3. Διαδικασίες Ανανέωσης και Ουρές

3.1. Διαδικασίες Ανανέωσης

Ορισμός 3.34. Μια *διαδικασία ανανέωσης* (*renewal process*) $N = \{N(t) : t \geq 0\}$ είναι μια στοχαστική διαδικασία συνεχούς χρόνου τέτοια ώστε

$$N(t) = \max\{n \colon T_n \leq t\},$$

όπου $T_0 = 0, T_n = X_1 + X_2 + \ldots + X_n$, για $n \geq 1$, και $\{X_i\}$ είναι μια ακολουθία ανεξάρτητων και όμοια κατανεμημένων τυχαίων μεταβλητών, που παίρνουν μη αρνητικές τιμές.

Συνήθως, αυτό που καταλαβαίνουμε με μια διαδικασία ανανέωσης είναι ότι οι τυχαίες μεταβλητές $N(t)$ αναπαριστούν το πλήθος των φορών, στις οποίες συμβαίνει κάποιο επαναλαμβανόμενο γεγονός κατά τη διάρκεια του χρονικού διαστήματος $[0, t]$, όπως, π.χ., το πλήθος αφίξεων ή άλλων χαρακτηριστικών συμβάντων. Με την έννοια αυτή, θα ονομάζουμε το T_n "*χρόνο της n-οστής άφιξης*" και το X_n "*n-οστό χρονικό διάστημα μεταξύ δυο διαδοχικών αφίξεων*".

Παρατήρηση 3.7. Για τον ορισμό μιας διαδικασίας ανανέωσης αρκεί να δίνεται απλώς μια στοχαστική διαδικασία συνεχούς χρόνου $N = \{N(t)\}$ (που παίρνει ακέραιες μη αρνητικές τιμές και είναι μη φθίνουσα), από την οποία μπορεί να ορισθούν η ακολουθία T_n των χρόνων των αφίξεων και η ακολουθία X_n των χρονικών διαστημάτων μεταξύ δυο διαδοχικών αφίξεων ως εξής:

$$T_n = \inf\{t \colon N(t) = n\}, \quad X_n = T_n - T_{n-1},$$

έτσι ώστε η $\{X_n\}$ να είναι μια ακολουθία ανεξάρτητων και όμοια κατανεμημένων τυχαίων μεταβλητών (μη αρνητικών τιμών).

Θεώρημα 3.23. *Κάθε διαδικασία Poisson είναι μια διαδικασία ανανέωσης. Επιπλέον, η διαδικασία Poisson είναι η μοναδική διαδικασία ανανέωσης, που είναι αλυσίδα Markov.*

Ορισμός 3.35. Η συνάρτηση ανανέωσης $m(t)$ ορίζεται σαν η μέση τιμή του $N(t)$, δηλαδή,

$$m(t) = E(N(t)).$$

Θεώρημα 3.24. *Αν $\mu = E(X_1) < \infty$, τότε*

$$\frac{1}{t} N(t) \xrightarrow{\sigma.\sigma.} \frac{1}{\mu}, \ \textit{καθώς } t \to \infty,$$

Παραπάνω, "*σ.σ.*" σημαίνει "*σχεδόν σίγουρα*", δηλαδή, ότι η πιθανότητα να συμβεί κάποιο γεγονός (εδώ, η σύγκλιση του παραπάνω ορίου) είναι 1.

Θεώρημα 3.25. *Αν $\mu = E(X_1) < \infty$, τότε*

$$\frac{1}{t} m(t) \to \frac{1}{\mu}, \ \textit{καθώς } t \to \infty.$$

Ας υποθέσουμε τώρα ότι τη χρονική στιγμή t αρχίζουμε να παρατηρούμε μια διαδικασία ανανέωσης N. Ως τότε έχουν ήδη λάβει χώρα $N(t)$ συμβάντα, ας πούμε, αφίξεις, και η ακριβώς επόμενη άφιξη θα είναι η $(N(t)+1)$-οστή. Με άλλα λόγια, ο χρόνος t ανήκει στο τυχαίο διάστημα $I_t = [T_{N(t)}, T_{N(t)+1})$. Τότε, έχει μεγάλο ενδιαφέρον να μπορούμε να προσδιορίσουμε τις εξής τρεις τυχαίες μεταβλητές:

(i) Ο *υπόλοιπος χρόνος ζωής* στο t του I_t: $E(t) = T_{N(t)+1} - t$.

(ii) Ο *παρών χρόνος ζωής* στο t του I_t: $C(t) = t - T_{N(t)}$.

(iii) Ο *συνολικός χρόνος ζωής* στο t του I_t: $D(t) = E(t) + C(t) = X_{N(t)+1}$.

3.2. Διαδικασίες Ανανέωσης-Αμοιβής και το Θεώρημα του Little

Έστω η διαδικασία ανανέωσης $N = N(t)$ ορισμένη ως $N(t) = \sup\{n\colon T_n \leq t\}$, όπου $T_n = X_1 + X_2 + \ldots + X_n$. Θεωρούμε τότε τα ζευγάρια $\{(X_i, R_i) : i \geq 1\}$ των ανεξάρτητων και όμοια κατανεμημένων τυχαίων μεταβλητών X_i και R_i με $X_i \geq 0$ (αλλά μεταξύ τους οι τυχαίες μεταβλητές X_i και R_i δεν υποτίθεται ότι είναι ανεξάρτητες). Η τυχαία μεταβλητή X_i είναι ο χρόνος μεταξύ δυο διαδοχικών αφίξεων της διαδικασίας ανανέωσης N και η τυχαία μεταβλητή R_i θεωρείται ότι είναι κάποια '*αμοιβή*,' που απονέμεται μεταξύ δυο διαδοχικών αφίξεων της N. Βέβαια, η αμοιβή θα μπορούσε να ήταν αρνητική, οπότε θα τη θεωρούσαμε σαν '*κόστος*'.

Ορίζουμε τη *διαδικασία συνολικής αμοιβής* $C = C(t)$ ως

$$C(t) = \sum_{i=1}^{N(t)} R_i$$

και τη *συνάρτηση αμοιβής* $c(t)$ ως τη μέση τιμή της $C(t)$, δηλαδή,

$$c(t) = E(C(t)).$$

Σε ό,τι ακολουθεί στην παράγραφο αυτή, θα συμβολίζουμε $X = X_1$ και $R = R_1$.

Θεώρημα 3.26. (**Θεώρημα Ανανέωσης-Αμοιβής**.) *Αν* $0 < E(X) < \infty$ *και* $E(|R|) < \infty$, *τότε*

$$\frac{C(t)}{t} \xrightarrow{\sigma.\sigma.} \frac{E(R)}{E(X)}, \quad \textit{καθώς } t \to \infty,$$
$$\frac{c(t)}{t} \longrightarrow \frac{E(R)}{E(X)}, \quad \textit{καθώς } t \to \infty.$$

Θα εφαρμόσουμε, στη συνέχεια, το θεώρημα ανανέωσης-αμοιβής για ένα σύστημα ουράς αναμονής. Παρότι θα ασχοληθούμε λεπτομερέστερα με διαδικασίες ουρών στην επόμενη παράγραφο, ας πούμε εδώ ότι με μια "*ουρά αναμονής και εξυπηρέτησης*" εννοούμε ένα σύστημα, στο οποίο οι πελάτες φθάνουν ένας-ένας μπροστά σε μια μονάδα (ή, γενικώς, περισσότερες) εξυπηρέτησης για να εξυπηρετηθούν με τη σειρά. Ο n-οστός πελάτης δαπανά ένα *χρόνο αναμονής* στο σύστημα (συμπεριλαμβανομένου και του χρόνου εξυπηρέτησης) και μετά αναχωρεί.

