

Φροντιστήριο 5

Άσκηση 1.

Σε μια πολύ μεγάλη ακολουθία DNA έχουν καταγραφεί οι εμπειρικές συχνότητες σύμφωνα με τις οποίες κάθε μια από τις βάσεις A, C, G, T ακολουθείται από μια άλλη. Αυτές φαίνονται στην παρακάτω πίνακα.

$$P = \begin{pmatrix} 0.200 & 0.255 & 0.285 & 0.260 \\ 0.302 & 0.322 & 0.178 & 0.198 \\ 0.392 & 0.139 & 0.292 & 0.177 \\ 0.148 & 0.198 & 0.346 & 0.308 \end{pmatrix}$$

Αν θεωρήσουμε (προσεγγιστικά) ότι η εναλλαγή των βάσεων ακολουθεί το μοντέλο μιας Μαρκοβιανής αλυσίδας με χώρο καταστάσεων $S = \{A, C, G, T\}$. Κάποιος συνάδελφός σας ισχυρίζεται ότι η οριακή κατανομή αυτής της αλυσίδας είναι η $[\pi_A, \pi_C, \pi_G, \pi_T] = [0.2639, 0.2245, 0.2772, 0.2344]$.

- α') Είναι σωστός ο ισχυρισμός του συναδέλφου σας; Βρείτε και ερμηνεύστε την οριακή κατανομή.
β') Βρείτε την πιθανότητα σε ένα τυχαίο απομακρυσμένο σημείο της ακολουθίας να υπάρχουν δύο διαδοχικές βάσεις A.

Λύση.

- α') Έστω X_n ο τύπος της n -οστής βάσης στην ακολουθία, με $X_n \in S = \{A, C, G, T\}$.

Εφόσον ο πίνακας μετάβασης $P > 0$ (όλα τα στοιχεία του είναι θετικά), η αλυσίδα έχει μια μοναδική επαναληπτική κλάση, η οποία είναι μοναδική. Υπό αυτές τις συνθήκες, η οριακή κατανομή $\underline{\pi} = [\pi_A, \pi_C, \pi_G, \pi_T]$ υπάρχει και είναι η μοναδική λύση των εξισώσεων ισορροπίας:

$$\begin{cases} \underline{\pi} = \underline{\pi} \cdot P \\ \underline{\pi} \cdot \underline{e} = 1 \end{cases}$$

Αντικαθιστώντας την κατανομή που μας δίνει ο συνάδελφός μας, διαπιστώνουμε ότι ικανοποιείται από τις εξισώσεις ισορροπίας, άρα είναι σωστός ο ισχυρισμός του.

Σημείωση: Θα μπορούσαμε να υπολογίσουμε και μόνοι μας την στατική κατανομή από τις εξισώσεις ισορροπίας, αλλά οι πιθανότητες στον παραπάνω στοχαστικό πίνακα δεν βοηθάνε για πράξεις στο χαρτί.

Ερμηνεία 1. (Οριακή Κατανομή) Αν πάρουμε μια βάση της αλυσίδα σε ένα σημείο πολύ μακριά από την αρχή, τότε ανεξάρτητα από την αρχική βάση, η πιθανότητα να βρούμε A τείνει στο 0.2639 κ.ο.κ.

Ερμηνεία 2. (Συχνότητες) Σε μια μεγάλη αλυσίδα, περίπου 26.39% των βάσεων είναι τύπου A, 22.45% είναι τύπου G κ.ο.κ.

β') Ζητείται η πιθανότητα $P(X_n = A, X_{n+1} = A)$, καθώς το $n \rightarrow \infty$.

$$\begin{aligned} \lim_{n \rightarrow \infty} P(X_n = A, X_{n+1} = A) &= \lim_{n \rightarrow \infty} P(X_{n+1} = A | X_n = A) P(X_n = A) \\ &= P_{AA} \cdot \pi_A = 0.200 \cdot 0.2639 \\ &= 0.0528 \end{aligned}$$

Άσκηση 2.

