονική στιγμή N . Χρησιμοποιώντας ότι $\alpha^k + \alpha^{-k} \leq \alpha^R + \alpha^{-R}$ για κάθε $k \in I_R$ και ότι $z \in (0, 1)$ έχουμε ότι

$$\mathbb{E}_x[M_N \mathbb{1}\{T_R > N\}] \leq (\alpha^R + \alpha^{-R})\mathbb{P}_x[T_R > N] \rightarrow 0, \text{ καθώς } N \rightarrow \infty.$$

Επιπλέον, από τον ορισμό του χρόνου διακοπής T_R , έχουμε ότι $M_{T_R} = (\alpha^R + \alpha^{-R})z^{T_R}$. Επομένως,

$$\begin{aligned} \mathbb{E}_x[M_{T_R} \mathbb{1}\{T_R \leq N\}] &= (\alpha^R + \alpha^{-R})\mathbb{E}_x[z^{T_R} \mathbb{1}\{T_R \leq N\}] = (\alpha^R + \alpha^{-R}) \sum_{k=0}^N z^k \mathbb{P}_x[T_R = k] \\ &\rightarrow (\alpha^R + \alpha^{-R}) \sum_{k=0}^{\infty} z^k \mathbb{P}_x[T_R = k] = (\alpha^R + \alpha^{-R})\mathbb{E}_x[z^{T_R}]. \end{aligned}$$

Καταλήγουμε έτσι στην απόδειξη του ισχυρισμού μας παίρνοντας το όριο καθώς $N \rightarrow \infty$ στη σχέση (4.13).

4.4 Martingales και μαρκοβιανές αλυσίδες

Στην παράγραφο αυτή θα δούμε πώς μπορούμε να υπολογίσουμε, με τη βοήθεια του Θεωρήματος επιλεκτικής διακοπής, πολλές ενδιαφέρουσες ποσότητες που εμφανίζονται όταν μελετάμε μαρκοβιανές αλυσίδες.

Ας ξεκινήσουμε με έναν απλό υπολογισμό. Έστω $\{X_n\}_{n \in \mathbb{N}_0}$ μια μαρκοβιανή αλυσίδα στον αριθμησιμο χώρο καταστάσεων \mathbb{X} . Θεωρούμε μια συνάρτηση $f : \mathbb{X} \rightarrow \mathbb{R}$, τέτοια ώστε $\mathbb{E}[|f(X_n)|] < \infty$ για κάθε $n \in \mathbb{N}_0$. Από τη μαρκοβιανή ιδιότητα έχουμε ότι $\mathbb{E}[f(X_{n+1}) | \mathcal{F}_n] = \mathbb{E}[f(X_{n+1}) | X_n]$. Αυτή η δεσμευμένη μέση τιμή είναι, όπως είδαμε, μια τυχαία μεταβλητή, η οποία είναι συνάρτηση της X_n , και η τιμή της στο ενδεχόμενο $\{X_n = x\}$ είναι

$$\mathbb{E}[f(X_{n+1}) | X_n = x] = \sum_{y \in \mathbb{X}} p(x, y) f(y).$$

Επομένως,

$$\mathbb{E}[f(X_{n+1}) | \mathcal{F}_n] = \sum_{y \in \mathbb{X}} p(X_n, y) f(y). \quad (4.14)$$

Θεώρημα 17 Αν $\mathbb{E}[|f(X_n)|] < \infty$ για κάθε $n \in \mathbb{N}_0$, τότε η στοχαστική διαδικασία $\{M_n^f\}_{n \in \mathbb{N}_0}$ με $M_0^f = f(X_0)$ και

$$M_n^f = f(X_n) - \sum_{k=0}^{n-1} Lf(X_k), \quad \text{για κάθε } n \in \mathbb{N}$$

είναι \mathcal{F}_n -martingale.

Απόδειξη: Παρατηρήστε αρχικά ότι

$$Lf(X_k) = \sum_{y \in \mathbb{X}} p(X_k, y)(f(y) - f(X_k)) = \sum_{y \in \mathbb{X}} p(X_k, y)f(y) - f(X_k) \sum_{y \in \mathbb{X}} p(X_k, y) = \mathbb{E}[f(X_{k+1}) | \mathcal{F}_k] - f(X_k).$$

Έχουμε επομένως

$$M_{n+1}^f - M_n^f = f(X_{n+1}) - f(X_n) - Lf(X_n) = f(X_{n+1}) - \mathbb{E}[f(X_{n+1}) | \mathcal{F}_n]$$

και

$$M_n^f = M_0^f + \sum_{k=1}^n (M_k^f - M_{k-1}^f) = f(X_0) + \sum_{k=1}^n f(X_k) - \mathbb{E}[f(X_k) | \mathcal{F}_{k-1}].$$

Από την τριγωνική ανισότητα παίρνουμε ότι $\mathbb{E}[|M_n^f|] < \infty$, ενώ

$$\mathbb{E}[M_{n+1}^f | \mathcal{F}_n] = M_n^f + \mathbb{E}[M_{n+1}^f - M_n^f | \mathcal{F}_n] = M_n^f. \quad \square$$

Χρησιμοποιώντας το Θεώρημα επιλεκτικής διακοπής, κάθε martingale που μπορούμε να γράψουμε μας δίνει μια πληροφορία για τη συμπεριφορά της στοχαστικής διαδικασίας που μελετάμε. Το Θεώρημα 17 μας προσφέρει έναν συστηματικό τρόπο για να κατασκευάζουμε martingale που σχετίζονται με μια μαρκοβιανή διαδικασία που θέλουμε να αναλύσουμε. Επιλέγοντας κατάλληλα τη συνάρτηση f , μπορούμε να κάνουμε χρήσιμους υπολογισμούς.

Αυτή η ιδέα ρίχνει καινούργιο φως στα αποτελέσματα του Κεφαλαίου 3. Για παράδειγμα, αν επιλέξουμε ως f τη λύση του ΠΣΤ 3.3

$$\begin{cases} Lf(x) = 0, & \text{αν } x \notin A \cup B \\ f(x) = 1, & \text{αν } x \in A \\ f(x) = 0, & \text{αν } x \in B \end{cases} \quad (4.15)$$

και εφαρμόσουμε το Θεώρημα επιλεκτικής διακοπής για τον χρόνο διακοπής $T_N = T_{A \cup B} \wedge N$, παίρνουμε ότι

$$f(x) = \mathbb{E}_x[M_{T_N}^f] = \mathbb{E}_x[f(X_{T_N}) - \sum_{k=0}^{T_N-1} Lf(X_k)] = \mathbb{E}_x[f(X_{T_N})], \quad \text{για κάθε } N \in \mathbb{N},$$

όπου στην τελευταία ισότητα χρησιμοποιήσαμε ότι πριν τον χρόνο T_N , η αλυσίδα οπωσδήποτε βρίσκεται εκτός των A και B , οπότε $Lf(X_k) = 0$ στο παραπάνω άθροισμα. Προσέξτε ότι η εξίσωση που η f ικανοποιεί εκτός των A και B , επιλέχθηκε ακριβώς ώστε να απλοποιήσει την παραπάνω έκφραση. Προσέξτε ακόμη ότι οι συνοριακές συνθήκες για την f έχουν έτσι επιλεγεί, ώστε περνώντας στο όριο (όταν αυτό είναι δυνατόν) το δεξί μέλος να μας δίνει την πιθανότητα $\mathbb{P}_x[T_A < T_B]$.

Ανάλογα μπορεί να σκεφτεί κανείς και για τη λύση f του ΠΣΤ 3.14

$$\begin{cases} Lf(x) = -1, & \text{αν } x \notin A \\ f(x) = 0, & \text{αν } x \in A. \end{cases} \quad (4.16)$$

Εδώ έχουμε επιλέξει την εξίσωση που ικανοποιεί η f εκτός του A , ώστε το άθροισμα

$$- \sum_{k=0}^{T_N-1} Lf(X_k) = T_N$$

να εμφανίζει την ποσότητα της οποίας την αναμενόμενη τιμή θέλουμε να βρούμε. Επιπλέον, οι μηδενικές συνοριακές συνθήκες εξασφαλίζουν ότι δεν θα έχουμε άλλους ανεπιθύμητους όρους.

Παράδειγμα 39 Στον απλό, συμμετρικό τυχαίο περίπατο θέλουμε να υπολογίσουμε τον αναμενόμενο αριθμό των επισκέψεων στο μηδέν πριν από τον χρόνο εξόδου $T_R = \inf\{k \geq 0 : |S_k| = R\}$, δηλαδή την

$$\mathbb{E}_x[V(T_R)] = \mathbb{E}_x\left[\sum_{k=0}^{T_R-1} \delta_0(X_k)\right].$$

Θα θέλαμε να χρησιμοποιήσουμε το Θεώρημα επιλεκτικής διακοπής για τη martingale $\{M_n^f\}_{n \in \mathbb{N}_0}$, όπου η f είναι μια συνάρτηση που ικανοποιεί το ΠΣΤ

$$\begin{cases} Lf(x) = -\delta_0(x), & \text{αν } |x| < R \\ f(x) = 0, & \text{αν } |x| > R. \end{cases}$$

Πράγματι, με αυτή την επιλογή έχουμε ότι

$$- \sum_{k=0}^{T_N-1} Lf(X_k) = \sum_{k=0}^{T_R-1} \delta_0(X_k)$$

Επιπλέον, λόγω των συνοριακών συνθηκών, θα έχουμε μηδενική συνεισφορά από όρους της μορφής $f(X_{T_R})$. Δεν είναι ιδιαίτερα δύσκολο να βρούμε μια συνάρτηση με αυτές τις ιδιότητες. Η συνθήκη

$$Lf(x) = 0 \Leftrightarrow \frac{f(x+1) + f(x-1)}{2} - f(x) = 0 \Leftrightarrow f(x+1) - f(x) = f(x) - f(x-1), \quad x = 1, 2, \dots, R-1,$$

επιβάλλει η f να είναι γραμμική για $x \geq 0$. Για να ικανοποιήσουμε ταυτόχρονα τη συνοριακή συνθήκη στο R , θα πρέπει επομένως να έχουμε $f(x) = A(R-x)$ για $x = 0, 1, \dots, R$. Ομοίως, θα πρέπει $f(x) = B(R+x)$, για $x = 0, -1, \dots, -R$. Θέτοντας $x = 0$ στους δύο κλάδους παίρνουμε ότι $f(0) = AR = BR$, επομένως έχουμε $A = B$ και

$$f(x) = A(R - |x|), \quad x \in \{-R, -R+1, \dots, R-1, R\}.$$

Θα υπολογίσουμε τώρα τη σταθερά A , ώστε $Lf(0) = -1$. Έχουμε όμως

$$Lf(0) = \frac{f(1) + f(-1)}{2} - f(0) = -A,$$

επομένως $f(x) = R - |x|$. Εφαρμόζουμε τώρα το Θεώρημα επιλεκτικής διακοπής για τη $\{M_n^f\}_n$ και τον χρόνο διακοπής $T_R \wedge N$.

$$\begin{aligned} R - |x| &= \mathbb{E}_x[M_{T_R \wedge N}^f] = \mathbb{E}_x[f(X_{T_R \wedge N})] + \mathbb{E}_x[V(T_R \wedge N)] \\ &= \mathbb{E}_x[(R - |X_N|) \mathbb{1}\{T_R > N\}] + \mathbb{E}_x\left[\sum_{k=0}^{T_R \wedge N - 1} \delta_0(X_k)\right]. \end{aligned}$$

Ο πρώτος όρος του δεξιού μέλους φράσσεται από την $R\mathbb{P}_x[T_R > N] \rightarrow R\mathbb{P}_x[T_R = \infty] = 0$, αφού η ακολουθία ενδεχομένων $C_N = \{T_R > N\}$ είναι φθίνουσα και $\bigcap_N C_N = \{T_R = \infty\}$. Για τον δεύτερο όρο, μπορούμε είτε να χρησιμοποιήσουμε το Θεώρημα μονότονης σύγκλισης (Πόρισμα 1.6 στο [8]), είτε να ελέγξουμε, με τη βοήθεια του Θεωρήματος Fubini-Tonelli, ότι

$$\mathbb{E}_x \left[\sum_{k=0}^{T_R \wedge N-1} \delta_0(X_k) \right] = \sum_{k=0}^{N-1} \mathbb{E}_x [\delta_0(X_k) \mathbb{1}\{T_R > k\}].$$

Έχουμε τότε ότι

$$\lim_{N \rightarrow \infty} \mathbb{E}_x \left[\sum_{k=0}^{T_R \wedge N-1} \delta_0(X_k) \right] = \sum_{k=0}^{\infty} \mathbb{E}_x [\delta_0(X_k) \mathbb{1}\{T_R > k\}] = \mathbb{E}_x \left[\sum_{k=0}^{T_R-1} \delta_0(X_k) \right].$$

Επομένως,

$$\mathbb{E}_x \left[\sum_{k=0}^{T_R-1} \delta_0(X_k) \right] = R - |x|.$$

Παράδειγμα 40 Μπορούμε να χρησιμοποιήσουμε τις παραπάνω ιδέες για να δείξουμε ότι ο απλός, συμμετρικός τυχαίος περίπατος στο \mathbb{Z}^d είναι παροδικός για $d > 2$. Θυμηθείτε ότι όταν $d = 1$ ή 2 , ο απλός, συμμετρικός, τυχαίος περίπατος είναι επαναληπτικός. Για $x = (x_1, \dots, x_d) \in \mathbb{Z}^d$ ορίζουμε

$$|x|^2 = \sum_{i=1}^d x_i^2$$

και

$$V(x) = |x|^{-\alpha}, \quad x \neq 0, \quad V(0) = 1.$$

Αν $0 < \alpha < d - 2$, μπορεί κανείς να ελέγξει ότι υπάρχει μια σταθερά $L = L(\alpha, d) > 0$, τέτοια ώστε $L|x|^{-\alpha} \leq 0$, όταν $|x| > L$. Αν ξεκινήσουμε τον περίπατο από κάποιο $x \in \mathbb{Z}^d$, με $|x| > L$ και ορίσουμε $T_L = \inf\{k \geq 0 : |X_k| \leq L\}$ να είναι ο χρόνος πρώτης άφιξης του περιπάτου στη σφαίρα με κέντρο το 0 και ακτίνα L , το Θεώρημα επιλεκτικής διακοπής δίνει ότι

$$|x|^{-\alpha} = V(x) = \mathbb{E}_x [V(X_{T_L \wedge N})] - \mathbb{E}_x \left[\sum_{k=0}^{T_L \wedge N-1} LV(X_k) \right] \geq \mathbb{E}_x [V(X_{T_L \wedge N})]$$

Επιπλέον,

$$\mathbb{E}_x [V(X_{T_L \wedge N})] = \mathbb{E}_x [V(X_{T_L}) \mathbb{1}\{T_L \leq N\}] + \mathbb{E}_x [V(X_N) \mathbb{1}\{T_L > N\}] \geq L^{-\alpha} \mathbb{P}_x [T_L \leq N].$$

Συνδυάζοντας τις δύο παραπάνω σχέσεις, παίρνουμε ότι

$$\mathbb{P}_x [T_L \leq N] \leq \left(\frac{L}{|x|} \right)^\alpha.$$

Εφόσον τώρα η ακολουθία ενδεχομένων $C_N = \{T_L \leq N\}$ είναι αύξουσα και $\bigcup_N C_N = \{T_L < \infty\}$, παίρνοντας το όριο καθώς $N \rightarrow \infty$ έχουμε ότι

$$\mathbb{P}_x [T_L < \infty] \leq \left(\frac{L}{|x|} \right)^\alpha < 1,$$

επομένως ο περίπατος είναι παροδικός.

Θεώρημα 18 Έστω $A \subset \mathbb{X}$ και $g : A \rightarrow \mathbb{R}$ μια φραγμένη συνάρτηση. Αν $\mathbb{P}_x[T_A < \infty] = 1$ για κάθε $x \in A^c$, τότε το ΠΣΤ

$$\begin{cases} Lf(x) = 0, & \text{αν } x \notin A \\ f(x) = g(x), & \text{αν } x \in A \end{cases} \quad (4.17)$$

έχει το πολύ μια φραγμένη λύση, που δίνεται από την

$$f(x) = \mathbb{E}_x[g(X_{T_A})].$$

Απόδειξη: Το ότι η $\mathbb{E}_x[g(X_{T_A})]$ λύνει το ΠΣΤ (4.17) αποδεικνύεται με ανάλυση πρώτου βήματος, με τρόπο εντελώς ανάλογο με το Θεώρημα 7 του Κεφαλαίου 3. Δείξτε το για εξάσκηση. Για τη μοναδικότητα, ας υποθέσουμε ότι η f είναι φραγμένη και ικανοποιεί το ΠΣΤ 4.17. Από το Θεώρημα επιλεκτικής διακοπής έχουμε ότι

$$\begin{aligned} f(x) &= \mathbb{E}_x[f(X_{T_A \wedge N})] - \mathbb{E}_x\left[\sum_{k=0}^{T_A \wedge N - 1} Lf(X_k)\right] = \mathbb{E}_x[f(X_{T_A \wedge N})] \\ &= \mathbb{E}_x[g(X_{T_A})\mathbb{1}\{T_A \leq N\}] + \mathbb{E}_x[f(X_N)\mathbb{1}\{T_A > N\}]. \end{aligned}$$

Από την υπόθεση έχουμε ότι $\mathbb{P}_x[T_A = \infty] = 0$. Αν λοιπόν η σταθερά M είναι ένα κοινό φράγμα των $|f|$ και $|g|$, τότε

$$\left| \mathbb{E}_x[g(X_{T_A})\mathbb{1}\{T_A > N\}] + \mathbb{E}_x[f(X_N)\mathbb{1}\{T_A > N\}] \right| \leq 2M\mathbb{P}_x[T_A > N] \rightarrow 0.$$

Επομένως, παίρνοντας $N \rightarrow \infty$ έχουμε ότι $f(x) = \mathbb{E}_x[g(X_{T_A})]$. □

Το παραπάνω Θεώρημα λειτουργεί προς δύο κατευθύνσεις. Αν μπορούμε να λύσουμε το ΠΣΤ 4.17, τότε έχουμε έναν τρόπο για να υπολογίσουμε την $\mathbb{E}_x[g(X_{T_A})]$. Αντίστροφα, αν θέλουμε να λύσουμε το ΠΣΤ 4.17, που συχνά είναι μια εξίσωση διαφορών σε πολλές διαστάσεις, μπορούμε να βρούμε τη λύση προσεγγιστικά, προσομοιώνοντας τροχιές μιας μαρκοβιανής αλυσίδας με γεννήτορα L και εκτιμώντας τη μέση τιμή $\mathbb{E}_x[g(X_{T_A})]$ με μεθόδους Monte Carlo.

4.5 Ασκήσεις

Άσκηση 62 Μπορούμε να φανταστούμε μια συνεχή συνάρτηση $X : [0, 1) \rightarrow \mathbb{R}$ ως μια τυχαία μεταβλητή ορισμένη στον δειγματικό χώρο $\Omega = [0, 1)$. Εφοδιάζουμε τον Ω με το σύννηθες μέτρο που δίνει στα διαστήματα $[a, b)$ πιθανότητα $b - a$. Ορίζουμε τώρα στον Ω την τυχαία μεταβλητή Y , με τιμές στο $\{0, 1, \dots, 9\}$ και τύπο $Y(x) = k$ αν $\frac{k}{10} \leq x < \frac{k+1}{10}$, $k = 0, 1, \dots, 9$. Υπολογίστε με τη βοήθεια του Θεωρήματος 11 την $\mathbb{E}[X | Y]$.

Άσκηση 63 Ρίχνουμε δύο ζάρια. Ανάλογα με το άθροισμα των ζαριών Y , επιλέγουμε τυχαία έναν αριθμό X από το σύνολο $\{1, 2, \dots, Y\}$. Υπολογίστε την $\mathbb{E}[X | Y]$ και επιβεβαιώστε την ιδιότητα 2 του Θεωρήματος 12.

Άσκηση 64 (ανισότητα Jensen) Αν η $f : \mathbb{R} \rightarrow \mathbb{R}$ είναι μια κυρτή συνάρτηση, η γραφική της παράσταση είναι πάνω από την εφαπτομένη της σε οποιοδήποτε σημείο $(x_0, f(x_0))$ του γραφήματος της f . Επομένως, για κάθε $x_0, x \in \mathbb{R}$,

$$f(x) \geq f(x_0) + f'_\pm(x_0)(x - x_0).$$

Με τη βοήθεια της παραπάνω ανισότητας δείξτε ότι, αν $\mathbb{E}[|X|] < \infty$ και Y είναι μια διακριτή τυχαία μεταβλητή, τότε

$$\mathbb{E}[f(X) | Y] \geq f(\mathbb{E}[X | Y]).$$

Ειδικές περιπτώσεις αυτής της ανισότητας είναι η ιδιότητα 5 του Θεωρήματος 12 και η (4.5), αφού οι $x \mapsto |x|$ και $x \mapsto x^2$ είναι κυρτές συναρτήσεις.

Άσκηση 65 Ας θεωρήσουμε τον δειγματικό χώρο $\Omega = [0, 1)$ και την τυχαία μεταβλητή $Z : \Omega \rightarrow \mathbb{R}$ με τύπο $Z(x) = x^2$, $x \in \Omega$. Ας θεωρήσουμε επίσης την ακολουθία τυχαίων μεταβλητών $\{X_n\}_{n \in \mathbb{N}_0}$, όπου

$$X_n(x) = \frac{k}{2^n}, \quad x \in \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right), \quad k = 0, 1, \dots, 2^n - 1.$$

Φτιάξτε τις γραφικές παραστάσεις των X_0, X_1, X_2, X_3 . Στη συνέχεια, με τη βοήθεια της Άσκησης 62, βρείτε τις $M_n = \mathbb{E}[Z | X_0, \dots, X_n]$ για $n = 1, 2, 3$ και ζωγραφίστε τις γραφικές παραστάσεις τους.

Άσκηση 66 Ορίζουμε τη στοχαστική διαδικασία $\{S_n\}_{n \in \mathbb{N}_0}$, με $S_0 = x > 0$ και $S_n = S_{n-1}X_n$, $n \in \mathbb{N}$, όπου η $\{X_n\}_{n \in \mathbb{N}}$ είναι μια ακολουθία από ανεξάρτητες τυχαίες μεταβλητές, με

$$\mathbb{P}[X_n = u] = q \in (0, 1), \quad \mathbb{P}[X_n = d] = 1 - q, \quad \text{για κάθε } n \in \mathbb{N},$$

όπου $u, d > 0$. Βρείτε την ικανή και αναγκαία συνθήκη που πρέπει να πληρούν οι παράμετροι u, d, q , ώστε η $\{S_n\}_{n \in \mathbb{N}_0}$ να είναι martingale. Είναι η στοχαστική διαδικασία $\{\log S_n\}_{n \in \mathbb{N}_0}$ martingale;

Άσκηση 67 Δίνεται ένας περίπατος $\{S_n\}_{n \in \mathbb{N}_0}$ στους ακεραίους, με $S_0 = x \in \mathbb{Z}$ και

$$S_n = x + \sum_{k=1}^n X_k, \quad n \in \mathbb{N},$$

όπου οι $\{X_k\}_{k \in \mathbb{N}}$ είναι ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με $\mathbb{E}[X_k] = \mu \neq 0$. Δείξτε ότι η διαδικασία $\{M_n\}_{n \in \mathbb{N}_0}$, με $M_n = S_n - \mu n$ είναι martingale.

Άσκηση 68 Ρίχνετε ένα ζάρι ώσπου το άθροισμα των ζαριών σας να φτάσει ή να ξεπεράσει τον αριθμό $\nu \in \mathbb{N}$. Αν $N \in \mathbb{N}$ είναι ο αριθμός των ζαριών που θα χρειαστεί να ρίξετε ώσπου να συμβεί αυτό, δείξτε ότι

$$\mathbb{E}[N] > \frac{2\nu}{7}.$$

Άσκηση 69 Κατασκευάζουμε ένα γράφο με την ακόλουθη διαδικασία. Στο i -οστό βήμα ($i \in \mathbb{N}$) προσθέτουμε την κορυφή v_i και τοποθετούμε τυχαία μια ακμή που συνδέει τη v_i με μία από τις κορυφές $\{v_1, \dots, v_i\}$ (επιτρέπουμε επομένως ένα βρόχο από τη v_i προς τον εαυτό της με πιθανότητα $\frac{1}{i}$). Έστω $\{X_n\}_{n \in \mathbb{N}}$ το μέγεθος της συνεκτικής συνιστώσας του γράφου που περιέχει τη v_1 μετά το n -οστό βήμα.

α) Δείξτε ότι η διαδικασία $\{M_n\}_{n \in \mathbb{N}}$ με $M_n = \frac{X_n}{n+1}$ είναι martingale.

β) Δείξτε γενικότερα ότι, για κάθε $p \in \mathbb{N}$, η διαδικασία $\{J_n^{(p)}\}_{n \in \mathbb{N}}$ με $J_n^{(p)} = \prod_{k=0}^{p-1} \frac{X_{n+k}}{n+k+1}$ είναι martingale.

γ) Υπολογίστε τη μέση τιμή και τη διασπορά της τυχαίας μεταβλητής X_n .

δ) Αν $\delta \in (0, 1)$, ορίζουμε $T_\delta = \inf\{n \in \mathbb{N} : X_n < \delta(n+1)\}$. Δείξτε ότι

$$\mathbb{P}[T_\delta < +\infty] \leq \frac{1}{2(1-\delta)}.$$

Άσκηση 70 Για τον απλό, συμμετρικό τυχαίο περίπατο στους ακεραίους $\{S_n\}_{n \in \mathbb{N}_0}$, υπολογίστε τη σταθερά α ώστε η διαδικασία $\{M_n\}_{n \in \mathbb{N}_0}$ με

$$M_n = S_n^3 - \alpha n S_n$$

να είναι martingale.

Άσκηση 71 N άτομα παρακολουθούν μια θεατρική παράσταση και έχουν αφήσει τα παλτά τους στη γκαρνταρόμπα. Στο τέλος της παράστασης τα παλτά θα μοιραστούν τυχαία στους θεατές. Όσοι τύχει να πάρουν το δικό τους παλτό τους θα φύγουν. Οι υπόλοιποι θα επιστρέψουν το λάθος παλτό στη γκαρνταρόμπα. Η διαδικασία θα επαναλαμβάνεται μέχρι να έχουν όλοι πάρει πίσω το παλτό τους. Αν X_n είναι το πλήθος των παλτών στη γκαρνταρόμπα μετά από n τυχαία μοιράσματα και $T = \inf\{k \geq 0 : X_k = 0\}$ είναι το πλήθος των μοιρασμάτων που θα χρειαστούν, δείξτε ότι η διαδικασία $\{M_n\}_{n \in \mathbb{N}_0}$ με $M_n = X_{n \wedge T} + n \wedge T$ είναι martingale. Ποια είναι η $\mathbb{E}[T]$;

Άσκηση 72 Έστω $z \in (0, 1)$. Για έναν απλό τυχαίο περίπατο στους ακεραίους, με $p(x, x+1) = p \in (0, 1)$ και $p(x, x-1) = 1-p$ για κάθε $x \in \mathbb{Z}$, δείξτε ότι η στοχαστική διαδικασία $\{M_n\}_{n \in \mathbb{N}_0}$ με

$$M_n = \alpha^{S_n} z^n, \quad n \in \mathbb{N}_0,$$

είναι martingale, αν και μόνο αν

$$\alpha = \frac{1 \pm \sqrt{1 - 4p(1-p)z^2}}{2pz}.$$

Στη συνέχεια, χρησιμοποιήστε το Θεώρημα επιλεκτικής διακοπής, για να δείξετε ότι, αν $S_0 = x \in \mathbb{N}$, η γεννήτρια συνάρτηση του χρόνου πρώτης άφιξης στο 0, $T_0 = \inf\{k \geq 0 : S_k = 0\}$ είναι η

$$\psi(z) = \mathbb{E}_x[z^{T_0}] = \left(\frac{1 - \sqrt{1 - 4p(1-p)z^2}}{2pz} \right)^x, \quad z \in (0, 1).$$

Η γεννήτρια συνάρτηση έχει όλη την πληροφορία για την κατανομή του χρόνου T_0 . Π.χ. έχουμε

$$\mathbb{P}_x[T_0 < \infty] = \lim_{z \uparrow 1} \psi(z) = \left(\frac{1 - |1 - 2p|}{2p} \right)^x = \begin{cases} 1 & , p \leq \frac{1}{2} \\ \left(\frac{1-p}{p}\right)^x & , p > \frac{1}{2}. \end{cases}$$

Υπολογίστε με τη βοήθεια της $\psi(z)$ και των παραγώγων της, τη μέση τιμή και τη διασπορά του T_0 .

Άσκηση 73 Μπορούμε να εμπλουτίσουμε τη γκάμα των martingale που σχετίζονται με μια μαρκοβιανή αλυσίδα $\{X_n\}_{n \in \mathbb{N}_0}$ θεωρώντας συναρτήσεις που εξαρτώνται και από τον χρόνο. Δείξτε κατ' αναλογία με το Θεώρημα 17 ότι, αν $\mathbb{E}[|f(n, X_n)|] < \infty$ για κάθε $n \in \mathbb{N}_0$, τότε η στοχαστική διαδικασία $\{M_n^f\}_{n \in \mathbb{N}_0}$, με $M_0^f = f(0, X_0)$ και

$$M_n^f = f(n, X_n) - \sum_{k=0}^{n-1} (f(n+1, X_k) - f(n, X_k) + Lf(X_k)), \quad \text{για } n \in \mathbb{N}$$

είναι \mathcal{F}_n -martingale.

Άσκηση 74 Για να βρούμε τη γεννήτρια συνάρτηση ενός χρόνου άφιξης $T_A = \inf\{k \geq 0 : X_k \in A\}$

$$\psi(z) = \mathbb{E}[z^{T_A}], \quad z \in (0, 1),$$

μπορούμε να θεωρήσουμε μια συνάρτηση $f : \mathbb{N}_0 \times X \rightarrow \mathbb{R}$ με τύπο $f(n, x) = z^n h(x)$. Ποιο ΠΣΤ πρέπει να ικανοποιεί η συνάρτηση $h : X \rightarrow \mathbb{R}$, ώστε εφαρμόζοντας το Θεώρημα επιλεκτικής διακοπής για τη martingale της προηγούμενης άσκησης να βρίσκουμε την $\psi(z)$; Μπορείτε τώρα να βρείτε μόνοι σας γιατί επιλέξαμε τη συγκεκριμένη martingale στο Παράδειγμα 38;

Άσκηση 75 Θεωρήστε τον απλό, συμμετρικό τυχαίο περίπατο στους ακεραίους και τη συνάρτηση $f : \mathbb{Z} \rightarrow \mathbb{Z}$, $f(k) = k^4$. Δείξτε ότι $Lf(k) = 6k^2 + 1$ και συμπεράνετε ότι η διαδικασία $\{M_n\}_{n \in \mathbb{N}_0}$, με $M_0 = 0$ και

$$M_n = X_n^4 - X_n^2 - 6 \sum_{k=0}^{n-1} X_k^2, \quad \text{για } n \in \mathbb{N},$$

είναι martingale. Αν $X_0 = 0$ και $T_R = \inf\{k \geq 0 : |X_k| = R\}$ είναι ο χρόνος πρώτης εξόδου από το $\{-R+1, \dots, R-1\}$, υπολογίστε την

$$\mathbb{E} \left[\sum_{k=1}^{T_N} X_k^2 \right].$$

Άσκηση 76 Παίζουμε ένα ευνοϊκό παιχνίδι στο οποίο σε κάθε γύρο με πιθανότητα $p \in (\frac{1}{2}, 1)$ διπλασιάζουμε το στοιχείμά μας και με πιθανότητα $1-p$ χάνουμε το ποσόν που στοιχηματήσαμε. Ξεκινάμε με αρχική περιουσία $S_0 > 0$ και αμέσως πριν τον n -οστό γύρο μπορούμε να στοιχηματήσουμε οποιοδήποτε ποσόν C_n μεταξύ του 0 και της περιουσίας μας μέχρι εκείνη τη στιγμή S_{n-1} . Επομένως, αμέσως μετά n -οστό γύρο η περιουσία μας θα είναι

$$S_n = S_{n-1} + C_n X_n,$$

όπου οι $\{X_k\}_{k \in \mathbb{N}}$ είναι ανεξάρτητες, ισόνομες τυχαίες μεταβλητές, που παίρνουν την τιμή $+1$ με πιθανότητα p και την τιμή -1 με πιθανότητα $1-p$. Η C_n πρέπει να ικανοποιεί την ανισότητα $0 \leq C_n \leq S_{n-1}$ και μπορεί να εξαρτάται από τα αποτελέσματα των προηγούμενων γύρων, θα πρέπει όμως οπωσδήποτε να είναι \mathcal{F}_{n-1} -μετρήσιμη. Ο σκοπός μας είναι να βρούμε μια στρατηγική που μεγιστοποιεί τον αναμενόμενο ρυθμό μεγέθυνσης

$$r_n = \frac{1}{n} \mathbb{E} [\log(S_n/S_0)].$$

Δείξτε ότι, ανεξαρτήτως της στρατηγικής βάσει της οποίας επιλέγουμε τα στοιχήματά μας, η στοχαστική διαδικασία $\{M_n\}_{n \in \mathbb{N}_0}$ με τύπο

$$M_n = \log(S_n) - nH(p), \quad \text{όπου } H(p) = p \log p + (1-p) \log(1-p) + \log 2$$

ικανοποιεί την ανισότητα

$$\mathbb{E}[M_{n+1} | \mathcal{F}_n] \leq M_n$$

και συμπεράνετε ότι $r_n \leq H(p)$ για κάθε $n \in \mathbb{N}$. Βρείτε τώρα μια στρατηγική (έναν τρόπο να επιλέγετε τα στοιχήματά σας $\{C_n\}_{n \in \mathbb{N}}$), τέτοια ώστε η $\{M_n\}_{n \in \mathbb{N}_0}$ να είναι martingale και δείξτε ότι με αυτή τη στρατηγική έχουμε

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \left(\frac{S_N}{S_0} \right) = H(p),$$

με πιθανότητα 1.

4.6 Αριθμητικά πειράματα

Άσκηση 77 Θεωρούμε το δισδιάστατο πλέγμα $h\mathbb{Z}^2 = \{hz : z \in \mathbb{Z}^2\}$. Αν $f : h\mathbb{Z}^2 \rightarrow \mathbb{R}$, η διακριτή Λαπλασιανή της f δίνεται από τον τύπο

$$\Delta_h f(x, y) = \frac{f(x+h, y) + f(x-h, y) + f(x, y+h) + f(x, y-h) - 4f(x, y)}{h^2}.$$

Θέλουμε να υπολογίσουμε τη λύση του προβλήματος συνοριακών τιμών

$$\begin{cases} \Delta_h f(x, y) = 0, & \text{αν } |x|, |y| < 1 \\ f(x, y) = |\sin(\pi x) - \sin(\pi y)|, & \text{αν } |x| = 1, \text{ ή } |y| = 1, \end{cases}$$

για $h = 0.01$, σύμφωνα με το αποτέλεσμα του Θεωρήματος 4.17. Η ρουτίνα `srw2d.py` προσομοιώνει $N = 10.000$ βήματα ενός απλού, συμμετρικού τυχαίου περιπάτου στο \mathbb{Z}^2 και υπολογίζει την αναμενόμενη τιμή $\mathbb{E}[|X_N|]$ με τη μέθοδο Monte Carlo. Προσομοιώνοντας μονοπάτια ενός απλού, συμμετρικού τυχαίου περιπάτου στο \mathbb{Z}^2 , εκτιμήστε την τιμή της λύσης του παραπάνω ΠΣΤ στο σημείο $(1/2, 1/4)$. Θα χρειαστεί να αλλάξετε τον κώδικα ώστε η προσομοίωση να σταματάει όχι τη συγκεκριμένη χρονική στιγμή N , αλλά στον χρόνο άφιξης του περιπάτου στο σύνορο.

Άσκηση 78 Σκοπός αυτής της άσκησης είναι να οπτικοποιήσουμε την κίνηση μιας μαρκοβιανής αλυσίδας. Κατεβάστε τον οδηγό και ακολουθήστε τα βήματα που περιγράφονται εκεί. Δείξτε τα διαδοχικά στιγμιότυπα ενός γκέιμ τένις, τρέχοντας τον κώδικα `tennis.py` που λύνει την Άσκηση 40.

Κεφάλαιο 5

Αναλλοίωτες κατανομές

5.1 Εισαγωγή

Οι αναλλοίωτες κατανομές είναι κατά κάποιο τρόπο οι φυσικές καταστάσεις μιας μαρκοβιανής αλυσίδας. Αν μια αλυσίδα ξεκινήσει από μια αναλλοίωτη κατανομή της, θα παραμείνει σε αυτήν για πάντα, ενώ η ασυμπτωτική συμπεριφορά μιας αλυσίδας μπορεί να χαρακτηριστεί μέσω αυτών. Επίσης, οι απαντήσεις σε πολλά προβλήματα μπορούν εύκολα να εκφραστούν μέσω των αναλλοίωτων κατανομών. Σε αυτό το κεφάλαιο θα δούμε ικανές και αναγκαίες συνθήκες για την ύπαρξη αναλλοίωτων κατανομών, θα μελετήσουμε τις ιδιότητες και τη δομή τους. Παρόμοιο υλικό μπορείτε να βρείτε στις αναφορές [7] και [6].

5.2 Αναλλοίωτες κατανομές

Θα συμβολίζουμε το σύνολο των κατανομών στον \mathbb{X} με $\mathcal{M}(\mathbb{X})$. Έτσι

$$\pi \in \mathcal{M}(\mathbb{X}) \Leftrightarrow \eta \pi \text{ είναι μια συνάρτηση } \pi : \mathbb{X} \rightarrow [0, 1] \text{ με } \sum_{x \in \mathbb{X}} \pi(x) = 1.$$

Έστω $\{X_n\}_{n \in \mathbb{N}}$ μια μαρκοβιανή αλυσίδα με αρχική κατανομή π_0 και πιθανότητες μετάβασης $\{p(x, y)\}_{x, y \in \mathbb{X}}$. Θα συμβολίζουμε την κατανομή της τυχαίας μεταβλητής X_n με π_n , δηλαδή

$$\pi_n(x) = \mathbb{P}[X_n = x], \quad \text{για } x \in \mathbb{X}. \quad (5.1)$$

Το κεντρικό ερώτημα αυτού του κεφαλαίου είναι αν η π_n συγκλίνει καθώς $n \rightarrow \infty$ και, αν ναι, ποιο είναι το όριο της π .

Ορισμός Αν $\{\pi_n\}_n$ είναι μια ακολουθία κατανομών στο $\mathcal{M}(\mathbb{X})$, θα λέμε ότι συγκλίνει στην κατανομή $\pi \in \mathcal{M}(\mathbb{X})$ αν και μόνο η πιθανότητα που αποδίδει η π_n σε κάθε κατάσταση συγκλίνει στην αντίστοιχη πιθανότητα που αποδίδει η π . Δηλαδή,

$$\pi_n \rightarrow \pi \in \mathcal{M}(\mathbb{X}) \Leftrightarrow \pi_n(x) \rightarrow \pi(x), \quad \text{για κάθε } x \in \mathbb{X}. \quad (5.2)$$

Παρατήρηση: Από ένα αποτέλεσμα της θεωρίας μέτρου, το λήμμα του Scheffé (βλ. [9], παρ. 5.10), ο παραπάνω ορισμός είναι ισοδύναμος με το φαινομενικά ισχυρότερο αποτέλεσμα

$$\pi_n \rightarrow \pi \in \mathcal{M}(\mathbb{X}) \Leftrightarrow \sum_{x \in \mathbb{X}} |\pi_n(x) - \pi(x)| \rightarrow 0. \quad (5.3)$$

Ορισμός: Αν η $\{\pi_n\}_n$ είναι η ακολουθία των κατανομών μιας μαρκοβιανής αλυσίδας $\{X_n\}$, αν δηλαδή η π_n ορίζεται όπως στην (5.1) για κάθε $n \in \mathbb{N}_0$ και $\pi_n \rightarrow \pi \in \mathcal{M}(\mathbb{X})$, θα λέμε ότι η π είναι η κατανομή

ισορροπίας (ή εναλλακτικά η ασυμπτωτική κατανομή) της $\{X_n\}$.

Το πρώτο μας αποτέλεσμα χαρακτηρίζει τις υποψήφιες κατανομές ισορροπίας μαρκοβιανών αλυσίδων.

Θεώρημα 19 Αν η ακολουθία $\{\pi_n\}_n$ των κατανομών μιας μαρκοβιανής αλυσίδας $\{X_n\}_{n \in \mathbb{N}}$ με πίνακα πιθανοτήτων μετάβασης P συγκλίνει στην κατανομή $\pi \in \mathcal{M}(\mathbb{X})$, τότε

$$\pi = \pi P, \quad \text{δηλαδή} \quad \pi(x) = \sum_{y \in \mathbb{X}} \pi(y)p(y, x), \quad \text{για κάθε } x \in \mathbb{X}. \quad (5.4)$$

Απόδειξη: Είδαμε στο δεύτερο κεφάλαιο ότι η κατανομή π_n μιας μαρκοβιανής αλυσίδας $\{X_n\}_n$ μετά από n βήματα, δηλαδή η κατανομή της τυχαίας μεταβλητής X_n , δίνεται από την

$$\pi_n = \pi_0 P^n.$$

Επομένως, για κάθε $n \in \mathbb{N}$ έχουμε

$$\pi_n = \pi_0 (P^{n-1} P) = (\pi_0 P^{n-1}) P = \pi_{n-1} P,$$

δηλαδή

$$\pi_n(x) = \sum_{y \in \mathbb{X}} \pi_{n-1}(y)p(y, x) \quad \text{για κάθε } x \in \mathbb{X}.$$

Παίρνοντας το όριο καθώς $n \rightarrow \infty$ στα δύο μέλη της προηγούμενης σχέσης, το αριστερό μέλος συγκλίνει στην $\pi(x)$, ενώ για το δεξί μέλος έχουμε

$$\left| \sum_{y \in \mathbb{X}} \pi_{n-1}(y)p(y, x) - \sum_{y \in \mathbb{X}} \pi(y)p(y, x) \right| \leq \sum_{y \in \mathbb{X}} |\pi_{n-1}(x) - \pi(x)| p(y, x) \leq \sum_{y \in \mathbb{X}} |\pi_{n-1}(x) - \pi(x)| \rightarrow 0$$

από την (5.3). Επομένως, $\pi(x) = \sum_{y \in \mathbb{X}} \pi(y)p(y, x)$, για κάθε $x \in X$. □

Σύμφωνα με το προηγούμενο Θεώρημα τα μόνα υποψήφια όρια των κατανομών μιας μαρκοβιανής αλυσίδας με πίνακα πιθανοτήτων μετάβασης P είναι εκείνες οι $\pi : \mathbb{X} \rightarrow [0, 1]$ για τις οποίες

$$\begin{cases} \pi = \pi P \\ \sum_{x \in \mathbb{X}} \pi(x) = 1. \end{cases} \quad (5.5)$$

Ορισμός: Αν μια κατανομή $\pi \in \mathcal{M}(\mathbb{X})$ ικανοποιεί την (5.4), θα λέμε ότι είναι *αναλλοίωτη* (ή *στάσιμη*) κατανομή για την αλυσίδα $\{X_n\}_n$ με πίνακα πιθανοτήτων μετάβασης P .

Μπορούμε να το αναδιατυπώσουμε το Θεώρημα 19 ως εξής. *Οι μόνες δυνατές κατανομές ισορροπίας μιας μαρκοβιανής αλυσίδας είναι οι αναλλοίωτες κατανομές της.* Η ορολογία αναλλοίωτη και στάσιμη εξηγείται από το παρακάτω Θεώρημα.

Θεώρημα 20 Αν η αρχική κατανομή $\pi_0 \in \mathcal{M}(\mathbb{X})$ μιας αλυσίδας στον \mathbb{X} με πίνακα πιθανοτήτων μετάβασης P ικανοποιεί την (5.4) τότε

$$\pi_n = \pi_0, \quad \text{για κάθε } n \in \mathbb{N}.$$

Επιπλέον, η $\{X_n\}_{n \in \mathbb{N}_0}$ είναι στάσιμη, δηλαδή για κάθε $n, k \in \mathbb{N}_0$ και κάθε $x_0, x_1, \dots, x_k \in \mathbb{X}$ έχουμε

$$\mathbb{P}[X_n = x_0, X_{n+1} = x_1, \dots, X_{n+k} = x_k] = \mathbb{P}[X_0 = x_0, X_1 = x_1, \dots, X_k = x_k].$$

Απόδειξη: Θα δείξουμε τον ισχυρισμό $\pi_n = \pi_0$ για κάθε $n \in \mathbb{N}$ επαγωγικά. Για $n = 1$ έχουμε

$$\pi_1 = \pi_0 P = \pi_0,$$

μια και για την π_0 έχουμε υποθέσει ότι ικανοποιεί την (5.4). Έστω τώρα ότι για κάποιο $n \in \mathbb{N}$ έχουμε $\pi_n = \pi_0$. Τότε

$$\pi_{n+1} = \pi_n P = \pi_0 P = \pi_0.$$

Επομένως, $\pi_n = \pi_0$ για κάθε $n \in \mathbb{N}$. Για τον δεύτερο ισχυρισμό έχουμε, από τη μαρκοβιανή ιδιότητα,

$$\begin{aligned} \mathbb{P}[X_n = x_0, \dots, X_{n+k} = x_k] &= \mathbb{P}[X_n = x_0] \mathbb{P}[X_{n+1} = x_1 | X_n = x_0] \cdots \mathbb{P}[X_{n+k} = x_k | X_{n+k-1} = x_{k-1}] \\ &= \pi_n(x_0) p(x_0, x_1) \cdots p(x_{k-1}, x_k) \\ &= \pi_0(x_0) p(x_0, x_1) \cdots p(x_{k-1}, x_k) \\ &= \mathbb{P}[X_0 = x_0, X_1 = x_1, \dots, X_k = x_k]. \end{aligned}$$

□

Παράδειγμα 41 Έστω $\{X_n\}_n$ μια μαρκοβιανή αλυσίδα σε έναν χώρο με δύο καταστάσεις $\mathbb{X} = \{\alpha, \beta\}$ και πίνακα πιθανοτήτων μετάβασης

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

με $p, q \in (0, 1)$. Από την (5.5) προκειμένου η $\pi = (\pi_\alpha, \pi_\beta)$ να είναι αναλλοίωτη κατανομή της αλυσίδας θα πρέπει να ικανοποιεί τις

$$\begin{cases} \pi_\alpha = (1-p)\pi_\alpha + q\pi_\beta \\ \pi_\beta = p\pi_\alpha + (1-q)\pi_\beta \\ \pi_\alpha + \pi_\beta = 1. \end{cases}$$

Λύνοντας το παραπάνω σύστημα βρίσκουμε ότι η

$$\pi = \left(\frac{q}{p+q}, \frac{p}{p+q} \right)$$

είναι αναλλοίωτη κατανομή της αλυσίδας. Επομένως, αν η αρχική κατανομή της αλυσίδας είναι η π , αν δηλαδή

$$\mathbb{P}[X_0 = \alpha] = \frac{q}{p+q} \quad \text{και} \quad \mathbb{P}[X_0 = \beta] = \frac{p}{p+q},$$

τότε $\pi_n = \pi$ για κάθε $n \in \mathbb{N}$. Προσέξτε ότι σε κάθε της βήμα η αλυσίδα αλλάζει κατάσταση με πιθανότητα p αν βρίσκεται στην κατάσταση α ή q αν βρίσκεται στην κατάσταση β . Αυτό που δεν αλλάζει είναι η κατανομή της X_n , δηλαδή η πιθανότητα να βρούμε την αλυσίδα σε καθεμία από τις δύο καταστάσεις.

5.3 Η δομή του $\mathcal{I}(P)$

Στη συνέχεια θα εξετάσουμε αν μπορούμε να βρούμε αναλλοίωτες κατανομές για μια αλυσίδα και, στην περίπτωση που μπορούμε, ποιες είναι αυτές. Θα συμβολίζουμε με $\mathcal{I}(P)$ το σύνολο των αναλλοίωτων κατανομών μιας αλυσίδας με πίνακα πιθανοτήτων μετάβασης P , δηλαδή

$$\mathcal{I}(P) = \{\pi \in \mathcal{M}(\mathbb{X}) : \pi = \pi P\}.$$

Το επόμενο λήμμα θα μας φανεί πολύ χρήσιμο.

