ΚΕΦΑΛΑΙΟ VI. ΑΣΥΜΠΤΩΤΙΚΑ ΘΕΩΡΗΜΑΤΑ

1 Εισαγωγή

Στο κεφάλαιο αυτό θα μελετήσουμε την ασυμπτωτική συμπεριφορά μαρκοβιανών αλυσίδων, πώς δηλαδή 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 την 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 [**;**] και [**;**].

2 Περιοδικότητα

Ας θεωρήσουμε μια αλυσίδα που κινείται στις κορυφές ενός τετραγώνου ΑΒΓΔ. Ξεκινώντας από την κορυφή Α, σε κάθε βήμα μεταβαίνει σε μια από τις δύο γειτονικές της κορυφές με πιθναότητα $1/2$. 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 $\pi_n \not\to \pi$. Πράγματι, εφόσον η αλυσίδα ξεκινά από το Α, θα βρίσκεται οπωσδήποτε στο Β ή στο Δ μετά από περιττό αριθμό βημάτων, και στο Α ή στο Γ μετά από άρτιο αριθμό βημάτων. Έτσι, π.χ.

$$\pi_{2n+1}(A) = \mathbb{P}\big[X_{2n+1} = A \,|\, X_0 = A\big] = 0 \not\to 1/4.$$

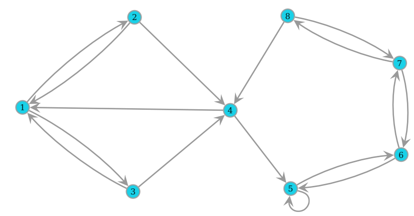
Βλέπουμε λοιπόν ότι στο παράδειγμά μας υπάρχουν υποσύνολα του χώρου καταστάσεων που είναι προσβάσιμα μόνο σε συγκεκριμένες χρονικές στιγμές. Αυτή η ιδιότητα αποτελεί κατά κάποιον τρόπο μια παθολογία που εμποδίζει την σύγκλιση της κατανομής της αλυσίδας π_n, και σε αυτήν την παράγραφο θα προσπαθήσουμε να την κατανοήσουμε.

Ορισμός: Για μια κατάσταση $x \in \mathbb{X}$ ορίζουμε το σύνολο των δυνατών χρόνων επιστροφής στο x

$$R(x) = \{n \in \mathbb{N} : p^n(x,x) > 0\}, \quad \text{όπου θυμίζουμε } p^{(n)}(x,x) = \mathbb{P}\big[X_n = x \,|\, X_0 = x\big].$$

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 1 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 $R(1) = R(2) = R(3) = R(6) = \{2,3,\ldots\}$, $R(4) = \{3,5,6,7,8,\ldots\}$, $R(5) = \mathbb{N}$, $R(7) = R(8) = \{2,4,5,6,\ldots\}$. 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

$$m, n \in R(x) \Rightarrow m + n \in R(x).$$

Διαισθητικά 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 $m, n \in \mathbb{R}(x)$. Για τέτοια σύνολα έχουμε το ακόλουθο λήμμα.

Λήμμα 1 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 $\{n_0 d, (n_0+1)d, (n_0+2)d, \ldots\} \subset R$. 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 $d\mathbb{N} = \{d, 2d, 3d, \ldots\}$. 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 $m_1, \ldots, m_k \in R$ και $p_1, p_2, \ldots, p_k \in \mathbb{Z} \setminus \{0\}$ τέτοια ώστε

$$\sum_{i=1}^{k} p_i m_i = d. \tag{1}$$

Έστω $I \subset \{1, 2, \ldots, k\}$ το σύνολο των δεικτών i για τους οποίους $p_i > 0$ και $I' = \{1, 2, \ldots, k\} \setminus I$ το σύνολο των δεικτών i για τους οποίους $p_i < 0$. Προφανώς, το I δεν μπορεί να είναι κενό αφού το δεξί μέλος της (1) είναι φυσικός αριθμός. Ούτε το I' μπορεί να είναι κενό όμως, αφού κάθε προσθετέος στο αριστερό μέλος της (1) είναι μεγαλύτερος ή ίσος από $2d$ κατ΄ απόλυτη τιμή στην περίπτωση ($d \notin R$) που εξετάζουμε. Μπορούμε επομένως να ξαναγράψουμε την (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 ακέραιας διαίρεσής του N με το n έχουμε

$$N = qn + v \qquad \text{για κάποια } q \geq \frac{n-1}{d},\ v \in \{0, 1, 2, \ldots, n-1\}. \tag{2}$$

Έχουμε τώρα

$$Nd = qnd + vd = qnd + v(m-n) = vm + (qd-v)n.$$

Οι v και $qd - v$ είναι μη αρνητικοί ακέραιοι λόγω της (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 1 για να δείξουμε ότι η περίοδος είναι χαρακτηριστικό κλάσης.

Θεώρημα 1 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 $\ell \in \mathbb{R}(x)$ 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 1, υπάρχει κάποιο $q \in \mathbb{N}$ τέτοιο ώστε $qd(x), (q+1)d(x) \in \mathbb{R}(x)$. 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 1, 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 περίοδός όλων των καταστάσεών της είναι 1. Το ακόλουθο πόρισμα είναι συχνά χρήσιμο γιατί δίνει μια ικανή συνθήκη ώστε να είναι μια αλυσίδα είναι απεριοδική που μπορεί να ελεγχθεί με απλή επισκόπηση.

Πόρισμα 1 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 περίδο με την x, δηλαδή 1. $\square$

Το ακόλουθο λήμμα θα μας φανεί επίσης χρήσιμο στις επόμενες παραγράφους.

Λήμμα 2 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 1 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

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

$\square$

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 καανομή 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 κατάστασεων 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,

$$\mathbb{P}\big[X_n \in C_j \,|\, X_0 \in C_j\big] \neq 0 \Leftrightarrow d/n.$$

Πιο συγκεκριμένα έχουμε το ακόλουθο θεώρημα, του οποίου η απόδειξη αφήνεται σαν άσκηση.