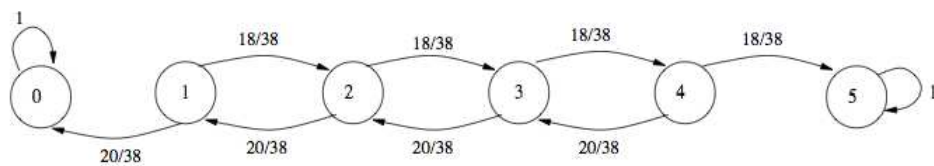
Ο Κώστας πηγαίνει στο Las Vegas. Δε θέλει να χάσει πάρα πολλά λεφτά οπότε αποφασίζει να παίζει μόνο με 3\$ και να σταματήσει να παίζει όταν θα τα έχει κάνει 5\$ ή όταν θα τα έχει χάσει όλα. Πλησιάζει μία ρουλέτα η οποία έχει 18 κόκκινες, 18 μαύρες και 2 πράσινες θέσεις. Αποφασίζει ότι πάντα θα ποντάρει 1\$ στο κόκκινο.

- α') Ορίστε τη Μαρκοβιανή αλυσίδα που περιγράφει πλήρως την εξέλιξη του παιχνιδιού του Κώστα. Ορίστε επακριβώς τις καταστάσεις, δώστε το γράφημα της αλυσίδας με τις πιθανότητες μετάβασης καθαρά γραμμένες στις ακμές και γράψτε τον αντίστοιχο πίνακα πιθανοτήτων μετάβασης. Ποιες από τις καταστάσεις είναι έμμονες και ποιες είναι μεταβατικές;
- β') Με τι πιθανότητα θα φύγει ο Κώστας χαρούμενος από το καζίνο; Με τι πιθανότητα θα φύγει χαμένος; Πώς μεταβάλλονται αυτές οι πιθανότητες αν αρχίζει το παιχνίδι με 2\$;
- γ') Πόσο θα διαρκέσει η διασκέδαση του Κώστα στη ρουλέτα; Με άλλα λόγια, ποιος είναι ο αναμενόμενος αριθμός προσπαθειών μέχρις ότου χάσει τα λεφτά του ή φύγει με 5\$ στην τσέπη;

Λύση.

- α') Ως κατάσταση του συστήματος ορίζουμε τον αριθμό των δολαρίων που ο Κώστας έχει στην τσέπη του $\{0, 1, 2, 3, 4, 5\}$.

Το γράφημα της αλυσίδας φαίνεται στο Σχήμα 1.



Σχήμα 1: Το διάγραμμα καταστάσεων για την Άσκηση 1.

Ο πίνακας πιθανοτήτων μετάβασης είναι ο ακόλουθος:

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{20}{38} & 0 & \frac{18}{38} & 0 & 0 & 0 \\ 0 & \frac{20}{38} & 0 & \frac{18}{38} & 0 & 0 \\ 0 & 0 & \frac{20}{38} & 0 & \frac{18}{38} & 0 \\ 0 & 0 & 0 & \frac{20}{38} & 0 & \frac{18}{38} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

- Έμμονες καταστάσεις: 0 και 5.
(Προσβάσιμες από όλες τις καταστάσεις της κλάσης τους, αλλά όχι από επαναληπτικές καταστάσεις άλλων κλάσεων.)
- Μεταβατικές καταστάσεις: 1, 2, 3 και 4.
(Δεν είναι προσβάσιμες από καμία έμμονη κατάσταση.)

β') Ζητούνται ουσιαστικά οι πιθανότητες απορρόφσεως του συστήματος από τις καταστάσεις $s = 0$ (φεύγει χαμένος) και $s = 5$ (φεύγει κερδισμένος), όταν αυτό ξεκινά από τις καταστάσεις $i = 3$ ή $i = 2$.

Σχετική θεωρία:

Έστω μία συγκεκριμένη κατάσταση απορρόφησης s . Ορίζουμε ως πιθανότητες απορρόφησης του συστήματος από την s , ξεκινώντας από την i :

$$\alpha_i \triangleq P(X_n \text{ τελικά ισούται με την κατάσταση απορρόφησης } s \mid X_0 = i)$$

Για μια αλυσίδα Markov όπου κάθε κατάσταση είναι είτε μεταβατική, είτε απορροφήσεως, οι πιθανότητες απορρόφησης α_i της αλυσίδας από την κατάσταση s ξεκινώντας από την i είναι οι μοναδικές λύσεις του γραμμικού συστήματος:

$$\begin{aligned} \alpha_s &= 1 \\ \alpha_i &= 0, \quad \text{για κάθε κατάσταση απορρόφησης } i \neq s \\ \alpha_j &= \sum_{i=1}^m p_{ij} \alpha_i, \quad \text{για κάθε μεταβατική κατάσταση } i \end{aligned}$$

