**Σημειώσεις Στοχαστικές Διεργασίες - 5ο Μάθημα**
**Τρίτη 31 Οκτωβρίου 2023**

**Παράδειγμα 1**
Σε ένα διάσημο πείραμα, ο Asch (1951) εξέτασε τον βαθμό στον οποίο οι άνθρωποι αποδέχονται την κοινωνική πίεση για συμμόρφωση. Φαινομενικά, το πείραμα αφορούσε την οπτική αντίληψη. Ένα άτομο οδηγούταν σε ένα δωμάτιο και καθόταν με κάποιους άλλους συμμετέχοντες. Ο πειραματιστής τους παρουσίαζε μια απλή οπτική εργασία - να καθορίσουν ποια από τις τρεις γραμμές είχε το ίδιο μήκος με μια γραμμή αναφοράς. Στη συνέχεια, ζητούνταν από κάθε συμμετέχοντα να δώσει την απάντησή του, με το υποκείμενο να απαντά τελευταίος. Η εργασία σχεδιάστηκε έτσι ώστε η σωστή απάντηση να είναι προφανής. Ωστόσο, εν αγνοία του συμμετέχοντα, όλα τα άλλα άτομα ήταν συνεργάτες του πειραματιστή και είχαν λάβει οδηγίες να δώσουν την ίδια εσφαλμένη απάντηση. Έτσι, όταν τελικά ερχόταν η σειρά του υποκειμένου να απαντήσει, το υποκείμενο έπρεπε να επιλέξει αν θα δώσει τη σωστή απάντηση (αγνοώντας την πίεση συμμόρφωσης με την ομάδα) ή τη λανθασμένη απάντηση (υποχώρηση σε αυτήν την πίεση).

Μετά την καταγραφή των απαντήσεων, ο πειραματιστής συνέχιζε με άλλη δοκιμασία όπου υπήρχε μία εξίσου προφανής επιλογή με τους συνεργάτες του να απαντούν πάλι εσκεμμένα λάθος και να αφήνουν πάλι το συμμετέχοντα να επιλέξει αν θα έλεγε τη σωστή απάντηση και όχι αυτήν που είχαν πει όλοι οι άλλοι.
Σε μια εκδοχή αυτού του πειράματος, αυτή η διαδικασία επαναλήφθηκε συνολικά 35 φορές. Έτσι, για κάθε υποκείμενο, τα δεδομένα μπορεί να αποτελούνται από μια σειρά απαντήσεων όπως π.χ

αααααβαβββαββββββββββββββββββββββββ

όπου το α δηλώνει τη σωστή απάντηση (αγνόηση των υπολοίπων) και το β τη λανθασμένη απάντηση (κοινωνική συμμόρφωση).

Αυτό το πείραμα επαναλήφθηκε με άλλα άτομα, καθένα από τα οποία δημιούργησε τη δική του ακολουθία απόκρισης. Βρέθηκε πως οι περισσότεροι ερωτώμενοι αρχικά αμφιταλαντεύονται μεταξύ των α και β, αλλά τελικά επέλεγαν μια σταθερή απάντηση (είτε το α είτε το β) χωρίς μεταβολή από επανάληψη σε επανάληψη της δοκιμασίας.

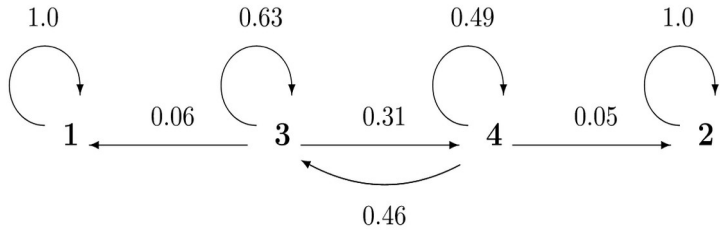
Για την μοντελοποίηση της συμπεριφοράς των ερωτώμενων, ο Cohen (1958) πρότεινε ένα μοντέλο αλυσίδας Markov με 4 καταστάσεις:
Κατάσταση 1: σταθερός αντικομφορμιστής (δεν επηρεάζεται από τους άλλους)
Κατάσταση 2: σταθερός κομφορμιστής (επιλέγει την άποψη των άλλων, αν και λανθασμένη)
Κατάσταση 3: προσωρινός αντικομφορμιστής (δεν επηρεάζεται από τους άλλους, προσωρινά)
Κατάσταση 4: προσωρινός κομφορμιστής (προσωρινά επιλέγει την άποψη των άλλων)
Τα υποκείμενα στην κατάσταση 1 ή 3 δίνουν σωστή απάντηση (α) αγνοώντας τους υπόλοιπους, ενώ τα άτομα στην κατάσταση 2 ή 4 δίνουν λανθασμένη απάντηση (β) υποκύπτοντας στην κοινωνική πίεση που τους ασκείται από τους υπόλοιπους.
Στο μοντέλο του Cohen[1], οι καταστάσεις 1 και 2 απορροφούν, ενώ οι μεταβάσεις μπορούν να συμβούν από την κατάσταση 3 στην 1 ή 4 και από την κατάσταση 4 στην 3 ή 2.

[1] Cohen, B. P. (1958). A Probability Model for Conformity. Sociometry, 21(1), 69–81. https://doi.org/10.2307/2786059

Χρησιμοποιώντας τα πειραματικά δεδομένα, ο Cohen υπολόγισε τον πίνακα μετάβασης της διεργασίας.

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0.06 & 0 & 0.63 & 0.31 \\ 0 & 0.05 & 0.46 & 0.49 \end{bmatrix}$$

Μακροπρόθεσμα, κάθε υποκείμενο γίνεται τελικά είτε μόνιμος μη κομφορμιστής (κατάσταση 1) είτε μόνιμος κομφορμιστής (κατάσταση 2).




Η θεωρία Μαρκοβιανών αλυσίδων μπορεί να μας ενημερώσει για τις πιθανότητες υιοθέτησης μίας από τις δύο απορροφητικές καταστάσεις, ανάλογα με την αρχική στάση ενός υποκειμένου.

Εάν η αλυσίδα ξεκινήσει στην 1η μεταβατική κατάσταση (κατάσταση 3), περιμένουμε ότι η αλυσίδα θα περάσει 11,06 περιόδους σε αυτήν την κατάσταση και 6,72 περιόδους στη 2η μεταβατική κατάσταση (κατάσταση 4) πριν να αφιχθεί στην 1η ή τη 2η απορροφητική κατάσταση αντίστοιχα.

