Κεφάλαιο 3 : Μοντέλα Θεωρίας Αναμονής

Τεχνικές Εκτίμησης Υπολογιστικών Συστημάτων

Γιάννης Γαροφαλάκης

Καθηγητής

Ορισμός συστημάτων αναμονής

- **Συστήματα αναμονής** (Queueing Systems): Συστήματα στα οποία οι αφίξεις «πελατών» δημιουργούν απαιτήσεις εξυπηρέτησης από πόρους πεπερασμένης δυνατότητας εξυπηρέτησης.
- Σχηματίζονται «ουρές», όταν δημιουργούνται απαιτήσεις σύγχρονης χρησιμοποίησης πόρων.

Ορισμός συστημάτων αναμονής (2)

- Οι ουρές επηρεάζονται από τη *μέση τιμή* και τη *στατιστική διακύμανση του ρυθμού αφίξεων.*
  - Ανεξέλεγκτες ουρές όταν: *μέση τιμή ρυθμού αφίξεων μεγαλύτερη από μέγιστη δυνατότητα εξυπηρέτησης*
  - Σχηματισμός ουρών και λόγω στατιστικών διακυμάνσεων των αφίξεων
- ***Θεωρία αναμονής*** *(Queueing Theory): ασχολείται με τη μελέτη συστημάτων, η απόδοση των οποίων επηρεάζεται από φαινόμενα αναμονής.*

Φορτίο εργασίας συστημάτων αναμονής (Μη-εκτελέσιμο)

- *Συνάρτηση Κατανομής Πιθανότητας* των *χρόνων μεταξύ διαδοχικών αφίξεων* (Χ.Α.)

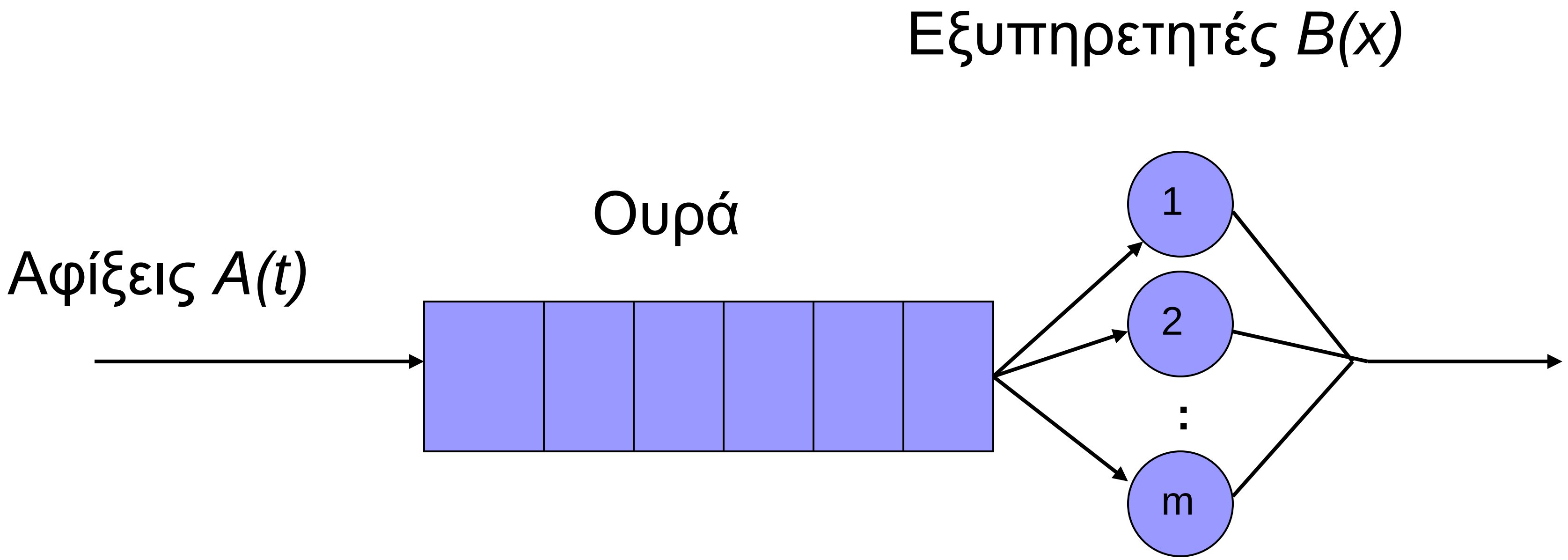
  $A(t)$ = Prob[χρόνος μεταξύ διαδοχικών αφίξεων $\leq t$]

- *Συνάρτηση Κατανομής Πιθανότητας* του *χρόνου εξυπηρέτησης ενός πελάτη* (Χ.Ε.)

  $B(x)$ = Prob[χρόνος εξυπηρέτησης $\leq x$]

- Συνήθως υποθέτουμε ότι οι παραπάνω Στοχαστικές Διαδικασίες (ΣΔ) συγκροτούνται από ανεξάρτητες, όμοια κατανεμημένες Τυχαίες Μεταβλητές (ΤΜ)

Ένα Σύστημα Αναμονής



Άλλα μεγέθη περιγραφής του συστήματος

- *Αριθμός εξυπηρετητών* (servers) στο σύστημα *m*.
- *Χωρητικότητα* του συστήματος (ουρά και servers) σε πελάτες *K* (default: *K* = ∞)
- *Πληθυσμός* υποψηφίων πελατών *M* (default: *M* = ∞)
- *Πολιτική εξυπηρέτησης*, δηλαδή ο τρόπος επιλογής πελατών από την ουρά για τον (τους) εξυπηρετητές. (default: FCFS ή FIFO)
- *Κλάσεις* πελατών (default: 1)
- *Ομάδες προτεραιότητας* πελατών (default: 1)
- *Διαθεσιμότητα* εξυπηρετητή (default: 100%)

Μετρικές απόδοσης

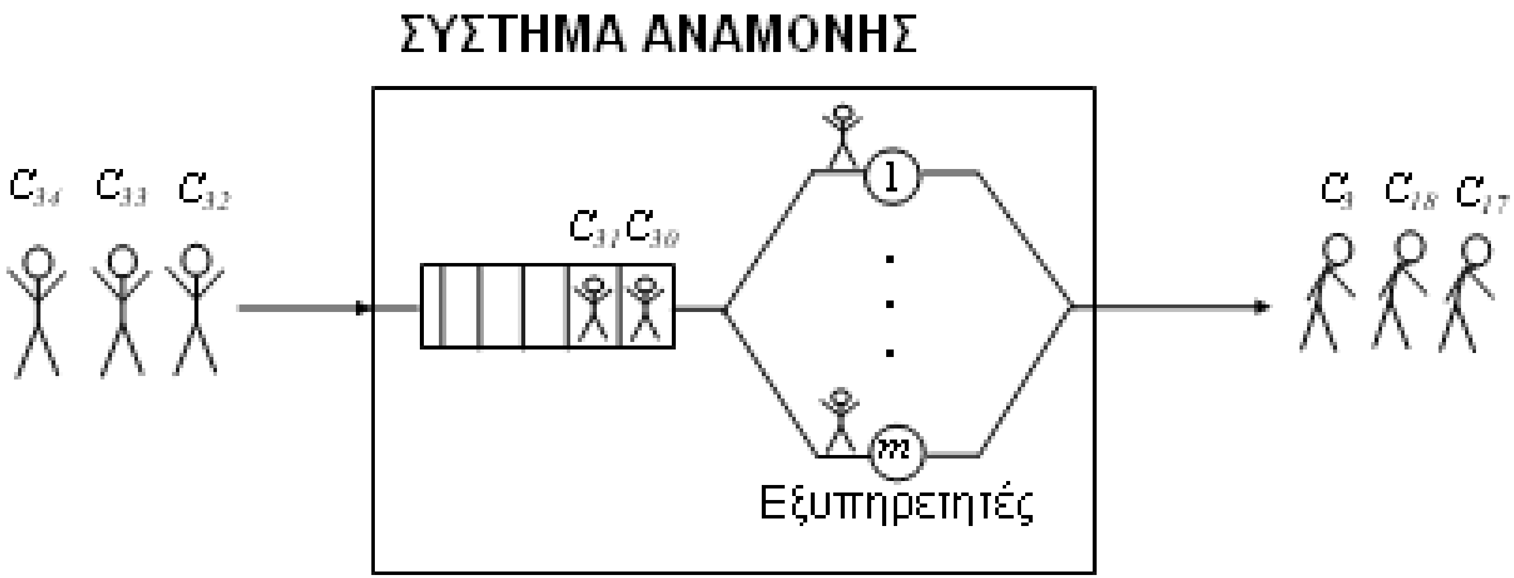
- *Χρόνος απόκρισης – response time* (συνολικός χρόνος στο σύστημα) για ένα πελάτη.
- *Χρόνος αναμονής* για ένα πελάτη.
- *Ρυθμαπόδοση* (throughput)
- *Αριθμός πελατών* στο σύστημα.
- *Χρησιμοποίηση* (Utilization) του συστήματος.

Συμβολισμός συστημάτων αναμονής

- *A/B/m*
  - A , B: Συναρτήσεις κατανομής πιθανότητας Χ.Α και Χ.Ε αντίστοιχα. Εκφράζονται ως
    - *M* (για την εκθετική κατανομή).
    - *D* (για τη ντετερμινιστική [σταθερή] κατανομή).
    - *Er* (για την κατανομή Erlang r-βαθμίδων).
    - *G* (για ΟΠΟΙΑΔΗΠΟΤΕ ΚΑΤΑΝΟΜΗ)
  - m: αριθμός εξυπηρετητών
- *A/B/m/K/M*
  - *K* : η χωρητικότητα του συστήματος
  - *M* : το μέγεθος του πληθυσμού των πελατών

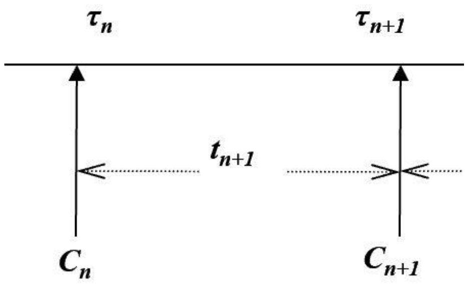
  όταν αυτά είναι διαφορετικά από ∞
- Παράδειγμα: D/M/2//200

Αναπαράσταση συστήματος αναμονής



- $A(t)$, $B(x)$: αυθαίρετα
- m εξυπηρετητές
- Αριθμούμε τους πελάτες με το δείκτη n και ορίζουμε C_n τον n-οστό πελάτη που εισέρχεται στο σύστημα

Συμβολισμοί βασικών μεγεθών



- **Χρόνοι μεταξύ διαδοχικών αφίξεων:**

$\tau_n \equiv$ χρονική στιγμή άφιξης του πελάτη C_n

$t_n \equiv$ χρόνος μεταξύ των αφίξεων των C_{n-1}, C_n

$= \tau_n - \tau_{n-1}$ για $n \geq 2 \quad (t_1 = \tau_1)$

Έστω ότι υπάρχει το: $\tilde{t} \triangleq \lim_{n\to\infty} t_n$

Δηλαδή το $\tilde{\boldsymbol{t}}$ είναι η Σ.Δ. του χρόνου μεταξύ διαδοχικών αφίξεων, ανεξάρτητο από το *n*.

$$\boldsymbol{A(t) = A_{\tilde{t}}(t) = \mathbf{Prob}[\tilde{t} \leq t]}$$

- $\bar{t}$ *μέσος χρόνος μεταξύ διαδοχικών αφίξεων*
- *Ρυθμός αφίξεων* (arrival rate) των πελατών: $\lambda = \dfrac{1}{\bar{t}}$

Συμβολισμοί βασικών μεγεθών (2)

- **Χρόνοι εξυπηρέτησης:**

  $x_n \equiv$ χρόνος εξυπηρέτησης του C_n . Και εδώ: $\tilde{x} \triangleq \lim_{n\to\infty} x_n$

  Prob[$\tilde{x}$ ≤ *x*] = *B(x)*

  $\overline{x}$: *μέσος χρόνος εξυπηρέτησης*

  Ρυθμός εξυπηρέτησης (service rate) των πελατών : $\mu = \dfrac{1}{\overline{x}}$

- **Χρόνος αναμονής ενός πελάτη στην ουρά:**

  $w_n \equiv$ *χρόνος αναμονής* (στην ουρά) του C_n. ($\tilde{w} \triangleq \lim_{n\to\infty} w_n$)

