Στοχαστικά Μοντέλα στην Επιχειρησιακή Έρευνα Μέρος 3 Θεωρία Ουρών Αναμονής

Αντώνης Οικονόμου

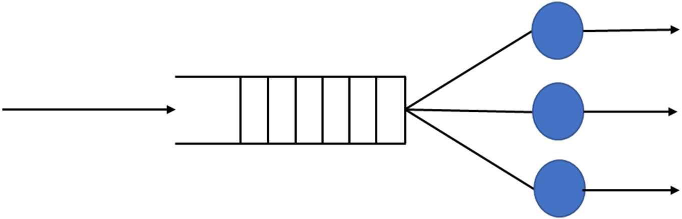
Μεταπτυχιακό Πρόγραμμα Σπουδών
Στατιστική και Επιχειρησιακή Έρευνα

19/10/2022-4/11/2022

Εισαγωγή στη Θεωρία Ουρών Αναμονής

Ουρά αναμονής - Βασική ιδέα

- Ουρά Αναμονής ή Σύστημα εξυπηρέτησης
  = Στοχαστικό σύστημα εισόδου-εξόδου με διακριτές μονάδες.



Σχήμα: Ουρά αναμονής

Βασικά χαρακτηριστικά - Ονοματολογία Kendall

- Βασικά χαρακτηριστικά:
  1. Διαδικασία αφίξεων (A).
  2. Χρόνοι εξυπηρέτησης (B).
  3. Αριθμός παράλληλων υπηρετών (c).
  4. Χωρητικότητα (k).
  5. Πειθαρχία ουράς ( ).
- Ονοματολογία Kendall:
  A/B/c/k( ).

Βασικά χαρακτηριστικά - τιμές

- Ενδιάμεσοι χρόνοι αφίξεων, χρόνοι εξυπηρέτησης: Θεωρούνται γενικά ανεξάρτητοι και ισόνομοι. A,B:
  1. G ή GI (General Independent) Γενική κατανομή.
  2. M (Markovian, Memoryless) Εκθετική κατανομή.
  3. D (Deterministic) Σταθεροί χρόνοι.
  4. E_k (Erlang) Erlang κατανομή.
  5. H_k (Hyperexponential) Υπερεκθετική κατανομή.
- Αριθμός παράλληλων υπηρετών: $c = 1, 2, \ldots$.
- Χωρητικότητα: $k = c, c + 1, c + 2, \ldots$.
- Πειθαρχία ουράς ( ):
  1. (FCFS) First-Come-First-Served.
  2. (LCFS) Last-Come-First-Served.
  3. (SIRO) Service-In-Random-Order.
  4. (SSTF) Shortest-Service-Time-First.

Βασικές παράμετροι

- Ρυθμός αφίξεων λ ή μέσος ενδιάμεσος χρόνος αφίξεων a. $\lambda = \frac{1}{a}$.
- Ρυθμός εξυπηρέτησης μ ή μέσος χρόνος εξυπηρέτησης b. $\mu = \frac{1}{b}$.
- Αριθμός υπηρετών c.
- Χωρητικότητα k.
- Πειθαρχία.
- Αν γνωρίζω μόνο μέσους χρόνους τότε χρησιμοποιώ για τη μοντελοποίηση την εκθετική κατανομή.
- Από όλες τις μη-αρνητικές τ.μ. με δοσμένη μέση τιμή, η εκθετική έχει την μικρότερη μεροληψία μοντελοποίησης (μέγιστη εντροπία).

Βασικές περιοχές Θεωρίας Ουρών Αναμονής

- Αποτίμηση απόδοσης. (Στοχαστικές Ανελίξεις).
- Σύγκριση συστημάτων. (Στοχαστικές Ανελίξεις + Στοχαστικές Διατάξεις).
- Βέλτιστος σχεδιασμός. (Στοχαστικές Ανελίξεις + Μη-γραμμικός, Ακέραιος Προγραμματισμός).
- Βέλτιστος δυναμικός έλεγχος. (Στοχαστικός Δυναμικός Προγραμματισμός).
- Στρατηγική συμπεριφορά. (Θεωρία Παιγνίων).

Στοχ. διαδικασίες αποτίμησης απόδοσης ουρών

- $Q(t)$: Αριθμός πελατών στο σύστημα τη στιγμή t.
- $Q_q(t)$: Αριθμός πελατών σε αναμονή τη στιγμή t.
- $Q_s(t)$: Αριθμός εξυπηρετούμενων πελατών τη στιγμή t.
- S_n: Χρόνος παραμονής n-οστού πελάτη.
- W_n: Χρόνος αναμονής n-οστού πελάτη.
- B_n: Χρόνος εξυπηρέτησης n-οστού πελάτη.
- I_n: n-οστή περίοδος αργίας.
- Y_n: n-οστή περίοδος συνεχούς λειτουργίας.
- Z_n: n-οστός κύκλος απασχόλησης (από πελάτη που φθάνει σε κενό σύστημα σε επόμενο πελάτη που φθάνει σε κενό σύστημα).

Εμφυτευμένες διαδικασίες

- $A_1 \leq A_2 \leq A_3 \leq \ldots$: Στιγμές διαδοχικών αφίξεων.
- $D_1 \leq D_2 \leq D_3 \leq \ldots$: Στιγμές διαδοχικών αναχωρήσεων.
- $Q_n^- = Q(A_n^-)$: Πλήθος πελατών που βρίσκει στο σύστημα η n-οστή άφιξη.
- $Q_n^+ = Q(D_n^+)$: Πλήθος πελατών που αφήνει στο σύστημα η n-οστή αναχώρηση.

Βασικά μέτρα απόδοσης ουρών

- Η διαδικασία $\{Q(t)\}$ είναι αναγεννητική.
- Στιγμές αναγέννησης: Σημεία έναρξης των κύκλων απασχόλησης (αφίξεις πελατών σε κενό σύστημα).
- Έστω Z ο πρώτος κύκλος απασχόλησης και έστω ότι έχει απεριοδική κατανομή.
- Λόγω της αναγεννητικότητας:

$$\begin{aligned} \lim_{t\to\infty} \frac{\int_0^t 1_{\{Q(u)=j\}} du}{t} &= \lim_{t\to\infty} \frac{E\left[\int_0^t 1_{\{Q(u)=j\}} du\right]}{t} \quad \text{με πιθ. } 1 \\ &= \lim_{t\to\infty} \frac{\int_0^t \Pr[\{Q(u)=j]\} du}{t} = \lim_{t\to\infty} \Pr[Q(t)=j] \\ &= \frac{E[\int_0^Z 1_{\{Q(u)=j\}} du]}{E[Z]} \overset{\text{ορσ}}{=} p_j = \Pr[Q=j]. \end{aligned}$$

Βασικά μέτρα απόδοσης ουρών (συνέχεια)

- Για την εμφυτευμένη διαδικασία σε στιγμές αφίξεων έχουμε ανάλογα:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} 1_{\{Q_k^- = j\}} = \lim_{n\to\infty} \frac{1}{n} E\left[\sum_{k=1}^{n} 1_{\{Q_k^- = j\}}\right] \text{ με πιθ. } 1$$

$$= \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \Pr[Q_k^- = j] = \lim_{n\to\infty} \Pr[Q_n^- = j]$$

$$= \frac{1}{E[A(Z)]} E\left[\sum_{k=1}^{A(Z)} 1_{\{Q_k^- = j\}}\right] \stackrel{\text{ορσ}}{=} a_j = \Pr[Q^- = j],$$

όπου Z ο πρώτος κύκλος λειτουργίας και $A(Z)$ το πλήθος των αφίξεων σε αυτόν.