Εδώ, έστω ότι ο Κώστας φεύγει χαμένος, τότε $s = 0$ και οι πιθανότητες απορρόφησης ικανοποιούν το γραμμικό σύστημα:

$$\begin{aligned} \alpha_0 &= 1 \\ \alpha_5 &= 0 \\ \alpha_1 &= p_{10} \cdot \alpha_0 + p_{12} \cdot \alpha_2 = \frac{20}{38} \cdot 1 + \frac{18}{38} \cdot \alpha_2 \\ \alpha_2 &= p_{21} \cdot \alpha_1 + p_{23} \cdot \alpha_3 = \frac{20}{38} \cdot \alpha_1 + \frac{18}{38} \cdot \alpha_3 \\ \alpha_3 &= p_{32} \cdot \alpha_2 + p_{34} \cdot \alpha_4 = \frac{20}{38} \cdot \alpha_2 + \frac{18}{38} \cdot \alpha_4 \\ \alpha_4 &= p_{43} \cdot \alpha_3 + p_{45} \cdot \alpha_5 = \frac{20}{38} \cdot \alpha_3 + \frac{18}{38} \cdot 0 \end{aligned}$$

Λύνοντας το σύστημα έχουμε:

$$\alpha = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.8398 \\ 0.6618 \\ 0.4640 \\ 0.2442 \\ 0 \end{bmatrix}$$

Η πιθανότητα λοιπόν, να φύγει χαμένος από το καζίνο, ενώ έχει μπει με 3\$ είναι $\alpha_3 = 0.4640$, ενώ η πιθανότητα να φύγει χαμένος αν ξεκινούσε με 2\$ είναι $\alpha_2 = 0.6618$.

Αντίστοιχα, η πιθανότητα ο Κώστας να φύγει χαρούμενος είναι η πιθανότητα απορρόφησης του συστήματος από την $s = 5$. Θα μπορούσαμε να κάνουμε ακριβώς την ίδια διαδικασία επίλυσης γραμμικού συστήματος όπως παραπάνω, αλλά αυτό δεν είναι απαραίτητο διότι έχουμε μόνο δύο έμμονες καταστάσεις, άρα οι ζητούμενες πιθανότητες είναι:

$$1 - \alpha = \begin{bmatrix} 0 \\ 0.1602 \\ 0.3382 \\ 0.5360 \\ 0.7558 \\ 1 \end{bmatrix}$$

Συνεπώς, η πιθανότητα να φύγει χαρούμενος από το καζίνο ενώ είχε μπει με 3\$ είναι $1 - \alpha_3 = 0.5360$, ενώ αν είχε μπει με 2\$ είναι $1 - \alpha_2 = 0.3382$.

γ') Ζητείται ο μέσος χρόνος απορρόφησης του συστήματος, ξεκινώντας από την κατάσταση $i = 3$.

Σχετική θεωρία:

Μέσος χρόνος απορροφήσεως μ_i , ξεκινώντας από την κατάσταση i είναι οι μοναδικές λύσεις του συστήματος:

$$\begin{aligned} \mu_i &= 0, \quad \text{για κάθε έμμονη κατάσταση } i \\ \mu_i &= 1 + \sum_{j=1}^m p_{ij} \mu_j, \quad \text{για κάθε μεταβατική κατάσταση } i \end{aligned}$$

Άρα εδώ:

$$\begin{aligned} \mu_0 &= \mu_5 = 0 \\ \mu_1 &= 1 + p_{10} \cdot \mu_0 + p_{12} \cdot \mu_2 = 1 + \frac{20}{38} \cdot 0 + \frac{18}{38} \cdot \mu_2 \\ \mu_2 &= 1 + p_{21} \cdot \mu_1 + p_{23} \cdot \mu_3 = 1 + \frac{20}{38} \cdot \mu_1 + \frac{18}{38} \cdot \mu_3 \\ \mu_3 &= 1 + p_{32} \cdot \mu_2 + p_{34} \cdot \mu_4 = 1 + \frac{20}{38} \cdot \mu_2 + \frac{18}{38} \cdot \mu_4 \\ \mu_4 &= 1 + p_{43} \cdot \mu_3 + p_{45} \cdot \mu_5 = 1 + \frac{20}{38} \cdot \mu_3 + \frac{18}{38} \cdot 0 \end{aligned}$$

Λύνουμε το σύστημα και βρίσκουμε το ζητούμενο μ_3 .

