UNIVERZITET U NOVOM SADU
PRIRODNO-MATEMATIČKI
FAKULTET
DEPARTMAN ZA MATEMATIKU
I INFORMATIKU

Elvira Klebečko

Redovi čekanja i njihove primene

Master rad

Novi Sad, 2012.

UNIVERZITET U NOVOM SADU
PRIRODNO-MATEMATIČKI
FAKULTET
DEPARTMAN ZA MATEMATIKU
I INFORMATIKU

Elvira Klebečko

Redovi čekanja i njihove primene

Master rad

Mentor: Prof. dr Danijela Rajter-Ćirić

Novi Sad, 2012.

Sadržaj

Predgovor

Čekanje u redu je sveprisutna pojava. Ovaj problem se pojavljuje u praksi kad određeni broj jedinica, bilo ljudi ili predmeta, koji traže odgovarajuću uslugu ili obradu moraju čekati, tj. provesti izvesno vreme u redu čekanja pre nego što budu usluženi. Čekanje u redu predstavlja stajanje u redu, a procesi čekanja u redu su stohastički procesi koji proizilaze iz fenomena stajanja u redu. Na primer, procesuiranje podataka u računarskom centru, usluživanje kupaca u supermarketu, kao i protok poslova kroz radionicu – sve to uključuje čekanje u redu.

Osnovni cilj ovog rada je da predstavi proces masovnog posluživanja i da ukaže na probleme na koje nailazimo prilikom proučavanja čekanja u redu.

U prvom poglavlju obrađene su osnovni pojmovi teorije verovatnoće i stohastičke analize. U ovom delu rada posebna pažnja je posvećena Poasonovoj i eksponencijalnoj raspodeli, sa akcentom na blisku povezanost Poasonovog procesa sa pomenutim raspodelama.

U drugom poglavlju razmatrano je nekoliko sistema čekanja u redu. Pretpostavljeno je da su procesi pristizanja u sisteme masovnog usluživanja Poasonovi procesi, kao i da su vremena usluživanja eksponencijalno raspodeljena. Detaljno je prikazan opšti model masovnog usluživanja po principu rađanja i umiranja. U nastavku, predstavljene su neke od široko primenjenih modela masovnog posluživanja, pre svega modeli sa jednim serverom, tj. $M/M/1$ i $M/M/1/c$ model, kao i sistemi sa više servera, odnosno $M/M/s/\infty$, $M/M/s/c$, $M/M/s/s$ i $M/M/\infty$ model, gde s i c predstavljaju broj servera i kapacitet sistema, redom. Za svaki od ovih modela predstavljena su osnovna svojstva i osobine, a većina je ilustrovana primerima. Na kraju, je urađena simulacija, gde smo kao primer posmatrali čekanje u redu u pošti.

Ovom prilikom se zahvaljujem svim profesorima i asistentima na ukazanom znanju tokom studiranja. Izuzetnu zahvalnost dugujem svom mentoru, prof. dr Danijeli Rajter-Ćirić, na sugestijama, savetima, strpljenju i pomoći prilikom izrade ovog master rada, kao i na zanimljivim predavanjima tokom studiranja koja su bila jedan od razloga da se opredelim za temu iz ove oblasti. Želela bih da se zahvalim svojoj porodici, prijateljima i svima koji su mi na bilo koji način pružili pomoć i podršku tokom studiranja i prilikom izrade ovog rada.

Novi Sad, oktobar 2012. Elvira Klebečko

Glava 1

1. Uvodni pojmovi

U ovoj glavi izložićemo osnovne pojmove teorije verovatnoće i osnovne pojmove stohastičkih procesa. S obzirom da ćemo u ovom radu posmatrati sisteme masovnog posluživanja, gde su pristizanja Poasonovi procesi, a vremena posluživanja eksponencijalna, poznavanje ove teorije je neophodno za razumevanje rada.

1.1 Osnovni pojmovi teorije verovatnoće

Teorija verovatnoće bavi se matematičkim modelima čiji ishodi zavise od slučaja. Skup svih mogućih ishoda označimo sa Ω. Elemente skupa Ω nazivamo elementarnim događajima i označavamo sa ω. Skup $A \subseteq \Omega$ nazivamo slučajnim događajem, dok sa $\bar{A}$ označavamo događaj suprotan događaju A. Skup Ω nazivamo sigurnim događajem, a $\emptyset$ je nemoguć događaj.

Definicija 1. Podskup $\mathcal{F}$ partitivnog skupa $\mathcal{P}(\Omega)$ je σ - *algebra* (σ - *polje*) nad Ω ako su zadovoljeni uslovi:

1. $\Omega \in \mathcal{F}$;
2. ako A $\in \mathcal{F}$, onda $\bar{A} \in \mathcal{F}$;
3. ako $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{F}$, onda $\bigcup_{i\in\mathbb{N}} A_i \in \mathcal{F}$.

Definicija 2. (*Aksioma verovatnoće*) Neka je Ω skup elementarnih događaja i $\mathcal{F}$ σ – algebra nad Ω. Funkcija $P : \mathcal{F} \rightarrow [0,1]$ se zove *verovatnoća* na prostoru $(\Omega, \mathcal{F})$ ako zadovoljava uslove:

1. $P(\Omega) = 1$;

2. Ako $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{F}$, $A_i \cap A_j = \emptyset$, $i \neq j$, $i, j = 1, 2, \ldots$, onda

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Prostor verovatnoća je uređena trojka $(\Omega, \mathcal{F}, P)$.

Definicija 3. Neka je dat prostor verovatnoća $(\Omega, \mathcal{F}, P)$ i neka su dati događaji $A, B \in \mathcal{F}$, pri čemu je $P(B) > 0$. *Uslovna verovatnoća* događaja A, pod uslovom realizacije događaja B, u oznaci $P(A|B)$, definiše se sa

$$P(A|B) = \frac{P(AB)}{P(B)}.$$

Drugi izraz za uslovnu verovatnoću je

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}.$$

Definicija 4. Familija $A_1, A_2, \ldots$ događaja iz $\mathcal{F}$ je *nezavisna* ako je

$$P\left(A_{k_1} \ldots A_{k_n}\right) = P\left(A_{k_1}\right) \cdots P\left(A_{k_n}\right),$$

za svaki konačni niz indeksa $k_1, \ldots, k_n$, takvih da je $k_1 < \ldots < k_n$.

Dva osnovna tipa slučajnih promenljivih su: *diskretne*, kod kojih možemo slučajnoj promenljivoj dodeliti konkretne vrednosti sa određenim verovatnoćama, i *apsolutno neprekidne*, kod kojih to nije moguće, tj. slučajna promenljiva kao vrednosti ima sve tačke nekog (konačnog ili beskonačnog) intervala.

Definicija 5. Skup vrednosti diskretne slučajne promenljive $\{x_1, x_2, \ldots\}$, zajedno sa odgovarajućim verovatnoćama $p(x_i) = P\{X = x_i\}$, $i = 1, 2, \ldots$, predstavlja *zakon raspodele verovatnoća* (ili kraće, *raspodela*) slučajne promenljive X. Zakon raspodele obično se predstavlja u obliku šeme

$$X : \begin{pmatrix} x_1 & x_2 & \cdots \\ p(x_1) & p(x_2) & \ldots \end{pmatrix}.$$

Definicija 6. *Funkcija raspodele* $F_X(x) : \mathbb{R} \to [0, 1]$ slučajne promenljive X je definisana kao

$$F_X(x) = P\{X < x\}.$$

Funkcija raspodele slučajne promenljive X pokazuje koliko je verovatno da X uzima vrednosti manje od x, i ona postoji i jedinstvena je za svaku slučajnu promenljivu.

Za apsolutno neprekidne slučajne promenljive, gustina raspodele slučajne promenljive se definiše na sledeći način.

Definicija 7. Slučajna promenljiva X je apsolutno neprekidnog tipa ako postoji nenegativna integrabilna funkcija $\varphi_X(x)$, $-\infty < x < \infty$, takva da je za svaki skup $S \in \mathcal{B}(\mathbb{R})$,

$$P\{X \in S\} = \int_S \varphi_X(x)\,dx.$$

Funkcija $\varphi_X(x)$ zove se *gustina raspodele verovatnoća* (*gustina raspodele*) slučajne promenljive X.

Definicija 8. Veza između funkcija raspodele i gustine slučajne promenljive X je data sa

$$F_X(x) = \int_{-\infty}^{x} \varphi_X(t)dt, \quad za\ svako\ x \in \mathbb{R}.$$

U svakoj tački neprekidnosti gustine $\varphi_X(x)$ važi

$$\varphi_X(x) = F_X'(x).$$

Definicija 9. *Matematičko očekivanje* $E(X)$ *diskretne slučajne promenljive* X sa raspodelom $p(x_k), k = 1, 2, \dots$, definiše se sa

$$E(X) := \sum_{k=1}^{\infty} x_k\, p(x_k),$$

i postoji ako i samo ako

$$\sum_{k=1}^{\infty} |x_k|\, p(x_k) < \infty.$$

Definicija 10. *Očekivanje apsolutno neprekidne slučajne promenljive* X sa gustinom $\varphi_X(x)$ je

$$E(X) = \int_{-\infty}^{\infty} x\varphi_X(x)dx,$$

i ono postoji ako gornji integral *apsolutno konvergira*, odnosno, ako

$$\int_{-\infty}^{\infty} |x| \varphi_X(x) dx < \infty.$$

Drugim rečima, $E(X)$ postoji ako i samo ako $E(|X|)$ postoji. U tom slučaju možemo imati predstavu o matematičkom očekivanju kao srednjoj vrednosti slučajne promenljive.
Vidimo da matematičko očekivanje ne mora da postoji za svaku slučajnu promenljivu.

Definicija 11. Ako slučajne promenljive $X_1, \ldots, X_n, n \in \mathbb{N}$, imaju očekivanja onda

$$E\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} E(X_k).$$

Definicija 12. *Momenat reda* $k, k \in \mathbb{N}$, slučajne promenljive X je $E(X^k)$. *Centralni momenat reda* $k, k \in \mathbb{N}$, slučajne promenljive X je $E\left(\left(X - E(X)\right)^k\right)$.
Ako slučajna promenljiva X ima momenat (centralni momenat) reda $n, n \geq 2$, onda ona ima i sve momente (centralne momente) reda $1, 2, \ldots, n-1$.

Matematičko očekivanje ne pruža dovoljno informacija o "rasturanju" vrednosti slučajne promenljive X oko $E(X)$. Zbog toga se sledećom definicijom uvodi još jedna važna numerička karakteristika slučajne promenljive.

Definicija 13. *Centralni momenat drugog reda* slučajne promenljive X zove se *disperzija* (*varijansa*) slučajne promenljive X i ona predstavlja očekivanu vrednost kvadrata odstupanja od srednje vrednosti. Označava se sa $D(X)$ ili $\sigma^2(X)$. Dakle,

$$D(X) = E\left(\left(X - E(X)\right)^2\right) = E(X^2) - E^2(X).$$

Definicija 14. *Standardna devijacija* (*standardno odstupanje, prosečno odstupanje*) slučajne promenljive X se definiše sa

$$\sigma(X) = \sqrt{D(X)},$$

gde je $D(X) \geq 0$.

Definicija 15. *Kovarijansa* slučajne promenljive X i Y je

$$cov(X, Y) = E\left[\left(X - E(X)\right)\left(Y - E(Y)\right)\right] = E(XY) - E(X)E(Y).$$

Ako je $cov(X, Y) = 0$, kažemo da su slučajne promenljive X i Y nekorelirane (nepovezane). Iz definicije vidimo da su nezavisne slučajne promenljive ujedno i nekorelirane.

1.2 Eksponencijalna raspodela

Definicija 16. Neprekidna slučajna promenljiva X ima *eksponencijalnu raspodelu* $\mathcal{E}(\lambda)$ sa parametrom $\lambda > 0$ ako je njena *funkcija gustine* data sa

$$\varphi_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0 \quad, & x < 0 \end{cases}$$

ili ekvivalentno, ako je *funkcija raspodele* data sa

$$F_X(x) = \int_{-\infty}^{x} \varphi_X(t)dt = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0 \quad, & x < 0 \end{cases}.$$

Očekivanje slučajne promenljive $X : \mathcal{E}(\lambda)$ je

$$E(X) = \int_{-\infty}^{\infty} x\varphi_X(x)dx = \int_{0}^{\infty} x\lambda e^{-\lambda x}\, dx = \frac{1}{\lambda}.$$

Disperzija slučajne promenljive $X : \mathcal{E}(\lambda)$ je

$$D(X) = E(X^2) - E^2(X) = \int_{0}^{\infty} x^2\lambda e^{-\lambda x}\, dx - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Eksponencijalna raspodela se često koristi u raznim primenama, na primer, pri analizi pouzdanosti rada sistema, kao model za vreme između dva kvara. U ovim situacijama recipročna vrednost parametra λ se javlja kao mera prosečnog vremena rada uređaja koji se ispituje.