Βασικά μέτρα απόδοσης ουρών (συνέχεια)

- Ομοίως ορίζουμε d_j την πιθανότητα j πελατών σε στιγμές αναχωρήσεων.
- Συνοπτικά:
  1. (p_j): Κατανομή ισορροπίας αριθμού πελατών στο σύστημα σε συνεχή χρόνο.
  2. (a_j): Κατανομή ισορροπίας αριθμού πελατών στο σύστημα σε στιγμές αφίξεων.
  3. (d_j): Κατανομή ισορροπίας αριθμού πελατών στο σύστημα σε στιγμές αναχωρήσεων.

Βασικά μέτρα απόδοσης ουρών (συνέχεια)

- Για την κατανομή του χρόνου παραμονής πελάτη έχουμε:

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}1_{\{S_k^-\le x\}} = \lim_{n\to\infty}\frac{1}{n}E\left[\sum_{k=1}^{n}1_{\{S_k^-\le x\}}\right] \text{ με πιθ. } 1$$

$$= \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\Pr[S_k^-\le x] = \lim_{n\to\infty}\Pr[S_n^-\le x]$$

$$= \frac{1}{E[A(Z)]}E\left[\sum_{k=1}^{A(Z)}1_{\{S_k^-\le x\}}\right] \overset{\text{ορσ}}{=} F_S(x) = \Pr[S\le x],$$

όπου Z ο πρώτος κύκλος λειτουργίας και $A(Z)$ το πλήθος των αφίξεων σε αυτόν.

Βασικά μέσα μέτρα απόδοσης ουρών

- Μέσος αριθμός πελατών στο σύστημα:

$$E[Q] \overset{\text{ορσ}}{=} \lim_{t\to\infty} \frac{\int_0^t Q(u)du}{t} = \lim_{t\to\infty} \frac{E\left[\int_0^t Q(u)du\right]}{t} \quad \text{με πιθ. } 1$$

$$= \lim_{t\to\infty} E[Q(t)] = \frac{E[\int_0^Z Q(u)du]}{E[Z]}.$$

- Μέσος χρόνος παραμονής πελάτη στο σύστημα:

$$E[S] \overset{\text{ορσ}}{=} \lim_{n\to\infty} \frac{\sum_{k=1}^n S_k}{n} = \lim_{n\to\infty} \frac{E\left[\sum_{k=1}^n S_k\right]}{n} \quad \text{με πιθ. } 1$$

$$= \lim_{n\to\infty} E[S_n] = \frac{E\left[\sum_{k=1}^{A(Z)} S_k\right]}{E[A(Z)]}.$$

4 Βασικά αποτελέσματα

1. Χαρακτηρισμός ευστάθειας σε συστήματα G/G/c.
2. Ιδιότητα μεμονωμένων μεταβάσεων (αφίξεων και αναχωρήσεων).
3. Ιδιότητα PASTA.
4. Νόμος του Little.

Χαρακτηρισμός ευστάθειας σε συστήματα G/G/c

- G/G/c με απεριοδική κατανομή χρόνων εξυπηρέτησης ή ενδιάμεσων χρόνων αφίξεων (ή και τα δύο).
- Ρυθμός συνωστισμού = Ρυθμός αφίξεων × Μέσος χρόνος εξυπηρέτησης:

$$\rho = \lambda \times b.$$

- $\rho < c \Leftrightarrow$ Ευστάθεια:
  $p_n, a_n, d_n > 0$ και $\sum_n p_n = \sum_n a_n = \sum_n d_n = 1.$
- $\rho \geq c \Leftrightarrow$ Αστάθεια:
  $p_n = a_n = d_n = 0$ και $\lim_{t\to\infty} Q(t) = \infty$ με πιθ. 1.

Ιδιότητα μεμονωμένων μεταβάσεων

- Μεμονωμένες αφίξεις και αναχωρήσεις $\Rightarrow (a_n) = (d_n)$.
- Βασικές τυχαίες μεταβλητές για την απόδειξη:
  1. $A(t)$: Πλήθος αφίξεων στο $(0, t]$.
  2. $D(t)$: Πλήθος αναχωρήσεων στο $(0, t]$.
  3. $A_j(t)$: Πλήθος αφίξεων στο $(0, t]$ που βρίσκουν j πελάτες.
  4. $D_j(t)$: Πλήθος αναχωρήσεων στο $(0, t]$ που αφήνουν j πελάτες.
- Είναι

$$
\begin{aligned}
a_j &= \lim_{t\to\infty} \frac{A_j(t)}{A(t)}, \\
d_j &= \lim_{t\to\infty} \frac{D_j(t)}{D(t)}.
\end{aligned}
$$

Ιδιότητα μεμονωμένων μεταβάσεων (αιτιολόγηση)

- Από ισότητα ρυθμών αφίξεων και αναχωρήσεων:

$$\lim_{t\to\infty} \frac{A(t)}{t} = \lim_{t\to\infty} \frac{D(t)}{t}.$$

- Από εναλλαγή άνω και κάτω διασχίσεων του ορίου των j πελατών:

$$|A_j(t) - D_j(t)| \leq 1 \Rightarrow \lim_{t\to\infty} \frac{A_j(t)}{t} = \lim_{t\to\infty} \frac{D_j(t)}{t}.$$

- Επομένως:

$$a_j = \lim_{t\to\infty} \frac{A_j(t)}{A(t)} = \lim_{t\to\infty} \frac{A_j(t)/t}{A(t)/t} = \lim_{t\to\infty} \frac{D_j(t)/t}{D(t)/t} = \lim_{t\to\infty} \frac{D_j(t)}{D(t)} = d_j.$$