Αντίστοιχα, εάν η αλυσίδα ξεκινήσει στη 2η μεταβατική κατάσταση (κατάσταση 4), μπορούμε να περιμένουμε ότι θα περάσει 9,98 περιόδους στην 1η μεταβατική κατάσταση και 8,03 περιόδους στη 2η μεταβατική κατάσταση πριν να αφιχθεί στην 1η ή τη 2η απορροφητική κατάσταση

**Κώδικας Gnu Octave (ή Matlab)**

```
>> P = [1 0 0 0; 0 1 0 0; 0.06 0 0.63 0.31; 0 0.05 0.46 0.49]    % Ο πίνακας μετάβασης της διεργασίας
>> R = P(3:4, 1:2)  % Ο πίνακας R που περιέχει τις πιθανότητες μετάβασης από τις μεταβατικές στις απορροφητικές.
>> Q = P(3:4, 3:4)  % Ο πίνακας Q που περιέχει τις πιθανότητες μετάβασης μεταξύ των μεταβατικών καταστάσεων.
>> N = inv(eye(2) – Q)          % Ο βασικός πίνακας της διεργασίας

>> sum(N, 2)     % Αναμενόμενο πλήθος μεταβάσεων πριν την άφιξη σε οποιαδήποτε απορροφητική κατάσταση.
>> N * R   % πιθανότητες μετάβασης από τις μεταβατικές στις απορροφητικές καταστάσεις
```

Εάν η αλυσίδα ξεκινήσει στην 1η μεταβατική κατάσταση (κατάσταση 3), περιμένουμε ότι η αλυσίδα θα αφιχθεί σε κάποια από τις δύο απορροφητικές καταστάσεις ύστερα από 17,8 περιόδους.

Αντίστοιχα, εάν η αλυσίδα ξεκινήσει στη 2η μεταβατική κατάσταση (κατάσταση 4), η άφιξη σε κάποια απορροφητική κατάσταση θα συμβεί ύστερα από 18 περιόδους.

Εάν το υποκείμενο ήταν αρχικά προσωρινός αντικομφορμιστής (κατάσταση 3), τότε με 66,38% πιθανότητα καταλήγει μόνιμος αντικομφορμιστής ενώ με 33,62% πιθανότητα καταλήγει μόνιμος κομφορμιστής.

Εάν αρχικά ήταν προσωρινά κομφορμιστής (κατάσταση 4), έχει 59,87% πιθανότητες να καταλήξει μόνιμος αντικομφορμιστής και 40,13% πιθανότητα να καταλήξει μόνιμος κομφορμιστής.

**Παράδειγμα 2**
Ένας τζογαδόρος έχει 3.000 € και αποφασίζει να παίξει 1.000 € τη φορά σε ένα τραπέζι Black Jack σε ένα καζίνο. Έχει αποφασίσει πως θα συνεχίζει να παίζει έως ότου χρεοκοπήσει ή κερδίσει 5.000 €. Η πιθανότητα να κερδίσει στο Black Jack είναι 0,40.
Να βρεθεί ο πίνακας μετάβασης και η πιθανότητα ότι ο τζογαδόρος θα χρεοκοπήσει
(α) έχοντας το αρχικό ποσό των 3.000 €
(β) με την υπόθεση ότι έχει βρεθεί να έχει 2.000 €.

**Λύση**
Πίνακας μετάβασης P =

$$\begin{array}{c} \\ 0 \\ 1K \\ 2K \\ 3K \\ 4K \\ 5K \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1K & 2K & 3K & 4K & 5K \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ .60 & 0 & .40 & 0 & 0 & 0 \\ 0 & .60 & 0 & .40 & 0 & 0 \\ 0 & 0 & .60 & 0 & .40 & 0 \\ 0 & 0 & 0 & .60 & 0 & .40 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Κανονική μορφή P =

$$\begin{array}{c} \\ 0 \\ 5K \\ 1K \\ 2K \\ 3K \\ 4K \end{array} \begin{array}{c} \begin{array}{cc|cccc} 0 & 5K & 1K & 2K & 3K & 4K \end{array} \\ \left[\begin{array}{cc|cccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline .60 & 0 & 0 & .40 & 0 & 0 \\ 0 & 0 & .60 & 0 & .40 & 0 \\ 0 & 0 & 0 & .60 & 0 & .40 \\ 0 & .40 & 0 & 0 & .60 & 0 \end{array}\right] \end{array}$$

**R** **Q**

Βασικός πίνακας **N** = (**I** – **Q**)$^{-1}$ =

$$\begin{array}{c} \\ 1K \\ 2K \\ 3K \\ 4K \end{array} \begin{array}{c} \begin{array}{cccc} 1K & 2K & 3K & 4K \end{array} \\ \begin{bmatrix} 1.54 & .90 & .47 & .19 \\ 1.35 & 2.25 & 1.18 & .47 \\ 1.07 & 1.78 & 2.25 & .90 \\ .64 & 1.07 & 1.35 & 1.54 \end{bmatrix} \end{array}$$

Πίνακας λύσεων **N** · **R**

$$\begin{bmatrix} 1.54 & .90 & .47 & .19 \\ 1.35 & 2.25 & 1.18 & .47 \\ 1.07 & 1.78 & 2.25 & .90 \\ .64 & 1.07 & 1.35 & 1.54 \end{bmatrix} \begin{bmatrix} .6 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & .4 \end{bmatrix} = \begin{bmatrix} .92 & .08 \\ .81 & .19 \\ .64 & .36 \\ .38 & .62 \end{bmatrix} =$$

Ο πίνακας **N** · **R** περιέχει τις πιθανότητες της απορρόφησης στην κατάσταση 0 € ή στην κατάσταση 5.000 € ξεκινώντας από οποιαδήποτε από τις τέσσερις μεταβατικές καταστάσεις (1K, 2K, 3K, 4K).
Εάν ο παίκτης έχει 3.000 €, τότε η πιθανότητα οικονομικής καταστροφής του(3.000 € → 0 €) είναι 64% και η πιθανότητα να φτάσει τα 5.000 € (3.000 € → 5.000 €) είναι 36%.
Εάν ο παίκτης έχει 2.000 €, τότε η πιθανότητα οικονομικής καταστροφής του(2.000 € → 0 €) είναι 81% και η πιθανότητα να φτάσει τα 5.000 € (2.000 € → 5.000 €) είναι 19%.