  $W = \overline{w}$ *μέσος χρόνος αναμονής*

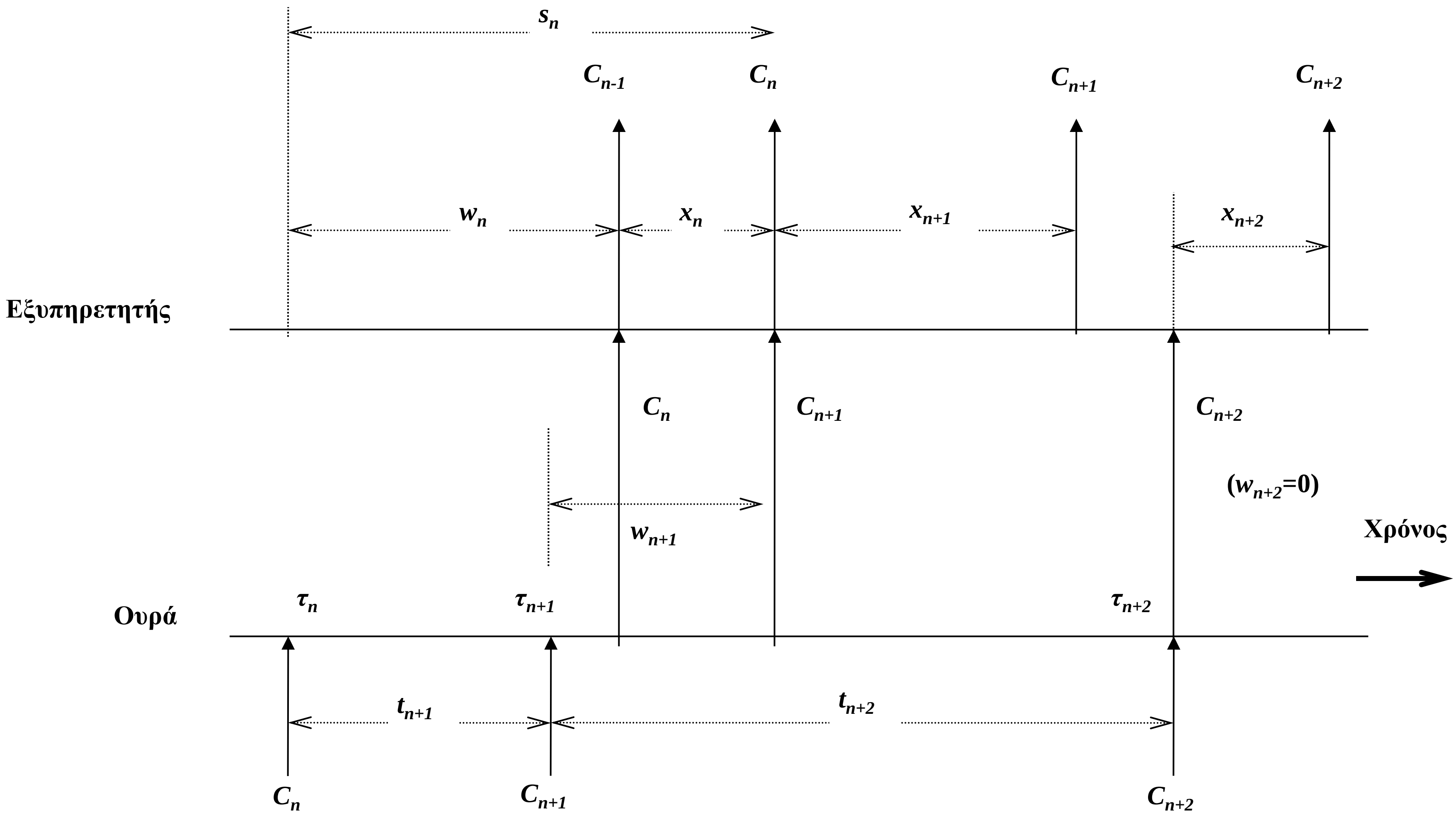
- **Συνολικός χρόνος ενός πελάτη στο σύστημα (χρόνος απόκρισης):**

  $s_n \equiv$ χρόνος συστήματος (ουρά + εξυπηρέτηση) του C_n

  $= w_n + x_n$ ($\tilde{s} \triangleq \lim_{n\to\infty} s_n$)

  $T = W + \overline{x}$ *μέσος χρόνος συστήματος* ($T \equiv \overline{s}$)

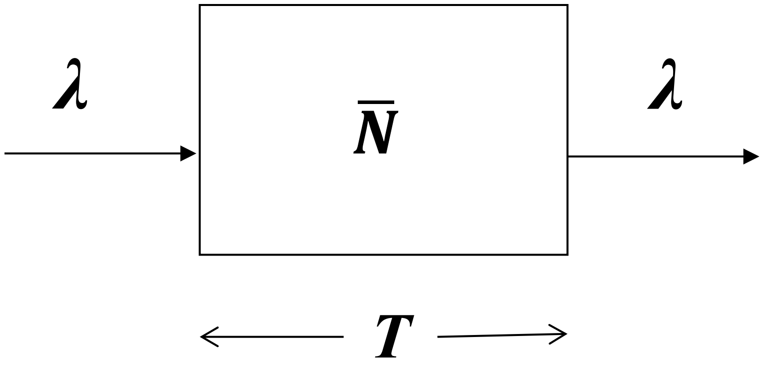
Χρονικό Διάγραμμα Συστήματος Αναμονής (1 εξυπηρετητής – FCFS)



Νόμος του Little

- Ο μέσος αριθμός πελατών σε ένα σύστημα αναμονής είναι ίσος με το μέσο ρυθμό αφίξεων πελατών στο σύστημα επί το μέσο χρόνο που ξοδεύει ένας πελάτης σ’ αυτό.

$$\bar{N} = \lambda \cdot T$$

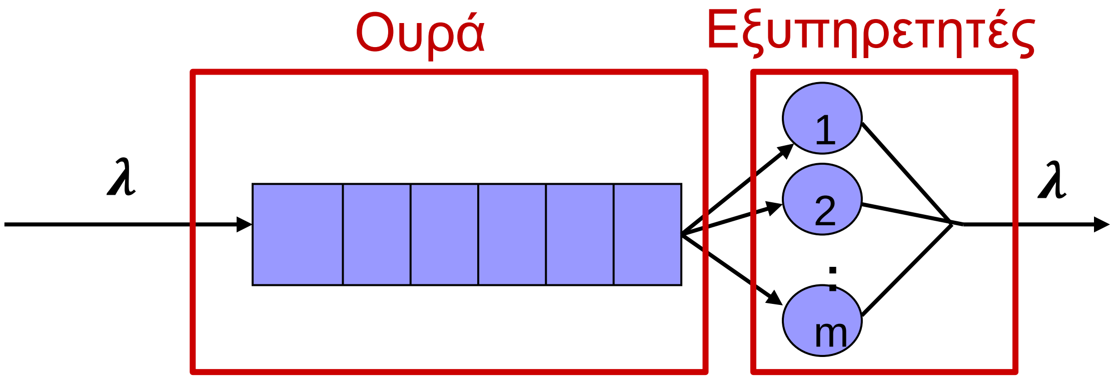


Νόμος του Little (2)

- *Διαισθητική αιτιολόγηση*: ένας πελάτης που φθάνει στο σύστημα θα *βρει* μέσα κατά μέσο όρο τον ίδιο αριθμό πελατών $\overline{N}$ που θα *υπάρχει* όταν φύγει. Όμως κατά το διάστημα της παρουσίας του ήρθαν $\lambda \cdot T$ πελάτες κατά μέσο όρο. Η τελευταία ποσότητα είναι οι πελάτες που αφήνει πίσω φεύγοντας.

- Ο Νόμος δίνει μια χρήσιμη σχέση μεταξύ ορισμένων βασικών μεγεθών ενός συστήματος αναμονής, αλλά δεν αποτελεί «λύση» στο γενικό μας πρόβλημα: Ουσιαστικά συνδέει ένα γνωστό μέγεθος εισόδου (λ), με δύο άγνωστα μεγέθη ($\overline{N}$, $\overline{T}$) τα οποία είναι μετρικές απόδοσης που *θέλουμε να βρούμε*.

Νόμος του Little (3)



- Για όρια του συστήματος μόνο στην ουρά

$$\bar{N}_q = \lambda \cdot W$$

- Για όρια συστήματος μόνο στον(-ους) εξυπηρετητή(-ές)

$$\bar{N}_s = \lambda \cdot \bar{x}$$

Χρησιμοποίηση (Utilization)

- Ο **συντελεστής απασχόλησης** ή **χρησιμοποίηση** *ρ*, ορίζεται ως ο λόγος του ρυθμού με τον οποίο εισέρχεται «δουλειά» στο σύστημα, προς το **μέγιστο** ρυθμό με τον οποίο το σύστημα μπορεί να εκτελέσει αυτή τη «δουλειά». Δηλαδή για 1 εξυπηρετητή:

ρ = (μέσος ρυθμός αφίξεων πελατών) x (μέσος χρόνος εξυπηρέτησης) / 1

$\rightarrow$ $\boxed{\rho = \lambda \cdot \bar{x}}$ για σύστημα με 1 εξυπηρετητή

- Στην περίπτωση *m* (ίδιων) εξυπηρετητών: $\boxed{\rho = \dfrac{\lambda \cdot \bar{x}}{m}}$
- *ρ* = {Μέση τιμή του ποσοστού εξυπηρετητών που είναι απασχολημένοι}. Διότι:

$$\rho = \frac{\lambda \cdot \bar{x}}{m} = \frac{\bar{N}_s}{m} \qquad \textit{(N. Little)}$$

Δηλαδή, για 1 εξυπηρετητή:
$\rho = \bar{N}_s$

Σταθερό σύστημα αναμονής

- **Σταθερό** σύστημα αναμονής, είναι αυτό στο οποίο δεν επιτρέπεται να δημιουργούνται ουρές ανεξέλεγκτου (άπειρου) μήκους.
- Σε ένα σταθερό σύστημα ισχύει $0 \leq \rho < 1$

G/G/1

- Έστω $\boldsymbol{\tau}$ ένα αυθαίρετα μεγάλο χρονικό διάστημα. Κατά τη διάρκεια αυτού του διαστήματος περιμένουμε ο αριθμός των αφίξεων **A** να είναι πολύ κοντά στην τιμή $\lambda \cdot \tau$ Επίσης, έστω p_0 η *πιθανότητα ο εξυπηρετητής να είναι άεργος* σε κάποιο τυχαία εκλεγμένο χρονικό διάστημα. Μπορούμε λοιπόν να πούμε ότι κατά τη διάρκεια του διαστήματος $\boldsymbol{\tau}$, ο εξυπηρετητής είναι απασχολημένος για $\tau - \tau \cdot p_0$ sec και άρα ο αριθμός των πελατών που εξυπηρετούνται **B** στο χρονικό διάστημα $\boldsymbol{\tau}$, είναι περίπου $\dfrac{(\tau - \tau \cdot p_0)}{\bar{x}}$
- **A** = **B**: $\lambda \cdot \tau \cong \dfrac{(\tau - \tau \cdot p_0)}{\bar{x}}$ οπότε για $\boldsymbol{\tau} \to \infty$, έχουμε: $\lambda\bar{x} = 1 - p_0$
- *Οπότε* $\boldsymbol{\rho = 1 - p_0}$

  όπου p_0 η *πιθανότητα ο εξυπηρετητής να είναι άεργος* σε κάποιο τυχαία εκλεγμένο χρονικό διάστημα

Στοχαστικές διαδικασίες

- *Στοχαστική Διαδικασία* (Σ.Δ.): ορίζεται ως μία οικογένεια Τυχαίων Μεταβλητών (Τ.Μ.), ***X(t)***, όπου οι Τ.Μ. έχουν δεικτοδοτηθεί με τη χρονική παράμετρο *t*.
- Παράγοντες ταξινόμησης στοχαστικών διαδικασιών
  - *ο χώρος καταστάσεων* (οι τιμές που παίρνουν οι ΤΜ)
    - πεπερασμένες ή αριθμήσιμες τιμές →Σ.Δ. *διακριτών-καταστάσεων* (*αλυσίδα*). *Ο χώρος καταστάσεων* ↔ {0, 1, 2, …}
    - τιμές από ένα πεπερασμένο ή άπειρο συνεχές διάστημα → Σ.Δ. *συνεχών-καταστάσεων*
  - η *χρονική παράμετρος* (επιτρεπτές χρονικές στιγμές αλλαγής κατάστασης)
    - Σ.Δ. Διακριτού-χρόνου [X_n – Στοχαστική Ακολουθία]
    - Σ.Δ. Συνεχούς χρόνου [X(t)]
  - η *στατιστική σχέση* μεταξύ των Τ.Μ.

Στατιστική σχέση μεταξύ των ΤΜ (1)

- Θέλουμε να προσδιορίσουμε την από κοινού PDF στις ΤΜ

  $\vec{X} = [\, X(t_1), X(t_2), \ldots]$, δηλαδή την:

$$F_{\vec{X}}\left(\vec{x}\,;\vec{t}\,\right) \equiv P\left[X\left(t_1\right) \le x_1,\ \ldots,\ X\left(t_n\right) \le x_n\right]$$

για όλα τα $\vec{x} = (x_1, x_2, \ldots, x_n)$, $\vec{t} = (t_1, t_2, \ldots, t_n)$ και n.

Στατιστική σχέση μεταξύ των ΤΜ (2)

1. Στάσιμες ΣΔ

   Αμετάβλητες στις ολισθήσεις στο χρόνο. Δηλαδή για οποιοδήποτε σταθερό *τ*, πρέπει: $F_{\vec{X}}\left(\vec{x};\vec{t}+\tau\right)=F_{\vec{X}}\left(\vec{x};\vec{t}\right)$

   όπου $\vec{t} + \tau = (t_1 + \tau, t_2 + \tau, \ldots, t_n + \tau)$.

2. Ανεξάρτητες ΣΔ

   Οι πιο απλές. Δεν υπάρχει καμία δομή ή εξάρτηση των Τ.Μ. τους:

$$f_{\vec{X}}\left(\vec{x}\,;\vec{t}\right) \equiv f_{X_1 \ldots X_n}\left(x_1,\ldots,x_n;t_1,\ldots,t_n\right) = f_{X1}\left(x_1\,;t_1\right)\ldots f_{Xn}\left(x_n\,;t_n\right)$$

Στατιστική σχέση μεταξύ των ΤΜ (3)

3. Διαδικασίες Markov

Για μια Αλυσίδα Markov $\{X(t)\}$, η πιθανότητα ότι η επόμενη τιμή $X(t_{n+1})$ θα είναι ίση με $\boldsymbol{x_{n+1}}$, εξαρτάται μόνο από την παρούσα τιμή $X(t_n) = \boldsymbol{x_n}$ και όχι από οποιαδήποτε προηγούμενη (ΙΔΙΟΤΗΤΑ ΑΜΝΗΣΙΑΣ).

Ιδιότητα Markov (για αλυσίδα Markov):

$$P\left[X(t_{n+1}) = x_{n+1} | X(t_n) = x_n, X(t_{n-1}) = x_{n-1}, \ldots, X(t_1) = x_1\right] =$$

$$= P\left[X(t_{n+1}) = x_{n+1} | X(t_n) = x_n\right]$$

όπου $t_1 < t_2 < \ldots < t_n < t_{n+1}$, ενώ τα x_i περιέχονται σε κάποιο διακριτό χώρο καταστάσεων.

- Θα αποδείξουμε αργότερα ότι ο **χρόνος παραμονής** σε μια κατάσταση ακολουθεί την:

  Εκθετική Κατανομή (διαδικασία συνεχούς χρόνου), ή την - ισοδύναμη -

  Γεωμετρική Κατανομή (διαδικασία διακριτού χρόνου).

Στατιστική σχέση μεταξύ των ΤΜ (4)

4. Διαδικασίες Γεννήσεων - Θανάτων

Κλάση των Διαδικασιών Markov: Οι αλλαγές κατάστασης γίνονται μόνο μεταξύ *γειτονικών* καταστάσεων.