Θεώρημα 2 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 δειγματοληπτούμε την $\{X_n\}_n$ κάθε d βήματα) είναι υποβιβάσιμη στις d απεριοδικές κλάσεις

$$C_v = \{y \in \mathbb{X} : p^{(dn+v)}(x,y) > 0 \text{ για κάποιο } n \in \mathbb{N}_0\}, \quad v \in 0,1,2,\dots d-1.$$

Επιπλέον, 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

$$\pi_v(y) = \begin{cases} d\,\pi(y) & \text{αν } y \in C_v \\ 0 & \text{αν } y \notin C_v. \end{cases} \tag{3}$$

Στην παράγραφο 4 θα δούμε ότι αν μια μη υποβιβάσιμη και γνησίως επαναληπτική αλυσίδα είναι απεριοδική, 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 απλά να κατασκευάσουμε μια μη υποβιβάσιμη, γνησίως επαναληπτική και απεριοδική αλυσίδα που έχει αναλλοίωτη κατανομή την π και να αφήσουμε αυτήν 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 8.

Για την απόδειξη της σύγκλισης $\pi_n \to \pi$ θα χρησιμοποιήσουμε μια πολύ ισχυρή πιθανοθεωρητική τεχνική, την *σύζευξη* (coupling). Λόγω της χρησιμότητας αυτής της τεχνικής θα αφιερώσουμε την επόμενη παράγραφο στο να την περιγράψουμε, μέσα από δύο παραδείγματα.

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

$$\mathbb{P}\big[A_1 \times \Omega_2\big] = \mathbb{P}_1\big[A_1\big] \quad \text{και} \quad \mathbb{P}\big[\Omega_1 \times A_2\big] = \mathbb{P}_2\big[A_2\big], \tag{4}$$

για κάθε $A_i \subset \Omega_1$ 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,

$$\mathbb{P}\big[\tilde{X} \in C\big] = \mathbb{P}\big[\{X \in C\} \times \Omega_2\big] = \mathbb{P}_1\big[X \in C\big],$$

και αντίστοιχα για τις $\tilde{Y}$ και Y. Ένα μέτρο πιθανότητας $\mathbb{P}$ που ικανοποιεί τις (4) ονομάζεται μέτρο σύζευξης των $\mathbb{P}_1$ και $\mathbb{P}_2$.

Παρατηρήστε ότι ενώ οι σχέσεις (4) που επιβάλλουμε στο $\mathbb{P}$ καθορίζουν την κατανομή καθεμιάς από τις $\tilde{X}$ και $\tilde{Y}$, δεν μας δίνουν καμμιά 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 2 Έστω X τυχαία μεταβλητή με τιμές στο $\mathbb{R}$. Αν οι συναρτήσεις $f, g : \mathbb{R} \to \mathbb{R}$ είναι και οι δύο αύξουσες ή και οι δύο φθίνουσες, τότε

$$\mathbb{E}\big[f(X)g(X)\big] \geq \mathbb{E}\big[f(X)\big]\mathbb{E}\big[g(X)\big],$$

οποτεδήποτε οι παραπάνω αναμενόμενες τιμές είναι καλά ορισμένες.

Σε αυτό το σημείο αξίζει να δοκιμάσετε να αποδείξετε τον ισχυρισμό του παραδειγματος 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 την X. Παρατηρήστε ότι εφόσον οι f, g έχουν το ίδιο είδος μονοτονίας τότε για κάθε $x, y \in \mathbb{R}$ έχουμε

$$\big(f(x) - f(y)\big)\big(g(x) - g(y)\big) \geq 0,$$

και επομένως η τυχαία μεταβλητή $\big(f(X) - f(Y)\big)\big(g(X) - g(Y)\big)$ παίρνει μόνο μη αρνητικές τιμές. Ειδικώτερα,

$$\begin{aligned} 0 &\leq \mathbb{E}\big[\big(f(X) - f(Y)\big)\big(g(X) - g(Y)\big)\big] \\ &= \mathbb{E}\big[f(X)g(X)\big] + \mathbb{E}\big[f(Y)g(Y)\big] - \mathbb{E}\big[f(X)g(Y)\big] - \mathbb{E}\big[f(Y)g(X)\big] \\ &= \mathbb{E}\big[f(X)g(X)\big] + \mathbb{E}\big[f(Y)g(Y)\big] - \mathbb{E}\big[f(X)\big]\mathbb{E}\big[g(Y)\big] - \mathbb{E}\big[f(Y)\big]\mathbb{E}\big[g(X)\big] \\ &= 2\big(\mathbb{E}\big[f(X)g(X)\big] - \mathbb{E}\big[f(X)\big]\mathbb{E}\big[g(X)\big]\big), \end{aligned}$$

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 την σύζευξη στο παραπάνω επιχείρημα; Λέγοντας 'θεωρήστε τώρα μια άλλη τυχαία μεταβλητή Y με την ίδια κατανομή όπως η X, και ανεξάρτητη από την 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

$$\mathbb{P}\big[A_1 \times A_2\big] = \mathbb{P}_1\big[A_1\big]\mathbb{P}_2\big[A_2\big],$$

για κάθε $A_i \subset \Omega_i$ στο πεδίο ορισμού $\mathcal{F}_i$ του P_i. Το $\mathbb{P}$ είναι μέτρο σύζευξης αφού $\mathbb{P}_i\big[\Omega_i\big] = 1$, 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

$$\mathbb{P}\big[\tilde{X} \in C_1, \tilde{Y} \in C_2\big] = \mathbb{P}\big[\{X \in C_1\} \times \{Y \in C_2\}\big] = \mathbb{P}_1\big[X \in C_1\big]\mathbb{P}_2\big[Y \in C_2\big] = \mathbb{P}\big[\tilde{X} \in C_1\big]\mathbb{P}\big[\tilde{Y} \in C_2\big].$$

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 για χάρη της απλότητας του συμβολισμού 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. Στην συνέχεια χρησιμοποιήσαμε την ανεξαρτησία τους στην πορεία της απόδειξης.