Definicija 17. Neka je X slučajna promenljiva sa eksponencijalnom raspodelom $\mathcal{E}(\lambda)$. Za svako $x \geq 0$ važi

$$P\{X > x\} = 1 - P\{X \leq x\} = 1 - \int_{0}^{x} \lambda e^{-\lambda t}\, dt = e^{-\lambda x}.$$

Koristeći definiciju uslovne verovatnoće, sledi da je

$$P\{X > s + t | X > s\} = \frac{P\{X > s + t, X > s\}}{P\{X > s\}} = \frac{P\{X > s + t\}}{P\{X > s\}} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P\{X > t\}$$

ili ekvivalentno

$$P\{X > s + t\} = P\{X > s\}P\{X > t\}.$$

Jednakost $P\{X > s + t | X > s\} = P\{X > t\}$ izražava osobinu odsustva memorije. Sve dok eksponencijalno raspodeljena slučajna promenljiva X zadovoljava ovu jednačinu važi da je ona "*bez memorije*" (tj. ne pamti šta se ranije desilo). Naime, ako slučajna promenljiva X predstavlja dužinu rada nekog uređaja bez kvara, tada nejednakost $X > s$ znači da je uređaj ispravan posle s sati rada. Prethodna jednakost izražava činjenicu da je verovatnoća da uređaj ispravno radi još bar t sati jednaka verovatnoći da je uređaj ispravan posle t sati od uključenja. Drugim rečima, kao da uređaj "ne zna" da je pre toga radio s sati.

Napomena 1. Jedino eksponencijalna raspodela poseduje ovu osobinu.

1.3 Poasonova raspodela

Ova raspodela je granični slučaj binomne raspodele pod uslovom da je broj opita n veliki, a verovatnoća p pojave događaja A u svakom pojedinačnom opitu mala.

Raspodela verovatnoća Bernulijeve slučajne promenljive S_n je

$$p(k) = P\{S_n = k\} = \binom{n}{k} p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n.$$

Teorema 1.1. *Ako u Bernulijevoj šemi broj nezavisnih ispitivanja neograničeno raste, a verovatnoća u svakom ispitivanju opada tako da je* $np_n \to \lambda = const > 0$, *gde je* p_n *verovatnoća pojavljivanja događaja* A *u* n *ponavljanja eksperimenta, tada*

$$P\{S_n = k\} \to \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, \dots; \; n \to \infty.$$

Definicija 18. Slučajna promenljiva S_∞, sa skupom vrednosti $\{0, 1, 2, \dots\}$ i raspodelom verovatnoća

$$P\{S_\infty = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, \qquad \lambda > 0$$

naziva se *Poasonova*[1] *slučajna promenljiva*, pišemo $X : \mathcal{P}(\lambda)$.

Poasonova raspodela vezana je za pojavu retkih događaja i ima široku primenu u telefoniji, saobraćaju, demografiji, biologiji, fizici, itd. Koristi se kao model za broj događaja koji se

[1] Ova raspodela nosi naziv po francuskom matematičaru Simeo Denis Poisson-u (1781-1840).

dešavaju u jedinici vremena, pri čemu parametar λ predstavlja srednju vrednost broja ovih događaja.

Očekivanje i *disperzija Poasonove slučajne promenljive* imaju istu vrednost

$$E(S_\infty) = D(S_\infty) = \lambda = n \cdot p.$$

1.4 Stohastički procesi

Slučajni (*stohastički*) *procesi* predstavljaju matematičke modele procesa čija je evolucija opisana zakonima verovatnoće. Teorija slučajnih procesa svoju primenu nalazi u raznovrsnim disciplinama kao što su finansijska matematika, telekomunikacije, teorija pouzdanosti, teorija opsluživanja i upravljanja, računarskim naukama i u mnogim drugim.

1.4.1 Pojam stohastičkog procesa

Zamislimo da se u svakom vremenskom trenutku t vremenskog intervala I posmatra neka karakteristika X nekog fizičkog sistema koja je slučajnog karaktera. Dakle, $X(t)$ je neka slučajna promenljiva za svako $t \in I$. To znači da na skup svih slučajnih promenljivih $\{X(t), t \in I\}$ možemo gledati kao na slučajnu veličinu koja se menja u vremenu, odnosno dobijamo jednu slučajnu funkciju vremena. U tom slučaju kažemo da je $\{X(t), t \in I\}$ jedan slučajni (stohastički) proces.

Definicija 19. *Stohastički proces* $\{X(t), t \in I\}$ je familija slučajnih promenljivih definisanih na istom prostoru verovatnoća $(\Omega, \mathcal{F}, P)$. Skup I zove se *parametarski skup*, a realni prostor $\mathbb{R}^d$ *skup stanja* procesa ($X : \Omega \to \mathbb{R}^d$).

Ako je parametarski skup I *konačan*, onda imamo konačno mnogo slučajnih promenljivih. Ako je I *prebrojiv*, govorimo o nizu, lancu slučajnih promenljivih. Izraz "proces" se najčešće koristi u slučaju kada parametarski skup *nije prebrojiv*.

Parametar $t \in I$ se u slučaju da je $I \subset \mathbb{R}$ obično interpretira kao vreme. Ukoliko je I diskretan podskup, tada imamo slučajni proces sa *diskretnim vremenom*, a ukoliko je I neprebrojiv skup, imamo slučajni proces sa *neprekidnim vremenom*.

Podrazumevaćemo da za svako $t \in I$ sve slučajne promenljive $X(t)$ uzimaju vrednosti iz skupa S koji ćemo zvati skup stanja.

Stohastički proces $\{X(t)\}_{t\in I} = \{X(t), t \in [t_0, T]\} = \{X(t,\omega), t \in [t_0, T], \omega \in \Omega\}$ je funkcija dva parametra, t i ω, ali se promenljiva ω najčešće izostavlja iz zapisa, pa stoga stohastički proces označavamo sa $X(t)$ ili X_t.

- Ako se fiksira $t \in [t_0, T]$, dobija se jedna slučajna promenljiva koja se zove *zasek* ili *sečenje* stohastičkog procesa $\{X(t)\}_{t\in I}$ u trenutku t.
- Ako se fiksira $\omega \in \Omega$, dobija se jedna realna funkcija definisana na intervalu $[t_0, T]$ i ta funkcija se zove *trajektorija* (*realizacija*, *staza*) stohastičkog procesa $\{X(t)\}_{t\in I}$.

Definicija 20. *Konačno-dimenzionalne raspodele* stohastičkog procesa $\{X(t), t \in [t_0, T]\}$ su date sa

$$\begin{aligned} F_{t_1}(x) = &\quad P\{X(t_1) < x\}, \\ F_{t_1,t_2}(x_1, x_2) = &\quad P\{X(t_1) < x_1, X(t_2) < x_2\}, \\ &\vdots \\ F_{t_1,\dots,t_n}(x_1, \dots, x_n) = &\quad P\{X(t_1) < x_1, \dots, X(t_n) < x_n\}, \\ &\vdots \end{aligned}$$

gde t, $t_i \in [t_0, T]$ i $x, x_i \in \mathbb{R}^d, i = 1, 2, \dots$.

Neka svojstva stohastičkog procesa:

- *Očekivanje* (*srednja vrednost*) stohastičkog procesa $X(t)$ je

$$m_X(t) = m(t) = E[X(t)],$$

- *Autokovarijansna funkcija* stohastičkog procesa $X(t)$ je

$$K_X(t,s) = K(t,s) = E\big[\big(X(t) - m(t)\big)\big(X(s) - m(s)\big)\big] = E[X(t)X(s)] - m(t)m(s),$$

- *Disperzija* stohastičkog procesa $X(t)$ je

$$D_X(t) = D(t) = K(t,t) = E[X^2(t)] - m^2(t).$$

Definicija 21. Za proces $\{X(t), t \in [t_0, T]\}$ za koji važi da su slučajne promenljive (tzv. priraštaji) $X(t_0)$, $X(t_1) - X(t_0)$, $X(t_2) - X(t_1), \dots$, $X(t_n) - X(t_{n-1}), \dots$ nezavisni za sve $t_0 < t_1 < \cdots < t_n < \cdots$, gde $t_0, t_1, \dots \in [t_0, T]$, kažemo da je *stohastički proces sa nezavisnim priraštajima*.

Definicija 22. Ako slučajne promenljive $X(t_2 + h) - X(t_1 + h)$ i $X(t_2) - X(t_1)$ imaju istu funkciju raspodele za svako $t_1 < t_2$ i h, tada je stohastički proces $\{X(t), t \in [t_0, T]\}$ *proces sa stacionarnim priraštajima*.

1.4.2 Lanci Markova

Lanci Markova predstavljaju korisne alate u statističkom modelovanju u praktično svim poljima primenjene matematike i imaju veliku primenu u opisivanju ponašanja sistema. Oni igraju glavnu ulogu u teoriji čekanja u redu.

Imati svojstvo Markova znači da pored datog trenutnog stanja, buduće stanje sistema ne zavisi od prošlih stanja. Drugim rečima, to znači, da opis sadašnjosti u potpunosti sadrži informaciju koja može uticati na buduće stanje procesa.

Definicija 23. Za stohastički proces $\{X(t), t \in I\}$ kažemo da je *proces Markova* ako za svaki događaj iz skupa A i za svaki vremenski trenutak $t_n < t_{n+1}$ važi

$$P\{X(t_{n+1}) \in A |\, X(t) = x_t, t \le t_n\} = P\{X(t_{n+1}) \in A | X(t_n) = x_{t_n}\}.$$

Prema tome, verovatnoća da će proces preći iz stanja x_{t_n} u kojem se nalazi u trenutku t_n, u neko drugo stanje iz skupa A, u trenutku t_{n+1}, ne zavisi od načina na koji je proces dospeo u stanje x_{t_n} iz stanja x_{t_0} u kojem se proces nalazio u početnom trenutku t_0.

Lanci Markova su posebna vrsta procesa Markova, gde se proces može nalaziti samo u konačnom broju stanja.

Razmotrićemo dva slučaja: slučaj kad je proces $\{X(t), t \in I\}$ sa *diskretnim vremenom* i slučaj sa *neprekidnim vremenom*, koji može uzimati samo konačno mnogo različitih vrednosti.

1.4.2.1 Diskretni slučaj lanaca Markova

U ovom slučaju stohastički proces je oblika $\{X_n, n \in \mathbb{N}_0\}$ i ima prebrojiv skup mogućih vrednosti $\{x_0, x_1, \dots\}$, koji se zove *skup stanja*. Posmatraćemo sistem koji se u vremenskim trenucima $t_0, t_1, \dots$ može nalaziti u nekom od mogućih n stanja. Ako se u trenutku t_k sistem nalazi u stanju i, pišemo $X_k = x_i$.

Definicija 24. Niz slučajnih promenljivih $X_0, X_1, \dots$ sa prebrojivim skupom mogućih vrednosti (skupom stanja) $S = \{x_0, x_1, x_2, \dots\}$ zove se *lanac Markova* ako važi

$$P\{X_{n+1} = x_{n+1} |\, X_0 = x_0, X_1 = x_1, \dots, X_n = x_n\} = P\{X_{n+1} = x_{n+1} |\, X_n = x_n\}.$$

Definicija 25. *Verovatnoća prelaza u jednom koraku* iz i-tog u j-to (susedno) stanje je

$$p_{i,j}^{n,n+1} = P\{X_{n+1} = x_j | X_n = x_i\}, \qquad n \ge 0.$$

Ako verovatnoća $p_{i,j}^{n,n+1}$ ne zavisi od broja koraka $n \geq 0$ (vremenskih trenutaka), tada je *lanac homogen* (*stacionaran*) i pišemo

$$p_{i,j} = P\{X_{n+1} = x_j | X_n = x_i\}.$$

Matricu $\boldsymbol{P} = \boldsymbol{P}_1 = [p_{i,j}]_{i,j}$ nazivamo *matricom prelaza za jedan korak.*

Analogno definišemo verovatnoću prelaza u n koraka iz stanje i u stanje j.

Definicija 26. *Verovatnoća prelaza u n koraka* iz i-tog u j-to stanje je

$$p_{i,j}(n) = P\{X_{k+n} = x_j | X_k = x_i\}, \qquad k, n, i, j \geq 0.$$

Matrica prelaza za n koraka je

$$\boldsymbol{P}_n = \boldsymbol{P}(n) = [p_{i,j}(n)]_{i,j}.$$

Zbir po vrstama u matrici prelaza treba da je uvek jednaka 1.

Izračunavanje ovih verovatnoća prelaza u n koraka omogućava nam jednačina Čepmen – Kolmogorova.

Teorema 1.2. *Jednačine Čepmen – Kolmogorova (1)*

$$p_{i,j}(n+m) = \sum_{k=0}^{\infty} p_{i,k}(n) p_{k,j}(m),$$

za svako $i, j, m, n \geq 0$.