Υποθέτουμε ότι υπάρχει σχεδόν σίγουρα (δηλαδή, με πιθανότητα 1) ένας πεπερασμένος (τυχαίος) χρόνος $T > 0$, ο οποίος ονομάζεται *χρόνος αναγέννησης* και είναι τέτοιος ώστε η διαδικασία, που αρχίζει το χρόνο T, να έχει την ίδια κατανομή με την αντίστοιχη διαδικασία, που αρχίζει το χρόνο 0. Επιπλέον, υποθέτουμε ότι ο αριθμός των πελατών στο σύστημα $Q(t)$ το χρόνο t είναι τέτοιος ώστε $Q(0) = Q(T) = 0$. Με άλλα λόγια, υπάρχει μια ακολουθία χρόνων $0 = T_0 < T_1 < T_2 < \ldots$, κάθε ένας από τους οποίους είναι χρόνος αναγέννησης της διαδικασίας, έτσι ώστε τα χρονικά διαστήματα μεταξύ δυο διαδοχικών χρόνων αναγέννησης $X_i = T_i - T_{i-1}$ να είναι ανεξάρτητα και όμοια κατανεμημένα. Με άλλα λόγια, η ουρά, που εξετάζουμε, είναι η διαδικασία ανανέωσης των παραπάνω χρόνων αναγέννησης.

Τα χρονικά διαστήματα $[T_{i-1}, T_i)$ συνήθως ονομάζονται *κύκλοι* αυτής της διαδικασίας. Παρατηρούμε ότι η στοχαστική διαδικασία $P_i = \{Q(t) : T_{i-1} \leq t < T_i\}, i \geq 1$, αποτελείται από ανεξάρτητες και όμοια κατανεμημένες τυχαίες μεταβλητές. Γράφουμε N_i για τον αριθμό των πελατών, που φθάνουν κατά τη διάρκεια του κύκλου $[T_{i-1}, T_i)$, και συμβολίζουμε $N = N_1, T = T_1$. Επιπλέον, υποθέτουμε ότι τα σημεία αναγέννησης επιλέγονται έτσι ώστε $N_i > 0$, για κάθε i. Τέλος, υποθέτουμε ότι

$$E(T) < \infty, \quad E(N) < \infty, \quad E(NT) < \infty.$$

Εφαρμόζοντας το θεώρημα ανανέωσης-αμοιβής τρεις φορές, καταλήγουμε στα εξής αποτελέσματα:

(A) Θεωρούμε τη διαδικασία ανανέωσης-αμοιβής με χρόνους αφίξεων $T_0, T_1, T_2, \ldots$, στην οποίαν αμοιβή στο χρονικό διάστημα μεταξύ δυο διαδοχικών αφίξεων $X_i = T_i - T_{i-1}$ ορίζεται ως:

$$R_i = \int_{T_{i-1}}^{T_i} Q(u)\, du.$$

Οι R_i έχουν την ίδιαν κατανομή με την $R = R_1 = \int_0^T Q(u)\, du$. Επιπλέον, $Q(u) \leq N$, όταν $0 \leq u \leq T$, και, άρα, $E(R) \leq E(NT) < \infty$. Επομένως, από το θεώρημα ανανέωσης-αμοιβής παίρνουμε:

$$\frac{1}{t} \int_0^t Q(u)\, du \xrightarrow{\sigma.\sigma.} \frac{E(R)}{E(T)}, \quad \text{καθώς } t \to \infty.$$

Ο λόγος $E(R)/E(T)$ ονομάζεται *ασυμπτωτικό μέσο μήκος ουράς* και συμβολίζεται με L.

(B) Θεωρούμε στη συνέχεια μια δεύτερη διαδικασία ανανέωσης-αμοιβής με χρόνους αφίξεων $T_0, T_1, T_2, \ldots$, στην οποίαν η αμοιβή στο χρονικό διάστημα μεταξύ δυο διαδοχικών αφίξεων ορίζεται τώρα να είναι ίση με τον αριθμό N_i των πελατών, που φθάνουν κατά τη διάρκεια του αντίστοιχου κύκλου. Πάλι από το θεώρημα ανανέωσης-αμοιβής παίρνουμε:

$$\frac{N(t)}{t} \xrightarrow{\sigma.\sigma.} \frac{E(N)}{E(T)}, \quad \text{καθώς } t \to \infty.$$

Ο λόγος $E(N)/E(T)$ ονομάζεται *ασυμπτωτικός ρυθμός αφίξεων* και συμβολίζεται με λ.

(C) Τέλος, θεωρούμε τη διαδικασία ανανέωσης-αμοιβής με χρονικά διαστήματα μεταξύ δυο διαδοχικών χρόνων αφίξεων $N_1, N_2, \ldots$, στην οποίαν η αμοιβή S_i στο διάστημα N_i μεταξύ δυο διαδοχικών αφίξεων ορίζεται ίση με το άθροισμα των χρόνων αναμονής των πελατών, που φθάνουν κατά τη διάρκεια του i-οστού κύκλου της ουράς. Πάλι από το θεώρημα ανανέωσης-αμοιβής παίρνουμε:

$$\frac{1}{n} \sum_{i=1}^{n} V_i \xrightarrow{\sigma.\sigma.} \frac{E(S)}{E(N)}, \quad \text{καθώς } t \to \infty.$$

Ο λόγος $E(S)/E(N)$ ονομάζεται *ασυμπτωτικός μέσος χρόνος αναμονής* και συμβολίζεται με W.

Θεώρημα 3.27. (Το Θεώρημα του Little.) *Για κάθε διαδικασία ανανέωσης-αμοιβής, έχουμε*

$$L = \lambda W.$$

3.3. Διαδικασίες Ουρών Αναμονής και Εξυπηρέτησης

Στη συνέχεια, θα εξετάσουμε τις διαδικασίες ουρών αναμονής σαν διαδικασίες ανανέωσης. Για το σκοπό αυτό, υποθέτουμε ότι οι πελάτες φθάνουν σε ένα σταθμό εξυπηρέτησης, στον οποίο λειτουργεί κάποιο συγκεκριμένο πλήθος μονάδων (ή συστημάτων) εξυπηρέτησης, για κάποιο καθορισμένο σκοπό, ανάλογα με την εφαρμογή, που μελετάμε. Επειδή, όταν φθάσει κάποιος πελάτης, είναι πιθανό όλες οι μονάδες εξυπηρέτησης να είναι κατειλημμένες από άλλους πελάτες, για αυτό, μπορεί οι πελάτες που φθάνουν στο σταθμό εξυπηρέτησης να πρέπει να περιμένουν για κάποιο χρονικό διάστημα μέχρις ότου να αδειάσει μια μονάδα εξυπηρέτησης και μετά αυτά να εξυπηρετηθούν (εκτός αν, όταν όλες οι μονάδες εξυπηρέτησης είναι κατειλημμένες, οι πελάτες είναι υποχρεωμένοι να φύγουν). Γενικώς πάντως, όταν έρθει η σειρά του, ένας πελάτης εξυπηρετείται από το σύστημα και αμέσως μετά φεύγει.