Λήμμα 6 Αν $\{X_n\}_{n \in \mathbb{N}_0}$ είναι μια μαρκοβιανή αλυσίδα με πίνακα πιθανοτήτων μετάβασης P και $\pi \in \mathcal{I}(P)$, τότε για κάθε $x, y \in \mathbb{X}$ έχουμε

$$\pi(y) \geq \pi(x) \mathbb{E}_x \left[\sum_{k=1}^{T_x^+} \mathbb{1}\{X_k = y\} \right]. \quad (5.6)$$

Απόδειξη: Έστω $\pi \in \mathcal{I}(P)$. Για κάθε $x, y \in \mathbb{X}$ έχουμε

$$\begin{aligned} \pi(y) &= \sum_{z \in \mathbb{X}} \pi(z) p(z, y) = \pi(x) p(x, y) + \sum_{z \neq x} \pi(z) p(z, y) \\ &= \pi(x) p(x, y) + \sum_{z \neq x} \left(\sum_{u \in \mathbb{X}} \pi(u) p(u, z) \right) p(z, y) \\ &= \pi(x) \left(p(x, y) + \sum_{z \neq x} p(x, z) p(z, y) \right) + \sum_{z, u \neq x} \pi(u) p(u, z) p(z, y). \end{aligned}$$

Ας επαναλάβουμε την παραπάνω διαδικασία ακόμα μια φορά γράφοντας

$$\pi(u) = \sum_{w \in \mathbb{X}} \pi(w) p(w, u) = \pi(x) p(x, u) + \sum_{w \neq x} \pi(w) p(w, u).$$

Παίρνουμε τότε ότι

$$\begin{aligned} \pi(y) &= \pi(x) \left(p(x, y) + \sum_{z \neq x} p(x, z) p(z, y) + \sum_{z, u \neq x} p(x, u) p(u, z) p(z, y) \right) \\ &\quad + \sum_{z, u, w \neq x} \pi(w) p(w, u) p(u, z) p(z, y). \end{aligned} \quad (5.7)$$

Προσέξτε ότι στον όρο $\sum_{z, u \neq x} p(x, u) p(u, z) p(z, y)$ αθροίζουμε τις πιθανότητες όλων των μονοπατιών μήκους 3 που ξεκινούν από το x και καταλήγουν στο y χωρίς ενδιάμεσα να έχουν ξαναπεράσει από το x . Άρα

$$\sum_{z, u \neq x} p(x, u) p(u, z) p(z, y) = \mathbb{P}_x[X_1 \neq x, X_2 \neq x, X_3 = y] = \mathbb{P}_x[X_3 = y, T_x^+ \geq 3].$$

Αντίστοιχα έχουμε

$$\sum_{z \neq x} p(x, z) p(z, y) = \mathbb{P}_x[X_1 \neq x, X_2 = y] = \mathbb{P}_x[X_2 = y, T_x^+ \geq 2]$$

ενώ

$$p(x, y) = \mathbb{P}_x[X_1 = y] = \mathbb{P}_x[X_1 = y, T_x^+ \geq 1].$$

Με αυτή την παρατήρηση μπορούμε να ξαναγράψουμε την (5.7) ως

$$\pi(y) = \pi(x) \left(\sum_{k=1}^3 \mathbb{P}_x[X_k = y, T_x^+ \geq k] \right) + \sum_{z, u, w \neq x} \pi(w) p(w, u) p(u, z) p(z, y).$$

Θα πρέπει να είναι τώρα φανερό ότι, αν επαναλάβουμε αυτήν τη διαδικασία n φορές, έχουμε ότι

$$\begin{aligned} \pi(y) &= \pi(x) \left(\sum_{k=1}^n \mathbb{P}_x[X_k = y, T_x^+ \geq k] \right) + \sum_{x_1, \dots, x_n \neq x} \pi(x_1) p(x_1, x_2) \cdots p(x_n, y) \\ &\geq \pi(x) \left(\sum_{k=1}^n \mathbb{P}_x[X_k = y, T_x^+ \geq k] \right), \quad \text{για κάθε } n \in \mathbb{N}. \end{aligned} \quad (5.8)$$

Περνώντας στο όριο $n \rightarrow \infty$ έχουμε λοιπόν

$$\begin{aligned}\pi(y) &\geq \pi(x) \left(\sum_{k=1}^{\infty} \mathbb{P}_x [X_k = y, T_x^+ \geq k] \right) \\ &= \pi(x) \left(\sum_{k=1}^{\infty} \mathbb{E}_x [\mathbb{1}\{X_k = y, T_x^+ \geq k\}] \right).\end{aligned}$$

Εφόσον οι δείκτριες συναρτήσεις που εμφανίζονται παραπάνω είναι μη αρνητικές, από το Θεώρημα Fubini-Tonelli μπορούμε να εναλλάξουμε την άθροιση ως προς k και την αναμενόμενη τιμή, παίρνοντας

$$\begin{aligned}\pi(y) &\geq \pi(x) \mathbb{E}_x \left[\sum_{k=1}^{\infty} \mathbb{1}\{X_k = y, T_x^+ \geq k\} \right] \\ &= \pi(x) \mathbb{E}_x \left[\sum_{k=1}^{T_x^+} \mathbb{1}\{X_k = y\} \right].\end{aligned}$$

□

Ορισμός: Θα λέμε μια κατάσταση $x \in \mathbb{X}$ *γνησίως επαναληπτική* αν $\mathbb{E}_x [T_x^+] < +\infty$.

Η γνήσια επαναληπτικότητα είναι μια έννοια ισχυρότερη της επαναληπτικότητας, αφού αν $\mathbb{E}_x [T_x^+] < +\infty$, τότε αναγκαστικά έχουμε $\mathbb{P}_x [T_x^+ < \infty] = 1$ και άρα η x είναι επαναληπτική. Το ακόλουθο Πρόρισμα μας εγγυάται ότι οποιαδήποτε αναλλοίωτη κατανομή μιας αλυσίδας στηρίζεται σε γνησίως επαναληπτικές καταστάσεις.

Πόρισμα 5 Αν $\pi \in \mathcal{I}(P)$ και $x \in \mathbb{X}$ τότε

$$\mathbb{E}_x [T_x^+] = +\infty \Rightarrow \pi(x) = 0.$$

Απόδειξη: Αθροίζοντας τις ανισότητες (5.6) για όλα τα $y \in \mathbb{X}$ έχουμε

$$1 = \sum_{y \in \mathbb{X}} \pi(y) \geq \pi(x) \sum_{y \in \mathbb{X}} \mathbb{E}_x \left[\sum_{k=1}^{T_x^+} \mathbb{1}\{X_k = y\} \right] = \pi(x) \mathbb{E}_x \left[\sum_{k=1}^{T_x^+} \sum_{y \in \mathbb{X}} \mathbb{1}\{X_k = y\} \right],$$

όπου χρησιμοποιήσαμε το Θεώρημα Fubini-Tonelli για να εναλλάξουμε τη σειρά των αθροίσεων και της αναμενόμενης τιμής. Προσέξτε όμως ότι στο άθροισμα ως προς y υπάρχει ακριβώς ένας όρος $\mathbb{1}\{X_k = y\}$ που είναι ίσος με 1 (αυτός που αντιστοιχεί στην κατάσταση της X_k), ενώ όλοι οι υπόλοιποι είναι μηδέν. Επομένως, η παραπάνω ανισότητα γίνεται

$$1 \geq \pi(x) \mathbb{E}_x \left[\sum_{k=1}^{T_x^+} 1 \right] = \pi(x) \mathbb{E}_x [T_x^+],$$

απ' όπου έπεται ο ισχυρισμός μας. □

Πόρισμα 6 Αν μια μαρκοβιανή αλυσίδα έχει αναλλοίωτη κατανομή, τότε έχει τουλάχιστον μία γνησίως επαναληπτική κατάσταση.

Απόδειξη: Αν $\pi \in \mathcal{I}(P)$, τότε $\pi(y) \geq 0$ για κάθε $y \in \mathbb{X}$ και $\sum_{y \in \mathbb{X}} \pi(y) = 1$. Θα υπάρχει επομένως κάποια κατάσταση x για την οποία $\pi(x) > 0$ και άρα από το Πόρισμα 5 θα πρέπει $\mathbb{E}_x[T_x^+] < \infty$. \square

Στο επόμενο θεώρημα θα δείξουμε ότι και το αντίστροφο του προηγούμενου πορίσματος είναι σωστό. Δηλαδή, αν μια αλυσίδα έχει μια γνησίως επαναληπτική κατάσταση x τότε έχει και αναλλοίωτη κατανομή, κατασκευάζοντας μια εκπεφρασμένα. Η ιδέα προέρχεται από το Λήμμα 6.

Θεώρημα 21 Έστω $x \in \mathbb{X}$ μια γνησίως επαναληπτική κατάσταση. Ορίζουμε

$$\pi_x(y) = \frac{1}{\mathbb{E}_x[T_x^+]} \mathbb{E}_x \left[\sum_{k=1}^{T_x^+} \mathbb{1}\{X_k = y\} \right]$$

Η π_x είναι αναλλοίωτη κατανομή της αλυσίδας $\{X_n\}_{n \in \mathbb{N}_0}$ και

$$\pi_x(x) = \frac{1}{\mathbb{E}_x[T_x^+]}. \quad (5.9)$$

Παρατήρηση: Μπορούμε να σκεφτούμε την κατανομή π_x ως εξής. Ξεκινώντας από τη x κάνουμε μια εκδρομή μέχρι να επιστρέψουμε πάλι στην κατάσταση x , δηλαδή μέχρι τον χρόνο διακοπής T_x^+ . Κατά τη διάρκεια αυτής της εκδρομής σημειώνουμε πόσες επισκέψεις κάναμε στην κατάσταση $y \in \mathbb{X}$. Το πλήθος αυτών των επισκέψεων είναι

$$\sum_{k=1}^{T_x^+} \mathbb{1}\{X_k = y\}$$

και είναι βέβαια μια τυχαία μεταβλητή. Το βάρος που δίνει η π_x στην κατάσταση y είναι ανάλογο προς το αναμενόμενο πλήθος των επισκέψεών μας στην κατάσταση y κατά τη διάρκεια αυτής της εκδρομής γύρω από τη x . Η σταθερά αναλογίας $\frac{1}{\mathbb{E}_x[T_x^+]}$ που εμφανίζεται είναι απλά ένας παράγοντας κανονικοποίησης ώστε η π_x να είναι κατανομή.

Απόδειξη: Ας δούμε πρώτα ότι $\pi_x \in \mathcal{M}(\mathbb{X})$. Πράγματι, είναι φανερό από τον ορισμό ότι $\pi_x(y) \geq 0$ για κάθε $y \in \mathbb{X}$, ενώ

$$\sum_{y \in \mathbb{X}} \pi_x(y) = \frac{1}{\mathbb{E}_x[T_x^+]} \sum_{y \in \mathbb{X}} \mathbb{E}_x \left[\sum_{k=1}^{T_x^+} \mathbb{1}\{X_k = y\} \right] = \frac{1}{\mathbb{E}_x[T_x^+]} \mathbb{E}_x \left[\sum_{k=1}^{T_x^+} \sum_{y \in \mathbb{X}} \mathbb{1}\{X_k = y\} \right] = \frac{\mathbb{E}_x[T_x^+]}{\mathbb{E}_x[T_x^+]} = 1.$$

Ο ισχυρισμός της (5.9) προκύπτει άμεσα από τον ορισμό του χρόνου $T_x^+ = \inf\{k \geq 1 : X_k = x\}$, αφού κατά τη διάρκεια της εκδρομής γύρω από τη x , δηλαδή για $k \in \{1, 2, \dots, T_x^+\}$, η αλυσίδα μας βρίσκεται στη x μόνο τη χρονική στιγμή T_x^+ .

Θα δείξουμε τώρα ότι η π_x είναι αναλλοίωτη. Πράγματι, για κάθε $y \in \mathbb{X}$ έχουμε

$$\begin{aligned}
\pi_x(y) &= \pi_x(x) \mathbb{E}_x \left[\sum_{k=1}^{T_x^+} \mathbb{1}\{X_k = y\} \right] \\
&= \pi_x(x) \mathbb{E}_x \left[\sum_{k=1}^{\infty} \mathbb{1}\{X_k = y, T_x^+ \geq k\} \right] \\
&= \pi_x(x) \mathbb{E}_x \left[\sum_{k=1}^{\infty} \sum_{z \in \mathbb{X}} \mathbb{1}\{X_{k-1} = z, X_k = y, T_x^+ \geq k\} \right] \\
&= \pi_x(x) \sum_{z \in \mathbb{X}} \sum_{k=1}^{\infty} \mathbb{E}_x \left[\mathbb{1}\{X_{k-1} = z, X_k = y, T_x^+ \geq k\} \right] \quad (\text{Θεώρημα Fubini-Tonelli}) \\
&= \pi_x(x) \sum_{z \in \mathbb{X}} \sum_{k=1}^{\infty} \mathbb{P}_x [X_{k-1} = z, X_k = y, T_x^+ \geq k]. \tag{5.10}
\end{aligned}$$

Παρατηρήστε τώρα ότι $\mathbb{1}\{T_x^+ \geq k\} = \mathbb{1}\{X_1 \neq x, X_2 \neq x, \dots, X_{k-1} \neq x\}$, επομένως το ενδεχόμενο $\{T_x^+ \geq k\}$ ανήκει στην κλάση \mathcal{F}_{k-1} . Έτσι, από τη μαρκοβιανή ιδιότητα έχουμε

$$\mathbb{P}_x [X_k = y | X_{k-1} = z, T_x^+ \geq k] = \mathbb{P}_x [X_k = y | X_{k-1} = z] = p(z, y)$$

και η σχέση (5.10) γίνεται

$$\begin{aligned}
\pi_x(y) &= \pi_x(x) \sum_{z \in \mathbb{X}} p(z, y) \sum_{k=1}^{\infty} \mathbb{P}_x [X_{k-1} = z, T_x^+ \geq k] \\
&= \pi_x(x) \sum_{z \in \mathbb{X}} p(z, y) \sum_{k=1}^{\infty} \mathbb{E}_x [\mathbb{1}\{X_{k-1} = z, T_x^+ \geq k\}] \\
&= \pi_x(x) \sum_{z \in \mathbb{X}} p(z, y) \mathbb{E}_x \left[\sum_{k=1}^{\infty} \mathbb{1}\{X_{k-1} = z, T_x^+ \geq k\} \right] \quad (\text{Θεώρημα Fubini-Tonelli}) \\
&= \pi_x(x) \sum_{z \in \mathbb{X}} p(z, y) \mathbb{E}_x \left[\sum_{k=1}^{T_x^+} \mathbb{1}\{X_{k-1} = z\} \right]. \tag{5.11}
\end{aligned}$$

Παρατηρήστε τέλος ότι

$$\mathbb{E}_x \left[\sum_{k=1}^{T_x^+} \mathbb{1}\{X_{k-1} = z\} \right] \stackrel{m=k-1}{=} \mathbb{E}_x \left[\sum_{m=0}^{T_x^+-1} \mathbb{1}\{X_m = z\} \right] = \mathbb{E}_x \left[\sum_{m=1}^{T_x^+} \mathbb{1}\{X_m = z\} \right],$$

όπου η τελευταία ισότητα προκύπτει γιατί η αλυσίδα μας με \mathbb{P}_x -πιθανότητα 1 βρίσκεται στη x τόσο τη χρονική στιγμή 0, όσο και τη χρονική στιγμή T_x^+ . Έτσι, η (5.11) γίνεται

$$\pi_x(y) = \sum_{z \in \mathbb{X}} p(z, y) \pi_x(x) \mathbb{E}_x \left[\sum_{m=1}^{T_x^+} \mathbb{1}\{X_m = z\} \right] = \sum_{z \in \mathbb{X}} p(z, y) \pi_x(z)$$

και επομένως $\pi_x \in \mathcal{I}(P)$. □

Όπως είδαμε μια γνησίως επαναληπτική κατάσταση $x \in \mathbb{X}$ είναι επαναληπτική και άρα ανήκει αναγκαστικά σε μια κλειστή κλάση C_x . Από την κατασκευή της κατανομής π_x είναι φανερό ότι

$$y \notin C_x \Rightarrow \pi_x(y) = 0,$$

αφού αν $y \notin C_x$ η αλυσίδα που ξεκινά από τη x αποκλείεται να επισκεφτεί την κατάσταση y ποτέ και ειδικότερα αποκλείεται να επισκεφτεί την y μέχρι να επιστρέψει στη x . Θα δείξουμε τώρα ότι η π_x δίνει θετικό βάρος σε όλες τις καταστάσεις της κλάσης C_x .

Θεώρημα 22 Αν η κατάσταση x είναι γνησίως επαναληπτική και $y \in C_x$, τότε $\pi_x(y) > 0$.

Απόδειξη: Εφόσον η y είναι προσβάσιμη από τη x υπάρχει $m \in \mathbb{N}$ τέτοιο ώστε $p^{(m)}(x, y) > 0$. Επίσης εύκολα μπορούμε να δείξουμε επαγωγικά ότι

$$\pi_x \in \mathcal{I}(P) \Rightarrow \pi_x = \pi_x P \Rightarrow \pi_x = \pi_x P^m.$$

Επομένως

$$\pi_x(y) = \sum_{z \in \mathbb{X}} \pi_x(z) p^{(m)}(z, y) \geq \pi_x(x) p^{(m)}(x, y) > 0.$$

Πόρισμα 7 Η γνήσια επαναληπτικότητα είναι ιδιότητα κλάσης, δηλαδή αν η κατάσταση x είναι γνησίως επαναληπτική και $y \in C_x$, τότε και η y είναι γνησίως επαναληπτική.

Απόδειξη: Προκύπτει αμέσως από το Πόρισμα 5 αφού

$$\pi_x \in \mathcal{I}(P) \text{ και } \pi_x(y) > 0 \Rightarrow \mathbb{E}_y [T_y^+] < +\infty.$$

□

Με βάση το παραπάνω Πόρισμα έχει νόημα να κάνουμε λόγο για γνησίως επαναληπτικές κλάσεις. Σημειώστε ότι, εφόσον οι ανοιχτές κλάσεις μιας αλυσίδας είναι παροδικές, τότε κάθε γνησίως επαναληπτική κλάση είναι κλειστή. Το ακόλουθο Πόρισμα συνοψίζει μια ικανή και αναγκαία συνθήκη για το πότε μια κλειστή κλάση είναι γνησίως επαναληπτική και αποτελεί συχνά έναν εύκολον τρόπο για να αποδείξει κανείς τη γνήσια επαναληπτικότητα μιας κλειστής κλάσης. Σημειώστε ότι για τον σκοπό αυτό αρκεί να περιορίσουμε την αλυσίδα σε αυτή την κλειστή κλάση, οπότε η αλυσίδα που θα προκύψει θα είναι μη υποβιβάσιμη.

Πόρισμα 8 Μια μη υποβιβάσιμη μαρκοβιανή αλυσίδα είναι γνησίως επαναληπτική, αν και μόνο αν έχει αναλλοίωτη κατανομή.

Απόδειξη: Το Θεώρημα 21 εξασφαλίζει την ύπαρξη μιας αναλλοίωτης κατανομής π_x αν υπάρχει μια επαναληπτική κατάσταση x . Αντίστροφα, αν η αλυσίδα έχει αναλλοίωτη κατανομή π , τότε, από το Πόρισμα 6, η αλυσίδα θα έχει μια γνησίως επαναληπτική κατάσταση. Εφόσον η αλυσίδα είναι μη υποβιβάσιμη, από το Πόρισμα 7 όλες οι καταστάσεις της θα είναι γνησίως επαναληπτικές. □

Πόρισμα 9 Μια μη υποβιβάσιμη μαρκοβιανή αλυσίδα σε έναν πεπερασμένο χώρο καταστάσεων είναι γνησίως επαναληπτική.

Απόδειξη: Θεωρούμε ένα $x \in \mathbb{X}$ και ορίζουμε $T_x^+ = \inf\{k > 0 : X_k = x\}$ τον χρόνο πρώτης επιστροφής στην κατάσταση x . Εφόσον η αλυσίδα είναι μη υποβιβάσιμη, ολόκληρος ο \mathbb{X} είναι μια κλειστή και πεπερασμένη, άρα επαναληπτική κλάση. Έτσι, $\mathbb{P}_x [T_x^+ < \infty] = 1$. Ορίζουμε την $\lambda : \mathbb{X} \rightarrow [0, \infty]$ με

$$\lambda(y) = \mathbb{E}_x \left[\sum_{k=1}^{T_x^+} \mathbb{1}\{X_k = y\} \right].$$

Από την απόδειξη του Θεωρήματος 21 έχουμε ότι $\lambda = \lambda P$, ενώ $\lambda(x) = 1$. Θα δείξουμε ότι $\lambda(y) < \infty$ για κάθε $y \in \mathbb{X}$. Εφόσον η x είναι προσβάσιμη από την y , επιλέγουμε $m \in \mathbb{N}$ τέτοιο ώστε $p^{(m)}(y, x) > 0$. Έτσι

$$1 = \lambda(x) = \sum_{z \in \mathbb{X}} \pi(z) p^{(m)}(z, x) \geq \lambda(y) p^{(m)}(y, x) \Rightarrow \lambda(y) < \frac{1}{p^{(m)}(y, x)} < +\infty.$$

Αν ορίσουμε τώρα $Z = \sum_{y \in \mathbb{X}} \lambda(y) \in (1, +\infty)$ και $\pi(y) = \lambda(y)/Z$ για κάθε $y \in \mathbb{X}$, εύκολα επιβεβαιώνουμε ότι $\pi \in \mathcal{I}(P)$ και άρα, από το Πρόβλημα 8, η αλυσίδα είναι γνησίως επαναληπτική. \square

Έχουμε ως τώρα δει ότι οποιαδήποτε αναλλοίωτη κατανομή δίνει μηδενικό βάρος σε κλάσεις που δεν είναι γνησίως επαναληπτικές (Πρόβλημα 5), ενώ για κάθε γνησίως επαναληπτική κλάση έχουμε κατασκευάσει μια αναλλοίωτη κατανομή π_x που στηρίζεται στις καταστάσεις αυτής της κλάσης (Θεωρήματα 21 και 22). Θα δείξουμε στη συνέχεια ότι η αναλλοίωτη κατανομή που στηρίζεται σε μια γνησίως επαναληπτική κλάση είναι μοναδική. Όπως και στο προηγούμενο Πρόβλημα, περιορίζοντας την αλυσίδα σε μια γνησίως επαναληπτική και επομένως κλειστή κλάση, αρκεί να υποθέσουμε ότι η αλυσίδα μας είναι μη υποβιβάσιμη.

Θεώρημα 23 Αν η $\{X_n\}_{n \in \mathbb{N}_0}$ είναι μια μη υποβιβάσιμη γνησίως επαναληπτική αλυσίδα, τότε έχει μοναδική αναλλοίωτη κατανομή.

Απόδειξη: Η ύπαρξη μιας τουλάχιστον αναλλοίωτης κατανομής εξασφαλίζεται από το Θεώρημα 21. Θα αποδείξουμε εδώ τη μοναδικότητα. Έστω λοιπόν π μια αναλλοίωτη κατανομή της $\{X_n\}_{n \in \mathbb{N}_0}$.

Θεωρούμε ένα $x \in \mathbb{X}$ και ορίζουμε την αναλλοίωτη κατανομή π_x όπως στο Θεώρημα 21. Θα δείξουμε ότι $\pi(y) = \pi_x(y)$ για κάθε κατάσταση $y \in \mathbb{X}$. Εφόσον η αλυσίδα είναι μη υποβιβάσιμη, η x θα είναι προσβάσιμη από την y . Υπάρχει επομένως $m \in \mathbb{N}$ τέτοιο ώστε $p^{(m)}(y, x) > 0$. Έχουμε τώρα

$$\begin{aligned} 0 &= \pi(x) - \pi(x) = \pi(x) - \frac{\pi(x)}{\pi_x(x)} \pi_x(x) \\ &= \sum_{z \in \mathbb{X}} \pi(z) p^{(m)}(z, x) - \frac{\pi(x)}{\pi_x(x)} \sum_{z \in \mathbb{X}} \pi_x(z) p^{(m)}(z, x) \quad (\pi, \pi_x \in \mathcal{I}(P)) \\ &= \sum_{z \in \mathbb{X}} \left(\pi(z) - \pi(x) \frac{\pi_x(z)}{\pi_x(x)} \right) p^{(m)}(z, x). \end{aligned}$$

Από το Λήμμα 6 κάθε προσθετός στο παραπάνω άθροισμα είναι μεγαλύτερος ή ίσος του μηδενός. Για να είναι το άθροισμα 0 θα πρέπει όλοι οι προσθετέοι να είναι ίσοι με 0. Ειδικότερα,

$$\left(\pi(y) - \frac{\pi_x(y)}{\pi_x(x)} \pi(x) \right) p^{(m)}(y, x) = 0 \Rightarrow \pi(y) = \frac{\pi(x)}{\pi_x(x)} \pi_x(y).$$

Εφόσον η y είναι αυθαίρετα επιλεγμένη έχουμε ότι

$$\pi(y) = \frac{\pi(x)}{\pi_x(x)} \pi_x(y), \quad \text{για κάθε } y \in \mathbb{X}. \quad (5.12)$$

Αθροίζοντας τις παραπάνω σχέσεις σε όλα τα $y \in \mathbb{X}$ έχουμε

$$1 = \sum_{y \in \mathbb{X}} \pi(y) = \sum_{y \in \mathbb{X}} \frac{\pi(x)}{\pi_x(x)} \pi_x(y) = \frac{\pi(x)}{\pi_x(x)}$$

και άρα η (5.12) γίνεται $\pi(y) = \pi_x(y)$, για κάθε $y \in \mathbb{X}$. \square

Παρατήρηση: Μια άμεση συνέπεια του Θεωρήματος 23 είναι ότι, αν οι καταστάσεις x, y ανήκουν σε μια γνήσιως επαναληπτική κλάση C , τότε $\pi_x \equiv \pi_y$. Έχει λοιπόν νόημα να μιλάμε για τη μοναδική αναλλοίωτη κατανομή π_C που αντιστοιχεί στην κλάση C . Το τελικό συμπέρασμα αυτής της παραγράφου είναι ότι κάθε αναλλοίωτη κατανομή είναι ένας κυρτός συνδυασμός τέτοιων κατανομών.

Θεώρημα 24 Έστω \mathcal{R} το σύνολο των γνήσιως επαναληπτικών κλάσεων μιας μαρκοβιανής αλυσίδας $\{X_n\}_{n \in \mathbb{N}_0}$ με πίνακα πιθανοτήτων μετάβασης P . Τότε

$$\mathcal{I}(P) = \text{co}\{\pi_C : C \in \mathcal{R}\} \equiv \left\{ \sum_{C \in \mathcal{R}} \alpha(C) \pi_C : \alpha(C) \geq 0 \forall C \in \mathcal{R}, \text{ και } \sum_{C \in \mathcal{R}} \alpha(C) = 1 \right\}.$$

Απόδειξη: Ας υποθέσουμε πρώτα ότι $\pi = \sum_{C \in \mathcal{R}} \alpha(C) \pi_C$ με $\alpha(C) \geq 0$ και $\sum_{C \in \mathcal{R}} \alpha(C) = 1$. Είναι φανερό ότι $\pi(x) \geq 0$ για κάθε $x \in \mathbb{X}$ ενώ

$$\sum_{x \in \mathbb{X}} \pi(x) = \sum_{x \in \mathbb{X}} \sum_{C \in \mathcal{R}} \alpha(C) \pi_C(x) = \sum_{C \in \mathcal{R}} \alpha(C) \sum_{x \in \mathbb{X}} \pi_C(x) = \sum_{C \in \mathcal{R}} \alpha(C) = 1.$$

Επομένως $\pi \in \mathcal{M}(\mathbb{X})$. Για να δείξουμε ότι $\pi \in \mathcal{I}(P)$ και άρα $\text{co}\{\pi_C : C \in \mathcal{R}\} \subset \mathcal{I}(P)$ παρατηρούμε ότι

$$\pi P = \left(\sum_{C \in \mathcal{R}} \alpha(C) \pi_C \right) P = \sum_{C \in \mathcal{R}} \alpha(C) (\pi_C P) = \sum_{C \in \mathcal{R}} \alpha(C) \pi_C = \pi.$$

Θα δείξουμε τώρα ότι $\mathcal{I}(P) \subset \text{co}\{\pi_C : C \in \mathcal{R}\}$. Για $\mu \in \mathcal{I}(P)$ και $C \in \mathcal{R}$ τέτοια ώστε $\mu[C] > 0$, θεωρούμε τον περιορισμό $\mu|_C$ του μ στην κλάση C . Συγκεκριμένα, για κάθε $A \subset \mathbb{X}$ έχουμε

$$\mu|_C [A] = \mu[A|C] = \frac{\mu[A \cap C]}{\mu[C]}.$$

Θα δείξουμε ότι $\mu|_C = \pi_C$. Η $\mu|_C$ είναι μια κατανομή που στηρίζεται στην κλειστή κλάση C και, από το Θεώρημα 23, για να την ταυτίσουμε με την π_C , αρκεί να δείξουμε ότι είναι αναλλοίωτη. Πράγματι, για κάθε $x \in C$ έχουμε $p(y, x) = 0$ για κάθε $y \notin C$ και άρα

$$\mu|_C(x) = \frac{\mu(x)}{\mu[C]} = \frac{1}{\mu[C]} \sum_{y \in \mathbb{X}} \mu(y) p(y, x) = \sum_{y \in C} \frac{\mu(y)}{\mu[C]} p(y, x) = \sum_{y \in C} \mu|_C(y) p(y, x).$$

Δείξαμε λοιπόν ότι αν $\mu[C] > 0$, τότε $\mu|_C \equiv \pi_C$. Αν $\mu[C] = 0$, ορίζουμε $\mu|_C = \pi_C$. Από το Πόρισμα 5 έχουμε ότι

$$\mu \left[\bigcup_{C \in \mathcal{R}} C \right] = 1.$$

Έτσι, για κάθε $A \subset \mathbb{X}$

$$\mu[A] = \mu \left[\bigcup_{C \in \mathcal{R}} (A \cap C) \right] = \sum_{C \in \mathcal{R}} \mu[A \cap C] = \sum_{C \in \mathcal{R}} \mu[C] \mu|_C[A] = \sum_{C \in \mathcal{R}} \mu[C] \pi_C[A]$$

και άρα

$$\mu = \sum_{C \in \mathcal{R}} \mu[C] \pi_C.$$

Παρατηρήστε ότι $\mu[C] \geq 0$ για κάθε $C \in \mathcal{R}$ και $\sum_{C \in \mathcal{R}} \mu[C] = \mu \left[\bigcup_{C \in \mathcal{R}} C \right] = 1$. Έχουμε γράψει το μ ως ένα κυρτό συνδυασμό των π_C , άρα $\mathcal{I}(P) \subset \text{co}\{\pi_C : C \in \mathcal{R}\}$. \square

5.4 Παραδείγματα

Στην προηγούμενη παράγραφο μελετήσαμε τη δομή των αναλλοίωτων κατανομών μιας μαρκοβιανής αλυσίδας. Είδαμε ότι, για κάθε γνησίως επαναληπτική κλάση C , υπάρχει μια μοναδική αναλλοίωτη κατανομή π_C της αλυσίδας που στηρίζεται στη C (δηλαδή $x \notin C \Rightarrow \pi_C(x) = 0$) και ότι κάθε αναλλοίωτη κατανομή της αλυσίδας μπορεί να γραφτεί ως κυρτός συνδυασμός των κατανομών π_C . Είδαμε όμως και αρκετούς τρόπους για να χαρακτηρίσουμε τις π_C . Έτσι, αν έχουμε μια μη υποβιβάσιμη γνησίως επαναληπτική μαρκοβιανή αλυσίδα $\{X_n\}_{n \in \mathbb{N}_0}$ με πίνακα πιθανοτήτων μετάβασης P (όπως είναι οποιαδήποτε μαρκοβιανή αλυσίδα περιορισμένη σε μια γνησίως επαναληπτική της κλάση), έχουμε τους ακόλουθους χαρακτηρισμούς για τη μοναδική αναλλοίωτη κατανομή της π .

1. Η π είναι λύση του προβλήματος

$$\begin{cases} \pi = \pi P \\ \sum_{x \in \mathbb{X}} \pi(x) = 1. \end{cases}$$

2. Για κάθε $x \in \mathbb{X}$, αν $T_x^+ = \inf\{k > 0 : X_k = x\}$, έχουμε

$$\pi(x) = \frac{1}{\mathbb{E}_x[T_x^+]}$$

3. Για κάθε $x, y \in \mathbb{X}$ έχουμε

$$\pi(y) = \pi(x) \mathbb{E}_x \left[\sum_{k=1}^{T_x^+} \mathbb{1}\{X_k = y\} \right].$$

Μπορεί κανείς να χρησιμοποιήσει έναν τρόπο για να υπολογίσει την αναλλοίωτη κατανομή της αλυσίδας και να πάρει από τους άλλους τρόπους ενδεχομένως χρήσιμες πληροφορίες.

Παράδειγμα 42 Ένα έντομο κινείται στις κορυφές V ενός κανονικού ν -γώνου. Σε κάθε βήμα του μετακινείται σε μία από τις δύο γειτονικές κορυφές με πιθανότητα $p \in (0, 1)$ κατά τη φορά των δεικτών του ρολογιού και με πιθανότητα $1 - p$ αντίστροφα από τους δείκτες του ρολογιού. Ποιος είναι ο αναμενόμενος αριθμός βημάτων μέχρι να επιστρέψει στην αρχική του θέση;

Η αλυσίδα που καταγράφει τη θέση του εντόμου είναι μη υποβιβάσιμη, αφού το έντομο μπορεί να επισκεφτεί όλες τις καταστάσεις και να επιστρέψει στην αρχική του κατάσταση κάνοντας ν βήματα κατά τη φορά των δεικτών του ρολογιού, ενδεχόμενο το οποίο έχει πιθανότητα $p^\nu > 0$. Εφόσον $|V| < +\infty$ η αλυσίδα θα είναι γνησίως επαναληπτική και άρα θα έχει μοναδική αναλλοίωτη κατανομή π . Στο συγκεκριμένο παράδειγμα δεν είναι δύσκολο να μαντέψει κανείς ότι λόγω συμμετρίας η μοναδική αναλλοίωτη κατανομή θα είναι η ομοιόμορφη κατανομή στις κορυφές του ν -γώνου

$$\pi(x) = \frac{1}{\nu} \quad \text{για κάθε } x \in V.$$

Πράγματι, για κάθε $x \in V$

$$\sum_{y \in V} \pi(y) p(y, x) = \sum_{y \in V} \frac{1}{\nu} p(y, x) = \frac{1}{\nu} (p + (1 - p)) = \frac{1}{\nu} = \pi(x).$$

Επομένως

$$\mathbb{E}_x[T_x^+] = \frac{1}{\pi(x)} = \nu.$$

□

Το προηγούμενο παράδειγμα μπορεί να γενικευτεί κατά τον ακόλουθο τρόπο. Θα λέμε ότι ένας στοχαστικός πίνακας P είναι διπλά στοχαστικός αν $\sum_{x \in \mathbb{X}} p(x, y) = 1$ για κάθε $y \in \mathbb{X}$. Έτσι σε έναν διπλά στοχαστικό πίνακα, εκτός από το άθροισμα των στοιχείων του κατά γραμμή, και το άθροισμα των στοιχείων του κατά στήλη είναι ίσο με 1. Παρατηρήστε ότι ένας συμμετρικός πίνακας πιθανοτήτων μετάβασης είναι πάντα διπλά στοχαστικός.

Παράδειγμα 43 Μια μη υποβιβάσιμη μαρκοβιανή αλυσίδα που κινείται σε έναν πεπερασμένο χώρο καταστάσεων \mathbb{X} και έχει διπλά στοχαστικό πίνακα πιθανοτήτων μετάβασης, έχει μοναδική αναλλοίωτη κατανομή την ομοιόμορφη κατανομή στον \mathbb{X} .

Από το Θεώρημα 23, η αλυσίδα έχει μοναδική αναλλοίωτη κατανομή. Αρκεί λοιπόν να επαληθεύσουμε ότι η ομοιόμορφη κατανομή $\pi : \mathbb{X} \rightarrow [0, 1]$ με $\pi(x) = \frac{1}{|\mathbb{X}|}$ είναι αναλλοίωτη. Πράγματι, για κάθε $x \in \mathbb{X}$ έχουμε

$$\sum_{y \in \mathbb{X}} \pi(y)p(y, x) = \frac{1}{|\mathbb{X}|} \sum_{y \in \mathbb{X}} p(y, x) = \frac{1}{|\mathbb{X}|} = \pi(x),$$

όπου η δεύτερη ισότητα προκύπτει επειδή ο πίνακας πιθανοτήτων μετάβασης είναι διπλά στοχαστικός. Επίσης,

$$\sum_{x \in \mathbb{X}} \pi(x) = \sum_{x \in \mathbb{X}} \frac{1}{|\mathbb{X}|} = \frac{1}{|\mathbb{X}|} |\mathbb{X}| = 1$$

και άρα $\pi \in \mathcal{I}(P)$. □

Παράδειγμα 44 Σε ένα ράφι της βιβλιοθήκης σας υπάρχουν τρία βιβλία: Algebra, Basic Topology, Calculus, που θα συμβολίζουμε με A, B, C για συντομία. Κάθε πρωί παίρνετε τυχαία ένα βιβλίο από τη θέση του με πιθανότητα p, q, r αντίστοιχα. Υποθέτουμε $p, q, r > 0$ με $p + q + r = 1$. Όταν τελειώνετε το διάβασμά σας για την ημέρα το ξαναβάζετε στο ράφι στην αριστερότερη θέση. Η διάταξη των βιβλίων είναι μια μαρκοβιανή αλυσίδα στον χώρο \mathbb{X} των μεταθέσεων των συμβόλων $\{A, B, C\}$. Αν κάποια στιγμή τα βιβλία είναι τοποθετημένα με αλφαβητική σειρά, βρείτε τον αναμενόμενο αριθμό ημερών μέχρι τα βιβλία να ξαναβρεθούν τοποθετημένα με αλφαβητική σειρά. Πόσες κατά μέση τιμή ημέρες θα διαβάσετε το βιβλίο Calculus ενδιάμεσα σε αυτές τις στιγμές;

Ας απαριθμήσουμε τις δυνατές καταστάσεις με την εξής σειρά

$$\mathbb{X} = \{ABC, CAB, BCA, BAC, ACB, CBA\}.$$

Ο πίνακας πιθανοτήτων μετάβασης της αλυσίδας είναι ο

$$P = \begin{pmatrix} p & r & 0 & q & 0 & 0 \\ 0 & r & q & 0 & p & 0 \\ p & 0 & q & 0 & 0 & r \\ p & 0 & 0 & q & 0 & r \\ 0 & r & 0 & q & p & 0 \\ 0 & 0 & q & 0 & p & r \end{pmatrix}.$$

Το ότι η αλυσίδα είναι μη υποβιβάσιμη είναι διαισθητικά φανερό. Μπορείτε να δείτε ότι, όποια κι αν είναι η αρχική της κατάσταση, αν επιλέξετε τις δύο πρώτες μέρες το δεξιότερο βιβλίο, την τρίτη μέρα το μεσαίο βιβλίο και τις δυο επόμενες πάλι το δεξιότερο βιβλίο (ενδεχόμενο το οποίο έχει θετική πιθανότητα), τότε η διάταξη των βιβλίων στο ράφι θα έχει περάσει από όλες τις δυνατές καταστάσεις της. Εφόσον ο χώρος

καταστάσεων είναι πεπερασμένος, όλες οι καταστάσεις θα είναι γνησίως επαναληπτικές. Για να βρούμε τη μοναδική αναλλοίωτη κατανομή π , λύνουμε το ομογενές σύστημα εξισώσεων $\pi = \pi P$. Ας δούμε τις εξισώσεις που αφορούν σε καταστάσεις με το A ως αριστερότερο βιβλίο:

$$\pi(ABC) = p\pi(ABC) + p\pi(BAC) + p\pi(BCA) \quad (5.13)$$

$$\pi(ACB) = p\pi(ACB) + p\pi(CAB) + p\pi(CBA). \quad (5.14)$$

Αν $S_A = \{ABC, ACB\}$, προσθέτοντας κατά μέλη βρίσκουμε $\pi(S_A) = p \sum_{x \in \mathbb{X}} \pi(x) = p$. Όμοια, αν $S_B = \{BAC, BCA\}$ και $S_C = \{CAB, CBA\}$, τότε $\pi(S_B) = q$, $\pi(S_C) = r$. Από τις (5.13), (5.14) παίρνουμε τελικά ότι

$$\pi(ABC) = \frac{p}{1-p} \pi(S_B) = \frac{pq}{1-p} \quad \text{και} \quad \pi(ACB) = \frac{p}{1-p} \pi(S_C) = \frac{pr}{1-p}. \quad (5.15)$$

Οι πιθανότητες των άλλων καταστάσεων βρίσκονται με ανάλογο τρόπο, οπότε τελικά

$$\pi = \left(\frac{pq}{1-p}, \frac{rp}{1-r}, \frac{qr}{1-q}, \frac{qp}{1-q}, \frac{pr}{1-p}, \frac{rq}{1-r} \right).$$

Έχοντας βρει την αναλλοίωτη κατανομή της αλυσίδας, μπορούμε εύκολα να απαντήσουμε τα υπόλοιπα ερωτήματα. Έτσι, αν $T_{ABC}^+ = \inf\{k > 0 : X_k = ABC\}$, τότε

$$\mathbb{E}[T_{ABC}^+ | X_0 = ABC] = \frac{1}{\pi(ABC)} = \frac{1-p}{pq}.$$

Κάθε φορά που διαβάζετε το βιβλίο Calculus το τοποθετείτε αριστερά οπότε η αλυσίδα περνά από μία από της καταστάσεις του S_C . Έτσι,

$$\mathbb{E} \left[\sum_{k=1}^{T_{ABC}^+} \mathbb{1}\{X_k \in S_C\} \mid X_0 = ABC \right] = \frac{\pi[S_C]}{\pi(ABC)} = \frac{(1-p)r}{pq}. \quad \square$$

5.5 Χρονική αντιστρεψιμότητα και ακριβής ισορροπία

Ας θεωρήσουμε μια μαρκοβιανή αλυσίδα $\{X_n\}_{n \in \mathbb{N}_0}$ που εξελίσσεται σε έναν αριθμησιμο χώρο καταστάσεων \mathbb{X} και βρίσκεται στην κατάσταση ισορροπίας της με κατανομή π . Όπως έχουμε εξηγήσει, η κατάσταση ισορροπίας μιας μαρκοβιανής αλυσίδας είναι μια κατάσταση δυναμικής ισορροπίας. Η αλυσίδα εξελίσσεται στον χρόνο αλλά με τέτοιο τρόπο ώστε η κατανομή της να μένει σταθερή. Ας κάνουμε τώρα το ακόλουθο νοητικό πείραμα. Ας υποθέσουμε πως κινηματογραφούμε την αλυσίδα για κάποιο χρόνο και έπειτα αναπαράγουμε την ταινία που τραβήξαμε ανάποδα. Πώς μοιάζει η δυναμική της αλυσίδας όταν ο χρόνος κυλάει ανάποδα; Την απάντηση σε αυτό το ερώτημα δίνει το παρακάτω θεώρημα.

Θεώρημα 25 Έστω $\{X_n\}_{n \in \mathbb{N}_0}$ μια μη υποβιβάσιμη στάσιμη μαρκοβιανή αλυσίδα με κατανομή ισορροπίας π . Για τον χρόνο $N \in \mathbb{N}$ ορίζουμε την ακολουθία $\{Y_n\}_{0 \leq n \leq N}$ με $Y_n = X_{N-n}$. Τότε η $\{Y_n\}_n$ είναι επίσης μια στάσιμη μαρκοβιανή αλυσίδα με κατανομή π αλλά με πιθανότητες μετάβασης $\hat{p}(\cdot, \cdot)$ που δίνονται από τη σχέση

$$\hat{p}(x, y) = \frac{\pi(y)}{\pi(x)} p(y, x), \quad \forall x, y \in \mathbb{X}. \quad (5.16)$$

Απόδειξη: Οι \hat{p} είναι καλά ορισμένες, αφού, όπως έχουμε δει, για την αναλλοίωτη κατανομή π μιας στάσιμης αλυσίδας έχουμε $\pi(x) > 0$, για κάθε $x \in \mathbb{X}$. Επίσης, είναι πιθανότητες μετάβασης, αφού $\hat{p}(x, y) \geq 0$ για κάθε $x, y \in \mathbb{X}$, ενώ

$$\sum_{y \in \mathbb{X}} \hat{p}(x, y) = \frac{1}{\pi(x)} \sum_{y \in \mathbb{X}} \pi(y) p(y, x) = 1, \quad \forall x \in \mathbb{X},$$

αφού η $\pi(\cdot)$ είναι αναλλοίωτη κατανομή για τη $\{X_n\}_n$. Θα δούμε τώρα ότι οι \hat{p} είναι οι πιθανότητες μετάβασης για την $\{Y_n\}$. Πράγματι, για $k \in \{1, 2, \dots, N\}$, έχουμε

$$\begin{aligned} \mathbb{P}[Y_0 = y_0, Y_1 = y_1, \dots, Y_k = y_k] &= \mathbb{P}[X_{N-k} = y_k, \dots, X_{N-1} = y_1, X_N = y_0] \\ &= \mathbb{P}[X_{N-k} = y_k] \mathbb{P}[X_{N-k+1} = y_{k-1} | X_{N-k} = y_k] \cdots \mathbb{P}[X_N = y_0 | X_{N-1} = y_1] \\ &= \pi(y_k) p(y_k, y_{k-1}) \cdots p(y_1, y_0) = \pi(y_0) \hat{p}(y_0, y_1) \cdots \hat{p}(y_{k-1}, y_k), \end{aligned}$$

όπου η τελευταία ισότητα προκύπτει από τον ορισμό (5.16) των \hat{p} . Τέλος, για κάθε $x \in \mathbb{X}$, έχουμε

$$\sum_{y \in \mathbb{X}} \pi(y) \hat{p}(y, x) = \pi(x) \sum_{y \in \mathbb{X}} p(x, y) = \pi(x),$$

άρα η κατανομή π είναι αναλλοίωτη για την $\{Y_n\}_n$ και, εφόσον η κατανομή της $Y_k = X_{N-k}$ για κάθε $k \in \{1, 2, \dots, N\}$ είναι π , η $\{Y_n\}_n$ είναι στάσιμη. \square

Αν η κατανομή π ικανοποιεί την

$$\pi(x) p(x, y) = \pi(y) p(y, x), \quad \forall x, y \in \mathbb{X}, \quad (5.17)$$

τότε, από τη σχέση (5.16), έχουμε ότι $\hat{p} = p$ και άρα η $\{Y_n\}_n$, η χρονική αντιστροφή της $\{X_n\}_n$, έχει την ίδια στάσιμη κατανομή και τις ίδιες πιθανότητες μετάβασης όπως και η $\{X_n\}_n$. Με άλλα λόγια, αν κινηματογραφήσουμε τη $\{X_n\}_n$, δεν θα μπορούμε με την αναπαραγωγή του φιλμ να διακρίνουμε τη φορά του χρόνου. Όταν ικανοποιείται η (5.17), λέμε ότι η κατανομή π και οι πιθανότητες μετάβασης p βρίσκονται σε *ακριβή ισορροπία* (detailed balance) και ότι η αλυσίδα είναι *χρονικά αντιστρέψιμη* (time reversible).

Τα περισσότερα ενδιαφέροντα φυσικά συστήματα ικανοποιούν τις συνθήκες ακριβούς ισορροπίας (5.17). Ένας από τους λόγους για τους οποίους οι συνθήκες αυτές είναι ιδιαίτερα χρήσιμες είναι ότι, όταν ικανοποιούνται, μας επιτρέπουν να εντοπίσουμε μια αναλλοίωτη κατανομή π πιο εύκολα από τις συνθήκες ορισμού της

$$\pi(x) = \sum_{y \in \mathbb{X}} \pi(y) p(y, x), \quad \forall x \in \mathbb{X}. \quad (5.18)$$

Πράγματι, όπως φαίνεται από την (5.17), οδηγούν σε απλούστερες εξισώσεις για την π , που σύμφωνα με το ακόλουθο Λήμμα όταν έχουν λύση, αυτή είναι πάντα αναλλοίωτη κατανομή.

Λήμμα 7 Αν η κατανομή π και οι πιθανότητες μετάβασης p βρίσκονται σε ακριβή ισορροπία, τότε η π είναι αναλλοίωτη κατανομή για τη μαρκοβιανή αλυσίδα με πιθανότητες μετάβασης p .

Απόδειξη: Για κάθε $x \in \mathbb{X}$ έχουμε

$$\sum_{y \in \mathbb{X}} \pi(y) p(y, x) \stackrel{(5.17)}{=} \sum_{y \in \mathbb{X}} \pi(x) p(x, y) = \pi(x) \sum_{y \in \mathbb{X}} p(x, y) = \pi(x). \quad \square$$

Παράδειγμα 45 Στο μοντέλο διάχυσης του Ehrenfest N σωματίδια τοποθετούνται σε ένα δοχείο με δύο διαμερίσματα, A και B. Σε κάθε βήμα επιλέγουμε τυχαία ένα από τα N σωματίδια και του αλλάζουμε διαμέρισμα.

Αν $\{X_n\}_{n \in \mathbb{N}_0}$ είναι η μαρκοβιανή αλυσίδα στον χώρο καταστάσεων $\mathbb{X} = \{0, 1, \dots, N\}$ που περιγράφει το πλήθος των σωματιδίων στο διαμέρισμα A, τότε οι πιθανότητες μετάβασης αυτής της αλυσίδας είναι

$$p(k, k+1) = \frac{N-k}{N} \quad \text{και} \quad p(k, k-1) = \frac{k}{N}, \quad \forall k \in \mathbb{X}.$$

Πράγματι, για να αυξηθεί κατά 1 το πλήθος των σωματιδίων στο διαμέρισμα A θα πρέπει να επιλέξουμε ένα από τα σωματίδια του διαμερίσματος B για να του αλλάξουμε διαμέρισμα. Αν στο διαμέρισμα A υπάρχουν k σωματίδια, τότε στο διαμέρισμα B υπάρχουν $N-k$ και άρα η πιθανότητα να επιλέξουμε ένα από τα σωματίδια του B είναι $\frac{N-k}{N}$. Ομοίως για να μειωθεί κατά 1 το πλήθος των σωματιδίων στο διαμέρισμα A θα πρέπει το σωματίδιο που θα επιλέξουμε για να του αλλάξουμε διαμέρισμα να προέρχεται από το A, ενδεχόμενο που έχει πιθανότητα $\frac{k}{N}$ όταν στο A υπάρχουν k σωματίδια.

Είναι φανερό ότι η $\{X_n\}_n$ είναι μη υποβιβάσιμη και άρα θα έχει μοναδική αναλλοίωτη κατανομή π . Αν επιχειρήσουμε να προσδιορίσουμε την π από τις εξισώσεις (5.18), θα οδηγηθούμε σε μια μη γραμμική αναδρομική εξίσωση δεύτερης τάξης που θέλει αρκετό κόπο για να λυθεί. Δοκιμάστε το για εξάσκηση. Αν αντίθετα επιχειρήσουμε να βρούμε κάποια κατανομή π στον \mathbb{X} που βρίσκεται σε ακριβή ισορροπία με τις πιθανότητες μετάβασης p , θα πάρουμε τις εξισώσεις

$$\pi(k) p(k, k+1) = \pi(k+1) p(k+1, k) \quad k \in \{0, 1, \dots, N-1\},$$

ή ισοδύναμα

$$\frac{\pi(k+1)}{\pi(k)} = \frac{N-k}{k} \quad k \in \{0, 1, \dots, N-1\}.$$

Πολλαπλασιάζοντας τις σχέσεις αυτές για $k = 0, 1, \dots, n-1$ λαμβάνουμε ότι

$$\pi(n) = \pi(0) \binom{N}{n}, \quad \forall n \in \mathbb{X},$$

ενώ από τη συνθήκη κανονικοποίησης $\sum_{n \in \mathbb{X}} \pi(n) = 1$ έχουμε ότι $\pi(0) \sum_{n=0}^N \binom{N}{n} = \pi(0) 2^N = 1$. Επομένως,

$$\pi(n) = 2^{-N} \binom{N}{n}, \quad \forall n \in \mathbb{X}. \quad \square$$

Παράδειγμα 46 : Τυχαίος περίπατος σε γράφο

Έστω $\{X_n\}_n$ ένας περίπατος στις κορυφές V ενός συνεκτικού, μη προσανατολισμένου γράφου. Σε κάθε βήμα ο περιπατητής επιλέγει τυχαία μία από τις ακμές της κορυφής που βρίσκεται και μεταβαίνει στην άλλη κορυφή αυτής της ακμής.