U matričnom zapisu ova jednačina ima sledeći oblik

$$\boldsymbol{P}_{n+m} = \boldsymbol{P}_n \cdot \boldsymbol{P}_m.$$

Ako uvrstimo $n = m = 1$, tada dobijamo

$$\boldsymbol{P}_2 = \boldsymbol{P}_1 \cdot \boldsymbol{P}_1 = \boldsymbol{P} \cdot \boldsymbol{P} = \boldsymbol{P}^2.$$

Dalje,

$$\boldsymbol{P}_3 = \boldsymbol{P}_2 \cdot \boldsymbol{P}_1 = \boldsymbol{P}^2 \cdot \boldsymbol{P} = \boldsymbol{P}^3,$$

$$\boldsymbol{P}_4 = \boldsymbol{P}_3 \cdot \boldsymbol{P}_1 = \boldsymbol{P}^3 \cdot \boldsymbol{P} = \boldsymbol{P}^4,$$

$$\vdots$$

$$\boldsymbol{P}_n = \boldsymbol{P}^n.$$

Definicija 27. Lanac Markova čiji je skup stanja $\{\ldots,-2,-1,0,1,2,\ldots.\}$ zove se *slučajan hod* ako za neko $p \in (0,1)$ i za $i \in \mathbb{Z}$ važi

$$p_{i,i+1} = p = 1 - p_{i,i-1}.$$

Definicija 28. Stanje x_j je *dostižno* iz stanja x_i ako $\exists n_0 \in \mathbb{N}$ tako da važi $p_{i,j}(n_0) > 0$.

Definicija 29. Ako je stanje x_i dostižno iz stanja x_j i ako je stanje x_j dostižno iz stanja x_i, tada kažemo da stanja x_i i x_j *komuniciraju*.

Definicija 30. Ako stanja x_i i x_j komuniciraju, tada su ona u istoj *klasi*.

Definicija 31. Neka je $\{X_n, n = 0,1,\ldots\}$ lanac Markova i neka je C podskup skupa stanja. Kažemo da je C *zatvoren skup* ako proces uvek ostaje u C, tj. ako važi

$$P\{X_{n+1} \in C | X_n = x_i \in C\} = 1 \ \ za\ svako\ x_i \in C.$$

Definicija 32. Lanac Markova je *nesvodljiv* ako sva stanja komuniciraju međusobno, odnosno kada skup stanja ne sadrži zatvoren podskup odvojeno od skupa svih stanja. Stanje koje sam po sebi formira nesvodljiv skup zove se *apsorbujuće stanje* (tj. stanje x_j je apsorbujuće ako je $p_{j,j} = 1$).

Definicija 33. Stanje x_j je *povratno* ako je verovatnoća da se sistem iz x_j bar jednom vrati u x_j jednaka 1.

Definicija 34. Ako postoji $n_0 \in \mathbb{N}$ takvo da matrica $\boldsymbol{P}_{n_0} = \boldsymbol{P}^{n_0}$ ima sve elemente strogo pozitivne, tada odgovarajući lanac Markova je *ergodičan*.

Teorema 1.3. *Ako je lanac Markova ergodičan i nesvodljiv, za svako i postoji* $p_j^* = lim_{n\to\infty} p_{i,j}(n)$ *i ove verovatnoće se zovu finalne (granične) verovatnoće (iz bilo kog stanja* i *se posle dovoljno dugo vremena prelazi u stanje* j *sa verovatnoćom* p_j^*).

Finalne verovatnoće računamo iz sistema:

$$\boldsymbol{p}^* = \boldsymbol{p}^*\boldsymbol{P},$$

$$\sum_{j=0}^{\infty} p_j^* = 1,$$

gde je $\boldsymbol{p}^* := (p_0^*, p_1^*, \ldots)$.

Može se pokazati da prethodni sistem ima jedinstveno pozitivno rešenje.

Napomena 2. Ako je skup stanja konačan i obuhvata k stanja, tada imamo k jednačina sledećeg oblika:

$$p_j^* = \sum_{i=0}^{k-1} p_i^* p_{i,j},$$

što zajedno sa uslovom $\sum_{j=0}^{k-1} p_j^* = 1$ daje $k+1$ jednačinu sa k nepoznatih.

1.4.2.2 Neprekidni slučaj lanaca Markova

Kod diskretnog slučaja lanaca Markova, vreme koje sistem provede u datom stanju je deterministički određeno. Kod neprekidnog slučaja lanaca Markova, za razliku od diskretnog, vreme koje sistem provede u datom stanju je slučajna promenljiva koja ima eksponencijalnu raspodelu.

Definicija 35. Neka je $\{X(t), t \geq 0\}$ stohastički proces sa neprekidnim vremenom i sa skupom stanja $\mathbb{N}_0 = \{0, 1, 2, \dots\}$. Kažemo da je $\{X(t), t \geq 0\}$ *neprekidni slučaj lanaca Markova* ako važi

$$P\{X(t+s) = x_j | X(s) = x_i, X(r) = x_r, 0 \leq r < s\} =$$

$$= P\{X(t+s) = x_j | X(s) = x_i\} = p_{i,j}(t)$$

za svako $s, t \geq 0$ i za svako $x_r \in \mathbb{N}_0$.

Teorema 1.4. *Jednačine Čepmen – Kolmogorova (2)*

$$p_{i,j}(t+s) = \sum_{k=0}^{\infty} p_{i,k}(t) p_{k,j}(s) = \sum_{k=0}^{\infty} p_{i,k}(s) p_{k,j}(t),$$

za svako $s, t \geq 0$.

Definicija 36. Veličine

$$\nu_{i,j} := \nu_i p_{i,j} \quad za\ svako\ i \neq j \in \{0, 1, \dots\} \tag{1.1}$$

nazivaju se *infinitezimalni parametri* ili *trenutne stope prelaza* za neprekidan slučaj lanaca Markova.

Napomena 3. Imamo da je

$$\sum_{j \neq i} \nu_{i,j} = \nu_i \sum_{j \neq i} p_{i,j} = \nu_i \quad (jer\ p_{i,i} = 0).$$

Ako označimo da je $\nu_{i,i} = -\nu_i$, sledi

$$\sum_{j=0}^{\infty} \nu_{i,j} = 0.$$

Definicija 37. Matrica

$$A = \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \end{array} \begin{bmatrix} \nu_{0,0} & \nu_{0,1} & \nu_{0,2} & \cdots \\ \nu_{1,0} & \nu_{1,1} & \nu_{1,2} & \cdots \\ \nu_{2,0} & \nu_{2,1} & \nu_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

se naziva *generator matrica* za neprekidan slučaj lanaca Markova $\{X(t), t \geq 0\}$.

Napomena 4. Ako znamo veličine $\nu_{i,j}$ za svako $i \neq j$, tada možemo da izračunamo stope ν_i i verovatnoće $p_{i,j}$ iz (1.1)

Napomena 5. Matrica A odgovara matrici prelaza P za diskretni slučaj lanaca Markova.

Sada ćemo navesti dva sistema diferencijalnih jednačina za određivanje verovatnoća $p_{i,j}$.

Teorema 1.4. *Verovatnoća da neprekidni slučaj lanaca Markova* $\{X(t), t \geq 0\}$ *napravi dva ili više prelaza u vrementskom intervalu dužine* δ *je jednaka* $o(\delta)$.

Teorema 1.5. (*Kolmogorova backward jednačina*) *Za svako stanje* $i, j \in \mathbb{N}_0$ *i za svako* $t \geq 0$, *važi*

$$p_{i,j}^{'}(t) = \sum_{k=0}^{\infty} \nu_{i,k} p_{k,j}(t).$$

Teorema 1.6. (*Kolmogorova forward jednačina*) *Za svako stanje* $i, j \in \mathbb{N}_0$ *i za svako* $t \geq 0$, *važi*

$$p_{i,j}^{'}(t) = \sum_{k=0}^{\infty} p_{i,k}(t) \nu_{k,j}.$$

1.4.3 Poasonov proces

Poasonov proces predstavlja stohastički proces u kojem se događaji dešavaju neprekidno i nezavisno jedan od drugog. Poasonov proces je neprekidan slučaj lanaca Markova. Ima primenu u modeliranju tzv. retkih događaja, odnosno, događaja koji su takvi da se u kratkim vremenskim intervalima može dogoditi najviše jedan takav događaj. U realne događaje koji se mogu opisivati Poasonovim procesom spadaju broj telefonskih poziva u centrali, broj dolazaka mušterija u samoposlugu, broj zahteva koje korisnik uputi nekom sistemu, … .

1.4.3.1 Procesi prebrajanja

Definicija 38. Stohastički proces $\{X(t), t \geq 0\}$ je *proces prebrajanja* ako $X(t)$ predstavlja ukupan broj događaja koji se dese do trenutka t, uključujući i t.

Osobine procesa prebrajanja:

1) $X(t) \geq 0$
2) Za fiksirano t slučajna promenljiva $X(t)$ uzima vrednosti iz $\mathbb{N}_0$.
3) Ako je $s < t$, onda $X(s) \leq X(t)$.
4) Za $s < t$, $X(t) - X(s)$ predstavlja broj događaja koji se dese u $(s, t]$.

Proces prebrajanja ima *nezavisne priraštaje* ako je broj događaja koji se dese u različitim vremenskim intervalima nezavisan.

Proces prebrajanja ima *stacionarne priraštaje* ako raspodela broja događaja koji se dese u proizvoljnom vremenskom intervalu zavisi samo od dužine vremenskog intervala, a ne od pozicije tog intervala na vremenskoj osi.

1.4.3.2 Pojam (homogenog) Poasonovog procesa

Definicija 39. Proces prebrajanja $\{X(t), t \geq 0\}$ se zove *Poasonov proces* sa stopom rasta $\lambda > 0$, ako su zadovoljeni sledeći uslovi:

1) $X(0) = 0$
2) Proces $X(t)$ ima nezavisne priraštaje.
3) Broj događaja u proizvoljnom intervalu dužine t ima Poasonovu raspodelu sa parametrom λt, tj.

$$P\{X(t+s) - X(s) = n\} = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \quad za\ svako\ s, t \geq 0\ i\ n = 0, 1, \dots .$$

Iz uslova 3) se vidi da Poasonov proces ima stacionarne priraštaje, tj. Poasonova raspodela zavisi samo od t, a ne i od s (raspodela broja događaja koji se pojavljuju u bilo kom intervalu t zavisi samo od dužine intervala t, a ne i od pozicije tog intervala na vremenskoj osi).

Druga definicija za Poasonov proces je data u nastavku i ona je ekvivalentna sa prethodnom definicijom.

Definicija 40. Proces prebrajanja $\{X(t), t \geq 0\}$ je Poasonov proces sa stopom rasta $\lambda > 0$, ako su ispunjeni sledeći uslovi:

1) $X(0) = 0$
2) Proces $X(t)$ ima stacionarne i nezavisne priraštaje.
3) $P\{X(t+h) - X(t) = 1\} = \lambda h + o(h)$,
4) $P\{X(t+h) - X(t) \geq 2\} = o(h)$,

pri čemo važi $\lim_{h\to 0} \frac{o(h)}{h} = 0$.

S obzirom da je Poasonov proces neprekidan slučaj lanaca Markova, možemo tvrditi da vreme τ_i koje proces provede u stanju $i \in \{0, 1, ..\}$ ima eksponencijalnu raspodelu sa parametrom $\lambda_i > 0$, i da su slučajne promenljive $\tau_0, \tau_1, \ldots$ nezavisne. Pošto Poasonov proces ima nezavisne i stacionarne priraštaje, sa tačke gledišta verovatnoće, proces može iznova da počne u bilo kom vremenskom trenutku. Sledi da svi τ_i, za $i = 0, 1, ...$, imaju istu raspodelu. Sada ćemo naći zajednički parametar slučajnih promenljivih τ_i.
Važi da je

$$P\{\tau_0 > t\} = P\{X(t) = 0\} = e^{-\lambda t}.$$

Sledi da je funkcija raspodele slučajne promenljive τ_0 dato sa

$$P\{\tau_0 \leq t\} = 1 - e^{-\lambda t},$$

a funkcija gustine sa

$$\varphi_{\tau_0}(t) = \left(1 - e^{-\lambda t}\right)' = \lambda e^{-\lambda t} \quad za\ t \geq 0,$$

odnosno možemo zaključiti da $\tau_0 : \mathcal{E}(\lambda)$.

Teorema 1.7. *Neka je $\{X(t), t \geq 0\}$ Poasonov proces sa stopom rasta λ, a τ_i vreme koje proces provede u stanju i, za $i = 0, 1, \ldots$. Slučajne promenljive τ_i, $i = 0, 1, \ldots$, su nezavisne sa eksponencijalnom raspodelom sa parametrom λ.*

Pre nego što navedemo sledeću lemu, definišimo gama raspodelu.

Definicija 41. Neprekidna slučajna promenljiva X sa funkcijom gustine

$$\varphi(x) = \begin{cases} \dfrac{\lambda e^{-\lambda x}\,(\lambda x)^{\alpha-1}}{\Gamma(\alpha)}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

se zove *gama slučajna promenljiva* sa parametrima $\lambda > 0$ i $\alpha > 0$.

$\Gamma(\alpha)$ se zove *gama funkcija* i definisana je na sledeći način

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\, dx.$$

Za $\alpha = n$, može se pokazati da je $\Gamma(n) = (n-1)!$, a gustina je data sa

$$\varphi(x) = \begin{cases} \dfrac{\lambda e^{-\lambda x}\,(\lambda x)^{n-1}}{(\mathrm{n}-1)!}, & x \geq 0 \\ 0, & x < 0 \end{cases}.$$

Lema 1. Neka su $X_1, \dots, X_n$ nezavisne slučajne promenljive sa eksponencijalnom raspodelom sa parametrom λ. Tada slučajna promenljiva $\sum_{i=1}^{n} X_i$ ima *gama raspodelu* sa parametrima n i λ.