Άσκηση 3.

Ένα πειραματικό ποντίκι είναι φυλακισμένο μέσα σε ένα κλουβί με τρεις πόρτες. Σε κάθε γύρο, το ποντίκι ανοίγει τυχαία μία πόρτα και τρώει ένα κομμάτι τυρί, εάν υπάρχει τέτοιο πίσω από την πόρτα που άνοιξε. Κάθε πόρτα επιλέγεται με ίση πιθανότητα σε κάθε γύρο, ανεξάρτητα από το που το ποντίκι βρήκε τυρί στον προηγούμενο γύρο. Αν δε βρέθηκε τυρί στον προηγούμενο γύρο, η πιθανότητα ότι το ποντίκι θα βρει τυρί σε αυτόν τον γύρο είναι $\frac{3}{4}$. Αν βρέθηκε τυρί στον προηγούμενο γύρο και η ίδια πόρτα επιλεγεί σε αυτόν τον γύρο, η πιθανότητα ότι το ποντίκι θα βρει τυρί είναι 0. Ενώ αν βρέθηκε τυρί στον προηγούμενο γύρο και διαφορετική πόρτα επιλεγεί σε αυτόν τον γύρο, η πιθανότητα ότι το ποντίκι θα βρει τυρί είναι 1.

- α') Μας ενδιαφέρει το αν το ποντίκι τρώει το τυρί του σε κάθε γύρο. Ορίστε τη Μαρκοβιανή αλυσίδα που περιγράφει πλήρως την εξέλιξη του πειράματος. Ορίστε επακριβώς τις καταστάσεις, δώστε το γράφημα της αλυσίδας με τις πιθανότητες μετάβασης καθαρά γραμμένες στις ακμές, και γράψτε τον αντίστοιχο πίνακα πιθανοτήτων μετάβασης.
- β') Αν παρατηρήσετε το ποντίκι για 1000 γύρους, σε ποιο ποσοστό αυτών θα το δείτε να τρώει το τυρί του;
- γ') Έστω ότι δε βρέθηκε τυρί στον προηγούμενο γύρο. Ποιος είναι ο μέσος αριθμός γύρων έως ότου το ποντίκι να φάει ένα κομμάτι τυρί;
- δ') Έστω ότι δε βρέθηκε τυρί στον προηγούμενο γύρο. Ποιος είναι ο μέσος αριθμός γύρων έως ότου το ποντίκι να φάει n κομμάτια τυρί;

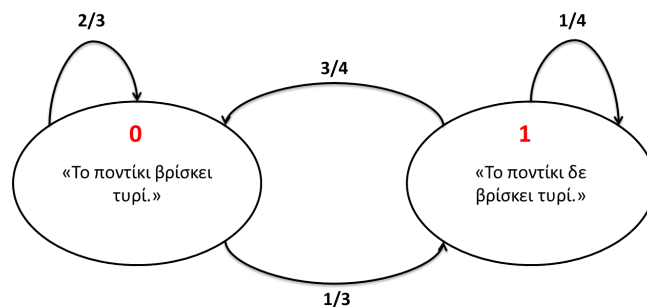
Βοήθεια: Στα ερωτήματα (γ') και (δ') χρησιμοποιείτε τις γνώσεις σας σχετικά με μέσους χρόνους ως την πρώτη επίσκεψη και μέσο χρόνο επανόδου, αντίστοιχα.

- ε') Κοιτάτε μέσα στο κλουβί και παρατηρείτε ότι το ποντίκι τρώει ένα κομμάτι τυρί πίσω από την πόρτα 1. Ποια είναι η πιθανότητα ότι 3 γύρους μετά, το ποντίκι θα βρεθεί να τρώει πάλι πίσω από την πόρτα 1;

Βοήθεια: Στο ερώτημα (ε'), χρειάζεται να ορίσετε μία νέα Μαρκοβιανή αλυσίδα που να περιγράφει και τις πόρτες πίσω από τις οποίες το ποντίκι βρίσκει τυρί.