Ιδιότητα PASTA

- Poisson διαδικασία αφίξεων $\Rightarrow (a_n) = (p_n)$.
- Αν $A(t, t+h)$ το πλήθος των αφίξεων στο $(t, t+h]$, τότε:

$$
\begin{aligned}
a_j &= \lim_{t\to\infty} \lim_{h\to 0^+} \Pr[Q(t) = j | A(t, t+h) > 0] \\
&= \lim_{t\to\infty} \lim_{h\to 0^+} \frac{\Pr[Q(t) = j, A(t, t+h) > 0]}{\Pr[A(t, t+h) > 0]} \\
&= \lim_{t\to\infty} \lim_{h\to 0^+} \frac{\Pr[Q(t) = j] \Pr[A(t, t+h) > 0 | Q(t) = j]}{\Pr[A(t, t+h) > 0]} \\
&= p_j.
\end{aligned}
$$

Νόμος του Little

- Μέσος αριθμός πελατών = Ρυθμός αφίξεων × Μέσος χρόνος παραμονής:

$$E[Q] = \lambda E[S].$$

- Θεμελιώδης ισότητα: Σε έναν κύκλο απασχόλησης Z με αριθμό αφίξεων $A(Z)$:

$$\int_0^Z Q(u)du = \sum_{k=1}^{A(Z)} S_k.$$

Νόμος του Little (απόδειξη)

- Έχουμε:

$$\begin{aligned} E[Q] &= \frac{E\left[\int_0^Z Q(u)du\right]}{E[Z]} = \frac{E[A(Z)]}{E[Z]} \cdot \frac{E\left[\int_0^Z Q(u)du\right]}{E[A(Z)]} \\ &= \frac{E[A(Z)]}{E[Z]} \cdot \frac{E\left[\sum_{k=1}^{A(Z)} S_k\right]}{E[A(Z)]} = \lambda E[S]. \end{aligned}$$

Ανάλυση Μέσης Τιμής

Άμεσες συνέπειες του Νόμου Little

- Εφαρμογή του Ν. Little στον χώρο αναμονής: Μέσος αριθμός πελατών σε αναμονή = Ρυθμός αφίξεων × Μέσος χρόνος αναμονής:

$$E[Q_q] = \lambda E[W].$$

- Εφαρμογή του Ν. Little στον χώρο εξυπηρέτησης: Μέσος αριθμός πελατών σε εξυπηρέτηση = Ρυθμός αφίξεων × Μέσος χρόνος εξυπηρέτησης:

$$E[Q_s] = \lambda b = \rho.$$

Άμεσες συνέπειες του Νόμου Little (συνέχεια)

- Από την
$$E[Q_s] = \lambda b = \rho,$$
έχουμε:
    1. Μέσος αριθμός απασχολημένων υπηρετών σε G/G/c σύστημα: ρ.
    2. Ποσοστό χρόνου που ο υπηρέτης είναι απασχολημένος σε G/G/c σύστημα: $\frac{\rho}{c}$.
    3. Πιθανότητα κενού συστήματος σε G/G/1 σύστημα: $p_0 = 1 - \rho$.
- Πράγματι, στην G/G/1 ουρά έχουμε

$$\rho = E[Q_s] = \Pr[Q_s = 1] = \Pr[Q \geq 1] = 1 - p_0.$$

Ερμηνείες του ρυθμού συνωστισμού ρ

- Ο ρυθμός συνωστισμού $\rho = \lambda b$ εκφράζει:
  1. τη μέση ποσότητα εργασίας που εισέρχεται στο σύστημα ανά χρονική μονάδα (μετρημένη σε χρόνο διεκπεραίωσης),
  2. τον αριθμό των απασχολημένων υπηρετών σε συστήματα G/G/c,
  3. το ποσοστό του χρόνου που ο υπηρέτης είναι απασχολημένος σε συστήματα G/G/1,
  4. την πιθανότητα μη-κενού συστήματος σε συστήματα G/G/1.

Ανάλυση Μέσης Τιμής - Μεθοδολογία

- Ανάλυση Μέσης Τιμής (AMT) $\rightarrow$ Γρήγορος υπολογισμός των $E[Q]$ και $E[S]$ ενός συστήματος.
- Βασική ιδέα: Επίλυση συστήματος δυο εξισώσεων με δυο αγνώστους.
- 1η εξίσωση: Ο Ν. Little:

$$E[Q] = \lambda E[S].$$

- 2η εξίσωση: Από τον υπολογισμό του χρόνου παραμονής πελάτη, S, δεσμεύοντας στον αριθμό πελατών που βρίσκει ο πελάτης + PASTA:

$$E[S] = \sum_j \Pr[Q^- = j] E[S|Q^- = j]$$

$$\rightarrow \quad E[S] = f(E[Q^-]) \stackrel{PASTA}{=} f(E[Q]).$$

Ανάλυση Μέσης Τιμής - Επεκτάσεις

- Εφαρμόζοντας τον Ν. Little σε διάφορα υποσυστήματα ενός συστήματος εξυπηρέτησης, μπορούμε να βρούμε όχι μόνο τις μέσες τιμές $E[Q]$, $E[S]$, αλλά και ολόκληρη την κατανομή (p_j).

AMT - M/G/1/1 ουρά

- N. Little:

$$E[Q] = \lambda E[S] \Rightarrow p_1 = \lambda E[S].$$

- Δέσμευση για $E[S]$ στην Q^-:

$$\begin{aligned} E[S] &= \Pr[Q^- = 0]E[S|Q^- = 0] + \Pr[Q^- = 1]E[S|Q^- = 1] \\ &= a_0 \cdot b + a_1 \cdot 0 \overset{PASTA}{=} bp_0. \end{aligned}$$

- Άρα $p_1 = \lambda b p_0 = \rho p_0$. Επίσης $p_0 + p_1 = 1$, οπότε:

$$p_0 = \frac{1}{1+\rho},\ p_1 = \frac{\rho}{1+\rho},\ E[Q] = \frac{\rho}{1+\rho},\ E[S] = \frac{1}{\mu(1+\rho)}.$$

AMT - M/M/1 ουρά - Μέσες τιμές

- N. Little:

$$E[Q] = \lambda E[S].$$

- Δέσμευση για $E[S]$ στην Q^-:

$$E[S] = E[E[S|Q^-]] = E[\sum_{i=1}^{Q^-+1} B_i],$$

όπου B_1 είναι ο υπολειπόμενος χρόνος εξυπηρέτησης του πελάτη που εξυπηρετείται (αν εξυπηρετείται κάποιος), $B_2, B_3, \dots, B_{Q^-}$ οι χρόνοι εξυπηρέτησης των πελατών στον χώρο αναμονής (αν υπάρχουν) και B_{Q^-+1} ο χρόνος εξυπηρέτησης του αφικνούμενου πελάτη.

