

Ανάλυση της Απόδοσης Πληροφοριακών Συστημάτων

Εισαγωγή στο σύστημα αναμονής M/G/1

Γιάννης Γαροφαλάκης

Το σύστημα αναμονής M/G/1

Θεωρούμε ένα σύστημα στο οποίο οι πελάτες φθάνουν κατά Poisson με ρυθμό λ , αλλά οι χρόνοι εξυπηρέτησης είναι **γενικά κατανεμημένοι**. Συγκεκριμένα, θεωρούμε πως οι πελάτες εξυπηρετούνται με την σειρά που φθάνουν και πως X_i είναι ο χρόνος εξυπηρέτησης της i -οστής άφιξης. Υποθέτουμε πως οι τυχαίες μεταβλητές (X_1, X_2, \dots) είναι όμοια κατανεμημένες, ανεξάρτητες μεταξύ τους, και ανεξάρτητες των χρόνων μεταξύ διαδοχικών αφίξεων. Έστω ότι ο μέσος χρόνος εξυπηρέτησης (δηλαδή η μέση τιμή των τυχαίων μεταβλητών X_i) είναι:

$$\mathbb{E}[X] = \frac{1}{\mu}$$

Η Pollaczek-Khintchine (P-K) φόρμουλα δίνει το μέσο χρόνο αναμονής στην ουρά συναρτήσει του ρυθμού αφίξεων, του ρυθμού εξυπηρέτησης και της ποσότητας: $\mathbb{E}[X^2]$ (δεύτερη ροπή του χρόνου εξυπηρέτησης):

$$\mathbb{E}[W] = \frac{\lambda \mathbb{E}[X^2]}{2(1 - \rho)} \quad (1)$$

όπου $\mathbb{E}[W]$ είναι ο μέσος χρόνος που περνά ένας πελάτης περιμένοντας στην ουρά και $\rho = \lambda/\mu = \lambda \mathbb{E}[X]$.

Όταν γνωρίζει κανείς την (P-K), προφανώς μπορεί πολύ εύκολα να πάρει όλες τις βασικές ενδιαφέρουσες μετρικές με χρήση του νόμου του Little.

Από τη θεωρία πιθανοτήτων:

■ Δεύτερη ροπή:

- Συνεχής κατανομή με pdf $a(x)$: $E[X^2] = \int_0^{\infty} x^2 a(x) dt$
- Διακριτή κατανομή με pmf p_k : $E[X^2] = \sum_{k=0}^{\infty} k^2 p_k$

■ Ισχύει:

$$\sigma_x^2 = E[X^2] - (E[X])^2 = E[X^2] - E[X]^2 \Rightarrow$$

$$E[X^2] = \sigma_x^2 + E[X]^2$$

■ Εκθετική κατανομή:

- $E[X] = \frac{1}{\mu}$
- $\sigma_x^2 = \frac{1}{\mu^2}$
- $E[X^2] = \frac{2}{\mu^2}$

Το σύστημα αναμονής M/G/1 (Μετρικές)

Μέσος Συνολικός Χρόνος Στο Σύστημα:

$$\mathbb{E}[T] = \underbrace{\mathbb{E}[X]}_{\text{μέσος χρόνος εξυπηρέτησης}} + \underbrace{\frac{\lambda \mathbb{E}[X^2]}{2(1-\rho)}}_{\text{μέσος χρόνος στην ουρά}} \quad (2)$$

Μέσος Αριθμός Πελατών Στην Ουρά:

$$\mathbb{E}[N_Q] = \lambda \times \mathbb{E}[W] = \frac{\lambda^2 \mathbb{E}[X^2]}{2(1-\rho)} \quad (3)$$

Μέσος Αριθμός Πελατών Στο Σύστημα:

$$\mathbb{E}[N] = \underbrace{\rho}_{\text{Μέσος αριθμός πελατών στην εξυπηρέτηση}} + \underbrace{\frac{\lambda^2 \mathbb{E}[X^2]}{2(1-\rho)}}_{\mathbb{E}[N_Q]} \quad (4)$$