Παράδειγμα 3 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 την 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. Παρόλα 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 την διαγράψει, και στην 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 στην δεύτερη με έναν συνεκτικό γράφο. Θα χρησιμοποιήσουμε την ιδέα της συζευξης 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 στην 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 διαγράφουμε κάθε ακμή ανεξάρτητα με πιθανότητα p, εφοδιάζουμε τον Ω με το μέτρο γινόμενο $\mathbb{P}_p$, κάτω από το οποίο οι τυχαίες μεταβητές $\{X_e\}_{e \in E}$ 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 $G(\omega)$, αλλά όχι από τον $G(\omega')$. Πράγματι,

$$\mathbb{P}_{p_1} \times \mathbb{P}_{p_2}\big[\omega_e = 0, \omega'_e = 1\big] = p_1(1 - p_2) > 0.$$

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

$$\mathbb{P}\big[\tilde{Y}_e \le \tilde{X}_e \text{ για κάθε } e \in E\big] = 1. \tag{5}$$

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 ανεξάρητες τυχαίες μεταβλητές με κατανομή Bernoulli$(1 - p_1)$ και οι $\{\tilde{Y}_e\}_{e \in E}$ να είναι ανεξάρητες 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

$$\mathbb{Q}\big[\{(0,0)\}\big] = p_1, \quad \mathbb{Q}\big[\{(1,0)\}\big] = p_2 - p_1, \quad \mathbb{Q}\big[\{(1,1)\}\big] = 1 - p_2, \quad \text{και } \mathbb{Q}\big[\{(0,1)\}\big] = 0.$$

Στην συνέχεια ορίζουμε για $(\omega,\omega')\in\Omega\times\Omega$

$$\mathbb{P}\big[\{(\omega,\omega')\}\big]=\prod_{e\in E}\mathbb{Q}\big[\{(\omega_e,\omega'_e)\}\big]$$

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 $\mathbb{P}\big[\tilde X_e=1\big]=\mathbb{Q}\big[\{(1,0),(1,1)\}\big]=1-p_1$ έχουμε ότι οι $\{\tilde X_e\}_{e\in E}$ είναι ανεξάρτητες ισόνομες τυχαίες μεταβλητές με κατανομή Bernoulli$(1-p_1)$. Αντίστοιχα, εφόσον $\mathbb{P}\big[\tilde Y_e=1\big]=\mathbb{Q}\big[\{(0,1),(1,1)\}\big]=1-p_2$, οι $\{\tilde Y_e\}_{e\in E}$ είναι ανεξάρτητες ισόνομες τυχαίες μεταβλητές με κατανομή 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

$$\mathbb{P}\big[\tilde Y_e\le\tilde X_e\big]=\mathbb{Q}\big[\{(0,0),(1,0),(1,1)\}\big]=p_1+(p_2-p_1)+(1-p_2)=1.$$

Επιπλέον, λόγω της ανεξαρτησίας των $\{(\tilde X_e,\tilde Y_e)\}_{e\in E}$ έχουμε ότι

$$\mathbb{P}\big[\tilde Y_e\le\tilde X_e \text{ για κάθε } e\in E\big]=\mathbb{P}\Big[\bigcap_{e\in E}\{\tilde Y_e\le\tilde X_e\}\Big]=\prod_{e\in E}\mathbb{P}\big[\tilde Y_e\le\tilde X_e\big]=1,$$

οπότε το μέτρο $\mathbb{P}$ πράγματι ικανοποιεί την (5). Θα δείξουμε τώρα ότι αν υπάρχει μέτρο σύζευξης $\mathbb{P}$ των $\mathbb{P}_{p_1}$ και $\mathbb{P}_{p_2}$ που ικανοποιεί την (5) τότε $\mathbb{P}_{p_1}\big[C\big]\ge\mathbb{P}_{p_2}\big[C\big]$.

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

$$\mathbb{P}_p\big[C\big]=\mathbb{P}_p\big[\{\omega\in\Omega: G(\omega)\text{ συνεκτικός}\}\big]=\mathbb{E}_p\big[F\big].$$

Επιπλέον αν $\omega,\omega'\in\Omega$ και $\omega'_e\le\omega_e$ για κάθε $e\in E$, τότε το σύνολο των ακμών του $G(\omega')$ είναι υποσύνολο των ακμών του $G(\omega)$ και άρα $F(\omega')\le F(\omega)$. 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 (5), τότε

$$\mathbb{P}\big[F_1\ge F_2\big]=\mathbb{P}\big[\{(\omega,\omega')\in\Omega\times\Omega: F(\omega)\ge F(\omega')\}\big]\ge\mathbb{P}\big[\omega'_e\le\omega_e \text{ για κάθε } e\in E\big]=1. \qquad (6)$$

Έχουμε τώρα

$$\mathbb{P}_{p_1}\big[C\big]=\mathbb{E}_{p_1}\big[F\big]=\mathbb{E}\big[F_1\big]\ge\mathbb{E}\big[F_2\big]=\mathbb{E}_{p_2}\big[F\big]=\mathbb{P}_{p_2}\big[C\big],$$

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).

Θα κλείσουμε αυτήν την παράγραφο με μια παρατήρηση. Στην απόδειξη που δώσαμε για τον ισχυρισμό του Παραδείγματος 2, 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 $\{X_e\}_{e\in E}$ είναι ανεξάρτητες τυχαίες μεταβλητές με κατανομή Bernoulli$(1-p_1)$, β) οι $\{Y_e\}_{e\in E}$ είναι ανεξάρτητες τυχαίες μεταβλητές με κατανομή Bernoulli$(1-p_2)$ και γ) $\mathbb{P}\big[Y_e\le X_e \text{ για κάθε } e\in E\big]=1$.

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 την 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

$$\mathbb{P}\big[X_e = 1\big] = \mathbb{P}\big[U_e > 1 - p_1\big] = 1 - p_1 \quad \text{και} \quad \mathbb{P}\big[Y_e = 1\big] = \mathbb{P}\big[U_e > 1 - p_2\big] = 1 - p_2,$$

άρα πράγματι οι (α) και (β) παραπάνω ικανοποιούνται. Τέλος η (γ) είναι προφανής από την κατασκευή των X_e, Y_e αφού $Y_e = 1 \Leftrightarrow U_e > p_2 \Rightarrow X_e = 1$.