Glava 2

2. Procesi čekanja (Jednostavni Markovljevi sistemi čekanja)

Mnoge pojave koje je potrebno matematički opisati uključuju i redove čekanja, bilo ljudi ili materijala. Čekanje u redu predstavlja stajanje u redu, a procesi čekanja u redu su stohastički procesi koji proizlaze iz fenomena stajanja u redu. Na primer, modeliranje procesa pristizanja kamiona sa žitaricama do elevatora, procesuiranje podataka u računarskom centru, kao i protok poslova kroz radionicu – sve to uključuje čekanje u redu. Mada je čekanje u redu sveprisutna pojava, ono se prilikom razvijanja determinističkih modela u cilju opisivanja sistema obično zanemaruje. Osim toga, zbog nasumičnih fluktuacija koja su svojstvena procesima masovnog posluživanja, sistemi se često ponašaju na neintuitivan način. Stoga je kod razvijanja sistemskih modela, kao i za razumevanje ponašanja sistema, izuzetno važno proučiti i čekanje u redu.

Sledeća četiri faktora utiču na (prosečnu) dužinu reda i u skladu sa tim na vreme čekanja:

- stopa pristizanja
- stopa usluživanja/obrade
- varijabilnost između vremena dolaska
- varijabilnost obrade zahteva

Intuitivno je jasno da se prosečan red povećava kada se svaki od stavki 1, 3 ili 4 povećava, i smanjuje se kada se stavka 2 povećava.

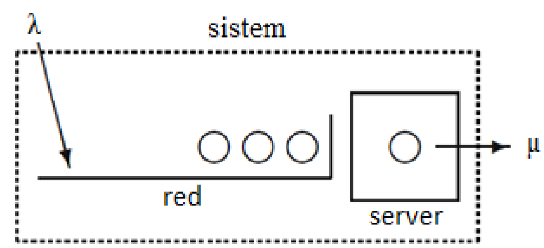
Razmotrićemo stohastičke procese $\{X(t),\ t > 0\}$ sa neprekidnim vremenom i diskretnim stanjima, gde $X(t)$ predstavlja *broj ljudi u sistemu čekanja u vremenskom trenutku* t.

Pretpostavljamo da korisnici koji ulaze u sistem dolaze po neku uslugu ili da bi obavili neki zadatak (na primer, da bi podigli novac sa automata). Može da postoji jedan ili više *servera* ili *uslužnih stanica*. Proces $\{X(t),\ t > 0\}$ predstavlja model za *masovno posluživanje* ili za *fenomen čekanja u redu*. Čekanje se odnosi na korisnike koji stoje u redu da bi bili posluženi, to jest koji *čekaju*, dok sistem masovnog posluživanja obuhvata sve korisnike u sistemu. Pošto *čekanje* predstavlja standardan izraz za ovaj tip procesa, ova dva izraza koristićemo kao sinonime. Osim toga, jasno je da se modeli čekanja u redu ne odnose samo na slučajeve kada nas interesuje broj osoba koje čekaju u redu. *Korisnici* sistema mogu da budu, na primer, i avioni koji čekaju na dozvolu za sletanje, mašine koje su poslate u radionicu na remont, itd.

2.1 Osnovne definicije i oznake

Proces masovnog posluživanja obuhvata pristizanje korisnika u uslužni objekat, kao i njihovo usluživanje. Za korisnike koji su pristigli, ali još nisu posluženi, kažemo da su na *čekanju*. Sistem čekanja u redu obuhvata sve klijente u redu i sve klijente u službi.

Slika 2.1. Prikaz jednog sistema masovnog usluživanja sa prosečnom stopom pristizanja λ i prosečnom stopom usluživanja μ, sa četiri klijenata u sistemu i tri klijenata u redu.



Tokom poslednjih nekoliko decenija razvijeno je više korisnih pravila koje pomažu pri specifikovanju pretpostavki koje se koriste u datoj analizi. U jednom istraživanju iz 1953. Kendall[2] je predložio način označavanja koji se i danas koristi za klasifikaciju raznih modela čekanja u redu, a koje je standardizovano 1971. godine (*Queueing Standardization Conference Report*, May 11, 1971). Radi se o jednom stenografu kojim je moguće ukratko i na brzinu izložiti pretpostavke datog modela masovnog posluživanja.

Najopštija *notacija* je oblika $A/S/s/c/p/D$, gde
A označava *raspodelu vremena između dva uzastopna dolaska* (proces dolaska),
S označava *raspodelu vremena usluživanja korisnika* (proces usluživanja),
s je *broj paralelnih servera u sistemu*,
c je *kapacitet sistema* (maksimalno dozvoljeni broj preuzimanja istovremenih poslova) ,
p je *veličina populacije* iz koje klijenti dolaze ,
D označava *politiku usluživanja*, pod nazivom *disciplina čekanja*, što predstavlja pravila izbora sledećeg posla/kupca za usluživanje.

[2]David George Kendall, penzionisani profesor Univerziteta u Kembridžu, u Engleskoj.

Pretpostavljamo da su vremena T_n između dolazaka uzastopnih klijenata (međuvreme dolaska) nezavisne i jednako raspodeljene slučajne promenljive. Slično tome, vremena usluživanja klijenata S_n su takođe nezavisne i jednako raspodeljene slučajne promenljive i nezavisne od T_n.

Najčešće korišćene raspodele za slučajne promenljive T_n i S_n, kao i odgovarajuće oznake za A ili S, susledeće:

M - *eksponencijalna* sa parametrima λ i μ;
E_k - *Erlangova* (ili *gama*) sa parametrima k i λ ili μ;
D - *deterministična* (ako su T_n i S_n konstante)
G - *opšta* (vremena mogu da slede bilo koju raspodelu).

Broj servera s je pozitivan ceo broj, ili ponekad beskonačan. Ako nije drugačije definisano, podrazumeva se da je kapacitet sistema beskonačan. Slično tome, veličina populacije iz koje klijenti dolaze pretpostavlja se da je beskonačna. Ako c (ili p) nije jednak beskonačnosti, njegova vrednost mora biti navedena. Sa druge strane, u slučaju kada je $c = p = \infty$, ove veličine možemo izostaviti u notaciji .

Disciplina čekanja, ako se ne naznači drugačije, je *FCFS* ("*first-come, first-served*"), tj. služi se prvi ko stigne. Ovaj slučaj se označava i kao *FIFO* ("*first in, first out*"), što znači "ko prvi ulazi, prvi izlazi", tj. služenje po redosledu pristizanja. Oznaku za ovu podrazumevanu discipilinu takođe možemo izostaviti u notaciji. U svim ostalim slučajevima, uslužna politika klijenata mora biti naznačena. Možemo imati *LCFS* ("*last come, first served*") koja odgovara principu štokovanja (skladištenja). Ova disciplina se takođe označava kao *LIFO* ("*last in, first out*"), tj. "ko poslednji ulazi, prvi izlazi". Klijenti mogu biti posluženi i nasumice, drugim rečima to znači da korisnici iz reda imaju istu verovatnoću da budu izabrani za opslugu. Ovu disciplinu označavamo sa *SIRO* (*"Service In Random Order*"), a druga skraćenica je *RS* (*Random Selection*).

U nekim slučajevima jedan ili više specijalnih korisnika dobijaju prioritet u posluživanju, što znači da pri dolasku posla/kupca sa većim prioritetom, tekući posao/kupac se može vratiti u red, da bi ustupio mesto poslu/kupcu sa većim prioritetom (npr. soba za urgentne slučajeve). Prve dve discipline uzimaju u razmatranje vreme nailaska, dok treća to ne razmatra i ne zahteva neku memoriju. Kao takva, može se koristiti kod prostijih tehničkih sistema.

Na primer, formula koja je razvijena za $G/D/1/\infty/FIFO$ čekanje biće formula koja se može koristiti za svaki opšti proces pristizanja, samo za determinističko vreme posluživanja, jedan server, neograničeni sistemski kapacitet, i uslužnu politiku koja je zasnovana na principu "ko prvi ulazi, prvi izlazi".

Cilj ovog rada je da predstavi proces masovnog posluživanja i da ukaže na probleme na koje nailazimo prilikom proučavanja čekanja u redu. Pretpostavićemo da su procesi pristizanja u sisteme masovnog posluživanja Poasonovi procesi, dok će vremena posluživanja biti eksponencijalna. Dakle, u ovom radu uglavnom ćemo se baviti $M/M/s/c$ sistemima pri različitim vrednostima s i c. Poasonova pristizanja i eksponencijalne usluge omogućavaju nam da primenimo Markovljeve modele masovnog posluživanja koji se lako mogu analizirati i čiji se rezultati mogu praktično primeniti. Istorijski gledano, ovi modeli su se u ranim fazama teorije redova čekanja koristili kao ispomoć u procesu odlučivanja u telefonskoj industriji. Osnovni Markovljev proces koji predstavlja broj korisnika u tim sistemima poznat je kao *proces rađanja i umiranja*, koji je u širokoj upotrebi kod modeliranja stanovništva. Terminologija rađanja/umiranja koristi se za predstavljanje rasta i pada broja stanovnika. Ove pojave u sistemima masovnog posluživanja predstavljaju *pristizanja* i *odlasci*.

2.2 Opšti model masovnog posluživanja po principu rađanja/smrti

Kada je veličina populacije n, i ovde se koristi terminologija rađanje (pristizanje) - umiranje (odlazak), dok su λ_n i μ_n *infinitezimalne stope prelaska* (*generatori*) rađanja, odnosno smrti. Kada stanovništvo predstavlja broj klijenata u sistemu, λ_n i μ_n ukazuju na to da stope dolazaka i odlazaka zavise od broja klijenata u sistemu. Na osnovu svojstava Poasonovog procesa, odnosno kada su dolasci u skladu sa Poasonovim procesom i vremena usluživanja eksponencijalna, možemo da damo sledeći izraz za verovatnoću prelaska tokom vremena $(t, t + \Delta t]$:

rađanje $(n \geq 0)$:

$$P(jedno\ rođenje) = \lambda_n \Delta t + o(\Delta t),$$

$$P(nema\ rađanja) = 1 - \lambda_n \Delta t + o(\Delta t),$$

$$P(više\ od\ jednog\ rođenja) = o(\Delta t),$$

umiranje $(n > 0)$:

$$P(jedna\ smrt) = \mu_n \Delta t + o(\Delta t),$$

$$P(nema\ umiranja) = 1 - \mu_n \Delta t + o(\Delta t),$$

$$P(više\ od\ jedne\ smrti) = o(\Delta t),$$

gde za $o(\Delta t)$ važi da $\frac{o(\Delta t)}{\Delta t} \to 0$ kada $\Delta t \to 0$. U ovim izrazima $o(\Delta t)$ ne određuje stvarne vrednosti. U svakom od navedena dva slučaja zbir izraza $o(\Delta t)$ je jednaka 0, tako da je ukupna verovatnoća ova tri događaja jednaka 1.

Neka $N(t)$ označava *broj klijenata u sistemu u vremenskom trenutku* t. Definišimo verovatnoću

$$P_{in}(t) = P\{N(t) = n | N(0) = i\}.$$

Uvrštavanjem verovatnoća prelaza tokom vremena $(t, t + \Delta t]$, kao što je navedeno iznad, dobijamo

$$\begin{aligned}
&P_{n,n+1}(\Delta t) = \lambda_n \Delta t + o(\Delta t), && n = 0, 1, 2, \dots,\\
&P_{n,n-1}(\Delta t) = \mu_n \Delta t + o(\Delta t), && n = 1, 2, 3, \dots,\\
&P_{nn}(\Delta t) = 1 - \lambda_n \Delta t - \mu_n \Delta t + o(\Delta t), && n = 1, 2, 3, \dots,\\
&P_{nj}(\Delta t) = o(\Delta t), && j \neq n-1, n, n+1.
\end{aligned} \tag{2.1}$$

Napomena 1. Proces rađanja i umiranja je *neprekidan slučaj lanaca Markova* kod kojeg se iz stanja n može preći jedino u stanje $n - 1$ (ako je $n > 0$) ili u stanje $n + 1$. U kratkom vremenskom intervalu dužine Δt, verovatnoća da će proces preći iz stanja n u stanje $n + 1$ je jednaka

$$\lambda_n \Delta t + o(\Delta t),$$

a verovatnoća da će preći iz stanja n u stanje $n - 1$ je jednaka

$$\mu_n \Delta t + o(\Delta t).$$

Prema tome, verovatnoća da će proces i dalje biti (ili će se vratiti) u stanje n nakon Δt jedinica vremena jednaka je

$$1 - (\lambda_n + \mu_n)\Delta t + o(\Delta t).$$

Kod izvođenja izraza na desnoj strani jednačina (2.1), iskoristićemo pojednostavljenja tipa

$$[\lambda_n \Delta t + o(\Delta t)][1 - \mu_n \Delta t + o(\Delta t)] = \lambda_n \Delta t + o(\Delta t),$$

$$[1 - \lambda_n \Delta t + o(\Delta t)][1 - \mu_n \Delta t + o(\Delta t)] = 1 - \lambda_n \Delta t - \mu_n \Delta t + o(\Delta t).$$

Infinitezimalne (beskrajno male) stope prelaza kod (2.1) vode do sledeće *generator matrice* za proces rađanja i umiranja kao modela sistema čekanja:

$$A = \begin{bmatrix} -\lambda_0 & \lambda_0 & & \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & \\ & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 \\ & & \ddots & \\ & & & \\ & & & \end{bmatrix}. \tag{2.2}$$

Generator matrica A vodi do sledeće *Kolmogorove forward* jednačine za $P_{in}(t)$. (Radi jednostavnijeg označavanja od sada pa nadalje, pisaćemo $P_{in}(t) = P_n(t)$ i naznačiti početno stanje i tamo gde je to potrebno).

$$P_0^{'}(t) = -\lambda_0 P_0(t) + \mu_1 P_1(t),$$

$$P_n^{'}(t) = -(\lambda_n + \mu_n)P_n(t) + \lambda_{n-1}P_{n-1}(t) + \mu_{n+1}P_{n+1}(t), \qquad n = 1, 2, \dots . \tag{2.3}$$

Ove jednačine predstavljaju skup diferencijalnih jednačina koje nazivamo Kolmogorovim diferencijalnim jednačinama. Njihovo rešenje je skup jednačina koji pokazuju kako se svaka verovatnoća menja sa vremenom.

