ΣΥΣΤΗΜΑΤΑ ΑΝΑΜΟΝΗΣ
Queuing Systems

Η Ουρά M/M/1/*N*
Σφαιρικές & Τοπικές Εξισώσεις Ισορροπίας

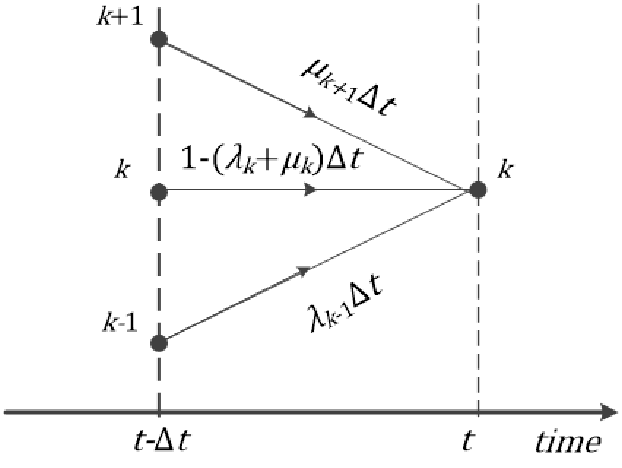
Βασίλης Μάγκλαρης

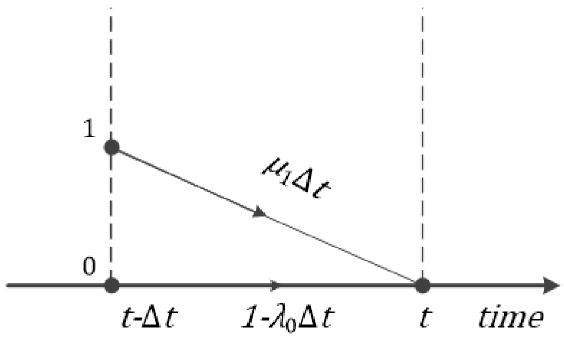
maglaris@netmode.ntua.gr

22/3/2017

ΔΙΑΔΙΚΑΣΙΑ ΓΕΝΝΗΣΕΩΝ – ΘΑΝΑΤΩΝ (1/4)
Birth – Death Processes *(Επανάληψη)*

- Παραδοχές:
  - Ανεξαρτησία γεννήσεων-θανάτων
  - Εξέλιξη της κατάστασης - πληθυσμού $n(t)$ βασισμένη μόνο στο παρόν (Ιδιότητα Markov)
- Σύστημα Διαφορικών εξισώσεων Διαφορών
  - Κατάσταση ισορροπίας (**steady state**)
  - Την χρονική στιγμή t το σύστημα καταλήγει σε πληθυσμό $n(t) = k$
  - Μπορεί να έχουν προηγηθεί οι ακόλουθες μεταβάσεις από την χρονική στιγμή $t - \Delta t, \Delta t \to 0$:
    - Μία άφιξη στο διάστημα Δt, με πιθανότητα $\lambda_{k-1}\Delta t$ αν $k > 0$
    - Μια αναχώρηση, με πιθανότητα $\mu_{k+1}\Delta t$ αν υπάρχει η $k+1$ (σε περίπτωση περιορισμού μέγιστου πληθυσμού K μπορούμε να θεωρήσουμε $\mu_{k+1} = 0$)
    - Τίποτα από τα δύο, με πιθανότητα $1 - (\lambda_k + \mu_k)\Delta t$ αν $k > 0$ ή $1 - \lambda_0 \Delta t$ αν $k = 0$
- Οι εξισώσεις μετάβασης (***Chapman - Kolmogorov***) προκύπτουν από τον τύπο συνολικής πιθανότητας:
  - $\mathrm{P}_k(t) = \lambda_{k-1}\Delta t \mathrm{P}_{k-1}(t - \Delta t) + \mu_{k+1}\Delta t \mathrm{P}_{k+1}(t - \Delta t) + [1 - (\lambda_k + \mu_k)\Delta \mathrm{t}]\mathrm{P}_k(\mathrm{t} - \Delta \mathrm{t})$
  - $\mathrm{P}_{\mathbf{0}}(t) = \mu_1 \Delta t \, \mathrm{P}_1(t - \Delta t) + (1 - \lambda_0 \Delta t) \, \mathrm{P}_{\mathbf{0}}(t - \Delta t)$
  - με αρχικές συνθήκες $\mathrm{P}_k(0)$ και οριακές συνθήκες $\sum_k \mathrm{P}_k(t) = 1, \forall t$





ΔΙΑΔΙΚΑΣΙΑ ΓΕΝΝΗΣΕΩΝ – ΘΑΝΑΤΩΝ (2/4)