Απόδειξη της P-K φόρμουλας I

Θεωρούμε τον i -οστό πελάτη που φθάνει στο σύστημα και συμβολίζουμε με W_i τον χρόνο αναμονής του στην ουρά. Έστω N_i ο αριθμός των πελατών που βρίσκει ο i να περιμένουν στην ουρά τη στιγμή της άφιξής του. Όταν $N_i > 0$ προφανώς στο σύστημα εκτός από τους N_i πελάτες που περιμένουν στην ουρά, υπάρχει κι' άλλος ένας που βρίσκεται στον εξυπηρετητή. Έστω R_i ο υπολειπόμενος χρόνος εξυπηρέτησης του πελάτη αυτού. Οπότε ο χρόνος W_i θα ισούται με:

$$W_i = R_i + \sum_{j=i-N_i}^{i-1} X_j \quad (5)$$

(Το R_i περιλαμβάνει και την περίπτωση να έχει τιμή 0)

Αυτό που μας ενδιαφέρει είναι η εύρεση της μέσης τιμής του χρόνου αναμονής, οπότε παίρνουμε μέσες τιμές της παραπάνω σχέσης:

$$\begin{aligned} \mathbb{E}[W_i] &= \mathbb{E}[R_i] + \mathbb{E}\left[\sum_{j=i-N_i}^{i-1} X_j\right] \\ &= \mathbb{E}[R_i] + \sum_{j=i-\mathbb{E}[N_i]}^{i-1} \mathbb{E}[X_j|N_i] \end{aligned}$$

Απόδειξη της P-K φόρμουλας II

Κάνοντας χρήση της υπόθεσης ανεξαρτησίας των N_i και $X_{i-1}, \dots, X_{i-N_i}$, καθώς και του γεγονότος ότι οι τυχαίες μεταβλητές X_j είναι όμοια κατανεμημένες, έχουμε:

$$\begin{aligned}\mathbb{E}[W_i] &= \mathbb{E}[R_i] + \sum_{j=i-\mathbb{E}[N_i]}^{i-1} \mathbb{E}[X_j] \\ &= \mathbb{E}[R_i] + \mathbb{E}[X]\mathbb{E}[N_i]\end{aligned}$$

οπότε θεωρώντας πως το σύστημα είναι εργοδικό και λειτουργεί σε κατάσταση ισορροπίας έχουμε:

$$\mathbb{E}[W] = \mathbb{E}[R] + \frac{1}{\mu}\mathbb{E}[N_Q] \quad (6)$$

όπου ως $\mathbb{E}[R]$ ορίζεται ο μέσος υπολειπόμενος χρόνος εξυπηρέτησης που βιώνει ένας τυχαίος πελάτης που φθάνει στο σύστημα όταν αυτό βρίσκεται σε ισορροπία. Από τον νόμο του Little, έχουμε πως ισχύει:

$$\mathbb{E}[N_Q] = \lambda\mathbb{E}[W]$$

■ ΣΥΝΕΠΩΣ:

$$E[W] = E[R] + \frac{1}{\mu} \lambda E[W] = E[R] + \rho E[W] \Rightarrow$$

$$\Rightarrow E[W](1 - \rho) = E[R] \Rightarrow$$

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

- **M/M/1**: $E[R] = \rho \cdot \bar{x} = \rho \cdot \frac{1}{\mu}$ (η πιθανότητα να υπάρχει κάποιος στο server = ρ)

Οπότε: $E[W] = \frac{\rho}{1 - \rho} \cdot \frac{1}{\mu}$ (γνωστό...)