Λύση.

- α') Εφόσον μας ενδιαφέρει αν το ποντίκι τρώει το τυρί του, ορίζουμε μία αλυσίδα Markov με δύο καταστάσεις, όπως φαίνεται στο Σχήμα 2.



Σχήμα 2: Το διάγραμμα καταστάσεων για την Άσκηση 2, Ερώτημα α'.

Ο αντίστοιχος πίνακας πιθανοτήτων μετάβασης:

$$\mathbf{P} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{3}{4} & \frac{1}{4} \end{bmatrix}$$

Εξήγηση επιμέρους πιθανοτήτων:

- $P(\Theta \text{α βρει} \mid \Delta \text{ε βρήκε πριν}) = \frac{3}{4}$ (δίδεται από την εκφώνηση)
- $P(\Delta \text{ε θα βρει} \mid \Delta \text{ε βρήκε πριν}) = 1 - P(\Theta \text{α βρει} \mid \Delta \text{ε βρήκε πριν}) = 1 - \frac{3}{4} = \frac{1}{4}$.
- $P(\Theta \text{α βρει} \mid \text{Βρήκε πριν}) = P(\text{Να επιλέξει διαφορετική πόρτα από αυτήν του προηγούμενου γύρου}) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$.
- $P(\Delta \text{ε θα βρει} \mid \text{Βρήκε πριν}) = P(\text{Να επιλέξει την ίδια πόρτα με αυτήν του προηγούμενου γύρου}) = \frac{1}{3}$.

β') Για να βρούμε σε πόσους από τους 1000 γύρους το ποντίκι τρώει το τυρί, θα πρέπει να υπολογίσουμε τη στατική κατανομή της αλυσίδας.

Οι εξισώσεις ισορροπίας παίρνουν τη μορφή:

$$\pi = \pi \cdot \mathbf{P} \quad \Rightarrow \quad \begin{bmatrix} \pi_0 & \pi_1 \end{bmatrix} = \begin{bmatrix} \pi_0 & \pi_1 \end{bmatrix} \cdot \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{3}{4} & \frac{1}{4} \end{bmatrix}$$

Επομένως,

$$\begin{aligned} \pi_0 &= \frac{2}{3}\pi_0 + \frac{3}{4}\pi_1 \\ \pi_1 &= \frac{1}{3}\pi_0 + \frac{1}{4}\pi_1 \\ \Rightarrow \quad \frac{1}{3}\pi_0 &= \frac{3}{4}\pi_1 \end{aligned}$$

που σε συνδυασμό με την $\pi_0 + \pi_1 = 1$ μας δίνει $\pi_0 = \frac{9}{13} = 0.692$ και $\pi_1 = \frac{4}{13}$.

Άρα σε κάποιο συγκεκριμένο γύρο στο μακρινό μέλλον, το ποντίκι θα τρώει τυρί με πιθανότητα 0.692 και αν παρατηρήσουμε το τυρί για 1000 γύρους, θα τρώει το τυρί στους $1000 \times 0.692 = 692$ από αυτούς.

γ') Μας ζητείται ο μέσος χρόνος ως την πρώτη επίσκεψη της κατάστασης 0, ξεκινώντας από την κατάσταση 1.

Σχετική θεωρία:

Ο μέσος αριθμός των βημάτων ή αναμενόμενος χρόνος μέχρι την πρώτη διέλευση του συστήματος από μία συγκεκριμένη έμμονη κατάσταση s , ξεκινώντας από οποιαδήποτε κατάσταση i , είναι οι μοναδικές λύσεις του συστήματος:

$$\begin{aligned} t_s &= 0 \\ t_i &= 1 + \sum_{j=1}^m p_{ij}t_j, \quad \text{για κάθε } i \neq s \end{aligned}$$

Άρα, εδώ:

$$\begin{aligned} t_0 &= 0, s = 0 \quad (\text{η έμμονη κατάσταση για την οποία ψάχνουμε την 1η διέλευση}) \\ t_1 &= 1 + p_{10} \cdot t_0 + p_{11} \cdot t_1 = 1 + \frac{3}{4} \cdot 0 + \frac{1}{4} \cdot t_1 \quad \Rightarrow \quad t_1 = \frac{4}{3} = 1.33 \end{aligned}$$