Birth – Death Processes *(Επανάληψη)*

Στο όριο, $\Delta t \approx dt \to 0$, $\frac{\mathrm{P}_k(t)-P_k(t-\Delta t)}{\Delta t} \to \frac{d\mathrm{P}_k(t)}{dt}$ και προκύπτει το γραμμικό σύστημα διαφορικών εξισώσεων διαφορών:

- $\frac{d\mathrm{P}_k(t)}{dt} = \lambda_{k-1}\mathrm{P}_{k-1}(t) + \mu_{k+1}\mathrm{P}_{k+1}(t) - (\lambda_k + \mu_k)\mathrm{P}_k(t),\ k > 0$
- $\frac{d\mathrm{P}_0(t)}{dt} = \mu_1\mathrm{P}_1(t) - \lambda_0\mathrm{P}_0(t)$
- με αρχικές συνθήκες $\mathrm{P}_k(0)$ και οριακές συνθήκες $\sum_k \mathrm{P}_k(t) = 1, \forall t$

Όταν $t \to \infty$ και κάτω από ορισμένες συνθήκες το σύστημα συγκλίνει σε σταθερή κατάσταση. Το μεταβατικό φαινόμενο παρέρχεται για καταστάσεις $n(t) = k$ απείρως επισκέψιμες - **positive recurrent**, ξεχνιέται η αρχική συνθήκη $\mathrm{P}_k(0)$ και οι πιθανότητές $\mathrm{P}_k(t)$ συγκλίνουν στις οριακές πιθανότητες $P_k > 0$:

Για $t \to \infty$, $\frac{d\mathrm{P}_k(t)}{dt} = 0$, $\mathrm{P}_k(t) \to P_k > 0$: ***Εργοδικές Οριακές Πιθανότητες***

Σημείωση: Ισχύει η ***εργοδική*** ιδιότητα και οι οριακές πιθανότητες μπορούν να προσεγγισθούν σαν $P_k = \lim_{T\to\infty}\{\frac{T_k}{T}\}$ όπου T_k είναι το σχετικό συνολικό χρονικό διάστημα T_k όταν $n(t) = k$ σε μεγάλο χρονικό ορίζοντα T ***μιας*** καταγραφής της ανέλιξης $n(t)$ σε ***ισορροπία***.

Οι εργοδικές οριακές πιθανότητες προκύπτουν από τις γραμμικά ανεξάρτητες ***Εξισώσεις Ισορροπίας***:

- $(\lambda_k + \mu_k)P_k = \lambda_{k-1}P_{k-1} + \mu_{k+1}P_{k+1},\ k > 1$
- $\lambda_0 P_0 = \mu_1 P_1$
- $P_0 + P_1 + \cdots + P_k + \cdots = 1$

ΔΙΑΔΙΚΑΣΙΑ ΓΕΝΝΗΣΕΩΝ – ΘΑΝΑΤΩΝ (3/4)

Birth – Death Processes *(Επανάληψη)*

Εφαρμογή σε Απλή Ουρά M/M/1

- Αφίξεις Poisson με μέσο ρυθμό λ αφίξεις/sec: $\lambda_k = \lambda, \;\; k = 0,1,2,3,\cdots$
- Χρόνοι εξυπηρέτησης εκθετικοί με μέση τιμή $\mathrm{E}(s) = \frac{1}{\mu}$ sec: $\mu_k = \mu, \; k = 1,2,3,\cdots$
- $\rho = \frac{\lambda}{\mu} < 1$ Erlang (συνθήκη για οριακή ισορροπία – εργοδικότητα)
- Η εξέλιξη των πιθανοτήτων $\mathrm{P}[n(t) = k] = \mathrm{P}_k(t)$ προκύπτει από το σύστημα διαφορικών εξισώσεων:
  - $\frac{d\mathrm{P}_k(t)}{dt} = \lambda \mathrm{P}_{k-1}(t) + \mu \mathrm{P}_{k+1}(t) - (\lambda + \mu)\mathrm{P}_k(t), \; k > 0$
  - $\frac{d\mathrm{P}_0(t)}{dt} = \mu \mathrm{P}_1(t) - \lambda \mathrm{P}_0(t)$
  - με αρχικές συνθήκες $\mathrm{P}_k(0)$ και οριακές συνθήκες $\sum_{k=0}^{\infty} \mathrm{P}_k(t) = 1 \;\; \forall t \geq 0$
- Στο όριο $t \to \infty$, $\frac{d\mathrm{P}_k(t)}{dt} = 0$, $\mathrm{P}_k(t) \to P_k > 0$, τις **εργοδικές πιθανότητες** που προκύπτουν από τις εξισώσεις ισορροπίας:
  - $\lambda P_0 = \mu P_1$ ή $P_1 = \left(\frac{\lambda}{\mu}\right) P_0 = \rho P_0$
  - $(\lambda + \mu)P_1 = \lambda P_0 + \mu P_2$ ή $P_2 = \rho^2 P_0$ και γενικά $P_k = \rho^k P_0, \;\; k > 0$
  - $P_0 + P_1 + \cdots + P_k + \cdots = 1 = P_0(1 + \rho + \rho^2 + \rho^3 + \cdots)$