Για να περιγράψουμε ένα τέτοιο σύστημα επακριβώς, πρέπει πρώτα να έχουμε περισσότερες πληροφορίες για το πώς γίνονται κάποιες λεπτομέρειες της διαδικασίας αυτής. Για παράδειγμα, χρειάζεται να γνωρίζουμε τέτοια ζητήματα, όπως πώς ακριβώς εισέρχονται στο σύστημα οι πελάτες, με ποια σειρά εξυπηρετούνται και πόσο διαρκεί η περίοδος της εξυπηρέτησής τους. Κάποιες γενικές απαντήσεις στα ζητήματα αυτά δίνονται με τις παρακάτω διευκρινίσεις:

(i) Ο αριθμός $N(t)$ των πελατών, που έχουν εισέλθει στο σύστημα μέχρι το χρόνο t, αποτελεί μια διαδικασία ανανέωσης. Δηλαδή, αν T_n είναι ο χρόνος άφιξης του n-οστού πελάτη (με τη σύμβαση ότι $T_0 = 0$), τότε τα *χρονικά διαστήματα μεταξύ δυο διαδοχικών αφίξεων* $X_n = T_n - T_{n-1}$ σχηματίζουν μια ανεξάρτητα και όμοια κατανεμημένη στοχαστική διαδικασία.

(ii) Οι τελευταία αφικνιόμενοι πελάτες εισέρχονται στο τέλος της ουράς αναμονής, δηλαδή, η εξυπηρέτηση των πελατών ακολουθεί τον κανόνα: "*αυτός που φθάνει πρώτος, εξυπηρετείται και πρώτος*" (στα αγγλικά "*First In, First Out*"). Με την έννοια (επειδή μπορεί να υπάρχουν περισσότερες της μιας μονάδες εξυπηρέτησης) ότι υπάρχει μια μοναδική ουρά ανομονής και, όταν μια μονάδα εξυπηρέτησης αφήνεται ελεύθερη, τότε ο πρώτος στην ουρά (σειρά) πελάτης πηγαίνει εκεί για να εξυπηρετηθεί.

(iii) Συμβολίζοντας με S_n το *χρόνο εξυπηρέτησης* του n-οστού πελάτη, υποθέτουμε ότι οι χρόνοι εξυπηρέτησης S_n είναι ανεξάρτητες και όμοια κατανεμημένες τυχαίες μεταβλητές, οι οποίες δεν εξαρτώνται από το χρόνο άφιξης των πελατών. Ακόμη, έστω V_n ο συνολικός *χρόνος αναμονής* στο σύστημα του n-οστού πελάτη (συμπεριλαμβανομένου του χρόνου εξυπηρέτησης S_n).

Έστω αριθμός $Q(t)$ των πελατών στο σύστημα το χρόνο t (συμπεριλαμβανομένων κι εκείνων που εξυπηρετούνται τότε). Έχει ιδιαίτερο ενδιαφέρον η τιμή του ορίου του $Q(t)$, καθώς $t \to \infty$. Προφανώς, δεν είναι επιθυμητές διαδικασίες, για τις οποίες $\lim_{t\to\infty} Q(t) = \infty$. Έτσι, μια ουρά ονομάζεται *ευσταθής*, αν υπάρχει το όριο $\lim_{t\to\infty} Q(t)$, και διαφορετικά *ασταθής*.

Συμβολίζοντας με $E(S)$ τη μέση τιμή της στοχαστικής διαδικασίας των χρόνων εξυπηρέτησης S_n και με $E(X)$ τη μέση τιμή της στοχαστικής διαδικασίας των χρονικών διαστημάτων μεταξύ δυο διαδοχικών αφίξεων X_n, ορίζουμε την *ένταση ρ της κυκλοφορίας της ουράς* σαν τον εξής λόγο:

$$\rho = \frac{E(S)}{E(X)}.$$

Θεώρημα 3.28. *Έστω $Q = \{Q(t) : t \geq 0\}$ μια ουρά σε ένα σύστημα με μοναδική μονάδα εξυπηρέτησης. Έστω ρ η ένταση κυκλοφορίας της ουράς αυτής.*

(i) Αν $\rho < 1$, τότε η ουρά Q είναι ευσταθής.

(ii) Αν $\rho > 1$, τότε η ουρά Q είναι ασταθής.

(iii) Αν $\rho = 1$ και τουλάχιστον μια από τις S_n και X_n έχει αυστηρά θετική μεταβλητότητα, τότε η ουρά Q είναι ασταθής.

3.4. Συμβολισμός Ουρών

Συνοψίζοντας τα προηγούμενα, μπορούμε να πούμε ότι σε κάθε ουρά αντιστοιχούν δυο ακολουθίες ανεξάρτητων και όμοια κατανεμημένων τυχαίων μεταβλητών: η ακολουθία $\{X_n : n \geq 1\}$ των χρονικών διαστημάτων μεταξύ δυο διαδοχικών αφίξεων των πελατών με κοινή συνάρτηση κατανομής F_X και η ακολουθία $\{S_n : n \geq 1\}$ των χρόνων εξυπηρέτησης των πελατών με κοινή συνάρτηση κατανομής F_S. Υποθέτουμε ότι οι αφίξεις των πελατών γίνονται όπως στις διαδικασίες ανανέωσης, δηλαδή, γνωρίζουμε τα χρονικά διαστήματα X_n μεταξύ δυο διαδοχικών αφίξεων, οπότε ο n-οστός πελάτης ξέρουμε ότι φθάνει το χρόνο $T_n = X_1 + X_2 + \ldots + X_n$. Κάθε αφικνιόμενος πελάτης στέκεται στη μοναδική ουρά εξυπηρέτησης, όπου μπροστά του βρίσκονται όλοι οι πελάτες που έχουν ήδη φθάσει πριν από αυτόν και περιμένουν να εξυπηρετηθούν από τον πρώτο *εξυπηρετητή* (server), που θα είναι διαθέσιμος. Όταν έρθει η σειρά του n-οστού πελάτη να εξυπηρετηθεί, δηλαδή, όταν αυτός βρίσκεται στην κεφαλή της ουράς και συγχρόνως αδειάζει κάποιος εξυπηρετητής, ο n-οστός πελάτης θα εξυπηρετηθεί μέσα στο χρονικό διάστημα S_n και αμέσως μετά θα φύγει από το σύστημα. Συμβολίζουμε με $Q(t)$ τον αριθμό των πελατών, που βρίσκονται στην ουρά της αναμονής εξυπηρέτησης το χρόνο t (συμπεριλαμβανομένων κι εκείνων που εξυπηρετούνται το χρόνο t), και, φυσικά, έχουμε $Q(0) = 0$. Προφανώς, η $Q = \{Q(t) : t \geq 0\}$ είναι μια στοχαστική διαδικασία, της οποίας οι κατανομές εξαρτώνται από τις συναρτήσεις κατανομών F_X και F_S. Επομένως, διάφορα ζητήματα, που μας ενδιαφέρουν να κατανοήσουμε (όπως, π.χ., αν η Q είναι ή όχι διαδικασία Markov, ποια είναι η ασυμπτωτική συμπεριφορά της $Q(t)$, καθώς $t \to \infty$, κ.ο.κ.) εξαρτώνται από τον επακριβή καθορισμό των συναρτήσεων κατανομών F_X και F_S.