δ') Εδώ, μας ζητείται ο αναμενόμενος χρόνος T , μέχρι την πρώτη διέλευση του συστήματος από την κατάσταση 0, ακολουθούμενος από $(n-1)$ επανόδους στην κατάσταση 0, ξεκινώντας από τη 0, δηλαδή:

$$T = t_1 + (n-1) \cdot t_0^*$$

Σχετική θεωρία:

Ο μέσος χρόνος επανάληψης της κατάστασης s , ορίζεται ως:

$$t_s^* = E[\text{αριθμός μεταβάσεων μέχρι την πρώτη επιστροφή στην } s, \text{ ξεκινώντας από την } s]$$

Μπορούμε να υπολογίσουμε το t_s^* , αν γνωρίζουμε τους χρόνους της 1ης επίσκεψης t_i , χρησιμοποιώντας την εξίσωση:

$$t_s^* = 1 + \sum_{j=1}^m p_{sj} t_j$$

Εδώ, γνωρίζουμε από το ερώτημα (γ') ότι $t_0 = 0$ και $t_1 = \frac{4}{3}$.

Έχουμε λοιπόν:

$$t_0^* = 1 + p_{01} \cdot t_1 + p_{00} \cdot t_0 = 1 + \frac{1}{3} \cdot \frac{4}{3} + \frac{2}{3} \cdot 0 = 1 + \frac{1}{3} \cdot \frac{4}{3} = \frac{13}{9}.$$

Συνεπώς,

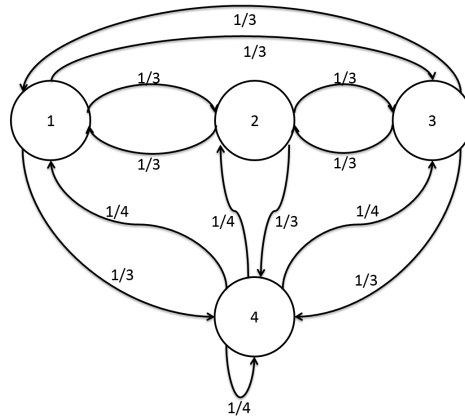
$$T = \frac{4}{3} + (n-1) \cdot \frac{13}{9}$$

ε') Εδώ, μας ενδιαφέρει πίσω από ποια πόρτα βρίσκει το ποντίκι να φάει. Επομένως, ορίζουμε κατάλληλα τις εξής καταστάσεις:

Κατάσταση i : ' Το ποντίκι βρίσκει τυρί πίσω από την πόρτα $i = 1, 2, 3$ '.

Κατάσταση 4: ' Το ποντίκι δε βρίσκει τυρί '.

Η αντίστοιχη αλυσίδα Markov φαίνεται στο Σχήμα 3.



Σχήμα 3: Το διάγραμμα καταστάσεων για την Άσκηση 2, Ερώτημα ε'.

Ο πίνακας πιθανοτήτων μετάβασης:

$$\mathbf{P} = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

Εξήγηση μεταβάσεων:

- Από κάθε πόρτα έχει πιθανότητα $\frac{1}{3}$ να επιλέξει την ίδια πόρτα και να μη βρει τυρί. Για κάθε πόρτα διαφορετική της τωρινής, υπάρχει πιθανότητα $\frac{1}{3}$ να επιλεγεί (και σίγουρα πίσω της θα υπάρχει τυρί).
- $P(\Delta\epsilon \text{ θα βρει τυρί} \mid \Delta\epsilon \text{ βρήκε τυρί πριν}) = \frac{1}{4}$ (από την εκφώνηση).
- $P(\Theta\alpha \text{ βρει τυρί} \mid \Delta\epsilon \text{ βρήκε τυρί πριν}) = 1 - P(\Delta\epsilon \text{ θα βρει τυρί} \mid \Delta\epsilon \text{ βρήκε τυρί πριν}) = 1 - \frac{1}{4} = \frac{3}{4} = \frac{1}{4} + \frac{1}{4} + \frac{1}{4}$ ($\frac{1}{4}$ για κάθε πόρτα).