Εφόσον $0 < \rho < 1$ η άπειρη δυναμοσειρά $(1 + \rho + \rho^2 + \rho^3 + \cdots) \to \frac{1}{1-\rho} \;\; \Rightarrow \;\; P_0(\frac{1}{1-\rho}) = 1$ και

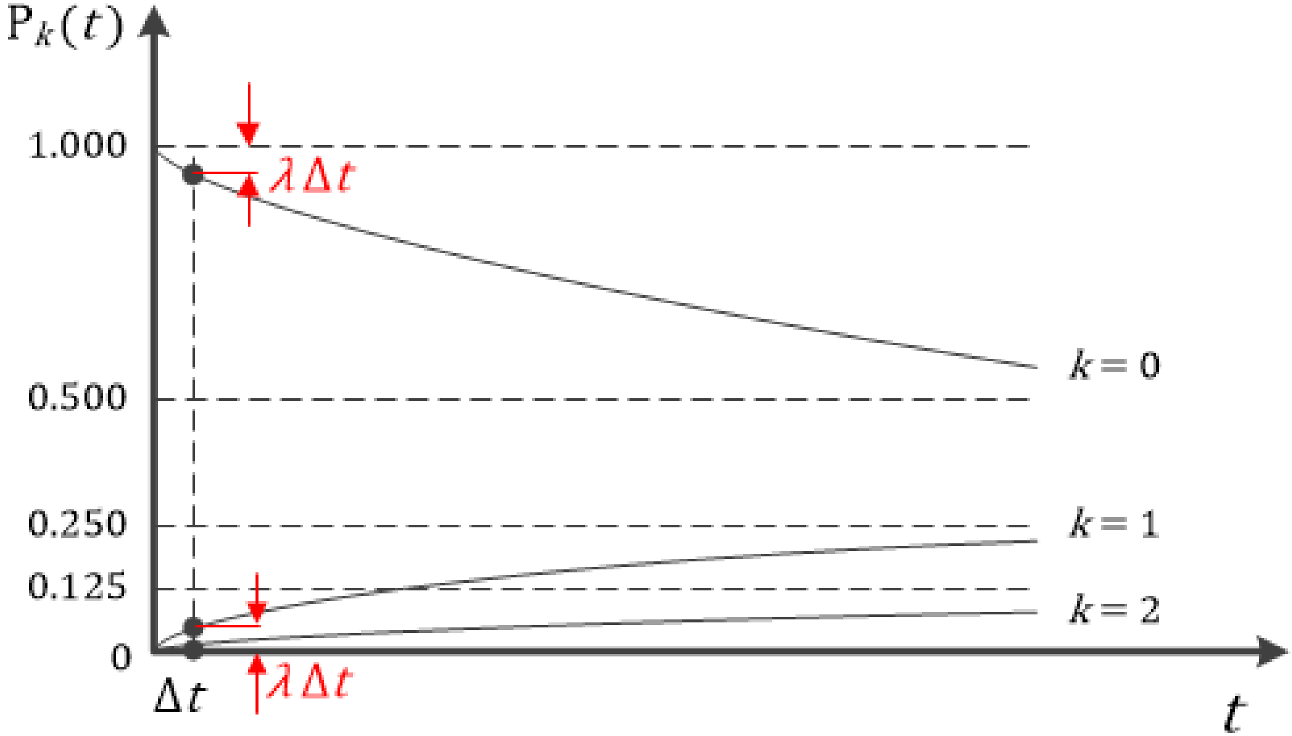
$$P_0 = (1 - \rho), \;\; P_k = (1 - \rho)\rho^k, \qquad k > 0$$

Μέσο μήκος ουράς M/M/1 σε ισορροπία: $E[n(t)] \triangleq \sum_{k=1}^{\infty} k P_k = \frac{\rho}{1-\rho}$

ΔΙΑΔΙΚΑΣΙΑ ΓΕΝΝΗΣΕΩΝ – ΘΑΝΑΤΩΝ (4/4)