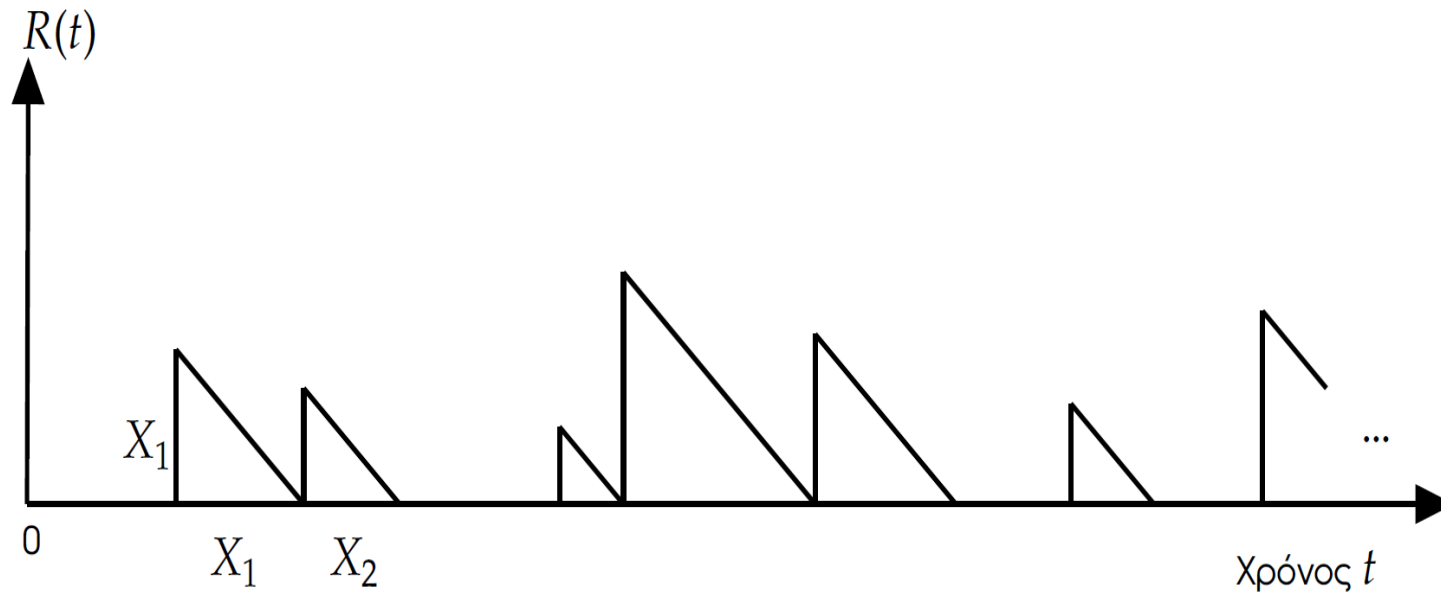
- **M/D/1**: $E[R] = \rho \cdot \frac{1}{2\mu}$ (διότι το R είναι ομοιόμορφα κατανομημένο μεταξύ 0 και $1/\mu$)

Οπότε: $E[W] = \frac{\rho}{1 - \rho} \cdot \frac{1}{2\mu}$ (το μισό από αυτό του M/M/1...)

Απόδειξη της P-K φόρμουλας III

$$\mathbb{E}[W] = \frac{\mathbb{E}[R]}{1 - \rho} \quad (7)$$

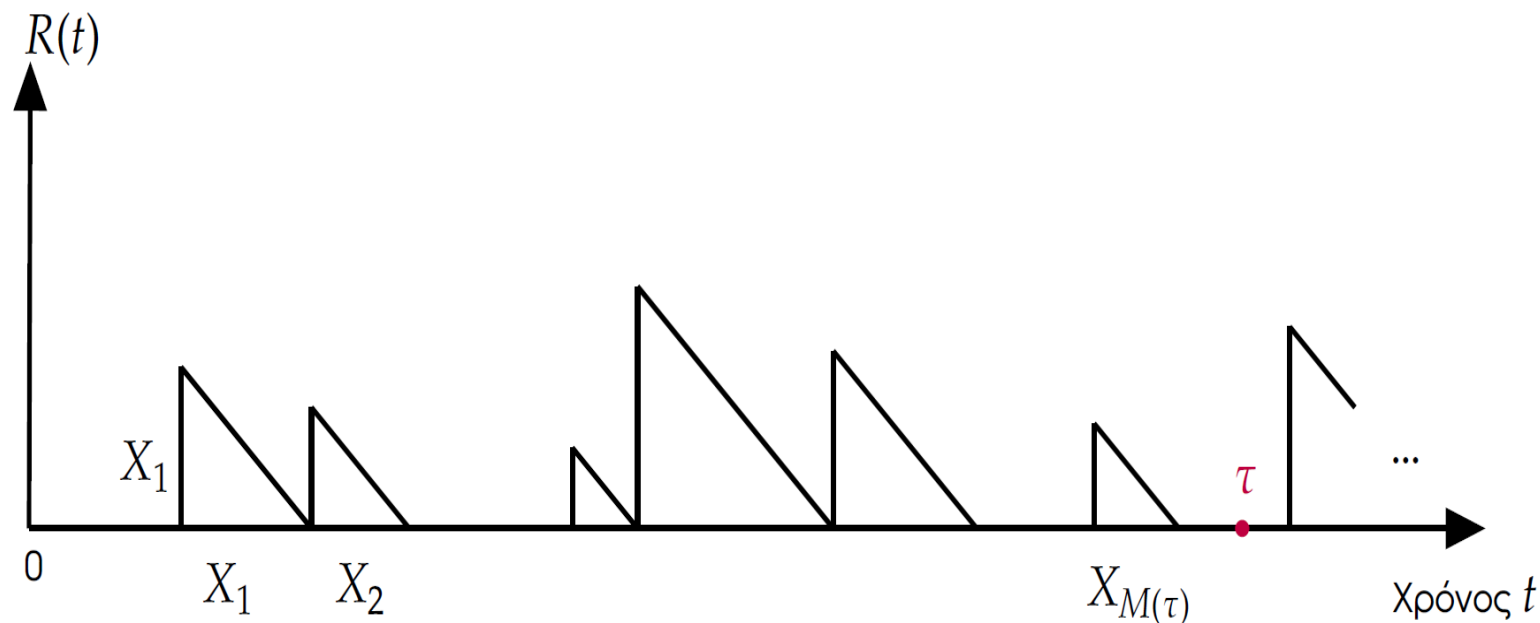
Μένει λοιπόν, να υπολογίσουμε το $\mathbb{E}[R]$. Σχεδιάζουμε τον υπολειπόμενο χρόνο εξυπηρέτησης, συναρτήσεως του t :