AMT - M/M/1 ουρά - Μέσες τιμές (συνέχεια)

- N. Little:

$$E[Q] = \lambda E[S].$$

- Δέσμευση για $E[S]$ στην Q^- + Αμνήμονη ιδιότητα εκθετικής + PASTA:

$$E[S] = E[Q^- + 1]E[B] = \frac{E[Q^-] + 1}{\mu} = \frac{E[Q] + 1}{\mu}.$$

- Λύνοντας το σύστημα:

$$E[Q] = \frac{\rho}{1-\rho}, \ E[S] = \frac{1}{\mu(1-\rho)},$$

όπου $\rho = \frac{\lambda}{\mu}$ είναι ο ρυθμός συνωστισμού.

AMT - M/M/1 ουρά - Κατανομή ισορροπίας

- Εφαρμόζοντας Ν. Little στη θέση j του συστήματος: Μέσο πλήθος πελατών στην θέση j $=$ Ρυθμός αφίξεων στη θέση j $\times$ Μέσος χρόνος παραμονής στη θέση j.
- Μέσο πλήθος πελατών στην θέση j $=$Πιθ. υπάρχει πελάτης στη θέση j $=\sum_{k=j}^{\infty} p_k$.
- Ρυθμός αφίξεων στη θέση j $=$ Ρυθμός αφίξεων $\times$ ποσοστό πελατών που περνάνε από τη θέση j $= \lambda \sum_{k=j-1}^{\infty} a_k \overset{\text{PASTA}}{=} \lambda \sum_{k=j-1}^{\infty} p_k$.
- Μέσος χρόνος παραμονής στη θέση $j$$= \frac{1}{\mu}$.
- Άρα

$$\sum_{k=j}^{\infty} p_k = \rho \sum_{k=j-1}^{\infty} p_k.$$

AMT - M/M/1 ουρά - Κατανομή ισορροπίας (συνέχεια)

- Αφαιρούμε την εξίσωση για την κατάσταση $j+1$ από την εξίσωση για την κατάσταση j:

$$\sum_{k=j}^{\infty} p_k = \rho \sum_{k=j-1}^{\infty} p_k \quad (\text{εξίσωση για } j)$$

$$- \quad \sum_{k=j+1}^{\infty} p_k = \rho \sum_{k=j}^{\infty} p_k \quad (\text{εξίσωση για } j+1)$$

- Άρα:

$$p_j = \rho p_{j-1}, \quad j \geq 1.$$

- Επίσης, για $j = 1$ έχουμε $1 - p_0 = \rho$, οπότε

$$p_j = (1-\rho)\rho^j, \quad j \geq 0.$$

AMT - M/G/1 ουρά - Μέσες τιμές

- N. Little:

$$E[Q] = \lambda E[S].$$

- Δέσμευση για $E[S]$ στην Q^-: $S = \sum_{i=1}^{Q^-+1} B_i$, αλλά ο B_1 εξαρτάται από το Q^-.
- Οπότε δεσμεύουμε στο κατά πόσον $Q^- = 0$ ή $Q^- > 0$:

$$\begin{aligned} E[S] &= \Pr[Q^- = 0] E\left[B_1 | Q^- = 0\right] \\ &\quad + \Pr[Q^- > 0] E\left[B_1 | Q^- > 0\right] \\ &\quad + \Pr[Q^- > 0] E\left[\sum_{i=2}^{Q^-+1} B_i | Q^- > 0\right]. \end{aligned}$$

AMT - M/G/1 ουρά - Μέσες τιμές (συνέχεια)

- Έχουμε:

$$\begin{aligned} E[S] &= \Pr[Q^- = 0] E\left[B_1 | Q^- = 0\right] \\ &\quad + \Pr[Q^- > 0] E\left[B_1 | Q^- > 0\right] \\ &\quad + \Pr[Q^- > 0] E\left[\sum_{i=2}^{Q^- + 1} B_i | Q^- > 0\right]. \end{aligned}$$

- $\Pr[Q^- = 0] = 1 - \rho$ (Πόρισμα N. Little - PASTA), $\Pr[Q^- > 0] = \rho$.
- $E\left[B_1 | Q^- = 0\right] = b$.
- $E\left[\sum_{i=2}^{Q^-+1} B_i | Q^- > 0\right] = E[Q^- | Q^- > 0] b$.
- $E\left[B_1 | Q^- > 0\right] = E[R_B] = \frac{E[B^2]}{2E[B]} = \frac{b_2}{2b} = \frac{b^2 + \sigma_B^2}{2b}$ (μέσος υπολειπόμενος ανανεωτ. διαδικασίας με ενδιάμεσους χρόνους τους χρόνους εξυπηρέτησης).

AMT - M/G/1 ουρά - Μέσες τιμές (συνέχεια)

- Τελικά:

$$\begin{aligned}
E[S] &= \Pr[Q^- = 0] E\left[B_1 | Q^- = 0\right] \\
&\quad + \Pr[Q^- > 0] E\left[B_1 | Q^- > 0\right] \\
&\quad + \Pr[Q^- > 0] E\left[\sum_{i=2}^{Q^-+1} B_i | Q^- > 0\right] \\
&= (1-\rho)b + \rho E[R_B] + \Pr[Q^- > 0] E[Q^- | Q^- > 0] b \\
&= (1-\rho)b + \rho E[R_B] + E[Q^- 1_{\{Q^- > 0\}}] b \\
&= (1-\rho)b + \rho E[R_B] + E[Q^-] b \\
&= (1-\rho)b + \rho E[R_B] + E[Q] b.
\end{aligned}$$

- Άρα:

$$E[Q] = (1-\rho)\lambda b + \lambda\rho E[R_B] + E[Q]\lambda b.$$

AMT - M/G/1 ουρά - Μέσες τιμές (συνέχεια)

- Επομένως:

$$E[Q] = \rho + \frac{\rho}{1-\rho}\lambda E[R_B],$$

και διαιρώντας με λ παίρνουμε και τον μέσο χρόνο παραμονής πελάτη στο σύστημα

$$E[S] = b + \frac{\rho}{1-\rho}E[R_B].$$

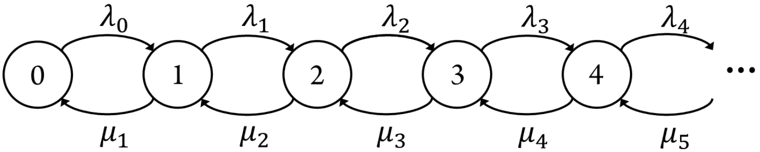
- Εύκολος μνημονικός τρόπος:
Μέσος χρόνος αναμονής στην M/G/1 ουρά = Μέσος αριθμός πελατών στην M/M/1 ουρά × μέση τιμή κατανομής ισορροπίας των χρόνων εξυπηρέτησης:

$$E[W] = \frac{\rho}{1-\rho}E[R_B].$$

Απλές Μαρκοβιανές Ουρές

Απλές Μαρκοβιανές Ουρές: Ορισμός

- Σύστημα εξυπηρέτησης όπου $\{Q(t)\}$ Μ.α.σ.χ = Μαρκοβιανή ουρά.
- Σύστημα εξυπηρέτησης όπου $\{Q(t)\}$ αλυσίδα γέννησης - θανάτου = Απλή Μαρκοβιανή ουρά.