Birth – Death Processes *(Επανάληψη)*

Χρονική Εξέλιξη Πιθανοτήτων Κατάστασης Απλής Ουράς M/M/1



$$\rho = \frac{\lambda}{\mu} = 0.5 \text{ Erlangs}$$

Αρχικές Συνθήκες: $P_0(0) = 1, \quad P_k(0) = 0$

Οριακές Εργοδικες Πιθανότητες:

$$\lim_{t\to\infty} \mathrm{P}_0(t) = P_0 = 1 - \rho = 0.500$$

$$\lim_{t\to\infty} \mathrm{P}_1(t) = P_1 = (1-\rho)\rho = 0.250$$

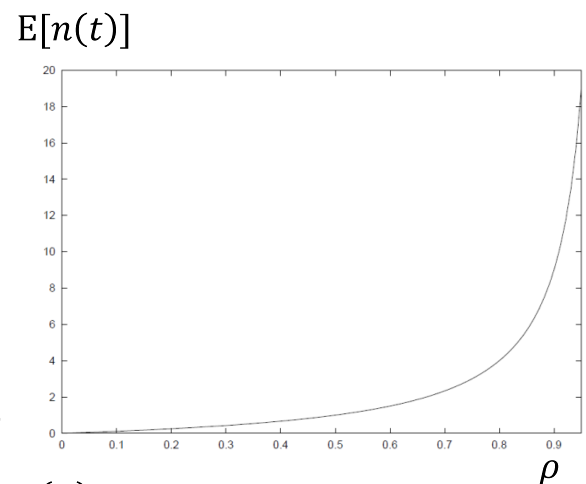
$$\lim_{t\to\infty} \mathrm{P}_2(t) = P_2 = (1-\rho)\rho^2 = 0.125$$

ΟΥΡΑ M/M/1 *(Επανάληψη)*

- Αφίξεις Poisson με μέσο ρυθμό λ αφίξεις/sec: $\lambda_k = \lambda = \gamma, \quad k = 0,1,2,3,\cdots$
- Χρόνοι εξυπηρέτησης εκθετικοί με μέση τιμή $\mathrm{E}(s) = \frac{1}{\mu}$ sec: $\mu_k = \mu, \ k = 1,2,3,\cdots$
- $\rho = u = \frac{\lambda}{\mu} < 1$ Erlang (συνθήκη για οριακή ισορροπία – εργοδικότητα)
- Οι **εργοδικές πιθανότητες** προκύπτουν από τις εξισώσεις ισορροπίας:
  - $\lambda P_0 = \mu P_1$ ή $P_1 = \left(\frac{\lambda}{\mu}\right) P_0 = \rho P_0$
  - $(\lambda + \mu) P_1 = \lambda P_0 + \mu P_2$ ή $P_2 = \rho^2 P_0$ και $P_k = \rho^k P_0, \quad k > 0$
  - $P_0 + P_1 + \cdots + P_k + \cdots = 1 = P_0(1 + \rho + \rho^2 + \rho^3 + \cdots)$

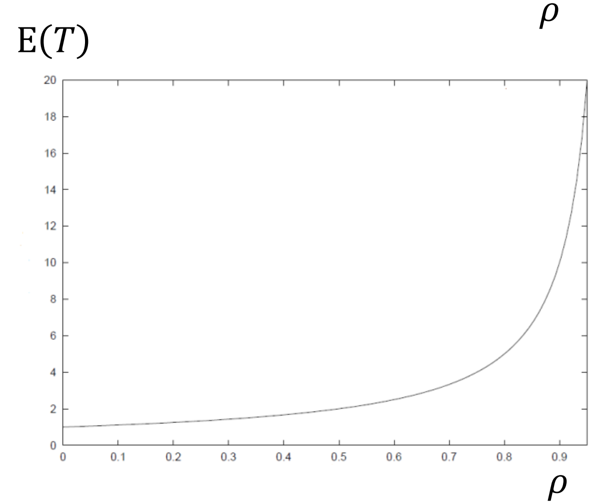
  Με $0 < \rho < 1$ η άπειρη δυναμοσειρά συγκλίνει, $P_0\left(\frac{1}{1-\rho}\right) = 1 \Rightarrow$

  $P_0 = (1-\rho), \quad P_k = (1-\rho)\rho^k, k > 0$ και $\mathrm{P}\{n(t) > 0\} = 1 - P_0 = \rho$



- **Μέση κατάσταση συστήματος** M/M/1 σε ισορροπία:

$$\mathrm{E}[n(t)] = \sum_{k=1}^{\infty} k P_k = \frac{\rho}{1-\rho}$$



- **Μέσος χρόνος καθυστέρησης**: Τύπος Little

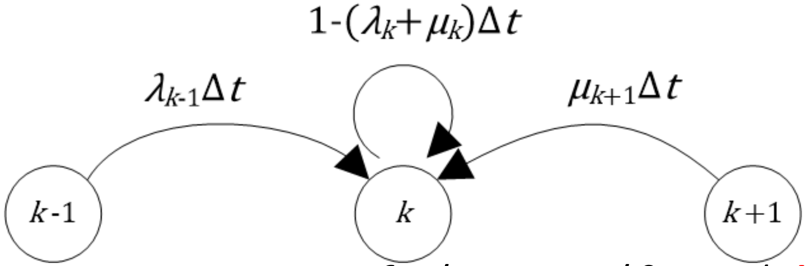
$$\mathrm{E}(T) = \frac{\mathrm{E}[n(t)]}{\gamma} = \frac{\mathrm{E}[n(t)]}{\lambda} = \frac{1/\mu}{1-\rho}$$

