

Στοχαστικά Μοντέλα
στην Επιχειρησιακή Έρευνα
Μέρος 3
Θεωρία Ουρών Αναμονής

Αντώνης Οικονόμου

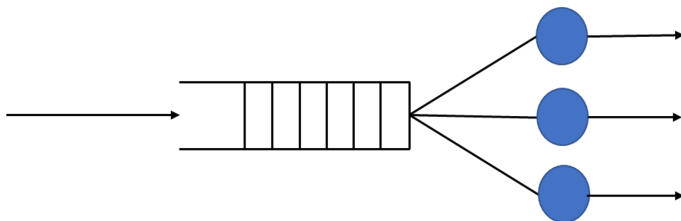
Μεταπτυχιακό Πρόγραμμα Σπουδών
Στατιστική και Επιχειρησιακή Έρευνα

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: $c = 1, 2, \dots$
- Χωρητικότητα: $k = c, c + 1, c + 2, \dots$
- 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 Ανελίζεις).
- Σύγκριση συστημάτων. (Στοχαστικές Ανελίζεις + Στοχαστικές Διατάξεις).
- Βέλτιστος σχεδιασμός. (Στοχαστικές Ανελίζεις + 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

- $A_1 \leq A_2 \leq A_3 \leq \dots$: Στιγμές διαδοχικών αφίξεων.
- $D_1 \leq D_2 \leq D_3 \leq \dots$: Στιγμές διαδοχικών αναχωρήσεων.
- $Q_n^- = Q(A_n^-)$: Πλήθος πελατών που βρίσκεται 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:

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

Βασικά μέτρα απόδοσης ουρών (συνέχεια)

- Για την εμφυτευμένη διαδικασία σε στιγμές αφίξεων έχουμε ανάλογα:

$$\begin{aligned}
 \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n 1_{\{Q_k^- = j\}} &= \lim_{n \rightarrow \infty} \frac{1}{n} E \left[\sum_{k=1}^n 1_{\{Q_k^- = j\}} \right] \quad \text{με πιθαν. 1} \\
 &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \Pr[Q_k^- = j] = \lim_{n \rightarrow \infty} \Pr[Q_n^- = j] \\
 &= \frac{1}{E[A(Z)]} E \left[\sum_{k=1}^{A(Z)} 1_{\{Q_k^- = j\}} \right] \stackrel{\text{ορσ}}{=} a_j = \Pr[Q^- = j],
 \end{aligned}$$

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:

$$\begin{aligned}
 \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n 1_{\{S_k^- \leq x\}} &= \lim_{n \rightarrow \infty} \frac{1}{n} E \left[\sum_{k=1}^n 1_{\{S_k^- \leq x\}} \right] \text{ με πιθαν. } 1 \\
 &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \Pr[S_k^- \leq x] = \lim_{n \rightarrow \infty} \Pr[S_n^- \leq x] \\
 &= \frac{1}{E[A(Z)]} E \left[\sum_{k=1}^{A(Z)} 1_{\{S_k^- \leq x\}} \right] \stackrel{\text{ορσ}}{=} F_S(x) = \Pr[S \leq x],
 \end{aligned}$$

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:

$$\begin{aligned}
 E[Q] &\stackrel{\text{ορσ}}{=} \lim_{t \rightarrow \infty} \frac{\int_0^t Q(u) du}{t} = \lim_{t \rightarrow \infty} \frac{E \left[\int_0^t Q(u) du \right]}{t} \quad \text{με πιθαν. 1} \\
 &= \lim_{t \rightarrow \infty} E[Q(t)] = \frac{E \left[\int_0^Z Q(u) du \right]}{E[Z]}.
 \end{aligned}$$

- Μέσος χρόνος παραμονής πελάτη στο σύστημα:

$$\begin{aligned}
 E[S] &\stackrel{\text{ορσ}}{=} \lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n S_k}{n} = \lim_{n \rightarrow \infty} \frac{E \left[\sum_{k=1}^n S_k \right]}{n} \quad \text{με πιθαν. 1} \\
 &= \lim_{n \rightarrow \infty} E[S_n] = \frac{E \left[\sum_{k=1}^{A(Z)} S_k \right]}{E[A(Z)]}.
 \end{aligned}$$

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 αperiοδική κατανομή χρόνων εξυπηρέτησης ή ενδιάμεσων χρόνων αφίξεων (ή και τα δύο).
- Ρυθμός συνωστισμού = Ρυθμός αφίξεων $\times$ 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 $\lim_{t \rightarrow \infty} Q(t) = \infty$ 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

$$a_j = \lim_{t \rightarrow \infty} \frac{A_j(t)}{A(t)},$$

$$d_j = \lim_{t \rightarrow \infty} \frac{D_j(t)}{D(t)}.$$

Ιδιότητα μεμονωμένων μεταβάσεων (αιτιολόγηση)

- Από ισότητα ρυθμών αφίξεων και αναχωρήσεων:

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

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

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

- Επομένως:

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

Ιδιότητα PASTA

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

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

Νόμος του Little

- Μέσος αριθμός πελατών = Ρυθμός αφίξεων $\times$ 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:

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

Άμεσες συνέπειες του Νόμου Little

- Εφαρμογή του N. Little στον χώρο αναμονής: Μέσος αριθμός πελατών σε αναμονή = Ρυθμός αφίξεων $\times$ Μέσος χρόνος αναμονής:

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

- Εφαρμογή του N. Little στον χώρο εξυπηρέτησης: Μέσος αριθμός πελατών σε εξυπηρέτηση = Ρυθμός αφίξεων $\times$ 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 $G/G/1$.