4 Ασυμπτωτική κατανομή

Η παράγραφος αυτή είναι αφιερωμένη στην απόδειξη του παρακάτω Θεωρήματος.

Θεώρημα 3 Έστω $\{X_n\}_{n\in\mathbb{N}_0}$ μια μη υποβιβάσιμη, γνησίως επαναληπτική και απεριοδική αλυσίδα στον χώρο καταστάσεων $\mathbb{X}$. Αν $\pi_n(x) = \mathbb{P}\big[X_n = x\big]$ 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

$$\lim_{n\to\infty} \pi_n = \pi.$$

Ειδικώτερα,

$$\lim_{n\to\infty} p^{(n)}(x, y) = \pi(y) \text{ για κάθε } x \in \mathbb{X}.$$

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 $\mathbb{P}\big[T < +\infty\big] = 1$.

Στον χώρο καταστάσεων $\mathbb{X} \times \mathbb{X}$ θεωρούμε την 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

$$\tilde{\pi}_0\big((x, u)\big) = \mathbb{P}\big[W_0 = (x, u)\big] = \mathbb{P}\big[X_0 = x, Y_0 = u\big] = \pi_0(x)\pi(u).$$

και πιθανότητες μετάβασης

$$\begin{aligned}\tilde{p}\big((x, u), (y, v)\big) &= \mathbb{P}\big[W_{n+1} = (y, v) \,|\, W_n = (x, u)\big] = \mathbb{P}\big[X_{n+1} = y, Y_{n+1} = v \,|\, X_n = x, Y_n = u\big] \\ &= \mathbb{P}\big[X_{n+1} = y \,|\, X_n = x\big]\mathbb{P}\big[Y_{n+1} = v \,|\, Y_n = u\big] = p(x, y)p(u, v).\end{aligned}$$

Επιπλέον η $\{W_n\}_n$ είναι μη υποβιβάσιμη. Πράγματι, εφόσον η $\{X_n\}$ είναι απεριοδική, από το Λήμμα 2 για κάθε $x, y, u, v \in \mathbb{X}$ υπάρχει ένα n_0 τέτοιο ώστε

$$n \geq n_0 \Rightarrow p^{(n)}(x, y) > 0 \text{ και } p^{(n)}(u, v) > 0.$$

Έτσι, αν $n \geq n_0$ έχουμε

$$\begin{aligned}\tilde{p}^{(n)}\big((x,u),(y,v)\big) &= \mathbb{P}\big[W_n = (y,v) \,|\, W_0 = (x,u)\big] = \mathbb{P}\big[X_n = y, Y_n = v \,|\, X_0 = x, Y_0 = u\big] \\ &= \mathbb{P}\big[X_n = y \,|\, X_0 = x\big]\mathbb{P}\big[Y_n = v \,|\, Y_0 = u\big] = p^{(n)}(x,y)p^{(n)}(u,v) > 0,\end{aligned}$$

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 $\tilde{\pi} : \mathbb{X} \times \mathbb{X} \to [0,1]$ με $\tilde{\pi}(y,v) = \pi(y)\pi(v)$ είναι αναλλοίωτη κατανομή για την $\{W_n\}_n$. 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,

$$\begin{aligned}\sum_{x,u \in \mathbb{X}} \tilde{\pi}(x,u)\tilde{p}\big((x,u),(y,v)\big) &= \sum_{x,u \in \mathbb{X}} \pi(x)\pi(u)p(x,y)p(u,v) \\ &= \sum_{x \in \mathbb{X}} \pi(x)p(x,y) \sum_{u \in \mathbb{X}} \pi(u)p(u,v) = \pi(y)\pi(v) = \tilde{\pi}(y,v).\end{aligned}$$

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 $\{W_n\}_n$ στην 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.

<u>Βήμα 2</u>: Ορίζουμε την 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 την $\{X_n\}_n$.

Αυτός ο ισχυρισμός είναι διαισθητικά φανερός. Η $\{Z_n\}$ παρακολουθεί την $\{X_n\}$ μέχρι τον χρόνο T κατά τον οποίο η $\{X_n\}_n$ συναντά για πρώτη φορά την $\{Y_n\}_n$, και στην 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 την 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 $\mathbb{P}\big[Z_0 = X_0\big] = 1$. Θα δείξουμε ακόμα ότι αν τα σημεία $z_0, z_1, \ldots, z_n \in \mathbb{X}$ είναι τέτοια ώστε $p(z_i, z_{i+1}) > 0$ για $i \in \{0, 1, \ldots, n-1\}$ τότε

$$\mathbb{P}\big[Z_{n+1} = y \,|\, Z_n = z_n, \ldots, Z_0 = z_0\big] = p(z_n, y). \tag{7}$$

Από την διαμέριση $\Omega = \bigcup_{k=0}^{n} \{T = k\} \cup \{T > n\}$ έχουμε

$$\begin{aligned}\mathbb{P}\big[Z_{n+1} = y \,|\, Z_n = z_n, \ldots, Z_0 = z_0\big] &= \sum_{k=0}^{n} \mathbb{P}\big[Z_{n+1} = y, T = k \,|\, Z_n = z_n, \ldots, Z_0 = z_0\big] \\ &\quad + \mathbb{P}\big[Z_{n+1} = y, T > n \,|\, Z_n = z_n, \ldots, Z_0 = z_0\big]\end{aligned} \tag{8}$$

Αν $k \in \{0, 1, \dots, n\}$ έχουμε

$$\begin{aligned}
&\mathbb{P}\big[Z_{n+1} = y, T = k \,|\, Z_n = z_n, \dots, Z_0 = z_0\big] \\
&\qquad = \frac{\mathbb{P}\big[X_0 = z_0 \neq Y_0, \dots, X_{k-1} = z_{k-1} \neq Y_{k-1}, X_k = Y_k = z_k, Y_{k+1} = z_{k+1}, \dots, Y_n = z_n, Y_{n+1} = y\big]}{\mathbb{P}\big[Z_0 = z_0, \dots, Z_n = z_n\big]} \\
&\qquad = \frac{\mathbb{P}\big[Y_{n+1} = y, Y_n = z_n, \dots, Y_k = z_k, Y_{k-1} \neq z_{k-1}, \dots, Y_0 \neq z_0\big]}{\mathbb{P}\big[Z_0 = z_0, \dots, Z_n = z_n\big]} \mathbb{P}\big[X_i = z_i, i = 0, \dots, k\big].
\end{aligned}$$