Αν ο περιπατητής μας βρίσκεται σε μια κορυφή $x \in V$ με βαθμό $v(x)$, τότε όλες οι $v(x)$ κορυφές που συνδέονται με την κορυφή x είναι ισοπίθανες ως επόμενος προορισμός. Επομένως, αν E είναι το σύνολο των ακμών του γράφου, οι πιθανότητες μετάβασης της αλυσίδας δίνονται από την

$$p(x, y) = \begin{cases} \frac{1}{v(x)}, & \text{αν } (x, y) \in E \\ 0, & \text{αν } (x, y) \notin E. \end{cases} \quad (5.19)$$

Εφόσον ο γράφος είναι συνεκτικός (ανάμεσα σε οποιοσδήποτε δυο κορυφές υπάρχει ένα μονοπάτι, του οποίου οι διαδοχικές κορυφές συνδέονται με ακμές), η αλυσίδα είναι μη υποβιβάσιμη και έχει επομένως μοναδική αναλλοίωτη κατανομή π , την οποία θέλουμε να προσδιορίσουμε. Από την (5.19) προκύπτει άμεσα ότι

$$v(x) p(x, y) = v(y) p(y, x) = \chi_E(x, y), \quad \forall x, y \in V,$$

όπου $\chi_E(x, y) = 1$ αν $(x, y) \in E$ και 0 διαφορετικά. Επομένως, αν για κάποια σταθερά c έχουμε $\pi(x) = cv(x)$, $\forall x \in V$, τότε η π βρίσκεται σε ακριβή ισορροπία με τις πιθανότητες μετάβασης p . Η σταθερά c θα προσδιοριστεί από τη συνθήκη κανονικοποίησης

$$1 = \sum_{x \in V} \pi(x) = c \sum_{x \in V} v(x) = 2c|E|,$$

όπου $|E|$ είναι το πλήθος των ακμών του γράφου. Η τελευταία ισότητα προκύπτει γιατί κάθε ακμή προστίθεται στο άθροισμα δύο φορές, μία για κάθε κορυφή στην οποία καταλήγει. Άρα

$$\pi(x) = \frac{v(x)}{2|E|}, \quad \forall x \in V. \quad \square$$

Ένας ακόμα πολύ σημαντικός λόγος για τον οποίο οι συνθήκες ακριβούς ισορροπίας είναι πολύ χρήσιμες είναι ότι μας δίνουν τη δυνατότητα να κατασκευάσουμε μια μαρκοβιανή αλυσίδα με επιθυμητή αναλλοίωτη κατανομή, όπως θα δούμε παρακάτω στον αλγόριθμο Metropolis-Hastings.

5.6 Ασκήσεις

Άσκηση 79 Η $\{X_n\}_{n \in \mathbb{N}_0}$ είναι μια μαρκοβιανή αλυσίδα στον χώρο καταστάσεων $\mathbb{X} = \{1, 2, 3, 4\}$ με πίνακα μετάβασης

$$P = \begin{pmatrix} 0 & 1/2 & 1/3 & 1/6 \\ 1/2 & 0 & 1/4 & 1/4 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 1/3 & 1/6 & 0 \end{pmatrix}.$$

α) Βρείτε την αναλλοίωτη κατανομή της αλυσίδας.

β) Αν $X_0 = 1$, υπολογίστε τον αναμενόμενο χρόνο πρώτης επιστροφής $T_1^+ = \inf\{k > 0 : X_k = 1\}$ στην κατάσταση 1.

γ) Υπολογίστε τον αναμενόμενο αριθμό επισκέψεων στην κατάσταση 3 μέχρι τη συμπλήρωση 93 επιστροφών στην κατάσταση 1.

Άσκηση 80 α) Βρείτε όλες τις αναλλοίωτες κατανομές της αλυσίδας $\{X_n\}$ που περιγράψαμε στην Άσκηση 5. Ίσως βρείτε χρήσιμο το Παράδειγμα 46.

β) Υπολογίστε τον αναμενόμενο αριθμό ημερών μέχρι την πρώτη επιστροφή του έντομου στο μπάνιο.

γ) Υπολογίστε τον αναμενόμενο αριθμό ημερών που θα περάσει το έντομο στο σαλόνι μέχρι την πρώτη του επιστροφή στο μπάνιο.

δ) Αν ένα άλλο έντομο αρχικά βρίσκεται στην κουζίνα και μετακινείται κάθε μέρα όπως το πρώτο, ποια είναι η πιθανότητα κάποια μέρα τα δύο έντομα να βρεθούν στο ίδιο δωμάτιο;

Άσκηση 81 Ένας παντοπώλης εφοδιάζεται κάθε πρωί με ένα πακέτο μπισκότα. Έχει παρατηρήσει ότι η ημερήσια ζήτηση είναι μια τυχαία μεταβλητή X με κατανομή $\mathbb{P}[X = 0] = \frac{1}{4}$, $\mathbb{P}[X = 1] = \frac{1}{2}$, $\mathbb{P}[X = 2] = \frac{1}{6}$, $\mathbb{P}[X = 3] = \frac{1}{12}$. Περιγράψτε την ποσότητα από μπισκότα που έχει στο παντοπωλείο κάθε βράδυ ως μια μαρκοβιανή αλυσίδα και βρείτε την αναλλοίωτη κατανομή της. Αν χτες το βράδυ δεν είχε μείνει κανένα πακέτο μπισκότα στο παντοπωλείο, ποιος είναι ο αναμενόμενος αριθμός ημερών μέχρι την επόμενη φορά που το παντοπωλείο θα μείνει χωρίς μπισκότα;

Άσκηση 82 Μια μαρκοβιανή αλυσίδα στον $\mathbb{X} = \mathbb{N}_0$ μετατοπίζεται ένα βήμα προς τα αριστερά όταν δεν βρίσκεται στο 0 και όταν φτάσει στο 0 κάνει ένα άλμα που ακολουθεί γεωμετρική κατανομή με παράμετρο q ($0 < q < 1$). Υπολογίστε την αναλλοίωτη κατανομή της με τρεις διαφορετικούς τρόπους.

Άσκηση 83 Μια μαρκοβιανή αλυσίδα στον $\mathbb{X} = \{0, 1, 2, \dots\}$ έχει πιθανότητες μετάβασης $p(k, k+1) = p < 1$, $p(k, 0) = 1 - p$, $\forall k \in \mathbb{X}$. Βρείτε την αναλλοίωτη κατανομή της. Ποια είναι η αναμενόμενη τιμή του χρόνου που μεσολαβεί ανάμεσα σε δύο διαδοχικές επισκέψεις στο 3; Ποιος είναι ο αναμενόμενος αριθμός επισκέψεων στο 0 ανάμεσα σε δύο διαδοχικές επισκέψεις στο 3;

Άσκηση 84 Ο κύριος X αντιμετωπίζει ένα σοβαρό πρόβλημα μνήμης. Κάθε νύχτα ξεχνά ένα μέρος από τα πρόσωπα που γνωρίζει. Συγκεκριμένα, αν θυμάται k πρόσωπα πριν πέσει για ύπνο, το πλήθος των προσώπων που εξακολουθεί να θυμάται μόλις ξυπνήσει μπορεί να είναι $0, 1, 2, \dots, k$, καθένα με πιθανότητα $1/(k+1)$. Ο γιατρός που τον παρακολουθεί του μαθαίνει κάθε μέρα ένα πρόσωπο, διαφορετικό από αυτά που θυμάται. Αν X_n είναι το πλήθος των προσώπων που ο κύριος X θυμάται το βράδυ της n -στης ημέρας:

- α) Βρείτε τις πιθανότητες μετάβασης $p(j, k)$, $j, k \in \mathbb{N}$ της αλυσίδας X_n .
- β) Δείξτε ότι η μοναδική αναλλοίωτη κατανομή της $\{X_n\}$ είναι η $\pi : \mathbb{N} \rightarrow [0, 1]$ με $\pi(k) = \frac{1}{e^{(k-1)!}}$.
- γ) Αν κάποιος βράδυ ο κύριος X θυμάται 3 πρόσωπα πριν πέσει για ύπνο, ποιος είναι ο αναμενόμενος χρόνος που θα μεσολαβήσει μέχρι το επόμενο βράδυ που θα θυμάται πάλι 3 πρόσωπα;

Άσκηση 85 Αν $\mathcal{R} \neq \emptyset$ είναι το σύνολο των γνησίως επαναληπτικών κλάσεων μιας αλυσίδας $\{X_n\}_n$ και $\pi \in \mathcal{I}(P)$, δείξτε ότι υπάρχει μόνο ένας τρόπος να γράψουμε την π ως

$$\pi = \sum_{C \in \mathcal{R}} \alpha(C) \pi_C.$$

Άσκηση 86 Δείξτε ότι για μια μαρκοβιανή αλυσίδα σε έναν πεπερασμένο χώρο καταστάσεων \mathbb{X} με γεννήτορα L και αναλλοίωτη κατανομή π έχουμε

$$\sum_{x \in \mathbb{X}} \pi(x) f(x) Lf(x) = -\frac{1}{2} \sum_{x, y \in \mathbb{X}} \pi(x) p(x, y) (f(y) - f(x))^2.$$

Συμπεράνετε ότι, αν η αλυσίδα είναι μη υποβιβάσιμη και $Lf(x) = 0$ για κάθε $x \in \mathbb{X}$, τότε η f είναι σταθερή συνάρτηση.

Άσκηση 87 (σταθερά του Kemeny) Έστω $\{X_n\}_n$ μια μη υποβιβάσιμη αλυσίδα σε έναν πεπερασμένο χώρο καταστάσεων \mathbb{X} με αναλλοίωτη κατανομή π . Θεωρούμε μια τυχαία μεταβλητή Y ανεξάρτητη από τις $\{X_n\}$ και με κατανομή π , δηλαδή $\mathbb{P}[Y = x] = \pi(x)$, $\forall x \in \mathbb{X}$. Ορίζουμε τώρα $T_Y = \inf\{k \geq 0 : X_k = Y\}$. Ποια είναι η κατανομή της τυχαίας μεταβλητής X_{T_Y} ; Δείξτε ότι η συνάρτηση $f(x) = \mathbb{E}[T_Y | X_0 = x]$ είναι σταθερή.

Άσκηση 88 Θεωρήστε μια αλυσίδα Markov στον $\mathbb{X} = \{0, 1, 2, \dots\}$ με $p(k, k+1) = \frac{1}{k+1}$, $p(k, k-1) = \frac{k}{k+1}$, για κάθε $k \in \mathbb{X}$. Βρείτε μια αναλλοίωτη κατανομή της. Είναι αυτή μοναδική; Υπόδειξη: δοκιμάστε να βρείτε μια κατανομή που ικανοποιεί τις συνθήκες ακριβούς ισορροπίας. Αφού τη βρείτε, δοκιμάστε να λύσετε το πρόβλημα με τον ορισμό της αναλλοίωτης κατανομής. Παρατηρήστε ότι είναι πολύ πιο εύκολο να βρει κανείς την αναλλοίωτη κατανομή π μέσω των συνθηκών ακριβούς ισορροπίας, όταν η π τις ικανοποιεί. Αυτό βέβαια συμβαίνει μόνο αν η αλυσίδα είναι χρονικά αντιστρέψιμη.

Άσκηση 89 Ένα αλογάκι ξεκινά από το κάτω αριστερά τετραγώνια μιας σκακιέρας. Σε κάθε βήμα, επιλέγει τυχαία μία από τις επιτρεπτές κινήσεις που έχει και μεταβαίνει στο αντίστοιχο τετράγωνο. Ποιος είναι ο αναμενόμενος χρόνος μέχρι την πρώτη επιστροφή του στο τετράγωνο απ' όπου ξεκίνησε;

Άσκηση 90 Ένας φίλος σας ισχυρίζεται ότι πολλές στατιστικές ιδιότητες των δεκαδικών ψηφίων του π προκύπτουν κάνοντας την υπόθεση ότι κάθε ψηφίο επιλέγεται ομοιόμορφα από τα $\{0, 1, \dots, 9\}$ και ανεξάρτητα από τα άλλα. Ο φίλος σας διάβασε ότι η πρώτη εμφάνιση 9 διαδοχικών ίδιων ψηφίων στο δεκαδικό ανάπτυγμα του π συμβαίνει ξεκινώντας από το 24.658.601 ψηφίο (οπότε εμφανίζονται 9 διαδοχικά 7άρια) και σάς ζητά να τον βοηθήσετε να ελέγξει κατά πόσον αυτή η παρατήρηση είναι συμβατή με τον ισχυρισμό του.

Θεωρήστε μια ακολουθία $\{\xi_n\}_{n \in \mathbb{N}}$ από ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με ομοιόμορφη κατανομή στο σύνολο $\{0, 1, \dots, 9\}$. Θα συμβολίζουμε με X_n του τρέχον μήκος του σερί από ίδια αποτελέσματα. Συγκεκριμένα

$$X_0 = 0 \quad \text{και} \quad X_n = \min \{k \geq 0 : \xi_{n-k} \neq \xi_n, n\}, \quad n \in \mathbb{N}.$$

Για παράδειγμα, αν $(\xi_1, \dots, \xi_7) = (3, 5, 5, 5, 9, 6, 6)$, τότε $(X_0, X_1, \dots, X_7) = (0, 1, 1, 2, 3, 1, 1, 2)$. Η $\{X_n\}_{n \in \mathbb{N}_0}$ είναι μια μαρκοβιανή αλυσίδα στον χώρο καταστάσεων \mathbb{N}_0 .

α) Περιγράψτε τις πιθανότητες μετάβασης της $\{X_n\}_{n \in \mathbb{N}_0}$.

β) Δείξτε ότι η αναλλοίωτη κατανομή της είναι η $\pi : \mathbb{N}_0 \rightarrow [0, 1]$ με $\pi(0) = 0$ και $\pi(n) = 9 \times 10^{-n}$, $n \in \mathbb{N}$.

γ) Για ένα σερί που ξεκινά τώρα δείξτε ότι η πιθανότητα να φτάσει σε μήκος 9 πριν σπάσει είναι 10^{-8} ενώ το αναμενόμενο μήκος του σερί είναι $10/9$.

δ) Αν N είναι ο αριθμός των ψηφίων που πρέπει να δούμε μέχρι να σχηματιστεί για πρώτη φορά ένα σερί από 9 ίδια ψηφία, δείξτε ότι $\mathbb{E}[N] = \frac{10^9 - 1}{9}$.

ε) Από την απάντηση στα ερωτήματα (γ,δ) φαίνεται ότι κάθε προσπάθεια για τον σχηματισμό ενός σερί από ίδια αποτελέσματα είναι μια δοκιμή Bernoulli με πολύ μικρή πιθανότητα επιτυχίας και χρειάζεται να κάνουμε πολλές τέτοιες προσπάθειες μέχρι να σημειώσουμε την πρώτη επιτυχία. Αυτό σημαίνει ότι η κατανομή της τυχαίας μεταβλητής $E = N/\mathbb{E}[N]$ είναι περίπου εκθετική $\mathcal{E}(1)$ (θεωρήστε το δεδομένο αλλά θυμηθείτε την Άσκηση 55). Με βάση αυτή την παρατήρηση πώς θα κρίνατε τον ισχυρισμό του φίλου σας;

5.7 Αριθμητικά πειράματα

Άσκηση 91 Κατεβάστε το πρόγραμμα `variance.py` και αποθηκεύστε το στον κατάλογο που θα δουλέψετε.

Το πρόγραμμα προσομοιώνει μια αλυσίδα στον χώρο καταστάσεων $\mathbb{X} = \{1, 2, 3, 4, 5\}$ και υπολογίζει με τη μέθοδο Monte Carlo τον αναμενόμενο χρόνο επιστροφής στην κατάσταση 1, $\mathbb{E}[T_1^+ | X_0 = 1]$, όπου

$$T_1^+ = \inf \{k > 0 : X_k = 1\}.$$

Η εκτίμηση για την $\mathbb{E}[T_1^+ | X_0 = 1]$ λαμβάνεται προσομοιώνοντας την αλυσίδα N φορές, παίρνοντας N ανεξάρτητα δείγματα t_1, \dots, t_N του χρόνου επιστροφής στο 1 και παίρνοντας τον μέσο όρο αυτών των δειγμάτων. Ο νόμος των μεγάλων αριθμών εγγυάται ότι ο μέσος όρος αυτών των N ανεξάρτητων δειγμάτων της τυχαίας μεταβλητής T_1^+ είναι για αρκετά μεγάλο N κοντά στην $\mathbb{E}[T_1^+ | X_0 = 1]$ με μεγάλη πιθανότητα. Θα καλούμε την

$$E_N = \frac{t_1 + t_2 + \dots + t_N}{N}$$

εκτιμήτρια Monte Carlo της αναμενόμενης τιμής $\mathbb{E}[T_1^+ | X_0 = 1]$. Φυσικά η E_N είναι μια τυχαία μεταβλητή και αν ξανακάνουμε το πείραμα θα πάρουμε μια διαφορετική εκτίμηση.

α) Τρέξτε το πρόγραμμα μερικές φορές και δείτε πόσο διαφέρουν οι εκτιμήσεις που παίρνουμε κάθε φορά για την $\mathbb{E}[T_1^+ | X_0 = 1]$. Επανάλάβετε για $N = 2^6, 2^7, \dots, 2^{12}$. Φαίνεται να διαφέρουν λιγότερο οι διαφορετικές εκτιμήσεις καθώς μεγαλώνει το N ;

Σκοπός αυτής της άσκησης είναι να βρούμε υπολογιστικά πώς επηρεάζεται η διασπορά της εκτιμήτριας E_N από το πλήθος των επαναλήψεων N .

β) Φτιάξτε έναν βρόχο που θα κάνει $M = 30$ εκτιμήσεις της $\mathbb{E}[T_1^+ | X_0 = 1]$ για κάθε δεδομένη τιμή του N και αποθηκεύστε αυτές τις M εκτιμήσεις στη λίστα `mcestimates`.

γ) Υπολογίστε τη δειγματική μέση τιμή και τη δειγματική διασπορά αυτών των εκτιμήσεων αφαιρώντας τον σχολιασμό από τις εντολές

```
sample_mean = float( sum(mcestimates) ) / M
squared_distance_from_mean = [ (e - sample_mean)**2 for e in mcestimates ]
sample_variance= float(sum( squared_distance_from_mean )) / (M-1)
```

που βρίσκονται στο τέλος του προγράμματος.

δ) Φτιάξτε έναν βρόχο που κάνει τον παραπάνω υπολογισμό για $N = 2^6, 2^7, \dots, 2^{14}$ και παραστήστε γραφικά πώς εξαρτάται η δειγματική μέση τιμή και η δειγματική τυπική απόκλιση από το N .

ε) Υπολογίστε θεωρητικά την $\mathbb{E}[T_1^+ | X_0 = 1]$. Συμφωνεί το αριθμητικό αποτέλεσμα που βρήκατε με τη θεωρητική τιμή;

στ) Σχεδιάστε σε λογαριθμική κλίμακα πώς εξαρτάται η δειγματική τυπική απόκλιση της εκτιμήτριας Monte Carlo από το N . Ποια είναι η κλίση της ευθείας στο λογαριθμικό διάγραμμα; Συμφωνεί αυτό που βρήκατε με το κεντρικό οριακό θεώρημα;

Άσκηση 92 Στην άσκηση αυτή θα επιβεβαιώσετε πειραματικά το Θεώρημα 21. Κατεβάστε τον κώδικα `set06_ex01.py`. Ο κώδικας αυτός υπολογίζει τον αναμενόμενο αριθμό επισκέψεων σε κάθε κατάσταση μιας μαρκοβιανής αλυσίδας στον $\mathbb{X} = \{1, 2, 3, 4\}$ κατά τη διάρκεια μιας εκδρομής γύρω από την κατάσταση 1.

- α) Τροποποιήστε τον κώδικα ώστε να υπολογίζει την αναλλοίωτη κατανομή π_1 του Θεωρήματος 21.
- β) Αλλάξτε τώρα την αρχική κατάσταση ώστε να υπολογίσετε τις κατανομές π_2, π_3, π_4 . Τι παρατηρείτε;
- γ) Υπολογίστε θεωρητικά την αναλλοίωτη κατανομή της αλυσίδας. Συμφωνεί το αριθμητικό αποτέλεσμα με τον θεωρητικό υπολογισμό;

Άσκηση 93 Αν επιλέξουμε τυχαία ένα σημείο s ένα τετράγωνο S πλευράς 2 με κέντρο το σημείο $(0,0)$ η πιθανότητα το σημείο αυτό να πέσει στο μοναδιαίο δίσκο $D = \{(x, y) \in S : x^2 + y^2 < 1\}$ είναι

$$p = \frac{|D|}{|S|} = \frac{\pi}{4}.$$

Ο κώδικας `pi_estimation.py` χρησιμοποιεί αυτή την απλή παρατήρηση για να εκτιμήσει αριθμητικά το π .

- α) Τροποποιήστε τον κώδικα ώστε να υπολογίσετε αριθμητικά το εμβαδό του χωρίου $D_1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq |x|\}$.
- β) Τροποποιήστε τον κώδικα ώστε να υπολογίσετε αριθμητικά τον όγκο της μοναδιαίας σφαίρας σε 3 διαστάσεις και συγκρίνετε με την θεωρητική τιμή $\frac{4\pi}{3}$.
- γ) Τροποποιήστε τον κώδικα ώστε να τυπώνει για διάσταση $d = 2, 3, \dots, 18$ τον όγκο της d -διάστατης μοναδιαίας σφαίρας όπως υπολογίζεται με Monte Carlo, τον πραγματικό όγκο της d -διάστατης μοναδιαίας σφαίρας και το σχετικό (%) σφάλμα της προσέγγισης. Μπορείτε να καταλάβετε γιατί το σχετικό σφάλμα μεγαλώνει καθώς μεγαλώνει η διάσταση d ;

Για το τελευταίο ερώτημα θα χρειαστεί να αφαιρέσετε το σύμβολο του σχολιασμού μπροστά από την εντολή που καλεί τη βιβλιοθήκη `math` και τις εντολές που ορίζουν τη συνάρτηση `volume_sphere`.

Άσκηση 94 Η άσκηση 93 μας δίνει έναν τρόπο να κάνουμε δειγματοληψία από μια τυχαία μεταβλητή με ομοιόμορφη κατανομή στον μοναδιαίο δίσκο D . Επιλέγουμε τυχαία ένα σημείο στο τετράγωνο S . Αν

ανήκει στον D , το κρατάμε. Διαφορετικά, επιλέγουμε τυχαία ένα άλλο σημείο στο S και συνεχίζουμε τη διαδικασία μέχρι να πάρουμε ένα σημείο του D . Αυτή η μέθοδος δειγματοληψίας ονομάζεται μέθοδος απόρριψης (rejection sampling). Ο κώδικας `s2_color.py` επιλέγει με αυτόν τον τρόπο 1.000 τυχαία σημεία στον D και επιστρέφει ένα διάγραμμα με τη θέση αυτών των σημείων (με πράσινο χρώμα) και τη θέση των σημείων που απορρίφθηκαν (με κόκκινο χρώμα).

α) Κατασκευάστε ένα διάγραμμα με τη θέση 10.000 τυχαία επιλεγμένων σημείων στον D . Δώστε τίτλο στο διάγραμμά σας. Πόσα δείγματα στο S χρειάστηκαν για να το κατασκευάσετε;

β) Γράψτε έναν κώδικα που εκτιμά με τη μέθοδο Monte Carlo την τιμή του ολοκληρώματος

$$\iint_D |x + y| \frac{dxdy}{\pi}.$$

(Υπόδειξη: το ολοκλήρωμα αυτό είναι η αναμενόμενη τιμή $\mathbb{E}[|X + Y|]$ όπου το ζεύγος (X, Y) ακολουθεί ομοιόμορφη κατανομή στον D).

γ) Πόσα περίπου δείγματα στον υπερκύβο διάστασης 15 θα χρειαζόμασταν προκειμένου να κρατήσουμε 1.000 τυχαία σημεία στη μοναδιαία σφαίρα διάστασης 15; Θα ήταν πρακτικό να χρησιμοποιήσουμε αυτή τη μέθοδο δειγματοληψίας για να υπολογίσουμε αριθμητικά ένα ολοκλήρωμα στη μοναδιαία σφαίρα διάστασης 15;

Στο επόμενο κεφάλαιο θα δούμε πώς μπορούμε να ξεπεράσουμε αυτό το πρόβλημα χρησιμοποιώντας ιδέες από τις μαρκοβιανές αλυσίδες.

Κεφάλαιο 6

Ασυμπτωτικά Θεωρήματα

6.1 Εισαγωγή

Στο κεφάλαιο αυτό θα μελετήσουμε την ασυμπτωτική συμπεριφορά μαρκοβιανών αλυσίδων, το πώς δηλαδή αυτές συμπεριφέρονται σε βάθος χρόνου. Η κατανόηση της ασυμπτωτικής συμπεριφοράς τους είναι σημαντική για δύο κυρίως λόγους. Αφενός, υπάρχουν συστήματα που μοντελοποιούνται από μαρκοβιανές αλυσίδες και τα οποία, αφού περάσουν μια παροδική φάση, τελικά λειτουργούν σε μια κατάσταση ισορροπίας. Αφετέρου, αν θέλουμε να χρησιμοποιήσουμε ιδέες από τις μαρκοβιανές αλυσίδες για να φτιάξουμε υπολογιστικούς αλγορίθμους, είναι σημαντικό να ξέρουμε τη συμπεριφορά του αλγορίθμου μετά από πολλές επαναλήψεις, να μπορούμε να εκτιμήσουμε πόσο γρήγορα συκλίνει και πώς εξαρτάται το σφάλμα της προσέγγισης από το πλήθος των επαναλήψεων που θα εκτελέσουμε. Παρόμοιο υλικό μπορείτε να βρείτε στις αναφορές [7] και [6].

6.2 Περιοδικότητα

Ας θεωρήσουμε μια αλυσίδα που κινείται στις κορυφές ενός τετραγώνου $AB\Gamma\Delta$. Ξεκινώντας από την κορυφή A , σε κάθε βήμα μεταβαίνει σε μία από τις δύο γειτονικές της κορυφές με πιθανότητα $1/2$. Εύκολα βλέπει κανείς ότι η μοναδική αναλλοίωτη κατανομή αυτής της αλυσίδας είναι η ομοιόμορφη κατανομή $\pi = (1/4, 1/4, 1/4, 1/4)$ και όπως είδαμε στο προηγούμενο κεφάλαιο αυτό είναι μόνο υποψήφιο όριο της κατανομής π_n της αλυσίδας μετά από n βήματα. Δεν είναι δύσκολο να δει κανείς όμως ότι $\pi_n \not\rightarrow \pi$. Πράγματι, εφόσον η αλυσίδα ξεκινά από το A , θα βρίσκεται οπωσδήποτε στο B ή στο Δ μετά από περιττό αριθμό βημάτων και στο A ή στο Γ μετά από άρτιο αριθμό βημάτων. Έτσι, π.χ.

$$\pi_{2n+1}(A) = \mathbb{P}[X_{2n+1} = A | X_0 = A] = 0 \not\rightarrow 1/4.$$

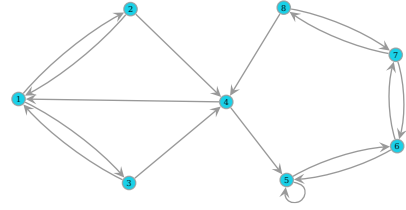
Βλέπουμε λοιπόν ότι στο παράδειγμά μας υπάρχουν υποσύνολα του χώρου καταστάσεων που είναι προσβάσιμα μόνο σε συγκεκριμένες χρονικές στιγμές. Αυτή η ιδιότητα αποτελεί κατά κάποιον τρόπο μια παθολογία που εμποδίζει τη σύγκλιση της κατανομής της αλυσίδας π_n , και σε αυτή την παράγραφο θα προσπαθήσουμε να την κατανοήσουμε.

Ορισμός: Για μια κατάσταση $x \in \mathbb{X}$ ορίζουμε το σύνολο των δυνατών χρόνων επιστροφής στο x

$$R(x) = \{n \in \mathbb{N} : p^n(x, x) > 0\}, \quad \text{όπου θυμίζουμε } p^{(n)}(x, x) = \mathbb{P}[X_n = x | X_0 = x].$$

Θα ονομάζουμε τον μέγιστο κοινό διαιρέτη του συνόλου $R(x)$ περίοδο της κατάστασης x και θα τον συμβολίζουμε με $d(x)$. Στην ειδική περίπτωση που $d(x) = 1$, θα λέμε ότι η κατάσταση x είναι απεριοδική.

Παράδειγμα 47 Στο διπλανό σχήμα φαίνονται οι καταστάσεις και οι δυνατές μεταβάσεις μιας μαρκοβιανής αλυσίδας που κινείται σ' έναν χώρο με οκτώ καταστάσεις. Μια προσανατολισμένη ακμή από το x στο y στον διπλανό γράφο αντιστοιχεί σε θετική πιθανότητα μετάβασης $p(x, y)$. Οι υπόλοιπες μεταβάσεις έχουν μηδενική πιθανότητα. Έχουμε $R(1) = R(2) = R(3) = R(6) = \{2, 3, \dots\}$, $R(4) = \{3, 5, 6, 7, 8, \dots\}$, $R(5) = \mathbb{N}$, $R(7) = R(8) = \{2, 4, 5, 6, \dots\}$. Στο παράδειγμα αυτό έχουμε $d(x) = 1$ για όλες τις καταστάσεις x της αλυσίδας, επομένως όλες οι καταστάσεις είναι απεριοδικές.



Παρατηρήστε ότι, για κάθε $x \in \mathbb{X}$, το σύνολο $R(x)$ είναι ένα υποσύνολο του \mathbb{N} που είναι κλειστό ως προς την πρόσθεση, δηλαδή

$$m, n \in R(x) \Rightarrow m + n \in R(x).$$

Διασθητικά αυτό σημαίνει ότι, αν είναι δυνατόν να επιστρέψουμε στο x μετά από m βήματα, αλλά και μετά από n βήματα, τότε είναι δυνατόν να επιστρέψουμε στο x και μετά από $m + n$ βήματα. Πράγματι, από τις εξισώσεις Chapman-Kolmogorov, έχουμε

$$p^{(m+n)}(x, x) = \sum_{y \in \mathbb{X}} p^{(m)}(x, y)p^{(n)}(y, x) \geq p^{(m)}(x, x)p^{(n)}(x, x) > 0,$$

εφόσον $m, n \in R(x)$. Για τέτοια σύνολα έχουμε το ακόλουθο λήμμα.

Λήμμα 8 Αν ένα σύνολο $R \subset \mathbb{N}$ είναι κλειστό ως προς την πρόσθεση και έχει μέγιστο κοινό διαιρέτη $d \in \mathbb{N}$, τότε υπάρχει $n_0 \in \mathbb{N}$ τέτοιο ώστε $\{n_0d, (n_0 + 1)d, (n_0 + 2)d, \dots\} \subset R$. Το R περιέχει επομένως τελικά όλα τα πολλαπλάσια του d .

Απόδειξη: Θα συμβολίζουμε με $d\mathbb{N}$ το σύνολο των φυσικών που είναι πολλαπλάσια του d , δηλαδή $d\mathbb{N} = \{d, 2d, 3d, \dots\}$. Αν $d \in R$, εφόσον το R είναι ένα κλειστό ως προς την πρόσθεση υποσύνολο του \mathbb{N} , έχουμε $R = d\mathbb{N}$. Αν πάλι $d \notin R$, τότε, εφόσον το d είναι ο μέγιστος κοινός διαιρέτης του R , θα υπάρχουν $m_1, \dots, m_k \in R$ και $p_1, p_2, \dots, p_k \in \mathbb{Z} \setminus \{0\}$ τέτοια ώστε

$$\sum_{i=1}^k p_i m_i = d. \quad (6.1)$$

Έστω $I \subset \{1, 2, \dots, k\}$ το σύνολο των δεικτών i για τους οποίους $p_i > 0$ και $I' = \{1, 2, \dots, k\} \setminus I$ το σύνολο των δεικτών i για τους οποίους $p_i < 0$. Προφανώς, το I δεν μπορεί να είναι κενό αφού το δεξί μέλος της (6.1) είναι φυσικός αριθμός. Ούτε το I' μπορεί να είναι κενό όμως, αφού κάθε προσθετέος στο αριστερό μέλος της (6.1) είναι μεγαλύτερος ή ίσος από $2d$ κατ' απόλυτη τιμή στην περίπτωση ($d \notin R$) που εξετάζουμε. Μπορούμε επομένως να ξαναγράψουμε την (6.1) ως

$$\sum_{i \in I} p_i m_i - \sum_{i \in I'} (-p_i) m_i = d.$$

Εφόσον το R είναι κλειστό ως προς την πρόσθεση και $p_i > 0$ για $i \in I$ έχουμε ότι $m = \sum_{i \in I} p_i m_i \in R$. Αντίστοιχα, έχουμε ότι $n = \sum_{i \in I'} (-p_i) m_i \in R$. Βρήκαμε επομένως δύο στοιχεία $m, n \in R$ τέτοια ώστε $m = n + d$. Θα δείξουμε ότι το R περιέχει κάθε φυσικό της μορφής Nd για $Nd \geq n(n-1)$. Πράγματι, αν γράψουμε την ταυτότητα της ακέριας διαίρεσης του N με το n έχουμε

$$N = qn + v, \quad \text{για κάποια } q \geq \frac{n-1}{d}, v \in \{0, 1, 2, \dots, n-1\}. \quad (6.2)$$

Έχουμε τώρα

$$Nd = qnd + vd = qnd + v(m - n) = vm + (qd - v)n.$$

Οι v και $qd - v$ είναι μη αρνητικοί ακέραιοι λόγω της (6.2) και τουλάχιστον ένας από αυτούς είναι διάφορος από το μηδέν αφού έχουν άθροισμα qd . Εφόσον $m, n \in R$ και το R είναι κλειστό ως προς την πρόσθεση, θα έχουμε $Nd \in R$ για κάθε φυσικό $N \geq n(n - 1)/d$. \square

Θα χρησιμοποιήσουμε τώρα το Λήμμα 8 για να δείξουμε ότι η περίοδος είναι χαρακτηριστικό κλάσης.

Θεώρημα 26 Αν δύο κατάστασεις ανήκουν στην ίδια κλάση επικοινωνίας, τότε έχουν την ίδια περίοδο.

Απόδειξη: Αν οι $x, y \in \mathbb{X}$ επικοινωνούν αμφίδρομα, υπάρχουν $m, n \in \mathbb{N}$ τέτοια ώστε

$$p^{(m)}(x, y) > 0 \quad \text{και} \quad p^{(n)}(y, x) > 0.$$

Επιπλέον, αν $\ell \in R(x)$, τότε θα έχουμε $m + n + \ell \in R(y)$. Πράγματι,

$$p^{(m+n+\ell)}(y, y) = \sum_{z, w \in \mathbb{X}} p^{(n)}(y, z)p^{(\ell)}(z, w)p^{(m)}(w, y) \geq p^{(n)}(y, x)p^{(\ell)}(x, x)p^{(m)}(x, y) > 0.$$

Από το Λήμμα 8, υπάρχει κάποιο $q \in \mathbb{N}$ τέτοιο ώστε $qd(x), (q + 1)d(x) \in R(x)$. Από την προηγούμενη παρατήρηση, θα έχουμε λοιπόν

$$m + n + qd(x) \in R(y) \quad \text{και} \quad m + n + (q + 1)d(x) \in R(y).$$

Επομένως, $d(y)/m + n + qd(x)$ και $d(y)/m + n + (q + 1)d(x)$ και άρα $d(y)/d(x)$. Εναλλάσσοντας τον ρόλο των x, y στο παραπάνω επιχείρημα, παίρνουμε ότι $d(x)/d(y)$ και άρα $d(x) = d(y)$. \square

Υπό το φως του Θεωρήματος 26 έχει νόημα να μιλάμε για την περίοδο μιας κλάσης ή για την περίοδο μιας μη υποβιβάσιμης αλυσίδας και να χαρακτηρίζουμε μια τέτοια αλυσίδα απεριοδική όταν η περίοδος όλων των καταστάσεων της είναι 1. Το ακόλουθο πόρισμα είναι συχνά χρήσιμο γιατί δίνει μια ικανή συνθήκη ώστε να είναι μια αλυσίδα απεριοδική που μπορεί να ελεγχθεί με απλή επισκόπηση.

Πόρισμα 10 Αν μια μη υποβιβάσιμη αλυσίδα έχει μια κατάσταση x για την οποία $p(x, x) > 0$, τότε η αλυσίδα είναι απεριοδική.

Απόδειξη: Εφόσον $p(x, x) > 0$ έχουμε $1 \in R(x)$ και άρα $d(x) = 1$. Επειδή η αλυσίδα είναι μη υποβιβάσιμη, όλες οι καταστάσεις της θα έχουν την ίδια περίοδο με την x , δηλαδή 1. \square

Το ακόλουθο λήμμα θα μας φανεί επίσης χρήσιμο στις επόμενες παραγράφους.

Λήμμα 9 Αν η $\{X_n\}_{n \in \mathbb{N}_0}$ είναι μια μη υποβιβάσιμη και απεριοδική αλυσίδα σε έναν χώρο καταστάσεων \mathbb{X} με πιθανότητες μετάβασης $p(\cdot, \cdot)$, τότε για κάθε $x, y \in \mathbb{X}$ υπάρχει $n_0 \in \mathbb{N}$ τέτοιο ώστε $p^{(n)}(x, y) > 0$, για κάθε $n \geq n_0$.

Απόδειξη: Εφόσον $d(x) = 1$ από το Λήμμα 8 υπάρχει ένα $n_1 \in \mathbb{N}$ τέτοιο ώστε $p^{(n)}(x, x) > 0$, για κάθε $n \geq n_1$. Εφόσον η $\{X_n\}_n$ είναι μη υποβιβάσιμη υπάρχει ακόμη ένα $m \in \mathbb{N}$ τέτοιο, ώστε $p^{(m)}(x, y) > 0$. Αν τώρα $n \geq n_1 + m$, έχουμε

$$p^{(n)}(x, y) = \sum_{z \in \mathbb{X}} p^{(n-m)}(x, z)p^{(m)}(z, y) \geq p^{(n-m)}(x, x)p^{(m)}(x, y) > 0.$$

□

Προσέξτε ότι στο παράδειγμα που είδαμε στην αρχή της παραγράφου, για το οποίο η κατανομή π_n της αλυσίδας μετά από n βήματα δεν συγκλίνει στην αναλλοίωτη κατανομή της αλυσίδας, η περίοδος της αλυσίδας είναι 2. Ο λόγος που η σύγκλιση αποτυγχάνει είναι γιατί αν μια μη υποβιβάσιμη αλυσίδα έχει περίοδο d , ο χώρος καταστάσεων διαμερίζεται σε d σύνολα $\mathbb{X} = \cup_{j=0}^{d-1} C_j$, σε καθένα από τα οποία η αλυσίδα μπορεί να επιστρέψει μόνο μετά από ένα αριθμό βημάτων που είναι πολλαπλάσιο του d . Δηλαδή,

$$\mathbb{P}[X_n \in C_j | X_0 \in C_j] \neq 0 \Leftrightarrow d/n.$$

Πιο συγκεκριμένα έχουμε το ακόλουθο θεώρημα, του οποίου η απόδειξη αφήνεται σαν άσκηση.

Θεώρημα 27 Έστω $\{X_n\}_{n \in \mathbb{N}_0}$ μια μη υποβιβάσιμη αλυσίδα σε έναν χώρο καταστάσεων $\mathbb{X} \ni x$ με πιθανότητες μετάβασης $p(\cdot, \cdot)$ και περίοδο d . Τότε η αλυσίδα $\{Y_n\}_{n \in \mathbb{N}_0}$ με $Y_n = X_{nd}$ (δηλαδή η αλυσίδα που προκύπτει αν δειγματοληπούμε την $\{X_n\}_n$ κάθε d βήματα) είναι υποβιβάσιμη στις d αperiοδικές κλάσεις

$$C_v = \{y \in \mathbb{X} : p^{(dn+v)}(x, y) > 0, \text{ για κάποιο } n \in \mathbb{N}_0\}, \quad v \in 0, 1, 2, \dots, d-1.$$

Αν η αλυσίδα $\{X_n\}_{n \in \mathbb{N}}$ είναι γνησίως επαναληπτική με αναλλοίωτη κατανομή π , τότε η αναλλοίωτη κατανομή π_v της αλυσίδας $\{Y_n\}_{n \in \mathbb{N}_0}$ που στηρίζεται στην κλάση C_v δίνεται από την

$$\pi_v(y) = \begin{cases} d\pi(y), & \text{αν } y \in C_v \\ 0, & \text{αν } y \notin C_v. \end{cases} \quad (6.3)$$

Στην παράγραφο 6.4 θα δούμε ότι αν μια μη υποβιβάσιμη και γνησίως επαναληπτική αλυσίδα είναι αperiοδική, τότε σε βάθος χρόνου η κατανομή της αλυσίδας π_n θα συγκλίνει στην αναλλοίωτη κατανομή της. Αυτό είναι ένα πολύ χρήσιμο αποτέλεσμα, γιατί εκτός του ότι περιγράφει την ασυμπτωτική συμπεριφορά της αλυσίδας, προσφέρει και έναν υπολογιστικό τρόπο για να πάρουμε δείγματα από μια κατανομή π . Χρειάζεται απλώς να κατασκευάσουμε μια μη υποβιβάσιμη, γνησίως επαναληπτική και αperiοδική αλυσίδα που έχει αναλλοίωτη κατανομή την π και να αφήσουμε αυτή την αλυσίδα να κάνει αρκετά βήματα ώστε η κατανομή της να πλησιάσει την π . Αυτή η ιδέα είναι η βάση της μεθόδου Markov Chain Monte Carlo (MCMC), που θα μελετήσουμε εκτενέστερα στο Κεφάλαιο 7.

Για την απόδειξη της σύγκλισης $\pi_n \rightarrow \pi$ θα χρησιμοποιήσουμε μια πολύ ισχυρή πιθανοθεωρητική τεχνική, τη σύζευξη (coupling). Λόγω της χρησιμότητας αυτής της τεχνικής θα αφιερώσουμε την επόμενη παράγραφο στο να την περιγράψουμε μέσα από δύο παραδείγματα.

6.3 Σύζευξη

Η σύζευξη είναι μια πολύ γενική μέθοδος με την οποία μπορούμε να πάρουμε χρήσιμα αποτελέσματα για την κατανομή τυχαίων μεταβλητών, ορισμένων εν γένει σε διαφορετικούς χώρους πιθανότητας, θεωρώντας έναν μεγαλύτερο χώρο πιθανότητας στον οποίο μπορούμε να ορίσουμε ταυτόχρονα τις τυχαίες μεταβλητές χωρίς να αλλάξουμε την κατανομή τους. Πιο συγκεκριμένα, αν η τυχαία μεταβλητή X έχει οριστεί στον χώρο πιθανότητας $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ και η τυχαία μεταβλητή Y έχει οριστεί στον χώρο πιθανότητας $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ μπορούμε να ορίσουμε τις \tilde{X}, \tilde{Y} στον κοινό δειγματικό χώρο $\Omega = \Omega_1 \times \Omega_2$ ως εξής

$$\tilde{X}(\omega_1, \omega_2) = X(\omega_1), \quad \tilde{Y}(\omega_1, \omega_2) = Y(\omega_2).$$

Αν εφοδιάσουμε τον Ω με ένα μέτρο πιθανότητας \mathbb{P} που, περιορισμένο στο Ω_i , ταυτίζεται με το \mathbb{P}_i για $i \in \{1, 2\}$, αν δηλαδή

$$\mathbb{P}[A_1 \times \Omega_2] = \mathbb{P}_1[A_1] \quad \text{και} \quad \mathbb{P}[\Omega_1 \times A_2] = \mathbb{P}_2[A_2], \quad (6.4)$$

για κάθε $A_i \subset \Omega_i$ στο πεδίο ορισμού \mathcal{F}_i του \mathbb{P}_i , τότε οι \tilde{X}, \tilde{Y} , που είναι από κοινού ορισμένες στον Ω , έχουν την ίδια κατανομή με τις X και Y , αντίστοιχα. Πράγματι,

$$\mathbb{P}[\tilde{X} \in C] = \mathbb{P}[\{X \in C\} \times \Omega_2] = \mathbb{P}_1[X \in C]$$

και αντίστοιχα για τις \tilde{Y} και Y . Ένα μέτρο πιθανότητας \mathbb{P} που ικανοποιεί τις (6.4) ονομάζεται μέτρο σύζευξης των \mathbb{P}_1 και \mathbb{P}_2 .

Παρατηρήστε ότι ενώ οι σχέσεις (6.4) που επιβάλλουμε στο \mathbb{P} καθορίζουν την κατανομή καθεμιάς από τις \tilde{X} και \tilde{Y} , δεν μας δίνουν καμιά πληροφορία για τον τρόπο που οι \tilde{X} και \tilde{Y} συνδέονται μεταξύ τους. Η τεχνική της σύζευξης συνίσταται στο να επιλέξουμε το μέτρο σύζευξης \mathbb{P} με τρόπο ώστε η από κοινού κατανομή των \tilde{X}, \tilde{Y} κάτω από το \mathbb{P} να μας οδηγεί σε χρήσιμα συμπεράσματα.

Ας δούμε πώς μπορούμε να εφαρμόσουμε το παραπάνω γενικό σχέδιο μέσα από δύο συγκεκριμένα παραδείγματα.

Παράδειγμα 48 Έστω X τυχαία μεταβλητή με τιμές στο \mathbb{R} . Αν οι συναρτήσεις $f, g : \mathbb{R} \rightarrow \mathbb{R}$ είναι και οι δύο αύξουσες ή και οι δύο φθίνουσες, τότε

$$\mathbb{E}[f(X)g(X)] \geq \mathbb{E}[f(X)]\mathbb{E}[g(X)]$$

οποτεδήποτε οι παραπάνω αναμενόμενες τιμές είναι καλά ορισμένες.

Σε αυτό το σημείο αξίζει να δοκιμάσετε να αποδείξετε τον ισχυρισμό του παραδείγματος χωρίς να χρησιμοποιήσετε την τεχνική της σύζευξης. Θα δείτε ότι δεν είναι καθόλου εύκολο. Θεωρήστε τώρα μια άλλη τυχαία μεταβλητή Y με την ίδια κατανομή όπως η X και ανεξάρτητη από τη X . Παρατηρήστε ότι, εφόσον οι f, g έχουν το ίδιο είδος μονοτονίας, τότε για κάθε $x, y \in \mathbb{R}$ έχουμε

$$(f(x) - f(y))(g(x) - g(y)) \geq 0$$

και επομένως η τυχαία μεταβλητή $(f(X) - f(Y))(g(X) - g(Y))$ παίρνει μόνο μη αρνητικές τιμές. Ειδικότερα,

$$\begin{aligned} 0 &\leq \mathbb{E}[(f(X) - f(Y))(g(X) - g(Y))] \\ &= \mathbb{E}[f(X)g(X)] + \mathbb{E}[f(Y)g(Y)] - \mathbb{E}[f(X)g(Y)] - \mathbb{E}[f(Y)g(X)] \\ &= \mathbb{E}[f(X)g(X)] + \mathbb{E}[f(Y)g(Y)] - \mathbb{E}[f(X)]\mathbb{E}[g(Y)] - \mathbb{E}[f(Y)]\mathbb{E}[g(X)] \\ &= 2(\mathbb{E}[f(X)g(X)] - \mathbb{E}[f(X)]\mathbb{E}[g(X)]), \end{aligned}$$

όπου η προτελευταία ισότητα ισχύει επειδή οι X, Y είναι ανεξάρτητες, ενώ η τελευταία ισότητα προκύπτει από το ότι οι X, Y έχουν την ίδια κατανομή.

Πού ακριβώς χρησιμοποιήσαμε τη σύζευξη στο παραπάνω επιχείρημα; Λέγοντας ‘θεωρήστε τώρα μια άλλη τυχαία μεταβλητή Y με την ίδια κατανομή όπως η X και ανεξάρτητη από τη X ’ αποσιωπήσαμε πώς μπορεί κανείς να ορίσει στον ίδιο χώρο πιθανότητας δύο τυχαίες μεταβλητές με δεδομένη κατανομή ώστε αυτές να είναι ανεξάρτητες. Αυτό γίνεται ακριβώς με την τεχνική της σύζευξης. Με τον συμβολισμό που περιγράψαμε πριν, αν η X είναι ορισμένη σε έναν χώρο πιθανότητας $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ και η Y είναι ορισμένη σε έναν χώρο πιθανότητας $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$, εφοδιάζουμε τον χώρο γινόμενο $\Omega_1 \times \Omega_2$ με το μέτρο γινόμενο $\mathbb{P} = \mathbb{P}_1 \times \mathbb{P}_2$, που ικανοποιεί την

$$\mathbb{P}[A_1 \times A_2] = \mathbb{P}_1[A_1]\mathbb{P}_2[A_2],$$

για κάθε $A_i \subset \Omega_i$ στο πεδίο ορισμού \mathcal{F}_i του \mathbb{P}_i . Το \mathbb{P} είναι μέτρο σύζευξης αφού $\mathbb{P}_i[\Omega_i] = 1$, επομένως οι \tilde{X} και \tilde{Y} που είναι ορισμένες στον Ω έχουν την κατανομή των X και Y , αντίστοιχα. Επιπλέον το \mathbb{P} κάνει τις \tilde{X} και \tilde{Y} ανεξάρτητες, αφού

$$\mathbb{P}[\tilde{X} \in C_1, \tilde{Y} \in C_2] = \mathbb{P}[\{X \in C_1\} \times \{Y \in C_2\}] = \mathbb{P}_1[X \in C_1]\mathbb{P}_2[Y \in C_2] = \mathbb{P}[\tilde{X} \in C_1]\mathbb{P}[\tilde{Y} \in C_2].$$

Κατασκευάσαμε λοιπόν ανεξάρτητες τυχαίες μεταβλητές \tilde{X}, \tilde{Y} , ορισμένες σε έναν κοινό χώρο πιθανότητας, με την ίδια κατανομή όπως οι X και Y αντίστοιχα. Έχοντας πλέον δει πώς γίνεται αυτό, μπορούμε χάριν απλότητας να συμβολίζουμε τις \tilde{X}, \tilde{Y} με X και Y αντίστοιχα.