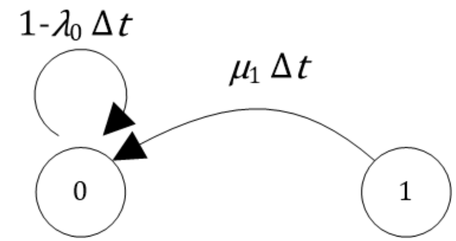
- **Μέσο μήκος ουράς & μέσος χρόνος αναμονής M/M/1**:

$$\mathrm{E}[n_q(t)] = \mathrm{E}[n(t)] - \rho, \qquad \mathrm{E}(W) = \mathrm{E}(T) - 1/\mu$$

ΕΞΙΣΩΣΕΙΣ ΙΣΟΡΡΟΠΙΑΣ (1/3)

- Birth-Death Process: Διάγραμμα ***Πιθανοτήτων Μεταβάσεων*** σε χρόνο $\Delta t \to 0$ προς $n(t) = k$



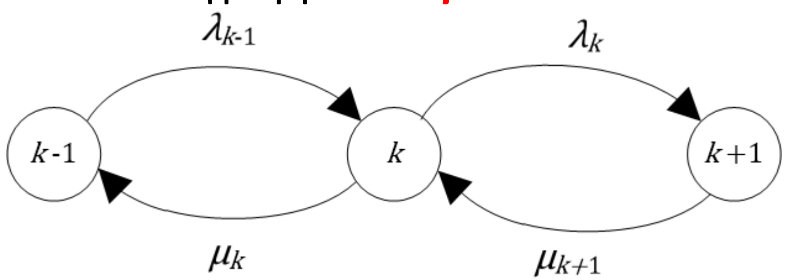


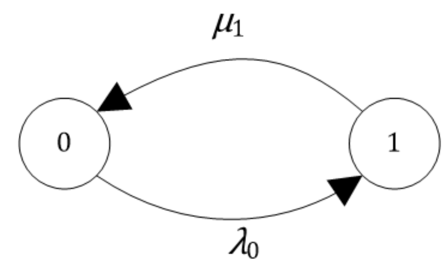
Εξισώσεις Μετάβασης (***Chapman - Kolmogorov***):

$$\mathrm{P}_k(t) = \lambda_{k-1}\Delta t \mathrm{P}_{k-1}(t-\Delta t) + \mu_{k+1}\Delta t \mathrm{P}_{k+1}(t-\Delta t) + [1-(\lambda_k+\mu_k)\Delta \mathrm{t}]\mathrm{P}_k(\mathrm{t}-\Delta \mathrm{t}), \qquad k \geq 1$$

$$\mathrm{P}_{\mathbf{0}}(t) \;=\; \mu_1\,\Delta t\;\mathrm{P}_1(t-\Delta t) \;+\; (1-\;\lambda_0\Delta t)\;\mathrm{P}_{\mathbf{0}}(t-\Delta t)$$

- Birth-Death Process: Διάγραμμα ***Ρυθμών Μεταβάσεων*** μεταξύ ***Εργοδικών*** Καταστάσεων





Εξισώσεις Ισορροπίας:

$(\lambda_k+\mu_k)P_k = \lambda_{k-1}P_{k-1} + \mu_{k+1}P_{k+1}$ για $k \geq 1$ και $\lambda_0 P_0 = \mu_1 P_1$

Σχετικές Πιθανότητες Μεταβάσεων $k \to (k+1), k \to (k-1)$:

$\mathrm{P}[k \to (k+1)/\text{μετάβαση}] = \lambda_k/(\lambda_k+\mu_k)$, $\mathrm{P}[k \to (k-1)/\text{μετάβαση}] = \mu_k/(\lambda_k+\mu_k)$

Dwell Time - Χρόνος Παραμονής στην $n(t) = k$ μέχρι την επόμενη μετάβαση

Εκθετική τυχαία μεταβλητή d_k **με μέσο** $1/(\lambda_k+\mu_k)$: Η μικρότερη δύο ανεξάρτητων εκθετικών τυχαίων μεταβλητών μέχρι *(1)* την επόμενη άφιξη με μέσο $1/\lambda_k$ ή *(2)* την ολοκλήρωση εξυπηρέτησης με μέσο $1/\mu_k$