Έτσι, για να περιγραφεί επακριβώς μια ουρά, χρησιμοποιούμε το συμβολισμό $A/B/s$, που εισήχηθκε από τον D. Kendall, όπου το A καθορίζει την F_X, το B την F_S και το s τον αριθμό των μονάδων εξυπηρέτησης (εξυπηρετητών) του συστήματος. Τυπικές περιπτώσεις για τα A και B είναι:

$$\begin{aligned} D(d) &\equiv \text{κατανομή σχεδόν σίγουρα συγκεντρωμένη στην τιμή } d \text{ (} D \text{ από το ‘}\textit{ντετερμινιστική}\text{’)},\\ M(\lambda) &\equiv \text{εκθετική κατανομή με παράμετρο } \lambda \text{ (} M \text{ από το ‘}\textit{Μαρκοβιανή}\text{’)},\\ \Gamma(\lambda, k) &\equiv \text{κατανομή γάμμα με παραμέτρους } \lambda \text{ και } k,\\ G &\equiv \text{κάποια γενική κατανομή, σταθερή αλλά μη καθορισμένη.} \end{aligned}$$

Αναφορικά τώρα με τον επακριβή τρόπο, με τον οποίον γίνεται η εξυπηρέτηση των πελατών, ας επανελάβουμε ότι υποθέτουμε πως ακολουθείται ο κανόνας *FIFO*, αρχικά προερχόμενα από την αγγλική έκφραση “*First In, First Out*”, με την οποίαν εννοούμε ότι “*αυτός που φθάνει πρώτος, εξυπηρετείται και πρώτος.*”

Μερικές φορές όμως, αντί του $A/B/s$, χρησιμοποιείται ο αναλυτικότερος συμβολισμός $A/B/s/K/N$, όπου K είναι η *χωρητικότητα* του συστήματος εξυπηρέτησης και N το συνολικό πλήθος των πελατών, που θέλουν να εξυπηρετηθούν από το σύστημα αυτό. Η χωρητικότητα K $(\geq s)$ του συστήματος καθορίζει το επιτρεπόμενο πλήθος πελατών, που μπορούν ταυτόχρονα να περιμένουν να εξυπηρετηθούν από τις μονάδες εξυπηρέτησης του συστήματος (συμπεριλαμβανομένων των s πελατών, που εξυπηρετούνται τον ίδιο χρόνο). Αυτό σημαίνει ότι, αν κάποια στιγμή βρίσκονται στο σύστημα K πελάτες, τότε το σύστημα διώχνει (είναι κλειστό για) όλους τους άλλους πελάτες, που έρχονται στη συνέχεια, πριν αδειάσει κάποιος εξυπηρετητής. Πάντως, ακολουθούμε την τυπική σύμβαση ότι, αν είναι απόν ένα από τα δυο τελευταία ορίσματα στον αναλυτικό συμβολισμό των ουρών, τότε υποθέτουμε πως το απόν όρισμα παίρνει την τίμη άπειρο.

Επιπλέον, ας παρατηρήσουμε ότι συνήθως η κατανομή $M(\lambda)$ στο συμβολισμό των ουρών γράφεται σε συντομία απλώς σαν M (για παράδειγμα, συνήθως γράφουμε $M/M/1$ αντί του πληρέστερου συμβολισμού $M(\lambda)/M(\mu)/1$).

3.5. Ουρές $M/M/1$

Στην ουρά $M/M/1$, που είναι συντόμευση του $M(\lambda)/M(\mu)/1$, τα χρονικά διαστήματα μεταξύ διαδοχικών αφίξεων κατανέμονται εκθετικά με παράμετρο λ και οι χρόνοι εξυπηρέτησης πάλι εκθετικά με παράμετρο μ. Δηλαδή, οι πελάτες φθάνουν σύμφωνα με μια διαδικασία Poisson με ένταση λ σε ένα σύστημα εξυπηρέτησης, στο οποίο λειτουργεί μια μόνο μονάδα εξυπηρέτησης. Η διαδικασία του αριθμού $Q = \{Q(t) : t \geq 0\}$ των εξυπηρετούμενων πελατών σχηματίζει μια αλυσίδα Markov συνεχούς χρόνου στο χώρο καταστάσεων $\{0, 1, 2, \ldots\}$, λόγω του ότι η εκθετική κατανομή έχει την ιδιότητα να στερείται μνήμης. Επιπλέον, η Q είναι μια διαδικασία γέννησης-θανάτου με τους παρακάτω ρυθμούς:

$$\lambda_n = \lambda\,, \quad \text{για } n = 0, 1, 2, \ldots\,, \qquad \mu_n = \begin{cases} \mu, & \text{αν } n \geq 1, \\ 0, & \text{αν } n = 0. \end{cases}$$

Οι πιθανότητες $p_n(t) = P(Q(t) = n)$ ικανοποιούν τις προς τα εμπρός εξισώσεις Kolmogorov:

$$\begin{aligned} \frac{dp_n}{dt} &= \lambda p_{n-1}(t) - (\lambda + \mu) p_n(t) + \mu p_{n+1}(t), \quad \text{για } n \geq 1, \\ \frac{dp_0}{dt} &= -\lambda p_0(t) + \mu p_1(t), \end{aligned}$$

οι οποίες πρέπει να λυθούν μαζί με τις αρχικές συνθήκες $p_n(0) = \delta_{0n}$, το δέλτα του Kronecker. Για τη λύση των εξισώσεων αυτών, μπορούμε να χρησιμοποιήσουμε το μετασχηματισμό Laplace της p_n, που δίνεται ως

$$\widehat{p}_n(\theta) = \int_0^\infty e^{-\theta t} p_n(t)\, dt.$$

Θεώρημα 3.29. *Ο μετασχηματισμός Laplace της λύσης των προς τα εμπρός εξισώσεων Kolmogorov είναι*

$$\widehat{p}_n(\theta) = \theta^{-1}[1 - \alpha(\theta)]\alpha(\theta)^n,$$

όπου

$$\alpha(\theta) = \frac{(\lambda + \mu + \theta) - \sqrt{(\lambda + \mu + \theta)^2 - 4\lambda\mu}}{2\mu}.$$

Παρότι μπορεί να εξαχθεί σαν πόρισμα του προηγούμενου αποτελέσματος, είναι ευκολότερο να μελετήσουμε απευθείας την ασυμπτωτική συμπεριφορά της $Q(t)$, καθώς $t \to \infty$, για να βρούμε:

Θεώρημα 3.30. *Έστω $\rho = \lambda/\mu$ η ένταση της κυκλοφορίας της ουράς $M/M/1$.*

(i) *Αν $\rho < 1$, τότε $\lim_{t\to\infty} P(Q(t) = n) = (1 - \rho)\rho^n = \pi_n$, για όλα τα $n \geq 0$, όπου $\pi = (\pi_n : n \geq 0)$ είναι η μοναδική στάσιμη κατανομή της Q.*

(ii) *Αν $\rho \geq 1$, τότε δεν υπάρχει στάσιμη κατανομή της Q και $\lim_{t\to\infty} P(Q(t) = n) = 0$, για όλα τα $n \geq 0$.*

3.6. Ουρές $M/M/s$ και $M/M/\infty$

Στην ουρά $M(\lambda)/M(\mu)/s$, οι πελάτες φθάνουν πάλι σύμφωνα με μια διαδικασία Poisson με ένταση λ σε ένα σύστημα εξυπηρέτησης, στο οποίο λειτουργούν τώρα s μονάδες εξυπηρέτησης, και εξυπηρετούνται σε χρόνους, που κατανέμονται κι αυτοί εκθετικά με παράμετρο μ. Η διαδικασία του αριθμού $Q = \{Q(t) : t \geq 0\}$ των εξυπηρετούμενων πελατών σχηματίζει πάλι μια αλυσίδα Markov συνεχούς χρόνου στο χώρο καταστάσεων $\{0, 1, 2, \ldots\}$, η οποία είναι μια διαδικασία γέννησης-θανάτου με τους παρακάτω ρυθμούς:

$$\begin{aligned} \lambda_n &= \lambda\,, \quad \text{για } n = 0, 1, 2, \ldots\,, \\ \mu_n &= \min\{n\mu, s\mu\} = \begin{cases} n\mu, & \text{αν } 0 \leq n \leq s, \\ s\mu, & \text{αν } n \geq s. \end{cases} \end{aligned}$$

Όπως προηγούμενα, παίρνουμε τώρα το παρακάτω αποτέλεσμα για την ασυμπτωτική συμπεριφορά της $Q(t)$, καθώς $t \to \infty$.

Θεώρημα 3.31. *Έστω $\rho = \lambda/(s\mu)$ η ένταση της κυκλοφορίας της ουράς $M/M/s$.*

(i) Αν $\rho < 1$, τότε η μοναδική στάσιμη κατανομή της Q είναι η $\pi = (\pi_n\colon\, n \geq 0)$, $\pi_n = \lim_{t\to\infty} P(Q(t) = n)$, για $n \geq 0$, με:

$$\pi_n = \begin{cases} \dfrac{(s\rho)^n}{n!}\,\pi_0, & \text{αν } \; n \leq s, \\[2ex] \dfrac{(\rho)^n s^s}{s!}\,\pi_0,, & \text{αν } \; n \geq s, \end{cases}$$

όπου

$$\pi_0 = \left[\sum_{n=0}^{s-1} \frac{(s\rho)^n}{n!} + \left(\frac{(s\rho)^s}{s!}\right)\left(\frac{1}{1-\rho}\right)\right]^{-1}.$$

(ii) Αν $\rho \geq 1$, τότε δεν υπάρχει στάσιμη κατανομή της Q και $\lim_{t\to\infty} P(Q(t) = n) = 0$, για όλα τα $n \geq 0$.

Από τους προηγούμενους τύπους, μπορούμε να εξάγουμε την πιθανότητα $P(\text{queueing})$ να μπει ένας πελάτης στην ουρά αναμονής (επειδή όλες οι μονάδες εξυπηρέτησης του συστήματος είναι κατειλημμένες):

$$\begin{aligned} P(\text{queueing}) &= \sum_{n=s}^{\infty} \pi_n = \\ &= \frac{\left(\frac{(s\rho)^s}{s!}\right)\left(\frac{1}{1-\rho}\right)}{\left[\sum_{n=0}^{s-1} \frac{(s\rho)^n}{n!} + \left(\frac{(s\rho)^s}{s!}\right)\left(\frac{1}{1-\rho}\right)\right]}. \end{aligned}$$

Η προηγούμενη έκφραση για την πιθανότητα $P(\text{queueing})$ να μπει ένας πελάτης στην ουρά αναμονής ονομάζεται *τύπος C του Erlang* και συχνά συμβολίζεται ως $C(s, \rho)$, για $\rho = \lambda/\mu$.

Ερχόμαστε τώρα στην ουρά $M(\lambda)/M(\mu)/\infty$. Η περίπτωση αυτή μπορεί εύκολα να αναλυθεί όπως προηγουμένως, παρατηρώντας ότι τώρα οι ρυθμοί της αντίστοιχης διαδικασίας γέννησης-θανάτου είναι:

$$\begin{aligned} \lambda_n &= \lambda\,, & \text{για } n = 0, 1, 2, \ldots\,, \\ \mu_n &= n\mu\,, & \text{για } n = 1, 2, 3, \ldots\,. \end{aligned}$$

Όπως προηγούμενα, παίρνουμε τώρα το παρακάτω αποτέλεσμα για την ασυμπτωτική συμπεριφορά της $Q(t)$, καθώς $t \to \infty$.

Θεώρημα 3.32. *Έστω $\rho = \lambda/\mu < \infty$ η ένταση της κυκλοφορίας της ουράς $M/M/\infty$. Τότε η μοναδική στάσιμη κατανομή της Q είναι η $\pi = (\pi_n\colon n \geq 0)$, $\pi_n = \lim_{t\to\infty} P(Q(t) = n)$, για $n \geq 0$, με:*

$$\pi_n = \frac{\rho^n}{n!}\, e^{-\rho}, \quad \text{για } \; n = 0, 1, 2, \ldots\,.$$

3.7. Ουρές $M/M/s/s$ με Απώλειες

Στην ουρά $M(\lambda)/M(\mu)/s/s$, λειτουργούν πάλι s μονάδες εξυπηρέτησης αλλά και η χωρητικότητα του συστήματος είναι ακριβώς s πελάτες. Αυτό σημαίνει ότι, όταν όλες οι s μονάδες εξυπηρέτησης του συστήματος είναι κατειλημμένες (δηλαδή, όταν υπάρχουν s πελάτες στο σύστημα), τότε υποχρεωτικά φεύγουν (χάνονται) όσοι τυχόν πελάτες έρθουν, πριν αδειάσει κάποιος εξυπηρετητής. Η περίπτωση αυτή μπορεί εύκολα να αναλυθεί όπως προηγουμένως, παρατηρώντας ότι τώρα οι ρυθμοί της αντίστοιχης διαδικασίας γέννησης-θανάτου είναι:

$$\lambda_n = \begin{cases} \lambda, & \text{αν } 0 \le n < s, \\ 0, & \text{αν } n \ge s, \end{cases}$$

$$\mu_n = \begin{cases} n\mu, & \text{αν } 1 \le n \le s, \\ 0, & \text{αν } n > s. \end{cases}$$

Επομένως, τώρα η διαδικασία του αριθμού $Q = \{Q(t) \colon t \ge 0\}$ των εξυπηρετούμενων πελατών σχηματίζει μια αλυσίδα Markov συνεχούς χρόνου στον πεπερασμένο χώρο καταστάσεων $\{0, 1, \dots, s\}$. Όπως προηγούμενα, παίρνουμε τώρα το παρακάτω αποτέλεσμα για την ασυμπτωτική συμπεριφορά της $Q(t)$, καθώς $t \to \infty$.