Παρατηρήστε ότι κάθε μέτρο σύζευξης δίνει στις \tilde{X}, \tilde{Y} την κατανομή των X, Y αντίστοιχα. Επιλέγοντας να εφοδιάσουμε τον Ω με το μέτρο γινόμενο \mathbb{P} καθορίσαμε επιπλέον πώς αυτές συνδέονται μεταξύ τους: είναι ανεξάρτητες. Στη συνέχεια χρησιμοποιήσαμε την ανεξαρτησία τους στην πορεία της απόδειξης.

Παράδειγμα 49 Θεωρήστε έναν συνεκτικό γράφο G με σύνολο κορυφών V ($|V| < +\infty$) και σύνολο ακμών E . Για κάθε ακμή του γράφου και ανεξάρτητα από τις άλλες ακμές, αποφασίζουμε αν θα τη διαγράψουμε με πιθανότητα $p \in (0, 1)$. Έστω C το ενδεχόμενο ο γράφος να παραμένει συνεκτικός μετά τις διαγραφές. Θα δείξουμε ότι η πιθανότητα του C είναι φθίνουσα συνάρτηση του p .

Είναι διαισθητικά αναμενόμενο ότι όσο πιο πιθανό είναι να διαγράψουμε κάθε ακμή τόσο λιγότερο πιθανό είναι το να παραμείνει συνεκτικός ο γράφος μετά τις διαγραφές. Παρ' όλα αυτά δεν είναι εύκολο να αποδείξει κανείς αυστηρά αυτόν τον ισχυρισμό χωρίς να χρησιμοποιήσει την ιδέα της σύζευξης. Ο λόγος είναι ότι, αν χρησιμοποιήσει κανείς ένα κέρμα που έχει πιθανότητα να φέρει κεφαλή p_1 και το στρίψει για κάθε ακμή του γράφου για να αποφασίσει αν θα τη διαγράψει και στη συνέχεια κάνει το ίδιο με ένα κέρμα που έχει πιθανότητα να φέρει κεφαλή $p_2 > p_1$, είναι δυνατόν τα αποτελέσματα των στριψιμάτων να είναι τέτοια ώστε στην πρώτη περίπτωση να καταλήξει με έναν μη συνεκτικό γράφο, ενώ στη δεύτερη με έναν συνεκτικό γράφο. Θα χρησιμοποιήσουμε την ιδέα της σύζευξης ώστε να κάνουμε τις διαγραφές με πιθανότητα p_1 και p_2 συντονισμένα και με τρόπο ώστε οι ακμές που διαγράφουμε στην πρώτη περίπτωση να είναι υποσύνολο εκείνων που διαγράφουμε στη δεύτερη.

Ο γράφος του πειράματος είναι μια τυχαία μεταβλητή G ορισμένη στον δειγματικό χώρο $\Omega = \{0, 1\}^E$ με τιμές στο σύνολο των γράφων που έχουν κορυφές τα σημεία του V και ακμές κάποιο υποσύνολο του E . Πράγματι, το $\omega = \{\omega_e\}_{e \in E} \in \Omega$ αντιστοιχεί στον γράφο $G(\omega)$ με σύνολο ακμών το $E(\omega) = \{e \in E : \omega_e = 1\}$. Προκειμένου να διαγράψουμε κάθε ακμή ανεξάρτητα με πιθανότητα p , εφοδιάζουμε τον Ω με το μέτρο γινόμενο \mathbb{P}_p , κάτω από το οποίο οι τυχαίες μεταβλητές $\{X_e\}_{e \in E}$ με $X_e(\omega) = \omega_e$ είναι ανεξάρτητες και ισόνομες με κατανομή Bernoulli(1-p).

Αν στρίψουμε δύο διαφορετικά κέρματα με πιθανότητες να φέρουν κεφαλή p_1, p_2 αντίστοιχα ($p_1 < p_2$), εφοδιάζουμε τον $\Omega \times \Omega$ με το μέτρο σύζευξης $\mathbb{P}_{p_1} \times \mathbb{P}_{p_2}$ κάτω από το οποίο οι $\tilde{X}_e(\omega, \omega') = \omega_e$ και $\tilde{Y}_e(\omega, \omega') = \omega'_e$ είναι ανεξάρτητες. Αυτό το μέτρο όμως επιτρέπει σε μια ακμή να διαγραφεί από τον $G(\omega)$, όχι όμως και από τον $G(\omega')$. Πράγματι,

$$\mathbb{P}_{p_1} \times \mathbb{P}_{p_2}[\omega_e = 0, \omega'_e = 1] = p_1(1 - p_2) > 0.$$

Για να εξασφαλίσουμε ότι όλες οι ακμές του $G(\omega')$ περιέχονται στις ακμές του $G(\omega)$ θα κατασκευάσουμε ένα μέτρο σύζευξης \mathbb{P} των \mathbb{P}_{p_1} και \mathbb{P}_{p_2} στον χώρο γινόμενο $\Omega \times \Omega$ τέτοιο ώστε

$$\mathbb{P}[\tilde{Y}_e \leq \tilde{X}_e \text{ για κάθε } e \in E] = 1. \quad (6.5)$$

Δεν είναι δύσκολο να δει κανείς ότι για να είναι το \mathbb{P} μέτρο σύζευξης των \mathbb{P}_{p_1} και \mathbb{P}_{p_2} θα πρέπει οι $\{\tilde{X}_e\}_{e \in E}$ να είναι ανεξάρτητες τυχαίες μεταβλητές με κατανομή Bernoulli(1 - p_1) και οι $\{\tilde{Y}_e\}_{e \in E}$ να είναι ανεξάρτητες τυχαίες μεταβλητές με κατανομή Bernoulli(1 - p_2).

Μπορούμε να κατασκευάσουμε ένα μέτρο \mathbb{P} στον $\Omega \times \Omega$ με τις παραπάνω ιδιότητες ως εξής. Αρχικά εφοδιάζουμε το $\{0, 1\} \times \{0, 1\}$ με το μέτρο πιθανότητας \mathbb{Q} για το οποίο

$$\mathbb{Q}[\{(0, 0)\}] = p_1, \quad \mathbb{Q}[\{(1, 0)\}] = p_2 - p_1, \quad \mathbb{Q}[\{(1, 1)\}] = 1 - p_2, \quad \text{και } \mathbb{Q}[\{(0, 1)\}] = 0.$$

Στη συνέχεια ορίζουμε για $(\omega, \omega') \in \Omega \times \Omega$

$$\mathbb{P}[\{(\omega, \omega')\}] = \prod_{e \in E} \mathbb{Q}[\{(\omega_e, \omega'_e)\}].$$

Παρατηρήστε ότι κάτω από το μέτρο \mathbb{P} οι $\{(\tilde{X}_e, \tilde{Y}_e)\}_{e \in E}$ είναι μια ακολουθία από ανεξάρτητες ισόνομες τυχαίες μεταβλητές με κατανομή \mathbb{Q} . Εφόσον $\mathbb{P}[\tilde{X}_e = 1] = \mathbb{Q}[\{(1, 0), (1, 1)\}] = 1 - p_1$ έχουμε ότι οι $\{\tilde{X}_e\}_{e \in E}$ είναι ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με κατανομή $\text{Bernoulli}(1 - p_1)$. Αντίστοιχα, εφόσον $\mathbb{P}[\tilde{Y}_e = 1] = \mathbb{Q}[\{(0, 1), (1, 1)\}] = 1 - p_2$, οι $\{\tilde{Y}_e\}_{e \in E}$ είναι ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με κατανομή $\text{Bernoulli}(1 - p_2)$. Επομένως το \mathbb{P} είναι πράγματι μέτρο σύζευξης. Κάτω από το \mathbb{P} οι \tilde{X}_e και \tilde{Y}_e δεν είναι μεταξύ τους ανεξάρτητες, αλλά για κάθε $e \in E$ έχουμε

$$\mathbb{P}[\tilde{Y}_e \leq \tilde{X}_e] = \mathbb{Q}[\{(0, 0), (1, 0), (1, 1)\}] = p_1 + (p_2 - p_1) + (1 - p_2) = 1.$$

Επιπλέον, λόγω της ανεξαρτησίας των $\{(\tilde{X}_e, \tilde{Y}_e)\}_{e \in E}$, έχουμε ότι

$$\mathbb{P}[\tilde{Y}_e \leq \tilde{X}_e \text{ για κάθε } e \in E] = \mathbb{P}\left[\bigcap_{e \in E} \{\tilde{Y}_e \leq \tilde{X}_e\}\right] = \prod_{e \in E} \mathbb{P}[\tilde{Y}_e \leq \tilde{X}_e] = 1,$$

οπότε το μέτρο \mathbb{P} πράγματι ικανοποιεί την (6.5). Θα δείξουμε τώρα ότι αν υπάρχει μέτρο σύζευξης \mathbb{P} των \mathbb{P}_{p_1} και \mathbb{P}_{p_2} που ικανοποιεί την (6.5), τότε $\mathbb{P}_{p_1}[C] \geq \mathbb{P}_{p_2}[C]$.

Αν ορίσουμε την τυχαία μεταβλητή

$$F(\omega) = \begin{cases} 1, & \text{αν ο } G(\omega) \text{ είναι συνεκτικός} \\ 0, & \text{αν ο } G(\omega) \text{ δεν είναι συνεκτικός,} \end{cases}$$

έχουμε προφανώς

$$\mathbb{P}_p[C] = \mathbb{P}_p[\{\omega \in \Omega : G(\omega) \text{ συνεκτικός}\}] = \mathbb{E}_p[F].$$

Επιπλέον αν $\omega, \omega' \in \Omega$ και $\omega'_e \leq \omega_e$ για κάθε $e \in E$, τότε το σύνολο των ακμών του $G(\omega')$ είναι υποσύνολο των ακμών του $G(\omega)$ και άρα $F(\omega') \leq F(\omega)$. Ορίζουμε τέλος τις συναρτήσεις F_1, F_2 στον $\Omega \times \Omega$ με τύπους $F_1(\omega, \omega') = F(\omega)$ και $F_2(\omega, \omega') = F(\omega')$. Παρατηρήστε ότι, αν το μέτρο \mathbb{P} ικανοποιεί την (6.5), τότε

$$\mathbb{P}[F_1 \geq F_2] = \mathbb{P}[\{(\omega, \omega') \in \Omega \times \Omega : F(\omega) \geq F(\omega')\}] \geq \mathbb{P}[\omega'_e \leq \omega_e, \text{ για κάθε } e \in E] = 1. \quad (6.6)$$

Έχουμε τώρα

$$\mathbb{P}_{p_1}[C] = \mathbb{E}_{p_1}[F] = \mathbb{E}[F_1] \geq \mathbb{E}[F_2] = \mathbb{E}_{p_2}[F] = \mathbb{P}_{p_2}[C],$$

όπου η δεύτερη και η προτελευταία ισότητα προκύπτουν από το γεγονός ότι το \mathbb{P} είναι μέτρο σύζευξης, ενώ η τρίτη ισότητα από την (6.6).

Θα κλείσουμε αυτή την παράγραφο με μια παρατήρηση. Στην απόδειξη που δώσαμε για τον ισχυρισμό του Παραδείγματος 49, χρησιμοποιήσαμε συμβολισμό συμβατό με αυτόν της εισαγωγής της παραγράφου, κατασκευάζοντας ένα κατάλληλο μέτρο σύζευξης στον χώρο γινόμενο $\{0, 1\}^E \times \{0, 1\}^E$. Διαβάζοντας όμως προσεκτικά την απόδειξη διαπιστώνουμε ότι είναι αρκετό να κατασκευάσουμε τυχαίες μεταβλητές $\{X_e, Y_e\}_{e \in E}$ σε έναν κοινό χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbb{P})$ έτσι ώστε: α) οι $\{X_e\}_{e \in E}$ να είναι ανεξάρτητες τυχαίες μεταβλητές με κατανομή $\text{Bernoulli}(1 - p_1)$, β) οι $\{Y_e\}_{e \in E}$ να είναι ανεξάρτητες τυχαίες μεταβλητές με κατανομή $\text{Bernoulli}(1 - p_2)$ και γ) $\mathbb{P}[Y_e \leq X_e \text{ για κάθε } e \in E] = 1$.

Αυτό μπορούμε να το πετύχουμε με τον εξής απλούστερο και ενδεχομένως πιο διδακτικό τρόπο. Θεωρούμε έναν χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbb{P})$ στον οποίο μπορούμε να ορίσουμε μια ακολουθία από ανεξάρτητες,

ισόνομες τυχαίες μεταβλητές $\{U_e\}_{e \in E}$ με ομοιόμορφη κατανομή στο $(0,1)$. Για κάθε $e \in E$ ορίζουμε τις X_e, Y_e με τη βοήθεια της U_e ως εξής

$$X_e = \mathbb{1}\{U_e > p_1\} \quad \text{και} \quad Y_e = \mathbb{1}\{U_e > p_2\}.$$

Λόγω της ανεξαρτησίας των $\{U_e\}_{e \in E}$ οι $\{(X_e, Y_e)\}_{e \in E}$ είναι ανεξάρτητες. Επιπλέον

$$\mathbb{P}[X_e = 1] = \mathbb{P}[U_e > 1 - p_1] = 1 - p_1 \quad \text{και} \quad \mathbb{P}[Y_e = 1] = \mathbb{P}[U_e > 1 - p_2] = 1 - p_2,$$

άρα πράγματι οι (α) και (β) παραπάνω ικανοποιούνται. Τέλος η (γ) είναι προφανής από την κατασκευή των X_e, Y_e αφού $Y_e = 1 \Leftrightarrow U_e > p_2 \Rightarrow X_e = 1$.

6.4 Ασυμπτωτική κατανομή

Η παράγραφος αυτή είναι αφιερωμένη στην απόδειξη του παρακάτω Θεωρήματος.

Θεώρημα 28 Έστω $\{X_n\}_{n \in \mathbb{N}_0}$ μια μη υποβιβάσιμη, γνησίως επαναληπτική και απεριοδική αλυσίδα στον χώρο καταστάσεων \mathbb{X} . Αν $\pi_n(x) = \mathbb{P}[X_n = x]$ για $x \in \mathbb{X}$, είναι η κατανομή της αλυσίδας μετά από $n \in \mathbb{N}_0$ βήματα και π είναι η αναλλοίωτη κατανομή της αλυσίδας, τότε

$$\lim_{n \rightarrow \infty} \pi_n = \pi.$$

Ειδικότερα,

$$\lim_{n \rightarrow \infty} p^{(n)}(x, y) = \pi(y), \quad \text{για κάθε } x \in \mathbb{X}.$$

Απόδειξη: Η απόδειξη βασίζεται στην τεχνική της σύζευξης. Όπως είδαμε στην παράγραφο της σύζευξης, μπορούμε να κατασκευάσουμε στον ίδιο χώρο πιθανότητας την αλυσίδα $\{X_n\}_{n \in \mathbb{N}_0}$ και μια άλλη ανεξάρτητη αλυσίδα $\{Y_n\}_{n \in \mathbb{N}}$ με αρχική κατανομή π και τις ίδιες πιθανότητες μετάβασης όπως η $\{X_n\}_{n \in \mathbb{N}_0}$. Ορίζουμε τον χρόνο

$$T = \inf\{k \geq 0 : X_k = Y_k\}$$

της πρώτης συνάντησης των δύο ανεξάρτητων αλυσίδων. Το υπόλοιπο της απόδειξης μπορεί να χωριστεί σε μια σειρά από βήματα.

Βήμα 1: Θα δείξουμε ότι $\mathbb{P}[T < +\infty] = 1$.

Στον χώρο καταστάσεων $\mathbb{X} \times \mathbb{X}$ θεωρούμε τη στοχαστική διαδικασία $\{W_n\}_{n \in \mathbb{N}_0}$ με $W_n = (X_n, Y_n)$. Είναι εύκολο να δει κανείς ότι αυτή είναι μια μαρκοβιανή αλυσίδα με αρχική κατανομή $\tilde{\pi}_0$ όπου

$$\tilde{\pi}_0((x, u)) = \mathbb{P}[W_0 = (x, u)] = \mathbb{P}[X_0 = x, Y_0 = u] = \pi_0(x)\pi(u)$$

και πιθανότητες μετάβασης

$$\begin{aligned} \tilde{p}((x, u), (y, v)) &= \mathbb{P}[W_{n+1} = (y, v) \mid W_n = (x, u)] = \mathbb{P}[X_{n+1} = y, Y_{n+1} = v \mid X_n = x, Y_n = u] \\ &= \mathbb{P}[X_{n+1} = y \mid X_n = x] \mathbb{P}[Y_{n+1} = v \mid Y_n = u] = p(x, y)p(u, v). \end{aligned}$$

Επιπλέον η $\{W_n\}_n$ είναι μη υποβιβάσιμη. Πράγματι, εφόσον η $\{X_n\}$ είναι απεριοδική, από το Λήμμα 9 για κάθε $x, y, u, v \in \mathbb{X}$ υπάρχει ένα n_0 τέτοιο ώστε

$$n \geq n_0 \Rightarrow p^{(n)}(x, y) > 0 \quad \text{και} \quad p^{(n)}(u, v) > 0.$$

Έτσι, αν $n \geq n_0$, έχουμε

$$\begin{aligned}\tilde{p}^{(n)}((x, u), (y, v)) &= \mathbb{P}[W_n = (y, v) | W_0 = (x, u)] = \mathbb{P}[X_n = y, Y_n = v | X_0 = x, Y_0 = u] \\ &= \mathbb{P}[X_n = y | X_0 = x] \mathbb{P}[Y_n = v | Y_0 = u] = p^{(n)}(x, y) p^{(n)}(u, v) > 0,\end{aligned}$$

επομένως οποιαδήποτε κατάσταση $(y, v) \in \mathbb{X} \times \mathbb{X}$ είναι προσβάσιμη από οποιαδήποτε άλλη κατάσταση $(x, u) \in \mathbb{X} \times \mathbb{X}$.

Θα δείξουμε τέλος ότι η $\tilde{\pi} : \mathbb{X} \times \mathbb{X} \rightarrow [0, 1]$ με $\tilde{\pi}(y, v) = \pi(y)\pi(v)$ είναι αναλλοίωτη κατανομή για τη $\{W_n\}_n$. Πράγματι, είναι φανερό ότι $\tilde{\pi}(y, v) \geq 0$ για κάθε $(y, v) \in \mathbb{X} \times \mathbb{X}$, ενώ

$$\sum_{y, v \in \mathbb{X}} \tilde{\pi}(y, v) = \sum_{y, v \in \mathbb{X}} \pi(y)\pi(v) = \sum_{y \in \mathbb{X}} \pi(y) \sum_{v \in \mathbb{X}} \pi(v) = 1.$$

Τέλος,

$$\begin{aligned}\sum_{x, u \in \mathbb{X}} \tilde{\pi}(x, u) \tilde{p}((x, u), (y, v)) &= \sum_{x, u \in \mathbb{X}} \pi(x)\pi(u) p(x, y) p(u, v) \\ &= \sum_{x \in \mathbb{X}} \pi(x) p(x, y) \sum_{u \in \mathbb{X}} \pi(u) p(u, v) = \pi(y)\pi(v) = \tilde{\pi}(y, v).\end{aligned}$$

Εφόσον η $\{W_n\}_n$ είναι μια μη υποβιβάσιμη αλυσίδα που έχει αναλλοίωτη κατανομή $\tilde{\pi}$, είναι και γνησίως επαναληπτική. Ο χρόνος T που ορίσαμε παραπάνω είναι ο χρόνος πρώτης άφιξης της $\{W_n\}_n$ στη διαγώνιο του $\mathbb{X} \times \mathbb{X}$, δηλαδή στο σύνολο $\Delta = \{(w, w) : w \in \mathbb{X}\}$. Από την επαναληπτικότητα της αλυσίδας προκύπτει ο ισχυρισμός του βήματος.

Βήμα 2: Ορίζουμε τη διαδικασία $\{Z_n\}_{n \in \mathbb{N}_0}$ ως εξής:

$$Z_n(\omega) = \begin{cases} X_n(\omega), & \text{αν } n \leq T(\omega) \\ Y_n(\omega), & \text{αν } n > T(\omega). \end{cases}$$

Θα δείξουμε ότι η $\{Z_n\}_n$ είναι μια μαρκοβιανή αλυσίδα στον \mathbb{X} με αρχική κατανομή π_0 και πιθανότητες μετάβασης $p(\cdot, \cdot)$, έχει δηλαδή την ίδια κατανομή με τη $\{X_n\}_n$.

Αυτός ο ισχυρισμός είναι διαισθητικά φανερός. Η $\{Z_n\}$ παρακολουθεί τη $\{X_n\}$ μέχρι τον χρόνο T κατά τον οποίο η $\{X_n\}_n$ συναντά για πρώτη φορά την $\{Y_n\}_n$ και στη συνέχεια παρακολουθεί την $\{Y_n\}_n$. Επειδή όμως, από την ισχυρή μαρκοβιανή ιδιότητα, τόσο η $\{X_{T+n}\}_{n \in \mathbb{N}_0}$ όσο και η $\{Y_{T+n}\}_{n \in \mathbb{N}_0}$ είναι μαρκοβιανές αλυσίδες που ξεκινούν από το σημείο συνάντησής τους, έχουν πιθανότητες μετάβασης $p(\cdot, \cdot)$ και είναι ανεξάρτητες από το παρελθόν του χρόνου T , περιμένουμε ότι θα είναι αδύνατο να ξεχωρίσει κανείς τη στατιστική συμπεριφορά των $\{X_n\}_n$ και $\{Z_n\}_n$.

Το ότι η Z_0 έχει κατανομή π_0 είναι φανερό, αφού $T \geq 0$ και άρα $\mathbb{P}[Z_0 = X_0] = 1$. Θα δείξουμε ακόμα ότι αν $z_0, z_1, \dots, z_n, z_{n+1} \in \mathbb{X}$, τότε

$$\mathbb{P}[Z_{n+1} = z_{n+1}, Z_n = z_n, \dots, Z_0 = z_0] = p(z_n, z_{n+1}) \mathbb{P}[Z_n = z_n, \dots, Z_0 = z_0]. \quad (6.7)$$

Από την (6.7) έπεται ότι η $\{Z_n\}_{n \in \mathbb{N}_0}$ είναι μαρκοβιανή, με πιθανότητες μετάβασης $\{p(x, y)\}_{x, y \in \mathbb{X}}$. Από τη διαμέριση $\Omega = \bigcup_{k=0}^n \{T = k\} \cup \{T > n\}$ έχουμε

$$\begin{aligned}\mathbb{P}[Z_{n+1} = z_{n+1}, Z_n = z_n, \dots, Z_0 = z_0] &= \sum_{k=0}^n \mathbb{P}[Z_{n+1} = z_{n+1}, Z_n = z_n, \dots, Z_0 = z_0, T = k] \\ &\quad + \mathbb{P}[Z_{n+1} = z_{n+1}, Z_n = z_n, \dots, Z_0 = z_0, T > n]. \quad (6.8)\end{aligned}$$

Για $k \in \{0, 1, \dots, n\}$, έχουμε

$$\{Z_{n+1} = z_{n+1}, \dots, Z_0 = z_0, T = k\} = \bigcap_{i=0}^k \{X_i = z_i\} \bigcap_{i=0}^{k-1} \{Y_i \neq z_i\} \bigcap_{i=k}^{n+1} \{Y_i = z_i\}.$$

Από την ανεξαρτησία των αλυσίδων $\{X_n\}_{n \in \mathbb{N}_0}$, $\{Y_n\}_{n \in \mathbb{N}_0}$ και τη μαρκοβιανή ιδιότητα της $\{Y_n\}_{n \in \mathbb{N}_0}$ παίρνουμε ότι

$$\begin{aligned} \mathbb{P}[Z_{n+1} = z_{n+1}, \dots, Z_0 = z_0, T = k] &= \mathbb{P}\left[\bigcap_{i=0}^k \{X_i = z_i\}\right] \mathbb{P}\left[\bigcap_{i=0}^{k-1} \{Y_i \neq z_i\} \bigcap_{i=k}^{n+1} \{Y_i = z_i\}\right] \\ &= \mathbb{P}\left[\bigcap_{i=0}^k \{X_i = z_i\}\right] \mathbb{P}\left[\bigcap_{i=0}^{k-1} \{Y_i \neq z_i\} \bigcap_{i=k}^n \{Y_i = z_i\}\right] p(z_n, z_{n+1}) \\ &= \mathbb{P}\left[\bigcap_{i=0}^k \{X_i = z_i\} \bigcap_{i=0}^{k-1} \{Y_i \neq z_i\} \bigcap_{i=k}^n \{Y_i = z_i\}\right] p(z_n, z_{n+1}) \\ &= \mathbb{P}[Z_n = z_n, \dots, Z_0 = z_0, T = k] p(z_n, z_{n+1}). \end{aligned} \quad (6.9)$$

Με αντίστοιχο τρόπο, έχουμε ότι

$$\{Z_{n+1} = z_{n+1}, \dots, Z_0 = z_0, T > n\} = \bigcap_{i=0}^{n+1} \{X_i = z_i\} \bigcap_{i=0}^n \{Y_i \neq z_i\},$$

ενώ από την ανεξαρτησία των δύο αλυσίδων και τη μαρκοβιανή ιδιότητα της $\{X_n\}_{n \in \mathbb{N}_0}$ παίρνουμε ότι

$$\mathbb{P}[Z_{n+1} = z_{n+1}, \dots, Z_0 = z_0, T > n] = \mathbb{P}[Z_n = z_n, \dots, Z_0 = z_0, T > n] p(z_n, z_{n+1}). \quad (6.10)$$

Αντικαθιστώντας τις (6.9) και (6.10) στην (6.8) παίρνουμε την (6.7).

Βήμα 3: Θα δείξουμε ότι, για κάθε $x \in \mathbb{X}$, έχουμε

$$|\pi_n(x) - \pi(x)| \leq \mathbb{P}[T > n]. \quad (6.11)$$

Εφόσον οι $\{X_n\}_n$ και $\{Z_n\}_n$ έχουν την ίδια κατανομή, έχουμε

$$\pi_n(x) = \mathbb{P}[X_n = x] = \mathbb{P}[Z_n = x], \text{ για κάθε } x \in \mathbb{X}, n \in \mathbb{N}_0.$$

Επειδή η αρχική κατανομή π της $\{Y_n\}_n$ είναι αναλλοίωτη κατανομή έχουμε

$$\pi(x) = \mathbb{P}[Y_n = x], \text{ για κάθε } x \in \mathbb{X}, n \in \mathbb{N}_0.$$

Συνδυάζοντας τις δύο παραπάνω σχέσεις έχουμε ότι, για κάθε $x \in \mathbb{X}$ και $n \in \mathbb{N}_0$,

$$\begin{aligned} |\pi_n(x) - \pi(x)| &= |\mathbb{P}[Z_n = x] - \mathbb{P}[Y_n = x]| \\ &= |\mathbb{P}[Z_n = x, T \leq n] + \mathbb{P}[Z_n = x, T > n] - \mathbb{P}[Y_n = x, T \leq n] - \mathbb{P}[Y_n = x, T > n]| \\ &= |\mathbb{P}[X_n = x, T > n] - \mathbb{P}[Y_n = x, T > n]| \\ &\leq \mathbb{P}[T > n]. \end{aligned}$$

Η ακολουθία ενδεχομένων $A_n = \{T > n\}$ είναι φθίνουσα, δηλαδή $A_{n+1} \subset A_n$, για κάθε $n \in \mathbb{N}_0$. Έχουμε λοιπόν

$$\lim_{n \rightarrow \infty} \mathbb{P}[T > n] = \mathbb{P}\left[\bigcap_n \{T > n\}\right] = \mathbb{P}[T = \infty] = 0$$

από το αποτέλεσμα του βήματος 1. Συνεπώς, για κάθε $x \in \mathbb{X}$ έχουμε $\pi_n(x) \rightarrow \pi(x)$ και άρα $\pi_n \rightarrow \pi$.

Τέλος, για να αποδείξουμε ότι $p^{(n)}(x, y) \rightarrow \pi(y)$ αρκεί να ξεκινήσουμε την αλυσίδα $\{X_n\}$ από την κατάσταση x , να πάρουμε δηλαδή $\mathbb{P}[X_0 = x] = 1$. Τότε $p^{(n)}(x, y) = \mathbb{P}[X_n = y] \rightarrow \pi(y)$, μια και ο ισχυρισμός του θεωρήματος ισχύει για οποιαδήποτε αρχική κατανομή π_0 . \square

Παρατηρήστε πώς χρησιμοποιήσαμε τη συνθήκη της απεριοδικότητας στην απόδειξη στο βήμα 1. Χωρίς την απεριοδικότητα της $\{X_n\}$ δεν είναι δυνατόν να συμπεράνουμε ότι η $\{W_n\}_n$ είναι μη υποβιβάσιμη. Αν για παράδειγμα θεωρήσουμε μια αλυσίδα που κινείται στις κορυφές ενός τετραγώνου A, B, C, D και σε κάθε βήμα μεταβαίνει με πιθανότητα $1/2$ σε μία από τις δύο γειτονικές της κορυφές, τότε η κατάσταση (A, A) δεν είναι προσβάσιμη για την $\{W_n\}_n$ από την κατάσταση (A, B) . Πράγματι,

$$\mathbb{P}[W_n = (A, A) \mid W_0 = (A, B)] = p^{(n)}(A, A)p^{(n)}(B, A) = 0,$$

αφού ο πρώτος όρος του γινομένου στο δεξί μέλος είναι 0 αν n περιττός, ενώ ο δεύτερος όρος είναι 0 αν n άρτιος. Επομένως, χωρίς την απεριοδικότητα της $\{X_n\}$ δεν μπορούμε να συμπεράνουμε τη συνθήκη $\mathbb{P}[T < +\infty] = 1$ που μας δίνει το ζητούμενο με τη βοήθεια της εκτίμησης του βήματος 3.

6.5 Το εργοδικό θεώρημα

Σε αυτή την παράγραφο θα αποδείξουμε το εργοδικό θεώρημα για μαρκοβιανές αλυσίδες. Το εργοδικό θεώρημα μας δίνει πληροφορίες για την ασυμπτωτική συμπεριφορά του χρονικού μέσου όρου συναρτησιακών της αλυσίδας και με αυτή την έννοια είναι μια γενίκευση του κλασικού νόμου των μεγάλων αριθμών στην περίπτωση μεταβλητών που δεν είναι ανεξάρτητες. Στην περίπτωση μη αρνητικών τυχαίων μεταβλητών που θα χρειαστούμε παρακάτω ο ισχυρός νόμος των μεγάλων αριθμών διατυπώνεται ως εξής.

Θεώρημα 29 Έστω $\{Z_n\}_{n \in \mathbb{N}}$ μια ακολουθία από ανεξάρτητες, ισόνομες, μη αρνητικές τυχαίες μεταβλητές με $\mathbb{E}[Z_i] = \mu \leq \infty$. Τότε

$$\mathbb{P}\left[\frac{1}{n} \sum_{k=1}^n Z_k \rightarrow \mu\right] = 1.$$

Η απόδειξη αυτού του θεωρήματος μπορεί να βρεθεί στην [8] και στην [9]. Προσέξτε ότι για κάθε $n \in \mathbb{N}$ ο χρονικός μέσος της ακολουθίας $\frac{1}{n} \sum_{k=1}^n Z_k$ είναι μια τυχαία μεταβλητή. Ο νόμος των μεγάλων αριθμών μάς δίνει την πληροφορία ότι η κατανομή αυτής της τυχαίας μεταβλητής συγκεντρώνεται ασυμπτωτικά γύρω από το μ . Η ανεξαρτησία των $\{Z_n\}_{n \in \mathbb{N}}$ παίζει αποφασιστικό ρόλο σε αυτό. Πράγματι, στην ακραία περίπτωση που πάρει κανείς όλες τις μεταβλητές ίσες με μια τυχαία μεταβλητή Z , ο μέσος όρος οσωνδήποτε από αυτές είναι η Z και φυσικά η κατανομή του δεν συγκεντρώνεται γύρω από έναν αριθμό. Μπορούμε να καταλάβουμε τον μηχανισμό με τον οποίον υπεισέρχεται η ανεξαρτησία στην περίπτωση $\mu < \infty$ ως εξής. Για ένα τυπικό $\omega \in \Omega$, κάποιες από τις τυχαίες μεταβλητές παίρνουν τιμές μεγαλύτερες από τη μέση τιμή τους μ και κάποιες μικρότερες. Επειδή οι μεταβλητές είναι ανεξάρτητες, οι προς τα πάνω αποκλίσεις από τη μέση τιμή σε μεγάλο βαθμό αλληλοαναιρούνται με τις προς τα κάτω αποκλίσεις, τόσο που οι αθροιστικές αποκλίσεις n προσθετών είναι πολύ μικρότερες από το n με το οποίο διαιρούμε και στο όριο $n \rightarrow \infty$ εξαφανίζονται.

Αν τώρα $\{X_n\}_{n \in \mathbb{N}_0}$ είναι μια μη υποβιβάσιμη μαρκοβιανή αλυσίδα σε έναν χώρο καταστάσεων \mathbb{X} και $f : \mathbb{X} \rightarrow \mathbb{R}$, μπορεί κανείς να μελετήσει τον χρονικό μέσο όρο των πραγματικών τυχαίων μεταβλητών $Z_n = f(X_n)$. Αυτές βέβαια δεν είναι ανεξάρτητες. Έχουμε όμως δει ότι η εξάρτηση της κατανομής μιας μη υποβιβάσιμης μαρκοβιανής αλυσίδας από τις προηγούμενες τιμές της εξασθενεί προϊόντος του χρόνου. Το εργοδικό Θεώρημα μας λέει ότι και σε αυτή την περίπτωση οι αθροιστικές αποκλίσεις n προσθετών

είναι πολύ μικρότερες από το n και οι χρονικοί μέσοι όροι συγκλίνουν καθώς $n \rightarrow \infty$. Στην απλούστερη μορφή του που θα δούμε πρώτα, η συνάρτηση f είναι η δείκτρια μιας κατάστασης $x \in \mathbb{X}$ και το εργοδικό θεώρημα μας δίνει πληροφορία για το ασυμπτωτικό ποσοστό του χρόνου που περνά η αλυσίδα στη x .

Για $x \in \mathbb{X}$ ορίζουμε $V_n(x)$ το πλήθος των επισκέψεων στη x μιας αλυσίδας $\{X_n\}_{n \in \mathbb{N}_0}$ πριν τη χρονική στιγμή n ,

$$V_n(x) = \sum_{k=0}^{n-1} \mathbb{1}\{X_k = x\}.$$

Θεώρημα 30 (εργοδικό Θεώρημα) Έστω $\{X_n\}_{n \in \mathbb{N}_0}$ μια μη υποβιβάσιμη μαρκοβιανή αλυσίδα σε έναν διακριτό χώρο καταστάσεων \mathbb{X} . Για οποιαδήποτε αρχική κατανομή της αλυσίδας, έχουμε

$$\mathbb{P}\left[\frac{V_n(x)}{n} \rightarrow \frac{1}{\mathbb{E}_x[T_x^+]}, \text{ για κάθε } x \in \mathbb{X}\right] = 1,$$

όπου $T_x^+ = \inf\{k > 0 : X_k = x\}$ είναι ο χρόνος πρώτης επιστροφής της αλυσίδας στην κατάσταση x .

Απόδειξη: Θα αποδείξουμε πρώτα ότι για κάθε $x \in \mathbb{X}$

$$\mathbb{P}\left[\frac{V_n(x)}{n} \rightarrow \frac{1}{\mathbb{E}_x[T_x^+]}\right] = 1. \quad (6.12)$$

Αξίζει να παρατηρήσουμε ότι το αν υπάρχει το όριο $V_n(x)/n$ και το ποιο είναι αυτό εξαρτάται μόνο από την τελική συμπεριφορά της αλυσίδας και όχι από τα οσαδήποτε πρώτα βήματά της. Πράγματι, αν $M \in \mathbb{N}$ και θεωρήσουμε τη διαδικασία $\{Y_n\}_{n \in \mathbb{N}_0}$ με $Y_n = X_{M+n}$, έχουμε

$$\sum_{k=0}^{n-1} \mathbb{1}\{X_k = x\} - \sum_{k=0}^{n-1} \mathbb{1}\{Y_k = x\} = \sum_{k=0}^{M-1} (\mathbb{1}\{X_k = x\} - \mathbb{1}\{X_{n+k} = x\})$$

και επομένως

$$\left|\frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}\{X_k = x\} - \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}\{Y_k = x\}\right| \leq \frac{M}{n} \rightarrow 0, \text{ καθώς } n \rightarrow \infty.$$

Μάλιστα, ακόμα κι αν η $M = M(\omega)$ είναι ένας τυχαίος χρόνος με $\mathbb{P}[M < +\infty] = 1$, μπορούμε να χρησιμοποιήσουμε αυτό το επιχείρημα για κάθε $\omega \in \{\omega \in \Omega : M(\omega) < +\infty\}$.

Αν η x είναι επαναληπτική κατάσταση και $T_x = \inf\{k \geq 0 : X_k = x\}$, τότε για οποιαδήποτε αρχική κατανομή της αλυσίδας έχουμε $\mathbb{P}[T_x < +\infty] = 1$. Επιπλέον, από την ισχυρή μαρκοβιανή ιδιότητα, η διαδικασία $\{Y_n\}_{n \in \mathbb{N}_0}$ με $Y_n = X_{T_x+n}$ είναι μια αλυσίδα που ξεκινά από το x και έχει τις ίδιες πιθανότητες μετάβασης όπως η $\{X_n\}_n$. Σύμφωνα με την προηγούμενη παρατήρησή μας, για να δείξουμε την (6.12) στην περίπτωση που η x είναι επαναληπτική, αρκεί να δείξουμε ότι

$$\mathbb{P}_x\left[\frac{V_n(x)}{n} \rightarrow \frac{1}{\mathbb{E}_x[T_x^+]}\right] = 1. \quad (6.13)$$

Ορίζουμε $S_0 = 0$ και επαγωγικά για κάθε $n \in \mathbb{N}$ τον χρόνο της n -στής επανόδου στην κατάσταση x ως εξής

$$S_n(\omega) = \inf\{k > S_{n-1}(\omega) : X_k = x\}.$$

Από την ισχυρή μαρκοβιανή ιδιότητα και την επαναληπτικότητα της x , έχουμε $\mathbb{P}_x[S_n < +\infty \text{ για κάθε } n \in \mathbb{N}] = 1$. Επιπλέον, όπως είδαμε στο Πρόρισμα 2, οι χρόνοι $\{S_n - S_{n-1}\}_{n \in \mathbb{N}}$, που μεσολαβούν ανάμεσα

σε διαδοχικές επισκέψεις στην κατάσταση x είναι ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με κοινή κατανομή αυτήν της $S_1 - S_0 = T_x^+$. Από τον νόμο των μεγάλων αριθμών (Θεώρημα 29) έχουμε επομένως

$$\mathbb{P}_x \left[\lim_{n \rightarrow \infty} \frac{S_n}{n} = \mathbb{E}_x [T_x^+] \right] = 1.$$

Από την επαναληπτικότητα της κατάστασης x έχουμε ακόμα

$$\mathbb{P}_x [V_n(x) \rightarrow \infty] = 1.$$

Αφού τα παραπάνω ενδεχόμενα συμβαίνουν με \mathbb{P}_x -πιθανότητα 1 και η τομή τους θα συμβαίνει με \mathbb{P}_x -πιθανότητα 1, επομένως

$$\mathbb{P}_x \left[\lim_{n \rightarrow \infty} \frac{S_{V_n(x)}}{V_n(x)} = \mathbb{E}_x [T_x^+] \right] = 1. \quad (6.14)$$

Παρατηρήστε όμως τώρα ότι, για μια αλυσίδα που ξεκινά από το x , ο χρόνος $S_{V_n(x)} = \inf \{k \geq n : X_k = x\}$ είναι ο χρόνος της πρώτης επίσκεψης στην x μετά τη χρονική στιγμή n για κάθε $n \in \mathbb{N}$, επομένως

$$\mathbb{P}_x [S_{V_n(x)-1} < n \leq S_{V_n(x)}, \quad \text{για κάθε } n \in \mathbb{N}] = 1$$

και άρα η (6.13) προκύπτει από την (6.14).

Αν η κατάσταση x είναι παροδική, η απόδειξη της (6.12) είναι ακόμα πιο εύκολη, αφού τότε $\mathbb{E}_x [T_x^+] = +\infty$, ενώ $\mathbb{P}[\lim_{n \rightarrow \infty} V_n(x) < +\infty] = 1$ και άρα $\mathbb{P}[\lim_{n \rightarrow \infty} V_n(x)/n = 0] = 1$.

Ο ισχυρισμός του θεωρήματος προκύπτει τώρα από την (6.12) επειδή ο \mathbb{X} έχει το πολύ αριθμήσιμο πλήθος από στοιχεία. Πράγματι, η (6.12) σημαίνει ότι για κάθε $x \in \mathbb{X}$ έχουμε $\mathbb{P}[U_x] = 1$, όπου

$$U_x = \left\{ \omega \in \Omega : \frac{V_n(x)}{n} \rightarrow \frac{1}{\mathbb{E}_x [T_x^+]} \right\}.$$

Εφόσον καθένα από τα U_x έχει πιθανότητα 1, το ίδιο θα συμβαίνει και για την αριθμήσιμη τομή τους $\bigcap_{x \in \mathbb{X}} U_x$, αφού

$$\mathbb{P} \left[\left(\bigcap_{x \in \mathbb{X}} U_x \right)^c \right] = \mathbb{P} \left[\bigcup_{x \in \mathbb{X}} U_x^c \right] \leq \sum_{x \in \mathbb{X}} \mathbb{P}[U_x^c] = \sum_{x \in \mathbb{X}} 0 = 0.$$

□

Παρατήρηση: Αν η αλυσίδα $\{X_n\}$ είναι εκτός από μη υποβιβάσιμη και γνησίως επαναληπτική, θα έχει μοναδική αναλλοίωτη κατανομή π , η οποία θα ικανοποιεί όπως είδαμε στο προηγούμενο κεφάλαιο την

$$\pi(x) = \frac{1}{\mathbb{E}_x [T_x^+]}, \quad \text{για κάθε } x \in \mathbb{X}.$$

Σε αυτή την περίπτωση το Θεώρημα 30 μάς λέει ότι σε βάθος χρόνου το ποσοστό του χρόνου που ξοδεύει η αλυσίδα σε κάθε κατάσταση $x \in \mathbb{X}$ είναι το βάρος που δίνει στη x η αναλλοίωτη κατανομή π :

$$\mathbb{P} \left[\frac{V_n(x)}{n} \rightarrow \pi(x), \quad \text{για κάθε } x \in \mathbb{X} \right] = 1. \quad (6.15)$$

Στην περίπτωση αυτή μπορούμε να διατυπώσουμε το εργοδικό Θεώρημα και με μια διαφορετική μορφή που πολλές φορές είναι πιο χρήσιμη για υπολογισμούς.

Θεώρημα 31 (εργοδικό Θεώρημα II) Έστω $\{X_n\}_{n \in \mathbb{N}_0}$ μια μη υποβιβάσιμη, γνησίως επαναληπτική μαρκοβιανή αλυσίδα σε έναν διακριτό χώρο καταστάσεων \mathbb{X} . Αν $f : \mathbb{X} \rightarrow \mathbb{R}$ είναι μια φραγμένη συνάρτηση, τότε για οποιαδήποτε αρχική κατανομή της αλυσίδας έχουμε

$$\mathbb{P} \left[\frac{1}{n} \sum_{k=0}^{n-1} f(X_k) \rightarrow \mathbb{E}^\pi[f] \right] = 1,$$

όπου $\mathbb{E}^\pi[f] = \sum_{x \in \mathbb{X}} f(x)\pi(x)$ είναι η μέση τιμή της τυχαίας μεταβλητής $f(X)$, αν η τυχαία μεταβλητή X ακολουθεί την αναλλοίωτη κατανομή της αλυσίδας π .

Απόδειξη: Επειδή κάθε χρονική στιγμή $k \in \mathbb{N}_0$ η αλυσίδα βρίσκεται σε ακριβώς μία από τις καταστάσεις του \mathbb{X} έχουμε

$$\sum_{x \in \mathbb{X}} \mathbb{1}\{X_k = x\} = 1.$$

Μπορούμε τώρα να γράψουμε τον χρονικό μέσο των τιμών της f κατά μήκος του μονοπατιού της αλυσίδας ως

$$\begin{aligned} \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) &= \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) \sum_{x \in \mathbb{X}} \mathbb{1}\{X_k = x\} \\ &= \frac{1}{n} \sum_{x \in \mathbb{X}} \sum_{k=0}^{n-1} f(X_k) \mathbb{1}\{X_k = x\} \\ &= \frac{1}{n} \sum_{x \in \mathbb{X}} \sum_{k=0}^{n-1} f(x) \mathbb{1}\{X_k = x\} \\ &= \sum_{x \in \mathbb{X}} f(x) \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}\{X_k = x\} \\ &= \sum_{x \in \mathbb{X}} f(x) \frac{V_n(x)}{n}. \end{aligned}$$

Στην περίπτωση που ο \mathbb{X} είναι πεπερασμένος, ο ισχυρισμός προκύπτει κατευθείαν από το Θεώρημα 30. Στην περίπτωση που ο \mathbb{X} έχει άπειρο αλλά αριθμήσιμο πλήθος από καταστάσεις, θεωρήστε $M > 0$ τέτοιο ώστε $|f(x)| \leq M$, για κάθε $x \in \mathbb{X}$. Εφόσον η σειρά $\sum_{x \in \mathbb{X}} \pi(x)$ συγκλίνει (στο 1), για κάθε $\epsilon > 0$ μπορούμε να βρούμε ένα πεπερασμένο σύνολο από καταστάσεις $A \subset \mathbb{X}$ τέτοιο ώστε $\sum_{x \notin A} \pi(x) < \frac{\epsilon}{2M}$. Έχουμε

$$\begin{aligned} \left| \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) - \mathbb{E}^\pi[f] \right| &= \left| \sum_{x \in \mathbb{X}} f(x) \left(\frac{V_n(x)}{n} - \pi(x) \right) \right| \\ &\leq M \sum_{x \in \mathbb{X}} \left| \frac{V_n(x)}{n} - \pi(x) \right| \\ &= M \sum_{x \in A} \left| \frac{V_n(x)}{n} - \pi(x) \right| + M \sum_{x \notin A} \left| \frac{V_n(x)}{n} - \pi(x) \right|. \end{aligned} \quad (6.16)$$

Για τους προσθετέους του τελευταίου όρου θα χρησιμοποιήσουμε την αλγεβρική ταυτότητα $|u| = u + 2u^-$, που ισχύει για κάθε $u \in \mathbb{R}$, όπου u^- είναι το αρνητικό μέρος του u και δίνεται από την $u^- = \max\{0, -u\}$.

Παίρνουμε έτσι

$$\begin{aligned}
\sum_{x \notin A} \left| \frac{V_n(x)}{n} - \pi(x) \right| &= \sum_{x \notin A} \left(\frac{V_n(x)}{n} - \pi(x) \right) + 2 \sum_{x \notin A} \left(\frac{V_n(x)}{n} - \pi(x) \right)^- \\
&= \sum_{x \in A} \left(\pi(x) - \frac{V_n(x)}{n} \right) + 2 \sum_{x \notin A} \left(\frac{V_n(x)}{n} - \pi(x) \right)^- \\
&\leq \sum_{x \in A} \left(\pi(x) - \frac{V_n(x)}{n} \right) + 2 \sum_{x \notin A} \pi(x) \\
&\leq \sum_{x \in A} \left(\pi(x) - \frac{V_n(x)}{n} \right) + \frac{\epsilon}{M},
\end{aligned}$$

όπου στο δεύτερο βήμα χρησιμοποιήσαμε ότι

$$\sum_{x \in \mathbb{X}} \frac{V_n(x)}{n} = 1 = \sum_{x \in \mathbb{X}} \pi(x),$$

ενώ στο τρίτο ότι, αν $u, v > 0$, τότε $(u - v)^- \leq v$. Με αντικατάσταση της παραπάνω σχέσης στην (6.16) έχουμε

$$\left| \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) - \mathbb{E}^\pi[f] \right| \leq 2M \sum_{x \in A} \left| \frac{V_n(x)}{n} - \pi(x) \right| + \epsilon.$$

Εφόσον το σύνολο A είναι πεπερασμένο, για τα $\omega \in \Omega$ για τα οποία $V_n(x)/n \rightarrow \pi(x)$ έχουμε

$$\limsup_{n \rightarrow \infty} \left| \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) - \mathbb{E}^\pi[f] \right| \leq \epsilon$$

και άρα από το Θεώρημα 30 έχουμε ότι για κάθε $\epsilon > 0$

$$\mathbb{P} \left[\limsup_{n \rightarrow \infty} \left| \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) - \mathbb{E}^\pi[f] \right| \leq \epsilon \right] = 1.$$

Αν πάρουμε τώρα $\epsilon = 1/N$, η ακολουθία ενδεχομένων

$$B_N = \left\{ \omega \in \Omega : \limsup_{n \rightarrow \infty} \left| \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) - \mathbb{E}^\pi[f] \right| \leq \frac{1}{N} \right\}$$

είναι φθίνουσα ($B_{N+1} \subset B_N$ για κάθε $N \in \mathbb{N}$) και $\mathbb{P}[B_N] = 1$ για κάθε $N \in \mathbb{N}$. Επομένως,

$$\mathbb{P} \left[\frac{1}{n} \sum_{k=0}^{n-1} f(X_k) \rightarrow \mathbb{E}^\pi[f] \right] = \mathbb{P} \left[\bigcap_N B_N \right] = \lim_{N \rightarrow \infty} \mathbb{P}[B_N] = 1,$$

που είναι και ο ισχυρισμός του θεωρήματος. □

6.6 Ασκήσεις

Άσκηση 95 Αποδείξτε το Θεώρημα 27.