- λ_i: Ρυθμός αφίξεων όταν υπάρχουν i πελάτες.
- μ_i: Ρυθμός αναχωρήσεων όταν υπάρχουν i πελάτες.

Απλές Μαρκοβιανές Ουρές: Κατανομή ισορροπίας αριθμού πελατών

- Το σύστημα είναι ευσταθές αν και μόνο αν

$$B^{-1} = 1 + \sum_{n=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n} < \infty.$$

- Όταν είναι ευσταθές, κατανομή ισορροπίας αριθμού πελατών (σε συνεχή χρόνο):

$$p_n = \begin{cases} B & \text{αν } n = 0, \\ B \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n} & \text{αν } n \geq 1. \end{cases}$$

Απλές Μαρκοβιανές Ουρές: Ρυθμοί αφίξεων/αναχωρήσεων, Εμφυτευμένες κατανομές

- Ρυθμοί αφίξεων, αναχωρήσεων:

$$\lambda^* = \sum_{n=0}^{\infty} \lambda_n p_n, \quad \mu^* = \sum_{n=1}^{\infty} \mu_n p_n.$$

- Κατανομή ισορροπίας σε στιγμές αφίξεων:

$$a_n = \frac{\lambda_n p_n}{\lambda^*}.$$

- Κατανομή ισορροπίας σε στιγμές αναχωρήσεων:

$$d_n = \frac{\mu_{n+1} p_{n+1}}{\lambda^*}.$$

Απλές Μαρκοβιανές Ουρές: Χρόνος παραμονής πελάτη

- S: Χρόνος παραμονής πελάτη υπό την πειθαρχία FCFS.
- Q^-: Πλήθος πελατών που βρίσκει ο πελάτης κατά την άφιξή του.
- Τότε:

$$F_S(x) = \Pr[S \leq x] = \sum_{n=0}^{\infty} a_n \Pr[S \leq x | Q^- = n]$$

- Για τη μέση τιμή, από Ν. Little:

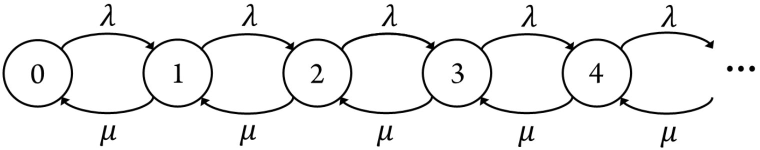
$$E[S] = \frac{E[Q]}{\lambda}.$$

Απλές Μαρκοβιανές Ουρές: Κύκλος απασχόλησης (λειτουργίας)

- I: Περίοδος αργίας.
- Y: Περίοδος συνεχούς λειτουργίας.
- $Z = Y + I$: Κύκλος απασχόλησης (λειτουργίας).
- $E[I] = \frac{1}{\lambda_0}$.
- $p_0 = \frac{E[I]}{E[Z]} \Rightarrow E[Z] = \frac{1}{\lambda_0 p_0}$.
- $E[Y] = E[Z] - E[I] = \frac{1-p_0}{\lambda_0 p_0}$.

Η M/M/1 ουρά

- Poisson διαδικασία αφίξεων με ρυθμό λ.
- Exp(μ) χρόνοι εξυπηρέτησης.
- 1 υπηρέτης.
- ∞ χωρητικότητα.
- FCFS πειθαρχία ουράς.
- Η $\{Q(t)\}$ είναι Μ.α.σ.χ. με διάγραμμα ρυθμών:



Η M/M/1 ουρά (συνέχεια)

- Ευστάθεια:
  Ευσταθής $\Leftrightarrow \rho < 1$ (από ευστάθεια G/G/1 ουράς).
- Υπολογισμός κατανομής ισορροπίας:

$$
\begin{aligned}
B^{-1} &= 1 + \sum_{n=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n} = \sum_{n=0}^{\infty} \rho^n \\
&= \begin{cases} \frac{1}{1-\rho} & \text{αν } \rho < 1 \text{ (ευστάθεια)} \\ \infty & \text{αν } \rho \geq 1 \text{ (αστάθεια).} \end{cases} \\
p_n &= \begin{cases} B & \text{αν } n = 0 \\ B \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n} & \text{αν } n \geq 1 \end{cases} \\
&= (1-\rho)\rho^n, \ n \geq 0.
\end{aligned}
$$

- Άρα $(p_n) \sim \mathrm{Geom}(\rho)$ στο $\mathbb{N}_0$.

Η Μ/Μ/1 ουρά (συνέχεια)

- PASTA+Ιδιότητα μεμονωμένων αφίξεων/αναχωρήσεων: $(d_n) = (a_n) = (p_n)$.
- Μέσο πλήθος πελατών:

$$E[Q^+] = E[Q^-] = E[Q] = \frac{\rho}{1-\rho}$$

(από Ν. Little ή $\sum_{n=0}^{\infty} np_n$).

- Μέσος χρόνος παραμονής/αναμονής πελάτη:

$$E[S] = \frac{1}{\mu(1-\rho)}, \quad E[W] = \frac{\rho}{\mu(1-\rho)}.$$

Η M/M/1 ουρά (συνέχεια)

- Κατανομή χρόνου παραμονής πελάτη:

$$\begin{aligned} F_S(x) &= \Pr[S \le x] = \sum_{n=0}^{\infty} \Pr[Q^- = n] \Pr[S \le x | Q^- = n] \\ &= \sum_{n=0}^{\infty} (1-\rho)\rho^n \int_0^x \frac{\mu^{n+1}}{n!} u^n e^{-\mu u} du \\ &= \int_0^x \mu(1-\rho) e^{-\mu u} \sum_{n=0}^{\infty} \frac{(\mu\rho u)^n}{n!} du \\ &= \int_0^x \mu(1-\rho) e^{-\mu u} e^{\mu\rho u} du = \int_0^x \mu(1-\rho) e^{-\mu(1-\rho)u} du \end{aligned}$$

- Άρα $S \sim \mathrm{Exp}(\mu(1-\rho)) = \mathrm{Exp}(\mu - \lambda)$.