Απόδειξη της P-K φόρμουλας IV

Όταν τη στιγμή t_i , ένας νέος πελάτης με χρόνο εξυπηρέτησης X_i εισέλθει στον εξυπηρετητή, η $R(t)$ ξεκινά από το σημείο (t_i, X_i) και μειώνεται γραμμικά για X_i μονάδες χρόνου (δηλαδή, μέχρι το σημείο $(t_i + X_i, 0)$).

Θεωρούμε μία χρονική στιγμή τ για την οποία $R(\tau) = 0$. (Γιατί υπάρχει?)



Απόδειξη της P-K φόρμουλας V

Έστω ότι μέχρι τη στιγμή τ έχουν εξυπηρετηθεί $M(\tau)$ πελάτες. Ο μέσος χρόνος του $R(t)$ στο διάστημα $[0, \tau]$ είναι:

$$\frac{1}{\tau} \int_0^{\tau} R(t) dt = \frac{1}{\tau} \sum_{i=1}^{M(\tau)} \frac{1}{2} X_i^2$$

Οπότε πολλαπλασιάζοντας και διαιρώντας με το $M(\tau)$ παίρνουμε

$$\begin{aligned} \frac{1}{\tau} \int_0^{\tau} R(t) dt &= \frac{1}{2} \frac{M(\tau)}{\tau} \frac{\sum_{i=1}^{M(\tau)} X_i^2}{M(\tau)} \\ \Rightarrow \underbrace{\lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^{\tau} R(t) dt}_{\text{Μέσος υπολειπόμενος χρόνος } R} &= \underbrace{\frac{1}{2} \lim_{\tau \rightarrow \infty} \frac{M(\tau)}{\tau}}_{\text{ο ρυθμός αναχωρήσεων}} \cdot \underbrace{\lim_{\tau \rightarrow \infty} \frac{\sum_{i=1}^{M(\tau)} X_i^2}{M(\tau)}}_{\mathbb{E}[X^2]} \end{aligned} \quad (8)$$

$\mathbb{E}[R] = \lambda$

Δειγματοληπτική ροπή k τάξης:

$$\mathbb{E}[X^k] = \frac{1}{n} \sum_{i=1}^n x_i^k$$

Απόδειξη της P-K φόρμουλας VI

Αλλά στην ισορροπία ο ρυθμός αναχωρήσεων από το σύστημα ισούται με τον ρυθμό αφίξεων στο σύστημα, λ , συνεπώς τελικά παίρνουμε:

$$\mathbb{E}[R] = \frac{1}{2}\lambda\mathbb{E}[X^2]$$

η οποία όταν αντικατασταθεί στην (7) μας δίνει το ζητούμενο:

$$\mathbb{E}[W] = \frac{\mathbb{E}[R]}{1 - \rho} \Rightarrow$$

$$\mathbb{E}[W] = \frac{\lambda\mathbb{E}[X^2]}{2(1 - \rho)}$$

Απόδειξη της P-K φόρμουλας VII

Παρατηρήσεις

- ▶ Η απόδειξη της P-K, παραπάνω, έκανε την παραδοχή πως οι πελάτες εξυπηρετούνται με την σειρά που φθάνουν, δηλαδή το πλήθος των πελατών που εξυπηρετούνται μεταξύ της χρονικής στιγμής άφιξης του i -οστού πελάτη και της εξυπηρέτησής του είναι ο αριθμός των πελατών στην ουρά κατά τη στιγμή άφιξης του i . Προκύπτει ωστόσο, πως η P-K ισχύει για οποιαδήποτε σειρά εξυπηρέτησης των πελατών, αρκεί η σειρά εξυπηρέτησης να μην εξαρτάται από τον χρόνο εξυπηρέτησης των πελατών στην ουρά.
- ▶ Για την αυστηρή απόδειξη της (P-K) στη βιβλιογραφία συνήθως χρησιμοποιούνται οι ιδιότητες των renewal στοχαστικών διαδικασιών, ή εναλλακτικά οι embedded αλυσίδες Markov διακριτού χρόνου με ανάλυση στο πεδίο του μετασχηματισμού Z .

Άσκηση

Θεωρήστε το $M/G/1$ σύστημα στο οποίο οι πελάτες περιμένουν για να παραλάβουν δέματα στο ταχυδρομείο. Οι πελάτες φθάνουν σύμφωνα με τη διαδικασία Poisson με ρυθμό 1 πελάτης κάθε 2 λεπτά. Ο κάθε πελάτης έχει να παραλάβει μέχρι k δέματα ($1 \leq k \leq 4$), και ισχύει

$$\Pr[k = 1] = 0.5, \quad \Pr[k = 2] = 0.25 \quad \Pr[k = 3] = 0.2 \quad \Pr[k = 4] = 0.05$$

Αν ο υπάλληλος στο ταχυδρομείο χρειάζεται ακριβώς 1 λεπτό για να βρεί το κάθε πακέτο, να βρεθεί με χρήση της P-K formula ο μέσος χρόνος αναμονής στην ουρά του ταχυδρομείου.

Απαντήσεις (1)

- Ζητείται το $E[W] = \frac{\lambda E[X^2]}{2(1-\rho)}$ (1)
- $\lambda = 0,5$ πελάτες/λεπτό.
- Χρειαζόμαστε ακόμα το $\rho = \lambda/\mu = \lambda E[X]$ και το $E[X^2]$, όπου X η Τ.Μ. του χρόνου εξυπηρέτησης.
- Η X είναι μια διακριτή Τ.Μ. με τιμές/pmf:
 - $X = 1$ λεπτό, με πιθανότητα $p_1 = 0,5$
 - $X = 2$ λεπτά, με πιθανότητα $p_2 = 0,25$
 - $X = 3$ λεπτά, με πιθανότητα $p_3 = 0,2$
 - $X = 4$ λεπτά, με πιθανότητα $p_4 = 0,05$
- $E[X] = \sum_{k=1}^{k=4} k \cdot p_k = 1 \cdot 0,5 + 2 \cdot 0,25 + 3 \cdot 0,2 + 4 \cdot 0,05 \Rightarrow$
 $\Rightarrow E[X] = 1,8$ λεπτά/πελάτη

Απαντήσεις (2)

- $\rho = \lambda \cdot E[X] = 0,5 \cdot 1,8 \Rightarrow \rho = 0,9$
- $E[X^2] = \sum_{k=1}^{k=4} k^2 p_k = 1^2 \cdot 0,5 + 2^2 \cdot 0,25 + 3^2 \cdot 0,2 + 4^2 \cdot 0,05 \Rightarrow$
 $\Rightarrow E[X^2] = 4,1 \text{ λεπτά}^2/\text{πελάτη}$
- Αντικαθιστώντας στην (1): $E[W] = \frac{\lambda E[X^2]}{2(1-\rho)} = \frac{0,5 \cdot 4,1}{2(1-0,9)} \Rightarrow$
 $\Rightarrow \boxed{E[W] = 10,25 \text{ λεπτά}}$

(Παρατήρηση: Αν ήταν M/M/1 με ίδια $\lambda, \rho, E[X] = 1/\mu$: $E[W] = \frac{\rho}{1-\rho} \cdot \frac{1}{\mu} = 16,2 \text{ λεπτά}$)

Επίσης:

- Πελάτες στην ουρά: $E[N_Q] = \lambda \cdot E[W] = 0,5 \cdot 10,25 = 5,125 \text{ πελάτες}$
- Πελάτες στο ταχ/μείο: $E[N] = \rho + E[N_Q] = 0,9 + 5,125 = 6,025 \text{ πελ.}$