Izraz (2.3) se takođe može izvesti direktno koristeći (2.1), bez prolaska kroz generator matricu kao što je prikazano u nastavku . Imajući u vidu prelaze procesa $N(t)$ tokom vremena $(t, t + \Delta t\,]$, imamo

$$P_0(t+\Delta t) = P_0(t)[1 - \lambda_0 \Delta t + o(\Delta t)] + P_1(t)[\,\mu_1 \Delta t + o(\Delta t)],$$

$$\begin{aligned} P_n(t+\Delta t) = {} & P_n(t)[1 - \lambda_n \Delta t - \mu_n \Delta t + o(\Delta t)] \\ & + P_{n-1}(t)[\,\lambda_{n-1}\Delta t + o(\Delta t)] \\ & + P_{n+1}(t)[\,\mu_{n+1}\Delta t + o(\Delta t)] \\ & + o(\Delta t), \qquad n = 1, 2, \dots \end{aligned} \tag{2.4}$$

Oduzimanjem $P_n(t)$ $(n = 0, 1, 2, \dots)$ sa obe strane odgovarajuće jednačine u (2.4) i deljenjem sa Δt dobijamo

$$\frac{P_0(t+\Delta t) - P_0(t)}{\Delta t} = -\lambda_0 P_0(t) + \mu_1 P_1(t) + \frac{o(\Delta t)}{\Delta t},$$

$$\frac{P_n(t+\Delta t) - P_n(t)}{\Delta t} = -(\lambda_n + \mu_n)P_n(t)$$

$$+\lambda_{n-1}P_{n-1}(t) + \mu_{n+1}P_{n+1}(t) + \frac{o(\Delta t)}{\Delta t}.$$

Kada $\Delta t \to 0$ sledi (2.3), tj.

$$P_0^{'}(t) = \lim_{\Delta t \to 0} \frac{P_0(t+\Delta t) - P_0(t)}{\Delta t} = -\lambda_0 P_0(t) + \mu_1 P_1(t),$$

$$P_n^{'}(t) = \lim_{\Delta t \to 0} \frac{P_n(t+\Delta t) - P_n(t)}{\Delta t} = -(\lambda_n + \mu_n) P_n(t) + \lambda_{n-1} P_{n-1}(t) + \mu_{n+1} P_{n+1}(t),$$

$n = 1, 2, \dots,$

Da bismo odredili $P_n(t)$ (tj. $P_{in}(t)$), treba da rešimo jednačine (2.3) zajedno sa početnim uslovima $P_i(0) = 1$, $P_n(0) = 0$ za $n \neq i$.
Nažalost, čak i u jednostavnim slučajevima kao što su $\lambda_n = \lambda$ i $\mu_n = \mu$, $n = 0, 1, 2, 3, \dots$, to je slučaj kada dolasci imaju Poasonovu raspodelu, a vremena usluživanja eksponencijalnu ($M/M/1$ model čekanja u redu), eksplicitno izvođenje $P_n(t)$ je mukotrpan proces. Osim toga, u većini aplikacija, poznavanje vremenski zavisnih ponašanja nije od ključne važnosti. Stoga je granični rezultat, koji je određen na osnovu (2.3) ako $t \to \infty$, je rezultat koji je u najširoj primeni. Opšti rezultat koji se odnosi na procese Markova je dat u nastavku.

Teorema 2.1.
(1) Ako je proces Markova nesvodljiv (moguće je preći iz bilo kog stanja u bilo koje stanje), tada granična raspodela $lim_{t\to\infty} P_n(t) = p_n$ postoji i nezavisna je od početnih uslova procesa. Granice $\{p_n, n \in S\}$ su takve da su one ili identički jednake 0 (tj., $p_n = 0$ za sve $n \in S$) ili su sve pozitivne i formiraju raspodelu verovatnoće (tj., $p_n > 0$ za sve $n \in S, \sum_{n\in S} p_n = 1$).
(2) Granice $\{p_n, n \in S\}$ nesvodljivog rekurentnog (povratnog) procesa Markova predstavlja jedinstveno rešenje jednačine $\boldsymbol{p}A = \boldsymbol{0}$ i $\sum_{j\in S} p_j = 1$, gde je $\boldsymbol{p} = (p_0, p_1, p_2, \dots)$.

Zbog pretpostavke o stacionarnosti nakon nekog perioda prelaska sistem će postati stabilan. Naravno, stanje će se stalno menjati, ali verovatnoće različitog broja klijenata u sistemu će biti konstantna. Drugim rečima, u stanju ravnoteže, poznatom kao *ustaljeno stanje*, ponašanje procesa je nezavisno od vremenskog parametra i od početne vrednosti, tj.

$$\lim_{t\to\infty} P_{in}(t) = p_n, \qquad n = 0, 1, 2, \dots .$$

Dakle, funkcija $P_n(t)$ postaje konstanta p_n i zbog toga

$$P_n^{'}(t) \to 0 \quad \text{kad} \quad t \to \infty.$$

Korišćenjem ovih rezultata kod (2.3), dobijamo skup algebarskih jednačina za stabilno stanje:

$$0 = -\lambda_0 p_0 + \mu_1 p_1,$$

$$0 = -(\lambda_n + \mu_n) p_n + \lambda_{n-1} p_{n-1} + \mu_{n+1} p_{n+1}, \qquad n = 1, 2, \dots . \tag{2.5}$$

Ove jednačine se mogu jednostavno rešiti pomoću rekurzije. Iz prve jednačina, dobijamo

$$p_1 = \frac{\lambda_0}{\mu_1} p_0. \tag{2.6}$$

Za $n = 1$, iz druge jednačine sledi

$$(\lambda_1 + \mu_1) p_1 = \lambda_0 p_0 + \mu_2 p_2.$$

Korišćenjem (2.6), ova jednačina se svodi na

$$\mu_2 p_2 = \lambda_1 p_1,$$

$$p_2 = \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2} p_0.$$

Nastavljajući ovu rekurziju za $n = 2, 3, \ldots$, dobijamo

$$\mu_n p_n = \lambda_{n-1} p_{n-1}, \tag{2.7}$$

iz čega sledi

$$p_n = \frac{\lambda_{n-1}}{\mu_n} p_{n-1} = \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n} p_0. \tag{2.8}$$

Teorema 2.1. takođe daje uslov normalizacije $\sum_{n \in S} p_n = 1$, koji kada se primeni na (2.8) daje

$$p_0 = \left[1 + \sum_{n=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n} \right]^{-1}, \tag{2.9}$$

što možemo zapisati na sledeći način

$$p_0 = \left[1 + \sum_{n=1}^{\infty} \prod_{k=0}^{n-1} \frac{\lambda_k}{\mu_{k+1}} \right]^{-1}.$$

Granična raspodela stanja modela rađanja i umiranja je $\{p_n, n = 0, 1, 2, \ldots\}$, kao što je dato sa (2.8) i (2.9). Primetimo da je $\{p_n, n = 0, 1, 2, \ldots\}$ različito od nule samo kada

$$1 + \sum_{n=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n} < \infty. \tag{2.10}$$

Da bismo dobili (2.5), nije potrebno prethodno izvoditi (2.3). Kako je navedeno u Teoremi 2.1., pomoću $\boldsymbol{p} = (p_0, p_1, p_2, \dots)$ i generator matrice A, (2.5) se može direktno dobiti iz

$$\boldsymbol{p}A = \boldsymbol{0}$$

i (2.11)

$$\sum_{n=1}^{\infty} p_n = 1.$$

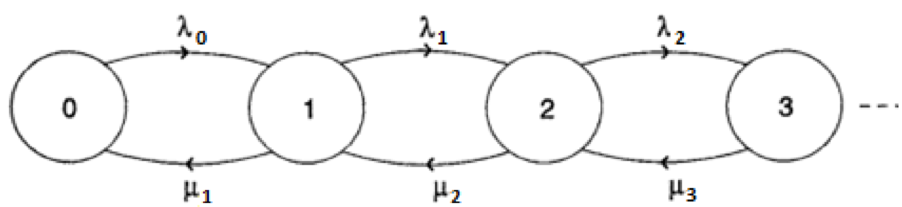
Za model masovnog posluživanja po principu rađanja/smrti, generator matrica A je data sa (2.2).

Drugi način za posmatranje (2.5) jeste posmatranje jednačina kao uslova razvnoteže između stanja. Preuređenjem (2.5) dobijamo

$$\lambda_0 p_0 = \mu_1 p_1,$$

$$(\lambda_n + \mu_n)p_n = \lambda_{n-1}p_{n-1} + \mu_{n+1}p_{n+1}. \tag{2.12}$$

Prelazi između stanja mogu na slikoviti način da se predstave kao što je to prikazano na Slici 2.2. Na dijagramu svako stanje je prikazano kružićem, a strelice predstavljaju prelaze sa pozitivnim verovatnoćama.



Slika 2.2. Dijagram prelaza

Uz napomenu da λ_s i μ_s predstavljaju infinitezimalne stope ulaska i izlaska iz stanja, jednakosti u (2.12) mogu se tumačiti kao

$$(dugoročna\ verovatnoća\ stanja\ n) \times (stopa\ izlaska\ iz\ stanja\ n) =$$

$$= \textstyle\sum_{i=n-1,n+1} (dugoročna\ verovatnoća\ stanja\ i) \times$$

$$\times (stopa\ prelaska\ iz\ stanja\ i\ u\ stanje\ n).$$

Jednačine ravnotežnog stanja se mogu lako napisati korišćenjem dijagrama prelaska što je prikazano na Slici 2.2.

Navešćemo prethodna tri načina koje smo izložili za utvrđivanje jednačina ravnotežnog stanja:

1. određivanje odgovarajućih ograničenja u forward Kolmogorovoj jednačini kad $t \to \infty$;
2. korišćenjem jednačine $\boldsymbol{p}A = \mathbf{0}$;
3. uz pomoć dijagrama prelaska.

Prilikom korišćenja ove poslednje metode, treba obratiti pažnju na to da li su svi prelasci uračunati. U našoj raspravi o specijalnim modelima, obično koristimo drugu metodu koja je zasnovana na generatorskoj matrici, osim ako ponašanje sistema nije potpunije objašnjeno dijagramom prelaska.

Postoje i dve druge teoreme kojima se ustanovljuju neka važna svojstva granične raspodele Markovljevog procesa sa nesvodljivim skupom stanja. Prva od njih bavi se konceptom stacionarnosti.

Definicija 2.1. Proces je *stacionaran* ako je raspodela stanja nezavisna od vremena, tj. ako

$$P_{\boldsymbol{n}}(0) = p_n, \qquad n = 0, 1, 2, \dots,$$

tada

$$P_n(t) = p_n \qquad za\ svako\ t.$$

Teorema 2.2. *Granična raspodela pozitivnog rekurentnog nesvodljivog procesa Markova je takođe stacionarna.*

Pošto se bavimo raspodelama prelaska koje su uslovljene početnim stanjem u stohastičkim procesima, stacionarnost označava da ako stacionarnu raspodelu koristimo kao početnu raspodelu stanja, od tog trenutka pa nadalje sve vremenski zavisne raspodele biće istovetne kao i one od kojih smo krenuli. Druga teorema nam omogućuje da tumačimo graničnu verovatnoću $p_n, n = 0, 1, 2, \dots$, kao deo vremena tokom koje se proces dugoročno nalazi u stanju n.