**Παράδειγμα 3**
Οι σπουδές ενός φοιτητή στο πανεπιστήμιο μπορούν να περιγραφούν σαν μια Μαρκοβιανή αλυσίδα με 6 καταστάσεις όπου η κάθε κατάσταση περιγράφεται παρακάτω:
Κατάσταση 1: ο φοιτητής εγκαταλείπει το πανεπιστήμιο.
Κατάσταση 2: ο φοιτητής παίρνει το πτυχίο του.
Κατάσταση 3: ο φοιτητής σπουδάζει στο 4$^{ο}$ έτος.
Κατάσταση 4: ο φοιτητής σπουδάζει στο 3$^{ο}$ έτος.
Κατάσταση 5: ο φοιτητής σπουδάζει στο 2$^{ο}$ έτος.
Κατάσταση 6: ο φοιτητής σπουδάζει στο 1$^{ο}$ έτος.
Ένας φοιτητής, κάθε χρονιά είτε συνεχίζει τις σπουδές του στο επόμενο έτος, είτε επαναλαμβάνει τα μαθήματα του ίδιου έτους, είτε εγκαταλείπει τις σπουδές του.
Αν p είναι η πιθανότητα εγκατάλειψης των σπουδών σε κάθε έτος, q η πιθανότητα επανάληψης του ίδιου έτους και r η πιθανότητα παρακολούθησης του επόμενου έτους τότε ο πίνακας μετάβασης της διεργασίας είναι:

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ p & r & q & 0 & 0 & 0 \\ p & 0 & r & q & 0 & 0 \\ p & 0 & 0 & r & q & 0 \\ p & 0 & 0 & 0 & r & q \end{pmatrix} \end{array}$$

R Q

Ο βασικός πίνακας N και ο πίνακας N · R που περιέχει τις πιθανότητες μετάβασης της διεργασίας είναι:

$$N = \frac{1}{(p+r)}\begin{pmatrix} 1 & 0 & 0 & 0 \\ t & 1 & 0 & 0 \\ t^2 & t & 1 & 0 \\ t^3 & t^2 & t & 1 \end{pmatrix} \qquad \begin{pmatrix} 1-t & t \\ 1-t^2 & t^2 \\ 1-t^3 & t^3 \\ 1-t^4 & t^4 \end{pmatrix}$$

όπου t = r / (p + r).
Εάν θέσουμε
p = 20%: πιθανότητα εγκατάλειψης των σπουδών σε κάθε έτος,
q = 10%: η πιθανότητα επανάληψης του ίδιου έτους,
r = 70%: η πιθανότητα παρακολούθησης του επόμενου έτους,
τότε t = r / (p + r) = 7/9 = 0,7777… και

$$N = \begin{pmatrix} 1.11 & 0 & 0 & 0 \\ 0.86 & 1.11 & 0 & 0 \\ 0.67 & 0.86 & 1.11 & 0 \\ 0.52 & 0.67 & 0.86 & 1.11 \end{pmatrix}, \quad N \cdot R = \begin{array}{c} \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \begin{array}{c} \begin{array}{cc} 1 & 2 \end{array} \\ \begin{pmatrix} 0.22 & 0.78 \\ 0.40 & 0.60 \\ 0.53 & 0.47 \\ 0.63 & 0.37 \end{pmatrix} \end{array}$$

Ενδεικτικά:

- Η πιθανότητα ένας 2ετής (κατάσταση 5) να πάρει πτυχίο (κατάσταση 2) μετά από πεπερασμένο πλήθος ετών είναι 47%.
- Η πιθανότητα ένας 2ετής (κατάσταση 5) να εγκαταλείψει τις σπουδές του(κατάσταση 1) μετά από πεπερασμένο πλήθος ετών είναι 53%.

**Κώδικας Gnu Octave (ή Matlab)**

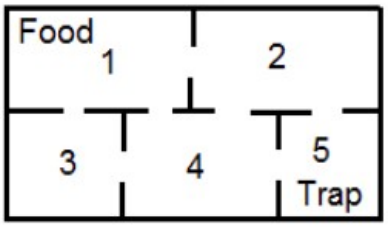
```
>> p =0.2; q = 0.1; r = 0.7
>> P = [1 0 0 0 0 0;0 1 0 0 0 0; p r q 0 0 0; p 0 r q 0 0; p 0 0 r q 0;p 0 0 0 r q]
>> R = P(3:6, 1:2)  % O πίνακας R που περιέχει τις πιθανότητες μετάβασης από τις μεταβατικές στις απορροφητικές.
>> Q = P(3:6, 3:6)  % O πίνακας Q που περιέχει τις πιθανότητες μετάβασης μεταξύ των μεταβατικών καταστάσεων.
>> N = inv(eye(4) – Q)  %  O βασικός πίνακας της διεργασίας
>> sum(N, 2)  % Αναμενόμενο πλήθος ετών πριν την αποφοίτηση ή την εγκατάλειψη των σπουδών.
>> N * R   % ποσοστά0 0αποφοίτησης ή εγκατάλειψης σπουδών από οποιοδήποτε έτος0.
```

**Κώδικας Python**

```
import numpy as np
p = 0.2; q = 0.1; r = 0.7
P = np.array([[1, 0, 0, 0, 0, 0], [0, 1, 0, 0, 0, 0], [p, r, q, 0, 0, 0],
                [p, 0, r, q, 0, 0], [p, 0, 0, r, q, 0],  [p, 0, 0, 0, r, q]])
R = P[2:6,0:2]
Q = P[2:6,2:6]
N = np.linalg.inv(np.identity(4) - Q)
print(np.dot(N, R))
```

**Παράδειγμα 4**

Ένα ποντίκι τοποθετείται στον λαβύρινθο που φαίνεται παρακάτω και μετακινείται από δωμάτιο σε δωμάτιο τυχαία. Από οποιοδήποτε δωμάτιο, το ποντίκι θα επιλέξει μια πόρτα στο επόμενο δωμάτιο με ίσες πιθανότητες. Μόλις το ποντίκι φτάσει στο δωμάτιο 1, βρίσκει φαγητό και δεν φεύγει ποτέ από αυτό το δωμάτιο. Και όταν φτάσει στο δωμάτιο 5, παγιδεύεται και δεν μπορεί να φύγει από αυτό το δωμάτιο. Ποια είναι η πιθανότητα το ποντίκι να καταλήξει στο δωμάτιο 5, αν είχε αρχικά τοποθετηθεί στο δωμάτιο 3;