Δηλαδή αν $X(t_n) = i$, τότε $X(t_{n+1}) = i - 1$ ή $X(t_{n+1}) = i + 1$ μόνο.

5. Διαδικασίες Semi Markov

- Επιτρέπουμε αυθαίρετη κατανομή του χρόνου που η διαδικασία μπορεί να παραμείνει σε μια κατάσταση.
- Η διαδικασία συμπεριφέρεται σαν Markov κατά τις χρονικές στιγμές αλλαγής κατάστασης, και στην πραγματικότητα *σε αυτές τις στιγμές* λέμε ότι έχουμε μια *συμπυκνωμένη (embedded) αλυσίδα Markov*.
- Υπερσύνολο των διαδικασιών Markov.

Στατιστική σχέση μεταξύ των ΤΜ (5)

6. Τυχαίοι περίπατοι

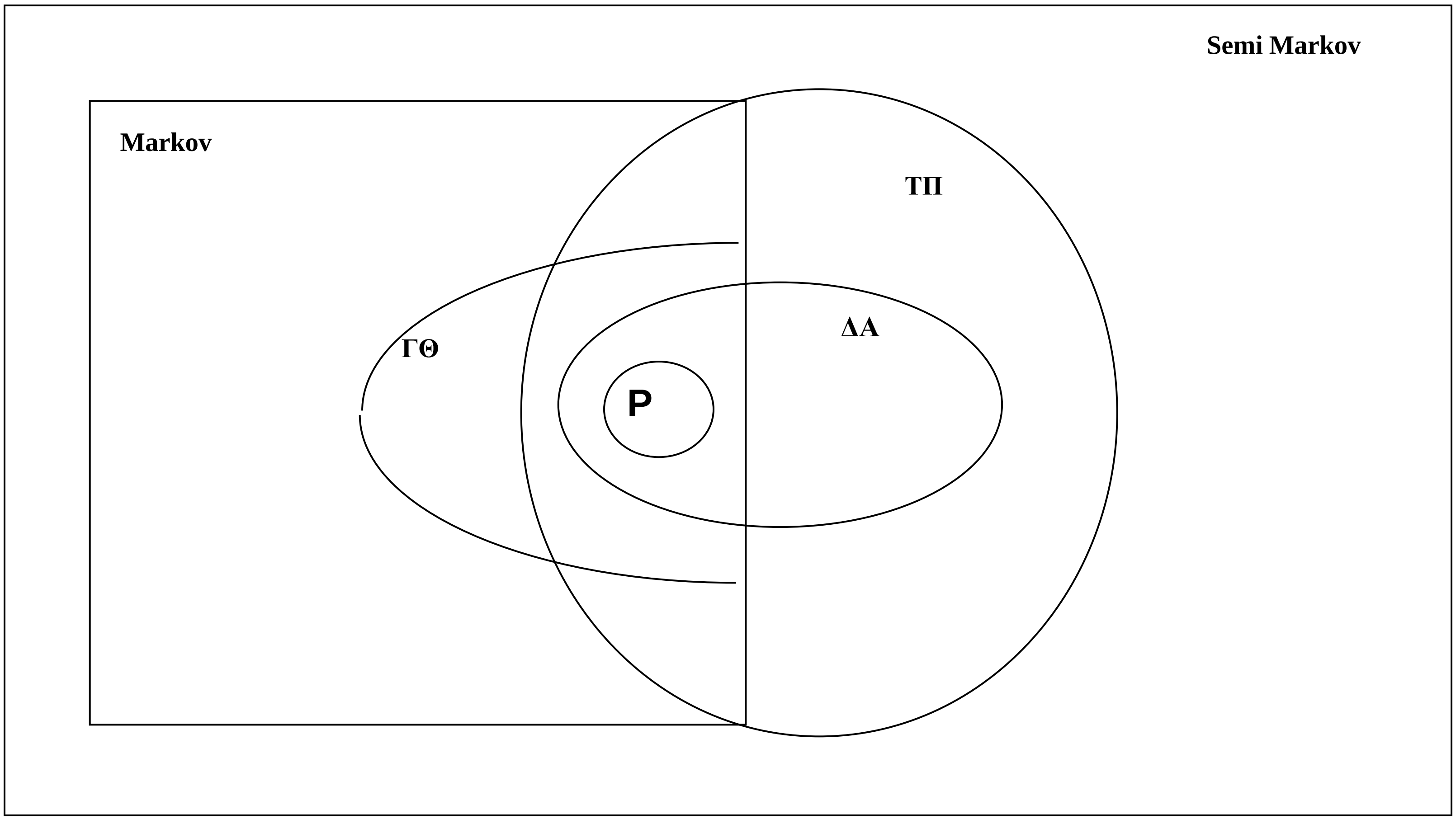
Η επόμενη θέση είναι ίση με την προηγούμενη θέση, συν μια τυχαία μεταβλητή

Δηλαδή, μια ακολουθία ΤΜ $\{Sn\}$ είναι τυχαίος περίπατος αν:
$S_n = X_1 + X_2 + \ldots + X_n$ όπου $n = 1, 2, \ldots,$ $S_0 = 0$ και $X_1, X_2, \ldots$ είναι ακολουθία ανεξάρτητων ΤΜ με κοινή κατανομή.

7. Διαδικασίες ανανέωσης

- Ειδική περίπτωση των τυχαίων περιπάτων.
- S_n είναι τώρα η ΤΜ που καθορίζει τη *χρονική στιγμή* στην οποία γίνεται η n-οστή μεταβολή κατάστασης και $\{X_n\}$ είναι ένα σύνολο ανεξάρτητων, όμοια κατανεμημένων ΤΜ, όπου η X_n αντιπροσωπεύει το χρόνο μεταξύ της $(n-1)$-οστής και n-οστής μεταβολής κατάστασης . Οι μεταβολές γίνονται μόνο μεταξύ γειτονικών καταστάσεων .

Σχέσεις των κλάσεων Στοχαστικών Διαδικασιών



P: Poisson

Αλυσίδες Markov διακριτού χρόνου

- Η ΣΔ καταλαμβάνει διακριτές θέσεις και οι αλλαγές μεταξύ αυτών των θέσεων γίνονται μόνο σε διακριτές χρονικές στιγμές
- Η υπό συνθήκη πιθανότητα να γίνει η μετάβαση της διαδικασίας από την *κατάσταση* E_i όπου είναι στο βήμα *(n-1)*, στην *κατάσταση* E_j κατά το βήμα *n*

$$P\left[X_n = j \mid X_1 = i_1, X_2 = i_2, \ldots, X_{n-1} = i_{n-1}\right] = P\left[X_n = j \mid X_{n-1} = i_{n-1}\right]$$

Πιθανότητα μετάβασης ενός βήματος

Ομογενείς αλυσίδες Markov

- Αν οι πιθανότητες μετάβασης ενός βήματος είναι *ανεξάρτητες του n*, τότε έχουμε μια *ομογενή αλυσίδα Markov. Ορίζουμε:*

$$p_{ij} \equiv P\left[X_n = j \mid X_{n-1} = i\right]$$

- Πιθανότητες μετάβασης *m*-βημάτων:

$$p_{ij}^{(m)} \equiv P\left[X_{n+m} = j \mid X_n = i\right]$$

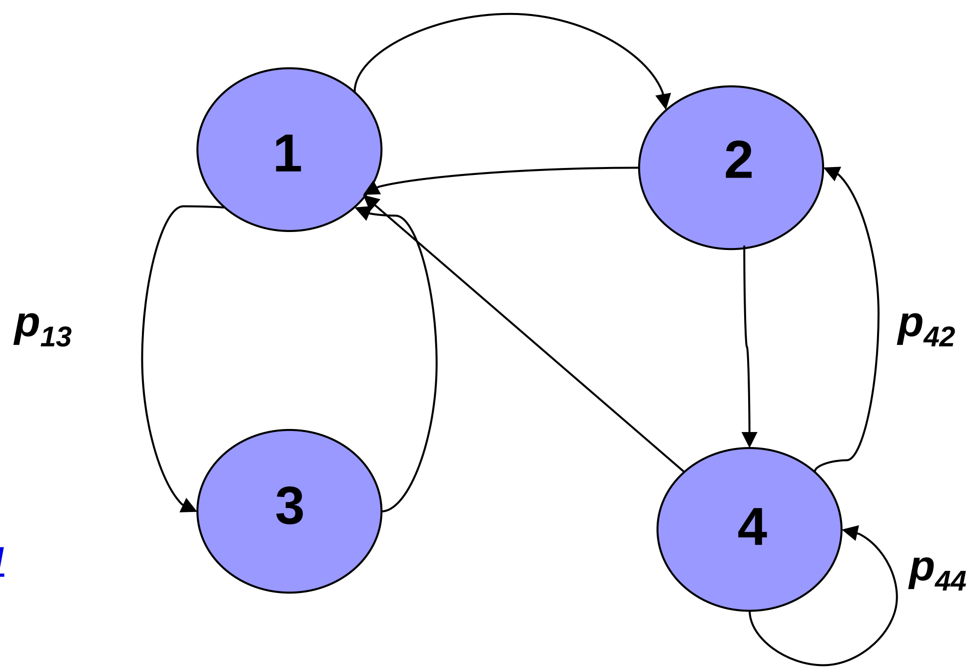
- Εύκολα βγαίνει: $p_{ij}^{(m)} = \sum_k p_{ik}^{(m-1)} p_{kj}$
- Δηλαδή, αν πρόκειται να «ταξιδέψουμε» από την E_i στην E_j μέσα σε *m* βήματα, πρέπει να το κάνουμε «ταξιδεύοντας» πρώτα από την E_i σε κάποια E_k μέσα σε *(m-1)* βήματα και μετά από την E_k στην E_j στο επόμενο βήμα

Ορισμοί για αλυσίδες Markov (1)

- **Αμείωτη**: κάθε κατάσταση της μπορεί να προσπελασθεί από όλες τις υπόλοιπες καταστάσεις. Δηλαδή, υπάρχει ένας ακέραιος m_0 για κάθε ζευγάρι καταστάσεων E_i , E_j :

$$p_{ij}^{(m_0)} > 0$$

ΑΜΕΙΩΤΗ ΑΛΥΣΙΔΑ

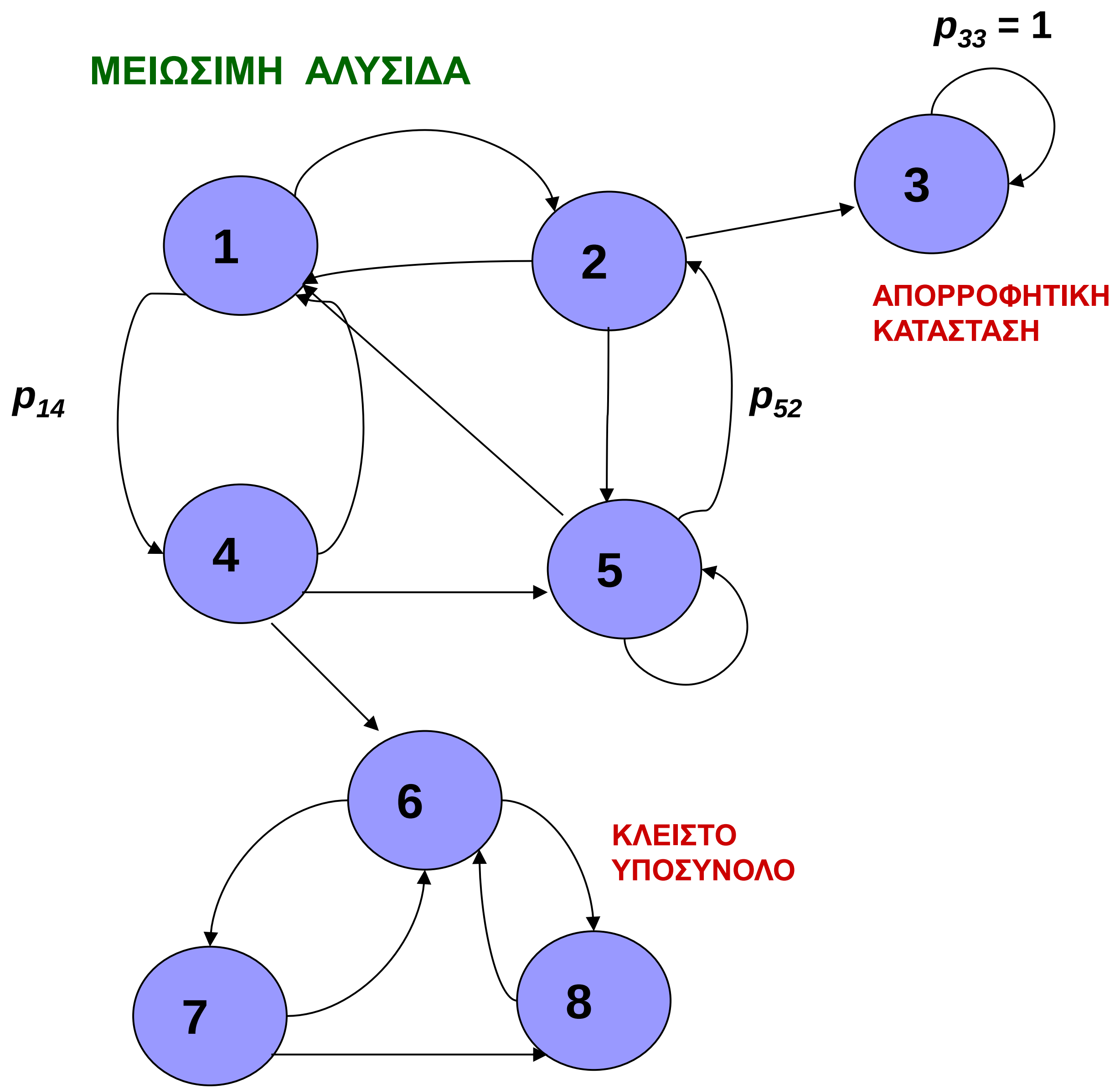


https://setosa.io/ev/markov-chains/

Ορισμοί για αλυσίδες Markov (2)