$d_k = \min(x,y)\,, F_{d_k}(\tau) = \mathrm{P}\{d_k \leq \tau\} = 1 - \mathrm{P}\{d_k > \tau\} = 1 - e^{-(\lambda_k+\mu_k)\tau}$ διότι

$$\mathrm{P}\{d_k > \tau\} = \mathrm{P}\{x > \tau, y > \tau\} = \mathrm{P}\{x > \tau\}\mathrm{P}\{y > \tau\} = e^{-\lambda_k\tau}e^{-\mu_k\tau} = e^{-(\lambda_k+\mu_k)\tau}$$

ΕΞΙΣΩΣΕΙΣ ΙΣΟΡΡΟΠΙΑΣ (2/3)

- Απείρως επισκέψιμες καταστάσεις $\boldsymbol{s} = n(t)$ ***positive recurrent states***: Με μη μηδενικές εργοδικές πιθανότητες $\mathrm{P}\{n(t) = k\} = P_k(t) \to P_k > 0,\ k = 0,1,2, \ldots$

- Σε μεγάλο χρονικό διάστημα παρατήρησης T ισορροπούν οι αριθμοί μεταβάσεων από και προς την κατάσταση $\boldsymbol{s}$:

$$\#\{\text{ΜΕΤΑΒΑΣΕΩΝ ΠΡΟΣ ΤΗΝ ΚΑΤΑΣΤΑΣΗ } \boldsymbol{s}\} = \#\{\text{ΕΚΤΟΣ ΤΗΣ } \boldsymbol{s}\}$$

Σφαιρική Ισορροπία, ***Global Balance Equations***

- Σε μεγάλο χρονικό διάστημα παρατήρησης T ισορροπούν οι αριθμοί μεταβάσεων μεταξύ δύο (όχι αναγκαστικά γειτονικών) καταστάσεων $\boldsymbol{s}_1$ και $\boldsymbol{s}_2$:

$$\#\{\text{ΜΕΤΑΒΑΣΕΩΝ } \boldsymbol{s}_1 \to \boldsymbol{s}_2\} = \#\{\text{ΜΕΤΑΒΑΣΕΩΝ } \boldsymbol{s}_2 \to \boldsymbol{s}_1\}$$

Τοπική Ισορροπία, Local Balance Equations

- Λόγω ***εργοδικότητας*** σε μεγάλο χρονικό διάστημα παρατήρησης T, με T_1 και T_2 τους συνολικούς χρόνους παραμονής στις $\boldsymbol{s}_1, \boldsymbol{s}_2$:

(1) $\#\{\text{ΜΕΤΑΒΑΣΕΩΝ } \boldsymbol{s}_1 \to \boldsymbol{s}_2\} = \mathrm{T}_1 \times r_{1,2}$

(2) $\#\{\text{ΜΕΤΑΒΑΣΕΩΝ } \boldsymbol{s}_2 \to \boldsymbol{s}_1\} = \mathrm{T}_2 \times r_{2,1}$

Όπου $r_{1,2}$ και $r_{2,1}$ οι μέσοι ρυθμοί μετάβασης από $\boldsymbol{s}_1 \to \boldsymbol{s}_2$ και $\boldsymbol{s}_2 \to \boldsymbol{s}_1$

- Λόγω ***ισορροπίας:*** (1) = (2) και $r_{1,2} \lim_{\mathrm{T}\to\infty} \frac{\mathrm{T}_1}{\mathrm{T}} = r_{2,1} \lim_{\mathrm{T}\to\infty} \frac{\mathrm{T}_2}{\mathrm{T}}$ ή

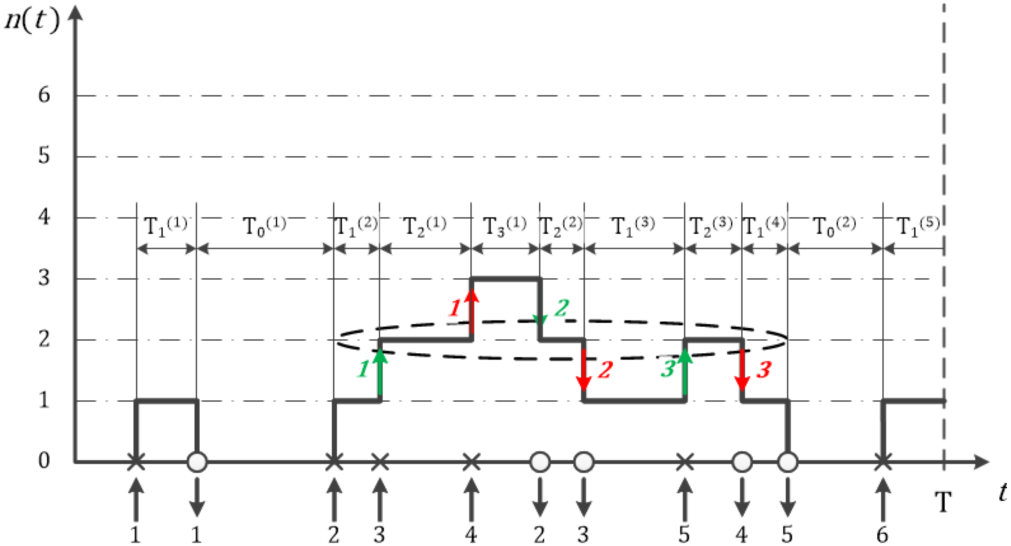
$$r_{1,2} P_1 = r_{2,1} P_2$$ ***Local Balance Equations***