Άσκηση 96 Υπολογίστε τη συνδιακύμανση των τυχαίων μεταβλητών \tilde{X}_e και \tilde{Y}_e του Παραδείγματος 49.

Άσκηση 97 Τέσσερις υπολογιστές είναι συνδεδεμένοι ανά δύο. Κάθε τέτοια σύνδεση είναι λειτουργική με πιθανότητα $p \in [0, 1]$ ανεξάρτητα από τις άλλες συνδέσεις. Υπολογίστε την πιθανότητα $f(p)$ να υπάρχει λειτουργικό κανάλι που να επιτρέπει την επικοινωνία δύο δεδομένων υπολογιστών (ενδεχομένως και μέσω των άλλων δύο υπολογιστών) και μελετήστε τη συνάρτηση $f : p \rightarrow f(p)$ ως προς τη μονοτονία της. Στη συνέχεια δείξτε ότι η f είναι αύξουσα με την τεχνική της σύζευξης και παρατηρήστε πόσο πιο εύκολος και γενικεύσιμος σε πιο πολύπλοκες συνδεσμολογίες είναι ο δεύτερος τρόπος.

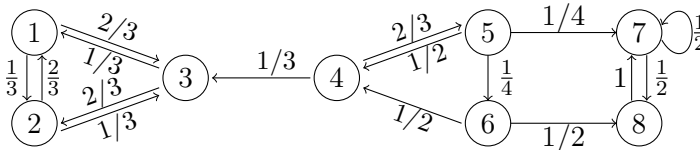
Άσκηση 98 Ρίχνουμε επαναλαμβανόμενα ένα τίμιο ζάρι και προσθέτουμε τα αποτελέσματα των ενδείξεων. Αν S_n είναι το άθροισμα των n πρώτων ζαριών μας, υπολογίστε το όριο

$$\lim_{n \rightarrow \infty} \mathbb{P}[S_n \text{ είναι πολλαπλάσιο του } 7].$$

Άσκηση 99 Στην Άσκηση 5, σε βάθος χρόνου, ποιο ποσοστό του χρόνου του ξοδεύει το έντομο στο σαλόνι;

Άσκηση 100 Στην Άσκηση 81, σε βάθος χρόνου, ποιο ποσοστό των ημερών το μπακάλικο δεν έχει μπισκότα την ώρα που κλείνει;

Άσκηση 101 Στο παρακάτω σχήμα φαίνονται οι καταστάσεις και οι πιθανότητες μετάβασης μιας αλυσίδας $\{X_n\}_{n \in \mathbb{N}_0}$.



α) Ταξινομήστε τις καταστάσεις σε κλάσεις επικοινωνίας.

β) Βρείτε όλες τις αναλλοίωτες κατανομές της αλυσίδας.

γ) Αν $X_0 = 5$, υπολογίστε την πιθανότητα η αλυσίδα να καταλήξει σε καθεμία από τις κλειστές κλάσεις.

δ) Υπολογίστε το όριο $\lim_{n \rightarrow \infty} \mathbb{P}[X_n = j | X_0 = i]$ για τα παρακάτω ζεύγη (i, j) : $(5, 5)$, $(8, 8)$ και $(5, 8)$

ε) Έστω $X_0 = 5$. Με ανάλυση πρώτου βήματος ή με οποιονδήποτε άλλον τρόπο, υπολογίστε την πιθανότητα η αλυσίδα να εγκαταλείψει την κλάση από την οποία ξεκινά έπειτα από έναν άρτιο αριθμό βημάτων.

Άσκηση 102 Αν η X_n είναι μια μαρκοβιανή αλυσίδα σ' ένα χώρο \mathbb{X} , τότε η $Y_{n+1} = (X_n, X_{n+1})$ είναι μια μαρκοβιανή αλυσίδα στον $\mathbb{X} \times \mathbb{X}$. Ποιες είναι οι πιθανότητες μετάβασης της Y ; Αν η X έχει μοναδική αναλλοίωτη κατανομή π , δείξτε ότι η y έχει μοναδική αναλλοίωτη κατανομή την $\pi_*(j, k) = \pi(j)p(j, k)$.

Άσκηση 103 Με τη βοήθεια της προηγούμενης άσκησης δείξτε ότι αν η αλυσίδα $\{X_n\}_{n \in \mathbb{N}}$ είναι μη υποβιβάζσιμη και κινείται σ' έναν πεπερασμένο χώρο καταστάσεων \mathbb{X} και $f : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}$, τότε με πιθανότητα 1 έχουμε

$$\lim_{n \rightarrow \infty} \frac{f(X_0, X_1) + f(X_1, X_2) + \cdots + f(X_{n-1}, X_n)}{n} = \sum_{(x,y) \in \mathbb{X} \times \mathbb{X}} f(x, y) \pi(x) p(x, y).$$

για οποιαδήποτε κατανομή της X_0 .

Άσκηση 104 Στην Άσκηση 79, σε βάθος χρόνου, ποιο ποσοστό των μεταβάσεων γίνονται μεταξύ των καταστάσεων 3 και 4; Ποιο είναι το όριο

$$\lim_{n \rightarrow \infty} \mathbb{P}[X_n = 4 | X_{n+1} = 3];$$

Άσκηση 105 Μια μαρκοβιανή αλυσίδα στο \mathbb{N} μπορεί να μεταβεί από έναν φυσικό $n \geq 2$ είτε στον $n + 1$ είτε στον $n - 1$. Όταν βρίσκεται στο 1 μπορεί είτε να πάει στο 2 είτε να παραμείνει στο 1. Οι πιθανότητες μετάβασης ικανοποιούν την

$$p(n, n + 1) = \frac{n}{n + 2} p(n + 1, n) > 0 \quad \text{για κάθε } n \in \mathbb{N}.$$

Με μια πρώτη ματιά φαίνεται ότι οι παραπάνω συνθήκες δεν καθορίζουν πλήρως τις πιθανότητες μετάβασης αλλά η $p(1, 1) = x$ είναι μια ελεύθερη παράμετρος. Οι άλλες πιθανότητες μετάβασης υπολογίζονται με τη βοήθεια της x ως εξής:

$$p(1, 2) = 1 - x, \quad p(2, 1) = \frac{1 + 2}{1} p(1, 2) = 3(1 - x), \quad p(2, 3) = 1 - p(2, 1) = 3x - 2 \quad \text{κ.ο.κ.}$$

α) Εκφράστε την $p(n, n + 1)$ ως συνάρτηση της x και δείξτε ότι $p(n, n + 1) \in (0, 1]$ για κάθε $n \in \mathbb{N}$ και μόνο αν $x = 4 \ln(2) - 2$.

Αν καταφέρατε να δείξετε τον προηγούμενο ισχυρισμό, σίγουρα θα κάνατε αρκετή δουλειά. Θα δούμε τώρα έναν πολύ απλό τρόπο για να υπολογίσουμε την $p(1, 1)$.

β) Βρείτε μια κατανομή π στο \mathbb{N} που ικανοποιεί τις συνθήκες ακριβούς ισορροπίας με τις $\{p(m, n)\}_{m, n \in \mathbb{N}}$.

γ) Χρησιμοποιώντας το εργοδικό θεώρημα δείξτε ότι το όριο $\lim_{n \rightarrow \infty} \frac{(-1)^{X_1 + \dots + (-1)^{X_n}}}{n}$ υπάρχει με πιθανότητα 1 και υπολογίστε το.

δ) Στο προηγούμενο άθροισμα υπάρχουν πολλές αλληλοαναιρέσεις ανάμεσα σε διαδοχικούς όρους της αλυσίδας. Προσπαθήστε να καταλάβετε ποιοι όροι δημιουργούν τη μη μηδενική συνεισφορά στο όριο και συμπεράνετε ότι $p(1, 1) = 4 \ln(2) - 2$.

Άσκηση 106 Ένας παίκτης του μπάσκετ προπονείται στα τρίποντα. Έχετε παρατηρήσει ότι η πιθανότητα να ευστοχήσει σε ένα σουτ είναι ίση με

- $1/2$ αν έχει ευστοχήσει και στα δύο προηγούμενα σουτ που έχει επιχειρήσει
- $1/3$ αν έχει ευστοχήσει σε ένα από τα δύο προηγούμενα σουτ που έχει επιχειρήσει
- $1/4$ αν έχει αστοχήσει και στα δύο προηγούμενα σουτ που έχει επιχειρήσει.

Για $n \in \mathbb{N}$ ορίζουμε $X_n = 1$ αν το n -οστό σουτ του παίκτη είναι εύστοχο και 0 διαφορετικά.

α. Είναι η ακολουθία $\{X_n\}$ μαρκοβιανή; Δικαιολογήστε την απάντησή σας.

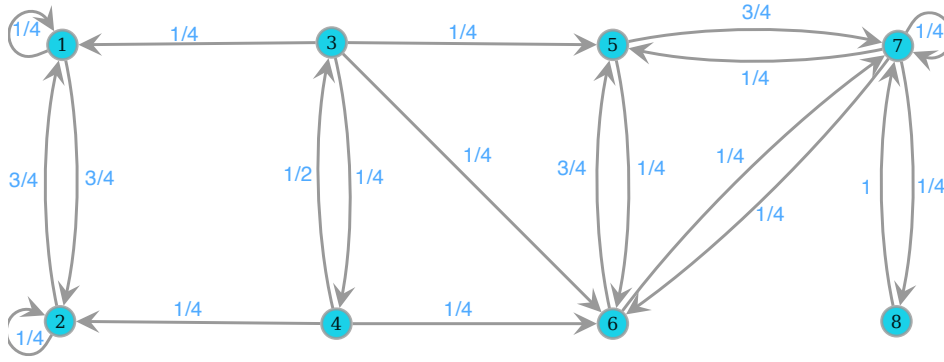
β. Είναι η αλυσίδα $Y_n = (X_n, X_{n+1})$ μια μαρκοβιανή αλυσίδα στον χώρο καταστάσεων $\mathbb{X} = \{0, 1\} \times \{0, 1\} = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$; Ποιες είναι οι πιθανότητες μετάβασης;

γ. Ποια είναι η αναλλοίωτη κατανομή αυτής της αλυσίδας;

δ. Σε βάθος χρόνου τι ποσοστό από τα σουτ του είναι εύστοχα;

Άσκηση 107 Επιλέγουμε τυχαία δύο n -ψήφιους αριθμούς και τους προσθέτουμε. Αν A_n είναι το πλήθος των κρατούμενων που μεταφέραμε κατά την πρόσθεση, τι συμβαίνει στο όριο $\lim_n \frac{A_n}{n}$;

Άσκηση 108 Στο παρακάτω σχήμα φαίνονται οι καταστάσεις και οι πιθανότητες μετάβασης μιας μαρκοβιανής αλυσίδας $\{X_n\}_n$.



- α) Βρείτε τις κλάσεις επικοινωνίας της αλυσίδας και χαρακτηρίστε τις ως προς την επαναληπτικότητα.
 β) Αν $X_0 = 3$, ποιος είναι ο αναμενόμενος χρόνος εξόδου της αλυσίδας από την κλάση που περιέχει την κατάσταση 3 και ποια η πιθανότητα να καταλήξει σε καθεμιά από τις άλλες κλάσεις;
 γ) Υπολογίστε το όριο $\lim_n \mathbb{P}[X_n = 8 \mid X_0 = 8]$.
 δ) Αν $X_0 = 8, T_1 = \inf\{k > 0 : X_k = 8\}$ και $T_2 = \inf\{k > T_1 : X_k = 8\}$, δηλαδή T_1 και T_2 είναι οι χρόνοι πρώτης και δεύτερης επανόδου στο 8 αντίστοιχα, υπολογίστε τις $\mathbb{E}[T_1]$ και $\mathbb{E}[T_2]$.
 ε) Έστω $X_0 = 3$. Αν κερδίζετε 1 ευρώ κάθε φορά που η αλυσίδα βρίσκεται σε κατάσταση με άρτιο δείκτη, τι μπορείτε να πείτε για το μέσο κέρδος σας ανά κίνηση σε βάρθος χρόνου; Ποιες τιμές μπορεί να πάρει; Με ποια πιθανότητα;

6.7 Αριθμητικά πειράματα

Άσκηση 109 Στην άσκηση αυτή γίνεται προσομοίωση της αλυσίδας της Άσκησης 83 για $p = 1/2$. Υπενθυμίζεται ότι αυτή η αλυσίδα στον χώρο καταστάσεων $\mathbb{N} \cup \{0\}$ είναι μη υποβιβάσιμη, γνησίως επαναληπτική, με αναλλοίωτη κατανομή π που δίνεται από την

$$\pi(k) = \frac{1}{2^{k+1}}, \quad k = 0, 1, 2, 3, \dots$$

Κατεβάστε και τρέξτε τον κώδικα `ergodic.py`. Ο κώδικας αυτός προσομοιώνει τα πρώτα $N = 10^6$ βήματα της αλυσίδας και επιστρέφει τον εργοδικό μέσο

$$\frac{X_1 + X_2 + \dots + X_N}{N}.$$

Από το εργοδικό Θεώρημα η παραπάνω ποσότητα συγκλίνει καθώς $N \rightarrow \infty$ με πιθανότητα 1 στο

$$\sum_{k=0}^{\infty} k\pi(k),$$

οπότε ο κώδικας υπολογίζει αριθμητικά την παραπάνω ποσότητα με τη μέθοδο Markov Chain Monte Carlo (MCMC).

- α) Υπολογίστε αναλυτικά το παραπάνω άθροισμα και επιβεβαιώστε ότι ο κώδικας προσφέρει μια καλή εκτίμησή του.
 β) Αλλάξτε τον κώδικα ώστε να υπολογίζει το παραπάνω άθροισμα 50 φορές και βρείτε τη διασπορά των αποτελεσμάτων.

γ) Αν θέλαμε να περιορίσουμε τη διασπορά των αποτελεσμάτων στο μισό, πόσο μεγάλο θα έπρεπε να πάρουμε το N ; Απαντήστε θεωρητικά και επιβεβαιώστε το αριθμητικά.

δ) Αλλάξτε τον κώδικα ώστε να υπολογίζει με τη μέθοδο MCMC το άθροισμα

$$\sum_{k=0}^{\infty} \frac{\cos(k + \cos(k))}{2^k}.$$

Άσκηση 110 Σ' ένα παιχνίδι έχετε το δικαίωμα να ρίξετε ένα ζάρι όσες φορές θέλετε και να εισπράξετε το άθροισμα των ζαριών σας αν δεν φέρετε κανέναν άσο. Αν σε κάποια ζαριά φέρετε άσο τότε το παιχνίδι τελειώνει χωρίς να κερδίσετε τίποτα. Θέλουμε να μεγιστοποιήσουμε το αναμενόμενο κέρδος από το παιχνίδι. Περιμένουμε ότι η βέλτιστη στρατηγική θα ήταν να σταματήσουμε μόλις το άθροισμα των ζαριών μας ξεπεράσει κάποιο κατώφλι k , το οποίο θέλουμε να προσδιορίσουμε αριθμητικά.

Ο κώδικας `opt_stop.py` προσομοιώνει για $k = 0, 1, \dots, 40$ $N = 1.000$ παρτίδες του παιχνιδιού με κατώφλι διακοπής k και υπολογίζει το μέσο κέρδος με κάθε στρατηγική. Στην έξοδό του επιστρέφει τη στρατηγική με την οποία σημειώθηκε το μεγαλύτερο κέρδος και ένα διάγραμμα με το υπολογιζόμενο κέρδος ως συνάρτηση του k .

α) Τρέξτε τον κώδικα και δείτε το διάγραμμα μερικές φορές. Παρατηρήστε ότι η μεγάλη διασπορά στην εκτίμηση του κέρδους δεν μας επιτρέπει να προσδιορίσουμε τη βέλτιστη στρατηγική.

Θα δούμε τώρα πώς μπορούμε να αντιμετωπίσουμε το πρόβλημα της διασποράς. Η μέθοδος Monte Carlo μας προσφέρει έναν αριθμητικό τρόπο υπολογισμού μιας ποσότητας x , αν μπορούμε να φανταστούμε τη x ως την αναμενόμενη τιμή κάποιας τυχαίας μεταβλητής X που μπορούμε να προσομοιώσουμε. Η εκτιμήτρια Monte Carlo της x

$$\bar{x}_N = \frac{X_1 + \dots + X_N}{N}$$

συγκλίνει στη x από το νόμο των μεγάλων αριθμών καθώς $N \rightarrow \infty$, ενώ η διασπορά της για πεπερασμένο N είναι $\frac{\sigma^2}{N}$, όπου σ^2 είναι η διασπορά της X . Στο παιχνίδι μας, το κέρδος σε μία παρτίδα είναι μια τυχαία μεταβλητή με μεγάλη διασπορά σ^2 . Παίρνει είτε την τιμή μηδέν, αν φέρουμε άσο πριν ξεπεράσουμε το κατώφλι k , είτε κάποια από τις τιμές $k + 2, k + 3, \dots, k + 6$. Έτσι η διασπορά της εκτιμήτριάς μας είναι κι αυτή σχετικά μεγάλη.

β) Ένας τρόπος να ελαττώσουμε τη διασπορά της εκτιμήτριας είναι να αυξήσουμε το N και επομένως τον υπολογιστικό χρόνο. Δεκαπλασιάστε το N και επαναλάβετε το προηγούμενο ερώτημα. Έλυσε αυτή η αλλαγή εντελώς το πρόβλημα που είχαμε;

Ένας άλλος τρόπος είναι να φανταστούμε το x ως τη μέση τιμή μιας τυχαίας μεταβλητής Y με μικρότερη διασπορά από τη X . Υπάρχουν πολλοί τρόποι να γίνει αυτό και αναφέρονται στη βιβλιογραφία ως τεχνικές *ελάττωσης διασποράς* (variance reduction). Θα δοκιμάσουμε στο παιχνίδι μας την τεχνική της *δημηματοληψίας κατά σημαντικότητα* (importance sampling).

Αν ρίχναμε ένα ζάρι χωρίς άσο, το κέρδος μας σε μία παρτίδα με κατώφλι διακοπής k θα έπαιρνε μόνο κάποια από τις τιμές $k + 2, k + 3, \dots, k + 6$ και θα είχε πολύ μικρότερη διασπορά. Θα προσπαθήσουμε λοιπόν να εκφράσουμε το μέσο κέρδος $\mathbb{E}[K]$ στο αρχικό μας παιχνίδι με κατώφλι διακοπής k μέσω του κέρδους που θα είχε αυτή η στρατηγική, αν το ζάρι μας δεν είχε άσο. Παρατηρούμε αρχικά ότι

$$\begin{aligned} \mathbb{E}[K] &= \sum_m \sum_{(x_0, x_1, \dots, x_m) \in J_k} x_m \mathbb{P}[X_1 = x_1, \dots, X_m = x_m] \\ &= \sum_m \sum_{(x_0, x_1, \dots, x_m) \in J_k} x_m \left(\frac{1}{6}\right)^m. \end{aligned}$$

όπου

$$J_k = \{(x_0, x_1, \dots, x_m) : x_0 = 0, x_{i+1} \in \{x_i + 2, \dots, x_i + 6\}, i = 0, 1, \dots, m-1, x_{m-1} \leq k, x_m > k\}$$

είναι εκείνες οι τροχιές που ξεπερνούν το k σε m ζαριές, χωρίς ενδιάμεσα να έχει έρθει άσος. Αν τώρα $\{Y_n\}_n$ είναι η αλυσίδα των σκορ μας όταν ρίχνουμε ένα ζάρι χωρίς άσο, έχουμε ότι

$$\mathbb{P}[Y_{n+1} = \ell | Y_n = j] = \frac{1}{5}, \quad \ell \in \{j+2, \dots, j+6\}.$$

Μπορούμε επομένως να γράψουμε

$$\mathbb{E}[K] = \sum_m \sum_{(x_0, x_1, \dots, x_m) \in J_k} x_m \left(\frac{5}{6}\right)^m \mathbb{P}[Y_1 = x_1, \dots, Y_m = x_m] = \mathbb{E}\left[K' \left(\frac{5}{6}\right)^M\right], \quad (6.17)$$

όπου K' είναι το κέρδος μας και M είναι ο αριθμός των ζαριών που ρίξαμε προκειμένου να ξεπεράσουμε το κατώφλι k με το πειραγμένο ζάρι.

γ) Χρησιμοποιήστε την (6.17) για να υπολογίσετε με Monte Carlo ποια είναι η βέλτιστη στρατηγική διακοπής και ποιο είναι το αναμενόμενο κέρδος μας από το παιχνίδι με αυτήν τη στρατηγική.

Κεφάλαιο 7

Εφαρμογές των μαρκοβιανών αλυσίδων

7.1 Εισαγωγή

Στο κεφάλαιο αυτό θα δούμε μερικές εφαρμογές των μαρκοβιανών αλυσίδων στις σύγχρονες επιστήμες και στην τεχνολογία. Θα δούμε γιατί η μηχανή αναζήτησης της εταιρείας Google κυριάρχησε στην αγορά, όταν εμφανίστηκε, αξιοποιώντας μια πολύ απλή ιδέα εμπνευσμένη από μαρκοβιανές αλυσίδες. Στη συνέχεια θα δούμε μια αναλογία μεταξύ μαρκοβιανών αλυσίδων και ηλεκτρικών κυκλωμάτων. Θα γνωρίσουμε ακόμα τη μέθοδο MCMC και τον πολύ σημαντικό αλγόριθμο Metropolis-Hastings που έφεραν επανάσταση στην υπολογιστική προσομοίωση. Τέλος θα γνωρίσουμε την τεχνική της *προσομοιωμένης απόπτωσης* (simulated annealing) και πώς μπορούμε να τη χρησιμοποιήσουμε για να προσεγγίσουμε δύσκολα προβλήματα βελτιστοποίησης, όπως π.χ. το διάσημο πρόβλημα του πλανόδιου πωλητή που προέρχεται από την πληροφορική.

7.2 Μια κατανομή αξίας δισεκατομμυρίων

Οι μεγαλύτεροι σε ηλικία αναγνώστες θα θυμούνται ότι πριν την εμφάνιση της Google οι μηχανές αναζήτησης δεν ήταν τόσο ακριβείς. Στα αποτελέσματα που επέστρεφαν υπήρχαν πολλές σελίδες που δεν είχαν καμία σχέση με την αναζήτηση και έτσι συχνά χρειαζόταν να δει κανείς αρκετές σελίδες αποτελεσμάτων για να βρει σελίδες με χρήσιμο περιεχόμενο. Όταν η Google εμφανίστηκε στην αγορά (1997) και εισήγαγε τη δική της μηχανή αναζήτησης, τα αποτελέσματα ήταν εντυπωσιακά καλύτερα από οτιδήποτε υπήρχε μέχρι τότε. Πολύ σύντομα η συντριπτική πλειονότητα των χρηστών του διαδικτύου χρησιμοποιούσε αυτήν, βάζοντας έτσι τις βάσεις για να γίνει η Google ο κολοσσός που γνωρίζουμε σήμερα.

Το μυστικό της επιτυχίας της συγκεκριμένης μηχανής ήταν ένας αλγόριθμος που χρησιμοποιούσε για να αξιολογεί τις ιστοσελίδες και να προτάσσει στα αποτελέσματα της αναζήτησης εκείνες που αξιολογούνταν ως πιο σημαντικές. Ο αλγόριθμος αυτός είναι ο PageRankTM και παρά τις τροποποιήσεις που έχει υποστεί από τότε η βασική του ιδέα είναι εξαιρετικά απλή και παραμένει επίκαιρη ακόμα και σήμερα (2015).

Φανταστείτε έναν περιηγητή του διαδικτύου που σε κάθε του βήμα επιλέγει τυχαία έναν από τους συνδέσμους της σελίδας στην οποία βρίσκεται και μεταβαίνει εκεί. Μπορούμε να φανταστούμε την περιήγηση

αυτή ως μια μαρκοβιανή αλυσίδα στο σύνολο V των ιστοσελίδων του διαδικτύου με πιθανότητες μετάβασης

$$q(x, y) = \begin{cases} \frac{1}{v_0(x)}, & \text{αν υπάρχει σύνδεσμος από τη σελίδα } x \text{ προς τη σελίδα } y, \\ 0, & \text{διαφορετικά,} \end{cases}$$

όπου $v_0(x)$ είναι το πλήθος των συνδέσμων από τη σελίδα x προς άλλες σελίδες. Ο χώρος καταστάσεων είναι τεράστιος αλλά πεπερασμένος. Ας υποθέσουμε για λίγο ότι η αλυσίδα αυτή είναι μη υποβιβάσιμη, οπότε θα έχει μοναδική αναλλοίωτη κατανομή π . Αυτή η υπόθεση είναι ισοδύναμη με το ότι ο περιηγητής μας μπορεί να φτάσει από οποιαδήποτε σελίδα του διαδικτύου σε οποιαδήποτε άλλη μέσα από ένα μονοπάτι διαδοχικών συνδέσμων. Κάτι τέτοιο δεν είναι ρεαλιστικό, αφού λ.χ. υπάρχουν σελίδες χωρίς εξωτερικούς συνδέσμους, θα μας βοηθήσει όμως να καταλάβουμε γιατί το βάρος $\pi(x)$ που αποδίδει στη σελίδα x η αναλλοίωτη κατανομή π είναι ένας πολύ καλός δείκτης σημαντικότητας της σελίδας x . Στη συνέχεια θα άρουμε τον περιορισμό της μη υποβιβασιμότητας.

Πράγματι, το εργοδικό θεώρημα εξασφαλίζει ότι το ποσοστό του χρόνου θα ξοδεύει ο περιηγητής μας στη σελίδα x είναι ασυμπτωτικά $\pi(x)$. Συγκεκριμένα, αν $N_n(x)$ είναι το πλήθος των επισκέψεων του περιηγητή στη σελίδα x στα πρώτα n βήματά του, τότε

$$\mathbb{P}\left[\lim_{n \rightarrow \infty} \frac{N_n(x)}{n} = \pi(x), \quad \forall x \in V\right] = 1.$$

Με αυτή την έννοια λοιπόν η κατανομή π είναι ένας δείκτης σημαντικότητας.

Μπορούμε επίσης να φανταστούμε ότι η σημαντικότητα $\sigma(x)$ μιας σελίδας x προσδιορίζεται εμμέσως από τη σημαντικότητα των σελίδων που έχουν συνδέσμους προς τη x . Αυτό είναι διαισθητικά επιθυμητό, αφού ένας σύνδεσμος από την ιστοσελίδα της Yahoo ενδεχομένως βαρύνει περισσότερο απ' ό,τι ένας σύνδεσμος από κάποιο άγνωστο ιστολόγιο. Θα θέλαμε επίσης η σημαντικότητα μιας σελίδας να επιμερίζεται στις σελίδες προς τις οποίες παρέχει συνδέσμους, ώστε σημαντικές σελίδες με πολλούς συνδέσμους να βαρύνουν συνολικά όσο εξίσου σημαντικές σελίδες με λίγους εξωτερικούς συνδέσμους. Τις προδιαγραφές αυτές μπορούμε να τις γράψουμε συνοπτικά ως εξής:

$$\sigma(x) = \sum_{y \rightarrow x} \frac{\sigma(y)}{v_0(y)}, \quad \forall x \in V. \quad (7.1)$$

Η παραπάνω σχέση μας λέει απλώς ότι η σημαντικότητα της σελίδας x προκύπτει, αν αθροίσουμε τις σημαντικότητες των σελίδων y που παρέχουν συνδέσμους προς τη x , με βάρη αντιστρόφως ανάλογα του πλήθους $v_0(y)$ των συνδέσμων τους. Αυτός ο ορισμός της σημαντικότητας είναι αυτοαναφορικός, αφού για να ορίσουμε τη σημαντικότητα μιας ιστοσελίδας χρησιμοποιούμε τη σημαντικότητα άλλων σελίδων. Δεν είναι λοιπόν εκ των προτέρων φανερό γιατί οι εξισώσεις (7.1) έχουν λύση. Μπορούμε όμως να ξαναγράψουμε τις εξισώσεις αυτές ως εξής

$$\sigma(x) = \sum_{y \in V} \sigma(y) q(y, x), \quad \forall x \in V.$$

Αν επιπλέον θελήσουμε να κανονικοποιήσουμε τη συνολική σημαντικότητα όλων των ιστοσελίδων στη μονάδα, δηλαδή

$$\sum_{x \in V} \sigma(x) = 1,$$

βλέπουμε ότι αυτές είναι ακριβώς οι εξισώσεις που ορίζουν τη μοναδική αναλλοίωτη κατανομή π της μαρκοβιανής αλυσίδας μας. Άρα όχι μόνο έχουν λύση, αλλά η λύση αυτή είναι μοναδική και είναι η π .

Ελπίζω τα παραπάνω επιχειρήματα να σας έχουν πείσει ότι, αν ο τυχαίος περίπατος στο διαδίκτυο που περιγράψαμε αντιστοιχούσε σε μια μη υποβιβάσιμη αλυσίδα, τότε η αναλλοίωτη κατανομή της θα ήταν ένας πολύ καλός δείκτης σημαντικότητας. Όπως προείπαμε όμως, αυτός ο περιορισμός είναι πολύ ισχυρός. Για να τον άρουμε τροποποιούμε την περιήγησή μας ως εξής. Πριν από κάθε μας βήμα στρίβουμε ένα κέρμα που φέρνει κεφαλή με πιθανότητα $\alpha \in (0, 1)$. Αν το κέρμα έρθει κεφαλή, τότε όπως και πριν επιλέγουμε τυχαία μία από τις ιστοσελίδες προς τις οποίες υπάρχει σύνδεσμος και μεταβαίνουμε σε αυτήν. Αν το κέρμα έρθει γράμματα, τότε επιλέγουμε τυχαία μία από όλες τις σελίδες του διαδικτύου και μεταβαίνουμε εκεί. Αν $|V|$ είναι το πλήθος των ιστοσελίδων του διαδικτύου, η παραπάνω περιήγηση αντιστοιχεί σε μια μαρκοβιανή αλυσίδα στο V με πιθανότητες μετάβασης

$$p(x, y) = \alpha q(x, y) + \frac{(1 - \alpha)}{|V|}, \quad \forall x, y \in V.$$

Η αλυσίδα αυτή είναι μη υποβιβάσιμη για κάθε $\alpha < 1$ και ο δείκτης PageRankTM είναι στην ουσία η αναλλοίωτη κατανομή αυτής της αλυσίδας για την επιλογή $\alpha = 0,85$.

7.3 Μαρκοβιανές αλυσίδες και ηλεκτρικά κυκλώματα

Σε αυτή την παράγραφο θα εξερευνήσουμε μια αναλογία που υπάρχει ανάμεσα σε χρονικά αντιστρέψιμες μαρκοβιανές αλυσίδες και σε ηλεκτρικά κυκλώματα. Μέσα από αυτή την αναλογία θα κατανοήσουμε την προέλευση της ορολογίας που χρησιμοποιήσαμε στο Κεφάλαιο 3 και θα αποδείξουμε την αρχή του Rayleigh.

Ας θεωρήσουμε ένα ηλεκτρικό κύκλωμα που αποτελείται από αντιστάσεις, αγωγούς και μία πηγή συνεχούς ρεύματος με διαφορά δυναμικού 1 ανάμεσα στους δύο πόλους της. Χωρίς βλάβη της γενικότητας, ας υποθέσουμε ότι το δυναμικό στον αρνητικό πόλο της πηγής είναι 0, ενώ στον θετικό πόλο είναι 1. Ας συμβολίζουμε με \mathbb{X} το σύνολο των κόμβων του κυκλώματος. Ένα υποσύνολο A από αυτούς τους κόμβους είναι απευθείας συνδεδεμένο στον θετικό πόλο της πηγής, επομένως το δυναμικό $\Phi(x)$ σε κάθε κόμβο $x \in A$ είναι 1. Αντίστοιχα, ένα υποσύνολο B των κόμβων του κυκλώματος είναι συνδεδεμένο στον αρνητικό πόλο της πηγής, επομένως το δυναμικό $\Phi(x)$ σε κάθε κόμβο $x \in B$ είναι 0. Δεν θέλουμε να βραχυκυκλώσουμε την πηγή, οπότε θα πρέπει $A \cap B = \emptyset$. Ποιο είναι όμως το δυναμικό στους υπόλοιπους κόμβους;

Ανάμεσα σε δύο κόμβους $x, y \in \mathbb{X}$ υπάρχει μια αγωγιμότητα $c(x, y) = c(y, x)$. Αν οι κόμβοι συνδέονται μέσω μιας αντίστασης, αυτή η αγωγιμότητα είναι απλά το αντίστροφο της αντίστασης. Διαφορετικά, η μεταξύ τους αγωγιμότητα είναι μηδέν. Οι νόμοι που καθορίζουν την τιμή του δυναμικού σε κάθε κόμβο είναι δύο. Ο Νόμος του Ohm ορίζει ότι το ρεύμα από τον κόμβο y προς τον κόμβο x δίνεται από τη σχέση

$$i(y, x) = c(x, y)(\Phi(y) - \Phi(x)), \quad \text{για κάθε } x, y \in \mathbb{X}.$$

Ο Νόμος του Kirchoff ορίζει ότι σε κάθε κόμβο x του κυκλώματος που δεν είναι συνδεδεμένος με την πηγή, το συνολικό ρεύμα που εισέρχεται από άλλους κόμβους είναι μηδέν, δηλαδή

$$\sum_{y \in \mathbb{X}} i(y, x) = 0, \quad \text{για κάθε } x \notin A \cup B.$$

Συνδυάζοντας τις δύο παραπάνω σχέσεις, έχουμε ότι η συνάρτηση που δίνει το δυναμικό Φ σε κάθε κόμβο λύνει το ΠΣΤ

$$\begin{cases} \sum_{y \in \mathbb{X}} c(x, y)(\Phi(y) - \Phi(x)) = 0, & \text{αν } x \notin A \cup B \\ \Phi(x) = 1, & \text{αν } x \in A \\ \Phi(x) = 0, & \text{αν } x \in B. \end{cases} \quad (7.2)$$

Το παραπάνω πρόβλημα μοιάζει με το ΠΣΤ (3.3) που δίνει τις πιθανότητες $\mathbb{P}_x[T_A < T_B]$, με τη διαφορά ότι οι αγωγιμότητες δεν είναι πιθανότητες μετάβασης, αφού δεν αθροίζονται απαραίτητα στη μονάδα. Αυτό όμως μπορούμε να το διορθώσουμε κανονικοποιώντας. Αν ορίσουμε

$$p(x, y) = \frac{c(x, y)}{\sum_{y \in \mathbb{X}} c(x, y)} = \frac{c(x, y)}{w(x)}, \quad (7.3)$$

τότε τα ΠΣΤ (3.3) και (7.2) ταυτίζονται και μάλιστα αν το κύκλωμα είναι συνεκτικό, αν δηλαδή οποιοδήποτε δύο κόμβοι επικοινωνούν μέσα από ένα μονοπάτι με θετικές αγωγιμότητες, αυτά τα ΠΣΤ έχουν μοναδική λύση. Επομένως, το δυναμικό στον κόμβο $x \in \mathbb{X}$ ταυτίζεται με την πιθανότητα $\mathbb{P}_x[T_A < T_B]$ για μια μαρκοβιανή αλυσίδα, με πιθανότητες μετάβασης που δίνονται από την (7.3).

Παρατηρήστε ότι οι πιθανότητες μετάβασης αυτής της αλυσίδας βρίσκονται σε ακριβή ισορροπία με τους παράγοντες κανονικοποίησης $w(x) = \sum_{y \in \mathbb{X}} c(x, y)$. Πράγματι, από την (7.3) βλέπουμε ότι

$$w(x)p(x, y) = c(x, y) = c(y, x) = w(y)p(y, x).$$

Επομένως, αν ορίσουμε

$$\pi(x) = \frac{w(x)}{\sum_{x \in \mathbb{X}} w(x)} = \frac{\sum_{y \in \mathbb{X}} c(x, y)}{\sum_{z, y \in \mathbb{X}} c(z, y)},$$

το Λήμμα 7 μας δίνει ότι η π είναι αναλλοίωτη κατανομή της μαρκοβιανής αλυσίδας με πιθανότητες μετάβασης $\{p(x, y)\}_{x, y \in \mathbb{X}}$, που δίνονται από την (7.3) και η αλυσίδα αυτή είναι χρονικά αντιστρέψιμη. Μάλιστα, αν το κύκλωμα είναι συνεκτικό, η αλυσίδα αυτή είναι μη υποβιβάσιμη και η π είναι η μοναδική αναλλοίωτη κατανομή της.

Αντίστροφα, σε κάθε χρονικά αντιστρέψιμη μαρκοβιανή αλυσίδα σ' έναν πεπερασμένο χώρο καταστάσεων \mathbb{X} , με πιθανότητες μετάβασης $\{p(x, y)\}_{x, y \in \mathbb{X}}$ και αναλλοίωτη κατανομή π , μπορούμε να αντιστοιχίσουμε ένα ηλεκτρικό κύκλωμα, θεωρώντας ένα κόμβο σε κάθε κατάσταση του \mathbb{X} και τοποθετώντας αγωγιμότητα

$$c(x, y) = \pi(x)p(x, y) = \pi(y)p(y, x)$$

ανάμεσα στους κόμβους που αντιστοιχούν στις καταστάσεις x, y . Αν τώρα ενώσουμε ένα σύνολο κόμβων A αυτού του κυκλώματος με τον θετικό πόλο μιας πηγής κι ένα σύνολο κόμβων B ($A \cap B = \emptyset$) με τον αρνητικό πόλο της πηγής, τότε το δυναμικό $\Phi(x)$ σε κάθε κόμβο x αυτού του κυκλώματος θα ταυτίζεται με την πιθανότητα $\mathbb{P}_x[T_A < T_B]$. Με άλλα λόγια μπορούμε να μετρήσουμε αυτές τις πιθανότητες μ' ένα πολύμετρο.

Η ισχύς που καταναλώνεται σε μια αγωγιμότητα $c(x, y)$ δίνεται σε σχέση με τη διαφορά δυναμικού στα άκρα της από την ποσότητα $c(x, y)(\Phi(y) - \Phi(x))^2$. Το ακόλουθο θεώρημα εξασφαλίζει ότι η κατανομή του δυναμικού σε κάθε κόμβο ενός ηλεκτρικού κυκλώματος γίνεται έτσι ώστε να ελαχιστοποιείται η συνολική ισχύς που καταναλώνεται στο κύκλωμα.

Θεώρημα 32 Από όλες τις συναρτήσεις στο σύνολο

$$F_{A,B} = \{\Psi : \mathbb{X} \rightarrow \mathbb{R} : \Psi(x) = 1 \text{ για κάθε } x \in A, \Psi(x) = 0 \text{ για κάθε } x \in B\},$$

η λύση του ΠΣΤ (7.2) είναι η μόνη που ελαχιστοποιεί την

$$\mathcal{D}(\Psi) = \frac{1}{2} \sum_{x, y \in \mathbb{X}} c(x, y)(\Psi(y) - \Psi(x))^2. \quad (7.4)$$

Απόδειξη: Έστω $\Psi : \mathbb{X} \rightarrow \mathbb{R}$ μια συνάρτηση για την οποία $\Psi(x) = 1$ για κάθε $x \in A$ και $\Psi(x) = 0$ για κάθε $x \in B$. Αν Φ είναι η λύση του ΠΣΤ (7.2), τότε η $h = \Psi - \Phi$ μηδενίζεται στα A και B . Αρκεί λοιπόν να δείξουμε ότι

$$\mathcal{D}(\Phi) \leq \mathcal{D}(\Phi + h)$$

για κάθε συνάρτηση $h : \mathbb{X} \rightarrow \mathbb{R}$ τέτοια ώστε $h(x) = 0$ για κάθε $x \in A \cup B$. Έχουμε όμως ότι

$$\begin{aligned} \mathcal{D}(\Phi + h) &= \frac{1}{2} \sum_{x,y \in \mathbb{X}} c(x,y) (\Phi(y) + h(y) - \Phi(x) - h(x))^2 \\ &= \mathcal{D}(\Phi) + \mathcal{D}(h) + \frac{1}{2} \sum_{x,y \in \mathbb{X}} c(x,y) (\Phi(y) - \Phi(x)) (h(y) - h(x)) \\ &= \mathcal{D}(\Phi) + \mathcal{D}(h) + \frac{1}{2} \sum_{x,y \in \mathbb{X}} c(x,y) (\Phi(y) - \Phi(x)) h(y) - \frac{1}{2} \sum_{x,y \in \mathbb{X}} c(x,y) (\Phi(y) - \Phi(x)) h(x). \end{aligned}$$

Αν εναλλάξουμε τον ρόλο των x και y στο πρώτο άθροισμα της τελευταίας έκφρασης και χρησιμοποιήσουμε ότι η αγωγιμότητα είναι συμμετρική, δηλαδή $c(x,y) = c(y,x)$ για κάθε $x,y \in \mathbb{X}$ παίρνουμε ότι

$$\begin{aligned} \mathcal{D}(\Phi + h) &= \mathcal{D}(\Phi) + \mathcal{D}(h) - \sum_{x,y \in \mathbb{X}} c(x,y) (\Phi(y) - \Phi(x)) h(x) \\ &= \mathcal{D}(\Phi) + \mathcal{D}(h) - \sum_{x \in \mathbb{X}} h(x) \sum_{y \in \mathbb{X}} c(x,y) (\Phi(y) - \Phi(x)). \end{aligned}$$

Από το ΠΣΤ (7.2) που ικανοποιεί η Φ έχουμε ότι, αν $x \notin A \cup B$, τότε $\sum_{y \in \mathbb{X}} c(x,y) (\Phi(y) - \Phi(x)) = 0$. Επιπλέον, αν $x \in A \cup B$, τότε $h(x) = 0$. Σε κάθε περίπτωση ο τελευταίος όρος της παραπάνω σχέσης μηδενίζεται, οπότε

$$\mathcal{D}(\Phi + h) = \mathcal{D}(\Phi) + \mathcal{D}(h) \geq \mathcal{D}(\Phi),$$

αφού η $\mathcal{D}(h)$ είναι εξ ορισμού μη αρνητική. Μάλιστα, για να έχουμε ισότητα στην παραπάνω σχέση θα πρέπει $\mathcal{D}(h) = 0$. Εφόσον η $\mathcal{D}(h)$ είναι άθροισμα μη αρνητικών όρων, αυτό σημαίνει ότι $h(y) = h(x)$, όποτε έχουμε $c(x,y) > 0$. Μια και το κύκλωμα είναι συνεκτικό, μπορούμε να συνδέσουμε οποιονδήποτε κόμβο του, y , με το A , μέσω ενός μονοπατιού κατά μήκος του οποίου συναντάμε θετικές αγωγιμότητες. Έτσι η h πρέπει να είναι σταθερή στους κόμβους αυτού του μονοπατιού και εφόσον μηδενίζεται στο A , έχουμε $h(y) = 0$. \square

Πόρισμα 11 (Αρχή του Rayleigh) Αν για δύο κυκλώματα με κόμβους \mathbb{X} και αγωγιμότητες $\{c(x,y)\}_{x,y \in \mathbb{X}}$ και $\{\tilde{c}(x,y)\}_{x,y \in \mathbb{X}}$ αντίστοιχα, έχουμε ότι $c(x,y) \leq \tilde{c}(x,y)$ για κάθε $x,y \in \mathbb{X}$, τότε η ισχύς που καταναλώνεται στο πρώτο είναι πάντα μικρότερη από την ισχύ που καταναλώνεται στο δεύτερο.

Απόδειξη: Για $\Psi \in F_{A,B}$ ορίζουμε την $\tilde{\mathcal{D}}(\Psi)$ κατ' αναλογία με την (7.4), χρησιμοποιώντας τις αγωγιμότητες $\tilde{c}(x,y)$. Από την υπόθεση ότι $c(x,y) \leq \tilde{c}(x,y)$ για κάθε $x,y \in \mathbb{X}$, έχουμε

$$\mathcal{D}(\Psi) \leq \tilde{\mathcal{D}}(\Psi), \quad \text{για κάθε } \Psi \in F_{A,B}.$$

Επομένως $\inf_{\Psi \in F_{A,B}} \mathcal{D}(\Psi) \leq \inf_{\Psi \in F_{A,B}} \tilde{\mathcal{D}}(\Psi)$. Από το Θεώρημα 32 όμως, το αριστερό μέλος αυτής της ανίσωσης είναι η ισχύς που καταναλώνεται στο πρώτο κύκλωμα, ενώ το δεξί μέλος είναι η ισχύς που καταναλώνεται στο δεύτερο. \square

Εφόσον η ποσότητα

$$C(A,B) = \mathcal{D}(\Phi) = \inf_{\Psi \in F_{A,B}} \mathcal{D}(\Psi)$$

είναι η ισχύς που καταναλώνει ολόκληρο το κύκλωμα όταν συνδέσουμε το σύνολο κόμβων A στον θετικό πόλο και το σύνολο κόμβων B στον αρνητικό πόλο μιας πηγής που δημιουργεί διαφορά δυναμικού 1, αυτή η ποσότητα είναι η ισοδύναμη αγωγιμότητα (ή χωρητικότητα) του κυκλώματος ανάμεσα στο A και στο B . Μάλιστα, εύκολα μπορεί να δει κανείς ότι $C(A, B) = C(B, A)$, αφού, αν η Φ λύνει το ΠΣΤ (7.2), τότε η $1-\Phi$ λύνει το αντίστοιχο ΠΣΤ με τον ρόλο των A και B εναλλαγμένο, ενώ $\mathcal{D}(\Phi) = \mathcal{D}(1 - \Phi)$.

Εντελώς ανάλογα αποτελέσματα με την ίδια ακριβώς απόδειξη έχουμε για μια χρονικά αντιστρέψιμη μαρκοβιανή αλυσίδα. Την αντίστοιχη ποσότητα

$$C(A, B) = \inf_{\Psi \in F_{A,B}} \mathcal{D}(\Psi) = \mathcal{D}(\Phi_{A,B}) = \frac{1}{2} \sum_{x,y \in \mathbb{X}} \pi(x)p(x,y)(\Phi_{A,B}(y) - \Phi_{A,B}(x))^2$$

ονομάζουμε *χωρητικότητα* (capacity) ανάμεσα στα σύνολα καταστάσεων A και B . Συμβολίζοντας για απλότητα τη συνάρτηση δυναμικού $\Phi_{A,B}$ με Φ , μπορούμε να ξαναγράψουμε τη χωρητικότητα, ως εξής.

$$C(A, B) = \frac{1}{2} \sum_{x,y \in \mathbb{X}} \pi(x)p(x,y)(\Phi(y) - \Phi(x))\Phi(y) - \frac{1}{2} \sum_{x,y \in \mathbb{X}} \pi(x)p(x,y)(\Phi(y) - \Phi(x))\Phi(x).$$

Αν εναλλάξουμε τον ρόλο των x και y στο πρώτο άθροισμα και χρησιμοποιήσουμε τη συνθήκη ακριβούς ισορροπίας, παίρνουμε περαιτέρω

$$\begin{aligned} C(A, B) &= \frac{1}{2} \sum_{x,y \in \mathbb{X}} \pi(y)p(y,x)(\Phi(x) - \Phi(y))\Phi(x) - \frac{1}{2} \sum_{x,y \in \mathbb{X}} \pi(x)p(x,y)(\Phi(y) - \Phi(x))\Phi(x) \\ &= - \sum_{x,y \in \mathbb{X}} \pi(x)p(x,y)(\Phi(y) - \Phi(x))\Phi(x) = - \sum_{x \in \mathbb{X}} \pi(x)\Phi(x)L\Phi(x). \end{aligned}$$

Όμως από το ΠΣΤ (3.3) που ικανοποιεί η Φ έχουμε ότι $L\Phi(x) = 0$ για $x \notin A \cup B$ και $\Phi(x) = 0$ για $x \in B$. Επομένως, στο τελευταίο άθροισμα επιζούν μόνο οι όροι για $x \in A$, για τους οποίους έχουμε $\Phi(x) = 1$. Έτσι,

$$\begin{aligned} C(A, B) &= - \sum_{x \in A} \pi(x)L\Phi(x) = \sum_{x \in A} \pi(x) \sum_{y \in \mathbb{X}} p(x,y)(\Phi(x) - \Phi(y)) \\ &= \sum_{x \in A} \pi(x) \sum_{y \in \mathbb{X}} p(x,y)(1 - \Phi(y)) = \sum_{x \in A} \pi(x) \sum_{y \in \mathbb{X}} p(x,y)\mathbb{P}_y[T_B < T_A] \\ &= \sum_{x \in A} \pi(x)\mathbb{P}_x[T_B^+ < T_A^+] = \sum_{x \in A} \pi(x)\mathbb{P}_x[T_B < T_A^+]. \end{aligned} \tag{7.5}$$

Στην προτελευταία ισότητα χρησιμοποιήσαμε ανάλυση πρώτου βήματος και τη μαρκοβιανή ιδιότητα και στην τελευταία ότι $A \cap B = \emptyset$. Στην ειδική περίπτωση που $A = \{x\}$ και $B = \{y\}$ έχουμε ότι

$$C(x, y) = \pi(x)\mathbb{P}_x[T_y < T_x^+].$$

Χρησιμοποιώντας τώρα τη συμμετρία της χωρητικότητας, $C(A, B) = C(B, A)$, φτάνουμε στο καθόλου τετριμμένο αποτέλεσμα ότι για οποιεσδήποτε καταστάσεις $x, y \in \mathbb{X}$ μιας χρονικά αντιστρέψιμης μαρκοβιανής αλυσίδας έχουμε

$$\pi(x)\mathbb{P}_x[T_y < T_x^+] = \pi(y)\mathbb{P}_y[T_x < T_y^+].$$

Μπορείτε να διαβάσετε περισσότερα για την αναλογία μαρκοβιανών αλυσίδων και ηλεκτρικών κυκλωμάτων στο εξαιρετικό βιβλίο [3]. Εκεί μπορείτε να βρείτε και μια απόδειξη πως ο απλός, συμμετρικός τυχαίος περίπατος στο \mathbb{Z}^d είναι επαναληπτικός όταν $d = 1$ ή 2 αλλά παροδικός όταν $d > 2$, χρησιμοποιώντας αυτή την αναλογία και την Αρχή του Rayleigh.