Teorema 2.3. *Počevši od stanja i, neka $N_{ij}(t)$ označava vreme provedeno po procesu Markova u stanju j tokom vremena $(0, t]$. Tada*

$$\lim_{t\to\infty}\left[\left|\frac{N_{ij}(t)}{t} - p_j\right| > \varepsilon\right] = 0.$$

Opšti model masovnog posluživanja po principu rađanja/umiranja obuhvata jednu široku lepezu specijalnih slučajeva. U nastavku rada razmatraćemo neke od široko primenjenih modela.

2.3 Sistemi masovnog posluživanja sa jednim serverom

Najjednostavniji sistemi čekanja za analizu su oni koji uključuju jedan eksponencijalni server i gde dolasci odgovaraju Poasonovom procesu. Ove sisteme relativno je lako proučavati, a mogu da posluže i za demonstriranje nekih opštih tehnika analize masovnog posluživanja koje se mogu proširiti i na složenije sisteme.

2.3.1 Sistemi sa jednim serverom i beskonačnim kapacitetom ($M/M/1$ model)

Razmotrimo prvo sistem čekanja sa jednim serverom usluživanja u kojoj klijenti pristižu u skladu sa Poasonovim procesom sa srednjom stopom λ i bivaju usluženi od strane jednog servera. Uzastopni vremenski intervali usluživanja su po pretpostavci nezavisne eksponencijalne slučajne promenjive sa srednjom vrednošću $1/\mu$. Pretpostavljamo da je kapacitet sistema beskonačan, kao i populacija iz koje klijenti dolaze. Usluživanje se vrši po principu "ko prvi dolazi, prvi odlazi". Stoga, ovaj sistem možemo jednostavno označiti sa $M/M/1$, ili ekvivalentno ovome kao $M/M/1/\infty/\infty/FIFO$ sistem. Svaki klijent po dolasku se direktno uslužuje ako je server slobodan, a ako nije, klijent staje u red koji je beskonačnog kapaciteta. Kada je korisnik poslužen, on napušta sistem, a korisnik koji je bio najduže na čekanju (u redu) momentalno ulazi u uslužnu stanicu i pružanje usluge počinje iz početka. Dakle, protok klijenata kroz sistem je proces Markova sa skupom stanja $\{0, 1, 2, ...\}$. Proces Markova označićemo sa $\{N(t), t \geq 0\}$, gde $N(t)$ označava *broj klijenata u sistemu u trenutku* t.

Verovatnoće stacionarnog (ustaljenog) stanja su

$$p_n = \lim_{t\to\infty} P\{N(t) = n\}.$$

Drugim rečima, p_n je *granična* ili *dugoročna verovatnoća* da će biti tačno n mušterija u sistemu. Često se dobija da je p_n proporcionalno jednako sa vremenom kad sistem ima tačno n mušterija. Na primer, ako je $p_0 = 0.4$, onda će sistem ostati bez mušterija za 40% vremena. Slično tome, ako je $p_1 = 0.3$ znači da će za 30% vremena u sistemu ostati tačno jedna mušterija.

Neka je N slučajna promenljiva sa raspodelom verovatnoća $\{p_0, p_1, p_2, ...\}$. N predstavlja *broj klijenata/poslova u sistemu* (u redu i u objektu usluživanja/obrade) u ustaljenom stanju i p_n predstavlja *dugoročnu verovatnoću da će biti tačno* n *klijenata u sistemu*. Neka slučajna promenljiva N_q označava *broj klijenata u redu* kad je sistem u stacionarnom stanju (broj kupaca koji su u redu i čekaju na usluživanje). Drugim rečima, ako je *sistem prazan,* $N_q = N$, a kada je *sistem zauzet,* $N_q = N - 1$.

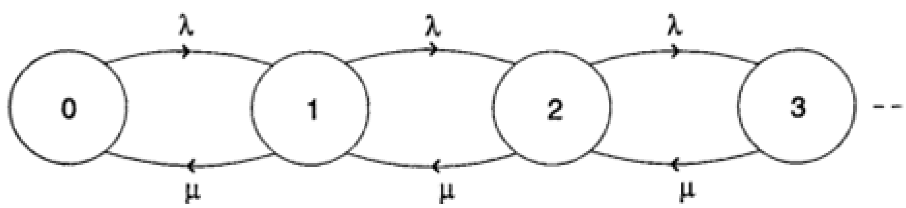
Izvođenje izraza za p_n, $n = 0, 1, \dots$, koji uključuje srednje stope dolaska i odlaska obično podrazumeva dva koraka:

(1) izvesti sistem jednačina za definisanja verovatnoća i

(2) rešiti sistem jednačina.

Korak (1) se relativno lako nalazi, međutim ono što je obično teže jeste korak (2). Jedna od intuitivnih metoda za dobijanje sistema jednačina jeste da se nacrta dijagram stanja i da se primeni metoda ravnoteže stopa, koja se sastoji od sistema jednačina formiranih izjednačavanjem "stope ulaska" sa "stopom izlaska" za svako stanje sistema čekanja.

Ovaj sistem jednačina koji izjednačava stope po kojima proces ulazi u određeno stanje sa stopom po kojoj proces izlazi iz tog stanja se naziva *ravnotežna jednačina*.



Slika 2.3. Dijagram stanja za $M/M/1$ sistem.

Za svako $n \geq 0$ stopa učestalosti po kojoj proces ulazi u stanje n je jednaka stopi učestalosti po kojoj proces izlazi iz stanja n. Za određivanje ovih stopa razmotrimo prvo stanje 0. Kada je u stanju 0, proces ga može napustiti samo ako klijent dođe, jer niko ne može otići iz sistema ako je on prazan. Pošto je stopa dolazaka λ, a verovatnoća za koju je sistem u stanju 0 je p_0, sledi da je stopa učestalosti po kojoj proces napušta stanje 0 jednaka λp_0. S druge strane u stanje 0 se može doći iz stanja 1 samo po odlasku. Odnosno ako je samo jedan klijent u sistemu i usluživanje se završi onda sistem ostaje prazan. Pošto je stopa odlazaka μ, a verovatnoća po kojoj je u sistemu jedan klijent je p_1, sledi da je stopa učestalosti po kojoj sistem ulazi u stanje 0 jednaka μp_1.

Dakle iz našeg principa jednakosti stopa dobijamo prvu jednačinu

$$\lambda p_0 = \mu p_1.$$

Sada razmatramo stanje 1. Proces može napustiti ovo stanje ili dolaskom (koji se dešava po stopi λ) ili odlaskom (koji se dešava po stopi μ). Zato kada je u stanju 1 proces napušta ovo stanje po stopi $\lambda + \mu$. Pošto je verovatnoća da je proces u stanju 1 jednaka p_1, stopa po kojoj proces napušta ovo stanje je $(\lambda + \mu)p_1$. Drugim rečima, u stanje 1 se može doći iz stanja 0 po dolasku klijenta ili iz stanja 2 po odlasku klijenta. Prema tome stopa po kojoj proces ulazi u stanje 1 je jednaka $\lambda p_0 + \mu p_2$. Pošto je način sličan za ostala stanja dobijamo sledeći *sistem jednačina za ravnotežno stanje sistema*:

stanje i	*stopa po kojoj proces napušta stanje i* ("stopa izlaska")	=	*stopa po kojoj proces ulazi u stanje i* ("stopa ulaska")
0	λp_0	=	μp_1
$n = 1, 2, \ldots$	$(\lambda + \mu) p_n$	=	$\lambda p_{n-1} + \mu p_{n+1}$

Da bi rešili ovaj sistem jednačina pišemo ga u sledećem obliku:

$$p_1 = \frac{\lambda}{\mu} p_0 \tag{2.13}$$

$$p_{n+1} = \frac{\lambda + \mu}{\mu} p_n - \frac{\lambda}{\mu} p_{n-1} \quad za\, n = 1, 2, \ldots\ .$$

Jedan od rigoroznijih pristupa za dobijanje sistema jednačina (2.13) je da se najpre odredi generator matrica za $M/M/1$ sistem. $M/M/1$ model je poseban slučaj opšteg modela masovnog posluživanja po principu rađanja i umiranja sa $\lambda_n = \lambda$ i $\mu_n = \mu$. Beskonačno dimenzionalna *generator matrica* sa prostorom stanja $\{0, 1, 2, \ldots\}$ je data sa

$$A = \begin{bmatrix} -\lambda & \lambda & & & \\ \mu & -(\lambda+\mu) & \lambda & & \\ & \mu & -(\lambda+\mu) & \lambda & \\ & & \ddots & & \\ & & & & \end{bmatrix}. \tag{2.14}$$

Sistem jednačina formiran sa $\boldsymbol{p}A = \boldsymbol{0}$ ponovo dovodi do jednačina (2.13).
Sistem jednačina može se rešiti sukcesivnom supstitucijom rešenja i izražavanjem svih promenljivih u odnosu na p_0. Pošto smo već odredili da je $p_1 = \frac{\lambda}{\mu} p_0$, ostaje nam da odredimo p_2, pa p_3:

$$p_2 = \frac{\lambda + \mu}{\mu} p_1 - \frac{\lambda}{\mu} p_0 = \frac{\lambda + \mu}{\mu}\left(\frac{\lambda}{\mu} p_0\right) - \frac{\lambda}{\mu}\frac{\mu}{\mu} p_0 = \frac{\lambda^2}{\mu^2} p_0 = \left(\frac{\lambda}{\mu}\right)^2 p_0$$

$$p_3 = \frac{\lambda + \mu}{\mu} p_2 - \frac{\lambda}{\mu} p_1 = \frac{\lambda + \mu}{\mu}\left(\frac{\lambda^2}{\mu^2} p_0\right) - \frac{\lambda}{\mu}\frac{\mu}{\mu}\left(\frac{\lambda}{\mu} p_0\right) = \frac{\lambda^3}{\mu^3} p_0 = \left(\frac{\lambda}{\mu}\right)^3 p_0. \tag{2.15}$$

U ovom trenutku počinje da se javlja obrazac i možemo tvrditi da

$$p_n = \left(\frac{\lambda}{\mu}\right)^n p_0 \quad za \;\; n \geq 0. \tag{2.16}$$

Ova tvrdnja se dokazuje matematičkom indukcijom , tj. koristeći indukcijsku hipotezu zaj edno sa opštom jednačinom u jednačini (2.13) dobijamo

$$p_{n+1} = \frac{\lambda+\mu}{\mu}\left(\frac{\lambda^n}{\mu^n}p_0\right) - \frac{\lambda}{\mu}\frac{\mu}{\mu}\left(\frac{\lambda^{n-1}}{\mu^{n-1}}p_0\right) = \frac{\lambda^{n+1}}{\mu^{n+1}}p_0 = \left(\frac{\lambda}{\mu}\right)^{n+1} p_0,$$

i na ovaj način smo dokazali da važi jednačina (2.16).

Odnos λ/μ koji ćemo označiti sa ρ za M/M/1 sistem se naziva *intenzitet saobraćaja* za sistem čekanja u redu, ili *stopa iskorišćenosti* sistema. (U opštijem slučaju ρ se obično definiše kao količnik stope dolaska i najveće stope usluživanja sistema). Sada imamo p_n za sve n izražene preko p_0. Da bismo odredili p_0 koristimo činjenicu da je zbir svih verovatnoća p_n jednak 1, pa je:

$$1 = \sum_{n=0}^{\infty} p_n = p_0 \sum_{n=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^n = p_0 \sum_{n=0}^{\infty} \rho^n = \frac{p_0}{1-\rho}. \tag{2.17}$$

Odavde sledi da je

$$p_0 = 1 - \rho. \tag{2.18}$$

Jednakost u prethodnom izrazu smo dobili koristeći geometrijsku progresiju, tako da ona važi samo za $\rho < 1$, jer u suprotnom bi $\sum_{n=0}^{\infty}\left(\frac{\lambda}{\mu}\right)^n$ bila beskonačna i sve verovatnoće p_n bi bile jednake nuli. Dakle, pretpostavićemo da je $\rho < 1$. Parametar ρ nam daje *prosečan broj pristizanja* u sistem tokom vremenskog perioda koji odgovara prosečnom vremenu koje je potrebno za posluživanje jednog proizvoljnog klijenta.

U slučaju da je $\rho \geq 1$, prosečan broj klijenata i vreme provedeno u sistemu bi se povećavalo bez ograničenja i ne bi postojale granične verovatnoće, tj. sistem bi postao *nestabilan*. Uslov $\rho < 1$ je ekvivalentan uslovu da srednje vreme usluživanja bude manje od srednjeg vremena uzastopnih dolazaka. Ovo je opšti uslov koji mora biti zadovoljen da bi postojale granične verovatnoće kod sistema sa jednim serverom.

Ako vrednost koju smo dobili za p_0 uvrstimo u jednačinu (2.16) tada dobijamo sledeći izraz za p_n za $M/M/1$ sistem:

$$p_n = (1-\rho)\rho^n \quad za \;\; n = 0, 1, \dots, \tag{2.19}$$

gde je $\rho = \lambda/\mu$ i $\rho < 1$.