ΕΞΙΣΩΣΕΙΣ ΙΣΟΡΡΟΠΙΑΣ (3/3)

Σχηματική Απεικόνιση

Εξισώσεις Ισορροπίας στην Εργοδική Κατάσταση

Χρόνος Παρατήρησης Δείγματος $n(t)$: T



Global Balance Equation

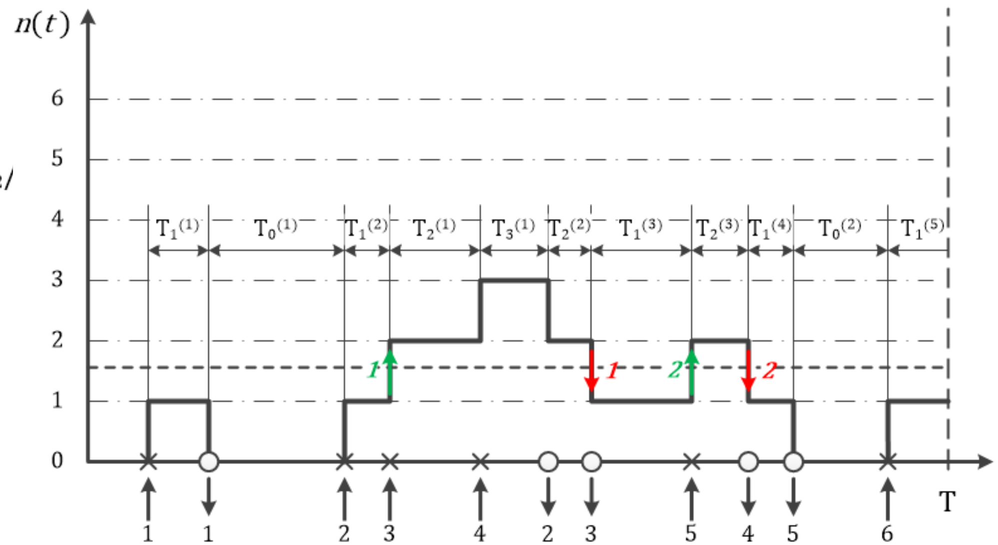
$T_1 = T_1^{(1)} + T_1^{(2)} + T_1^{(3)} + T_1^{(4)} + T_1^{(5)}$

$T_2 = T_2^{(1)} + T_2^{(2)} + T_2^{(3)}$, $T_3 = T_3^{(1)}$

(State 2: Transitions In = Transitions Out)

$\lambda T_1 + \mu T_3 \sim \lambda T_2 + \mu T_2 \Rightarrow (\lambda T_1 + \mu T_3)/T \sim (\lambda + \mu)\ (T_2/$

$\Rightarrow \lambda P_1 + \mu P_3 = (\lambda + \mu) P_2$



Local Balance Equation

$T_1 = T_1^{(1)} + T_1^{(2)} + T_1^{(3)} + T_1^{(4)} + T_1^{(5)}$

$T_2 = T_2^{(1)} + T_2^{(2)} + T_2^{(3)}$

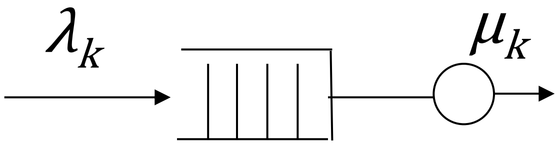
(Transitions 1->2 = Transitions 2->1)