- Ένα υποσύνολο καταστάσεων *A1* λέγεται *κλειστό* αν δεν είναι δυνατή καμία μετάβαση ενός βήματος από οποιαδήποτε κατάσταση του *A1* σε οποιαδήποτε κατάσταση εκτός του *A1*.
- Αν το *A1* αποτελείται από μια μόνο κατάσταση, έστω E_i , τότε αυτή καλείται *απορροφητική* κατάσταση. Μια αναγκαία και ικανή συνθήκη ώστε να είναι η E_i απορροφητική, είναι $p_{ii} = 1$.
- Αν μία αλυσίδα περιέχει κλειστά υποσύνολα, η αλυσίδα λέγεται *μειώσιμη.*

Παράδειγμα


$p_{33} = 1$
ΜΕΙΩΣΙΜΗ ΑΛΥΣΙΔΑ
1
2
3
ΑΠΟΡΡΟΦΗΤΙΚΗ
ΚΑΤΑΣΤΑΣΗ
p_{14}
p_{52}
4
5
6
ΚΛΕΙΣΤΟ
ΥΠΟΣΥΝΟΛΟ
7
8

Ορισμοί για αλυσίδες Markov (3)

- $f_j^{(n)} \equiv$ *Prob [Η πρώτη επιστροφή στην E_j γίνεται μετά από n βήματα από την αναχώρηση από την E_j]*

- $f_j = \sum_{n=1}^{\infty} f_j^{(n)}$ *= Prob[Κάποτε να επιστρέψουμε στην E_j]*

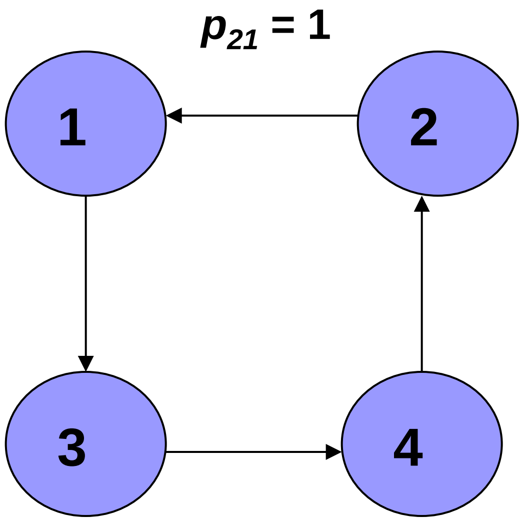
- Αν η πιθανότητα να επιστρέψουμε κάποτε στην κατάσταση E_j, f_j, *είναι* $f_j = 1$, η κατάσταση E_j λέγεται *επαναληπτική*.

- Αν $f_j < 1$, λέγεται *μεταβατική*.

Ορισμοί για αλυσίδες Markov (4)

- Αν τα μόνα δυνατά βήματα κατά τα οποία μπορούμε να επιστρέψουμε στην E_j είναι *γ, 2γ, 3γ, . . ., (γ* ο μεγαλύτερος τέτοιος ακέραιος*)* τότε η E_j λέγεται *περιοδική* με περίοδο *γ*. Αν *γ = 1* τότε η E_j είναι *μη-περιοδική*.

ΠΕΡΙΟΔΙΚΗ ΑΛΥΣΙΔΑ



Ορισμοί για αλυσίδες Markov (5)

- Για τις καταστάσεις με $f_j = 1$, ορίζουμε το Μέσο Χρόνο Επανάληψης της (επιστροφής στην) E_j :

$$M_j \equiv \sum_{n=1}^{\infty} n f_j^{(n)}$$

- Αν ο μέσος χρόνος επιστροφής στην E_j , M_j, *είναι* $M_j = \infty$, η E_j λέγεται *μηδενικά επαναληπτική*, ενώ αν *είναι* $M_j < \infty$, η E_j λέγεται *βέβαια επαναληπτική*.

Θεώρημα 1

- *Οι καταστάσεις μιας αμείωτης αλυσίδας Markov είναι είτε* ***όλες*** *μεταβατικές, είτε όλες βέβαια επαναληπτικές ή όλες μηδενικά επαναληπτικές. Αν είναι περιοδικές, τότε όλες οι καταστάσεις έχουν την ίδια περίοδο γ.*

Πιθανότητες μόνιμης κατάστασης

- $\pi_j^{(n)} \equiv P[X_n = j]$: Πιθανότητα να βρεθεί το σύστημα (η αλυσίδα Markov) στην κατάσταση E_j κατά το n-στο βήμα.
- $\{\pi_j\}$: στάσιμη κατανομή πιθανοτήτων που περιγράφει την πιθανότητα να βρεθεί το σύστημα στην κατάσταση E_j κάποια χρονική στιγμή στο απώτερο μέλλον.

  Πιθανότητες Μόνιμης Κατάστασης: $\pi_j = \lim_{n\to\infty} \pi_j^{(n)}$
- Στην στάσιμη κατανομή, η επίδραση της κατανομής αρχικής κατάστασης $\{\pi_j^{(0)}\}$ έχει εξαφανιστεί
- Το να βρούμε τα $\{\pi_j\}$ είναι το πιο σημαντικό τμήμα της ανάλυσης των αλυσίδων Markov

Θεώρημα 2

Σε μια αμείωτη και μη-περιοδική ομογενή αλυσίδα Markov, οι πιθανότητες μόνιμης κατάστασης $\pi_j = \lim_{n\to\infty} \pi_j^{(n)}$ υπάρχουν πάντα , και είναι ανεξάρτητες από την κατανομή της αρχικής κατάστασης.

Επίσης ισχύει:

1. Είτε όλες οι καταστάσεις είναι μεταβατικές ή όλες είναι μηδενικά επαναληπτικές, οπότε $\pi_j = 0$ και δεν υπάρχει κατανομή μόνιμης κατάστασης.
2. Είτε όλες οι καταστάσεις είναι βέβαια επαναληπτικές και τότε $\pi_j > 0$ για όλα τα j, στην οποία περίπτωση το σύνολο $\{\pi_j\}$ είναι μια κατανομή μόνιμης κατάστασης και

$$\pi_j = \frac{1}{M_j}$$

Στην τελευταία περίπτωση οι ποσότητες π_j καθορίζονται κατά μοναδικό τρόπο από τις εξής εξισώσεις:

$$1 = \sum_i \pi_i \qquad \pi_j = \sum_i \pi_i p_{ij}$$

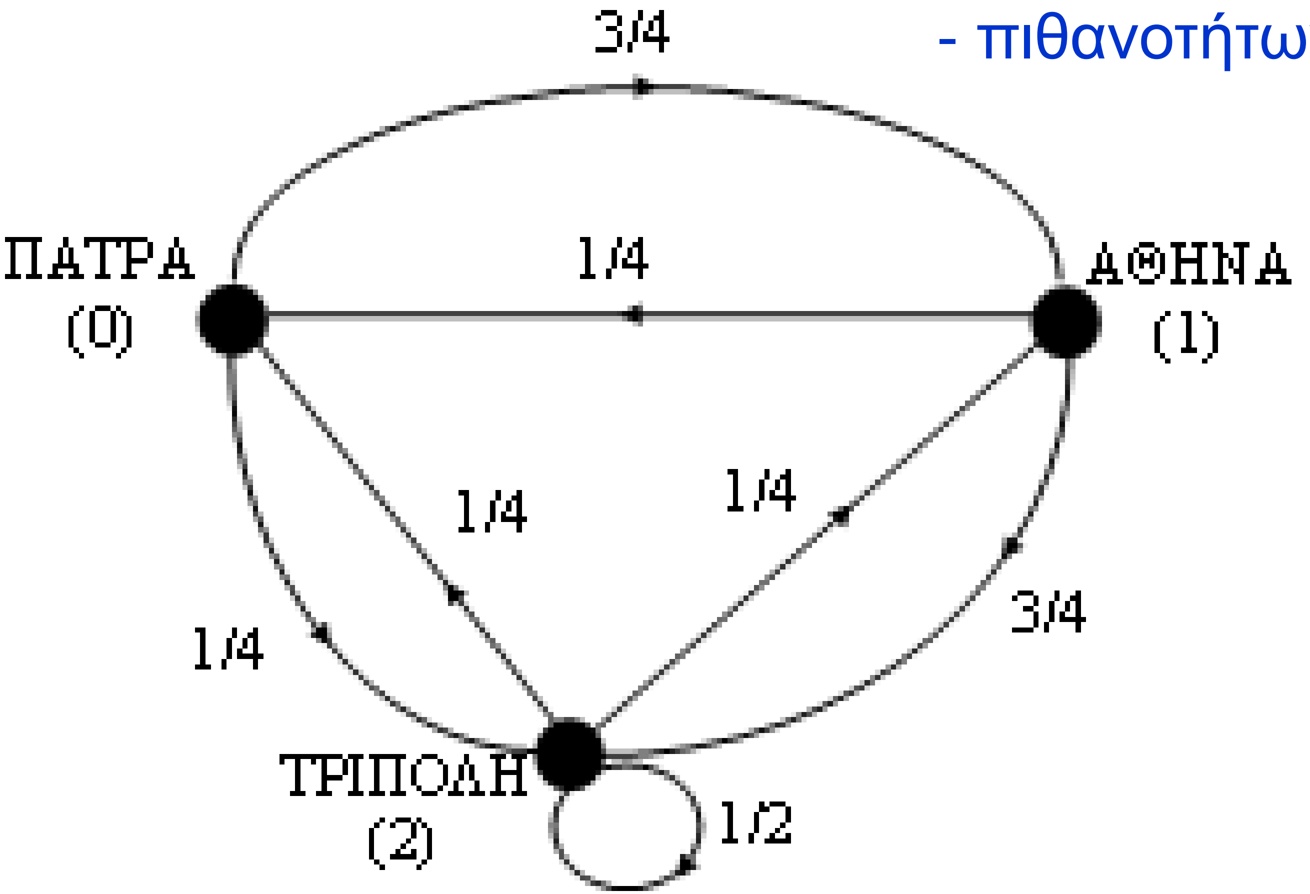
Ορισμοί Markov αλυσίδων (συνέχεια)

- Μια κατάσταση E_j λέγεται **εργοδική**, αν είναι μη-περιοδική και βέβαια επαναληπτική.

  Δηλαδή αν $f_j = 1$, $M_j < \infty$ και $\gamma = 1$.
- Αν όλες οι καταστάσεις μιας αλυσίδας Markov είναι εργοδικές, τότε η αλυσίδα Markov λέγεται και η ίδια *εργοδική*.

Παράδειγμα

Διάγραμμα καταστάσεων -
- πιθανοτήτων μεταβάσεων



Υπολογισμός πιθανοτήτων μόνιμης κατάστασης

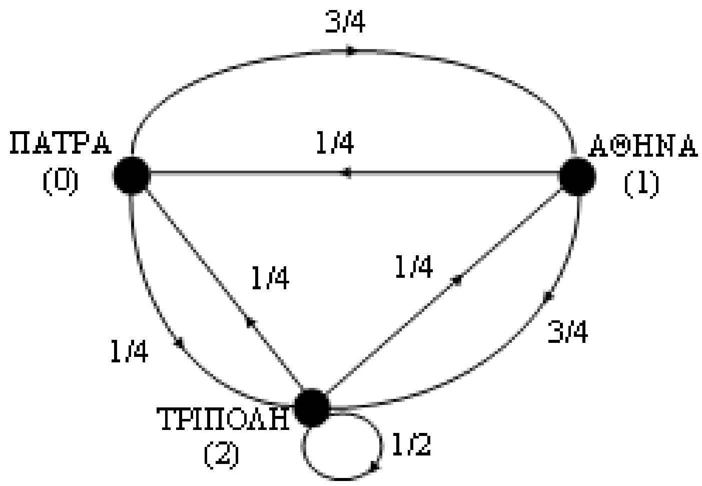
- $\vec{\mathbf{P}} = \left[p_{ij} \right]$ πίνακας μεταβάσεων
- $\vec{\boldsymbol{\pi}} = [\pi_0, \pi_1, \pi_2, \ldots,]$ διάνυσμα πιθανοτήτων
- Από το θεώρημα 2: $\vec{\boldsymbol{\pi}} = \vec{\boldsymbol{\pi}} \cdot \vec{\mathbf{P}}$
- Στο παράδειγμα

$$\vec{\mathbf{P}} = \begin{bmatrix} 0 & 3/4 & 1/4 \\ 1/4 & 0 & 3/4 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}$$

Υπολογισμός πιθανοτήτων μόνιμης κατάστασης (2)

- Λύνουμε τις εξισώσεις

$$\pi_0 = 0\cdot\pi_0 + \frac{1}{4}\cdot\pi_1 + \frac{1}{4}\cdot\pi_2$$
$$\pi_1 = \frac{3}{4}\cdot\pi_0 + 0\cdot\pi_1 + \frac{1}{4}\cdot\pi_2$$
$$\pi_2 = \frac{1}{4}\cdot\pi_0 + \frac{3}{4}\cdot\pi_1 + \frac{1}{2}\cdot\pi_2$$
$$\pi_0 + \pi_1 + \pi_2 = 1$$

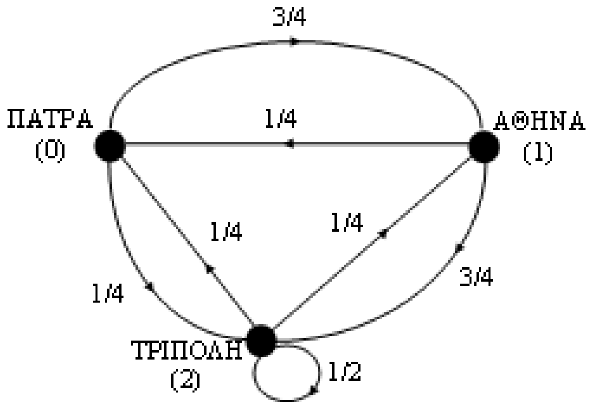


Οι παραπάνω εξισώσεις προέρχονται από τις:

$$\pi_0^{(n+1)} = 0\cdot\pi_0^{(n)} + \frac{1}{4}\cdot\pi_1^{(n)} + \frac{1}{4}\cdot\pi_2^{(n)}$$
$$\pi_1^{(n+1)} = \frac{3}{4}\cdot\pi_0^{(n)} + 0\cdot\pi_1^{(n)} + \frac{1}{4}\cdot\pi_2^{(n)}$$
$$\pi_2^{(n+1)} = \frac{1}{4}\cdot\pi_0^{(n)} + \frac{3}{4}\cdot\pi_1^{(n)} + \frac{1}{2}\cdot\pi_2^{(n)}$$
$$\pi_0^{(n)} + \pi_1^{(n)} + \pi_2^{(n)} = 1$$