7.4 Ο αλγόριθμος Metropolis-Hastings και το μοντέλο Ising

Η κλασική μέθοδος Monte Carlo χρησιμοποιεί τον νόμο των μεγάλων αριθμών για να υπολογίσει τη μέση τιμή βάσει κάποιας κατανομής π ως ένα μέσο όρο ανεξάρτητων δειγμάτων. Συγκεκριμένα, αν X_1, X_2, \dots είναι ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με κατανομή π , ο νόμος των μεγάλων αριθμών εξασφαλίζει ότι

$$\mathbb{P}\left[\lim_{N \rightarrow \infty} \frac{f(X_1) + f(X_2) + \dots + f(X_N)}{N} = \mathbb{E}^\pi[f]\right] = 1.$$

Συνήθως χρησιμοποιεί κανείς τη μέθοδο Monte Carlo για να υπολογίσει την αναμενόμενη τιμή $\mathbb{E}^\pi[f]$ προσομοιώνοντας ανεξάρτητα δείγματα X_1, X_2, \dots, X_N από την κατανομή π και παίρνοντας τον μέσο όρο των $f(X_i)$. Ο νόμος των μεγάλων αριθμών μας εξασφαλίζει ότι με πιθανότητα 1, όταν το N είναι μεγάλο, οι δύο ποσότητες είναι κοντά, ενώ το κεντρικό οριακό θεώρημα μπορεί να μας δώσει εκτιμήσεις για το τυπικό σφάλμα της προσέγγισης.

Σε πολλές ενδιαφέρουσες εφαρμογές το να πάρει κανείς δείγματα από την κατανομή π είναι εξαιρετικά δύσκολο. Αυτό μπορεί να συμβεί γιατί ο χώρος \mathbb{X} στον οποίο έχει οριστεί η κατανομή π είναι πολύπλοκος και η κατανομή π δεν είναι ακριβώς γνωστή. Η μέθοδος MCMC προσομοιώνει την π ως την κατανομή ισορροπίας μιας μαρκοβιανής αλυσίδας. Η προσομοίωση της μαρκοβιανής αλυσίδας είναι συνήθως πολύ εύκολη, αφού η δεσμευμένη κατανομή της X_{n+1} δεδομένης της X_n δίνεται από τις πιθανότητες μετάβασης της αλυσίδας. Έχουμε ήδη δει πώς γίνεται αυτό στα αριθμητικά πειράματα. Χρειάζεται όμως να κατασκευάσουμε κατάλληλες πιθανότητες μετάβασης που θα μας εξασφαλίζουν ότι η $\{X_n\}$ έχει κατανομή ισορροπίας την π . Το πρόβλημα αυτό δεν έχει μοναδική λύση, αλλά μία από τις πιο επιτυχημένες προσεγγίσεις είναι ο αλγόριθμος Metropolis-Hastings.

Ο αλγόριθμος αυτός μπορεί να περιγραφεί ως εξής. Χωρίς βλάβη της γενικότητας μπορούμε να υποθέσουμε ότι $\pi(x) > 0$ για κάθε $x \in \mathbb{X}$. Θεωρούμε έναν περίπατο στο \mathbb{X} με πιθανότητες μετάβασης $r(x, y)$ από την κατάσταση x στην κατάσταση y . Αν κάποια στιγμή βρισκόμαστε στην κατάσταση $x \in \mathbb{X}$ επιλέγουμε την υποψήφια επόμενη κατάσταση y με πιθανότητα $r(x, y)$. Στη συνέχεια αποφασίζουμε αν θα μεταβούμε στην y με πιθανότητα να το κάνουμε ίση με

$$\frac{\pi(y)r(y, x)}{\pi(x)r(x, y)} \wedge 1 = \min\left\{\frac{\pi(y)r(y, x)}{\pi(x)r(x, y)}, 1\right\}.$$

Η πιθανότητα μετάβασης από την x στην $y \neq x$ είναι το γινόμενο της πιθανότητας να επιλέξουμε την y ως υποψήφια επόμενη κατάσταση με την πιθανότητα να πραγματοποιήσουμε τη μετάβαση, δεδομένου ότι επιλέξαμε την y . Συγκεκριμένα,

$$p(x, y) = r(x, y) \left(\frac{\pi(y)r(y, x)}{\pi(x)r(x, y)} \wedge 1 \right), \quad y \neq x. \quad (7.6)$$

Επομένως, η πιθανότητα να παραμείνουμε στην κατάσταση x είναι

$$p(x, x) = 1 - \sum_{z \neq x} p(x, z) \geq 1 - \sum_{z \neq x} r(x, z) = r(x, x).$$

Από την παραπάνω ανισότητα βλέπουμε ότι, αν $r(x, x) > 0$ για κάποιο $x \in \mathbb{X}$, τότε η αλυσίδα που κατασκευάσαμε είναι αυτόματα απεριοδική. Ακόμα όμως κι αν επιλέξουμε $r(x, x) = 0$ για κάθε $x \in \mathbb{X}$, αν οι πιθανότητες μετάβασης $\{r(x, y)\}_{x, y \in \mathbb{X}}$ δεν είναι σε ακριβή ισορροπία με την π , τότε υπάρχουν $x, y \in \mathbb{X}$ με $\pi(y)r(y, x) < \pi(x)r(x, y)$. Επομένως,

$$p(x, x) > 1 - \sum_{z \neq x} r(x, z) = r(x, x) \geq 0$$

και πάλι η αλυσίδα με πιθανότητες μετάβασης $\{p(x, y)\}_{x, y \in \mathbb{X}}$ είναι απεριοδική. Η κατανομή π που θέλουμε να προσομοιώσουμε και οι πιθανότητες μετάβασης $\{p(x, y)\}_{x, y \in \mathbb{X}}$ βρίσκονται σε ακριβή ισορροπία. Πράγματι, για $x \neq y$ έχουμε

$$\pi(x) p(x, y) = \pi(x) r(x, y) \left(\frac{\pi(y) r(y, x)}{\pi(x) r(x, y)} \wedge 1 \right) = \pi(y) r(y, x) \wedge \pi(x) r(x, y) = \pi(y) p(y, x).$$

Σύμφωνα με το Λήμμα 1, η π είναι αναλλοίωτη κατανομή της αλυσίδας $\{X_n\}$ με πιθανότητες μετάβασης $\{p(x, y)\}_{x, y \in \mathbb{X}}$. Επιπλέον, η $\{X_n\}$ είναι μη υποβιβάζσιμη, όταν ο περίπατος που ορίζουν οι πιθανότητες μετάβασης $\{r(x, y)\}_{x, y \in \mathbb{X}}$ είναι μη υποβιβάζσιμος. Επομένως, όπως είδαμε στο προηγούμενο Κεφάλαιο 6, έχουμε

$$\mathbb{P}[X_n = x] \xrightarrow{n \rightarrow \infty} \pi(x).$$

Μπορούμε λοιπόν να πάρουμε δείγματα από την κατανομή π , προσομοιώνοντας την $\{X_n\}_{n \in \mathbb{N}_0}$ για αρκετό χρόνο ώστε να έρθει κοντά στην κατανομή ισορροπίας της.

Ένα κλασικό παράδειγμα κατανομών που παρουσιάζουν μεγάλο ενδιαφέρον και συχνά είναι δύσκολο να προσομοιωθούν είναι οι κατανομές Gibbs. Αυτές εμφανίζονται φυσικά σε συστήματα που βρίσκονται σε κατάσταση θερμοδυναμικής ισορροπίας σε θερμοκρασία T , οπότε η πιθανότητα να βρούμε ένα τέτοιο σύστημα σε μια κατάσταση x με ενέργεια $H(x)$ είναι

$$\pi(x) = \frac{1}{Z(\beta)} e^{-\beta H(x)}.$$

Στην παραπάνω σχέση $\beta = \frac{1}{kT}$ είναι ο παράγοντας Boltzmann. Η $Z(\beta)$, η οποία στη βιβλιογραφία της Φυσικής αναφέρεται ως *συνάρτηση επιμερισμού* (partition function), είναι μια σταθερά κανονικοποίησης που δίνεται από τη σχέση

$$Z(\beta) = \sum_{x \in \mathbb{X}} e^{-\beta H(x)}.$$

Η συνάρτηση επιμερισμού είναι ό,τι θα ήθελε να ξέρει ένας Στατιστικός Φυσικός για το σύστημα που μελετά. Πολλές θερμοδυναμικές παράμετροι του συστήματος, όπως η ελεύθερη ενέργεια, η εντροπία ή η πίεση, μπορούν να υπολογιστούν από την $Z(\cdot)$ και τις παραγώγους της.

Συνήθως όμως είναι δύσκολο να υπολογιστεί το παραπάνω άθροισμα. Σ' ένα απλό μοντέλο μαγνήτισης για παράδειγμα, στις κορυφές ενός τετραγωνικού πλέγματος $L \times L$ τοποθετούμε σωματίδια που έχουν σπιν $+1$ ή -1 . Η κατάσταση του συστήματος x περιγράφεται από το σπιν σε κάθε κορυφή του πλέγματος, επομένως ο χώρος καταστάσεων \mathbb{X} έχει πληθώραριθμό 2^{L^2} . Για ένα πλέγμα 100×100 ατόμων ο \mathbb{X} αριθμεί $2^{10000} \geq 10^{3000}$ καταστάσεις. Αυτός είναι ένας τεράστιος αριθμός. Τα άτομα στο σύμπαν εκτιμάται ότι είναι της τάξης του 10^{80} . Παρότι λοιπόν η συνάρτηση ενέργειας του συστήματος H μπορεί να είναι γνωστή, είναι συχνά δύσκολο να πάρει κανείς δείγματα από την κατανομή π γιατί δεν μπορεί να υπολογίσει το άθροισμα στον ορισμό της συνάρτησης επιμερισμού.

Με τον αλγόριθμο Metropolis-Hastings έχουμε το πλεονέκτημα ότι δεν χρειάζεται να γνωρίζουμε τη Z για να προσομοιώσουμε την αλυσίδα. Συγκεκριμένα, επιλέγοντας για απλότητα μια συμμετρική r , $r(x, y) = r(y, x)$ για κάθε $x, y \in \mathbb{X}$, οι πιθανότητες μετάβασης είναι για $x \neq y$

$$p(x, y) = r(x, y) \left(\frac{\pi(y)}{\pi(x)} \wedge 1 \right) = r(x, y) \left(e^{-\beta(H(y)-H(x))} \wedge 1 \right) = r(x, y) e^{-\beta(H(y)-H(x))}.$$

Μπορούμε να επιλέξουμε τις $r(x, y)$ να είναι 0, εκτός αν οι καταστάσεις x, y διαφέρουν μόνο ως προς το σπιν μίας από τις L^2 θέσεις, οπότε $r(x, y) = 1/L^2$. Σε αυτή την περίπτωση ο αλγόριθμος Metropolis-Hastings θα εφαρμοζόταν ως εξής:

1. Επιλέγουμε τυχαία ένα σωματίδιο στο πλέγμα.
2. Υπολογίζουμε πόσο μεταβάλλεται η ενέργεια του συστήματος, αν αλλάξουμε το σπιν αυτού του σωματιδίου. Αν x είναι η παρούσα κατάσταση και y η κατάσταση που προκύπτει, αν αλλάξουμε το σπιν του σωματιδίου που επιλέξαμε στο βήμα 1, υπολογίζουμε τη διαφορά $\Delta H = H(y) - H(x)$.
3. Αν $\Delta H \leq 0$, αλλάζουμε το σπιν του σωματιδίου και η νέα κατάσταση γίνεται η y . Αν $\Delta H > 0$, τότε αλλάζουμε το σπιν του σωματιδίου με πιθανότητα $e^{-\beta\Delta H}$, διαφορετικά παραμένουμε στην κατάσταση x .
4. Επιστρέφουμε στο βήμα 1.

Στον κώδικα `ising.py` υλοποιούμε σε γλώσσα Python τον αλγόριθμο για το μοντέλο Ising, σύμφωνα με το οποίο η ενέργεια του συστήματος στην κατάσταση x δίνεται από τη σχέση

$$H(x) = - \sum_{i \sim j} x(i)x(j).$$

Στην παραπάνω σχέση $x(i), x(j) \in \{-1, +1\}$ είναι τα σπιν της κατάστασης x στις θέσεις i, j αντίστοιχα, ενώ ο συμβολισμός $i \sim j$ δείχνει ότι η άθροιση εκτείνεται σε ζευγάρια σωματιδίων που γειτνιάζουν στο πλέγμα. Το πλέγμα που χρησιμοποιούμε στον κώδικα έχει $L = 128$, ενώ ο παράγοντας Boltzmann είναι $\beta = 1/3$. Προσέξτε ότι εφόσον $\beta > 0$, το μέτρο Gibbs π αποδίδει μεγαλύτερη πιθανότητα σε καταστάσεις που έχουν χαμηλή ενέργεια. Στον υπολογισμό της ενέργειας μιας κατάστασης x στο μοντέλο Ising κάθε ζευγάρι γειτονικών σωματιδίων συνεισφέρει -1 , αν τα σπιν τους έχουν τον ίδιο προσανατολισμό, και $+1$, αν έχουν αντίθετο. Άρα καταστάσεις όπου γειτονικά σωματίδια έχουν το ίδιο σπιν είναι πιο πιθανές, γ' αυτό και το μοντέλο Ising είναι ένα μοντέλο σιδηρομαγνητισμού.

Παρατηρήστε ότι, επειδή $r(x, y) \neq 0$ μόνο αν τα σπιν των καταστάσεων x, y διαφέρουν σε μία μόνο θέση, δεν είναι ανάγκη να υπολογίσουμε ολόκληρη τη συνάρτηση ενέργειας στο 2ο βήμα παραπάνω. Συγκεκριμένα, αν η y προκύπτει από τη x αλλάζοντας το σπιν του σωματιδίου που βρίσκεται στη θέση k , δηλαδή

$$y(j) = \begin{cases} x(j) & j \neq k \\ -x(j) & j = k, \end{cases}$$

τότε

$$\Delta H = H(y) - H(x) = - \sum_{i \sim j} y(i)y(j) + \sum_{i \sim j} x(i)x(j) = 2x(k) \sum_{j: j \sim k} x(j).$$

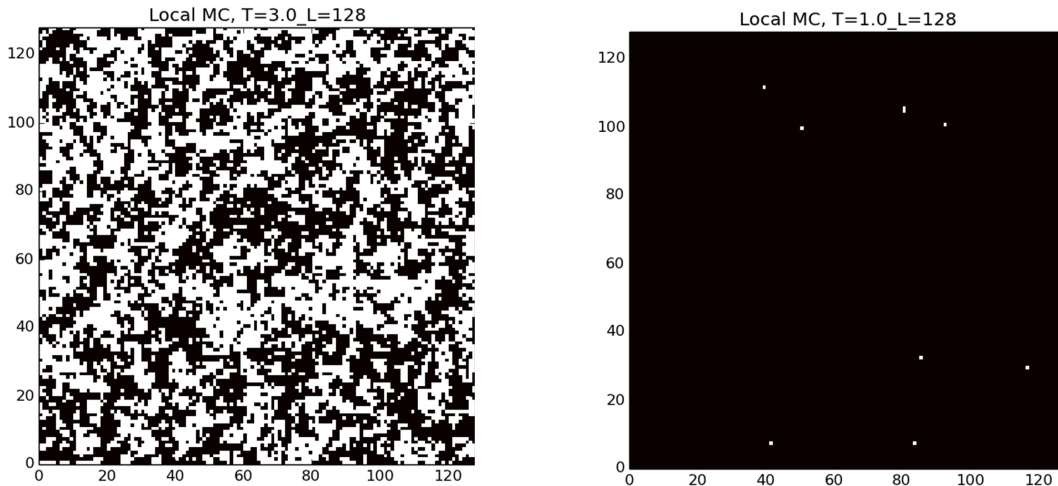
Επομένως, προκειμένου να υπολογίσουμε τη μεταβολή στην ενέργεια που προκύπτει από την αλλαγή του σπιν ενός σωματιδίου, αρκεί να υπολογίσουμε το αθροιστικό σπιν των γειτόνων αυτού του σωματιδίου.

Ας μελετήσουμε τώρα τον ρόλο του παράγοντα Boltzmann. Θυμηθείτε ότι

$$\pi(x) = \frac{e^{-\beta H(x)}}{\sum_{y \in \mathbb{X}} e^{-\beta H(y)}}.$$

Για $\beta \rightarrow \infty$, δηλαδή όταν η θερμοκρασία $T \downarrow 0$, η κατανομή π τείνει να συγκεντρωθεί στις καταστάσεις με τη μικρότερη δυνατή ενέργεια. Στο μοντέλο Ising αυτές είναι οι δύο καταστάσεις x_+ και x_- όπου όλα τα σπιν έχουν τον ίδιο προσανατολισμό, όλα $+1$ στην κατάσταση x_+ , ή όλα -1 στην κατάσταση x_- . Συγκεκριμένα,

$$\pi(x_+) \rightarrow \frac{1}{2}, \quad \pi(x_-) \rightarrow \frac{1}{2} \quad \text{καθώς } \beta \rightarrow \infty.$$



Σχήμα 7.1: Τυπικές διαμορφώσεις των σπιν s ένα πλέγμα διαστάσεων 128×128 για μια μεγάλη θερμοκρασία (αριστερά) και για μια θερμοκρασία κοντά στο 0 (δεξιά).

Αντίθετα για $\beta = 0$ (άπειρη θερμοκρασία) όλες οι καταστάσεις έχουν την ίδια πιθανότητα $\pi(x) = \frac{1}{|\mathbb{X}|} = 2^{-L^2}$. Βλέπουμε λοιπόν ότι, καθώς θερμαίνουμε το σύστημα από τη θερμοκρασία $T = 0$ στη θερμοκρασία $T = \infty$, το σύστημα μεταβαίνει από μια κατάσταση πλήρους οργάνωσης (όλα τα σπιν προσανατολισμένα) σε μια κατάσταση πλήρους αταξίας, όπου όλες οι διαμορφώσεις σπιν είναι ισοπίθανες.

Στο Σχήμα 7.4 βλέπουμε δείγματα από την κατανομή π που έχουμε πάρει με τη βοήθεια του αλγορίθμου μας για $\beta = 1/3$ (αριστερά) και για $\beta = 1$ (δεξιά). Τα μαύρα pixel αντιστοιχούν σε σωματίδια με σπιν $+1$ ενώ τα άσπρα σε σωματίδια με σπιν -1 . Επιβεβαιώνουμε αριθμητικά έτσι ότι σε χαμηλές θερμοκρασίες (δεξιά) τα σπιν τείνουν να είναι σε συμφωνία, ενώ σε υψηλές θερμοκρασίες (αριστερά) τα σπιν φαίνονται τυχαία προσανατολισμένα.

Στα αριθμητικά πειράματα θα εξετάσουμε πώς πραγματοποιείται η μετάβαση από τη μία εικόνα στην άλλη καθώς κατεβάζουμε τη θερμοκρασία και θα δούμε ότι αυτή η μετάβαση συμβαίνει απότομα, όταν η θερμοκρασία πέσει κάτω από μια κρίσιμη τιμή. Περισσότερα για τον αλγόριθμο Metropolis-Hastings και για τη μέθοδο MCMC μπορείτε να διαβάσετε στις αναφορές [4] και [5].

7.5 Προσομοιωμένη ανόπτηση (simulated annealing)

Στην προηγούμενη παράγραφο είδαμε ότι καθώς $\beta \rightarrow \infty$ (καθώς ψύχουμε το σύστημα), η κατανομή Gibbs π τείνει να συγκεντρωθεί στις καταστάσεις εκείνες όπου η ενέργεια του συστήματος ελαχιστοποιείται. Στο μοντέλο Ising οι καταστάσεις που ελαχιστοποιούν την ενέργεια είναι προφανείς. Υπάρχουν όμως παραδείγματα όπου το ελάχιστο της συνάρτησης ενέργειας H είναι εξαιρετικά δύσκολο να υπολογιστεί, όπως για παράδειγμα στο πρόβλημα του πλανόδιου πωλητή που πρέπει να επισκεφτεί μια σειρά από πόλεις και θέλει να βρει τη σειρά με την οποία θα τις επισκεφτεί, ώστε να ελαχιστοποιήσει τη συνολική απόσταση που θα διανύσει.

Η μέθοδος της προσομοιωμένης ανόπτησης χρησιμοποιεί αυτή την ιδιότητα των κατανομών Gibbs για

να υπολογίσει το ελάχιστο μιας συνάρτησης. Η ιδέα οφείλει το όνομά της σε μια τεχνική επεξεργασίας μετάλλων με σταδιακή ψύξη τους (ανόπτηση) και είναι πολύ απλή. Χρησιμοποιούμε τον αλγόριθμο Metropolis-Hastings για κάποιο αρχικό $\beta > 0$. Μετά από έναν αριθμό βημάτων ώστε η αλυσίδα μας να έχει προσεγγίσει την κατάσταση ισορροπίας που αντιστοιχεί σε αυτήν τη θερμοκρασία, ανεβάζουμε την τιμή του β (ψύχουμε το σύστημα). Συνήθως ρίχνουμε τη θερμοκρασία κατά έναν πολλαπλασιαστικό παράγοντα ψύξης (cooling factor). Στην καινούργια θερμοκρασία αφήνουμε πάλι την αλυσίδα μας να προσεγγίσει την ισορροπία. Η συνάρτηση $r(\cdot, \cdot)$ στον αλγόριθμο Metropolis-Hastings μπορεί να αλλάζει καθώς αλλάζουμε τη θερμοκρασία. Αυτό συνήθως γίνεται επιτρέποντας όλο και πιο εντοπισμένες μεταβάσεις όσο η θερμοκρασία κατεβαίνει, ώστε ο αλγόριθμος να εντοπίζει το ελάχιστο με μεγαλύτερη ακρίβεια. Η επιλογή του παράγοντα ψύξης και της r ονομάζεται πρόγραμμα ψύξης (cooling schedule) και έχει, όπως θα δούμε στις προγραμματιστικές ασκήσεις, μεγάλη σημασία στην αποτελεσματική λειτουργία του αλγορίθμου. Καθώς επαναλαμβάνουμε τη διαδικασία, ανεβάζοντας σταδιακά την τιμή του β , οι καταστάσεις με χαμηλή ενέργεια γίνονται όλο και πιο πιθανές και στο όριο καθώς $\beta \rightarrow \infty$ η αλυσίδα μας θα βρίσκεται με μεγάλη πιθανότητα σε κάποια από τις καταστάσεις με τη χαμηλότερη ενέργεια.

Πιο συγκεκριμένα, ας υποθέσουμε χωρίς βλάβη της γενικότητας (αφού αν μετατοπίσουμε τη συνάρτηση ενέργειας κατά μία σταθερά, το μέτρο π δεν αλλάζει) ότι το ελάχιστο της συνάρτησης H στο \mathbb{X} είναι το 0. Ορίζουμε το σταθμικό σύνολο της H

$$E_t = \{x \in \mathbb{X} : H(x) \leq t\}.$$

Το σύνολο E_0 αποτελείται από τα σημεία του \mathbb{X} στα οποία η H λαμβάνει την ελάχιστη τιμή της. Έστω $|E_0| = M$ ο πληθάνριθμος του E_0 . Για $\epsilon > 0$ το σύνολο E_ϵ^c είναι εκείνες οι καταστάσεις με ενέργεια τουλάχιστον ϵ . Μπορούμε να εκτιμήσουμε την $\pi[E_\epsilon^c]$ ως εξής:

$$\pi[E_\epsilon^c] = \frac{\sum_{x \in E_\epsilon^c} e^{-\beta H(x)}}{\sum_{x \in E_\epsilon^c} e^{-\beta H(x)} + \sum_{x \in E_\epsilon} e^{-\beta H(x)}} \leq \frac{\sum_{x \in E_\epsilon^c} e^{-\beta H(x)}}{\sum_{x \in E_\epsilon^c} e^{-\beta H(x)} + M} \leq \frac{|\mathbb{X}|}{M} e^{-\beta \epsilon}.$$

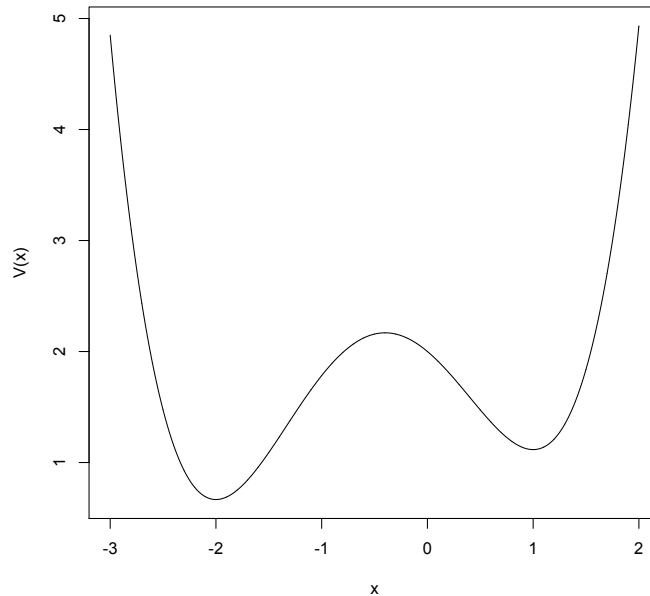
Βλέπουμε λοιπόν ότι, καθώς $\beta \rightarrow \infty$, η $\pi[E_\epsilon^c]$ τείνει στο 0 και μάλιστα εκθετικά γρήγορα. Επομένως, έχοντας αφήσει την αλυσίδα να ισορροπήσει σε μια χαμηλή θερμοκρασία, με πολύ μεγάλη πιθανότητα αυτή θα βρεθεί σε κάποιο σημείο του \mathbb{X} , που δίνει στη συνάρτηση H τιμή κοντά στο ελάχιστό της.

Ο λόγος που ψύχουμε σταδιακά το σύστημα είναι για να μην παγιδευτούμε σε τοπικά ελάχιστα της συνάρτησης H . Αυτός είναι ο κυριότερος λόγος που ο αλγόριθμος της προσομοιωμένης ανόπτησης δουλεύει πολύ καλύτερα από κλασικές μεθόδους κλίσης, στην περίπτωση που τα σταθμικά σύνολα της H , $E_t = \{x \in \mathbb{X} : H(x) \leq t\}$ έχουν πολύπλοκη γεωγραφία.

Παράδειγμα 50 Ας δούμε τώρα τη μέθοδο της προσομοιωμένης ανόπτησης μέσα από ένα εύκολο παράδειγμα που μπορούμε να λύσουμε και αναλυτικά. Η συνάρτηση

$$V(x) = \frac{x^4}{4} + \frac{7x^3}{15} - \frac{4x^2}{5} - \frac{4x}{5} + 2 \quad (7.7)$$

έχει ελάχιστο στο $x = -2$ με $V(-2) = \frac{2}{3}$ και τοπικό ελάχιστο στο $x = +1$ με $V(1) = 67/60$. Στο Σχήμα 7.2 φαίνεται η γραφική παράσταση της συνάρτησης $V(\cdot)$. Ο κώδικας `potential_siman.py` υλοποιεί σε Python τον αλγόριθμο της προσομοιωμένης ανόπτησης. Η σταθερά `gamma` στον κώδικα καθορίζει τον παράγοντα ψύξης, αφού όταν κατεβάζουμε τη θερμοκρασία η νέα θερμοκρασία T είναι η προηγούμενη πολλαπλασιασμένη με `1-gamma`. Το πρόγραμμα επαναλαμβάνει `iter` φορές τον αλγόριθμο και υπολογίζει το ποσοστό p των φορών που επιτυγχάνει να βρει το ολικό ελάχιστο, δηλαδή το -2.



Σχήμα 7.2: Η γραφική παράσταση της συνάρτησης V του Παραδείγματος 50

Στα αριθμητικά πειράματα θα έχετε την ευκαιρία να δείτε ένα κινούμενο σχέδιο σχετικά με το πώς λειτουργεί ο αλγόριθμος της προσομοιωμένης ανόπτησης στο συγκεκριμένο παράδειγμα και πώς επηρεάζεται η αποτελεσματικότητα του αλγορίθμου στο να εντοπίζει το ελάχιστο από το πρόγραμμα ψύξης.

Με αντίστοιχο τρόπο μπορούμε να επιχειρήσουμε να βρούμε το ελάχιστο και στην περίπτωση του προβλήματος του πλανόδιου πωλητή. Με δεδομένες τις θέσεις των πόλεων μπορούμε να πάρουμε ως χώρο των καταστάσεων \mathbb{X} της αλυσίδας μας τις κυκλικές μεταθέσεις των πόλεων που θα παριστάνουν τη σειρά με την οποία ο πωλητής επισκέπτεται τις πόλεις. Σαν ενέργεια μιας τέτοιας κατάστασης μπορούμε να πάρουμε το συνολικό μήκος του μονοπατιού που διανύουμε για επισκεπτούμε τις πόλεις με τη συγκεκριμένη σειρά. Με αυτό το πρόβλημα θα ασχοληθείτε εκτενώς στα αριθμητικά πειράματα. Μπορείτε να βρείτε περισσότερες πληροφορίες για την προσομοιωμένη ανόπτηση στην αναφορά [5].

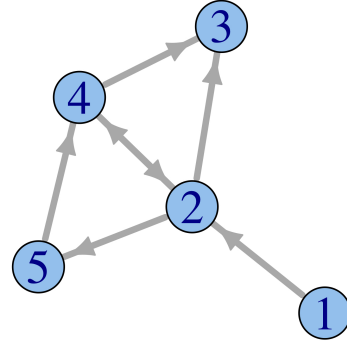
7.6 Ασκήσεις

Άσκηση 111 Στο δίκτυο του διπλανού σχήματος υπολογίστε τον δείκτη σημαντικότητας που κατασκευάσαμε στην παράγραφο 7.2 για κάθε κορυφή του.

Άσκηση 112 Αν διατηρούσατε τη σελίδα 1 του διπλανού σχήματος πώς θα μπορούσατε να αυξήσετε τον δείκτη σημαντικότητάς της ώστε να εμφανίζεται πιο ψηλά σε σχετικές αναζητήσεις;

Άσκηση 113 Κατασκευάστε μια μαρκοβιανή αλυσίδα στον χώρο $\mathbb{X} = \{A, B, C, D, E, F, G\}$, ώστε να έχει κατανομή ισορροπίας

$$\pi = (\pi(A), \pi(B), \pi(C), \pi(D), \pi(E), \pi(F), \pi(G)) = (1/4, 1/8, 1/6, 1/16, 3/16, 1/12, 1/8).$$



Άσκηση 114 Δείξτε ότι αν $B_1 \subset B_2$, τότε $C(A, B_1) \leq C(A, B_2)$.

Άσκηση 115 Στον αλγόριθμο Metropolis-Hastings παίρνουμε τις προτεινόμενες μεταβάσεις $\{r(x, y)\}_{x, y \in \mathbb{X}}$ ανεξάρτητες από την κατάσταση αφετηρίας, δηλαδή

$$r(x, y) = \rho(y), \quad x, y \in \mathbb{X},$$

όπου η $\rho(\cdot)$ είναι μια σ.μ.π. στον X , τέτοια ώστε $\rho(x) > 0$ για κάθε $x \in \mathbb{X}$. Ορίζουμε $w(x) = \frac{\rho(x)}{\pi(x)}$. Με αυτή την επιλογή οι πιθανότητες μετάβασης της αλυσίδας του αλγορίθμου γίνονται

$$p(x, y) = \rho(y) \left(\frac{w(x)}{w(y)} \wedge 1 \right).$$

Αν ταξινομήσουμε τις καταστάσεις του $\mathbb{X} = \{x_1, \dots, x_N\}$, ώστε

$$w(x_1) \leq w(x_2) \leq \dots \leq w(x_N),$$

δείξτε ότι οι ιδιοτιμές και τα αντίστοιχα ιδιοδιανύσματα του πίνακα πιθανοτήτων μετάβασης $P = \{p(x, y)\}_{x, y \in \mathbb{X}}$ είναι $\lambda_1 = 1$, $v_1 = (1, 1, \dots, 1)^T$ και για $k \in \{1, 2, \dots, N-1\}$

$$\lambda_{k+1} = \sum_{j=k+1}^N \pi(x_j) (w(x_j) - w(x_k)), \quad v_{k+1} = (0, \dots, 0, \sum_{j=k+1}^N \pi(x_j), -\pi(x_k), \dots, -\pi(x_k))^T.$$

Δείξτε ότι $1 = \lambda_1 > \lambda_2 = 1 - w(x_1) > \lambda_3 > \dots > \lambda_N$.

7.7 Αριθμητικά πειράματα

Άσκηση 116 Ο αλγόριθμος Metropolis-Hastings μας επιτρέπει να επιλέξουμε τυχαία ένα σημείο από κάποιο χωρίο $D \subset \mathbb{R}^n$, αν πάρουμε ως συνάρτηση ενέργειας την

$$H(x) = \begin{cases} 0, & \text{αν } x \in D \\ +\infty, & \text{αν } x \notin D \end{cases}.$$

Ο κώδικας `disc_sampler.py` χρησιμοποιεί τον αλγόριθμο Metropolis-Hastings για να πάρει 1.000 δείγματα από μια τυχαία μεταβλητή με ομοιόμορφη κατανομή στον μοναδιαίο δίσκο

$$D_2 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}.$$

Η αλυσίδα ξεκινά από το μηδέν. Η προτεινόμενη μετατόπιση είναι ομοιόμορφη στο $(-1,1)$ παράλληλη προς έναν από τους άξονες. Αφήνουμε την αλυσίδα να κάνει $N = 100$ βήματα πριν πάρουμε ένα δείγμα.

α) Αλλάζετε τον κώδικα ώστε για διάφορες τιμές του δ να υπολογίζει ποιο ποσοστό από τα προτεινόμενα βήματα πραγματοποιήθηκαν και να κατασκευάζει το γράφημα των δειγμάτων που πήρατε από τον δίσκο. Μπορείτε να εξηγήσετε γιατί ο αλγόριθμος δεν δίνει καλά αποτελέσματα αν επιλέξουμε το δ πολύ κοντά στο μηδέν ή πολύ μεγάλο;

β) Αλλάζετε τον κώδικα ώστε να δίνει 1.000 τυχαία δείγματα από τη μοναδιαία σφαίρα D_d στις d διαστάσεις.

γ) Προκειμένου να πάρουμε ένα τυχαίο δείγμα από τον κύλινδρο

$$S_{d+1} = \{(x_1, \dots, x_{d+1}) \in \mathbb{R}^{d+1} : (x_1, \dots, x_d) \in D_d, |x_{d+1}| < 1\},$$

αρκεί να επιλέξουμε το (x_1, \dots, x_d) ομοιόμορφα από τη D_d και τη x_{d+1} ομοιόμορφα από το $(-1,1)$. Παρατηρήστε ότι $|S_{d+1}| = 2 \times |D_d|$ (επιφάνεια βάσης επί ύψος) και ξανακάνετε την Άσκηση 93, χρησιμοποιώντας αυτή τη φορά δείγματα από τον κύλινδρο S_{d+1} και μετρώντας πόσα από αυτά έπεσαν στη D_{d+1} . Είναι τώρα ικανοποιητική η εκτίμησή σας ακόμα και σε αρκετά μεγάλες διαστάσεις;

Άσκηση 117 Στην άσκηση αυτή θα χρησιμοποιήσουμε τον αλγόριθμο Metropolis-Hastings για να κάνουμε δειγματοληψία από την αναλλοίωτη κατανομή π του μοντέλου Ising. Καταβάστε και ανοίξτε τον κώδικα `ising.py`. Στον κώδικά μας θεωρούμε $L = 64$, οπότε ο χώρος των δυνατών διαμορφώσεων έχει πληθώρα αριθμό $2^{4096} > 10^{1200}$. Καταλαβαίνει κανείς ότι η δειγματοληψία από έναν τόσο μεγάλο χώρο είναι πρακτικά αδύνατη με συμβατικές μεθόδους. Ο κώδικας χρησιμοποιεί τον αλγόριθμο Metropolis-Hastings για να προσομοιώσει μια μαρκοβιανή αλυσίδα στον \mathbb{X} που ξεκινά από μια τυχαία αρχική διαμόρφωση και έχει κατανομή ισορροπίας την π .

α) Τρέξτε τον κώδικα για τη δοσμένη θερμοκρασία $T = 30$ και δείτε τη διαμόρφωση στην οποία καταλήγει ο αλγόριθμος. Επαναλάβετε μερικές φορές.

β) Αλλάζετε τον κώδικα ώστε η αρχική διαμόρφωση των σπιν να είναι $+1$ παντού. Ξανατρέξτε τον κώδικα μερικές φορές. Αλλάζουν τα ποιοτικά χαρακτηριστικά της διαμόρφωσης που επιστρέφει ο αλγόριθμος;

γ) Αρχίστε τώρα να κατεβάζετε τη θερμοκρασία, θέτοντας διαδοχικά $T = 10.0, 3.0, 1.0, 0.3, 0.1, 0.03$ και περιγράψτε τι παρατηρείτε.

δ) Εξηγήστε γιατί η γραμμή του κώδικα

```
if random.uniform(0.0,1.0) < math.exp(-deltaE/Temperature): Spin[k] *=-1
```

υλοποιεί τον αλγόριθμο Metropolis-Hastings.

ε) Η μαγνήτιση του συστήματος ορίζεται ως ο μέσος όρος των σπιν του πλέγματος Λ , δηλαδή

$$m(\sigma) = \frac{1}{|\Lambda|} \sum_{k \in \Lambda} \sigma(k).$$

Εξετάστε πώς συμπεριφέρεται η $m(\sigma)$ για τις καταστάσεις σ που δίνει ο αλγόριθμος καθώς κατεβάζετε τη θερμοκρασία, όπως στο ερώτημα (β).

ε) Για τη θερμοκρασία $T = 0.03$ αλλάζετε την τιμή της παραμέτρου `nsteps`. Πώς μοιάζουν τα ενδιάμεσα στάδια από τα οποία περνάει η αλυσίδα μέχρι να καταλήξουμε στην εικόνα που πήραμε για

`nsteps = 100 × L × L;`

στ) Η θεωρία προβλέπει ότι υπάρχει μια απότομη μεταβολή φάσης από την παραμαγνητική κατάσταση (υψηλή θερμοκρασία) στη σιδηρομαγνητική (χαμηλή θερμοκρασία) όταν $\beta \simeq 0.44$. Μπορείτε να επιβεβαιώσετε αριθμητικά αυτό το γεγονός;

Άσκηση 118 Στην άσκηση αυτή θα χρησιμοποιήσετε τη μέθοδο της προσομοιωμένης ανόπτωσης για να βρείτε το ελάχιστο μιας συνάρτησης $V : \mathbb{R} \rightarrow \mathbb{R}$. Ο αλγόριθμος έχει ως εξής. Ένα σωματίδιο είναι αρχικά τοποθετημένο τυχαία στο διάστημα $(-3.5, +3.5)$ σε θερμοκρασία $T = 128$. Σε κάθε βήμα επιχειρεί μια μετατόπιση που ακολουθεί ομοιόμορφη κατανομή στο διάστημα $(-\delta, +\delta)$ και πραγματοποιεί τη μετατόπιση σύμφωνα με τον αλγόριθμο Metropolis-Hastings. Κάθε 100 βήματα ρίχνουμε τη θερμοκρασία και επαναλαμβάνουμε μέχρι η θερμοκρασία να πέσει κάτω από το $T = 0.01$.

Κατεβάστε τον κώδικα `sim_annealing.py` και μελετήστε τον. Η παράμετρος `minima` σας επιτρέπει να επιλέξετε αν θα ελαχιστοποιήσετε μια συνάρτηση με 1, 2 ή 3 τοπικά ακρότατα. Η αρχική επιλογή είναι 3. Η παράμετρος `gamma` σας επιτρέπει να ελέγξετε τον ρυθμό ψύξης. Η αρχική επιλογή είναι να ρίχνουμε τη θερμοκρασία στο μισό κάθε 100 βήματα. Η παράμετρος `iterations` σας επιτρέπει να επαναλάβετε τη διαδικασία εύρεσης του ελαχίστου όσες φορές θέλετε. Η αρχική επιλογή της είναι 1. Η παράμετρος `delta` ελέγχει το μέγεθος της επιχειρούμενης μετατόπισης σε κάθε βήμα. Η αρχική της επιλογή είναι 3. Τέλος η λογική παράμετρος `animation` ελέγχει αν θέλουμε να δούμε μια οπτική αναπαράσταση της προσομοιωμένης ανόπτωσης. Η αρχική της επιλογή είναι `True`.

α) Τρέξτε τον κώδικα. Παρατηρήστε πώς σε μεγάλες θερμοκρασίες το σωματίδιο κινείται σχετικά ελεύθερα στον χώρο και πώς καθώς ψύχουμε το σύστημα το σωματίδιο εντοπίζεται κοντά στο ολικό ελάχιστο της V . Αλλάζτε την παράμετρο `minima` και επαναλάβετε.

β) Αλλάζτε την παράμετρο `iterations` σε 1000 και την παράμετρο `animation` σε `False` ώστε να επιταχύνετε την εκτέλεση του προγράμματος. Για τις διάφορες τιμές της παραμέτρου `minima` βρείτε σε τι ποσοστό επιτυγχάνει ο αλγόριθμος να εντοπίσει το ελάχιστο της V .

γ) Αλλάζτε την τιμή της παραμέτρου `delta` σε 0.1 και επαναλάβετε το προηγούμενο ερώτημα. Ποιο είναι τώρα το ποσοστό επιτυχούς εύρεσης του ολικού ελαχίστου; Γιατί πιστεύετε ότι συμβαίνει αυτό; Αν θέλετε να έχετε μια εικόνα του τι συμβαίνει, αλλάζτε την παράμετρο `iterations` σε 1 και την παράμετρο `animation` σε `True` για να δείτε πώς κινείται το σωματίδιο και τρέξτε τον κώδικα μέχρι να δείτε το σωματίδιο να καταλήγει κοντά σε ένα τοπικό ελάχιστο, διαφορετικό από το ολικό.

δ) Αλλάζτε τον κώδικα ώστε να βρίσκει το ελάχιστο της συνάρτησης V της (7.7).

ε) Μεταβάλετε το `gamma` και δείτε πώς συμπεριφέρεται το ποσοστό p των επιτυχημένων προσπαθειών ως συνάρτηση του `gamma`.

- Επαληθεύστε ότι για σχετικά μεγάλες τιμές του `gamma` (π.χ. $> 1/2$) ο ο αλγόριθμος δεν βρίσκει πάντα το ολικό ελάχιστο.

- Πού τερματίζει ο αλγόριθμος στις περιπτώσεις που δεν βρίσκει το ολικό ελάχιστο; Τι παρατηρείτε;

- Παραστήστε γραφικά την πιθανότητα p ως συνάρτηση του `gamma`.

- Επαληθεύστε ότι για πολύ μικρές τιμές του `gamma` έχουμε p πολύ κοντά στο 1. Δηλαδή για κατάλληλα αργό ρυθμό ανόπτωσης ο αλγόριθμος εντοπίζει το ολικό ελάχιστο με πιθανότητα κοντά στο 1.

Άσκηση 119 Στην άσκηση αυτή θα χρησιμοποιήσουμε τον αλγόριθμο της προσομοιωμένης ανόπτωσης για να λύσουμε το πρόβλημα του πλανόδιου πωλητή. Κατεβάστε το αρχείο `eu.csv` που περιέχει το γεωγραφικό μήκος και πλάτος 27 ευρωπαϊκών πόλεων. Κατεβάστε τους κώδικες `siman_tsp.py` και `plot_lib.py`. Ο πρώτος περιλαμβάνει τη βασική ρουτίνα που θα χρησιμοποιήσουμε, ενώ ο δεύτερος εισάγεται ως βιβλιοθήκη και χρησιμοποιείται για τα διαγράμματα που θα κατασκευάσουμε. Ο κώδικας `siman_tsp.py` διαβάζει από το αρχείο `eu.csv` τις πόλεις και τα γεωγραφικά μήκη και πλάτη τους και χρησιμοποιεί τη ρουτίνα `geodesic.distance` για να υπολογίσει τη γεωδαισιακή τους απόσταση (την απόστασή τους πάνω στη

γήνη σφαίρα). Στη συνέχεια χρησιμοποιεί τον αλγόριθμο της προσομοιωμένης ανόπτωσης για να υπολογίσει μια προσεγγιστική λύση για το πρόβλημα του πλανόδιου πωλητή, δηλαδή με ποια σειρά πρέπει να επισκεφτεί κανείς αυτές τις πόλεις ώστε να ελαχιστοποιήσει τη συνολική απόσταση που θα διανύσει.

Στην έξοδό του ο κώδικας επιστρέφει τρία διαγράμματα. Το πρώτο έχει ένα μονοπάτι όπου οι επισκέψεις στις πόλεις γίνονται με τυχαία σειρά. Το δεύτερο έχει το βέλτιστο μονοπάτι που υπολογίζει ο αλγόριθμος της προσομοιωμένης ανόπτωσης, ενώ το τρίτο δείχνει πώς εξελίσσεται η συνολική απόσταση για το τρέχον μονοπάτι του αλγορίθμου (μπλε γραμμή) και για το βέλτιστο μονοπάτι που έχει βρεθεί μέχρι τότε (κόκκινη γραμμή), καθώς κατεβάζουμε τη θερμοκρασία.

α) Τρέξτε τον κώδικα `siman_tsp.py` και δείτε το μονοπάτι που επιστρέφει ο αλγόριθμος της προσομοιωμένης ανόπτωσης και την εξέλιξη της συνολικής απόστασης καθώς ψύχουμε το σύστημα. Τυπώστε αυτά τα διαγράμματα και εξηγήστε την εικόνα που βλέπετε στο δεύτερο.

β) Εντοπίστε στον κώδικα τις εντολές που υλοποιούν τον αλγόριθμο Metropolis-Hastings και εξηγήστε πώς ακριβώς αλλάζει το μονοπάτι σε κάθε βήμα του αλγορίθμου.

γ) Αλλάξτε τον κώδικα ώστε να διαβάζει από το αρχείο `europa.csv` τις συντεταγμένες 35 ευρωπαϊκών πόλεων και λύστε το πρόβλημα του πλανόδιου πωλητή που θέλει να επισκεφτεί και τις 35 αυτές πόλεις. Στο διάγραμμα δεν τυπώνονται τα ονόματα των πόλεων για να είναι πιο ευκρινές το μονοπάτι.

δ) Πόσα είναι τα δυνατά μονοπάτια που περνούν από όλες τις 35 πόλεις μία μόνο φορά; Αν σε έναν υπολογιστή ο χρόνος που απαιτείται για τον υπολογισμό του συνολικού μήκους ενός τέτοιου μονοπατιού ήταν 1μsec, πόσο χρόνο θα χρειαζόμασταν για να βρούμε το βέλτιστο μονοπάτι με έναν εξαντλητικό αλγόριθμο που θα υπολόγιζε τη συνολική απόσταση που θα διανύσει ο πλανόδιος πωλητής κατά μήκος κάθε μονοπατιού και θα επέστρεφε το μονοπάτι με την ελάχιστη συνολική απόσταση;

Κεφάλαιο 8

Διαδικασίες Poisson

8.1 Εισαγωγή

Σ' αυτό το κεφάλαιο θα ορίσουμε τις διαδικασίες Poisson και θα μελετήσουμε τις βασικές τους ιδιότητες. Οι διαδικασίες αυτές είναι ίσως οι απλούστερες μη τετριμμένες διαδικασίες συνεχούς χρόνου. Οι τροχιές τους αυξάνουν στους ακέραιους και οι χρόνοι που μεσολαβούν ανάμεσα σε διαδοχικές αυξήσεις είναι ανεξάρτητες εκθετικές τυχαίες μεταβλητές. Οι διαδικασίες αυτές είναι ένα καλό μοντέλο για διακριτές αφίξεις τυχαίων συμβάντων σε συνεχή χρόνο, όπως π.χ. οι κλήσεις που δέχεται το τηλεφωνικό κέντρο του ΕΚΑΒ, οι διασπάσεις των πυρήνων ενός ραδιενεργού υλικού, οι αφίξεις φωτονίων σε μια φωτοδίοδο, τα αιτήματα που δέχεται ένας εξυπηρετητής και πολλά άλλα. Υπάρχουν πολλά καλά βιβλία στα οποία μπορεί να βρει κανείς μια εισαγωγή στις διαδικασίες Poisson, μερικά απ' αυτά είναι το [1] και το [7]

8.2 Ορισμός και ιδιότητες

Συνήθως σ' αυτό το βιβλίο περιγράφουμε εξ αρχής μια στοχαστική διαδικασία μέσω των κατανομών πεπερασμένης διάστασης. Σ' αυτό το κεφάλαιο θα κάνουμε μια εξαίρεση. Θα δώσουμε πρώτα έναν κατασκευαστικό, διαισθητικό ορισμό για τις διαδικασίες Poisson και στη συνέχεια θα υπολογίσουμε τις κατανομές πεπερασμένης διάστασης.