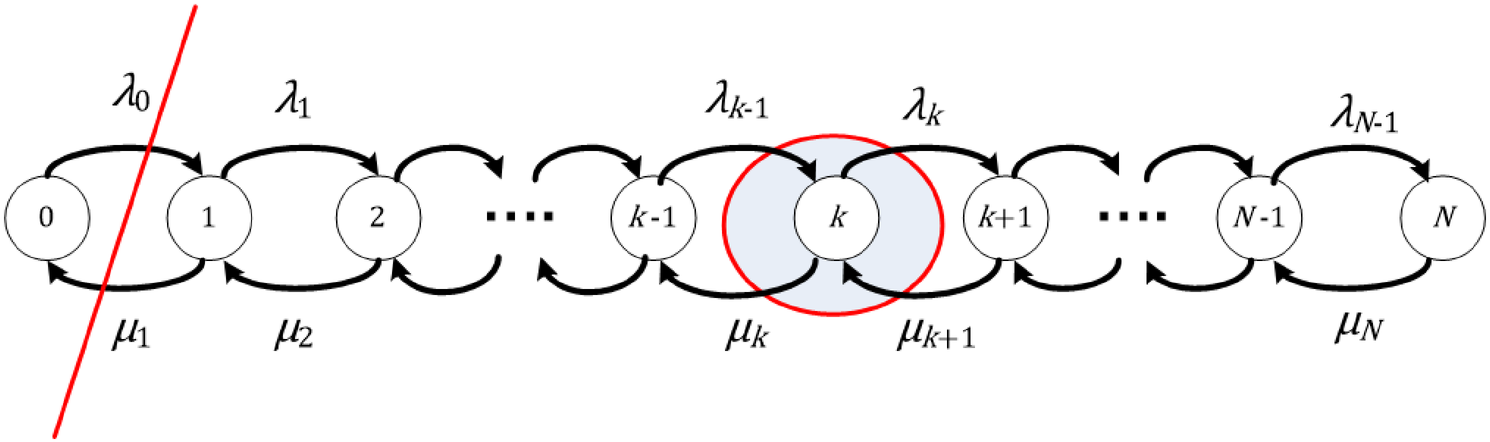
$\lambda T_1 \sim \mu T_2 \Rightarrow \lambda (T_1/T) \sim \mu (T_2/T) \Rightarrow \lambda P_1 = \mu P_2$

ΟΥΡΑ M/M/1/*N* (1/2)

- Συστήματα M/M/1/*N* με ρυθμούς άφιξης και ρυθμούς εξυπηρέτησης εξαρτώμενους από τον αριθμό των πελατών στο σύστημα (από την παρούσα κατάσταση του συστήματος)

  (State Dependent M/M/1/N Queues)





Local Balance Equation

$$\lambda_0 P_0 = \mu_1 P_1$$

$$\lambda_{k-1} P_{k-1} = \mu_k P_k \qquad k = 1, 2, \ldots N$$

Global Balance Equation

$$\lambda_0 P_0 = \mu_1 P_1$$

$$(\lambda_k + \mu_k) P_k = \lambda_{k-1} P_{k-1} + \mu_{k+1} P_{k+1} \qquad k = 1, \ldots, N$$

Κανονικοποίηση Εργοδικών Πιθανοτήτων

$$P_0 + \cdots + P_N = 1$$

ΟΥΡΑ M/M/1/N (2/2)

- Σταθεροί μέσοι ρυθμοί αφίξεων (γεννήσεων)

  $\lambda_k = \lambda$, Poisson, $k = 1,2,\ldots,N$
- Σταθεροί μέσοι ρυθμοί εξυπηρέτησης (θανάτων)

  $\mu_k = \mu, \; k = 1,2,\ldots,N$

  Εκθετικοί ανεξάρτητοι χρόνοι εξυπηρέτησης $s, E(s) = 1/\mu$
- Εργοδικές πιθανότητες καταστάσεων

  $P_k = \rho^k P_0, \;\; k = 0,1,2,\ldots,N$

  $P_0 + P_1 + \cdots + P_{N-1} + P_N = 1$

  $\rho = \lambda/\mu$ Erlangs (η M/M/1/N είναι **πάντα ευσταθής** γιατί υπερβολικό φορτίο δεν προωθείται)
- Αντικαθιστώντας με τον τύπο πεπερασμένου αθροίσματος γεωμετρικής προόδου:

$$P_0 = \frac{1-\rho}{1-\rho^{N+1}}, \qquad \rho \neq 1$$

$$P_0 = \frac{1}{N+1}, \qquad \rho = 1$$

- Χρησιμοποίηση Εξυπηρετητή (Server Utilization) $U = 1 - P_0$
- Ρυθμαπόδοση (throughput) $\gamma = \lambda(1 - P_N) = \mu(1 - P_0) = \mu U$
- Πιθανότητα απώλειας $P_{blocking} = P_N$

- Στάσιμος Εργοδικός μέσος όρος πληθυσμού – κατάστασης

$$\mathrm{E}[n(t)] \to \mathrm{E}(k) = \sum_{k=1}^{N} k P_k = \rho \frac{1-(N+1)\rho^N + N\rho^{N+1}}{(1-\rho)(1-\rho^{N+1})}$$

- Νόμος του Little: $\mathrm{E}(T) = \mathrm{E}(k)/\gamma = \mathrm{E}(k)/[\lambda(1 - P_N)]$