Η M/M/1 ουρά (συνέχεια)

- I: Περίοδος αργίας.
- Y: Περίοδος συνεχούς λειτουργίας.
- $Z = Y + I$: Κύκλος απασχόλησης (λειτουργίας).
- $E[I] = \frac{1}{\lambda}$.
- $p_0 = \frac{E[I]}{E[Z]} \Rightarrow E[Z] = \frac{1}{\lambda(1-\rho)}$.
- $E[Y] = E[Z] - E[I] = \frac{\rho}{\lambda(1-\rho)} = \frac{1}{\mu(1-\rho)} = \frac{1}{\mu-\lambda}$.
- Παράδοξο: $E[Y] = E[S] = \frac{1}{\mu-\lambda}$.
  Μια πρώτη σκέψη θα έλεγε ότι το $E[Y]$ θα έπρεπε να είναι μεγαλύτερο από το $E[S]$, αφού κάθε περίοδος συνεχούς λειτουργίας είναι μεγαλύτερη από τον χρόνο παραμονής κάθε πελάτη μέσα σε αυτή.

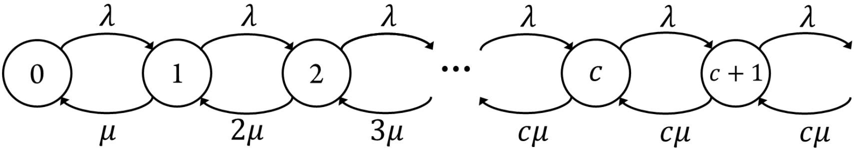
Η M/M/1 ουρά (συνέχεια)

- Σε μια ευσταθή M/M/1 ουρά ποιο είναι το ποσοστό των κύκλων απασχόλησης στους οποίους εξυπηρετείται μόνο ένας πελάτης;
- Είναι

$$\Pr[\mathrm{Exp}(\mu) < \mathrm{Exp}(\lambda)] = \frac{\mu}{\lambda + \mu} > \frac{\mu}{\mu + \mu} = \frac{1}{2}.$$

Η M/M/c ουρά

- Poisson διαδικασία αφίξεων με ρυθμό λ.
- Exp(μ) χρόνοι εξυπηρέτησης.
- c υπηρέτες.
- ∞ χωρητικότητα.
- FCFS πειθαρχία ουράς.
- Η $\{Q(t)\}$ είναι Μ.α.σ.χ. με διάγραμμα ρυθμών:



Η M/M/c ουρά (συνέχεια)

- Ευστάθεια:
  Ευσταθής $\Leftrightarrow \rho < c$ (από ευστάθεια G/G/c ουράς).
- Υπολογισμός κατανομής ισορροπίας:

$$
\begin{aligned}
B^{-1} &= 1 + \sum_{n=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n} = \sum_{n=0}^{c} \frac{\rho^n}{n!} + \sum_{n=c+1}^{\infty} \frac{\rho^n}{c! c^{n-c}} \\
&= \begin{cases} \sum_{n=0}^{c} \frac{\rho^n}{n!} + \frac{\rho^c}{c!} \frac{\rho}{\rho - c} & \text{αν } \rho < c \text{ (ευστάθεια)}, \\ \infty & \text{αν } \rho \geq c \text{ (αστάθεια)}. \end{cases} \\
p_n &= \begin{cases} B \frac{\rho^n}{n!} & \text{αν } 0 \leq n \leq c, \\ B \frac{\rho^n}{c! c^{n-c}} & \text{αν } n \geq c+1. \end{cases}
\end{aligned}
$$

Η M/M/c ουρά (συνέχεια)

- Πιθανότητα καθυστέρησης πελάτη:

$$
\begin{aligned}
C(c,\rho) &= \sum_{n=c}^{\infty} a_n = \sum_{n=c}^{\infty} p_n = \Pr[Q \geq c] = B\frac{\rho^c}{c!}\frac{c}{c-\rho} \\
&= \frac{\frac{\rho^c}{c!}\frac{c}{c-\rho}}{\sum_{j=0}^{c-1}\frac{\rho^j}{j!} + \frac{\rho^c}{c!}\frac{c}{c-\rho}}
\end{aligned}
$$

τύπος καθυστερήσεων του Erlang ή τύπος Erlang C (Erlang C formula).

Η M/M/c ουρά (συνέχεια)

- Κατανομή ισορροπίας πελατών στον χώρο αναμονής, δεδομένου ότι όλοι οι πελάτες είναι απασχολημένοι:

$$\begin{aligned}\Pr[Q_q = n | Q \geq c] &= \Pr[Q = n + c | Q \geq c] \\ &= \frac{B \frac{\rho^{c+n}}{c! c^n}}{\sum_{m=c}^{\infty} B \frac{\rho^m}{c! c^{m-c}}} \\ &= \left(1 - \frac{\rho}{c}\right)\left(\frac{\rho}{c}\right)^n, \ n \geq 0,\end{aligned}$$

- Άρα $(Q_q | Q \geq c) \sim \mathrm{Geom}(\frac{\rho}{c})$ στο $\mathbb{N}_0$.
- Επομένως:

$$E[Q_q | Q \geq c] = \frac{\rho/c}{1 - \rho/c}.$$

Η M/M/c ουρά (συνέχεια)

- Μέσος αριθμός πελατών στην ουρά:

$$\begin{aligned} E[Q_q] &= \Pr[Q < c]E[Q_q|Q < c] + \Pr[Q \geq c]E[Q_q|Q \geq c] \\ &= \Pr[Q < c] \cdot 0 + \Pr[Q \geq c] \cdot \frac{\rho/c}{1-\rho/c} \\ &= B\frac{\rho^{c+1}}{(c-1)!(c-\rho)^2}. \end{aligned}$$

- Μέσος αριθμός πελατών στο σύστημα:

$$E[Q] = B\frac{\rho^{c+1}}{(c-1)!(c-\rho)^2} + \rho.$$

Η M/M/c ουρά (συνέχεια)

- Μέσο χρόνο αναμονής πελάτη:

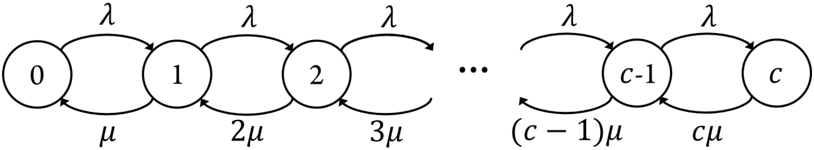
$$E[W] = \frac{E[Q_q]}{\lambda} = B\frac{\rho^c}{(c-1)!(c-\rho)^2}\frac{1}{\mu}.$$