**Κώδικας Gnu Octave (ή Matlab)**

```
>> P = [1 0 0 0 0; 1/3 0 0 1/3 1/3; 1/2 0 0 1/2 0;1/4 1/4 1/4 0 1/4;0 0 0 0 1]
>> P2 = [1 0 0 0 0; 0 1 0 0 0; 1/2 0 0 1/2 0;1/4 1/4 1/4 0 1/4;1/3 1/3 0 1/3 0]
>> R = P2(3:5, 1:2)
>> Q = P2(3:5, 3:5)
>> N = inv(eye(3) – Q)  %  Ο βασικός πίνακας της διεργασίας
>> sum(N, 2)  % Αναμενόμενο πλήθος κινήσεων πριν την άφιξη στην τροφή ή την
παγίδα.
>> N * R   % πιθανότητα άφιξης στην τροφή ή στην παγίδα.
ans =
   0.78947   0.21053
   0.57895   0.42105
   0.52632   0.47368
```

Η πιθανότητα το ποντίκι να καταλήξει στο δωμάτιο 5, αν είχε αρχικά τοποθετηθεί στο δωμάτιο 3 είναι **21,1%**.

**Παράδειγμα 5**
Δύο ισοδύναμοι τενίστες, ο Andre και ο Vilay ο καθένας με 2 € στην τσέπη, αποφασίζουν να στοιχηματίσουν από 1 €, για κάθε παιχνίδι που παίζουν. Συνεχίζουν να παίζουν μέχρι να χρεοκοπήσει ένας από αυτούς.
(α) Γράψτε τον πίνακα μετάβασης για τον Andre.
(β) Προσδιορίστε τις απορροφητικές καταστάσεις.
(γ) Γράψτε τον πίνακα λύσης.
(δ) Αν ο Andre έχει 1 €, ποια είναι η πιθανότητα να τα χάσει όλα τελικά;

**Κώδικας Gnu Octave (ή Matlab)**

```
>> P = [1 0 0 0 0; 1/2 0 1/2 0 0; 0 1/2 0 1/2 0; 0 0 1/2 0 1/2; 0 0 0 0 1]
>> P2 = [1 0 0 0 0; 0 1 0 0 0; 1/2 0 0 1/2 0;0 0 1/2 0 1/2;0 1/2 0 1/2 0]
>> R = P2(3:5, 1:2)
>> Q = P2(3:5, 3:5)
>> N = inv(eye(3) – Q)  %  Ο βασικός πίνακας της διεργασίας
>> sum(N, 2)  % Αναμενόμενο πλήθος κινήσεων πριν την άφιξη στην τροφή ή την
παγίδα.
>> N * R   % πιθανότητα άφιξης στην τροφή ή στην παγίδα.
ans =
   0.75000   0.25000
   0.50000   0.50000
   0.25000   0.75000
```

Αν ο Antre έχει 1 €, η πιθανότητα να τα χάσει όλα τελικά  είναι **75%**.

**Παράδειγμα 5 (εκδοχή β)**
Δύο ισοδύναμοι τενίστες, ο Andre και ο Vilay ο καθένας με 2 € στην τσέπη, αποφασίζουν να στοιχηματίσουν από 1 €, για κάθε παιχνίδι που παίζουν. Συνεχίζουν να παίζουν μέχρι να χρεοκοπήσει ένας από αυτούς.
Η πιθανότητα νίκης για τον Andre είναι 0,4 και για τον Vilay 0,6.
(α) Γράψτε τον πίνακα μετάβασης για τον Andre.
(β) Προσδιορίστε τις απορροφητικές καταστάσεις.
(γ) Γράψτε τον πίνακα λύσης.
(δ) Αν ο Andre έχει 1 €, ποια είναι η πιθανότητα να τα χάσει όλα τελικά;

**Παράδειγμα 5 (εκδοχή γ)**
Αρχικά ο Andre έχει 3 € και ο Vilay έχει 2 €. Αποφασίζουν να στοιχηματίσουν από 1 €, για κάθε παιχνίδι που παίζουν. Συνεχίζουν να παίζουν μέχρι να χρεοκοπήσει ένας από αυτούς.
Η πιθανότητα νίκης για τον Andre είναι 0,4 και για τον Vilay 0,6.
(α) Γράψτε τον πίνακα μετάβασης για τον Andre.
(β) Προσδιορίστε τις απορροφητικές καταστάσεις.
(γ) Γράψτε τον πίνακα λύσης.
(δ) Αν ο Andre έχει41 €, ποια είναι η πιθανότητα να τα χάσει όλα τελικά;

**Παράδειγμα 6**
Για ένα νόμισμα γνωρίζουμε πως φέρνει κορώνα (Κ) με πιθανότητα 0,6 και γράμματα (Γ) με πιθανότητα 0,4. Το νόμισμα ρίχνεται 100 φορές. Αν λ το μήκος της μεγαλύτερης ακολουθίας συνεχόμενων Κ στις 100 ρίψεις, να υπολογιστεί η πιθανότητα ότι το λ υπερβαίνει το 5, δηλαδή P(λ > 5).

**Λύση**
Θεωρούμε τη Μαρκοβιανή διαδικασία $\{W_n\}$ όπου η $W_n$ μετράει το πλήθος των συνεχόμενων επιτυχιών (Κ) και παίρνει τις τιμές 0, 1, 2.., 5. Τις καταστάσεις 0, 1, ..., 4 τις θεωρούμε μεταβατικές και την κατάσταση 5 ως απορροφητική και αναζητούμε την πιθανότητα η αλυσίδα να απορροφηθεί στην κατάσταση 5 στις 100 επαναλήψεις του πειράματος (δηλαδή την πιθανότητα να υπάρξει τουλάχιστον μία σειρά από 5 επιτυχίες στις 100 επαναλήψεις). Η πιθανότητα επιτυχίας σε ένα πείραμα και άρα και η πιθανότητα να έχω μία παραπάνω επιτυχία είναι 0,6 ενώ 1-0,6 = 0,4 είναι η πιθανότητα να επιστρέψει στην 1η κατάσταση (0 επιτυχίες).