Θεώρημα 3.33. *Έστω $\rho = \lambda/\mu < \infty$ η ένταση της κυκλοφορίας της ουράς $M/M/s/s$. Τότε η μοναδική στάσιμη κατανομή της Q είναι η $\pi = (\pi_n \colon n \ge 0)$, $\pi_n = \lim_{t\to\infty} P(Q(t) = n)$, για $n \ge 0$, με:*

$$\pi_n = \begin{cases} \dfrac{\rho^n}{n!}\,\pi_0, & \text{για } \ 0 \le n \le s, \\ \\ 0, & \text{για } \ n > s. \end{cases}$$

όπου

$$\pi_0 = \left[\sum_{n=0}^{s} \frac{\rho^n}{n!}\right]^{-1}.$$

Από το παραπάνω αποτέλεσμα, μπορούμε εύκολα να υπολογίσουμε το κλάσμα του χρόνου, στο οποίο όλοι οι s εξυπηρετητές του συστήματος είναι κατειλημμένοι. Το κλάσμα αυτό δίνεται από την πιθανότητα π_s, η οποία είναι:

$$\pi_s = \frac{\dfrac{\rho^s}{s!}}{\displaystyle\sum_{n=0}^{s} \frac{\rho^n}{n!}}.$$

Αυτή η έκφραση ονομάζεται *τύπος B των απωλειών του Erlang* και συχνά συμβολίζεται ως $B(s, \rho)$, για $\rho = \lambda/\mu$.

3.8. Ουρές $M/G/1$

Στην ουρά $M(\lambda)/G/1$, οι πελάτες φθάνουν σύμφωνα με μια διαδικασία Poisson έντασης λ. Έστω D_n ο χρόνος αναχώρησης (από τον εξυπηρετητή) του n-οστού πελάτη (εννοείται, φυσικά, αφού αυτός εξυπηρετηθεί) και έστω $Q(D_n)$ ο αριθμός των πελατών, που αυτός αφήνει πίσω του στην ουρά αμέσως μετά την αναχώρησή του.

Θεώρημα 3.34. *Η διαδικασία $Q(D) = \{Q(D_n) \colon n \ge 1\}$ είναι μια αλυσίδα Markov με πίνακα πιθανοτήτων μεταβάσεων:*

$$P_D = \begin{pmatrix} \delta_0 & \delta_1 & \delta_2 & \dots \\ \delta_0 & \delta_1 & \delta_2 & \dots \\ 0 & \delta_0 & \delta_1 & \dots \\ 0 & 0 & \delta_0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

όπου

$$\delta_j = E\biggl(\frac{(\lambda S)^j}{j\,!}\, e^{-\lambda S}\biggr)$$

και S είναι ένας τυπικός χρόνος εξυπηρέτησης.

Θεώρημα 3.35. *Έστω $\rho = \lambda E(S)$ η ένταση της κυκλοφορίας της ουράς $M/G/1$.*

(i) Αν $\rho < 1$, τότε η αλυσίδα $Q(D)$ είναι εργοδική και έχει μια μοναδική στάσιμη κατανομή π με γεννήτρια συνάρτηση

$$G(s) = \sum_j \pi_j s^j = (1-\rho)(s-1)\frac{M_S(\lambda(s-1))}{s - M_S(\lambda(s-1))},$$

όπου M_S είναι η ροπο-γεννήτρια συνάρτηση ενός τυπικού χρόνου εξυπηρέτησης.

(ii) Αν $\rho > 1$, τότε η $Q(D)$ είναι παροδική.

(iii) Αν $\rho = 1$, τότε η $Q(D)$ είναι μηδενικά επαναφερόμενη.

Θεώρημα 3.36. *Όταν $\rho < 1$ και η ουρά $M/G/1$ βρίσκεται σε ισορροπία (δηλαδή, έχει κατανομή π), ο χρόνος αναμονής W των πελατών, πριν αυτοί αρχίσουν να εξυπηρετούνται, έχει τη ροπο-γεννήτρια συνάρτηση*

$$M_W(s) = \frac{(1-\rho)s}{\lambda + s - \lambda M_S(s)}.$$

3.9. Ουρές $G/M/1$

Στην ουρά $G/M(\mu)/1$, οι πελάτες εξυπηρετούνται σύμφωνα με μια διαδικασία Poisson έντασης μ. Έστω $Q(A_n)$ ο αριθμός των πελατών, που βρίσκονται μπροστά από τον n-οστό πελάτη, τη στιγμή που μόλις αυτός εισέρχεται στον εξυπηρετητή.

Θεώρημα 3.37. *Η διαδικασία $Q(A) = \{Q(A_n) \colon n \geq 1\}$ είναι μια αλυσίδα Markov με πίνακα πιθανοτήτων μεταβάσεων:*

$$P_A = \begin{pmatrix} 1-\alpha_0 & \alpha_0 & 0 & 0 & \dots \\ 1-\alpha_0-\alpha_1 & \alpha_1 & \alpha_0 & 0 & \dots \\ 1-\alpha_0-\alpha_1-\alpha_2 & \alpha_2 & \alpha_1 & \alpha_0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

όπου

$$\alpha_j = E\biggl(\frac{(\mu X)^j}{j\,!}\, e^{-\mu X}\biggr)$$

και X είναι ένα τυπικό χρονικό διάστημα μεταξύ διαδοχικών αφίξεων.

Θεώρημα 3.38. *Έστω* $\rho = \{\mu E(X)\}^{-1}$ *η ένταση της κυκλοφορίας της ουράς* $G/M/1$.

(i) *Αν* $\rho < 1$, *τότε η αλυσίδα* $Q(A)$ *είναι εργοδική και έχει μια μοναδική στάσιμη κατανομή* $\pi = (\pi_j : j \geq 0)$,

$$\pi_j = (1-\eta)\eta^j, \quad \text{για } j \geq 0,$$

όπου η *είναι η μικρότερη θετική ρίζα της εξίσωσης* $\eta = M_X(\mu(\eta - 1))$ *και* M_X *είναι η ροπογεννήτρια συνάρτηση της* X.