Φανταστείτε ότι επαναλαμβάνετε ανεξάρτητες δοκιμές Bernoulli με πιθανότητα επιτυχίας p και σημειώνετε τις επιτυχείς δοκιμές. Ανάμεσά τους μεσολαβεί ένας τυχαίος αριθμός προσπαθειών, που ακολουθεί γεωμετρική κατανομή με παράμετρο p . Αν σε μια μονάδα φυσικού χρόνου προλαβαίνετε να κάνετε ν δοκιμές, τότε ο ρυθμός με τον οποίο έρχονται επιτυχίες είναι νp . Ας κάνουμε ξανά το πείραμα, αυτή τη φορά όμως ας κάνουμε τις δοκιμές με διπλάσια ταχύτητα αλλά με τη μισή πιθανότητα επιτυχίας. Αυτή τη φορά οι προσπάθειες που μεσολαβούν ανάμεσα σε διαδοχικές επιτυχίες είναι στατιστικά περισσότερες, αφού ακολουθούν γεωμετρική κατανομή με παράμετρο $p/2$. Επειδή όμως στη μονάδα του χρόνου κάνουμε διπλάσιες προσπάθειες, ο ρυθμός με τον οποίο σημειώνονται επιτυχίες στον πραγματικό χρόνο θα παραμείνει νp . Μπορούμε να συνεχίσουμε να πολλαπλασιάζουμε την ταχύτητα με την οποία εκτελούμε τις δοκιμές και να μειώνουμε αντίστοιχα την πιθανότητα επιτυχίας μας. Αν κάνουμε $N\nu$ δοκιμές ανά μονάδα χρόνου με πιθανότητα επιτυχίας σε κάθε δοκιμή p/N , τότε καθώς το N μεγαλώνει, θα μεσολαβούν όλο και περισσότερες προσπάθειες ανάμεσα σε διαδοχικές επιτυχίες. Επειδή όμως και το πλήθος των προσπαθειών που προλαβαίνουμε να κάνουμε ανά μονάδα χρόνου μεγαλώνει ανάλογα, ο ρυθμός με τον οποίον σημειώνονται επιτυχίες σε πραγματικό χρόνο δεν θα αλλάξει. Στο όριο καθώς το $N \rightarrow \infty$, ο πραγματικός χρόνος ανάμεσα με διαδοχικές επιτυχίες ακολουθεί εκθετική κατανομή και η διαδικασία που μετρά τις επιτυχίες μας στον πραγματικό χρόνο είναι μια διαδικασία Poisson με ρυθμό νp .

Ας θεωρήσουμε έναν χώρο πιθανότητας Ω , στον οποίο είναι ορισμένη μια ακολουθία $\{E_j\}_{j \in \mathbb{N}}$ από ανεξάρτητες, εκθετικές, τυχαίες μεταβλητές, με παράμετρο λ . Επομένως,

$$\mathbb{P}[E_j > t] = e^{-\lambda t}, \quad \text{για κάθε } t \geq 0.$$

Μια καθοριστική ιδιότητα που έχουν οι εκθετικές τυχαίες μεταβλητές είναι η *ιδιότητα της απώλειας μνήμης* (memoryless property). Για κάθε $s, t \geq 0$: $\mathbb{P}[E_j > t + s \mid E_j > t] = \mathbb{P}[E_j > s]$. Ορίζουμε ακόμα

$$S_0 = 0 \quad \text{και} \quad S_n = S_{n-1} + E_n, \text{ για κάθε } n \in \mathbb{N}.$$

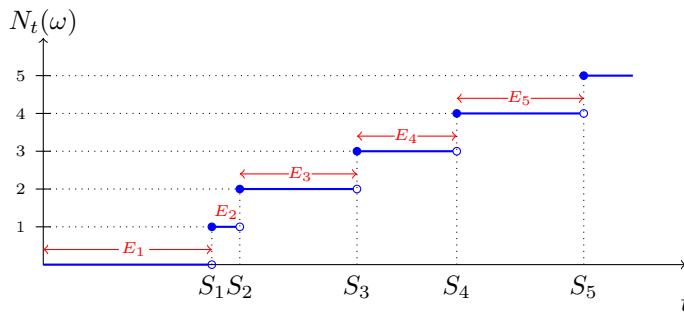
Οι μεταβλητές $\{E_j\}_{j \in \mathbb{N}}$ μοντελοποιούν τον χρόνο ανάμεσα σε διαδοχικές αφίξεις, οπότε οι μεταβλητές $\{S_j\}_{j \in \mathbb{N}}$ είναι οι χρόνοι στους οποίους αυτές οι αφίξεις συμβαίνουν. Η διαδικασία Poisson $\{N_t\}_{t \geq 0}$ με ρυθμό λ μετρά πόσες τέτοιες αφίξεις έχουν συμβεί μέχρι την εκάστοτε χρονική στιγμή, δηλαδή

$$N_t(\omega) = \sup\{k \in \mathbb{N}_0 : S_k(\omega) \leq t\}, \quad t \geq 0, \omega \in \Omega. \quad (8.1)$$

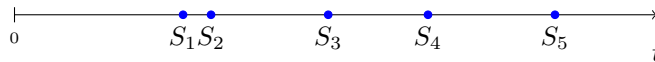
Μπορούμε επίσης να γράψουμε ισοδύναμα ότι

$$N_t(\omega) = \sum_{k \in \mathbb{N}} \mathbb{1}\{S_k(\omega) \leq t\}, \quad t \geq 0, \omega \in \Omega. \quad (8.2)$$

Στο παρακάτω σχήμα φαίνεται η αρχή μιας τυπικής τροχιάς διαδικασίας Poisson.



Εφόσον τα άλματα μιας διαδικασίας Poisson έχουν μέγεθος 1, μπορούμε να αναπαραστήσουμε μια τροχιά της δίνοντας απλά τα χρονικά σημεία $\{S_n\}_{n \in \mathbb{N}}$ στα οποία η διαδικασία αυξάνει. Για παράδειγμα, μπορούμε να αναπαραστήσουμε την προηγούμενη τυπική τροχιά και ως εξής.



Θεώρημα 33 Η διαδικασία Poisson της (8.1) έχει τις ακόλουθες ιδιότητες.

1. Οι τροχιές της είναι αύξουσες, δηλαδή $0 \leq t_1 \leq t_2 \Rightarrow N_{t_1}(\omega) \leq N_{t_2}(\omega)$.
2. $\mathbb{P}[N_t \in \mathbb{X} = \mathbb{N}_0, \text{ για κάθε } t \geq 0] = 1$.
3. Οι τροχιές της είναι δεξιά συνεχείς.
4. Για κάθε $m \in \mathbb{N}$ η διαδικασία $\{\tilde{N}_t\}_{t \geq 0}$ με $\tilde{N}_t = N_{S_m+t} - m$ είναι κι αυτή μια διαδικασία Poisson με ρυθμό λ και είναι ανεξάρτητη από τις E_1, \dots, E_m .

Απόδειξη: Από τον ορισμό της διαδικασίας Poisson είναι προφανές ότι οι τροχιές της είναι αύξουσες, αφού, αν $0 \leq t_1 \leq t_2$, τότε

$$\{k \in \mathbb{N}_0 : S_k(\omega) \leq t_1\} \subset \{k \in \mathbb{N}_0 : S_k(\omega) \leq t_2\}.$$

Για την (2) αρκεί να δείξουμε ότι

$$\mathbb{P}[N_t < +\infty, \text{ για κάθε } t \geq 0] = \mathbb{P}\left[\bigcap_{t \geq 0} \{N_t < +\infty\}\right] = 1. \quad (8.3)$$

Έχουμε όμως ότι

$$\{S_k \rightarrow \infty\} \subset \bigcap_{t \geq 0} \{N_t < +\infty\}.$$

Πράγματι, αν $\omega \in \{S_k \rightarrow \infty\}$, για κάθε $t \geq 0$ υπάρχει ένα $k_0(\omega, t)$ τέτοιο ώστε $S_k(\omega) > t$ για όλα τα $k > k_0(\omega, t)$. Επομένως, $N_t(\omega) \leq k_0(\omega, t)$ και ειδικότερα $N_t(\omega) < +\infty$. Από τον νόμο των μεγάλων αριθμών έχουμε ότι

$$\mathbb{P}[S_n/n \rightarrow 1/\lambda] = \mathbb{P}\left[\frac{1}{n} \sum_{j=1}^n E_j \rightarrow 1/\lambda\right] = 1,$$

επομένως $\mathbb{P}[S_k \rightarrow \infty] = 1$, απ' όπου προκύπτει η (8.3).

Για την (3) προσέξτε ότι, αν για κάποιο $\omega \in \Omega$, $t \geq 0$ είχαμε ότι $N_{t+\frac{1}{n}}(\omega) \neq N_t(\omega)$, τότε θα έπρεπε να υπάρχει κάποιο $k \in \mathbb{N}_0$ τέτοιο ώστε

$$N_t(\omega) = k \text{ και } N_{t+\frac{1}{n}}(\omega) \geq k+1, \text{ για κάθε } n \in \mathbb{N}.$$

Όμως $N_t(\omega) = k \Rightarrow S_{k+1}(\omega) > t$, ενώ $N_{t+\frac{1}{n}}(\omega) \geq k+1$ για κάθε $n \in \mathbb{N} \Rightarrow S_{k+1}(\omega) \leq t + \frac{1}{n}$ για κάθε $n \in \mathbb{N}$. Θα έπρεπε επομένως

$$S_{k+1}(\omega) \in \bigcap_{n \in \mathbb{N}} \left(t, t + \frac{1}{n}\right] = \emptyset.$$

Θα δείξουμε τέλος την (4), η οποία συνεπάγεται ότι μια διαδικασία Poisson ξαναγεννιέται μετά από κάθε άλμα της και το μέλλον της είναι ανεξάρτητο του παρελθόντος της. Η ιδέα είναι ότι μετά το m -οστό άλμα της, η διαδικασία μετρά τις αφίξεις που υπαγορεύονται από τις E_{m+1}, E_{m+2}, \dots . Συγκεκριμένα έχουμε ότι

$$\begin{aligned} N_{S_m+t} &= \sup\{k \in \mathbb{N} : S_k \leq S_m + t\} = \sup\{k \in \mathbb{N} : \sum_{j=m+1}^k E_j \leq t\} \\ &= \sup\{k \in \mathbb{N} : \sum_{j=1}^{k-m} E_{j+m} \leq t\} = m + \sup\{r \in \mathbb{N} : \sum_{j=1}^r E_{j+m} \leq t\}. \end{aligned}$$

Επομένως,

$$\tilde{N}_t = \sup\{r \in \mathbb{N} : \sum_{j=1}^r E_{j+m} \leq t\}.$$

Όμως οι E_{m+1}, E_{m+2}, \dots είναι ανεξάρτητες, ισόνομες, εκθετικές, τυχαίες μεταβλητές με παράμετρο λ και είναι ανεξάρτητες από τις E_1, \dots, E_m , οπότε το ζητούμενο προκύπτει από τον ορισμό της διαδικασίας Poisson. \square

Θα υπολογίσουμε τώρα τις κατανομές πεπερασμένης διάστασης της $\{N_t\}_{t \geq 0}$. Θα ξεκινήσουμε από τις μονοδιάστατες κατανομές.

Λήμμα 10 Για κάθε $t \geq 0$ η τυχαία μεταβλητή N_t ακολουθεί κατανομή Poisson με ρυθμό λt .

Απόδειξη: Για $k = 0$ έχουμε $\mathbb{P}[N_t = 0] = \mathbb{P}[E_1 > t] = e^{-\lambda t}$. Για $k \in \mathbb{N}$ έχουμε

$$\mathbb{P}[N_t = k] = \mathbb{P}[S_k \leq t, S_{k+1} > t] = \mathbb{P}[S_k \leq t] - \mathbb{P}[S_k \leq t, S_{k+1} \leq t] = \mathbb{P}[S_k \leq t] - \mathbb{P}[S_{k+1} \leq t]. \quad (8.4)$$

Εφόσον όμως οι $\{E_j\}_{j \in \mathbb{N}}$ είναι ανεξάρτητες, με $E_j \sim \exp(\lambda) = G(\lambda, 1)$, έχουμε ότι

$$S_k = \sum_{j=1}^k E_j \sim G(\lambda, k)$$

και η συνάρτηση πυκνότητας πιθανότητάς της είναι η

$$f_k(x) = \begin{cases} \frac{\lambda^k}{(k-1)!} x^{k-1} e^{-\lambda x}, & \text{αν } x \geq 0 \\ 0, & \text{αν } x < 0. \end{cases}$$

Επομένως, η (8.4) γίνεται

$$\mathbb{P}[N_t = k] = \int_0^t \frac{\lambda^k}{(k-1)!} x^{k-1} e^{-\lambda x} dx - \int_0^t \frac{\lambda^{k+1}}{k!} x^k e^{-\lambda x} dx = \left[\frac{\lambda^k}{k!} x^k e^{-\lambda x} \right]_0^t = e^{-\lambda t} \frac{(\lambda t)^k}{k!}.$$

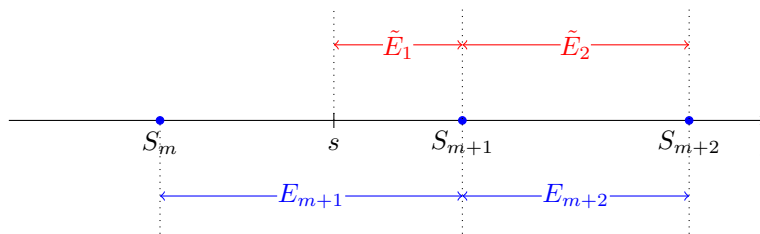
Συμπεραίνουμε λοιπόν ότι $N_t \sim Po(\lambda t)$. □

Θα δούμε στη συνέχεια ότι η διαδικασία Poisson έχει ανεξάρτητες προσαυξήσεις. Η απόδειξη θα γίνει σε δύο βήματα.

Λήμμα 11 Έστω $\{N_t\}_{t \geq 0}$ μια διαδικασία Poisson με ρυθμό λ . Για κάθε $s \geq 0$ και κάθε $m \in \mathbb{N}_0$, με δεδομένο ότι $N_s = m$, η διαδικασία $\{\tilde{N}_t\}_{t \geq 0}$ με τύπο $\tilde{N}_t = N_{s+t} - N_s$ είναι ομοίως μια διαδικασία Poisson με ρυθμό λ και είναι ανεξάρτητη από τις τυχαίες μεταβλητές E_1, \dots, E_m .

Απόδειξη: Παρατηρήστε ότι η διαδικασία $\{\tilde{N}_t\}_{t \geq 0}$ μετρά τις αφίξεις μετά τη χρονική στιγμή s . Με δεδομένο ότι $N_s = m$, οι χρόνοι ανάμεσα στα διαδοχικά άλματα της $\{\tilde{N}_t\}_{t \geq 0}$ είναι

$$\tilde{E}_1 = S_{m+1} - s, \quad \tilde{E}_2 = E_{m+2}, \dots, \tilde{E}_k = E_{m+k}, \dots$$



Το ενδεχόμενο ως προς το οποίο δεσμεύουμε $\{N_s = m\} = \{S_m \leq s, S_{m+1} > s\}$ εξαρτάται μόνο από τις E_1, \dots, E_{m+1} . Επομένως, με δεδομένο το $\{N_s = m\}$, οι $\tilde{E}_2, \tilde{E}_3, \dots$ συνεχίζουν να είναι ανεξάρτητες, ισόνομες, εκθετικές τυχαίες μεταβλητές, με παράμετρο λ και να είναι ανεξάρτητες από τις E_1, \dots, E_m αλλά και την \tilde{E}_1 που εξαρτάται μόνο από τις E_1, \dots, E_{m+1} . Μένει να δείξουμε πως, δεδομένου ότι $N_s = m$, η \tilde{E}_1 είναι ανεξάρτητη από τις E_1, \dots, E_m και ακολουθεί εκθετική κατανομή με παράμετρο λ . Θεωρήστε ένα ενδεχόμενο $B = \{(E_1, \dots, E_m) \in F \subset \mathbb{R}^m\}$ που εξαρτάται μόνο από τις E_1, \dots, E_m . Χρειάζεται να δείξουμε ότι

$$\mathbb{P}[\{\tilde{E}_1 > t\} \cap B \mid N_s = m] = e^{-\lambda t} \mathbb{P}[B \mid N_s = m], \quad \text{για κάθε } t \geq 0.$$

Παρατηρήστε τώρα ότι $\tilde{E}_1 = S_{m+1} - s = S_m + E_{m+1} - s$ και $\{N_s = m\} = \{S_m \leq s, S_m + E_{m+1} > s\}$. Επομένως για κάθε $t \geq 0$ έχουμε

$$\begin{aligned} \mathbb{P}[\{\tilde{E}_1 > t\} \cap B \mid N_s = m] &= \mathbb{P}[\{E_{m+1} + S_m > t + s\} \cap B \mid S_m \leq s, E_{m+1} + S_m > s] \\ &= \frac{\mathbb{P}[\{E_{m+1} + S_m > t + s, S_m \leq s\} \cap B]}{\mathbb{P}[N_s = m]}. \end{aligned} \quad (8.5)$$

Ας συμβολίσουμε με $g : \mathbb{R}^m \rightarrow [0, \infty)$ την από κοινού πυκνότητα πιθανότητας των E_1, \dots, E_m . Θα χρησιμοποιήσουμε τώρα ότι η E_{m+1} είναι ανεξάρτητη από τις E_1, \dots, E_m για να γράψουμε

$$\begin{aligned} \mathbb{P}[\{E_{m+1} + S_m > t + s, S_m \leq s\} \cap B] &= \int_{F \cap \{\sum_{i=1}^m t_i \leq s\}} g(t_1, \dots, t_m) \mathbb{P}[E_{m+1} > t + s - \sum_{i=1}^m t_i] dt_1 \cdots dt_m \\ &= e^{-\lambda t} \int_{F \cap \{\sum_{i=1}^m t_i \leq s\}} g(t_1, \dots, t_m) \mathbb{P}[E_{m+1} > s - \sum_{i=1}^m t_i] dt_1 \cdots dt_m \\ &= e^{-\lambda t} \mathbb{P}[\{E_{m+1} + S_m > s, S_m \leq s\} \cap B] \\ &= e^{-\lambda t} \mathbb{P}[B \cap \{N_s = m\}], \end{aligned}$$

όπου η δεύτερη ισότητα προέκυψε από την απώλεια μνήμης της εκθετικής κατανομής. Το ζητούμενο έπεται τώρα άμεσα από την (8.5). \square

Θεώρημα 34 (μαρκοβιανή ιδιότητα) Έστω $\{N_t\}_{t \geq 0}$ μια διαδικασία Poisson με ρυθμό λ . Για κάθε $s \geq 0$ η διαδικασία $\{\tilde{N}_t\}_{t \geq 0}$ με τύπο $\tilde{N}_t = N_{s+t} - N_s$ είναι ομοίως μια διαδικασία Poisson με ρυθμό λ και είναι ανεξάρτητη από τις $\{N_r\}_{0 \leq r \leq s}$.

Απόδειξη: Χρειάζεται να δείξουμε ότι αν \tilde{A} είναι ένα ενδεχόμενο της μορφής $\{\tilde{N}_{t_1} = m_1, \dots, \tilde{N}_{t_k} = m_k\}$ όπου $k \in \mathbb{N}_0$, $t_1, \dots, t_k \geq 0$ και $m_1, \dots, m_k \in \mathbb{N}_0$ και B είναι ένα ενδεχόμενο της μορφής $\{N_{r_1} = n_1, \dots, N_{r_j} = n_j\}$ όπου $j \in \mathbb{N}_0$, $r_1, \dots, r_j \in [0, s]$ και $n_1, \dots, n_j \in \mathbb{N}_0$, τότε $\mathbb{P}[\tilde{A} \cap B] = \mathbb{P}[\tilde{A}] \mathbb{P}[B]$, όπου $A = \{N_{t_1} = m_1, \dots, N_{t_k} = m_k\}$. Αυτό θα μας εξασφαλίσει άμεσα ότι η $\{\tilde{N}_t\}_{t \geq 0}$ είναι ανεξάρτητη από τις $\{N_r\}_{0 \leq r \leq s}$. Επιπλέον, παίρνοντας $j = 0 \Leftrightarrow B = \Omega$ θα μας εξασφαλίσει επίσης ότι

$$\mathbb{P}[\tilde{A}] = \mathbb{P}[A] \Leftrightarrow \mathbb{P}[\tilde{N}_{t_1} = m_1, \dots, \tilde{N}_{t_k} = m_k] = \mathbb{P}[N_{t_1} = m_1, \dots, N_{t_k} = m_k]$$

και επομένως ότι η $\{\tilde{N}_t\}_{t \geq 0}$ έχει τις κατανομές πεπερασμένης διάστασης μιας διαδικασίας Poisson με ρυθμό λ . Από τον τύπο της ολικής πιθανότητας έχουμε ότι

$$\mathbb{P}[\tilde{A} \cap B] = \sum_{m=0}^{\infty} \mathbb{P}[\tilde{A} \cap B \mid N_s = m] \mathbb{P}[N_s = m]. \quad (8.6)$$

Παρατηρήστε τώρα ότι στο το ενδεχόμενο $\{N_s = m\}$ έχουμε

$$N_r = \sum_{k=1}^m \mathbb{1}\{S_k \leq r\}, \quad \text{για κάθε } r \leq s.$$

Αυτό συμβαίνει γιατί γνωρίζοντας ότι $N_s = m$ μπορούμε να περιορίσουμε το άθροισμα στη σχέση (8.2) στους m πρώτους όρους. Από το Λήμμα 11 έχουμε ότι δεδομένου ότι $N_s = m$, η $\{\tilde{N}_t\}_{t \geq 0}$ είναι ανεξάρτητη

από τις E_1, \dots, E_m . Επομένως, δεδομένου ότι $N_s = m$, η $\{\tilde{N}_t\}_{t \geq 0}$ θα είναι ανεξάρτητη και από τις $\{N_r\}_{0 \leq r \leq s}$ και η (8.6) θα γίνει

$$\mathbb{P}[\tilde{A} \cap B] = \sum_{m=0}^{\infty} \mathbb{P}[\tilde{A} \mid N_s = m] \mathbb{P}[B \mid N_s = m] \mathbb{P}[N_s = m]. \quad (8.7)$$

Από το Λήμμα 11 έχουμε ακόμα πως, δεδομένου ότι $N_s = m$, η $\{\tilde{N}_t\}_{t \geq 0}$ είναι διαδικασία Poisson με ρυθμό λ . Επομένως, $\mathbb{P}[\tilde{A} \mid N_s = m] = \mathbb{P}[A]$ και η (8.7) δίνει πράγματι ότι $\mathbb{P}[\tilde{A} \cap B] = \mathbb{P}[A] \mathbb{P}[B]$. \square

Πόρισμα 12 Η διαδικασία Poisson έχει ανεξάρτητες και χρονικά ομοιογενείς προσαυξήσεις. Συγκεκριμένα, για κάθε $s, t \geq 0$ η προσαύξηση $N_{t+s} - N_s$ είναι ανεξάρτητη από τις προηγούμενες τιμές της διαδικασίας και ακολουθεί την ίδια κατανομή όπως η N_t , δηλαδή $Po(\lambda t)$.

Θα δούμε τώρα πώς η πληροφορία που περιέχεται στο προηγούμενο πόρισμα καθορίζει τις κατανομές πεπερασμένης διάστασης της $\{N_t\}_{t \geq 0}$.

Θεώρημα 35 Έστω $t_0 = 0 \leq t_1 \leq t_2 \leq \dots \leq t_n$. Για οποιουδήποτε ακεραίους $k_0 = 0 \leq k_1 \leq k_2 \leq \dots \leq k_n$ έχουμε

$$\mathbb{P}[N_{t_1} = k_1, \dots, N_{t_n} = k_n] = e^{-\lambda t_n} \lambda^{k_n} \prod_{i=0}^{n-1} \frac{(t_{i+1} - t_i)^{k_{i+1} - k_i}}{(k_{i+1} - k_i)!}. \quad (8.8)$$

Από το Θεώρημα 34 έχουμε

$$\begin{aligned} \mathbb{P}[N_{t_1} = k_1, \dots, N_{t_n} = k_n] &= \mathbb{P}[N_{t_1} = k_1, N_{t_2} - N_{t_1} = k_2 - k_1, \dots, N_{t_n} - N_{t_{n-1}} = k_n - k_{n-1}] \\ &= \mathbb{P}[N_{t_1} = k_1] \mathbb{P}[N_{t_2} - N_{t_1} = k_2 - k_1] \cdots \mathbb{P}[N_{t_n} - N_{t_{n-1}} = k_n - k_{n-1}] \\ &= \mathbb{P}[N_{t_1} = k_1] \mathbb{P}[N_{t_2 - t_1} = k_2 - k_1] \cdots \mathbb{P}[N_{t_n - t_{n-1}} = k_n - k_{n-1}] \\ &= \prod_{i=0}^{n-1} e^{-\lambda(t_{i+1} - t_i)} \frac{(\lambda(t_{i+1} - t_i))^{k_{i+1} - k_i}}{(k_{i+1} - k_i)!} \\ &= e^{-\lambda t_n} \lambda^{k_n} \prod_{i=0}^{n-1} \frac{(t_{i+1} - t_i)^{k_{i+1} - k_i}}{(k_{i+1} - k_i)!}. \end{aligned}$$

Πόρισμα 13 Αν μια διαδικασία $\{N_t\}_{t \geq 0}$ έχει ανεξάρτητες προσαυξήσεις με $N_{s+t} - N_s \sim Po(\lambda t)$, τότε η $\{N_t\}_{t \geq 0}$ έχει την κατανομή μιας διαδικασίας Poisson με ρυθμό λ .

Απόδειξη: Είδαμε στην απόδειξη του Θεωρήματος 35 ότι μια διαδικασία $\{N_t\}_{t \geq 0}$ η οποία έχει ανεξάρτητες προσαυξήσεις με $N_{s+t} - N_s \sim Po(\lambda t)$ έχει τις κατανομές πεπερασμένης διάστασης που φαίνονται στη σχέση (8.8). Επομένως έχει τις ίδιες κατανομές πεπερασμένης διάστασης με μια διαδικασία Poisson με ρυθμό λ . Ο ισχυρισμός έπεται από το γεγονός ότι οι κατανομές πεπερασμένης διάστασης καθορίζουν την κατανομή μιας στοχαστικής διαδικασίας. \square

Προκειμένου να καταλάβουμε διαισθητικά τη σχέση (8.8) ας θυμηθούμε την πολυωνυμική κατανομή, την οποία μάλλον έχετε συναντήσει σε κάποιο εισαγωγικό μάθημα Πιθανοτήτων. Η πολυωνυμική κατανομή με παραμέτρους n, p_1, \dots, p_m είναι μια διακριτή κατανομή σε m διαστάσεις με σ.μ.π.

$$p(n_1, \dots, n_m) = \binom{n}{n_1, \dots, n_m} p_1^{n_1} \cdots p_m^{n_m}, \quad n_1, \dots, n_m \in \mathbb{N}_0,$$

όπου

$$\binom{n}{n_1, \dots, n_m} = \frac{n!}{n_1! \cdots n_m!}, \quad \text{αν } n_1 + \cdots + n_m = n,$$

και μηδέν, αν $n_1 + \cdots + n_m \neq n$. Εμφανίζεται φυσικά όταν μελετάμε την κατανομή σφαιρών σε κελιά. Συγκεκριμένα, αν τοποθετήσουμε n σφαίρες σε m κελιά, και επιλέγουμε το κελί i με πιθανότητα p_i ανεξάρτητα για κάθε σφαίρα, τότε η πιθανότητα να τοποθετηθούν n_1 σφαίρες στο κελί 1, \dots , n_m σφαίρες στο κελί m είναι ίση με $p(n_1, \dots, n_m)$.

Μπορούμε τώρα να παρατηρήσουμε, διαιρώντας τα δύο μέλη της (8.8) με $\mathbb{P}[N_{t_n} = k_n] = e^{-\lambda t_n} \frac{(\lambda t_n)^{k_n}}{k_n!}$, ότι

$$\mathbb{P}[N_{t_1} = k_1, \dots, N_{t_n} = k_n \mid N_{t_n} = k_n] = \binom{k_n}{k_1, k_2 - k_1, \dots, k_n - k_{n-1}} \prod_{i=0}^{n-1} \left(\frac{t_{i+1} - t_i}{t_n} \right)^{k_{i+1} - k_i}.$$

Επομένως, δεδομένου του πλήθους των αφίξεων k_n στο διάστημα $(0, t_n]$, οι αφίξεις που σημειώθηκαν στα διαστήματα

$$I_1 = (0, t_1], I_2 = (t_1, t_2], I_3 = (t_2, t_3], \dots, I_n = (t_{n-1}, t_n]$$

ακολουθούν πολυωνυμική κατανομή με πιθανότητα επιτυχίας για κάθε διάστημα $p_k = \frac{t_k - t_{k-1}}{t_n}$, ίση δηλαδή με το κλάσμα του μήκους του διαστήματος I_k στο συνολικό μήκος t_n . Με διαφορετικά λόγια, δεδομένου ότι μέχρι τη χρονική στιγμή t_n είχαμε k_n αφίξεις, η κατανομή αυτών των αφίξεων στα διαστήματα I_1, I_2, \dots, I_n είναι ίδια όπως αν τοποθετούσαμε τις k_n αφίξεις ανεξάρτητα και ομοιόμορφα στο διάστημα $(0, t_n]$. Αυτό ακριβώς είναι το περιεχόμενο του επόμενου θεωρήματος.

Θεώρημα 36 Έστω $\{N_t\}_{t \geq 0}$ μια διαδικασία Poisson με ρυθμό αφίξεων λ . Δεδομένου ότι $N_t = n$, οι χρόνοι S_1, S_2, \dots, S_n , στους οποίους η διαδικασία Poisson αλλάζει κατάσταση, ακολουθούν την ίδια κατανομή όπως μια διατεταγμένη n -άδα σημείων, ομοιόμορφα επιλεγμένων στο $(0, t]$. Συγκεκριμένα, έχουν από κοινού συνάρτηση πυκνότητας πιθανότητας που δίνεται από την

$$g(s_1, s_2, \dots, s_n) = n! t^{-n} \mathbb{1}\{0 \leq s_1 \leq s_2 \leq \cdots \leq s_n \leq t\}. \quad (8.9)$$

Απόδειξη: Η συνάρτηση πυκνότητας πιθανότητας του διανύσματος $E = (E_1, E_2, \dots, E_n, E_n + 1)$ είναι η

$$f(t_1, t_2, \dots, t_n, t_{n+1}) = \lambda^{n+1} e^{-(\lambda t_1 + \cdots + \lambda t_{n+1})} \mathbb{1}\{t_1, \dots, t_{n+1} \geq 0\}.$$

Αν ορίσουμε $S = (S_1, S_2, \dots, S_n, S_{n+1})$, έχουμε ότι $S = \Phi(E)$, όπου

$$\Phi(t_1, \dots, t_{n+1}) = (t_1, t_1 + t_2, \dots, t_1 + \cdots + t_{n+1}), \quad (t_1, \dots, t_{n+1}) \in \mathbb{R}_+^{n+1}.$$

Ο γραμμικός μετασχηματισμός Φ έχει ορίζουσα 1 και ο αντίστροφός του Φ^{-1} δίνεται από την

$$\Phi^{-1}(s_1, \dots, s_{n+1}) = (s_1, s_2 - s_1, \dots, s_{n+1} - s_n), \quad 0 \leq s_1 \leq s_2 \leq \cdots \leq s_{n+1}.$$

Επομένως, η συνάρτηση πυκνότητας πιθανότητας του S είναι

$$f(s_1, s_2 - s_1, \dots, s_{n+1} - s_n) = \lambda^{n+1} e^{-\lambda s_{n+1}} \mathbb{1}\{0 \leq s_1 \leq s_2 \leq \cdots \leq s_{n+1}\}.$$

Έχουμε τώρα ότι για κάθε Borel σύνολο $A \subset \mathbb{R}^n$

$$\begin{aligned} \mathbb{P}[(S_1, \dots, S_n) \in A, N_t = n] &= \mathbb{P}[(S_1, \dots, S_n) \in A, S_n \leq t, S_{n+1} > t] \\ &= \lambda^{n+1} \int_A \int_t^\infty e^{-\lambda s_{n+1}} \mathbb{1}\{0 \leq s_1 \leq s_2 \leq \cdots \leq s_n \leq t\} ds_{n+1} ds_1 \cdots ds_n \\ &= \lambda^n e^{-\lambda t} \int_A \mathbb{1}\{0 \leq s_1 \leq s_2 \leq \cdots \leq s_n \leq t\} ds_1 \cdots ds_n. \end{aligned}$$

Από το Λήμμα 10 έχουμε όμως ότι $\mathbb{P}[N_t = n] = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$. Επομένως,

$$\mathbb{P}[(S_1, \dots, S_n) \in A \mid N_t = n] = \int_A n! t^{-n} \mathbb{1}\{0 \leq s_1 \leq s_2 \leq \dots \leq s_n \leq t\} ds_1 \cdots ds_n.$$

Η παραπάνω σχέση σημαίνει ακριβώς ότι η δεσμευμένη συνάρτηση πυκνότητας πιθανότητας των S_1, \dots, S_n , δεδομένου ότι $N_t = n$, είναι η g της (8.9). Θα δείξουμε τώρα ότι η g είναι επίσης η πυκνότητα της μιας διατεταγμένης κατά μέγεθος n -άδας ανεξάρτητων, ισόνομων τυχαίων μεταβλητών, με ομοιόμορφη κατανομή στο $[0, t]$.

Έστω U_1, \dots, U_n ακολουθία από ανεξάρτητες, ισόνομες, τυχαίες μεταβλητές με ομοιόμορφη κατανομή στο διάστημα $[0, t]$. Αν $\pi : \{1, 2, \dots, n\} \mapsto \{1, 2, \dots, n\}$ είναι μια μετάθεση του συνόλου $\{1, 2, \dots, n\}$, θα συμβολίζουμε με πU το διάνυσμα $(U_{\pi(1)}, U_{\pi(2)}, \dots, U_{\pi(n)})$. Θα συμβολίζουμε επίσης με S το σύνολο των μεταθέσεων του συνόλου $\{1, 2, \dots, n\}$. Επειδή η ομοιόμορφη κατανομή είναι συνεχής, με πιθανότητα 1 οι U_1, \dots, U_n θα λαμβάνουν διαφορετικές τιμές. Επομένως θα υπάρχει μια μοναδική μετάθεση $\pi^<$ για την οποία ισχύει

$$U_{\pi^<(1)} \leq U_{\pi^<(2)} \leq \dots \leq U_{\pi^<(n)}.$$

Φυσικά η μετάθεση $\pi^<$ για την οποία ισχύει η παραπάνω σχέση είναι τυχαία, αφού εξαρτάται από την πραγματοποίηση των U_1, \dots, U_n . Θα δείξουμε τώρα ότι το τυχαίο διάνυσμα

$$\pi^<U = (U_{\pi^<(1)}, \dots, U_{\pi^<(n)}),$$

το οποίο είναι η μετάθεση του διανύσματος $U = (U_1, \dots, U_n)$ που τοποθετεί τις συντεταγμένες του διανύσματος U σε αύξουσα διάταξη, έχει συνάρτηση πυκνότητας πιθανότητας την g της (8.9). Πράγματι, αν A είναι ένα σύνολο Borel του $[0, t]^n$, θα έχουμε

$$\mathbb{P}[\pi^<U \in A] = \sum_{\pi \in S} \mathbb{P}[\pi^<U \in A, \pi^< = \pi] = \sum_{\pi \in S} \mathbb{P}[\pi U \in A, \pi^< = \pi].$$

Επειδή για κάθε $\pi \in S$ τα διανύσματα πU και U έχουν την ίδια κατανομή (οι συντεταγμένες τους είναι ανεξάρτητες και ομοιόμορφες τυχαίες μεταβλητές στο $[0, t]$), οι πιθανότητες στο παραπάνω άθροισμα είναι όλες ίσες, δηλαδή

$$\mathbb{P}[\pi U \in A, \pi^< = \pi] = \mathbb{P}[U \in A, \pi^< = id], \quad \forall \pi \in S,$$

όπου $id \in S$ είναι η ταυτοτική μετάθεση $id(1) = 1, \dots, id(n) = n$. Επομένως,

$$\begin{aligned} \mathbb{P}[\pi^<U \in A] &= |S| \mathbb{P}[U \in A, \pi^< = id] = n! \mathbb{P}[U \in A, U_1 \leq U_2 \leq \dots \leq U_n] \\ &= \int \cdots \int_A n! t^{-n} \mathbb{1}\{0 \leq s_1 \leq s_2 \leq \dots \leq s_n \leq t\} ds_1 \cdots ds_n, \end{aligned}$$

δηλαδή το διάνυσμα $\pi^<U$ έχει συνάρτηση πυκνότητας πιθανότητας την g . □

Το διαισθητικό συμπέρασμα του Θεωρήματος 36 είναι ότι, με δεδομένο ότι σε ένα χρονικό διάστημα $[0, t]$ έχουμε n αφίξεις μιας διαδικασίας Poisson, μπορούμε να ανακατασκευάσουμε την κατανομή των χρόνων άφιξης επιλέγοντας τυχαία (ανεξάρτητα και ομοιόμορφα) n σημεία στο διάστημα $[0, t]$.

Προσέξτε ότι το αποτέλεσμα είναι ανεξάρτητο από τον ρυθμό των αφίξεων της διαδικασίας. Ο ρυθμός επηρεάζει οπωσδήποτε το πλήθος των αφίξεων μέχρι τη στιγμή t . Με δεδομένο όμως το πλήθος των αφίξεων, ο ρυθμός δεν έχει καμία άλλη πληροφορία για το πώς κατανέμονται αυτές οι αφίξεις στο διάστημα $[0, t]$. Όπως είδαμε, αρκεί να τις τοποθετήσουμε τυχαία σ' αυτό το διάστημα.

Σύμφωνα με το προηγούμενο Θεώρημα, θα μπορούσαμε να προσομοιώσουμε ένα μονοπάτι μιας διαδικασίας Poisson με ρυθμό λ στο χρονικό διάστημα $[0, t]$ ως εξής. Αρχικά επιλέγουμε το πλήθος των αφίξεων

που θα τοποθετήσουμε συνολικά σ' αυτό το διάστημα. Από το Λήμμα 10 το πλήθος των αφίξεων στο διάστημα $[0, t]$ ακολουθεί κατανομή Poisson με παράμετρο λt . Έχοντας επιλέξει το πλήθος των αφίξεων n , επιλέγουμε ομοιόμορφα και ανεξάρτητα n σημεία στο $[0, t]$. Αυτά είναι και τα σημεία που θα κάνει άλμα κατά 1 το μονοπάτι της διαδικασίας Poisson.

Παράδειγμα 51 Τα τηλεφωνήματα που δέχεται το ΕΚΑΒ μεταξύ 6 το πρωί και 12 το μεσημέρι είναι μια διαδικασία Poisson με ρυθμό 10/ώρα. Αν από τις 6 μέχρι τις 9 έγιναν δέκα τηλεφωνήματα, ποια είναι η πιθανότητα να έγινε ένα ακριβώς την πρώτη ώρα;

Αν $\{N_t\}_{t \geq 0}$ είναι η διαδικασία που μετρά τα τηλεφωνήματα μετά τις 6, ψάχνουμε για την

$$\mathbb{P}[N_s = 1 \mid N_t = 10],$$

όπου $s=1$ ώρα και $t=3$ ώρες. Μπορούμε να λύσουμε το πρόβλημα με δύο τρόπους. Εκμεταλλευόμενοι τις ανεξάρτητες προσαυξήσεις της διαδικασίας έχουμε

$$\begin{aligned} \mathbb{P}[N_s = 1 \mid N_t = 10] &= \frac{\mathbb{P}[N_s = 1, N_t = 10]}{\mathbb{P}[N_t = 10]} = \frac{\mathbb{P}[N_s = 1, N_t - N_s = 9]}{\mathbb{P}[N_t = 10]} \\ &= \frac{e^{-\lambda s} \frac{(\lambda s)^1}{1!} e^{-\lambda(t-s)} \frac{(\lambda(t-s))^9}{9!}}{e^{-\lambda t} \frac{(\lambda t)^{10}}{10!}} = 10 \times 2^9 \times 3^{-10}. \end{aligned}$$

Στο ίδιο αποτέλεσμα θα μπορούσαμε να φτάσουμε και χρησιμοποιώντας το Θεώρημα 36. Δεδομένου ότι έχουμε 10 τηλεφωνήματα, αυτά κατανέμονται ανεξάρτητα και ομοιόμορφα στο διάστημα $(0, t]$. Έχουμε λοιπόν 10 ανεξάρτητες δοκιμές Bernoulli με πιθανότητα επιτυχίας (να συμβούν κατά την πρώτη ώρα) ίση με $s/t = 1/3$. Το πλήθος των επιτυχιών σ' αυτές τις δοκιμές ακολουθεί διωνυμική κατανομή $\text{bin}(10, 1/3)$, επομένως

$$\mathbb{P}[N_s = 1 \mid N_t = 10] = \binom{10}{1} \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^9.$$

Σκεφτείτε τώρα την πολυπλοκότητα των δύο μεθόδων, αν θέλαμε να βρούμε την πιθανότητα να έχουμε ακριβώς μία άφιξη από τις δέκα, στα πρώτα είκοσι δευτερόλεπτα κάθε λεπτού. Ο πρώτος τρόπος μάλλον δεν είναι ελκυστικός, ο δεύτερος τρόπος όμως μας δίνει εύκολα ότι η απάντηση παραμένει ίδια!

Παράδειγμα 52 Ένας φωτεινός σηματοδότης επιτρέπει τη διέλευση πεζών από μια διασταύρωση τα πρώτα 10 δευτερόλεπτα κάθε λεπτού. Οι πεζοί φτάνουν στη διασταύρωση ως μια διαδικασία Poisson. Αν σε διάστημα μιας ώρας έφτασαν στη διασταύρωση 500 πεζοί, ποια είναι προσεγγιστικά η πιθανότητα το πολύ 90 από αυτούς να διέσχισαν τη διασταύρωση χωρίς να χρειαστεί να περιμένουν;

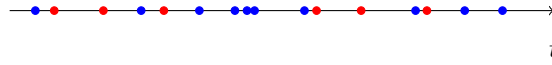
Σύμφωνα με το Θεώρημα 36 οι χρόνοι άφιξης των 500 πεζών έχουν την ίδια κατανομή όπως 500 ανεξάρτητες ισόνομες τυχαίες μεταβλητές με ομοιόμορφη κατανομή στο διάστημα $[0, T]$, όπου $T = 60\text{min}$. Επομένως, για καθέναν από αυτούς η πιθανότητα να έφτασε στο σηματοδότη κάποια στιγμή που η ένδειξη του σηματοδότη ήταν πράσινη είναι ίση με το ποσοστό του χρόνου που η ένδειξη του σηματοδότη ήταν πράσινη στη διάρκεια μιας ώρας, δηλαδή $\frac{1}{6}$. Το πλήθος N των πεζών που έφτασαν στη διασταύρωση κάποια στιγμή που ένδειξη του σηματοδότη ήταν πράσινη και τη διέσχισαν χωρίς να χρειαστεί να περιμένουν ακολουθεί λοιπόν διωνυμική κατανομή $\text{bin}(500, 1/6)$. Εφόσον $500 \times \frac{1}{6} \geq 10$, μπορούμε να χρησιμοποιήσουμε την προσέγγιση από το Κεντρικό Οριακό Θεώρημα ώστε να εκτιμήσουμε την πιθανότητα

$$\mathbb{P}[N \leq 90] = \mathbb{P}\left[\frac{N - 500 \times \frac{1}{6}}{\sqrt{500 \frac{1}{6} \frac{5}{6}}} \leq \frac{90 - 500 \times \frac{1}{6}}{\sqrt{500 \frac{1}{6} \frac{5}{6}}}\right] \simeq \Phi\left(\frac{4}{5}\right) \simeq 0,78814.$$

8.3 Πρόσθεση και εκλέπτυνση

Σ' αυτήν την παράγραφο θα δούμε δύο χρήσιμους χειρισμούς, στους οποίους μπορούμε να υποβάλλουμε διαδικασίες Poisson και το αποτέλεσμα είναι πάλι μια διαδικασία Poisson, την πρόσθεση ανεξάρτητων διαδικασιών και την εκλέπτυνση (thinning).

Στο παρακάτω σχήμα φαίνεται μια πραγματοποίηση των χρόνων άφιξης δύο ανεξάρτητων διαδικασιών Poisson $\{X_t\}_{t \geq 0}$ (με ρυθμό 2 σε μπλε χρώμα) και $\{Y_t\}_{t \geq 0}$ (με ρυθμό 1 σε κόκκινο χρώμα).



Το πρώτο αποτέλεσμα που θα δείξουμε έχει να κάνει με τη διαδικασία των αφίξεων που βλέπουμε, αν αγνοήσουμε το χρώμα τους.

Θεώρημα 37 Αν $\{X_t\}_{t \geq 0}$ και $\{Y_t\}_{t \geq 0}$ είναι δύο ανεξάρτητες διαδικασίες Poisson με ρυθμούς λ και μ αντίστοιχα, τότε η διαδικασία $\{N_t\}_{t \geq 0}$ με $N_t = X_t + Y_t$, είναι μια διαδικασία Poisson με ρυθμό $\lambda + \mu$.

Απόδειξη: Σύμφωνα με το Πρόσχημα 13, αρκεί να δείξουμε ότι για κάθε s, t με $0 \leq s \leq t$, η τυχαία μεταβλητή $N_t - N_s$ είναι ανεξάρτητη από τις $\{N_r\}_{0 \leq r \leq s}$ και ακολουθεί κατανομή $Po((\lambda + \mu)(t - s))$. Έχουμε όμως

$$N_t - N_s = (X_t - X_s) + (Y_t - Y_s).$$

Η $X_t - X_s$ είναι ανεξάρτητη από τις $\{X_r\}_{0 \leq r \leq s}$, αφού η $\{X_t\}_{t \geq 0}$ είναι διαδικασία Poisson. Επίσης είναι ανεξάρτητη από τις $\{Y_r\}_{0 \leq r \leq s}$, αφού οι διαδικασίες που προσθέτουμε είναι ανεξάρτητες. Επομένως, η $X_t - X_s$ είναι ανεξάρτητη από τις $\{N_r\}_{0 \leq r \leq s}$. Με τον ίδιο τρόπο βλέπουμε ότι και η $Y_t - Y_s$ είναι ανεξάρτητη από τις $\{N_r\}_{0 \leq r \leq s}$. Επομένως η $\{N_t\}_{t \geq 0}$ έχει ανεξάρτητες προσαυξήσεις.

Σχετικά με την κατανομή της $N_t - N_s$, προσέξτε ότι η $X_t - X_s$ ακολουθεί κατανομή $Po(\lambda(t - s))$, η $Y_t - Y_s$ ακολουθεί κατανομή $Po(\mu(t - s))$, ενώ οι $X_t - X_s$ και $Y_t - Y_s$ είναι ανεξάρτητες. Συνεπώς, το άθροισμά τους ακολουθεί κατανομή Poisson με ρυθμό το άθροισμα των δύο ρυθμών, δηλαδή $(\lambda + \mu)(t - s)$.

□

Παρατήρηση: Το παραπάνω Θεώρημα γενικεύεται εύκολα με επαγωγή, σε αθροίσματα n ανεξάρτητων διαδικασιών Poisson. Το άθροισμά τους είναι μια διαδικασία Poisson με ρυθμό το άθροισμα των επί μέρους ρυθμών.

Το επόμενο Θεώρημα περιγράφει την κατανομή του πλήθους των αφίξεων κάθε είδους, με δεδομένο τον συνολικό αριθμό αφίξεων.

Θεώρημα 38 Αν $\{X_t^{(i)}\}_{t \geq 0}, i = 1, 2, \dots, m$ είναι ανεξάρτητες διαδικασίες Poisson με ρυθμούς $\lambda_i, i = 1, 2, \dots, m$ και $\{N_t\}_{t \geq 0}$ είναι το άθροισμά τους, τότε για κάθε $k_1, \dots, k_m \in \mathbb{N}_0$ έχουμε

$$\mathbb{P}[X_t^{(1)} = k_1, \dots, X_t^{(m)} = k_m | N_t = n] = \binom{n}{k_1, \dots, k_m} \left(\frac{\lambda_1}{\lambda}\right)^{k_1} \dots \left(\frac{\lambda_m}{\lambda}\right)^{k_m}, \quad (8.10)$$

όπου $\lambda = \lambda_1 + \dots + \lambda_m$. Επομένως, αν ξέρουμε τον συνολικό αριθμό αφίξεων σε κάποιο διάστημα, μπορούμε να ανακατασκευάσουμε το είδος των αφίξεων αποδίδοντας κάθε άφιξη στο είδος i με πιθανότητα $p_i = \lambda_i/\lambda$.

Απόδειξη: Αρκεί να αποδείξουμε τον ισχυρισμό όταν $k_1 + \dots + k_m = n$, αφού διαφορετικά και τα δυο μέλη της (8.10) είναι ίσα με μηδέν. Από το Θεώρημα 37 η διαδικασία $\{N_t\}_{t \geq 0}$ είναι μια διαδικασία Poisson με ρυθμό λ . Από την ανεξαρτησία των διαδικασιών $\{X_t^{(i)}\}_{t \geq 0}$ και το Λήμμα 10 έχουμε

$$\begin{aligned} \mathbb{P}[X_t^{(1)} = k_1, \dots, X_t^{(m)} = k_m \mid N_t = n] &= \frac{\mathbb{P}[X_t^{(1)} = k_1, \dots, X_t^{(m)} = k_m]}{\mathbb{P}[N_t = n]} \\ &= \frac{\mathbb{P}[X_t^{(1)} = k_1] \cdots \mathbb{P}[X_t^{(m)} = k_m]}{\mathbb{P}[N_t = n]} \\ &= \frac{e^{-\lambda_1 t} \frac{(\lambda_1 t)^{k_1}}{k_1!} \cdots e^{-\lambda_m t} \frac{(\lambda_m t)^{k_m}}{k_m!}}{e^{-\lambda t} \frac{(\lambda t)^n}{n!}} \\ &= \frac{n!}{k_1! \cdots k_m!} \frac{\lambda_1^{k_1} \cdots \lambda_m^{k_m}}{\lambda^n} \\ &= \binom{n}{k_1, \dots, k_m} \left(\frac{\lambda_1}{\lambda}\right)^{k_1} \cdots \left(\frac{\lambda_m}{\lambda}\right)^{k_m}. \end{aligned}$$

Για να δείξουμε τον τελευταίο ισχυρισμό, αρκεί να παρατηρήσουμε ότι το δεξί μέλος της τελευταίας σχέσης είναι η σ.μ.π. της πολυωνυμικής κατανομής με παραμέτρους n, p_1, \dots, p_m . \square

Πόρισμα 14 Αν $\{X_t^{(i)}\}_{t \geq 0}, i = 1, 2, \dots, m$ είναι ανεξάρτητες διαδικασίες Poisson με ρυθμούς $\lambda_i, i = 1, 2, \dots, m$ και $\{N_t\}_{t \geq 0}$ είναι το άθροισμά τους, τότε

$$\mathbb{P}[X_t^{(i)} = k \mid N_t = n] = \binom{n}{k} \left(\frac{\lambda_i}{\lambda}\right)^k \left(1 - \frac{\lambda_i}{\lambda}\right)^{n-k},$$

όπου $\lambda = \lambda_1 + \dots + \lambda_m$. Επομένως, με δεδομένο τον συνολικό αριθμό αφίξεων n σε ένα διάστημα, οι αφίξεις του είδους i ακολουθούν διωνυμική κατανομή $\text{bin}(n, \frac{\lambda_i}{\lambda})$.