Ο 6 x 6 πίνακας μετάβασης της διαδικασίας είναι ο

$$P = \begin{bmatrix} 0,4 & 0,6 & 0 & 0 & 0 & 0 \\ 0,4 & 0 & 0,6 & 0 & 0 & 0 \\ 0,4 & 0 & 0 & 0,6 & 0 & 0 \\ 0,4 & 0 & 0 & 0 & 0,6 & 0 \\ 0,4 & 0 & 0 & 0 & 0 & 0,6 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Ο πίνακας $P^{100}$ περιέχει τις πιθανότητες μετάβασης μεταξύ όλων των καταστάσεων ($1^η$: 0 επιτυχίες έως 6η: 5 επιτυχίες) σε 100 βήματα. Εμάς, μας ενδιαφέρει η πιθανότητα μετάβασης στην απορροφητική κατάσταση 5 σε 100 βήματα, ξεκινώντας από την 1η κατάσταση (0 επιτυχίες). Αυτό υπολογίζεται από το γινόμενο

$$[1, 0, 0, ..., 0] \cdot P^{100}.$$

Ο κώδικας του Octave που το υπολογίζει είναι ο εξής:

```
N = 100
P = [0.4, 0.6, 0, 0, 0, 0; 0.4, 0, 0.6, 0, 0, 0;  0.4, 0, 0, 0.6, 0, 0;  0.4, 0, 0, 0, 0.6, 0;  0.4, 0, 0, 0, 0,
0.6;  0, 0, 0, 0, 0, 1]
[1, 0, 0, 0, 0, 0]*P^N
```

```
ans =
   1.0043e-02   6.2616e-03   3.9041e-03   2.4342e-03   1.5177e-03   9.7584e-01
```

**Αποτέλεσμα**
**Στις 100 ρίψεις θα υπάρχουν 5 συνεχόμενες κορώνες (Κ) με πιθανότητα 97,6%.**

Σημείωση: Εναλλακτικά, υπάρχει τύπος που αναπτύχθηκε από τον de Moivre το 1738 και οδηγεί στο ίδιο αποτέλεσμα. Ο τύπος όπως και μία απόδειξη αυτού είναι διαθέσιμη εδώ:
https://math.stackexchange.com/a/59749/664787

**10. Επαναληπτικές και παροδικές καταστάσεις**

Ορίζουμε $f_{ii}^{(n)} = P(X_n = i, X_{n-1} \neq i, X_{n-2} \neq i, \dots, X_1 \neq i \mid X_0 = i)$ και

$$f_{ii} = P(\text{υπάρχει } n \in N \text{ ώστε } X_n = i \mid X_0 = i) = P(U_{n\geq 1}\{X_n = i\} \mid X_0 = i) = \Sigma_{n\geq 1} f_{ii}^{(n)}$$

Η κατάσταση i ϵ S λέγεται **επαναληπτική (recurrent)** όταν $\mathbf{f_{ii} = 1}$**,** δηλαδή όταν η αλυσίδα, αν βρεθεί σε αυτήν, επιστρέφει κάποια στιγμή στο μέλλον.

Η κατάσταση i ϵ S λέγεται **μεταβατική ή παροδική (transient)** όταν $\mathbf{f_{ii} < 1}$**.**

Αποδεικνύεται ότι:

**Θεώρημα**

- Αν η κατάσταση i ϵ S είναι **επαναληπτική** και i ↔ j, τότε και η κατάσταση j είναι **επαναληπτική**.
- Αν η κατάσταση i ϵ S είναι **παροδική** και i ↔ j,, τότε και η κατάσταση j είναι **παροδική**.

Το τελευταίο θεώρημα υποδεικνύει ότι **οι έννοιες της επαναληπτικότητας και της περιοδικότητας είναι ουσιαστικά ιδιότητες των κλάσεων ισοδυναμίας** ως προς τη σχέση της επικοινωνίας.

Σημείωση: Μία απόδειξη για το θεώρημα είναι διαθέσιμη εδώ:
https://www.uni-ulm.de/fileadmin/website_uni_ulm/mawi.inst.110/lehre/ws13/Stochastik_II/Skript_2B.pdf
(Theorem 2.7.6)

**Ορισμός**

Μία κλάση ισοδύναμων καταστάσεων ως προς τη σχέση της επικοινωνίας λέγεται **επαναληπτική** όταν περιέχει τουλάχιστον μία επαναληπτική κατάσταση.

**Ορισμός**

Μία *αδιαχώριστη* Μαρκοβιανή αλυσίδα ονομάζεται **επαναληπτική** όταν περιέχει τουλάχιστον μία επαναληπτική κατάσταση.

Μία *αδιαχώριστη* Μαρκοβιανή αλυσίδα ονομάζεται **παροδική** όταν όλες οι καταστάσεις της είναι παροδικές.

**Παράδειγμα**

Αν $\{X_n, n \geq 0\}$ Μαρκοβιανή αλυσίδα με χώρο καταστάσεων S = {1, 2, 3, 4, 5} και πίνακα μετάβασης

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

Οι καταστάσεις 2, 3, 4, 5 επικοινωνούν μεταξύ τους και είναι μεταβατικές αφού η απορροφητική κατάσταση 1 είναι προσιτή από αυτές. Το σύνολο {2, 3, 4, 5} είναι αδιαχώριστο, αλλά όχι κλειστό.

Αποδεικνύεται ότι:

**Θεώρημα**

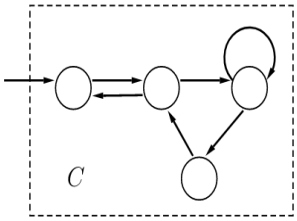
Η κατάσταση i ϵ S είναι **επαναληπτική** αν και μόνο αν $\Sigma_{n \geq 1} p_{ii}^{(n)} = +\infty$.

Η κατάσταση i ϵ S είναι **παροδική** αν και μόνο αν $\Sigma_{n \geq 1} p_{ii}^{(n)} < +\infty$.

Σημείωση: Μία απόδειξη για το θεώρημα είναι διαθέσιμη εδώ: https://www.uni-ulm.de/fileadmin/website_uni_ulm/mawi.inst.110/lehre/ws13/Stochastik_II/Skript_2B.pdf (Theorem 2.7.5) ή εδώ: https://www.sjsu.edu/faculty/guangliang.chen/Math263/lec3classification.pdf (σελ. 16).

**Ορισμός**

Μία κλάση καταστάσεων ονομάζεται **κλειστή (closed)** όταν το σύστημα δεν μπορεί να απομακρυνθεί από αυτήν σε περίπτωση που βρεθεί μέσα της.




**Θεώρημα**

Κάθε επαναληπτική κλάση είναι κλειστή.