(ii) *Αν* $\rho > 1$, *τότε η* $Q(A)$ *είναι παροδική.*

(iii) *Αν* $\rho = 1$, *τότε η* $Q(A)$ *είναι μηδενικά επαναφερόμενη.*

Θεώρημα 3.39. *Όταν* $\rho < 1$ *και η ουρά* $G/M/1$ *βρίσκεται σε ισορροπία (δηλαδή, έχει κατανομή* π*), ο χρόνος αναμονής* W *των πελατών, πριν αυτοί αρχίσουν να εξυπηρετούνται, έχει την κατανομή*

$$P(W \leq x) = \begin{cases} 0, & \text{αν } x < 0, \\ 1 - \eta\, e^{-\mu(1-\eta)x}, & \text{αν } x \geq 0. \end{cases}$$

3.10. Ουρές $G/G/1$

Οι ουρές $G/G/1$ αποτελούν τη γενική περίπτωση ουρών, στην οποίαν ούτε τα χρονικά διαστήματα μεταξύ των διαδοχικών αφίξεων ούτε οι χρόνοι εξυπηρέτησης δεν κατανέμονται εκθετικά. Συμβολίζοντας με W_n το χρόνο αναμονής του n-οστού πελάτη, μπορεί να βρεθεί η παρακάτω σχέση μεταξύ του W_n και του W_{n+1} συναρτήσει του χρόνου εξυπηρέτησης S_n του n-οστού πελάτη και του χρονικού διαστήματος X_{n+1} μεταξύ της άφιξης του n-οστού και του $(n+1)$-οστού πελάτη.

Θεώρημα 3.40. (**Η εξίσωση του Lindley**.) *Ισχύει ότι*

$$W_{n+1} = \max\{0, W_n + S_n - X_{n+1}\}.$$

Η εξίσωση του Lindley συνεπάγεται τη σύγκλιση της ακολουθίας των συναρτήσεων κατανομής $F_n(x) = P(W_n \leq x)$ των χρόνων αναμονής W_n, καθώς $n \to \infty$, σύμφωνα με το επόμενο αποτέλεσμα.

Θεώρημα 3.41. *Έστω* $F_n(x) = P(W_n \leq x)$ *οι συναρτήσεις κατανομής των χρόνων αναμονής* W_n. *Τότε*

$$F_{n+1}(x) = \begin{cases} 0, & \text{αν } x < 0, \\ \displaystyle\int_{-\infty}^{0} F_n(x-y)\, dG(y), & \text{αν } x \geq 0, \end{cases}$$

όπου G *είναι η συνάρτηση κατανομής των ανεξάρτητων και όμοια κατανεμημένων τυχαίων μεταβλητών* $U_n = S_n - X_{n+1}$. *Επομένως, υπάρχει το όριο* $F(x) = \lim_{n\to\infty} F_n(x)$ *και ικανοποιεί την εξίσωση Wiener-Hopf*

$$F(x) = \int_{-\infty}^{0} F(x-y)\, dG(y), \quad \text{για } x \geq 0.$$

Θεώρημα 3.42. *Έστω* $\rho = E(S)/E(X)$ *η ένταση μιας ουράς* $G/G/1$ *και έστω* $F(x)$ *μια λύση της εξίσωσης Wiener-Hopf.*

(i) Αν $\rho < 1$, *τότε η* $F(x)$ *είναι μια συνάρτηση κατανομής.*

(ii) Αν είτε $\rho > 1$ *ή* $\rho = 1$ *και* $\mathrm{var}(U) > 0$, *τότε* $F(x) = 0$, *για όλα τα* x.

3.11. Δίκτυα Ουρών

Είναι δυνατόν, όταν ένας πελάτης βγαίνει από μια ουρά, αμέσως μετά, να μπαίνει σε μια άλλη. Μάλιστα, αν αυτό συνεχισθεί, είναι δυνατόν ο πελάτης να θελήσει να ξαναμπεί σε μια ουρά, από την οποίαν είχε ήδη περάσει προηγουμένως. Έτσι, σχηματίζονται τα ονομαζόμενα "*δίκτυα ουρών*," με τη βοήθεια των οποίων μπορούν να μελετηθούν διαδικασίες διαδοχικών περασμάτων σε ουρές, που είναι εντελώς φυσικές σε πολλές πραγματικές καταστάσεις (π.χ., στην κυκλοφορία πακέτων δεδομένων, που διακινούνται μεταξύ διάφορων κόμβων ενός δικτύου υπολογιστών).

Ας υποθέσουμε λοιπόν ότι υπάρχει ένα πεπερασμένο σύνολο S αποτελούμενο από c σταθμούς εξυπηρέτησης, $s_1, s_2, \ldots, s_c$, από τους οποίους περνούν N πελάτες. Αν το χρόνο t η ουρά του σταθμού i περιλαμβάνει $Q_i(t)$ πελάτες, τότε οι πελάτες στο χρόνο t, οι οποίοι βρίσκονται στο συνολικό δίκτυο ουρών του συστήματος, αναπαρίστανται από το διάνυσμα $Q(t) = (Q_1(t), Q_2(t), \ldots, Q_c(t))$. Χάριν απλότητας, υποθέτουμε ότι η διαδικασία Q είναι Μαρκοβιανή με αριθμήσιμο σύνολο καταστάσεων $\mathcal{N} = \{n = (n_1, n_2, \ldots, n_c) \colon n_i = 0, 1, 2, \ldots, \text{ για } 1 \le i \le c\}$.

Έστω $Q(t) = n \in \mathcal{N}$. Συμβολίζουμε με $e_k = (0, 0, \ldots, 0, 1, 0, \ldots, 0)$ το διάνυσμα σειράς με c συνιστώσες, όλες από τις οποίες είναι 0 εκτός από την k-οστή, που είναι 1. Τότε, για όλα τα $i \neq j$, έχουμε:

$$Q(t+h) = \begin{cases} n - e_i + e_j, & \text{με πιθανότητα } \lambda_{ij}\phi_i(n_i)h + o(h), \\ n + e_j, & \text{με πιθανότητα } \nu_j h + o(h), \\ n - e_i, & \text{με πιθανότητα } \mu_i \phi_i(n_i)h + o(h), \end{cases}$$

όπου $\lambda_{ij}, \nu_j, \mu_i$ είναι σταθερές και ϕ_i είναι συναρτήσεις τέτοιες ώστε $\phi_i(0) = 0$ και $\phi_i(n) > 0$, για $n \ge 1$. Επιπλέον, ας υποθέσουμε ότι $\lambda_{ii} = 0$, για όλα τα i. Με άλλα λόγια, οι πελάτες κινούνται από το σταθμό εξυπηρέτησης i στο σταθμό j ($\neq i$) με ρυθμό $\lambda_{ij}\phi_i(n_i)$, οι νέες άφιξεις στον j γίνονται με ρυθμό ν_j και οι πελάτες φεύγουν από τον i με ρυθμό $\mu_i\phi_i(n_i)$. Παρατηρούμε λοιπόν ότι οι αναχωρήσεις από ένα σταθμό εξυπηρέτησης γίνονται με ρυθμούς, που εξαρτώνται από τον αριθμό των πελατών στο σταθμό αυτό.

Τα δίκτυα ουρών, που ορίζονται με τους προηγούμενους γενικούς κανόνες, συνήθως ονομάζονται *διαδικασίες διακινήσεων* ή *δίκτυα Jackson*. Επιπλέον, ένα τέτοιο δίκτυο ουρών λέγεται ότι αποτελεί μια *κλειστή διαδικασία διακινήσεων*, όταν $\nu_j = \mu_j = 0$, για όλα τα j, διότι στην περίπτωση αυτή δεν μπορεί να γινεί καμία άφιξη από ή αναχώρηση προς τον εξωτερικό κόσμο. Όταν κάποια από τα ν_j ή μ_j είναι θετικά, τότε το δίκτυο ουρών λέγεται ότι αποτελεί μια *ανοικτή διαδικασία διακινήσεων*.