Υπάρχουν 7 διαδρομές μήκους 3 που ξεκινούν και καταλήγουν στην κατάσταση 1:

- $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$: Πιθανότητα $\frac{1}{27}$
- $1 \rightarrow 2 \rightarrow 4 \rightarrow 1$: Πιθανότητα $\frac{1}{36}$
- $1 \rightarrow 3 \rightarrow 2 \rightarrow 1$: Πιθανότητα $\frac{1}{27}$
- $1 \rightarrow 3 \rightarrow 4 \rightarrow 1$: Πιθανότητα $\frac{1}{36}$
- $1 \rightarrow 4 \rightarrow 2 \rightarrow 1$: Πιθανότητα $\frac{1}{36}$
- $1 \rightarrow 4 \rightarrow 3 \rightarrow 1$: Πιθανότητα $\frac{1}{36}$
- $1 \rightarrow 4 \rightarrow 4 \rightarrow 1$: Πιθανότητα $\frac{1}{48}$

Συνεπώς, η ζητούμενη πιθανότητα είναι:

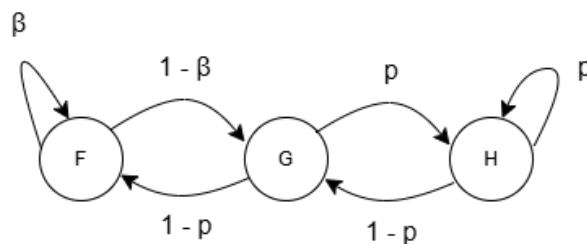
$$2 \cdot \frac{1}{27} + 4 \cdot \frac{1}{36} + \frac{1}{48} = 0.206.$$

Άσκηση 4.

Διαβάζετε ένα επιστημονικό βιβλίο που περιέχει τρία κεφάλαια με τίτλους, F, G και H. Αν μία τυχαία μέρα διαβάζετε το κεφάλαιο F, με πιθανότητα β διαβάζετε το ίδιο κεφάλαιο την επόμενη μέρα. Αλλιώς διαβάζετε το κεφάλαιο G. Αν μία τυχαία μέρα διαβάζετε το κεφάλαιο G, με πιθανότητα p διαβάζετε το κεφάλαιο H την επόμενη μέρα. Αλλιώς διαβάζετε το κεφάλαιο F. Αν μία τυχαία μέρα διαβάζετε το κεφάλαιο H, με πιθανότητα p διαβάζετε το ίδιο κεφάλαιο την επόμενη μέρα. Αλλιώς διαβάζετε το κεφάλαιο G.

- α') Ορίστε τη Μαρκοβιανή αλυσίδα που περιγράφει πλήρως τη διαδικασία ανάγνωσης του βιβλίου. Ορίστε επακριβώς τις καταστάσεις, δώστε το γράφημα της αλυσίδας με τις πιθανότητες μετάβασης γραμμένες στις ακμές, και γράψτε τον αντίστοιχο πίνακα πιθανοτήτων μετάβασης. Τι είδους αλυσίδα προκύπτει;
- β') Αν $\beta < 1$, τι ποσοστό του χρόνου θα διαβάζετε καθένα από τα τρία κεφάλαια (σε βάθος χρόνου);
- γ') Αν διαβάζετε το κεφάλαιο G μία μέρα, ποιο είναι το αναμενόμενο πλήθος των ημερών που θα σας πάρει για να μεταβείτε για πρώτη φορά στο κεφάλαιο F;
- δ') Ποιο είναι το αναμενόμενο πλήθος των ημερών που θα σας πάρει για να επανέρθετε για πρώτη φορά στο κεφάλαιο F δεδομένου ότι ξεκινήσατε από το κεφάλαιο F;
- ε') Ορίζουμε τη στοχαστική διαδικασία Y_k , $k = 1, 2, 3, \dots$ ως εξής: Y_1 είναι το πλήθος των ημερών που θα σας πάρει για να επανέρθετε για πρώτη φορά στο κεφάλαιο F δεδομένου ότι ξεκινήσατε από το κεφάλαιο F την μέρα 0. Y_k είναι το πλήθος των ημερών που θα σας πάρει για να επανέρθετε για k -οστή φορά στο κεφάλαιο F δεδομένου ότι ξεκινήσατε από το κεφάλαιο F την μέρα 0. Είναι οι τ.μ. $Y_1, Y_2 - Y_1, \dots, Y_k - Y_{k-1}$, $k > 1$ ανεξάρτητες; Εξηγήστε γιατί ναι ή γιατί όχι.
- ς') Ποιο είναι το όριο ως προς την πιθανότητα της ακολουθίας των τ.μ. Y_k/k ;

Λύση.