Από την μαρκοβιανή ιδιότητα της $\{Y_n\}_n$ έχουμε περαιτέρω ότι

$$\begin{aligned}
&\mathbb{P}\big[Z_{n+1} = y, T = k \,|\, Z_n = z_n, \dots, Z_0 = z_0\big] \\
&\qquad = p(z_n, y) \frac{\mathbb{P}\big[X_0 = z_0 \neq Y_0, \dots, X_{k-1} = z_{k-1} \neq Y_{k-1}, X_k = Y_k = z_k, Y_{k+1} = z_{k+1}, \dots, Y_n = z_n\big]}{\mathbb{P}\big[Z_0 = z_0, \dots, Z_n = z_n\big]} \\
&\qquad = p(z_n, y) \mathbb{P}\big[T = k \,|Z_0 = z_0, \dots, Z_n = z_n\big]. \qquad (9)
\end{aligned}$$

Με παρόμοιο τρόπο, έχουμε

$$\begin{aligned}
\mathbb{P}\big[Z_{n+1} = y, T > n \,|\, Z_n = z_n, \dots, Z_0 = z_0\big] &= \frac{\mathbb{P}\big[X_i = z_i \neq Y_i, i = 0, 1, \dots, n, X_{n+1} = y\big]}{\mathbb{P}\big[Z_0 = z_0, \dots, Z_n = z_n\big]} \\
&= p(z_n, y) \frac{\mathbb{P}\big[X_i = z_i \neq Y_i, i = 0, 1, \dots, n\big]}{\mathbb{P}\big[Z_0 = z_0, \dots, Z_n = z_n\big]} \\
&= p(z_n, y) \mathbb{P}\big[T > n \,|Z_0 = z_0, \dots, Z_n = z_n\big]. \qquad (10)
\end{aligned}$$

Αντικαθιστώντας τις (9) και (10) στην (8) παίρνουμε την (7).

Βήμα 3: Θα δείξουμε ότι για κάθε $x \in \mathbb{X}$ έχουμε

$$|\pi_n(x) - \pi(x)| \leq \mathbb{P}\big[T > n\big]. \qquad (11)$$

Εφόσον οι $\{X_n\}_n$ και $\{Z_n\}_n$ έχουν την ίδια κατανομή έχουμε

$$\pi_n(x) = \mathbb{P}\big[X_n = x\big] = \mathbb{P}\big[Z_n = x\big] \text{ για κάθε } x \in \mathbb{X},\ n \in \mathbb{N}_0.$$

Επειδή η αρχική κατανομή π της $\{Y_n\}_n$ είναι αναλλοίωτη κατανομή έχουμε

$$\pi(x) = \mathbb{P}\big[Y_n = x\big] \text{ για κάθε } x \in \mathbb{X},\ n \in \mathbb{N}_0.$$

Συνδυάζοντας τις δύο παραπάνω σχέσεις έχουμε ότι για κάθε $x \in \mathbb{X}$ και $n \in \mathbb{N}_0$

$$\begin{aligned}
\big|\pi_n(x) - \pi(x)\big| &= \big|\ \mathbb{P}\big[Z_n = x\big] - \mathbb{P}\big[Y_n = x\big]\ \big| \\
&= \big|\ \mathbb{P}\big[Z_n = x, T \leq n\big] + \mathbb{P}\big[Z_n = x, T > n\big] - \mathbb{P}\big[Y_n = x, T \leq n\big] - \mathbb{P}\big[Y_n = x, T > n\big]\ \big| \\
&= \big|\ \mathbb{P}\big[X_n = x, T > n\big] - \mathbb{P}\big[Y_n = x, T > n\big]\ \big| \\
&\leq \mathbb{P}\big[T > n\big].
\end{aligned}$$

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

$$\lim_{n \to \infty} \mathbb{P}\big[T > n\big] = \mathbb{P}\big[\bigcap_n \{T > n\}\big] = \mathbb{P}\big[T = \infty\big] = 0,$$

από το αποτέλεσμα του βήματος 1. Συνεπώς, για κάθε $x \in \mathbb{X}$ έχουμε $\pi_n(x) \to \pi(x)$ και άρα $\pi_n \to \pi$.

Τέλος, για να αποδείξουμε ότι $p^{(n)}(x,y) \to \pi(y)$ αρκεί να ξεκινήσουμε την αλυσίδα $\{X_n\}$ από την κατάσταση x, να πάρουμε δηλαδή $\mathbb{P}\big[X_0 = x\big] = 1$. Τότε $p^{(n)}(x,y) = \mathbb{P}\big[X_n = y\big] \to \pi(y)$ μια και ο ισχυρισμός του θεωρήματος ισχύει για οποιαδήποτε αρχική κατανομή π_0. □

Παρατηρήστε πώς χρησιμοποιήσαμε την 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 μια 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,

$$\mathbb{P}\big[W_n = (A,A) \,|\, W_0 = (A,B)\big] = p^{(n)}(A,A)p^{(n)}(B,A) = 0,$$

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 την συνθήκη $\mathbb{P}\big[T < +\infty\big] = 1$ που μας δίνει το ζητούμενο με την βοήθεια της εκτίμησης του βήματος 3.