Koraci u dobijanju jednačine (2.19) isti su kao i kod mnogih drugih sistema Markova. Korisno je da se osvrnemo na te korake kako bismo se upoznali sa njima. Oni su sledeći:

1. Formiranje generator matrice Markova, A (Jednačina 2.14).
2. Određivanje sistema jednačina rešavanjem $\boldsymbol{p}A = \mathbf{0}$ (Jednačina 2.13).
3. Rešavanje sistema jednačina u funciji p_0 pomoću sukcesivne supstitucije ili indukcijom ako je moguće (Jednačina 2.16).
4. Određivanje p_0 pomoću normirane jednačine (Jednačina 2.17).

Jednačine (2.19) se mogu direktno koristiti za određivanje sledećih verovatnoća:

$$P\{server\ je\ slobodan\} = P\{nema\ čekanja\ po\ dolasku\} = p_0 = 1 - \rho$$

$$P\{server\ je\ zauzet\} = P\{čeka\ se\ po\ dolasku\} = 1 - p_0 = \rho$$

$$P\{n\ klijenata\ u\ sistemu\} = p_n = (1-\rho)\rho^n$$

$$P\{n\ ili\ više\ klijenata\ u\ sistemu\} = $$
$$= \sum_{m=n}^{\infty} p_m = \sum_{k=0}^{\infty} \rho^{n+k}(1-\rho) = (1-\rho)\,\rho^n \sum_{k=0}^{\infty} \rho^k = \rho^n$$

$$P\{manje\ od\ n\ klijenata\ u\ sistemu\} = 1 - \rho^n$$

Primer 2.1. Operater malog elevatora na raspolaganju ima samo jednu platformu za istovar žitarica. Dolasci kamiona u špicu sezone čine Poasonov proces sa prosečnom stopom pristizanja od 5 kamiona na sat. Zbog različite veličine tovara, vreme koje svaki kamion provodi pred platformom za istovar je slučajna promenljiva sa eksponencijalnom raspodelom i prosečnim vremenom od 11 minuta. Uz pretpostavku da je prostor za parkiranje neograničen, redovi čekanja koji se formiraju opisuju se kao $M/M/1$ sistem. Shodno ovome imamo:

(a) Iskorišćenost platforme za istovar žitarica:

$$\lambda = stopa\ pristizanja = 5/č,$$

$$\mu = stopa\ usluživanja = \frac{60}{11}/č,$$

$$\rho = stopa\ iskorišćenosti = \frac{\lambda}{\mu} = 0.9166.$$

(b) Verovatnoća da je platforma za istovar slobodna:

$$p_0 = 1 - \rho = 0.0834.$$

(c) Verovatnoća da tačno četiri kamiona čekaju na istovar:

$$P\{N_q = 4\} = P\{N = 5\} = p_5 = \rho^5 \cdot p_0 = 0.9166^5 \cdot 0.0834 = 0.054,$$

odnosno 5.4 %.

(d) Verovatnoća da u sistemu ima pet ili više kamiona:

$$P\{N \geq 5\} = \sum_{n=5}^{\infty} p_n = \rho^5 = 0.9166^5 = 0.647,$$

tj. 64.7 %. □

Za opisivanje *efikasnosti* sistema redova čekanja postoji više kriterijuma (mera). Najčešće se za tu meru uzima *očekivani broj klijenata u sistemu*, u oznaci L, i očekivani *broj klijenata u redu*, u oznaci L_q. Ove očekivane vrednosti dobijaju se korišćenjem jednačine (2.19) i pomoću izvoda geometrijske progresije, što rezultira izrazom za očekivani broj u $M/M/1$ sistemu kao:

$$L = E(N) = \sum_{n=0}^{\infty} np_n = \sum_{n=1}^{\infty} np_n = \sum_{n=1}^{\infty} n(1-\rho)\rho^n = (1-\rho)\rho \sum_{n=1}^{\infty} n\rho^{n-1} =$$

$$= (1-\rho)\rho \left(\sum_{n=0}^{\infty} \rho^n \right)' = (1-\rho)\rho \left(\frac{1}{1-\rho} \right)' = \frac{\rho}{1-\rho}, \tag{2.20}$$

što možemo još napisati na sledeći način

$$L = E(N) = \frac{\lambda}{\mu - \lambda}.$$

Očekivani broj klijenata koji čekaju u redu u jednom M/M/1 sistemu se izvodi na sličan način:

$$L_q = E(N_q) = 0 \cdot (p_0 + p_1) + \sum_{n=1}^{\infty} np_{n+1} = \sum_{n=1}^{\infty} n\,(1-\rho)\rho^{n+1} = (1-\rho)\rho^2 \sum_{n=1}^{\infty} n\rho^{n-1} =$$

$$= (1-\rho)\rho^2 \left(\sum_{n=0}^{\infty} \rho^n \right)' = (1-\rho)\rho^2 \left(\frac{1}{1-\rho} \right)' = \frac{\rho^2}{1-\rho}, \tag{2.21}$$

što takođe možemo još napisati kao

$$L_q = E(N_q) = \frac{\lambda^2}{\mu(\mu - \lambda)}.$$

Vidimo da je $L_q = L \cdot \rho$ i iz ove formule možemo izračunati L_q pod uslovom da znamo L i obrnuto.

Prilikom izvođenja L_q koristili smo sledeću relaciju između N i N_q:

$$N_q = \begin{cases} 0 & za \ N = 0,1 \\ N-1 & za \ N \geq 2 \end{cases} \tag{2.22}$$

Faktor iskorišćenja ρ je verovatnoća da je server z auzet kada je sistem u ravnotežnom stanju, stoga on nam daje *očekivani broj klijenata u servisu* (službi), tj. $E(N_s)$. Sa ovakvom tumačenjem možemo dati očigledno objašnjenje za (2.21) kao:

$$E(broj\ klijenata\ u\ sistemu\ ) \\ = E(broj\ klijenata\ u\ redu) + E(broj\ klijenata\ u\ servisu),$$

tj.

$$E(N) = E\left(N_q\right) + E(N_s),$$

iz čega dobijamo

$$E(N_s) = E(N) - E\left(N_q\right) = \frac{\rho}{1-\rho} - \frac{\rho^2}{1-\rho} = \rho.$$

Ovo važi za sve redove čekanja, budući sa svaki klijent koji je prisutan u sistemu ili čeka u redu ili je u procesu pružanja usluge.

Prilikom opisivanja slučajne promenljive nije poželjno koristiti samo srednju vrednost. Iz ovog razloga navešćemo i *disperziju broja klijenata u sistemu* i *u redu*. Disperzija slučajne promenljive N koja predstavlja broj klijenata u sistemu koji je u ravnotežnom stanju se lako određuje ako primetimo u jednačini (2.19) da

$$N + 1 :\ \mathcal{G}(1-\rho).$$

Sledi da

$$D(N) = D(N+1) = \frac{\rho}{(1-\rho)^2} = \frac{\lambda\mu}{(\ \mu - \lambda\ )^2}. \tag{2.23}$$

Kako je $E(N+1) = E(N) + 1$, primetimo da se L može izraziti na sledeći način:

$$L = E(N) = \frac{1}{1-\rho} - 1 = \frac{\rho}{1-\rho}.$$

Slučajna promenljiva N_s koja predstavlja *broj klijenata u servisu* ima Bernulijevu raspodelu sa parametrom $\rho = 1 - p_0$. Možemo odrediti njenu disperziju:

$$D(N_s) = (1-\rho)\rho = \left(1-\frac{\lambda}{\mu}\right)\frac{\lambda}{\mu}.$$

Za određivanje disperzije slučajne promenljive N_q koristićemo relaciju (2.22) između N_q i N. Iz toga sledi

$$E\left(N_q^2\right) = \sum_{n=1}^{\infty} n^2 p_{n+1} = \sum_{n=1}^{\infty} n^2 (1-\rho)\rho^{n+1} = \rho^2 \sum_{n=1}^{\infty} n^2 p_{n-1} = \rho^2 E(Z^2),$$

gde $Z : \mathcal{G}(1-\rho)$, iz čega možemo zaključiti da

$$E\left(N_q^2\right) = \rho^2\left(D(Z) + E^2(Z)\right) = \rho^2\left(\frac{\rho}{(1-\rho)^2} + \frac{1}{(1-\rho)^2}\right) = \rho^2\frac{\rho+1}{(1-\rho)^2}.$$

Srednja vrednost za N_q je dato sa (2.21), iz čega dobijamo

$$D\left(N_q\right) = E\left(N_q^2\right) - E^2\left(N_q\right) = \rho^2\frac{\rho+1}{(1-\rho)^2} - \frac{\rho^4}{(1-\rho)^2} = \frac{\rho^2(1+\rho-\rho^2)}{(1-\rho)^2}. \tag{2.24}$$

Napomena 2. S obzirom da je $N = N_q + N_s$, možemo koristiti formulu

$$D(N) = D\left(N_q\right) + D(N_s) + 2cov\left(N_q, N_s\right)$$

za izračunavanje kovarijanse (i koeficijenta korelacije) slučajne promenljive N_q i N_s.

Neka od osnovnih veličina (mera) vezana za modele redova čekanja pored N i N_q su i *ukupno vreme koje klijent provede u sistemu* (koji je u ravnotežnom stanju) i *u redu*, u oznaci T i T_q, respektivno. Ukupno vreme provedeno u sistemu (T) jednako je $T_q + T_s$, gde T_s predstavlja *vreme potrebno za uslugu*. Obično je teško pronaći raspodelu za slučajnu promenljivu T. Za $M/M/1$ model čekanja u redu ovu raspodelu možemo odrediti na eksplicitan način. Da bismo to učinili, potrebno nam je rezultat sledeće tvrdnje koju ćemo dokazati u nastavku.

Teorema 2.4. *Neka α_n predstavlja verovatnoću da će proizvoljan klijent zateći n klijenata (stanje n) u sistemu kada dođe. Ako klijenti pristižu u skladu sa Poasonovim procesom, tada*

$$p_n = \alpha_n \quad za\ svako\ n \geq 0.$$

Dokaz: Koristićemo činjenicu da Poasonov proces ima nezavisne priraštaje. Pretpostavimo da klijent stiže u trenutku t. Označimo sa F_ε klijenta koji stiže u intervalu $[t, t+\varepsilon)$. Imamo da je

$$\alpha_n = \lim_{t\to\infty}\lim_{\varepsilon\to 0} P\{X(t^-) = n \mid F_\varepsilon\} = \lim_{t\to\infty}\lim_{\varepsilon\to 0} \frac{P[\{X(t^-) = n\} \cap F_\varepsilon]}{P(F_\varepsilon)} =$$

$$= \lim_{t\to\infty}\lim_{\varepsilon\to 0} \frac{P\{F_\varepsilon | X(t^-) = n\} P\{X(t^-) = n\}}{P(F_\varepsilon)} = \lim_{t\to\infty}\lim_{\varepsilon\to 0} \frac{P(F_\varepsilon) P\{X(t^-) = n\}}{P(F_\varepsilon)} =$$

$$= \lim_{t\to\infty} \lim_{\varepsilon\to 0} P\{X(t^-) = n\} = \lim_{t\to\infty} P\{X(t^-) = n\} = p_n \qquad \blacksquare$$

Napomena 3. Takođe može se pokazati da u bilo kom sistemu gde klijenti dolaze jedan po jedan, i uslužuju se jedan po jedan, važi da je

$$\alpha_n = \delta_n \quad za\, svako\, n \geq 0,$$

gde δ_n predstavlja *dugoročan udeo vremena u kome po odlasku klijenta iz sistema iza njega ostavi* n *klijenata*. Ovo proizilazi iz činjenice da je stopa prelazaka iz stanja n u stanje $n+1$ jednaka stopi prelazaka iz stanja $n+1$ u stanje n tokom dugog vremenskog perioda, ili ekvivalentno tome stopa po kojoj klijenti koji dolaze vide stanje n je jednaka stopi po kojoj klijenti koje odlaze vide stanje n. Drugim rečima, tvrđenje sledi iz činjenice da stopa ukupnih dolazaka mora biti jednaka stopi ukupnih odlazaka (onaj ko uđe mora jednom izaći).