Απόδειξη: Θεωρήστε τη διαδικασία $\{Y_t\}_{t \geq 0}$ με $Y_t = \sum_{j \neq i} X_t^{(j)}$, η οποία είναι μια διαδικασία Poisson με ρυθμό $\mu = \sum_{j \neq i} \lambda_j = \lambda - \lambda_i$ (Θεώρημα 37) και είναι ανεξάρτητη από την $\{X_t^{(i)}\}_{t \geq 0}$. Η (8.10) δίνει ότι

$$\mathbb{P}[X_t^{(i)} = k \mid N_t = n] = \mathbb{P}[X_t^{(i)} = k, Y_t = n - k \mid N_t = n] = \binom{n}{k} \left(\frac{\lambda_i}{\lambda}\right)^k \left(1 - \frac{\lambda_i}{\lambda}\right)^{n-k}.$$

\square

Παράδειγμα 53 Ένα Τμήμα Επειγόντων Περιστατικών εφημερεύει μόνο για παθολογικά, χειρουργικά και καρδιολογικά περιστατικά. Τα παθολογικά περιστατικά καταφτάνουν όπως μια διαδικασία Poisson με ρυθμό $\lambda = 1/10\text{min}$, τα χειρουργικά περιστατικά όπως μια διαδικασία Poisson με ρυθμό $\mu = 1/15\text{min}$, ενώ τα καρδιολογικά περιστατικά όπως μια διαδικασία Poisson με ρυθμό $\nu = 1/30\text{min}$. Οι τρεις διαδικασίες είναι ανεξάρτητες. Αν 20 λεπτά μετά την έναρξη της εφημερίας είχαν προσέλθει 6 ασθενείς, ποια είναι οι πιθανότητες να είχαμε 2 ασθενείς από κάθε είδος περιστατικού; Αν κατά τη διάρκεια της εφημερίας εμφανίστηκαν 300 ασθενείς, ποια είναι προσεγγιστικά η πιθανότητα να υπήρχαν ανάμεσά τους τουλάχιστον 50 καρδιολογικά περιστατικά;

Οι συνολικές αφίξεις στο Τμήμα Επειγόντων είναι επομένως μια διαδικασία Poisson $\{N_t\}_{t \geq 0}$ με ρυθμό $\lambda + \mu + \nu = 1/5\text{min}$. Από το Θεώρημα 38 έχουμε ότι

$$\mathbb{P}[\Pi_t = 2, X_t = 2, K_t = 2 \mid N_t = 6] = \frac{6!}{(2!)^3} \left(\frac{1}{2}\right)^2 \left(\frac{1}{3}\right)^2 \left(\frac{1}{6}\right)^2 = \frac{5}{72}.$$

Από το Πρόγραμμα 14, με δεδομένο ότι κατά τη διάρκεια της εφημερίας εμφανίστηκαν 300 ασθενείς, το πλήθος K_T των καρδιολογικών περιστατικών ακολουθεί διωνυμική κατανομή $\text{bin}(300, \frac{1}{6})$. Επομένως,

$$\mathbb{P}[K_T \geq 50 | N_T = 300] = \sum_{k=50}^{300} \binom{300}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{300-k}.$$

Μπορούμε να εκτιμήσουμε το άθροισμα αυτό χρησιμοποιώντας την προσέγγιση της διωνυμικής κατανομής με βάση το Κεντρικό Οριακό Θεώρημα. Συγκεκριμένα,

$$\mathbb{P}[K_T \geq 50 | N_T = 300] = \mathbb{P}\left[\frac{K_T - 300 \times \frac{1}{6}}{\sqrt{300 \times \frac{1}{6} \times \frac{5}{6}}} \geq 0 | N_T = 300\right] \simeq 1 - \Phi(0) = \frac{1}{2}.$$

□

Παράδειγμα 54 Στο προηγούμενο παράδειγμα, υπολογίστε την πιθανότητα το πρώτο περιστατικό που θα εμφανιστεί στην εφημερία να είναι παθολογικό.

Έστω X, Y, Z οι χρόνοι αναμονής μέχρι την άφιξη του πρώτου παθολογικού, χειρουργικού ή καρδιολογικού περιστατικού, αντίστοιχα. Θέλουμε να υπολογίσουμε την πιθανότητα του ενδεχομένου $\{X \leq Y\} \cap \{X \leq Z\} = \{X \leq Y \wedge Z\}$. Μπορούμε να κάνουμε τον υπολογισμό απ' ευθείας, χρησιμοποιώντας την από κοινού συνάρτηση πυκνότητας πιθανότητας των X, Y, Z . Μπορούμε όμως και να παρατηρήσουμε ότι η $Y \wedge Z$ είναι εκθετική με ρυθμό $\mu + \nu$. Πράγματι,

$$\mathbb{P}[Y \wedge Z > t] = \mathbb{P}[\{Y > t\} \cap \{Z > t\}] = \mathbb{P}[Y > t] \mathbb{P}[Z > t] = e^{-\mu t} e^{-\nu t} = e^{-(\mu+\nu)t}.$$

Αυτό θα έπρεπε να το περιμένουμε. Η $Y \wedge Z$ είναι ο χρόνος αναμονής μέχρι την πρώτη άφιξη της διαδικασίας $\{X_t + K_t\}_{t \geq 0}$, η οποία σύμφωνα με το Θεώρημα 37 είναι μια διαδικασία Poisson με ρυθμό $\mu + \nu$.

Εφόσον οι Y, Z είναι ανεξάρτητες από την X , οι $Y \wedge Z$ και X θα είναι επίσης ανεξάρτητες. Επομένως,

$$\mathbb{P}[X \leq Y \wedge Z] = \int_0^\infty \lambda e^{-\lambda x} \int_x^\infty (\mu + \nu) e^{-(\mu+\nu)v} dv dx = \lambda \int_0^\infty e^{-(\lambda+\mu+\nu)x} dx = \frac{\lambda}{\lambda + \mu + \nu} = \frac{1}{2}.$$

Αντίστοιχα, η πιθανότητα το πρώτο περιστατικό να είναι χειρουργικό είναι $\frac{\mu}{\lambda+\mu+\nu} = \frac{1}{3}$ και η πιθανότητα το πρώτο περιστατικό να είναι καρδιολογικό είναι $\frac{\nu}{\lambda+\mu+\nu} = \frac{1}{6}$.

Θα μπορούσαμε άραγε να ισχυριστούμε ότι κάθε περιστατικό που εμφανίζεται είναι παθολογικό, χειρουργικό ή καρδιολογικό με πιθανότητα $\frac{1}{2}, \frac{1}{3}$ και $\frac{1}{6}$ αντίστοιχα, ανεξάρτητα από τα άλλα περιστατικά; Κάτι τέτοιο υποδεικνύει και το περιεχόμενο του Θεωρήματος 38. Θα μπορούσαμε πράγματι να το ισχυριστούμε, αν γενικεύαμε το Θεώρημα 34 ώστε να εξασφαλίσουμε την ισχυρή μαρκοβιανή ιδιότητα για τον χρόνο διακοπής $T = X \wedge Y \wedge Z$. Η άφιξη του δεύτερου ασθενούς είναι η πρώτη άφιξη μετά τον χρόνο T , οπότε το πρόβλημα της δεύτερης άφιξης ανάγεται στο πρόβλημα που μόλις λύσαμε, ενώ μπορούμε να επεκτείνουμε τον παραπάνω συλλογισμό επαγωγικά. Στο ίδιο αποτέλεσμα μπορούμε να καταλήξουμε χρησιμοποιώντας την έννοια της εκλέπτυνσης μιας διαδικασίας Poisson, με την οποία θα ασχοληθούμε τώρα. □

Ορισμός: Η p -εκλέπτυνση μιας διαδικασίας Poisson $\{N_t\}_{t \geq 0}$ είναι η διαδικασία αφίξεων που προκύπτει αν διαγράψουμε κάθε άφιξη της $\{N_t\}_{t \geq 0}$ με πιθανότητα $1 - p$ ανεξάρτητα από τις άλλες αφίξεις και τη διαδικασία $\{N_t\}_{t \geq 0}$.

Για να κατανοήσουμε καλύτερα την εκλέπτυνση μας κάνουμε το ακόλουθο νοητικό πείραμα. Θα θεωρήσουμε μια ακολουθία $\{X_i\}_{i \in \mathbb{N}}$ από ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με κατανομή Bernoulli(p), που είναι ανεξάρτητες από τη διαδικασία αφίξεων $\{N_t\}_{t \geq 0}$ και θα χρωματίσουμε την i -οστή άφιξη κόκκινη, αν $X_i = 1$,

ή μπλε, αν $X_i = 0$. Η διαδικασία που περιγράφει τις κόκκινες αφίξεις είναι μια p -εκλέπτυνση της $\{N_t\}_{t \geq 0}$. Το πλήθος των κόκκινων αφίξεων μέχρι τη χρονική στιγμή t είναι $\sum_{i=1}^{N_t} X_i$. Αντίστοιχα, η διαδικασία που περιγράφει τις μπλε αφίξεις είναι μια $(1-p)$ -εκλέπτυνση της $\{N_t\}_{t \geq 0}$ και το πλήθος των μπλε αφίξεων μέχρι τη χρονική στιγμή t είναι $\sum_{i=1}^{N_t} (1 - X_i)$. Το επόμενο Θεώρημα χαρακτηρίζει τις διαδικασίες των κόκκινων και μπλε αφίξεων.

Θεώρημα 39 Έστω $\{N_t\}_{t \geq 0}$ μια διαδικασία Poisson με ρυθμό λ και $\{X_i\}_{i \in \mathbb{N}}$ μια ακολουθία από ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με κατανομή Bernoulli(p), που είναι ανεξάρτητες από τη διαδικασία αφίξεων $\{N_t\}_{t \geq 0}$. Τότε οι διαδικασίες $\{N_t^+\}_{t \geq 0}$ και $\{N_t^-\}_{t \geq 0}$ με

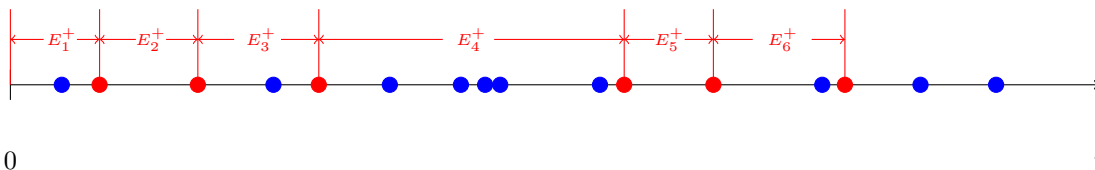
$$N_t^+ = \sum_{i=1}^{N_t} X_i \quad \text{και} \quad N_t^- = N_t - N_t^+ = \sum_{i=1}^{N_t} (1 - X_i),$$

είναι δύο ανεξάρτητες διαδικασίες Poisson με ρυθμούς λp και $\lambda(1-p)$ αντίστοιχα.

Απόδειξη: Έστω M_1, M_2, \dots , η ακολουθία των επιτυχημένων προσπαθειών, δηλαδή

$$M_1 = \inf\{k \geq 1 : X_k = 1\}, \quad M_{n+1} = \inf\{k > M_n : X_k = 1\}, n \in \mathbb{N}.$$

Για παράδειγμα, στο παρακάτω σχήμα, όπου έχουμε χρωματίσει την k -οστή άφιξη κόκκινη, αν $X_k = 1$, ή μπλε, αν $X_k = 0$, έχουμε $M_1 = 2, M_2 = 3, M_3 = 5, M_4 = 11, M_5 = 12, M_6 = 14$.



Οι $M_1, M_2 - M_1, M_3 - M_2, \dots$ είναι ανεξάρτητες, ισόνομες τυχαίες μεταβλητές, που ακολουθούν γεωμετρική κατανομή με παράμετρο p , δηλαδή

$$\mathbb{P}[M_1 = k] = (1-p)^{k-1}p, \quad k \in \mathbb{N}. \quad (8.11)$$

Αν $\{E_k\}_{k \in \mathbb{N}}$ είναι η ακολουθία των χρόνων που μεσολαβούν ανάμεσα στις διαδοχικές αφίξεις της διαδικασίας $\{N_t\}_{t \geq 0}$ και ορίσουμε $M_0 = 0$, τότε οι χρόνοι ανάμεσα στις διαδοχικές αφίξεις της $\{N_t^+\}_{t \geq 0}$ είναι η ακολουθία $\{E_k^+\}_{k \in \mathbb{N}}$ με

$$E_n^+ = \sum_{k=M_{n-1}+1}^{M_n} E_k.$$

Εφόσον η $\{M_n\}_{n \in \mathbb{N}}$ εξαρτάται μόνο από την $\{X_i\}_{i \in \mathbb{N}}$ και η $\{E_k\}_{k \in \mathbb{N}}$ μόνο από τη διαδικασία αφίξεων, οι ακολουθίες αυτές είναι ανεξάρτητες. Επιπλέον, η ανεξαρτησία και ισονομία των $M_1, M_2 - M_1, \dots$ συνεπάγεται την ανεξαρτησία και ισονομία των E_1^+, E_2^+, \dots . Θα δείξουμε ότι οποιεσδήποτε δύο από αυτές είναι ανεξάρτητες και ισόνομες καθώς το επιχείρημα στη γενική περίπτωση είναι εντελώς ανάλογο. Για $n, m \in \mathbb{N}$ με $n < m$ έχουμε

$$\begin{aligned} I &= \mathbb{P}[E_n^+ > t, E_m^+ > s] = \mathbb{P}\left[\sum_{k=M_{n-1}+1}^{M_n} E_k > t, \sum_{k=M_{m-1}+1}^{M_m} E_k > s\right] \\ &= \sum_{p_1, p_2 \in \mathbb{N}} \sum_{q_2 \geq q_1 + p_1} \mathbb{P}\left[\sum_{k=q_1+1}^{q_1+p_1} E_k > t, \sum_{k=q_2+1}^{q_2+p_2} E_k > s, M_{n-1} = q_1 = M_n - p_1, M_{m-1} = q_2 = M_m - p_2\right]. \end{aligned}$$

Από την ανεξαρτησία των ακολουθιών $\{M_n\}_{n \in \mathbb{N}}$ και $\{E_n\}_{n \in \mathbb{N}}$ παίρνουμε

$$I = \sum_{p_1, p_2 \in \mathbb{N}} \sum_{q_2 \geq q_1 + p_1} \mathbb{P} \left[\sum_{k=q_1+1}^{q_1+p_1} E_k > t, \sum_{k=q_2+1}^{q_2+p_2} E_k > s \right] \mathbb{P} [M_{n-1} = q_1 = M_n - p_1, M_{m-1} = q_2 = M_m - p_2].$$

Από την ανεξαρτησία και την ισονομία των τυχαίων μεταβλητών $E_n, n = 1, 2, \dots$ παίρνουμε στη συνέχεια

$$\begin{aligned} I &= \sum_{p_1, p_2 \in \mathbb{N}} \sum_{q_2 \geq q_1 + p_1} \mathbb{P} \left[\sum_{k=1}^{p_1} E_k > t \right] \mathbb{P} \left[\sum_{k=1}^{p_2} E_k > s \right] \times \\ &\quad \times \mathbb{P} [M_{n-1} = q_1, M_n - M_{n-1} = p_1, M_{m-1} = q_2, M_m - M_{m-1} = p_2] \\ &= \sum_{p_1, p_2 \in \mathbb{N}} \mathbb{P} \left[\sum_{k=1}^{p_1} E_k > t \right] \mathbb{P} \left[\sum_{k=1}^{p_2} E_k > s \right] \mathbb{P} [M_n - M_{n-1} = p_1, M_m - M_{m-1} = p_2] \end{aligned}$$

Τέλος, από το γεγονός ότι οι $M_n - M_{n-1}$ και $M_m - M_{m-1}$ είναι ανεξάρτητες και έχουν την ίδια κατανομή με τη M_1 παίρνουμε

$$\begin{aligned} I &= \sum_{p_1, p_2 \in \mathbb{N}} \mathbb{P} \left[\sum_{k=1}^{p_1} E_k > t \right] \mathbb{P} \left[\sum_{k=1}^{p_2} E_k > s \right] \mathbb{P} [M_1 = p_1] \mathbb{P} [M_1 = p_2] \\ &= \left(\sum_{p_1 \in \mathbb{N}} \mathbb{P} \left[\sum_{k=1}^{p_1} E_k > t \right] \mathbb{P} [M_1 = p_1] \right) \left(\sum_{p_2 \in \mathbb{N}} \mathbb{P} \left[\sum_{k=1}^{p_2} E_k > s \right] \mathbb{P} [M_1 = p_2] \right) \\ &= \left(\sum_{p_1 \in \mathbb{N}} \mathbb{P} \left[\sum_{k=1}^{M_1} E_k > t, M_1 = p_1 \right] \right) \left(\sum_{p_2 \in \mathbb{N}} \mathbb{P} \left[\sum_{k=1}^{M_1} E_k > s, M_1 = p_2 \right] \right) \\ &= \mathbb{P} [E_1^+ > t] \mathbb{P} [E_1^+ > s]. \end{aligned}$$

Συνοψίζοντας το παραπάνω επιχείρημα, δείξαμε ότι για κάθε $m, n \in \mathbb{N}$ με $m > n$ και κάθε $s, t > 0$ έχουμε

$$\mathbb{P} [E_n^+ > t, E_m^+ > s] = \mathbb{P} [E_1^+ > t] \mathbb{P} [E_1^+ > s].$$

Επομένως οι τυχαίες μεταβλητές E_n, E_m είναι ανεξάρτητες και έχουν την ίδια κατανομή όπως η E_1^+ . Προκειμένου να βρούμε την κατανομή τους παρατηρούμε όπως πριν ότι

$$\mathbb{P} [E_1^+ > t] = \sum_{n=1}^{\infty} \mathbb{P} \left[\sum_{k=1}^n E_k > t \right] \mathbb{P} [M_1 = n]. \quad (8.12)$$

Ας θυμηθούμε τώρα ότι η τυχαία μεταβλητή $\sum_{k=1}^n E_k$ ακολουθεί κατανομή $G(\lambda, n)$. Επομένως,

$$\mathbb{P} \left[\sum_{k=1}^n E_k > t \right] = \int_t^{\infty} \frac{\lambda^n x^{n-1}}{(n-1)!} e^{-\lambda x} dx.$$

Αν αντικαταστήσουμε την παραπάνω σχέση και την (8.11) στην 8.12 και χρησιμοποιήσουμε το Θεώρημα Fubini-Tonelli για να εναλλάξουμε τη σειρά της άθροισης και της ολοκλήρωσης, παίρνουμε

$$\begin{aligned}
\mathbb{P}[E_1^+ > t] &= \sum_{n=1}^{\infty} \int_t^{\infty} \frac{\lambda^n x^{n-1}}{(n-1)!} e^{-\lambda x} dx (1-p)^{n-1} p \\
&= \lambda p \int_t^{\infty} e^{-\lambda x} \left(\sum_{n=1}^{\infty} \frac{(\lambda(1-p)x)^{n-1}}{(n-1)!} \right) dx \\
&= \lambda p \int_t^{\infty} e^{-\lambda x} e^{\lambda(1-p)x} dx \\
&= \lambda p \int_t^{\infty} e^{-\lambda p x} dx = e^{-\lambda p t}.
\end{aligned}$$

Επομένως οι $\{E_k^+\}_{k \in \mathbb{N}}$ είναι ανεξάρτητες, ισόνομες και ακολουθούν εκθετική κατανομή με παράμετρο λp , άρα η διαδικασία $\{N_t^+\}_{t \geq 0}$ είναι μια διαδικασία Poisson με ρυθμό λp .

Εφόσον η $\{1 - X_i\}_{i \in \mathbb{N}}$ είναι μια ακολουθία από ανεξάρτητες, ισόνομες τυχαίες μεταβλητές, με κατανομή Bernoulli($1-p$), παίρνουμε αμέσως ότι η $\{N_t^-\}_{t \geq 0}$ είναι επίσης μια διαδικασία Poisson με ρυθμό $\lambda(1-p)$.

Θα δείξουμε τώρα ότι οι $\{N_t^+\}_{t \geq 0}$ και $\{N_t^-\}_{t \geq 0}$ είναι ανεξάρτητες. Θα ξεκινήσουμε δείχνοντας ότι οι τυχαίες μεταβλητές N_t^+ και N_t^- είναι ανεξάρτητες για κάθε $t \geq 0$.

$$\mathbb{P}[N_t^+ = m, N_t^- = n] = \mathbb{P}[N_t^+ = m, N_t = m+n] = \mathbb{P}\left[\sum_{i=1}^{m+n} X_i = m, N_t = m+n\right] \quad (8.13)$$

$$= \mathbb{P}\left[\sum_{i=1}^{m+n} X_i = m\right] \mathbb{P}[N_t = m+n]. \quad (8.14)$$

Στο τελευταίο βήμα χρησιμοποιήσαμε ότι η ακολουθία $\{X_i\}_{i \in \mathbb{N}}$ είναι ανεξάρτητη από τη διαδικασία αφίξεων $\{N_t\}_{t \geq 0}$. Από το Λήμμα 10 έχουμε ότι

$$\mathbb{P}[N_t = m+n] = e^{-\lambda t} \frac{(\lambda t)^{m+n}}{(m+n)!}.$$

Επιπλέον, εφόσον οι X_1, X_2, \dots είναι ανεξάρτητες τ.μ. Bernoulli με πιθανότητα επιτυχίας p έχουμε ότι $\sum_{i=1}^{m+n} X_i \sim \text{bin}(m+n, p)$ και

$$\mathbb{P}\left[\sum_{i=1}^{m+n} X_i = m\right] = \binom{n+m}{m} p^m (1-p)^n.$$

Αν αντικαταστήσουμε τις δύο παραπάνω ισότητες στην (8.13), έχουμε ότι

$$\begin{aligned}
\mathbb{P}[N_t^+ = m, N_t^- = n] &= e^{-\lambda t} \frac{(\lambda t)^{m+n}}{(m+n)!} \binom{n+m}{m} p^m (1-p)^n \\
&= e^{-\lambda p t} \frac{(\lambda p t)^m}{m!} e^{-\lambda(1-p)t} \frac{(\lambda(1-p)t)^n}{n!}. \quad (8.15)
\end{aligned}$$

Επομένως οι τυχαίες μεταβλητές N_t^+ και N_t^- είναι ανεξάρτητες για κάθε $t \geq 0$. Για να δείξουμε ότι οι διαδικασίες $\{N_t^+\}_{t \geq 0}$ και $\{N_t^-\}_{t \geq 0}$ είναι ανεξάρτητες, χρειάζεται να δείξουμε κάτι αντίστοιχο για τις από

κοινού κατανομές πεπερασμένης διάστασης. Συγκεκριμένα, αν $0 \leq t_1 \leq t_2 \leq \dots \leq t_k$ και οι $\{m_i\}_{1 \leq i \leq k}$ και $\{n_i\}_{1 \leq i \leq k}$ είναι αύξουσες ακολουθίες ακεραίων, θέλουμε να δείξουμε ότι

$$\mathbb{P}\left[\bigcap_{i=1}^k \{N_{t_i}^+ = m_i, N_{t_i}^- = n_i\}\right] = \mathbb{P}\left[\bigcap_{i=1}^k \{N_{t_i}^+ = m_i\}\right] \mathbb{P}\left[\bigcap_{i=1}^k \{N_{t_i}^- = n_i\}\right]. \quad (8.16)$$

Μπορούμε να χρησιμοποιήσουμε το γεγονός ότι οι διαδικασίες Poisson έχουν ανεξάρτητες και χρονικά ομοιογενείς προσauζήσεις, ώστε να αναγάγουμε αυτό το ερώτημα στο ερώτημα για τις μονοδιάστατες κατανομές της διαδικασίας που ήδη απαντήσαμε. Πράγματι, αν θέσουμε $t_0 = m_0 = n_0 = 0$, έχουμε

$$\begin{aligned} \mathbb{P}\left[\bigcap_{i=1}^k \{N_{t_i}^+ = m_i, N_{t_i}^- = n_i\}\right] &= \mathbb{P}\left[\bigcap_{i=1}^k \{N_{t_i}^+ - N_{t_{i-1}}^+ = m_i - m_{i-1}, N_{t_i}^- - N_{t_{i-1}}^- = n_i - n_{i-1}\}\right] \\ &= \mathbb{P}\left[\bigcap_{i=1}^k \left\{ \sum_{k=N_{t_{i-1}}^+ + 1}^{N_{t_i}^+} X_k = m_i - m_{i-1}, \sum_{k=N_{t_{i-1}}^- + 1}^{N_{t_i}^-} (1 - X_k) = n_i - n_{i-1} \right\}\right] \\ &= \prod_{i=1}^k \mathbb{P}\left[N_{t_i - t_{i-1}}^+ = m_i - m_{i-1}, N_{t_i - t_{i-1}}^- = n_i - n_{i-1}\right]. \end{aligned}$$

Από την (8.15) έχουμε τώρα ότι

$$\mathbb{P}\left[\bigcap_{i=1}^k \{N_{t_i}^+ = m_i, N_{t_i}^- = n_i\}\right] = \prod_{i=1}^k \mathbb{P}\left[N_{t_i - t_{i-1}}^+ = m_i - m_{i-1}\right] \mathbb{P}\left[N_{t_i - t_{i-1}}^- = n_i - n_{i-1}\right]$$

και αντιστρέφοντας τα βήματα του τελευταίου επιχειρήματος φτάνουμε στην (8.16). \square

Πόρισμα 15 Θεωρήστε δύο ανεξάρτητες διαδικασίες Poisson $\{X_t\}_{t \geq 0}$ και $\{Y_t\}_{t \geq 0}$ με ρυθμούς λ και μ αντίστοιχα. Αν $\{N_t\}_{t \geq 0}$ είναι το άθροισμά τους και $\{\tilde{X}_t\}_{t \geq 0}$ είναι μια $\left(\frac{\lambda}{\lambda + \mu}\right)$ -εκλέπτυνση της $\{N_t\}_{t \geq 0}$, τότε τα ζευγάρια διαδικασιών

$$(\{X_t\}_{t \geq 0}, \{Y_t\}_{t \geq 0}) \quad \text{και} \quad (\{\tilde{X}_t\}_{t \geq 0}, \{N_t - \tilde{X}_t\}_{t \geq 0})$$

έχουν την ίδια κατανομή. Ειδικότερα, κάθε άφιξη της $\{N_t\}_{t \geq 0}$ οφείλεται σε αύξηση της $\{X_t\}_{t \geq 0}$ με πιθανότητα $\frac{\lambda}{\lambda + \mu}$, ανεξάρτητα από τις άλλες αφίξεις και τη διαδικασία αφίξεων $\{N_t\}_{t \geq 0}$.

Πόρισμα 16 Με βάση τα Θεωρήματα 37 και 39 μπορούμε να προσομοιώσουμε ταυτόχρονα δύο ανεξάρτητες διαδικασίες Poisson $\{X_t\}_{t \geq 0}$ και $\{Y_t\}_{t \geq 0}$ με ρυθμούς λ και μ αντίστοιχα, κάνοντας εκλέπτυνση στο άθροισμά τους. Συγκεκριμένα, προκειμένου να προσομοιώσουμε τις δύο διαδικασίες στο διάστημα $[0, T]$ μπορούμε να κάνουμε τα ακόλουθα βήματα.

1. Προσομοιώνουμε το συνολικό πλήθος αφίξεων $N_T = X_T + Y_T \sim Po((\lambda + \mu)T)$.
2. Αν $N_T = n$ προσομοιώνουμε τους χρόνους των n αφίξεων επιλέγοντας n ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με ομοιόμορφη κατανομή στο $[0, T]$.
3. Αποδίδουμε ανεξάρτητα κάθε μια από τις n αφίξεις είτε στη $\{X_t\}_{t \geq 0}$ (με πιθανότητα $p_X = \frac{\lambda}{\lambda + \mu}$) είτε στην $\{Y_t\}_{t \geq 0}$ (με πιθανότητα $p_Y = \frac{\mu}{\lambda + \mu}$).

Παράδειγμα 55 Οι άνθρωποι που προσέρχονται σε ένα αναψυκτήριο ένα καλοκαιρινό μεσημέρι φτάνουν ως μια διαδικασία Poisson με ρυθμό 3/min. Καθένας από αυτούς επιλέγει είτε χυμό πορτοκάλι είτε χυμό μήλο με πιθανότητα 5/6 και 1/6 αντίστοιχα και ανεξάρτητα ο ένας από τον άλλον. Αν κάποια στιγμή υπάρχουν αποθέματα πορτοκαλιών για 240 χυμούς και μήλων για 39 χυμούς, ποια είναι η πιθανότητα να εξυπηρετηθούν όλοι οι άνθρωποι που θα φτάσουν μέσα στην επόμενη μιάμιση ώρα;

Σύμφωνα με το προηγούμενο Θεώρημα, οι παραγγελίες για χυμό πορτοκαλιού φτάνουν ως μια διαδικασία Poisson με ρυθμό $3/\text{min} \times 5/6 = 2,5/\text{min}$. Επομένως, από το Θεώρημα 10, οι παραγγελίες για χυμό πορτοκαλιού την επόμενη ώρα είναι μια τυχαία μεταβλητή N , που ακολουθεί κατανομή Poisson με παράμετρο $2,5/\text{min} \times 90\text{min} = 225$. Αντίστοιχα, οι παραγγελίες για χυμό μήλου φτάνουν ως μια διαδικασία Poisson με ρυθμό $3/\text{min} \times 1/6 = 1/2\text{min}$. Επομένως, οι παραγγελίες για χυμό μήλου την επόμενη ώρα, είναι μια τυχαία μεταβλητή, M , που ακολουθεί κατανομή Poisson, με παράμετρο $1/2\text{min} \times 90\text{min} = 45$. Μάλιστα, εφόσον οι δύο διαδικασίες είναι ανεξάρτητες, θα είναι ανεξάρτητες και οι τυχαίες μεταβλητές M και N . Εμείς θέλουμε να υπολογίσουμε την πιθανότητα

$$\mathbb{P}[\{N \leq 240\} \cap \{M \leq 39\}] = \mathbb{P}[N \leq 240] \mathbb{P}[M \leq 39].$$

Προκειμένου να ολοκληρώσουμε τον υπολογισμό, μπορούμε να υπολογίσουμε αυτές τις ποσότητες είτε με τη βοήθεια κάποιου στατιστικού πακέτου, όπως η R, είτε να βρούμε προσεγγιστικά την απάντηση με τη βοήθεια του Κεντρικού Οριακού Θεωρήματος. Συγκεκριμένα, αν η τυχαία μεταβλητή X ακολουθεί κατανομή $Po(\lambda)$, τότε

$$\mathbb{P}\left[\frac{X - \lambda}{\sqrt{\lambda}} \leq x\right] \rightarrow \Phi(x), \quad \text{καθώς } \lambda \rightarrow \infty,$$

όπου η συνάρτηση Φ στο δεξί μέλος είναι η συνάρτηση κατανομής πιθανότητας της τυπικής κανονικής κατανομής. Στην πράξη, η προσέγγιση είναι αρκετά καλή για $\lambda \geq 30$, οπότε μπορούμε να εκτιμήσουμε

$$\mathbb{P}[N \leq 240] = \mathbb{P}\left[\frac{N - 225}{\sqrt{225}} \leq \frac{240 - 225}{\sqrt{225}}\right] \simeq \Phi(1) \simeq 0,8414.$$

Ομοίως,

$$\mathbb{P}[M \leq 39] = \mathbb{P}\left[\frac{M - 45}{\sqrt{45}} \leq \frac{39 - 45}{\sqrt{45}}\right] \simeq \Phi\left(-\frac{2}{\sqrt{5}}\right) = 0,1856.$$

Επομένως η ζητούμενη πιθανότητα είναι περίπου 0,1561. Η απάντηση που θα παίρναμε με τη βοήθεια του υπολογιστή, χωρίς την προσέγγιση του Κεντρικού Οριακού Θεωρήματος, θα ήταν 0,1769. \square

Παράδειγμα 56 Αν $\{X_t\}_{t \geq 0}$ και $\{Y_t\}_{t \geq 0}$ είναι δύο ανεξάρτητες διαδικασίες Poisson με ρυθμούς λ και μ αντίστοιχα, ποια κατανομή ακολουθεί το πλήθος Z των αφίξεων της $\{X_t\}_{t \geq 0}$ ανάμεσα σε δύο διαδοχικές αφίξεις της $\{Y_t\}_{t \geq 0}$;

Η διαδικασία $\{N_t\}_{t \geq 0}$ με $N_t = X_t + Y_t$ είναι μια διαδικασία Poisson με ρυθμό $\lambda + \mu$. Το ενδεχόμενο $\{Z \geq k\}$ είναι το ενδεχόμενο οι πρώτες k αφίξεις της $\{N_t\}_{t \geq 0}$ μετά από μια άφιξη της $\{Y_t\}_{t \geq 0}$ να οφείλονται σε αυξήσεις της $\{X_t\}_{t \geq 0}$. Επομένως

$$\mathbb{P}[Z \geq k] = \left(\frac{\lambda}{\lambda + \mu}\right)^k, \quad k = 0, 1, \dots$$

Επομένως για κάθε $k \in \mathbb{N}$

$$\mathbb{P}[Z + 1 = k] = \mathbb{P}[Z \geq k - 1] - \mathbb{P}[Z \geq k] = \left(\frac{\lambda}{\lambda + \mu}\right)^{k-1} \frac{\mu}{\lambda + \mu},$$

δηλαδή η $Z + 1$ ακολουθεί γεωμετρική κατανομή με παράμετρο $\frac{\mu}{\lambda + \mu}$. \square

8.4 Ασκήσεις

Άσκηση 120 Σωματίδια προσπίπτουν σ' έναν ανιχνευτή σύμφωνα με μια διαδικασία Poisson. Αν στα πρώτα δύο λεπτά λειτουργίας του ανιχνευτή προσέπεσαν 14 σωματίδια, ποια είναι η πιθανότητα τα 5 από αυτά να προσέπεσαν κατά το πρώτο λεπτό;

Άσκηση 121 Σε μια διαδικασία Poisson με ρυθμό αφίξεων μ , αν S_N είναι ο χρόνος της N -οστής άφιξης, υπολογίστε τα όρια

$$\mathbb{P}\left[\lim_{N \rightarrow \infty} \frac{S_N}{N} = \frac{1}{\mu}\right] \quad \text{και} \quad \lim_{N \rightarrow \infty} \mathbb{P}\left[S_N \leq \frac{N}{\mu}\right].$$

Άσκηση 122 Σε μια διαδικασία Poisson με ρυθμό αφίξεων λ , βρείτε την κατανομή που ακολουθεί το πλήθος των αφίξεων στο διάστημα $(0, s]$ δεδομένου ότι έχουμε N αφίξεις στο διάστημα $(0, t]$ με $t \geq s$.

Άσκηση 123 Οι αφίξεις φοιτητών στη βιβλιοθήκη του ΕΜΠ μεταξύ 8πμ και 2μμ είναι μια διαδικασία Poisson με ρυθμό αφίξεων $1/5\text{min}$. Οι αφίξεις καθηγητών την ίδια περίοδο είναι μια διαδικασία Poisson με ρυθμό αφίξεων $1/30\text{min}$. Αν στο διάστημα 11πμ και 12μμ ήταν στη βιβλιοθήκη 10 άτομα, ποια είναι η πιθανότητα να μπήκαν το πολύ δύο καθηγητές; Η πιθανότητα να δανειστεί βιβλίο κάποιος καθηγητής που επισκέπτεται τη βιβλιοθήκη είναι $4/5$ ενώ η πιθανότητα να δανειστεί βιβλίο ένας φοιτητής που επισκέπτεται τη βιβλιοθήκη είναι $1/6$. Υποθέστε ότι δεν μεσολαβεί χρόνος ανάμεσα στην άφιξη ενός ατόμου που θέλει να δανειστεί βιβλίο και στον δανεισμό του βιβλίου. Πώς θα περιγράφατε τη διαδικασία που μετράει τους δανεισμούς βιβλίων κατά την παραπάνω περίοδο; Ποια είναι η πιθανότητα το πρώτο βιβλίο της ημέρας να το δανειστεί ένας φοιτητής;

Άσκηση 124 Σ' ένα δισκάδικο εξυπηρετούν δύο υπάλληλοι. Ο χρόνος εξυπηρέτησης ενός πελάτη ακολουθεί εκθετική κατανομή με ρυθμό $1/2\text{min}$. Μπαίνετε στο δισκάδικο και παρατηρείτε ότι και οι δύο υπάλληλοι εξυπηρετούν, ενώ υπάρχει ακόμα ένας πελάτης πριν από σας που περιμένει να εξυπηρετηθεί. Ποια είναι η κατανομή του χρόνου που θα σας πάρει μέχρι να εξυπηρετηθείτε; Ποια είναι η πιθανότητα να τελειώσετε πριν από τον πελάτη που περιμένει; Ποια είναι η πιθανότητα να τελειώσετε πριν από κάποιον πελάτη που εξυπηρετείτο όταν μπήκατε;

Άσκηση 125 Στις αναμετρήσεις των ποδοσφαιρικών ομάδων A και B τα τέρματα της ομάδας A σημειώνονται σύμφωνα με μια διαδικασία Poisson με ρυθμό αφίξεων $1/45\text{min}$, ενώ εκείνα της ομάδας B σύμφωνα με μια ανεξάρτητη από την προηγούμενη διαδικασία Poisson με ρυθμό αφίξεων $1/60\text{min}$. Ποια είναι η πιθανότητα στο ημίχρονο να προηγείται η ομάδα B με 0-1 αλλά στο τέλος να κερδίσει η ομάδα A με 3-1; Ποια είναι η πιθανότητα να κερδίσει η ομάδα A με 2-1, δεδομένου ότι το σκορ ημιχρόνου είναι 0-1;

Άσκηση 126 Σωματίδια τύπου A και τύπου B φτάνουν σ' έναν ανιχνευτή σύμφωνα με δύο ανεξάρτητες διαδικασίες Poisson με ρυθμούς αφίξεων λ και μ αντίστοιχα. Ποια κατανομή ακολουθεί το πλήθος των σωματιδίων τύπου A που έχουν ανιχνευτεί μέχρι την ανίχνευση του πρώτου σωματιδίου τύπου B; Ποια κατανομή ακολουθεί το πλήθος των σωματιδίων τύπου A που έχουν ανιχνευτεί μέχρι την ανίχνευση του τρίτου σωματιδίου τύπου B;

Άσκηση 127 Σ' έναν μελλοντικό κόσμο διαστημικά λεωφορεία αναχωρούν από τη Γη για τη Σελήνη ως μια διαδικασία Poisson $\{D_t : t \geq 0\}$ με ρυθμό $\lambda = 4$ ανά ημέρα. Οι ταξιδιώτες προσέρχονται ως μια διαδικασία Poisson $\{A_t : t \geq 0\}$ με ρυθμό $\mu = 50$ ανά ώρα, ανεξάρτητη από την $\{D_t : t \geq 0\}$ και φεύγουν με το πρώτο διαθέσιμο διαστημικό λεωφορείο. Αυτή τη στιγμή η ώρα είναι 12 το μεσημέρι.

- Ποια είναι η πιθανότητα το μεθεπόμενο λεωφορείο να αναχωρήσει πριν τα μεσάνυχτα;
- Ποια είναι προσεγγιστικά η πιθανότητα να έχουν φτάσει τουλάχιστον 640 επιβάτες μέχρι τα μεσάνυχτα;

β) Αν N είναι το πλήθος των επιβατών σε μία πτήση υπολογίστε την πιθανότητα $\mathbb{P}[N = k]$, για $k = 0, 1, 2, \dots$

στ) Αν Λ είναι το πλήθος των λεωφορειών που θα έχουν αναχωρήσει μέχρι τη στιγμή που θα φτάσει ο χιλιοστός ταξιδιώτης (μετρώντας από τώρα και μετά) υπολογίστε την πιθανότητα $\mathbb{P}[\Lambda = k]$, για $k = 0, 1, 2, \dots$

Άσκηση 128 Αυτοκίνητα και μοτοσυκλέτες προσέρχονται σε έναν σταθμό διοδίων ως δύο ανεξάρτητες διαδικασίες Poisson με ρυθμούς 3/λεπτό και 1/λεπτό αντίστοιχα.

α) Υπολογίστε την πιθανότητα το επόμενο λεπτό να περάσουν τουλάχιστον δύο οχήματα από τον σταθμό.

β) Υπολογίστε την πιθανότητα το επόμενο λεπτό να περάσουν μόνο μοτοσυκλέτες από τον σταθμό.

γ) Αν το προηγούμενο λεπτό πέρασαν από τον σταθμό δύο οχήματα, υπολογίστε την πιθανότητα οι αφίξεις τους να απείχαν χρονικά λιγότερο από 1/2 λεπτό.

δ) Υπολογίστε την πιθανότητα το επόμενο λεπτό να μην περάσουν διαδοχικά οχήματα του ίδιου τύπου.

Άσκηση 129 Σε μια πειραματική διάταξη μια δέσμη laser περνά από ένα διαχωριστή που τη χωρίζει σε δύο δέσμες A,B. Η προσπίπτουσα δέσμη αποτελείται από φωτόνια που φτάνουν στον διαχωριστή ως μια διαδικασία Poisson με ρυθμό λ . Κάθε φωτόνιο, κατευθύνεται είτε στη δέσμη A είτε στη δέσμη B με πιθανότητα 1/2, ανεξάρτητα για κάθε φωτόνιο. Οι δύο δέσμες A, B κατευθύνονται στη συνέχεια σε δύο φωτοδιόδους. Για κάθε φωτόνιο που προσπίπτει στη φωτοδίοδο απελευθερώνεται ένα ηλεκτρόνιο. Τα ηλεκτρόνια που απελευθερώνονται δημιουργούν τα ρεύματα I_1 και I_2 αντίστοιχα. Στη συνέχεια τα δύο αυτά ρεύματα αφαιρούνται, οπότε προκύπτει ένα ηλεκτρικό ρεύμα $I = I_1 - I_2$, που οδηγείται σε μια συσκευή μέτρησης. Ποια είναι η μέση τιμή και ποια είναι η διασπορά του αριθμού των ηλεκτρονίων που φτάνει στη συσκευή μέτρησης σε χρόνο Δt ;

Άσκηση 130 Αρσενικά και θηλυκά μπαρμπούνια πέφτουν στα δίχτυα ενός ψαρά σύμφωνα με δύο ανεξάρτητες διαδικασίες Poisson με ρυθμούς $\alpha = 1/20\text{min}$ και $\theta = 1/30\text{min}$ αντίστοιχα.

α) Αν ο ψαράς αφήσει τα δίχτυα για 2 ώρες ποια είναι η κατανομή του πλήθους των μπαρμπουνιών που θα πιάσει;

β) Αν σε 2 ώρες έπιασε 10 μπαρμπούνια ποια είναι η πιθανότητα να μην έπιασε κανένα στο διάστημα 30min έως 90min;

γ) Αν σε 2 ώρες έπιασε 10 μπαρμπούνια ποια είναι η πιθανότητα να μην έπιασε κανένα θηλυκό μεταξύ 30min και 90min;

δ) Αν T είναι η πρώτη χρονική στιγμή (από την ώρα που έριξε τα δίχτυα) που έχουν πιαστεί τουλάχιστον 2 αρσενικά και τουλάχιστον 2 θηλυκά μπαρμπούνια, ποια κατανομή ακολουθεί η τυχαία μεταβλητή T ;

Άσκηση 131 Σε μια διαδικασία Poisson με ρυθμό αφίξεων λ , ποια είναι η κατανομή του διαστήματος ανάμεσα σε διαδοχικές αφίξεις που περιέχει το $t > 0$; Γιατί η κατανομή ΔΕΝ είναι εκθετική με ρυθμό λ ; Ποιο είναι το όριο αυτής της κατανομής όταν $t \rightarrow \infty$;

Βιβλιογραφία

- [1] Bertsekas, Dimitri P. and Tsitsiklis, John N. *Introduction to probability*. Athena Scientific, Belmont, 2002.
- [2] Brémaud, Pierre. *Markov chains : Gibbs fields, Monte Carlo simulation and queues*. Texts in applied mathematics. Springer, New York, NY, 1999.
- [3] Doyle, Peter G. *Random Walks and Electrical Networks*. Carus Mathematical Monographs. John Wiley & sons, New York, NY, 1985.
- [4] Häggström, Olle. *Finite Markov chains and algorithmic applications*. London Mathematical Society student texts. Cambridge University Press, Cambridge, UK, 2002.
- [5] Krauth, Werner. *Statistical Mechanics: Algorithms and Computations*. Oxford Master Series in Physics. Oxford University Press, Oxford, UK, 2006.
- [6] Levin, David Asher and Peres, Yuval and Wilmer, Elizabeth Lee. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009.
- [7] Norris, James R. *Markov chains*. Cambridge series in statistical and probabilistic mathematics, Cambridge, UK. Cambridge University Press, 1998.
- [8] Varadhan, S. R. S. *Probability theory*. Courant Lecture Notes in Mathematics. American Mathematical Society, Providence, RI, 2001.
- [9] Williams, David. *Probability With Martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, UK, 1991.
- [10] Ουρανία Χρυσ αφίνου. *Εισαγωγή στις Στοχαστικές Ανελιξεις, β' έκδοση*. Σοφία, Θεσσαλονίκη, 2012.

Ευρετήριο

- ακριβής ισορροπία, 85
αλγόριθμος Metropolis-Hastings, 118
αλγόριθμος Pagerank, 112
ανάλυση πρώτου βήματος, 34
αναλλοίωτη κατανομή, 73
ανοιχτή κλάση, 18
αντιστρέψιμη διαδικασία, 85
απεριοδική, 92
αρχή του Rayleigh, 116
γεννήτορας, 34
γνησίως επαναληπτική κατάσταση, 76
δειγματοληψία κατά σημαντικότητα, 110
δεσμευμένη μέση τιμή, 52
διαδικασία Poisson, 129
εκλέπτυνση, 139
ελάττωση διασποράς, 110
επαναληπτική κατάσταση, 23
επαναληπτική κλάση, 26
εργοδικό θεώρημα, 103, 105
εξισώσεις Chapman-Kolmogorov, 16
θεώρημα επιλεκτικής διακοπής, 60
θεώρημα συνέπειας, 6
ιδιότητα της απώλειας μνήμης, 129
ισχυρή μαρκοβιανή ιδιότητα, 20, 22
κίνηση Brown, 4
κατανομές πεπερασμένης διάστασης, 4
κατανομή ισορροπίας, 73
κατανομή Gibbs, 119
κλάσεις επικοινωνίας, 18
κλειστή κλάση, 18
λήμμα του Scheffé, 72
μέτρο σύζευξης, 96
μαρκοβιανή αλυσίδα, 6
μαρκοβιανή ιδιότητα, 10
μετασχηματισμός martingale, 59
μη υποβιβάσιμη αλυσίδα, 18
μοντέλο του Ising, 120
παροδική κατάσταση, 23
παροδική κλάση, 26
περίοδος κατάστασης, 92
πίνακας πιθανοτήτων μετάβασης, 7
πιθανότητες απορρόφησης, 33
πιθανότητες μετάβασης, 7
πιθανότητες μετάβασης ανώτερης τάξης, 16
πρόβλημα του πλανόδιου πωλητή, 126
πρόγραμμα ψύξης, 122
προσαρμοσμένη διαδικασία, 58
προσβάσιμη κατάσταση, 18
προσομοιωμένη ανόπτηση, 121
σύγκλιση κατανομών, 72
σύζευξη, 95
σταθερά του Kemeny, 88
στοχαστική διαδικασία, 1
στοχαστικός πίνακας, 7
συνάρτηση δυναμικού, 34
συνάρτηση επιμερισμού, 119
τύπος του Stirling, 28
τυχαίος περίπατος, 27, 30
τυχαίος περίπατος σε γράφο, 86
τυχαίος περίπατος σε d διαστάσεις, 66
χρόνος διακοπής, 21
χρόνος πρώτης άφιξης, 21
χωρητικότητα, 117
χώρος καταστάσεων, 1
Markov chain Monte Carlo, 109
Monte Carlo, 31
martingale, 58

Γλωσσάρι τεχνικών όρων

ακριβής ισορροπία	detailed balance
αναλλοίωτη κατανομή	invariant distribution
απεριοδική	aperiodic
γεννήτορας	generator
γνησίως επαναληπτική	positive recurrent
δειγματοληψία κατά σημαντικότητα	importance sampling
δεσμευμένη μέση τιμή	conditional expectation
δυναμικό	potential
ελάττωση διασποράς	variance reduction
εκλέπτυνση	thinning
επαναληπτική	recurrent
εργοδικό θεώρημα	ergodic theorem
θεώρημα επιλεκτικής διακοπής	optional stopping theorem
ισχυρή μαρκοβιανή ιδιότητα	strong Markov property
κατανομές πεπερασμένη διάστασης	finite dimensional distributions
κλάσεις επικοινωνίας	communication classes
μαρκοβιανή αλυσίδα	Markov chain
μη υποβιβάσιμη	irreducible
παροδική	transient
πιθανότητα απορρόφησης	absorption probability
πιθανότητες μετάβασης	transition probabilities
πίνακας πιθανοτήτων μετάβασης	transition matrix
πρόγραμμα ψύξης	cooling schedule
προσαρμοσμένη διαδικασία	adapted process
προσομοιωμένη ανόπτηση	simulated annealing
στοχαστική διαδικασία	stochastic process
συνάρτηση επιμερισμού	partition function
σύζευξη	coupling
τυχαίος περίπατος	random walk
χρονικά αντιστρέψιμη διαδικασία	time reversible process
χρόνος διακοπής	stopping time
χρόνος πρώτης άφιξης	hitting time
χωρητικότητα	capacity
χώρος καταστάσεων	state space

THE END