5 Το εργοδικό θεώρημα

Σε αυτήν 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 αυτήν 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 4 Έστω $\{Z_n\}_{n\in\mathbb{N}}$ μια ακολουθία από ανεξάρτητες ισόνομες μη αρνητικές τυχαίες μεταβλητές με $\mathbb{E}\big[Z_i\big] = \mu \le \infty$. Τότε

$$\mathbb{P}\Big[\ \frac{1}{n}\sum_{k=1}^{n} Z_k \to \mu\Big] = 1.$$

Η απόδειξη αυτού του θεωρήματος μπορεί να βρεθεί στην [;] και στην [;]. 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 μας 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 πάλι 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 την μέση τιμή τους μ, και κάποιες μικρότερες. Επειδή οι μεταβλητές είναι ανεξάρτητες, οι προς τα πάνω αποκλίσεις απο την μέση τιμή σε μεγάλο βαθμό αλληλοαναιρούνται με τις προς τα κάτω αποκλίσεις, τόσο που οι αθροιστικές αποκλίσεις n προσθετέων είναι πολύ μικρότερες από το n με το οποίο διαιρούμε, και στο όριο $n \to \infty$ εξαφανίζονται.

Αν τώρα $\{X_n\}_{n\in\mathbb{N}_0}$ είναι μια μη υποβιβάσιμη μαρκοβιανή αλυσίδα σε έναν χώρο καταστάσεων $\mathbb{X}$ και $f : \mathbb{X} \to \mathbb{R}$ 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 αυτήν την περίπτωση οι αθροιστικές αποκλίσεις n προσθετέων είναι πολύ μικρότερες από το n και οι χρονικοί μέσοι όροι συγκλίνουν καθώς $n \to \infty$. 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 στην x.

Για $x \in \mathbb{X}$ ορίζουμε $V_n(x)$ το πλήθος των επισκέψεων στην x μιας αλυσίδας $\{X_n\}_{n\in\mathbb{N}_0}$ πριν την χρονική στιγμή n,

$$V_n(x) = \sum_{k=0}^{n-1} \mathbb{1}\{X_k = x\}.$$

Θεώρημα 5 (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

$$\mathbb{P}\Big[\ \frac{V_n(x)}{n} \to \frac{1}{\mathbb{E}_x\big[T_x^+\big]},\ \text{για κάθε } x \in \mathbb{X}\Big] = 1,$$

όπου $T_x^+ = \inf\{k > 0 : X_k = x\}$ είναι ο χρόνος πρώτης επανόδου της αλυσίδας στην κατάσταση x.

Απόδειξη: Θα αποδείξουμε πρώτα ότι για κάθε $x \in \mathbb{X}$

$$\mathbb{P}\Big[\ \frac{V_n(x)}{n} \to \frac{1}{\mathbb{E}_x\big[T_x^+\big]}\Big] = 1. \tag{12}$$

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 την διαδικασία $\{Y_n\}_{n\in\mathbb{N}_0}$ με $Y_n = X_{M+n}$ έχουμε

$$\sum_{k=0}^{n-1} \mathbb{1}\{X_k = x\} - \sum_{k=0}^{n-1} \mathbb{1}\{Y_k = x\} = \sum_{k=0}^{M-1} \big(\mathbb{1}\{X_k = x\} - \mathbb{1}\{X_{n+k} = x\}\big)$$

και επομένως

$$\Big|\frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}\{X_k = x\} - \frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}\{Y_k = x\}\Big| \leq \frac{M}{n} \to 0 \quad \text{καθώς } n \to \infty.$$

Μάλιστα, ακόμα κι αν η $M = M(\omega)$ είναι ένας τυχαίος χρόνος με $\mathbb{P}\big[M < +\infty\big] = 1$ 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 $\mathbb{P}\big[T_x < +\infty\big] = 1$. 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 (12) στην περίπτωση που η x είναι επαναληπτική, αρκεί να δείξουμε ότι

$$\mathbb{P}_x\Big[\ \frac{V_n(x)}{n} \to \frac{1}{\mathbb{E}_x\big[T_x^+\big]}\Big] = 1. \tag{13}$$

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 $\mathbb{P}_x\big[S_n < +\infty \text{ για κάθε } n \in \mathbb{N}\big] = 1$. Επιπλέον, όπως είδαμε στο Πόρισμα **;;**, οι χρόνοι που μεσολαβούν ανάμεσα σε διαδοχικές επισκέψεις στην κατάσταση x, $\{S_n - S_{n-1}\}_{n\in\mathbb{N}}$ είναι ανεξάρτητες, ισόνομες, τυχαίες μεταβλητές με κοινή κατανομή αυτήν της $S_1 - S_0 = T_x^+$. Από τον νόμο των μεγάλων αριθμών (Θεώρημα 4) έχουμε επομένως

$$\mathbb{P}_x\Big[\lim_{n\to\infty}\frac{S_n}{n} = \mathbb{E}_x\big[T_x^+\big]\Big] = 1.$$

Από την επαναληπτικότητα της κατάστασης x έχουμε ακόμα

$$\mathbb{P}_x\big[V_n(x)\to\infty\big] = 1.$$

Αφού τα παραπάνω ενδεχόμενα συμβαίνουν με $\mathbb{P}_x$-πιθανότητα 1, και η τομή τους θα συμβαίνει με $\mathbb{P}_x$-πιθανότητα 1, και επομένως

$$\mathbb{P}_x\Big[\lim_{n\to\infty}\frac{S_{V_n(x)}}{V_n(x)} = \mathbb{E}_x\big[T_x^+\big]\Big] = 1. \tag{14}$$

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 την χρονική στιγμή n για κάθε $n\in\mathbb{N}$, επομένως

$$\mathbb{P}_x\Big[S_{V_n(x)-1} < n \leq S_{V_n(x)} \quad \text{για κάθε } n\in\mathbb{N}\Big] = 1,$$

και άρα η (13) έπεται από την (14).

Αν η κατάσταση x είναι παροδική, η απόδειξη της (12) είναι ακόμα πιο εύκολη αφού τότε $\mathbb{E}_x\big[T_x^+\big] = +\infty$, ενώ $\mathbb{P}\big[\lim_{n\to\infty}V_n(x) < +\infty\big] = 1$ και άρα $\mathbb{P}\big[\lim_{n\to\infty}V_n(x)/n = 0\big] = 1$.