Παρατήρηση 3.8. Ας υποθέσουμε ότι σε κάθε σταθμό i υπάρχουν r εξυπηρετητές και ότι οι πελάτες στο σταθμό αυτό εξυπηρετούνται σύμφωνα με την εκθετική κατανομή με παράμετρο γ_i. Φεύγοντας από το σταθμό i, οι πελάτες κατευθύνονται προς το σταθμό j ($\neq i$) με πιθανότητα p_{ij} ή βγαίνουν εξ ολοκλήρου εκτός του συνολικού δικτύου των σταθμών με πιθανότητα $q_i = 1 - \sum_{j\colon j \neq i} p_{ij}$. Αν υποτεθεί η σύνηθης ανεξαρτησία στο σύστημα, τότε η αντίστοιχη διαδικασία διακινήσεων έχει παραμέτρους $\phi_i = \min\{n, r\}, \lambda_{ij} = \gamma_i p_{ij}, \mu_i = \gamma_i q_i, \nu_j = 0$.

Ερχόμαστε τώρα να διερευνήσουμε τη στάσιμη κατανομή μιας κλειστής διαδικασίας διακινήσεων. Έτσι, υποθέτουμε εδώ ότι $\nu_j = \mu_j = 0$, για όλα τα j, και συμβολίζουμε με N το συνολικό πλήθος των πελατών, που βρίσκονται σε ολόκληρο το δίκτυο ουρών.

Πρώτα, θα ασχοληθούμε με την περίπτωση $N = 1$ και, χάριν απλότητας, θα υποθέσουμε ότι $\phi_i = 1$, για κάθε i. Όταν τώρα ο μοναδικός πελάτης βγεί από το σταθμό i, θα κατευθυνθεί στον j με ρυθμό λ_{ij}. Επομένως, η θέση του πελάτη αυτού είναι μια αλυσίδα Markov συνεχούς χρόνου με γεννήτορα $H = \{H_{ij}\}$, που δίνεται ως εξής:

$$H_{ij} = \begin{cases} \lambda_{ij}, & \text{αν } i \neq j, \\ -\sum_k \lambda_{ik}, & \text{αν } i = j. \end{cases}$$

Η αλυσίδα H έχει στάσιμη κατανομή $\alpha = (\alpha_i : i \in S)$, που ικανοποιεί τη σχέση $aH = 0$, δηλαδή, τις εξισώσεις:

$$\sum_j \alpha_j \lambda_{ji} = \alpha_i \sum_j \lambda_{ij}, \quad \text{για } i \in S,$$

εξισώσεις που θα υποθέτουμε στη συνέχεια ότι ικανοποιούνται από τις συνιστώσες του α. Επιπλέον, υποθέτοντας ότι η αλυσίδα είναι αδιαχώριστη, θα έχουμε $\alpha_i > 0$, για κάθε i (ενώ πάντα ισχύει η κανονικοποίηση $\sum_i \alpha_i = 1$).

Συμβολίζοντας με $\mathcal{N}_N$ τα διανύσματα του συνόλου $\mathcal{N}$ με άθροισμα των συνιστωσών τους ίσο προς N και θεωρώντας τα άδεια γινόμενα να είναι ίσα προς 1, παίρνουμε το παρακάτω αποτέλεσμα για τη στάσιμη κατανομή αυτής της διαδικασίας.

Θεώρημα 3.43. *Η στάσιμη κατανομή μιας αδιαχώριστης κλειστής διαδικασίας διακινήσεων N πελατών είναι*

$$\pi(n) = B_N \prod_{i=1}^{c} \left\{ \frac{\alpha_i^{n_i}}{\prod_{r=1}^{n_i} \phi_i(r)} \right\}, \quad \text{για } n \in \mathcal{N}_N,$$

όπου B_N είναι μια κατάλληλη σταθερά για την κανονικοποίηση των συνιστωσών της π.

Στη συνέχεια, θα εξετάσουμε τη γενική περίπτωση, στην οποία οι πελάτες μπορούν να μπαίνουν και να βγαίνουν από το συνολικό δίκτυο ουρών. Όπως και στην προηγούμενη περίπτωση της κλειστής διαδικασίας διακινήσεων, χρειάζεται να δούμε πρώτα τι γίνεται με μια βοηθητική διαδικασία, στην οποία συμμετέχει μόνον ένας πελάτης. Τώρα, δίπλα στους c σταθμούς εξυπηρέτησης, προσθέτουμε έναν ακόμη, που τον συμβολίζουμε με ∞, οπότε παίρνουμε μια κλειστή διαδικασία διακινήσεων στο σύνολο σταθμών $S \cup \{\infty\}$. Στη διαδικασία αυτή, ο μοναδικός πελάτης, μετά από το σταθμό i, κατευθύνεται στο σταθμό j ($\neq i$) με ρυθμό:

$$\begin{cases} \lambda_{ij}, & \text{αν } 1 \leq i, j \leq c, \\ \mu_i, & \text{αν } j = \infty, \\ \nu_j, & \text{αν } i = \infty. \end{cases}$$

Υποθέτοντας ότι η βοηθητική αυτή διαδικασία είναι αδιαχώριστη, έστω J ο γεννητόρας της και $\beta = (\beta_1, \beta_2, \ldots, \beta_c, \beta_\infty)$ η μοναδική στάσιμη κατανομή της, που ικανοποιεί τη σχέση $\beta J = 0$, δηλαδή, τις εξισώσεις:

$$\beta_\infty \nu_i + \sum_{j \in S} \beta_j \lambda_{ji} = \beta_i \bigg(\mu_i + \sum_{j \in S} \lambda_{ij} \bigg), \quad \text{για } i \in S.$$

Παρατηρούμε ότι $\beta_i > 0$, για $i \in S \cup \{\infty\}$. Θέτοντας $\alpha_i = \beta_i / \beta_\infty$, παίρνουμε ένα διάνυσμα $\alpha = (\alpha_i : i \in S)$ τέτοιο ώστε $\alpha_i > 0$, για $i \in S$, και

$$\nu_i + \sum_{j \in S} \alpha_j \lambda_{ji} = \alpha_i \bigg(\mu_i + \sum_{j \in S} \lambda_{ij} \bigg), \quad \text{για } i \in S.$$

Έστω

$$D_i = \sum_{n=0}^{\infty} \frac{\alpha_i^n}{\prod_{r=1}^{n} \phi_j(r)}.$$

Θεώρημα 3.44. *Αν η παραπάνω βοηθητική διαδικασία είναι αδιαχώριστη και $D_i < \infty$, για κάθε $i \in S$, τότε η στάσιμη κατανομή της αρχικής ανοικτής διαδικασίας διακινήσεων είναι*

$$\pi(n) = \prod_{i=1}^{c} \pi_i(n_i), \quad \text{για } n \in \mathcal{N},$$

όπου

$$\pi_i(n_i) = D_i^{-1} \frac{\alpha_i^{n_i}}{\prod_{r=1}^{n_i} \phi_i(r)}.$$