- Μέσος χρόνος παραμονής πελάτη:

$$E[S] = E[W] + \frac{1}{\mu} = \left(B\frac{\rho^c}{(c-1)!(c-\rho)^2} + 1\right)\frac{1}{\mu}.$$

Η $M/M/c/c$ ουρά

- Poisson διαδικασία αφίξεων με ρυθμό λ.
- $\mathrm{Exp}(\mu)$ χρόνοι εξυπηρέτησης.
- c υπηρέτες.
- c χωρητικότητα (καθόλου χώρος αναμονής).
- FCFS πειθαρχία ουράς.
- Η $\{Q(t)\}$ είναι Μ.α.σ.χ. με διάγραμμα ρυθμών:



Η M/M/c/c ουρά (συνέχεια)

- Ευστάθεια:Πάντα ευσταθής λόγω πεπερασμένου χώρου καταστάσεων.
- Υπολογισμός κατανομής ισορροπίας:

$$
\begin{aligned}
B^{-1} &= 1+\sum_{n=1}^{c}\frac{\lambda_0\lambda_1\cdots\lambda_{n-1}}{\mu_1\mu_2\cdots\mu_n}=\sum_{n=0}^{c}\frac{\rho^n}{n!},\\
p_n &= \frac{\frac{\rho^n}{n!}}{\sum_{m=0}^{c}\frac{\rho^m}{m!}},\quad 0\le n\le c.
\end{aligned}
$$

Η M/M/c/c ουρά (συνέχεια)

- Πιθανότητα απώλειας πελάτη:

$$B(c,\rho) = a_c = p_c = \frac{\frac{\rho^c}{c!}}{\sum_{m=0}^{c} \frac{\rho^m}{m!}}$$

  τύπος απωλειών του Erlang ή Erlang B (Erlang B formula).
- Ο τύπος αυτός ισχύει και για το M/G/c/c σύστημα (ιδιότητα μη-ευαισθησίας).

Η M/M/c ουρά με αποθαρρυνόμενους πελάτες

- Poisson διαδικασία αφίξεων με ρυθμό λ.
- Exp(μ) χρόνοι εξυπηρέτησης.
- c υπηρέτες.
- ∞ χωρητικότητα.
- FCFS πειθαρχία ουράς.
- Κάθε αφικνούμενος πελάτης που βρίσκει n πελάτες στο σύστημα αναχωρεί άμεσα με πιθανότητα q_n.
- Συνήθως: $q_n \uparrow n$ και $\lim_{n\to\infty} q_n = 1$.
- $\{Q(t)\}$ αλυσίδα γέννησης-θανάτου με ρυθμούς

$$\begin{aligned} \lambda_n &= \lambda(1-q_n), \quad n \geq 0, \\ \mu_n &= \min(n,c)\mu, \quad n \geq 1. \end{aligned}$$

Η $M/M/c$ ουρά με ανυπόμονους πελάτες

- Poisson διαδικασία αφίξεων με ρυθμό λ.
- $\text{Exp}(\mu)$ χρόνοι εξυπηρέτησης.
- c υπηρέτες.
- ∞ χωρητικότητα.
- FCFS πειθαρχία ουράς.
- Κάθε αφικνούμενος πελάτης εισέρχεται σε αυτό αλλά έχει έναν χρόνο υπομονής $\text{Exp}(\theta)$ και αποχωρεί από το σύστημα αν ο χρόνος εκπνεύσει πριν να αρχίσει την εξυπηρέτησή του.
- $\{Q(t)\}$ αλυσίδα γέννησης-θανάτου με ρυθμούς

$$\begin{aligned}\lambda_n &= \lambda, \quad n \geq 0,\\ \mu_n &= \min(n,c)\mu + \max(n-c,0)\theta, \quad n \geq 1.\end{aligned}$$

Σύγκριση M/M/c και αντίστοιχων M/M/c συστημάτων

- Κοινή ουρά για όλους τους υπηρέτες ή μια ουρά για κάθε υπηρέτη;
- Πολλοί αργοί ή λίγοι γρήγοροι υπηρέτες;

Κοινή ουρά για όλους τους υπηρέτες ή μια ουρά για κάθε υπηρέτη;

- Πρόκειται να σχεδιάσουμε σύστημα εξυπηρέτησης με c όμοιους υπηρέτες.
- Είναι προτιμότερο να έχουμε ξεχωριστή ουρά για κάθε υπηρέτη ή μια κοινή ουρά για όλους (pooling);
- Συγκεκριμένο πλαίσιο:
  Poisson διαδικασία αφίξεων.
  2 διαθέσιμοι υπηρέτες με $\mathrm{Exp}(\mu)$ χρόνους εξυπηρέτησης.
  Απεριόριστη χωρητικότητα.
- Σύστημα 1: Δυο παράλληλες $M/M/1$ ουρές με ρυθμό αφίξεων $\frac{\lambda}{2}$ και ρυθμό μ εξυπηρέτησης έκαστη.
- Σύστημα 2: Μια $M/M/2$ ουρά με ρυθμό αφίξεων λ και ρυθμό εξυπηρέτησης μ.

Κοινή ουρά για όλους τους υπηρέτες ή μια ουρά για κάθε υπηρέτη; (συνέχεια)

- Θέτουμε $\rho = \frac{\lambda}{\mu}$.
- Σύστημα 1: Δυο παράλληλες M/M/1 ουρές με ρυθμό αφίξεων $\frac{\lambda}{2}$ και ρυθμό μ έκαστη. Μέσος χρόνος παραμονής:

$$E[S_1] = \frac{1}{\mu - \frac{\lambda}{2}} = \frac{2}{\mu(2-\rho)}.$$

- Σύστημα 2: Μια M/M/2 ουρά με ρυθμό αφίξεων λ και ρυθμό μ. Μέσος χρόνος παραμονής:

$$E[S_2] = \frac{4}{\mu(2+\rho)(2-\rho)}.$$

Κοινή ουρά για όλους τους υπηρέτες ή μια ουρά για κάθε υπηρέτη; (συνέχεια)

- Έχουμε

$$\begin{aligned} E[S_1] - E[S_2] &= \frac{2}{\mu(2-\rho)} - \frac{4}{\mu(2+\rho)(2-\rho)} \\ &= \frac{2\rho}{\mu(2+\rho)(2-\rho)} > 0. \end{aligned}$$

- Συμπέρασμα: Προτιμότερο να υπάρχει μια κοινή ουρά για όλους τους υπηρέτες, αν ενδιαφερόμαστε για τον χρόνο παραμονής ενός πελάτη στο σύστημα.
- Το ίδιο ισχύει και για τον χρόνο αναμονής στην ουρά, αφού έχουμε $E[W_i] = E[S_i] - \frac{1}{\mu}$, $i = 1, 2$, οπότε

$$E[W_1] - E[W_2] = E[S_1] - E[S_2] > 0.$$

Κοινή ουρά για όλους τους υπηρέτες ή μια ουρά για κάθε υπηρέτη; (συνέχεια)

- Έχουμε:

$$\lim_{\rho \to 2^-} (E[S_1] - E[S_2]) = \lim_{\rho \to 2^-} \frac{2\rho}{\mu(2+\rho)(2-\rho)} = \infty,$$

δηλαδή η κοινή ουρά μειώνει σημαντικά τον μέσο χρόνο παραμονής σε συστήματα με ρυθμό συνωστισμού κοντά στην κρίσιμη τιμή για αστάθεια.