Ο ισχυρισμός του θεωρήματος προκύπτει τώρα από την (12) επειδή ο $\mathbb{X}$ έχει το πολύ αριθμήσιμο πλήθος από στοιχεία. Πράγματι, η (12) σημαίνει ότι για κάθε $x\in\mathbb{X}$ έχουμε $\mathbb{P}\big[U_x\big] = 1$, όπου

$$U_x = \big\{\omega\in\Omega : \frac{V_n(x)}{n} \to \frac{1}{\mathbb{E}_x\big[T_x^+\big]}\big\}.$$

Εφόσον καθένα από τα U_x έχει πιθανότητα 1, το ίδιο θα συμβαίνει και για την αριθμήσιμη τομή τους $\cap_{x\in\mathbb{X}}U_x$, αφού

$$\mathbb{P}\Big[\big(\bigcap_{x\in\mathbb{X}}U_x\big)^c\Big] = \mathbb{P}\Big[\bigcup_{x\in\mathbb{X}}U_x^c\Big] \leq \sum_{x\in\mathbb{X}}\mathbb{P}\big[U_x^c\big] = \sum_{x\in\mathbb{X}}0 = 0.$$

□

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

$$\pi(x) = \frac{1}{\mathbb{E}_x\big[T_x^+\big]} \quad \text{για κάθε} x\in\mathbb{X}.$$

Σε αυτήν την περίπτωση το Θεώρημα 6 μας 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 στην x η αναλλοίωτη κατανομή π:

$$\mathbb{P}\Big[\frac{V_n(x)}{n} \to \pi(x) \quad \text{για κάθε } x\in\mathbb{X}\Big] = 1. \tag{15}$$

Στην περίπτωση αυτή μπορούμε να διατυπώσουμε το εργοδικό θεώρημα και με μια διαφορετική, πολλές φορές πιο χρήσιμη για υπολογισμούς μορφή.

Θεώρημα 6 (Εργοδικό Θεώρημα ΙΙ) Έστω $\{X_n\}_{n\in\mathbb{N}_0}$ μια μη υποβιβάσιμη γνησίως επαναληπτική μαρκοβιανή αλυσίδα σε έναν διακριτό χώρο καταστάσεων $\mathbb{X}$, Αν $f:\mathbb{X}\to\mathbb{R}$ είναι μια φραγμένη συνάρτηση, τότε για οποιαδήποτε αρχική κατανομή της αλυσίδας έχουμε

$$\mathbb{P}\Big[\,\frac{1}{n}\sum_{k=0}^{n-1} f(X_k)\to\mathbb{E}^{\pi}\big[f\big]\Big]=1,$$

όπου $\mathbb{E}^{\pi}\big[f\big]=\sum_{x\in\mathbb{X}} f(x)\pi(x)$ 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 μια 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 6. Στην περίπτωση που ο $\mathbb{X}$ έχει άπειρο αλλά αριθμήσιμο πλήθος από καταστάσεις θεωρήστε $M>0$ τέτοιο ώστε $|f(x)|\le M$ 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

$$\begin{aligned}
\Big|\,\frac{1}{n}\sum_{k=0}^{n-1} f(X_k)-\mathbb{E}^{\pi}\big[f\big]\Big| &= \Big|\sum_{x\in\mathbb{X}} f(x)\big(\frac{V_n(x)}{n}-\pi(x)\big)\Big|\\
&\le M\sum_{x\in\mathbb{X}}\Big|\frac{V_n(x)}{n}-\pi(x)\Big|\\
&= M\sum_{x\in A}\Big|\frac{V_n(x)}{n}-\pi(x)\Big|+M\sum_{x\notin A}\Big|\frac{V_n(x)}{n}-\pi(x)\Big|. \qquad (16)
\end{aligned}$$

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

$$\begin{aligned}
\sum_{x\notin A}\Big|\frac{V_n(x)}{n}-\pi(x)\Big| &= \sum_{x\notin A}\Big(\frac{V_n(x)}{n}-\pi(x)\Big)+2\sum_{x\notin A}\Big(\frac{V_n(x)}{n}-\pi(x)\Big)^{-}\\
&= \sum_{x\in A}\Big(\pi(x)-\frac{V_n(x)}{n}\Big)+2\sum_{x\notin A}\Big(\frac{V_n(x)}{n}-\pi(x)\Big)^{-}\\
&\leq \sum_{x\in A}\Big(\pi(x)-\frac{V_n(x)}{n}\Big)+2\sum_{x\notin A}\pi(x)\\
&\leq \sum_{x\in A}\Big(\pi(x)-\frac{V_n(x)}{n}\Big)+\frac{\epsilon}{M},
\end{aligned}$$

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 $(u-v)^{-}\leq v$. Αντικαθιστώντας τώρα στην (16) έχουμε

$$\Big|\frac{1}{n}\sum_{k=0}^{n-1}f(X_k)-\mathbb{E}^{\pi}\big[f\big]\Big|\leq 2M\sum_{x\in A}\Big|\frac{V_n(x)}{n}-\pi(x)\Big|+\epsilon.$$

Εφόσον το σύνολο A είναι πεπερασμένο, για τα $\omega\in\Omega$ για τα οποία $V_n(x)/n\to\pi(x)$ έχουμε

$$\limsup_{n\to\infty}\Big|\frac{1}{n}\sum_{k=0}^{n-1}f(X_k)-\mathbb{E}^{\pi}\big[f\big]\Big|\leq\epsilon,$$

και άρα από το Θεώρημα 6 έχουμε ότι για κάθε $\epsilon>0$

$$\mathbb{P}\Big[\limsup_{n\to\infty}\Big|\frac{1}{n}\sum_{k=0}^{n-1}f(X_k)-\mathbb{E}^{\pi}\big[f\big]\Big|\leq\epsilon\Big]=1.$$

Αν πάρουμε τώρα $\epsilon=1/N$ η ακολουθία ενδεχομένων

$$B_N=\{\omega\in\Omega:\limsup_{n\to\infty}\Big|\frac{1}{n}\sum_{k=0}^{n-1}f(X_k)-\mathbb{E}^{\pi}\big[f\big]\Big|\leq\frac{1}{N}\}$$

είναι φθίνουσα ($B_{N+1}\subset B_N$ για κάθε $N\in\mathbb{N}$) και $\mathbb{P}\big[B_N\big]=1$ για κάθε $N\in\mathbb{N}$. Επομένως,

$$\mathbb{P}\Big[\frac{1}{n}\sum_{k=0}^{n-1}f(X_k)\to\mathbb{E}^{\pi}\big[f\big]\Big]=\mathbb{P}\big[\bigcap_{N}B_N\big]=\lim_{N\to\infty}\mathbb{P}\big[B_N\big]=1,$$

που είναι και ο ισχυρισμός του θεωρήματος. □

6 Ασκήσεις

Άσκηση 1 Αποδείξτε το Θεώρημα 2

Άσκηση 2 Υπολογίστε την συνδιακύμανση των τυχαίων μεταβλητών $\tilde{X}_e$ και $\tilde{Y}_e$ του Παραδείγματος 3.