Πράγματι, αν το σύστημα βρεθεί σε μία κατάσταση μίας επαναληπτικής κλάσης ισοδυναμίας ως προς τη σχέση της επικοινωνίας, τότε οποιαδήποτε μετακίνησή του σε άλλη κατάσταση θα είναι εξ ορισμού σε κατάσταση που θα επικοινωνεί με την πρώτη, δηλαδή πάλι σε κατάσταση της κλάσης. Συνεπώς, δεν θα μπορέσει ποτέ να βγει από αυτήν.

**Θεώρημα**

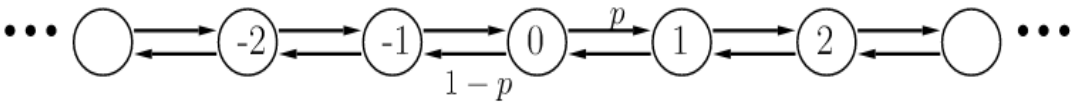
Κάθε πεπερασμένη κλειστή κλάση ισοδύναμων καταστάσεων είναι επαναληπτική.

Συμπεραίνουμε ότι σε μία πεπερασμένη Μαρκοβιανή αλυσίδα, ισχύει ότι:

- **Μία κλάση ισοδυναμίας είναι επαναληπτική αν και μόνον αν είναι κλειστή.**
- **Μία κλάση ισοδυναμίας είναι παροδική αν και μόνον αν δεν είναι κλειστή.**

Σημείωση: Μία απόδειξη είναι διαθέσιμη εδώ: https://www.uni-ulm.de/fileadmin/website_uni_ulm/mawi.inst.110/lehre/ws13/Stochastik_II/Skript_2B.pdf (Theorem 2.7.9 και Theorem 2.7.10.)

**Άσκηση 1**
Δίνεται η Μαρκοβιανή αλυσίδα




με $p_{i,i+1} = p$ και $p_{i,i-1} = 1 – p$. Όλες οι καταστάσεις επικοινωνούν μεταξύ τους άρα όλες τους είναι είτε επαναληπτικές είτε παροδικές. Βρείτε τι από τα δύο ισχύει συναρτήσει της πιθανότητας p.

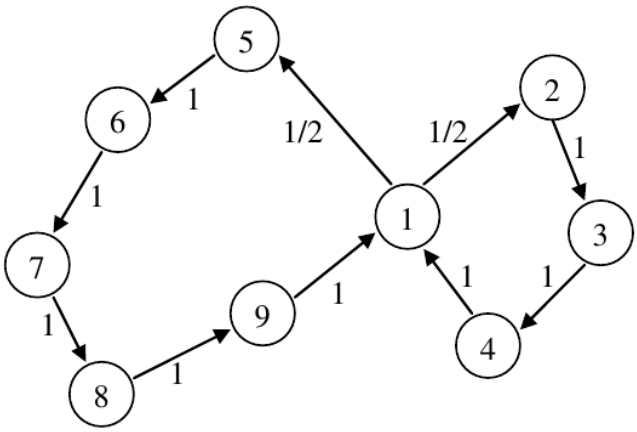
**Υπόδειξη:** Αρκεί να υπολογίσετε το $\Sigma_{n\geq 0} p_{00}^{(n)}$ και να βρείτε αν είναι πεπερασμένο ή άπειρο.

## 11. Περιοδικές καταστάσεις

**Ορισμός**
Περίοδος π(i) μιας κατάστασης i είναι ο μέγιστος κοινός διαιρέτης (ΜΚΔ) όλων των $n \geq 1$, για τους οποίους $p_{ii}^{(n)} > 0$.

Δηλαδή, η κατάσταση i έχει περίοδο d εάν οι παρακάτω δύο συνθήκες ισχύουν
(i) $p_{ii}^{(n)} = 0$ για κάθε n εκτός εάν n = md για κάποιο θετικό ακέραιο m.
(ii) d είναι ο μέγιστος ακέραιος με την ιδιότητα (i).

**Δραστηριότητα**
Να βρεθεί η περίοδος των καταστάσεων της παρακάτω αλυσίδας




Αν για την κατάσταση i είναι π(i) > 1, τότε η κατάσταση i λέγεται **περιοδική**.
Αν για την κατάσταση i είναι π(i) = 1, τότε η κατάσταση i λέγεται **απεριοδική**.

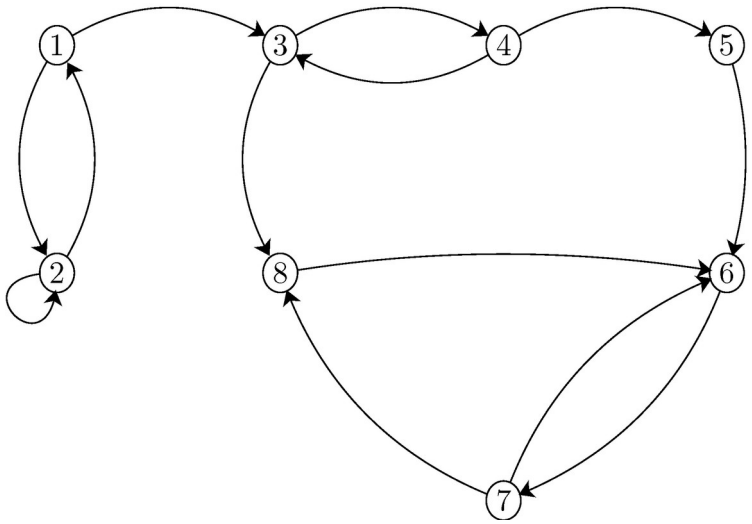
Ισχύει το εξής:

**Θεώρημα**
Εάν $i \leftrightarrow j$ τότε και π(i) = π(j).

- Μία **κλάση** ισοδύναμων καταστάσεων λέγεται **απεριοδική**, αν όλα τα στοιχεία της είναι απεριοδικά.
- Μία **Μαρκοβιανή αλυσίδα** λέγεται **απεριοδική**, αν όλα τα στοιχεία της είναι απεριοδικά.

Για να βρούμε αν μία αδιαχώριστη αλυσίδα είναι απεριοδική, αρκεί να βρούμε $p_{ii}^{(n)} > 0$ και $p_{ii}^{(m)} > 0$, με n, m πρώτους μεταξύ τους (ΜΚΔ(m,n) = 1).