- Επίσης:

$$\frac{E[S_2]}{E[S_1]} = \frac{2}{2+\rho} \in \left(\frac{1}{2}, 1\right), \quad \rho \in (0, 2),$$

δηλαδή η κοινή ουρά μειώνει στο μισό τον χρόνο παραμονής για ρυθμό συνωστισμού κοντά στην κρίσιμη τιμή για αστάθεια.

Κοινή ουρά για όλους τους υπηρέτες ή μια ουρά για κάθε υπηρέτη; (συνέχεια)

- Η χρήση κοινής ουράς (pooling) βελτιώνει την απόδοση του συστήματος σε πολύ γενικότερες καταστάσεις.
- Η χρήση κοινής ουράς έχει υιοθετηθεί σε πολλές πρακτικές εφαρμογές των ουρών σε πραγματικά συστήματα τα τελευταία χρόνια.
- Αν οι πελάτες είναι ετερογενείς μπορεί να είναι προτιμότερο να έχουμε διαφορετικές ουρές για διαφορετικές κατηγορίες πελατών.
- Αν οι υπηρέτες δεν είναι όμοιοι, το ερώτημα θα πρέπει να εξεταστεί εξαρχής με προσοχή.

Πολλοί αργοί ή λίγοι γρήγοροι υπηρέτες;

- Πρόκειται να σχεδιάσουμε σύστημα εξυπηρέτησης με συγκεκριμένη δυναμικότητα (συνολικό ρυθμό) εξυπηρέτησης.
- Είναι προτιμότερο να έχουμε έναν υπηρέτη με όλη τη δυναμικότητα ή να τη μοιράσουμε σε περισσότερους;
- Συγκεκριμένο πλαίσιο:
  Poisson διαδικασία αφίξεων.
  Δυνατότητα για έναν υπηρέτη με $\text{Exp}(2\mu)$ χρόνους εξυπηρέτησης ή δυο υπηρετών με $\text{Exp}(\mu)$ χρόνους εξυπηρέτησης.
- Σύστημα 1: $M/M/1$ ουρά με ρυθμό αφίξεων λ και ρυθμό εξυπηρέτησης 2μ.
- Σύστημα 2: $M/M/2$ ουρά με ρυθμό αφίξεων λ και ρυθμό εξυπηρέτησης μ ανά υπηρέτη.

Πολλοί αργοί ή λίγοι γρήγοροι υπηρέτες; (συνέχεια)

- Θέτουμε $\rho = \frac{\lambda}{\mu}$.
- Σύστημα 1: M/M/1 ουρά με ρυθμό αφίξεων λ και ρυθμό εξυπηρέτησης 2μ. Μέσος χρόνος παραμονής:

$$E[S_1] = \frac{1}{2\mu - \lambda} = \frac{1}{\mu(2-\rho)}.$$

- Σύστημα 2: M/M/2 ουρά με ρυθμό αφίξεων λ και ρυθμό εξυπηρέτησης μ ανά υπηρέτη. Μέσος χρόνος παραμονής:

$$E[S_2] = \frac{4}{\mu(2+\rho)(2-\rho)}.$$

Πολλοί αργοί ή λίγοι γρήγοροι υπηρέτες; (συνέχεια)

- Έχουμε

$$\begin{aligned} E[S_1] - E[S_2] &= \frac{1}{\mu(2-\rho)} - \frac{4}{\mu(2+\rho)(2-\rho)} \\ &= \frac{\rho - 2}{\mu(2+\rho)(2-\rho)} = -\frac{1}{\mu(2+\rho)} < 0. \end{aligned}$$

- Συμπέρασμα: Προτιμότερο να ανατεϑεί όλη η δυναμικότητα σε έναν υπηρέτη, αν ενδιαφερόμαστε για τον χρόνο παραμονής ενός πελάτη στο σύστημα.

Πολλοί αργοί ή λίγοι γρήγοροι υπηρέτες; (συνέχεια)

- Εδώ:

$$\frac{E[S_1]}{E[S_2]} = \frac{\rho+2}{4} \in \left(\frac{1}{2}, 1\right), \quad \rho \in (0,2),$$

  δηλαδή η ανάθεση όλης της δυναμικότητας εξυπηρέτησης σε έναν υπηρέτη βελτιώνει σημαντικά την λειτουργία του συστηματος όσο μικρότερος είναι ο συντελεστής συνωστισμού.
- Συγκεκριμένα οδηγεί σε μείωση του μέσου χρόνου παραμονής των πελατών με ένα συντελεστή που φθάνει κοντά στο $\frac{1}{2}$ για μικρές τιμές του ρυθμού συνωστισμού.

Πολλοί αργοί ή λίγοι γρήγοροι υπηρέτες; (συνέχεια)

- Για τους χρόνους αναμονής έχουμε:
  Για την M/M/1 ουρά:

$$E[W_1] = E[S_1] - \frac{1}{2\mu} = \frac{\rho}{2\mu(2-\rho)},$$

ενώ για την M/M/2 ουρά:

$$E[W_2] = E[S_2] - \frac{1}{\mu} = \frac{4}{\mu(2+\rho)(2-\rho)} - \frac{1}{\mu} = \frac{\rho^2}{\mu(2+\rho)(2-\rho)}.$$

Πολλοί αργοί ή λίγοι γρήγοροι υπηρέτες; (συνέχεια)

- Επομένως:

$$\begin{aligned} E[W_1] - E[W_2] &= \frac{\rho}{2\mu(2-\rho)} - \frac{\rho^2}{\mu(2+\rho)(2-\rho)} \\ &= \frac{\rho}{2\mu(2+\rho)} > 0. \end{aligned}$$

- Προτιμότερο να μοιραστεί η δυναμικότητα εξυπηρέτησης σε δυο υπηρέτες, αν ενδιαφερόμαστε για τον χρόνο αναμονής ενός πελάτη στην ουρά.
- Ισχύει το αντίθετο από αυτό που είδαμε για τον χρόνο παραμονής.