Άσκηση 3 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 την συνάρτηση $f : p \to f(p)$ ως προς την μονοτονία της. Στην 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 4 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

$$\lim_{n\to\infty} \mathbb{P}\big[S_n \text{ είναι πολλαπλάσιο του } 7\big].$$

Άσκηση 5 Στην Άσκηση **;;**, σε βάθος χρόνου, ποιο ποσοστό του χρόνου του ξοδεύει το έντομο στο σαλόνι;

Άσκηση 6 Στην Άσκηση **;;**, σε βάθος χρόνου, ποιο ποσοστό των ημερών το μπακάλικο δεν έχει μπισκότα την ώρα που κλείνει;

Άσκηση 7 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 8 Με την βοήθεια και της προηγούμενης άσκησης δείξτε ότι αν η αλυσίδα $\{X_n\}_{n\in\mathbb{N}}$ είναι μη υποβιβάσιμη και κινείται σ′ έναν πεπερασμένο χώρο καταστάσεων $\mathbb{X}$ και $f : \mathbb{X}\times\mathbb{X} \to \mathbb{R}$ τότε με πιθανότητα 1 έχουμε

$$\lim_{n\to\infty} \frac{f(X_0,X_1) + f(X_1,X_2) + \cdots + f(X_{n-1},X_n)}{n} = \sum_{(x,y)\in\mathbb{X}\times\mathbb{X}} f(x,y)\pi(x)p(x,y).$$

για οποιαδήποτε κατανομή της X_0.

Άσκηση 9 Στην Άσκηση **;;**, σε βάθος χρόνου, ποιο ποσοστό των μεταβάσεων γίνονται μεταξύ των καταστάσεων 3 και 4; Ποιο είναι το όριο

$$\lim_{n\to\infty} \mathbb{P}\big[X_n = 4 | X_{n+1} = 3\big]?$$

Άσκηση 10 Ένας παίκτης του μπάσκετ προπονείται στα τρίποντα. Έχετε παρατηρήσει ότι η πιθανότητα να ευστοχήσει σε ένα σουτ είναι ίση με

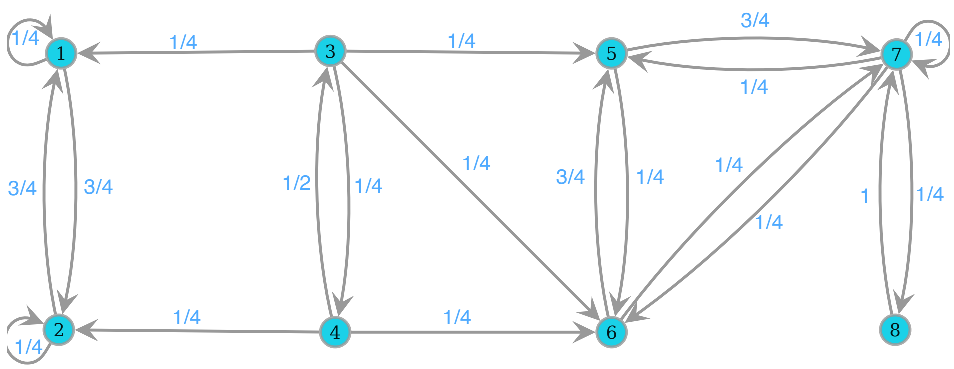
- $1/2$ αν έχει ευστοχήσει και στα δύο προηγούμενα σουτ που έχει επιχειρήσει
- $1/3$ αν έχει ευστοχήσει σε 1 από τα δύο προηγούμενα σουτ που έχει επιχειρήσει
- $1/4$ αν έχει αστοχήσει και στα δύο προηγούμενα σουτ που έχει επιχειρήσει.

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 11 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 12 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) Ψπολογίστε το όριο $\lim_n \mathbb{P}\big[X_n = 8 \,|\, X_0 = 8\big]$.
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 $\mathbb{E}\big[T_1\big]$ και $\mathbb{E}\big[T_2\big]$.
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 βάθος χρόνου; ποιες τιμές μπορεί να πάρει; με ποια πιθανότητα;

7 Αριθμητικά πειράματα

Άσκηση 13 Στην άσκηση αυτή γίνεται προσομοίωση της αλυσίδας της άσκησης 5 του Κεφαλαίου 6, 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

$$\pi(k) = \frac{1}{2^{k+1}}, \quad k = 0, 1, 2, 3, \ldots$$

Κατεβάστε και τρέξτε τον κώδικα `ergodic.py`. Ο κώδικας αυτός προσομοιώνει τα πρώτα $N = 10^6$ βήματα της αλυσίδας και επιστρέφει τον εργοδικό μέσο

$$\frac{X_1 + X_2 + \cdots + X_N}{N}.$$

Από το εργοδικό θεώρημα η παραπάνω ποσότητα συγκλίνει καθώς $N \to \infty$ με πιθανότητα 1 στο

$$\sum_{k=0}^{\infty} k\pi(k),$$

οπότε ο κώδικας υπολογίζει αριθμητικά την παραπάνω ποσότητα με την 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 την διασπορά των αποτελεσμάτων.

γ) Αν θέλαμε να περιορίσουμε την 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 την μέθοδο MCMC το άθροισμα

$$\sum_{k=0}^{\infty} \frac{\cos(k + \cos(k))}{2^k}.$$