για $n \to \infty$

Υπολογισμός πιθανοτήτων μόνιμης κατάστασης (3)



- Αποτέλεσμα:

  $\pi_0 = 1/5 = 0.20$

  $\pi_1 = 7/25 = 0.28$

  $\pi_2 = 13/25 = 0.52$

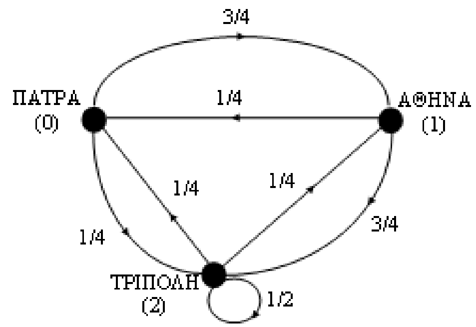
Πιθανότητες Μόνιμης Κατάστασης

Ανάλυση μεταβατικής συμπεριφοράς συστήματος (1)

- Υπολογισμός πιθανοτήτων $\pi_j^{(n)}$: η πιθανότητα να βρεθούμε στην κατάσταση E_j τη χρονική στιγμή n.
- $\vec{\boldsymbol{\pi}}^{(n)} \equiv \left[\pi_0^{(n)}, \pi_1^{(\mathrm{n})}, \pi_2^{(\mathrm{n})}, \quad \ldots\right]$ διάνυσμα πιθανοτήτων στο βήμα n
- Ισχύει ότι $\vec{\boldsymbol{\pi}}^{(n)} = \vec{\boldsymbol{\pi}}^{(n-1)} \cdot \vec{\mathbf{P}}$

$$\vec{\boldsymbol{\pi}}^{(n)} = \vec{\boldsymbol{\pi}}^{(0)} \cdot (\vec{\mathbf{P}})^n$$

Ανάλυση μεταβατικής συμπεριφοράς συστήματος (2)



- Στο παράδειγμα των πόλεων, έστω ότι η αρχική κατανομή είναι η $\vec{\pi}^{(0)} = [1,0,0]$, δηλαδή αρχική πόλη είναι η Πάτρα.
- Στον παρακάτω πίνακα φαίνεται η ακολουθία τιμών των πιθανοτήτων σε κάθε βήμα.

n	0	1	2	3	4	. . . .	∞
$\pi_0^{(n)}$	1	0	0.250	0.187	0.203	. . . .	0.20
$\pi_1^{(n)}$	0	0.75	0.062	0.359	0.254	. . . .	0.28
$\pi_2^{(n)}$	0	0.25	0.688	0.454	0.543	. . . .	0.52

- Οι ποσότητες συγκλίνουν πολύ γρήγορα προς τις οριακές τιμές της μόνιμης κατάστασης.

Χρόνος παραμονής σε μια κατάσταση

Prob [Το σύστημα να παραμείνει στην E_i για ακριβώς m επιπλέον βήματα, δεδομένου ότι έχει μόλις εισέλθει στην E_i] = $(1 - p_{ii})\, p_{ii}^{m}$

Γεωμετρική κατανομή

(Ιδιότητα αμνησίας)

Αλυσίδες Markov συνεχούς χρόνου (1)

Τα απλούστερα συστήματα: *M/M/m/K*

- *Εκθετικά κατανεμημένοι* χρόνοι μεταξύ διαδοχικών αφίξεων $\tilde{\boldsymbol{t}}$ (Στοχαστική Διαδικασία «όμοιων» αφίξεων)

$$A(t) = A_{\tilde{t}}(t) = Prob[\tilde{t} \leq t] = 1 - e^{-\lambda t}, \quad t \geq 0$$

- *Εκθετικά κατανεμημένοι* χρόνοι εξυπηρέτησης $\tilde{\boldsymbol{x}}$

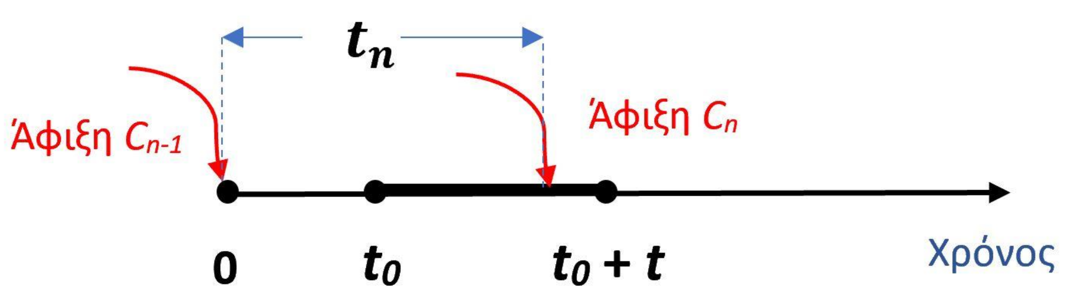
$$B(x) = B_{\tilde{x}}(x) = Prob[\tilde{x} \leq x] = 1 - e^{-\mu x}, \quad x \geq 0$$

Αλυσίδες Markov συνεχούς χρόνου (2)

- *Ιδιότητα της αμνησίας*: «ο χρόνος ως το επόμενο γεγονός, είναι ανεξάρτητος από το χρόνο που έχει περάσει από το τελευταίο γεγονός».
- ΑΦΙΞΕΙΣ:

  Αν έχει περάσει χρόνος t_0 από την τελευταία άφιξη (του C_{n-1})

  $$Prob[\, t_n \le t + t_0 \mid t_n > t_0 \,] = Prob[\, t_n \le t \,]$$



- ΑΝΑΧΩΡΗΣΕΙΣ:

  Αν έχει περάσει χρόνος x_0 εξυπηρέτησης του πελάτη C_n

  $$Prob[\, x_n \le x + x_0 \mid x_n > x_0 \,] = Prob[\, x_n \le x \,]$$

Αλυσίδες Markov συνεχούς χρόνου (3)

- $P_k(t)$ = *Prob*[*k* πελάτες στο σύστημα τη χρονική στιγμή t]

  για $0 \le k \le K, \quad t \ge 0$

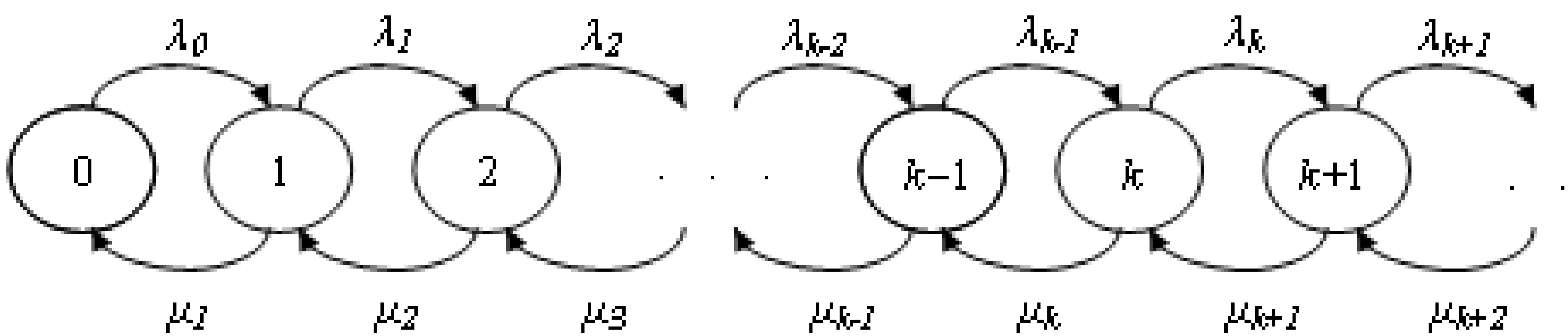
- $p_k = \lim_{t \to \infty} P_k(t)$ = *Prob*[*k* πελάτες στο σύστημα κάποια χρονική στιγμή *στο μέλλον*]
- *Κατανομή μόνιμης κατάστασης*.
- Υπάρχει, αν το σύστημα είναι σταθερό ($0 \le \rho < 1$)

Νόμος ισορροπίας της ροής πιθανότητας

Στη μόνιμη κατάσταση, ο «ρυθμός ροής πιθανότητας» μιας αλυσίδας Markov *από* κάθε κατάσταση, είναι ίσος με το «ρυθμό ροής πιθανότητας» *προς* την κατάσταση.

Αλυσίδες Markov Γεννήσεων – Θανάτων (1)

- Αν το σύστημα βρίσκεται στην κατάσταση j, τότε στην επόμενη αλλαγή κατάστασης θα βρεθεί σε μια από τις καταστάσεις $j-1$ ή $j+1$.
- λ_k : ρυθμός αφίξεων όταν υπάρχουν k πελάτες στο σύστημα
- μ_k : ρυθμός εξυπηρέτησης όταν υπάρχουν k πελάτες στο σύστημα



Διάγραμμα καταστάσεων-ρυθμών μεταβάσεων

Αλυσίδες Markov Γεννήσεων – Θανάτων (2)

- {Ρυθμός ροής πιθανότητας ***από*** την κατάσταση *k*} =

$$p_k \cdot (\lambda_k + \mu_k)$$

- {Ρυθμός ροής πιθανότητας ***προς*** την κατάσταση *k*} =

$$p_{k-1} \cdot \lambda_{k-1} + p_{k+1} \cdot \mu_{k+1}$$

- Με βάση το νόμο ισορροπίας ροής
  - Για $k \geq 1$ $\quad p_k \cdot (\lambda_k + \mu_k) = p_{k-1} \cdot \lambda_{k-1} + p_{k+1} \cdot \mu_{k+1}$ **(1)**
  - Για $k = 0$ $\quad p_0 \cdot \lambda_0 = p_1 \cdot \mu_1$ **(2)**

Αλυσίδες Markov Γεννήσεων – Θανάτων (3)

- Ισχύει πάντα ότι $$\sum_{k=0}^{\infty} p_k = 1$$ **(3)**

- Λύνοντας τις εξισώσεις (1), (2), (3), παίρνουμε:

$$p_k = p_0 \cdot \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}}$$

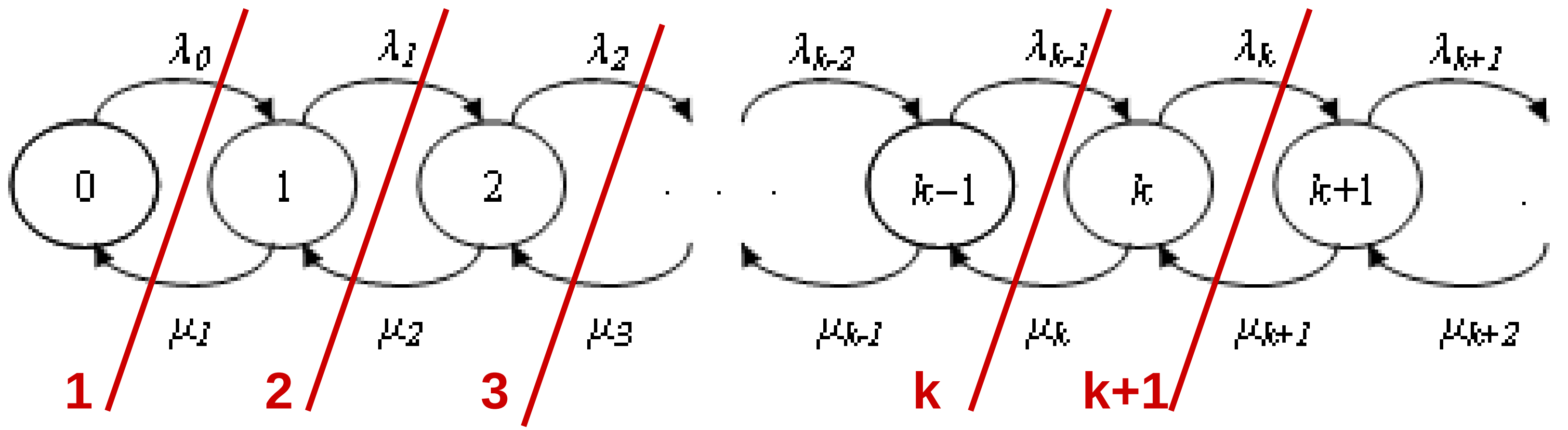
$$p_0 = \frac{1}{1 + \sum_{k=1}^{\infty} \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}}}$$ **(4)**

- Η παραπάνω λύση υπάρχει (δηλαδή, υπάρχει μόνιμη κατάσταση), αν $p_0 > 0$, δηλαδή αν ο παρονομαστής της σχέσης (4) είναι μικρότερος από ∞. Για να ισχύει το τελευταίο, θα πρέπει η ακολουθία λ_k / μ_k να συγκλίνει, δηλαδή θα πρέπει να υπάρχει κάποιο k_0 τέτοιο ώστε: $\frac{\lambda_k}{\mu_k} < 1$ για όλα τα $k \geq k_0$

Αλυσίδες Markov Γεννήσεων – Θανάτων (4)

ΕΝΑΛΛΑΚΤΙΚΑ

Ο Νόμος διατήρησης της ροής εφαρμόζεται και σε κάθε «σύνορο» της αλυσίδας Markov:



1: $p_0 \cdot \lambda_0 = p_1 \cdot \mu_1$

2: $p_1 \cdot \lambda_1 = p_2 \cdot \mu_2$

:

k: $p_{k-1} \cdot \lambda_{k-1} = p_k \cdot \mu_k$

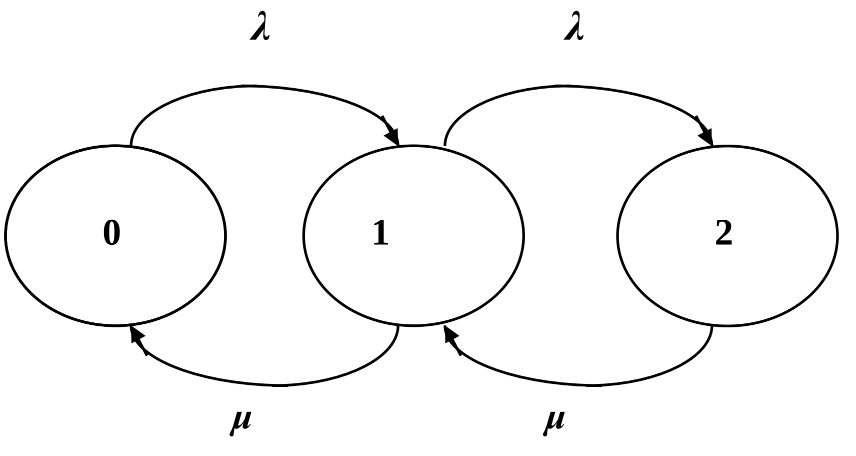
Ίδια Αποτελέσματα

Αλυσίδες Markov Γεννήσεων – Θανάτων (5)

ΠΑΡΑΔΕΙΓΜΑ

Μας δίνεται μια αλυσίδα Markov γεννήσεων – θανάτων, η οποία έχει μόνο τρεις καταστάσεις {0, 1, 2}, ενώ ισχύει:

$\lambda_k = \lambda$ για $k = 0,1$ και $\mu_k = \mu$ για $k = 1,2$



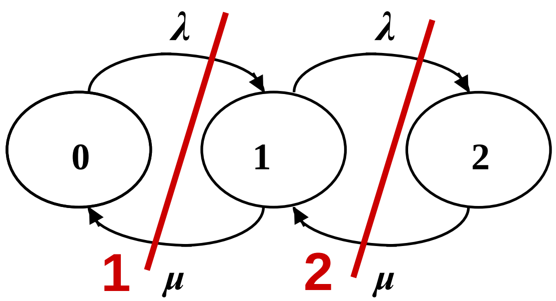
Διάγραμμα Καταστάσεων – Ρυθμών Μεταβάσεων

Αλυσίδες Markov Γεννήσεων – Θανάτων (6)

- Για την κατάσταση 0: $p_0 \cdot \lambda = p_1 \cdot \mu$
- Για την κατάσταση 1: $p_1 \cdot (\lambda + \mu) = p_0 \cdot \lambda + p_2 \cdot \mu$
- Για την κατάσταση 2: $p_2 \cdot \mu = p_1 \cdot \lambda$

Από τις παραπάνω 3 σχέσεις, μόνο οι 2 είναι ανεξάρτητες.

ΕΝΑΛΛΑΚΤΙΚΑ:



1: $p_0 \cdot \lambda = p_1 \cdot \mu$

2: $p_1 \cdot \lambda = p_2 \cdot \mu$

……ανεξάρτητες.

Αλυσίδες Markov Γεννήσεων – Θανάτων (6)

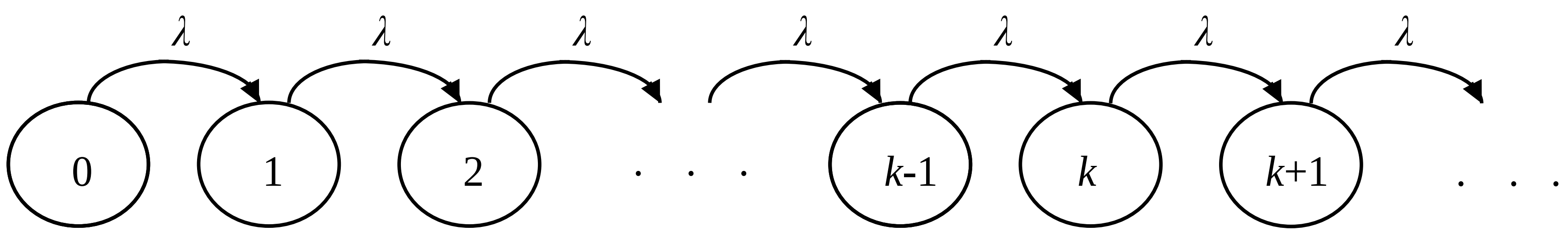
- Χρησιμοποιούμε την $p_0 + p_1 + p_2 = 1$ με 2 από τις παραπάνω, και παίρνουμε την τελική λύση:

$$p_0 = \frac{1}{1 + \lambda/\mu + (\lambda/\mu)^2} \qquad p_1 = \frac{\lambda/\mu}{1 + \lambda/\mu + (\lambda/\mu)^2} \qquad p_2 = \frac{(\lambda/\mu)^2}{1 + \lambda/\mu + (\lambda/\mu)^2}$$

  - Η αλυσίδα αυτή αντιστοιχεί στο σύστημα ***M/M/1/2***. Γιατί;
  - Στο σύστημα αυτό επιτρέπεται ***λ/μ* ≥ 1**. Γιατί;

Διαδικασίες Poisson (1)

- Ειδική περίπτωση Γεννήσεων-Θανάτων (μόνο αφίξεις)
  - $\lambda_\kappa = \lambda$ για όλα τα k
  - $\mu_\kappa = 0$ για όλα τα k



Διάγραμμα Καταστάσεων – Ρυθμών Μεταβάσεων Διαδικασίας Poisson

Δεν είναι εργοδικό σύστημα. Όλες οι καταστάσεις είναι *μεταβατικές*.

Διαδικασίες Poisson (2)

- Έστω το σύστημα ξεκινά τη στιγμή $t = 0$, άδειο. Δηλαδή:

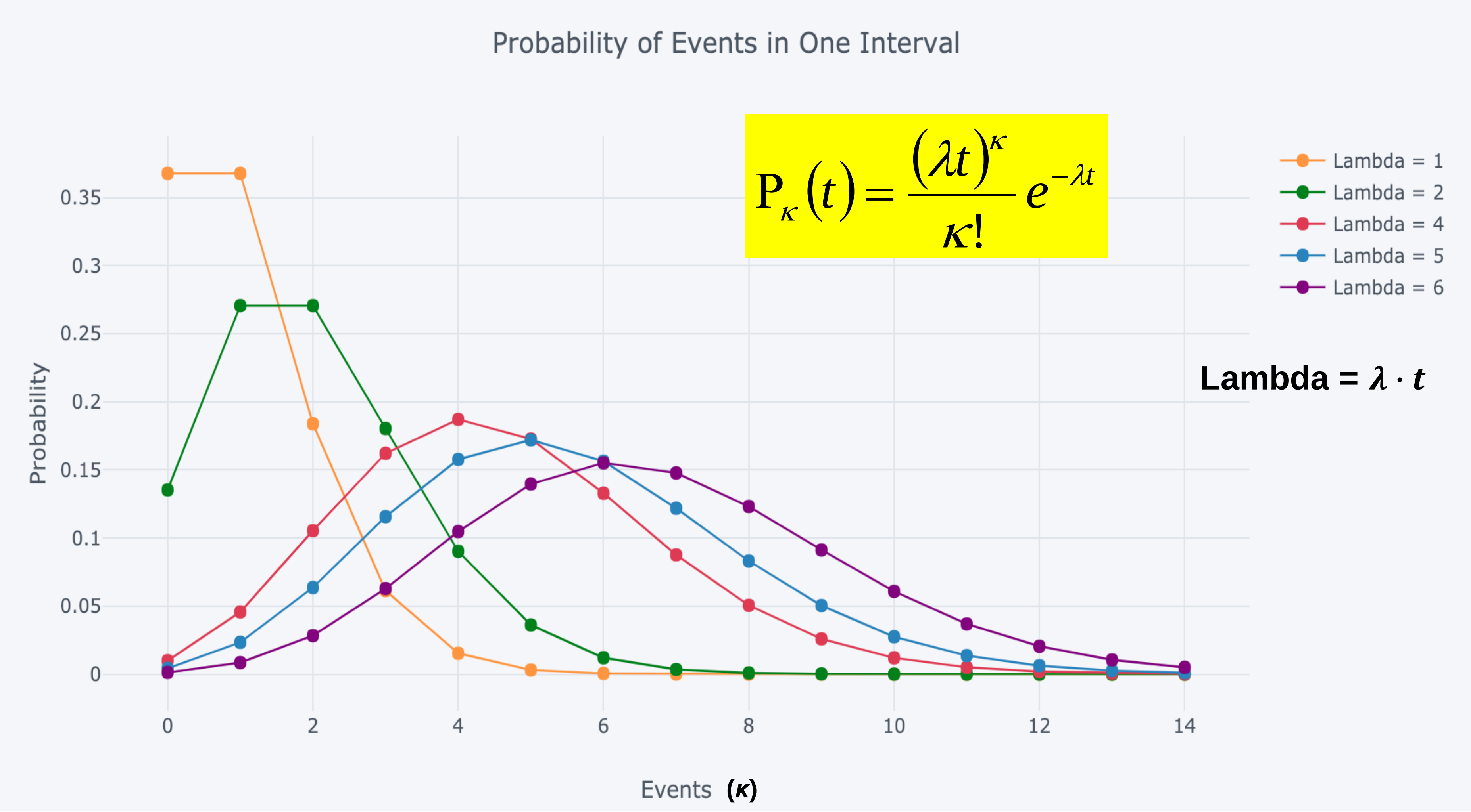
$$\mathrm{P}_{\kappa}(0) = \begin{cases} 1, & \kappa = 0 \\ 0, & \kappa \neq 0 \end{cases}$$

- Τη χρονική στιγμή t: $\mathrm{P}_{\kappa}(t) = \dfrac{(\lambda t)^{\kappa}}{\kappa!} e^{-\lambda t}$ για $k \geq 0$, $t \geq 0$

  Κατανομή Poisson

- *Μέση Τιμή* και *Διακύμανση* (αριθμού αφίξεων στο $[\,0, t\,]$), ίσα με $\boldsymbol{\lambda t}$. (αναμενόμενο).
- **Δηλαδή, στο M/M/1, η διαδικασία μόνο των αφίξεων, είναι Poisson**

Η κατανομή Poisson



$$P_\kappa(t) = \frac{(\lambda t)^\kappa}{\kappa!} e^{-\lambda t}$$

Lambda = $\lambda \cdot t$

Poisson αφίξεις → Εκθετικοί χρόνοι μεταξύ αφίξεων

$A(t) = \mathrm{P}[\tilde{t} \le t]$

$\mathrm{P}_\kappa(t) = \frac{(\lambda t)^\kappa}{\kappa!} e^{-\lambda t}$

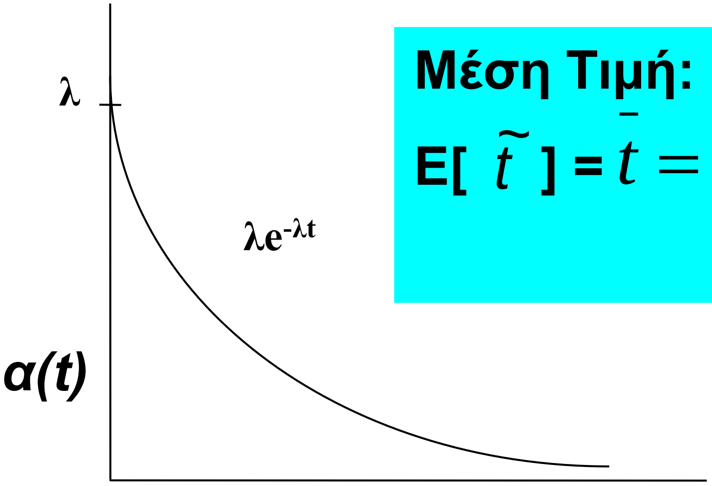
- $\tilde{t}$ = ΤΜ για το χρόνο μεταξύ αφίξεων, με PDF *A(t)* και pdf *α(t)*

$A(t) = 1 - P[\tilde{t} > t] = 1 - \underbrace{P_0(t)}_{\text{Poisson}} = \mathbf{1 - e^{-\lambda t}}$, $t \ge 0$ (PDF Εκθετικής)

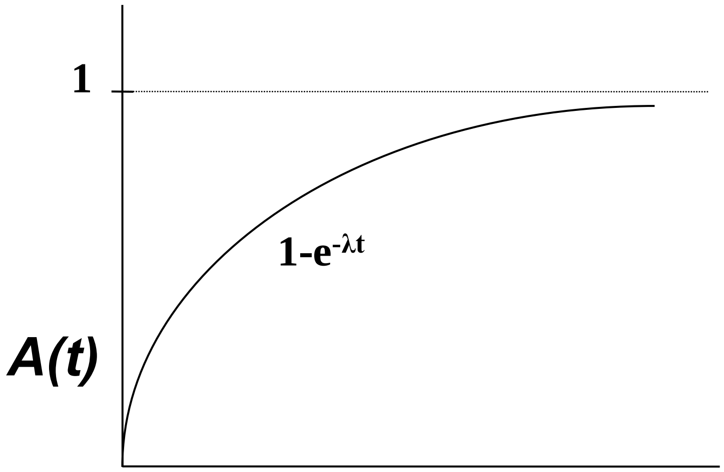
Παράγωγος ως προς t: $\boldsymbol{\alpha(t) = \lambda e^{-\lambda t}}$, $t \ge 0$ (pdf Εκθετικής)

Μέση Τιμή:

$\mathrm{E}[\tilde{t}] = \bar{t} = \frac{1}{\lambda}$



(a) pdf

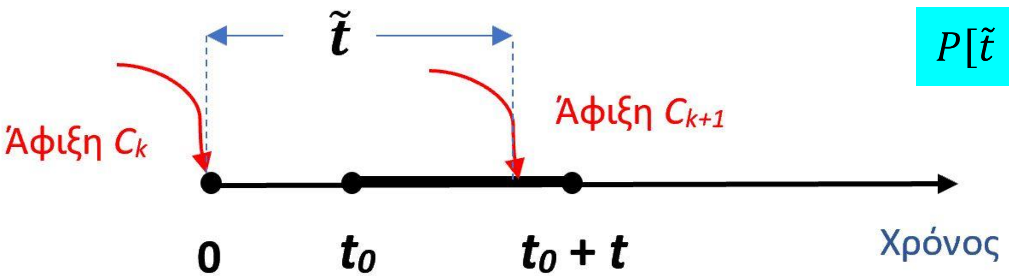


(b) PDF

Η εκθετική κατανομή

Ιδιότητα Αμνησίας της Εκθετικής Κατανομής

Έστω ότι γίνεται μια άφιξη τη χρονική στιγμή **0**. Τώρα, έστω ότι πέρασαν $\boldsymbol{t_0}$ δευτερόλεπτα κατά τη διάρκεια των οποίων δεν έγινε άφιξη. Αν αυτή τη στιγμή $\boldsymbol{t_0}$ ρωτήσουμε *«ποια είναι η πιθανότητα η επόμενη άφιξη να γίνει μέσα στα επόμενα* ***t*** *δευτερόλεπτα από τώρα»*, η απάντηση θα είναι:



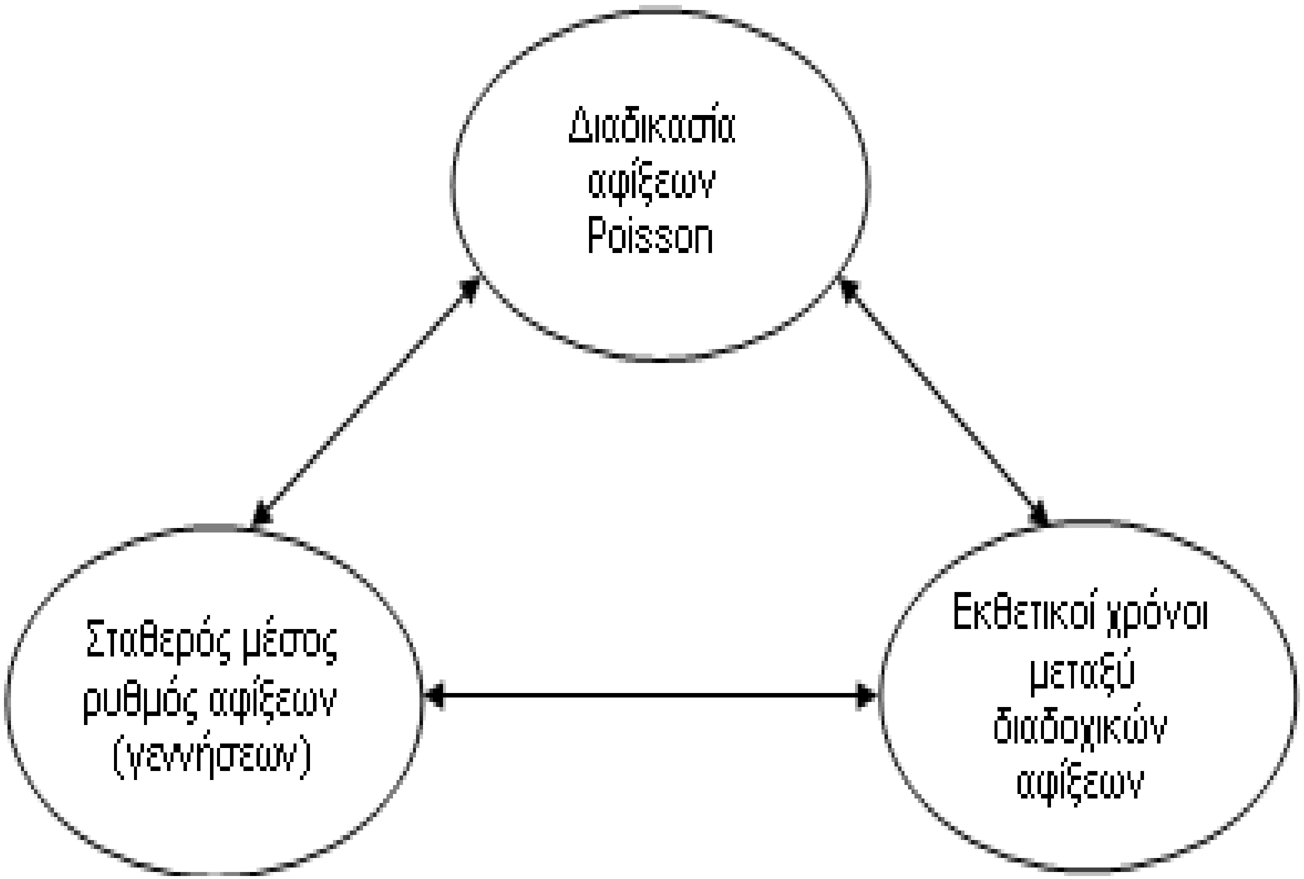
$$P[\tilde{t} \le t] = A(t) = 1 - e^{-\lambda t}$$

$$P\left[\tilde{t} \le t + t_0 \;\middle|\; \tilde{t} > t_0\right] = \frac{P\left[t_0 < \tilde{t} \le t + t_0\right]}{P\left[\tilde{t} > t_0\right]} = \frac{P\left[\tilde{t} \le t + t_0\right] - P\left[\tilde{t} \le t_0\right]}{P\left[\tilde{t} > t_0\right]}$$

$$\Leftrightarrow \quad P\left[\tilde{t} \le t + t_0 \;\middle|\; \tilde{t} > t_0\right] = \frac{1 - e^{-\lambda(t+t_0)} - \left(1 - e^{-\lambda t_0}\right)}{1 - \left(1 - e^{-\lambda t_0}\right)} \quad \Leftrightarrow$$

$$P\left[\tilde{t} \le t + t_0 \;\middle|\; \tilde{t} > t_0\right] = 1 - e^{-\lambda t} \quad \Leftrightarrow \quad P\left[\tilde{t} \le t + t_0 \;\middle|\; \tilde{t} > t_0\right] = P\left[\tilde{t} \le t\right]$$

Σχέσεις



Διαδικασία
αφίξεων
Poisson
Σταθερός μέσος
ρυθμός αφίξεων
(γεννήσεων)
Εκθετικοί χρόνοι
μεταξύ
διαδοχικών
αφίξεων

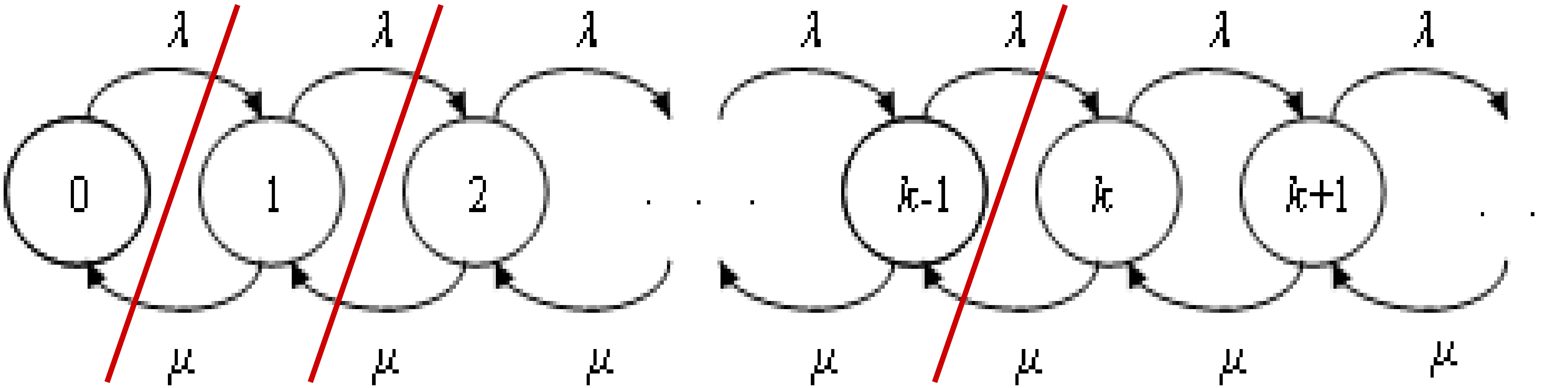
Το κλασικό Σύστημα Αναμονής *M/M/1*

- *Εκθετική* κατανομή της διαδικασίας των χρόνων μεταξύ διαδοχικών αφίξεων
- *Εκθετική* κατανομή των χρόνων εξυπηρέτησης
- Ένας εξυπηρετητής
- Άπειρο μήκος ουράς
- Αλυσίδα Markov Γεννήσεων – Θανάτων
- Με τη συνηθισμένη παραδοχή ότι οι *ρυθμοί* αφίξεων και εξυπηρέτησης δεν εξαρτώνται από την κατάσταση του συστήματος (αριθμός παρόντων πελατών), ισχύει:

$$\lambda_k = \lambda \quad \text{για} \quad k = 0, 1, 2, \ldots$$

$$\mu_k = \mu \quad \text{για} \quad k = 1, 2, 3, \ldots$$

Το Σύστημα Αναμονής *M/M/1 (συνέχεια)*



Διάγραμμα καταστάσεων - ρυθμών μεταβάσεων για το σύστημα *M/M/1*.

Για τη λύση:

$p_0 \cdot \lambda = p_1 \cdot \mu$

$p_1 \cdot \lambda = p_2 \cdot \mu$

:

$p_{k-1} \cdot \lambda = p_k \cdot \mu$

:

και $$\sum_{k=0}^{\infty} p_k = 1$$

Λύση συστήματος *M/M/1*

- Χρησιμοποίηση (G/G/1): $$\rho = \lambda \cdot \bar{x} = \frac{\lambda}{\mu}$$
- Συνθήκη σταθερότητας: $$0 < \rho = \frac{\lambda}{\mu} < 1$$
- Από τη γενική λύση των διαδικασιών Γ-Θ (ή την προσέγγιση της προηγούμενης διαφάνειας):

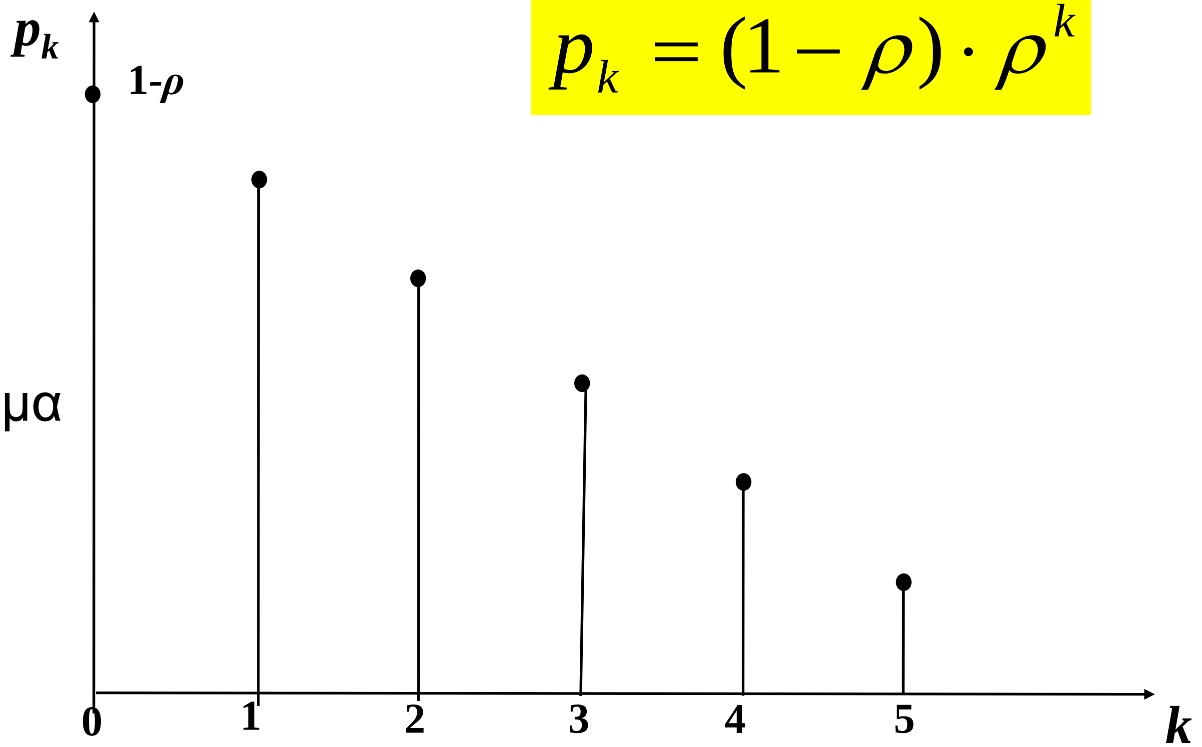
$$p_k = (1-\rho) \cdot \rho^k$$ για $k = 0, 1, 2, \ldots$

- Περιέχεται το: $p_0 = 1 - \rho$

Λύση συστήματος *M/M/1* (συν)

- Τα p_k ακολουθούν τη *Γεωμετρική Κατανομή*
- Εξαρτώνται από τα *λ* και *μ*, μόνο μέσω του λόγου τους ***ρ***

$$p_k = (1-\rho)\cdot\rho^k$$

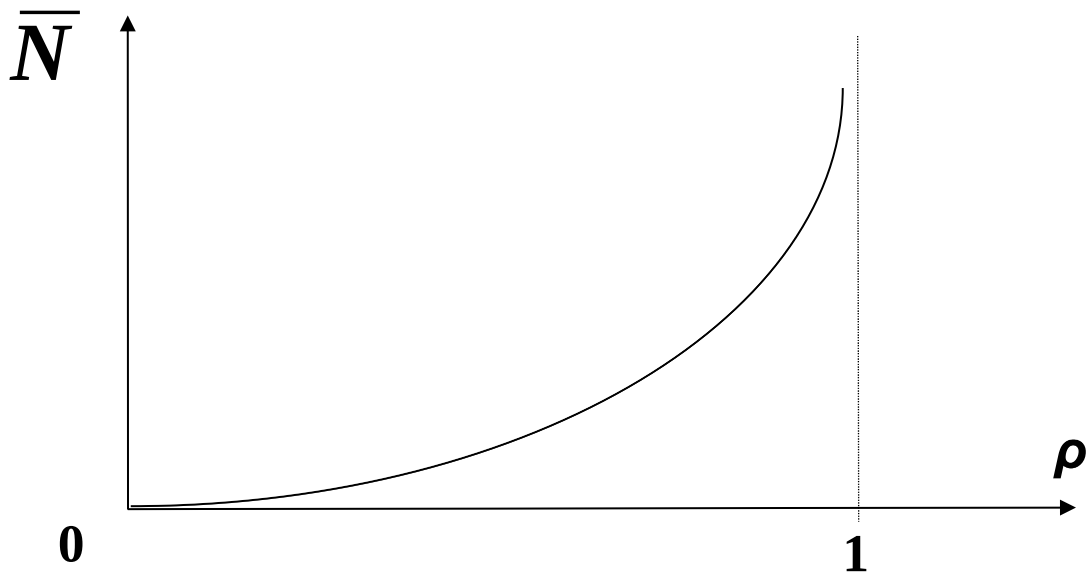


Τα p_k στο σύστημα *M/M/1*.