**Δραστηριότητα**
Στην αλυσίδα του σχήματος, βρείτε αν οι κλάσεις {1, 2}, {3, 4}, {6, 7, 8} είναι περιοδικές ή απεριοδικές.




**Ορισμός**
Μία Μαρκοβιανή αλυσίδα που είναι αδιαχώριστη, επαναληπτική και απεριοδική ονομάζεται **εργοδική**.

**Δραστηριότητα**
Να δείξετε ότι η αλυσίδα του σχήματος είναι εργοδική

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

## 12. Στάσιμη συμπεριφορά

Αν P ο πίνακας μετάβασης μίας ΜΑΔΧ, $\pi_i = P(X_0 = i)$, i = 1, 2, .., n και $\pi = (\pi_0, \pi_1, \ldots, \pi_n)$, η κατανομή της μεταβλητής $X_0$, τότε:

- η κατανομή της $X_1$ είναι η $\pi^{(1)} = \pi \cdot P$.
- η κατανομή της $X_2$ είναι η $\pi^{(2)} = \pi \cdot P^{(2)} = \pi \cdot P^2$.
- η κατανομή της $X_n$ είναι η $\pi^{(n)} = \pi \cdot P^{(n)} = \pi \cdot P^n$.

Ένα ερώτημα που φυσιολογικά προκύπτει είναι το κατά πόσο μία Μαρκοβιανή αλυσίδα έχει στάσιμη συμπεριφορά, δηλαδή αν, καθώς το $n \to \infty$, η κατανομή $\pi^{(n)}$ συγκλίνει σε ένα διάνυσμα $\pi = (\pi_0, \pi_1, \pi_2, \ldots \pi_\kappa)$, όπου $\pi_j = \lim_{n \to \infty} P(X_n = j)$.

Φανερά, όταν υπάρχει το $P^\infty = \lim_{n\to\infty} P^{(n)}$, τότε υπάρχει οριακή (στάσιμη) συμπεριφορά και υπολογίζεται εύκολα, καθώς $\pi_j = \lim_{n \to \infty} [P(X_n = j)] = \pi_0 P^\infty$. Ωστόσο, είναι δυνατό να υπάρχει στάσιμη συμπεριφορά, χωρίς να υπάρχει το όριο $\lim_{n\to\infty} P^{(n)}$. Ένα παράδειγμα είναι η αλυσίδα με πίνακα μετάβασης

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

για την οποία δεν υπάρχει το $\lim_{n\to\infty} P^{(n)}$ αλλά φανερά $\pi = [½, ½]$.

Ανακύπτουν δύο ερωτήματα:

**Ερώτημα Α: Υπάρχει το διάνυσμα $\pi = (\pi_0, \pi_1, \pi_2, \ldots \pi_\kappa)$;**

**Απάντηση**

Αποδεικνύεται ότι για κάθε ομογενής Μαρκοβιανή αλυσίδα με πεπερασμένο χώρο καταστάσεων το διάνυσμα

$$\pi = (\pi_0, \pi_1, \pi_2, \ldots \pi_\kappa), \quad \pi_j = \lim_{n \to \infty} P(X_n = j), \quad j = 1, 2, \ldots, \kappa,$$

υπάρχει χωρίς ωστόσο να είναι απαραίτητα μοναδικό. Αν επιπλέον, η αλυσίδα είναι αδιαχώριστη και απεριοδική, τότε αποδεικνύεται πως είναι μοναδικό.

Σημειώσεις
1. Μία απόδειξη για την ύπαρξη είναι διαθέσιμη εδώ:
https://math.uchicago.edu/~may/REU2017/REUPapers/Freedman.pdf
ή εδώ: http://homepages.math.uic.edu/~furman/preprints/whatis-published.pdf
ή εδώ: https://mpaldridge.github.io/math2750/S10-stationary-distributions.html#exist-unique (Theorem 10.4).

Ειδικότερα, αν δεν υπάρχουν μηδενικές καταχωρήσεις είναι συνέπεια του Θεωρήματος Perron-Frobenius.

2. Αν υπάρχει παραπάνω από μία οριακή κατανομή τότε κάθε γραμμικός τους συνδυασμός είναι επίσης οριακή κατανομή, δηλαδή τελικά είναι άπειρες σε πλήθος.

3. Κάθε απορροφητική αλυσίδα έχει παραπάνω από μία οριακή κατανομή.

**Ερώτημα Β: Αν υπάρχει η οριακή κατανομή π, τότε πώς υπολογίζεται;**

**Απάντηση**

Θέτουμε $p_j^{(n)} = P(X_n = j)$ και υποθέτουμε ότι το διάνυσμα $\pi = (\pi_0, \pi_1, \pi_2, \dots \pi_\kappa)$ ορίζεται καλώς. Από τη σχέση

$$P(X_n = j) = \Sigma_{i \in S} P(X_n = j \mid X_{n-1} = i) \cdot P(X_{n-1} = i)$$

παίρνουμε

$$p_j^{(n)} = \Sigma_{i \in S}\, p_{ij}\, p_i^{(n-1)}$$

άρα θα είναι και $\lim_{n \to \infty} p_j^{(n)} = \lim_{n \to \infty} \Sigma_{i \in S}\, p_{ij}\, p_i^{(n-1)} = \Sigma_{i \in S}\, p_{ij} \lim_{n \to \infty} p_i^{(n-1)}$

ή $\pi_j = \Sigma_{i \in S}\, p_{ij} \cdot \pi_i = \Sigma_{i \in S}\, \pi_i \cdot p_{ij}$, $j \in S$,.

Βρήκαμε, πως αν $\boldsymbol{\pi} = [\pi_1, \pi_2, \dots]$, τότε

$$\boldsymbol{\pi} = \boldsymbol{\pi} \cdot \mathbf{P}$$

Δηλαδή: Η επίλυση του συστήματος των εξισώσεων

$$\boldsymbol{\pi} = \boldsymbol{\pi} \cdot \mathbf{P}$$

$$\Sigma_{j \in S} \pi_j = 1$$

μπορεί να οδηγήσει στον εντοπισμό της στάσιμης κατανομής (και άρα και των οριακών πιθανοτήτων), δηλαδή των πιθανοτήτων με τις οποίες λαμβάνεται κάθε διαφορετική κατάσταση καθώς το πλήθος των βημάτων τείνει στο άπειρο.

Επιπλέον, η κατανομή **π** είναι αυτή στην οποία αν βρεθεί το σύστημα τότε παραμένει σε αυτήν στο διηνεκές και για το λόγο αυτό **εκτός από οριακή είναι και στάσιμη** κατανομή.