Σχήμα 4: Γράφος της αλυσίδας

- α') Πίνακας πιθανοτήτων μετάβασης:

$$P = \begin{bmatrix} \beta & 1 - \beta & 0 \\ 1 - p & 0 & p \\ 0 & 1 - p & p \end{bmatrix}$$

Παρατηρούμε ότι πρόκειται για μια αλυσίδα γεννήσεως-θανάτου.

β') Για $\beta < 1$, έχουμε μία κλάση επικοινωνίας που αποτελείται από τις καταστάσεις F, G, H. Οι τοπικές εξισώσεις ισορροπίας δίνουν:

$$\begin{aligned}\pi_F(1 - \beta) &= \pi_G(1 - p) \Rightarrow \pi_F = \frac{1 - p}{1 - \beta} \pi_G \\ \pi_G \cdot p &= \pi_H(1 - p) \Rightarrow \pi_H = \frac{p}{1 - p} \pi_G \\ \pi_F + \pi_G + \pi_H &= 1 \Rightarrow \frac{1 - p}{1 - \beta} \pi_G + \pi_G + \frac{p}{1 - p} \pi_G = 1 \\ &\Rightarrow \pi_G \cdot \frac{(1 - p)^2 + (1 - \beta)(1 - p) + p(1 - \beta)}{(1 - p)(1 - \beta)} = 1 \\ &\Rightarrow \begin{bmatrix} \pi_G = \frac{(1 - p)(1 - \beta)}{2 - 2p - \beta + p^2} \\ \pi_F = \frac{(1 - p)^2}{2 - 2p - \beta + p^2} \\ \pi_H = \frac{p(1 - \beta)}{2 - 2p - \beta + p^2} \end{bmatrix}\end{aligned}$$

γ') Ζητάμε το μέσο χρόνο έως την πρώτη επίσκεψη. Έστω t_G και t_H ο μέσος αριθμός βημάτων (ημερών) μέχρι την πρώτη επίσκεψη στην κατάσταση F όταν ξεκινάμε από τις καταστάσεις G και H, αντίστοιχα. Έχουμε ότι:

$$\begin{cases} t_G = 1 + p \cdot t_H + (1 - p) \cdot 0 \\ t_H = 1 + (1 - p) \cdot t_G + p \cdot t_H \end{cases} \Rightarrow \begin{cases} t_H = \frac{1}{(1 - p)^2} \\ t_G = \frac{p}{(1 - p)^2} \end{cases}$$

δ')

$$t_F^* = 1 + (1 - \beta) \cdot t_G + 0 \cdot t_H = 1 + \frac{p \cdot (1 - \beta)}{(1 - p)^2}$$

ε') Η διαφορά $Y_k - Y_{k-1}$ είναι το πλήθος των ημερών που παίρνει για να πάμε από το κεφάλαιο F στο F, την k -οστή φορά που συμβαίνει αυτό. Από την Μαρκοβιανή ιδιότητα, το παρελθόν και το μέλλον είναι ανεξάρτητα δεδομένης της παρούσας κατάστασης. Επομένως, δεδομένου ότι είμαστε τώρα στην κατάσταση F, το πλήθος των ημερών για να επιστρέψουμε στην F είναι ανεξάρτητο από οτιδήποτε έχει συμβεί στο παρελθόν. Άρα οι τ.μ. $(Y_k - Y_{k-1})$ είναι ανεξάρτητες και όμοια κατανομημένες.

Ϝ') Από προηγούμενα ερωτήματα γνωρίζουμε ότι οι i.i.d. τ.μ. $Z_i = Y_i - Y_{i-1}$ έχουν μέση τιμή:

$$\mathbb{E}[Z_i] = t_F^* = 1 + \frac{p \cdot (1 - \beta)}{(1 - p)^2}$$

Έχουμε ότι:

$$\frac{Y_k}{k} = \frac{1}{k} \sum_{i=1}^k Z_i.$$

Άρα από τον νόμο των μεγάλων αριθμών έχουμε ότι η ακολουθία $\frac{Y_k}{k}$ συγκλίνει ως προς την πιθανότητα στο $\mathbb{E}[Z_i] = 1 + \frac{p \cdot (1 - \beta)}{(1 - p)^2}$.