Μετρικές απόδοσης στο *M/M/1*

- Μέσος αριθμός πελατών στο σύστημα

$$\overline{N} = \sum_{k=0}^{\infty} k p_k = (1-\rho)\sum_{k=0}^{\infty} k\rho^k = (1-\rho)\frac{\rho}{(1-\rho)^2} = \frac{\rho}{1-\rho}$$

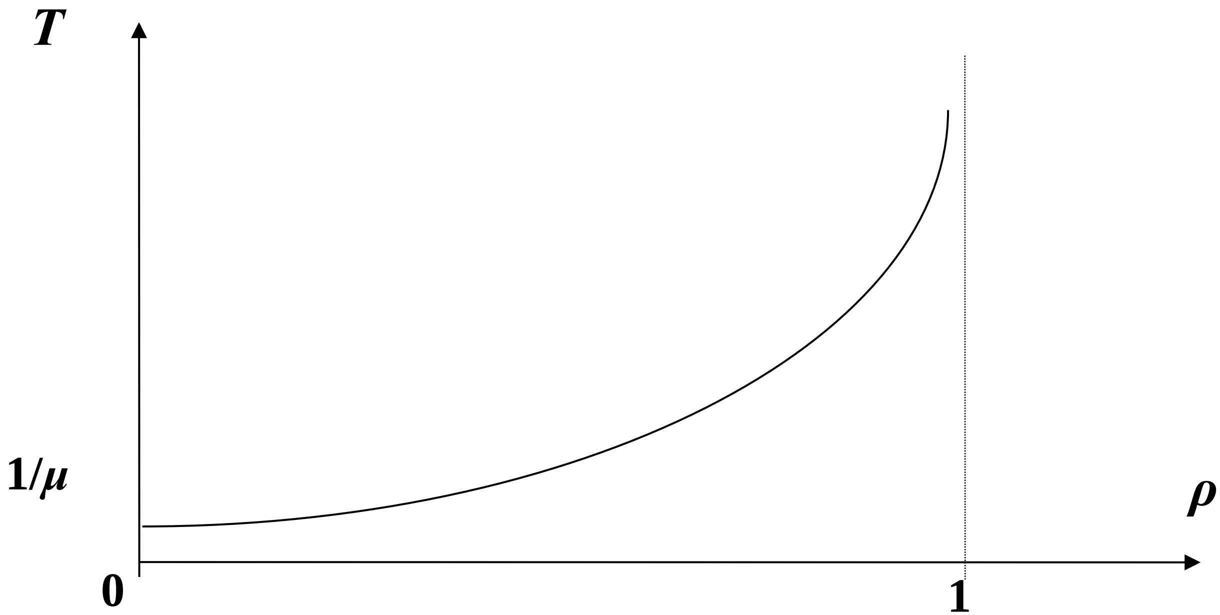


Μετρικές απόδοσης στο *M/M/1 (συν.)*

- Μέσος χρόνος ενός πελάτη στο σύστημα (*Response Time*)

  Με χρήση του *Νόμου του Little*:

$$T = \frac{\overline{N}}{\lambda} = \frac{1/\mu}{1-\rho}$$



Μετρικές απόδοσης στο *M/M/1* (συν.)

- Μέσος χρόνος αναμονής ενός πελάτη στην ουρά

$$W = T - \bar{x} = T - 1/\mu = \frac{\rho}{\mu \cdot (1-\rho)}$$

- Μέσος αριθμός πελατών στην ουρά

$$\overline{N}_q = \overline{N} - \rho = \frac{\rho^2}{1-\rho}$$

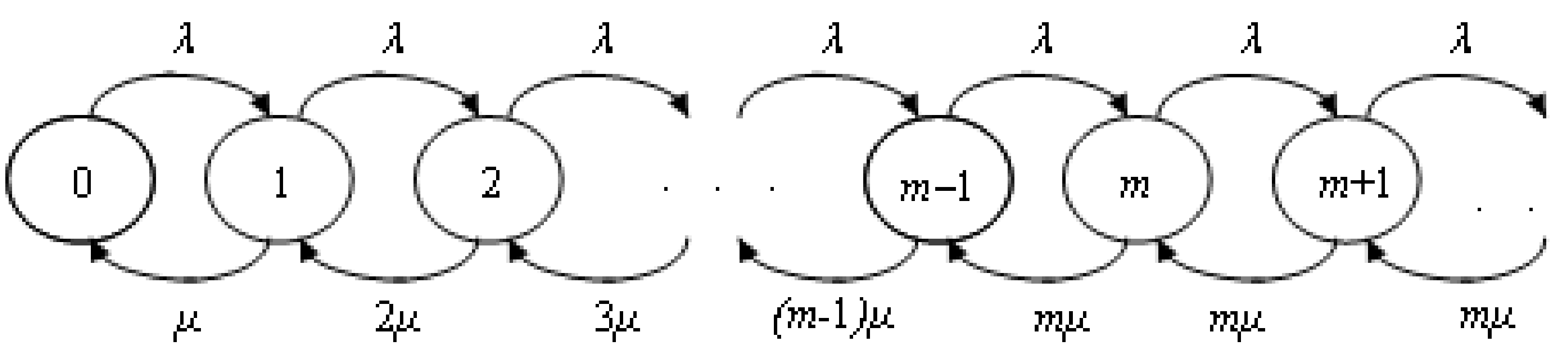
- Πιθανότητα να υπάρχουν τουλάχιστον *n* πελάτες στο σύστημα

$P_{(n)}$ = *Prob*[*n* ή περισσότεροι πελάτες στο σύστημα]

$$P_{(n)} = \sum_{k=n}^{\infty} p_k = (1-\rho)\sum_{k=n}^{\infty} \rho^k = (1-\rho)\rho^n \sum_{k=0}^{\infty} \rho^k = (1-\rho)\rho^n \frac{1}{1-\rho} = \rho^n$$

Το σύστημα αναμονής *M/M/m*

- ***m*** ίδιοι εξυπηρετητές
- Ο καθένας με ρυθμό εξυπηρέτησης ***μ***
- Τα υπόλοιπα χαρακτηριστικά, ίδια με του *M/M/1*
- $\lambda_k = \lambda$ για $k = 0, 1, 2, \ldots$
- $$\mu_k = \begin{cases} k\mu & \text{για} \quad 1 \le k \le m \\ m\mu & \text{για} \quad m \le k \end{cases}$$



Το σύστημα αναμονής *M/M/m* (συν)

- Χρησιμοποίηση

$$\rho = \frac{\lambda \bar{x}}{m} = \frac{\lambda}{m\mu}$$

Συνθήκη Σταθερότητας

$$0 < \rho = \frac{\lambda}{m\mu} < 1$$

- Λύση μόνιμης κατάστασης

$$p_k = \begin{cases} p_0 \dfrac{(m\rho)^k}{k!} & \text{για} \quad 1 \le k \le m \\ p_0 \dfrac{\rho^k m^m}{m!} & \text{για} \quad k \ge m \end{cases}$$

$$p_0 = \left[\sum_{k=0}^{m-1} \frac{(m\rho)^k}{k!} + \left(\frac{(m\rho)^m}{m!} \right) \left(\frac{1}{1-\rho} \right) \right]^{-1}$$

Μετρικές απόδοσης στο *M/M/m*

- Πιθανότητα να χρειαστεί να περιμένει στην ουρά ένας πελάτης:

Π = *Prob*[*m* ή περισσσότεροι πελάτες στο σύστημα]

$$\Pi = \sum_{k=m}^{\infty} p_k = \sum_{k=m}^{\infty} p_0 \frac{\rho^k m^m}{m!} = p_0 \frac{m^m}{m!} \sum_{k=m}^{\infty} \rho^k = p_0 \frac{m^m}{m!} \rho^m \frac{1}{1-\rho} = p_0 \frac{(m\rho)^m}{m!\,(1-\rho)}$$

- Μέσος αριθμός πελατών στο σύστημα

$$\overline{N} = \sum_{k=0}^{\infty} k p_k = m\rho + \frac{\rho \Pi}{1-\rho}$$

Μετρικές απόδοσης στο *M/M/m* (συν)

- Μέσος χρόνος ενός πελάτη στο σύστημα (response time)

$$T = \frac{\overline{N}}{\lambda} = \frac{1}{\mu}\left(1 + \frac{\Pi}{m(1-\rho)}\right)$$

(N. Little)

- Μέσος χρόνος αναμονής ενός πελάτη στην ουρά

$$W = T - \overline{x} = T - 1/\mu = \frac{\Pi}{m\mu(1-\rho)}$$

- Μέσος αριθμός πελατών στην ουρά

$$\overline{N}_q = \overline{N} - m\rho = \frac{\rho\Pi}{1-\rho}$$

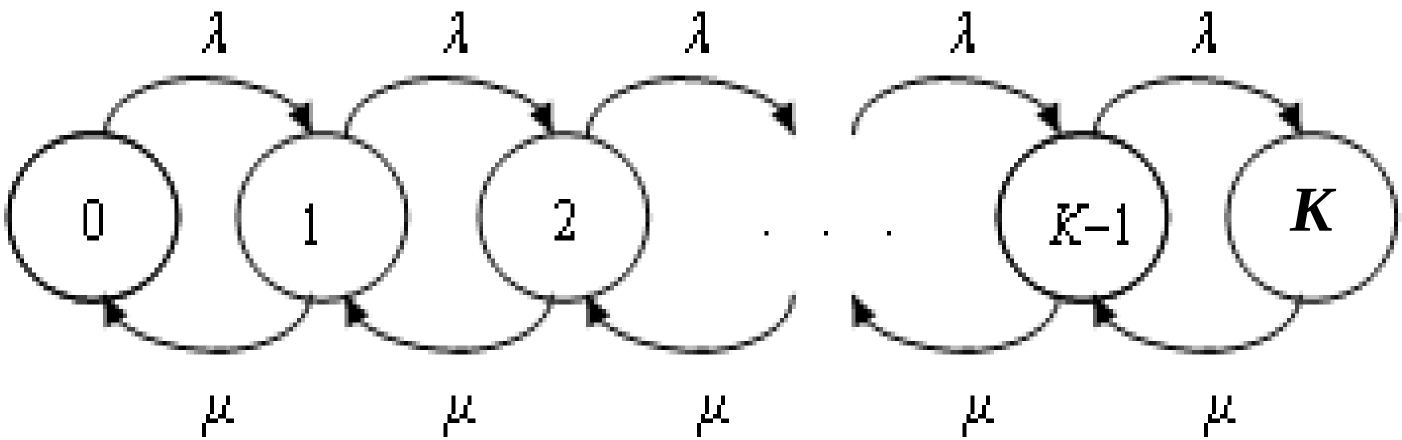
Το σύστημα αναμονής *M/M/1/K*

- Ίδια χαρακτηριστικά με το *M/M/1*, αλλά περιορισμένη χωρητικότητα σε πελάτες.
- Στο σύστημα μπορούν να βρίσκονται το πολύ ***K*** πελάτες (στην ουρά και στον εξυπηρετητή).
- Πελάτες που φθάνουν και βρίσκουν γεμάτο το σύστημα, χάνονται.
- Οι ρυθμοί αφίξεων και εξυπηρέτησης του *M/M/1/K*:

$$\lambda_k = \begin{cases} \lambda & \text{για} \quad 0 \le k < K \\ 0 & \text{για} \quad k \ge K \end{cases}$$

$$\mu_k = \begin{cases} \mu & \text{για} \quad 1 \le k \le K \\ 0 & \text{για} \quad k > K \end{cases}$$

Το σύστημα αναμονής *M/M/1/K* (συν)



- Λύση συστήματος

$$p_k = \begin{cases} \dfrac{1-\lambda/\mu}{1-(\lambda/\mu)^{K+1}}\left(\dfrac{\lambda}{\mu}\right)^k & \text{για} \quad 0 \le k \le K \\ 0 & \text{αλλιώς} \end{cases}$$

Το σύστημα αναμονής *M/M/1/K* (συν)

ΜΕΤΡΙΚΕΣ

- Μέσος αριθμός πελατών στο σύστημα

$$\overline{N} = \sum_{k=0}^{K} k p_k = \frac{\lambda/\mu}{1-\lambda/\mu} - \frac{(K+1)(\lambda/\mu)^{K+1}}{1-(\lambda/\mu)^{K+1}}$$

- Μέσος αριθμός πελατών στην ουρά

$$\overline{N}_q = \sum_{k=2}^{K} (k-1) p_k = \frac{\lambda/\mu}{1-\lambda/\mu} - (\lambda/\mu) \cdot \frac{1+K(\lambda/\mu)^{K}}{1-(\lambda/\mu)^{K+1}}$$

Το σύστημα αναμονής *M/M/1/K* (συν)

- **Παράδειγμα**: Το μοντέλο μιας τηλεφωνικής συσκευής χωρίς κράτηση κλήσεων (παλιό αναλογικό σύστημα):

M/M/1/1

$$p_k = \begin{cases} \dfrac{1}{1+\lambda/\mu} & \text{για} \quad k=0 \\ \dfrac{\lambda/\mu}{1+\lambda/\mu} & \text{για} \quad k=1 \\ 0 & \text{αλλιώς} \end{cases}$$

p_0 : Πιθανότητα να μιλήσει, κάποιος που καλεί

p_1 : Πιθανότητα να βρει κατειλημμένη τη συσκευή, κάποιος που καλεί

λ : Μέσος ρυθμός με τον οποίο γίνονται κλήσεις στη συσκευή

$\bar{x} = 1/\mu$: Μέση χρονική διάρκεια μιας συνομιλίας