Στην περίπτωση όπου υπάρχει στάσιμη κατανομή $\pi = (\pi_1, \pi_2, \dots, \pi_\kappa)$, αποδεικνύεται ότι ο λόγος

$$E_j = 1/\pi_j$$

είναι το μέσο πλήθος βημάτων που απαιτούνται για να επανέλθει το σύστημα στην κατάσταση j αν υποτεθεί ότι ξεκίνησε από αυτήν.

Σημείωση: Αν κατά μέσο όρο, η αλυσίδα βρίσκεται στην κατάσταση j μία φορά κάθε $E_j$ βήματα, πρέπει να είναι $\pi_j = 1 / E_j$, δηλαδή $E_j = 1/\pi_j$.

**Παράδειγμα 1**
Έστω $0 \le \alpha, \beta \le 1$ και $\{X_n, n \ge 0\}$ μια Αλυσίδα Markov με καταστάσεις $S = \{0, 1\}$ και πίνακα πιθανοτήτων μετάβασης

$$\begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$$

Να βρεθεί η στάσιμη κατανομή (αν υπάρχει).

**Λύση**

Είναι $\pi = \pi \cdot P$ ή $[\pi_0, \pi_1] = [(1-\alpha)\pi_0 + b\pi_1, \alpha\pi_0 + (1-b)\pi_1]$. Επιπλέον, $\pi_0 + \pi_1 = 1$.

Λύνουμε το σύστημα και βρίσκουμε $[\pi_0, \pi_1] = [b/(\alpha + b), \alpha/(\alpha + b)]$.

**Παρατήρηση**
Προσέξτε πως η στάσιμη κατανομή είναι η ίδια με αυτήν που αποδείξαμε ως οριακή σε προηγούμενο πρόβλημα.

**Παράδειγμα 2**

Έστω $S = \{0, 1\}$ και

$$P = \begin{pmatrix} 0.5 & 0.5 \\ 0.4 & 0.6 \end{pmatrix}$$

Αν $\pi = (\pi_0, \pi_1)$ η οριακή (στάσιμη) κατανομή πιθανοτήτων τότε

$$\left.\begin{array}{r} \pi = \pi \cdot P \\ \pi_0 + \pi_1 = 1 \end{array}\right\} \Rightarrow \left.\begin{array}{r} \pi_0 = 0{,}5\pi_0 + 0{,}4\pi_1 \\ \pi_1 = 0{,}5\pi_0 + 0{,}6\pi_1 \\ \pi_0 + \pi_1 = 1 \end{array}\right\} \Rightarrow \left.\begin{array}{r} \pi_0 = \frac{4}{9} \\ \pi_1 = \frac{5}{9} \end{array}\right\}$$

Δηλαδή, το σύστημα, ανεξάρτητα από την αρχική του κατάσταση, οριακά θα λαμβάνει την κατάσταση 0 με πιθανότητα 4/9 και την κατάσταση 1 με πιθανότητα 5/9.

Επιπλέον: Αν το σύστημα βρεθεί στην κατάσταση 1 τότε θα ξαναβρεθεί σε αυτήν μετά από 9/5 = 1,8 βήματα ενώ μεταξύ δύο επισκέψεων στην κατάσταση 0 μεσολαβούν 9/4 = 2,25 βήματα.

**Παράδειγμα 3**

Έστω S = {0, 1, 2} και $P = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/4 & 1/2 & 1/4 \\ 1/6 & 1/3 & 1/2 \end{bmatrix}$

(α) Να βρεθεί η στάσιμη κατανομή

(β) Να βρεθεί το μέσο πλήθος βημάτων που απαιτούνται μεταξύ δύο επισκέψεων στις καταστάσεις της αλυσίδας.

**Λύση**

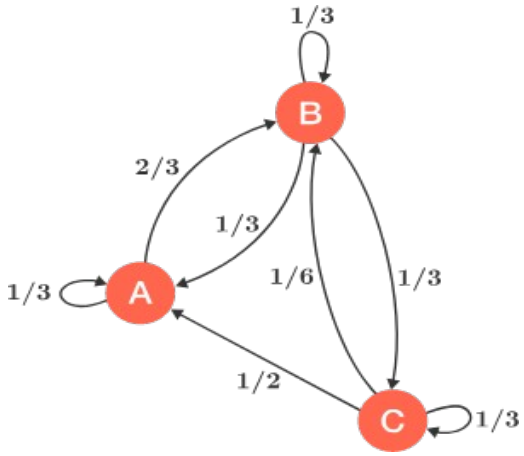
Αν π = ($\pi_0$, $\pi_1$, $\pi_2$) η οριακή στάσιμη κατανομή πιθανοτήτων τότε $\pi_0 + \pi_1 + \pi_2 = 1$ και

$$\begin{aligned} \pi_0 &= \frac{\pi_0}{3} + \frac{\pi_1}{4} + \frac{\pi_2}{6} \\ \pi_1 &= \frac{\pi_0}{3} + \frac{\pi_1}{2} + \frac{\pi_2}{3} \\ \pi_2 &= \frac{\pi_0}{3} + \frac{\pi_1}{4} + \frac{\pi_2}{2} \end{aligned}$$

Από την λύση του παραπάνω συστήματος προκύπτει ($\pi_0$, $\pi_1$, $\pi_2$) = (6/25, 2/5, 9/25).
Μεταξύ δύο επισκέψεων στην κατάσταση 0 μεσολαβούν 25/6 = 4,25 βήματα.
Μεταξύ δύο επισκέψεων στην κατάσταση 1 μεσολαβούν 5/2 = 2,5 βήματα.
Μεταξύ δύο επισκέψεων στην κατάσταση 2 μεσολαβούν 25/9 = 2,78 βήματα.

**Δραστηριότητα**

Να βρεθεί η στάσιμη κατανομή πιθανοτήτων της αλυσίδας που περιγράφεται από το διάγραμμα:



**Δραστηριότητα**

Στο διάγραμμα φαίνονται οι πιθανότητες μετάβασης ενός χρηματιστηρίου από την κατάσταση σημαντικής αύξησης (>20%) των τιμών (bull market) στην κατάσταση στασιμότητας (stagnant market) και στην κατάσταση σημαντικής μείωσης των τιμών (bear market). Να βρεθεί η στάσιμη κατανομή πιθανοτήτων της αλυσίδας.