Teorema 2.5. *Za* $M/M/1$ *model čekanja u redu važi*

$$T : \mathcal{E}(\mu - \lambda).$$

Dokaz: Neka R označava broj klijenata u sistemu kada stigne novi klijent. Tada imamo

$$P\{T \leq t\} = \sum_{n=0}^{\infty} P\{T \leq t \mid R = n\} P\{R = n\}.$$

Štaviše, zbog odsustva memorije što karakteriše eksponencijalnu raspodelu možemo pisati

$$T \mid \{R = n\} : \Gamma(n+1, \mu).$$

S obzirom da klijenti dolaze u skladu sa Poasonovim procesom na osnovu prethodne teoreme imamo

$$P\{R = n\} \equiv \alpha_n = p_n = (1-\rho)\rho^n = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n.$$

Iz toga sledi da

$$P\{T \leq t\} = \sum_{n=0}^{\infty} \left[\int_0^t \mu e^{-\mu s} \frac{(\mu s)^n}{n!} ds\right] \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n = \sum_{n=0}^{\infty} \int_0^t (\mu - \lambda) e^{-\mu s} \frac{(\lambda s)^n}{n!} ds.$$

Kada suma i integral zamene mesta i primenimo Tejlorov red za $e^{\lambda s}$ dobijamo

$$P\{T \leq t\} = \int_0^t (\mu - \lambda) e^{-\mu s} \sum_{n=0}^{\infty} \frac{(\lambda s)^n}{n!} ds = \int_0^t (\mu - \lambda) e^{-\mu s} e^{\lambda s} ds =$$

$$= \int_0^t (\mu - \lambda) e^{-(\mu-\lambda)s} ds = 1 - e^{-(\mu-\lambda)t},$$

iz čega sledi da je *gustina raspodele slučajne promenljive* T jednaka

$$\varphi_T(t) = (\mu - \lambda) e^{-(\mu-\lambda)t} \;\; za\; t \geq 0,$$

odnosno *ukupno vreme koje klijent provede u sistemu* je *eksponencijalna slučajna promenljiva* sa stopom $\mu - \lambda$, to zapisujemo $T : \mathcal{E}(\mu - \lambda)$. ■

Raspodela vremena čekanja, T_q, proizvoljnog klijenta koji dođe u sistem je slučajna promenljiva *mešovitog tipa*. Ako klijent stigne dok je sistem prazan, imamo da je $T_q = 0$. S druge strane, ako postoje $R = n \geq 1$ korisnika u sistemu po dolasku, onda T_q ima $G(n, \mu)$ raspodelu. Pošto $P\{R = 0\} = p_0 = 1 - \rho = 1 - \frac{\lambda}{\mu}$, po postupku kao gore, nalazimo da je

$$P\{T_q \leq t\} = \sum_{n=0}^{\infty} P\{T_q \leq t \mid R = n\} P\{R = n\} =$$

$$= 1 - \frac{\lambda}{\mu} + \left(\frac{\lambda}{\mu}\right)\left(1 - e^{-(\mu-\lambda)t}\right) = 1 - \left(\frac{\lambda}{\mu}\right) e^{-(\mu-\lambda)t}.$$

Tako imamo da je

$$P\{T_q = 0\} = 1 - \frac{\lambda}{\mu}$$

i

$$P\{0 < T_q \leq t\} = \left(\frac{\lambda}{\mu}\right)\left(1 - e^{-(\mu-\lambda)t}\right) \;\; ako\; je\; t > 0.$$

Dalje računamo

$$P\{T_q \leq t \mid T_q > 0\} = \frac{P\{0 < T_q \leq t\}}{P\{T_q > 0\}} = 1 - e^{-(\mu-\lambda)t} \;\; ako\; je\; t > 0.$$

Dobijamo da je

$$T_q \mid \{T_q > 0\} : \ \mathcal{E}(\mu - \lambda).$$

Možemo pisati

$$T_q \mid \{T_q > 0\} = T,$$

što direktno sledi iz činjenice da je

$$P\{R = n \mid R > 0\} = \frac{P\{R = n, R > 0\}}{P\{R > 0\}} = \frac{p_n}{1 - p_0} = \frac{\left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n}{\frac{\lambda}{\mu}} = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^{n-1} =$$

$$= p_{n-1} = P\{R = n - 1\} \quad za\ svako\ n \geq 1.$$

Napomena 4. Kako su po pretpostavci slučajne promenljive T_q i T_s (vreme opsluživanja) nezavisne, činjenicu da T ima eksponencijalnu raspodelu sa parametrom $\mu - \lambda$ možemo proveriti i pomoću funkcije raspodele verovatnoća slučajne promenljive T_q i T_s. S obzirom da je slučajna promenljiva T_q mešovitog tipa, lakše je da koristimo sledeću formulu (gde je $\rho = \frac{\lambda}{\mu}$)

$$F_T(t) = \int_0^\infty F_{T_q}(t - s)\varphi_{T_s}(s)ds = \int_0^t \left(1 - \rho e^{-(\mu-\lambda)(t-s)}\right)\mu e^{-\mu s}\,ds =$$

$$= \int_0^t \left(\mu e^{-\mu s} - \rho\mu e^{-(\mu-\lambda)t-\lambda s}\right)ds = \int_0^t \mu e^{-\mu s}\,ds - \rho\mu\, e^{-(\mu-\lambda)t}\int_0^t e^{-\lambda s}\,ds =$$

$$= 1 - e^{-\mu t} - e^{-(\mu-\lambda)t}\left(1 - e^{-\lambda t}\right) =$$

$$= 1 - e^{-(\mu-\lambda)t} \quad za\ svako\ t \geq 0.$$

Napomenimo da sve dok je server uposlen kada u sistemu ima korisnika, broj korisnika u sistemu ne zavisi od redosleda u kojem se oni poslužuju. Međutim, kod vremena čekanja redosled opsluživanja ima veoma važnu ulogu. U sistemu gde se usluživanje vrši po principu "ko prvi ulazi, prvi izlazi", vreme čekanja na uslugu (T_q) klijenta koji dođe predstavlja količinu vremena koji je potrebno da bi se poslužili svi klijenti koji su već bili u sistemu. Kada u sistemu ima n klijenata, s obzirom da vremena usluge imaju eksponencijalnu raspodelu sa parametrom μ, *ukupno vreme usluge* n *klijenata* ima *Erlangovu raspodelu*[3] *sa gustinom raspodele verovatnoća*

$$\varphi_n(x) = \frac{\mu^n x^{n-1}}{(n-1)!} e^{-\mu x}.$$

Neka je $F_{T_q}(t) = P\{T_q \le t\}$ *funkcija raspodele slučajne promenljive* T_q. Jasno je da

$$F_{T_q}(0) = P\{T_q = 0\} = P\{N = 0\} = 1 - \rho.$$

Napomenimo da zbog nezavisnih priraštaja eksponencijalne raspodele, preostalo vreme usluživanja klijenata u servisu je takođe eksponencijalna sa parametrom μ. Možemo pisati da je

$$dF_{T_q}(t) = P\{t < T_q \le t + dt\} \ za \ t > 0.$$

Imamo

$$dF_{T_q}(t) = \sum_{n=1}^{\infty} p_n \frac{\mu^n t^{n-1}}{(n-1)!} e^{-\mu t} \, dt = (1-\rho) \sum_{n=1}^{\infty} \rho^n \frac{\mu^n t^{n-1}}{(n-1)!} e^{-\mu t} dt =$$

$$= \lambda(1-\rho) \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\mu t} dt = \lambda(1-\rho) e^{\lambda t} e^{-\mu t} dt =$$

$$= \lambda(1-\rho) e^{-(\mu-\lambda)t} dt = \lambda(1-\rho) e^{-\mu(1-\rho)t} dt.$$

Zbog prekida u 0 kod raspodele T_q, imamo sledeće

$$F_{T_q}(t) = P\{T_q = 0\} + \int_0^t d\, F_{T_q}(t) = 1 - \rho e^{-\mu(1-\rho)t}. \tag{2.25}$$

Neka je $E(T_q) = W_q$ i $E(T) = W$. Iz (2.25) lako možemo izraziti o*čekivanje* i *disperziju za* T_q

$$W_q = E(T_q) = \frac{\rho}{\mu(1-\rho)} = \frac{\lambda}{\mu(\mu-\lambda)} \tag{2.26}$$

i

$$D(T_q) = \frac{\rho(2-\rho)}{\mu^2(1-\rho)^2}. \tag{2.27}$$

---

[3] *Erlangova raspodela* je specijalan slučaj gama raspodele, gde je parametar k ceo broj. Gustina raspodele verovatnoća je $\varphi(x) = (\lambda^k x^{k-1} e^{-\lambda x})/(k-1)!$, gde je $k > 0, \lambda > 0 \ i \ k \in \mathbb{Z}$.

Pošto je *ukupno vreme provedeno u sistemu*, T, jednak vremenu provedenom u redu uvećanim za vreme usluživanja, tj. zbiru T_q i T_s, za sve redove čekanja važi relacija:

$$W = W_q + W_s,$$

odnosno

$$W = E(T) = \frac{\lambda}{\mu(\mu-\lambda)} + \frac{1}{\mu} = \frac{1}{\mu-\lambda}. \tag{2.28}$$

Ako uporedimo (2.20) i (2.28), primećujemo sledeću vezu

$$L = \lambda W. \tag{2.29}$$

Slično, poređenjem (2.21) i (2.26) možemo utvrditi sledeću povezanost

$$L_q = \lambda W_q.$$

Rezultat (2.29) je poznat kao Little-ov zakon. Little[4] je 1961. godine pokazao da skoro kod svakog stacionarnog sistema čekanja postoji jedna jednostavna veza između prosečnog broja klijenata u sistemu, prosečnog vremena čekanja i stopa dolazaka.

Osobina 1. (*Little-ov zakon) Posmatrajmo jedan sistem čekanja u redu koji je u ravnotežnom stanju. Neka* $L = E(N)$ *označava dugoročan prosečan broj klijenata/poslova u sistemu,* $W = E(T)$ *dugoročno prosečno vreme čekanja u sistemu, a* λ_e *označava prosečnu stopu pristizanja klijenata/poslova u sistem. Dalje, neka* $L_q = E(N_q)$ *i* $W_q = E(T_q)$ *označavaju analogne količine koje se odnose na red. Tada*

$$L = \lambda_e W$$

i

$$L_q = \lambda_e W_q.$$

λ_e se odnosi na *efektivnu stopu dolaska* klijenata u sistem, dok se λ odnosi na *prosečnu stopu dolaska* u sistem. Drugim rečima, λ uključuje one klijente koji dođu do sistema, ali iz nekog razloga ne uđu u njega, npr. u slučaju da je sistem sa konačnim kapacitetom pun. λ_e broji samo one klijente koji dođu na server i budu usluženi. Kod $M/M/1$ sistema efektivna stopa dolazaka je ista kao i prosečna stopa dolazaka (tj. $\lambda_e = \lambda$).

Kada su dolasci u skladu sa Poasonovim procesom, postoji *uopštenje Little-ove formule* koja važi za *disperzije*.

---

[4] John D.C. Little, profesor na Masačusetskom institutu za tehnologiju u Sjedinjenim Državama.

Osobina 2. *Posmatrajmo jedan sistem čekanja u redu koji je u ravnotežnom stanju sa Poasonovim tokom pristizanja korisnika koji ulaze u sistem. Neka N označava broj klijenata u sistemu, T označava vreme čekanja u sistemu, a λ_e prosečnu stopu pristizanja klijenata u sistem. Takođe, neka N_q i T_q označavaju analogne veličine koje se odnose na red. Tada važi:*

$$D(N) - E(N) = \lambda_e^2 D(T)$$

i

$$D(N_q) - E(N_q) = \lambda_e^2 D(T_q).$$

Little-ov zakon (Osobina 1.) je veoma moćan rezultat zbog svoje opštosti . Može se korisiti za skoro sve modele redova čekanja bez obzira na proces dolaženja, broj servera, ili uređenje sistema. Međutim, to nije slučaj kod verzije primenjene na disperzije (Osobina 2.) zbog ograničenja na Poasonov dolazak. Primenom Osobine 2. na $M/M/1$ sistem, dobijamo

$$D(T) = \frac{D(N) - E(N)}{\lambda^2} = \frac{\frac{\rho}{(1-\rho)^2} - \frac{\rho}{1-\rho}}{\lambda^2} = \frac{\rho^2}{\lambda^2(1-\rho)^2} =$$
$$= \frac{1}{\mu^2(1-\rho)^2} = \frac{1}{(\mu-\lambda)^2}$$

i

$$D(T_q) = \frac{D(N_q) - E(N_q)}{\lambda^2} = \frac{\frac{\rho^2(1+\rho-\rho^2)}{(1-\rho)^2} - \frac{\rho^2}{1-\rho}}{\lambda^2} = \frac{\frac{2\rho^3-\rho^4}{(1-\rho)^2}}{\lambda^2} =$$
$$= \frac{2\rho-\rho^2}{(\mu-\lambda)^2} = \frac{\rho(2-\rho)}{\mu^2(1-\rho)^2}.$$

Osobina 3. *Posmatrajmo jedan sistem čekanja u redu koji je u ravnotežnom stanju. Označimo sa W dugoročno prosečno vreme čekanja u sistemu, sa W_q dugoročno prosečno vreme čekanja u redu i sa μ prosečnu stopu usluživanja. Tada važi*

$$W = W_q + \frac{1}{\mu}.$$

Za ovu osobinu važi isto što i za Little-ov zakon, odnosno ona važi i za neeksponencijalne sisteme, sisteme sa više servera kao i za sisteme sa konačnim i beskonačnim kapacitetom.

Primer 2.2. Na aerodromu postoji samo jedna pista. Avioni pristižu brzinom od 15 letelica na sat. Procenjuje se da svako sletanje traje 3 minuta. Uzimajući da pristizanja predstavljaju Poasonov proces, a vremena sletanja imaju eksponencijalnu raspodelu, na osnovu $M/M/1$ modela odredimo sledeće karakteristike.

(a) Iskorišćenost piste:

$$\lambda = stopa\, pristizanja = 15/č,$$

$$\mu = stopa\, usluživanja = \frac{60}{3}/č = 20/č,$$

$$\rho = stopa\, iskorišćenosti = \frac{\lambda}{\mu} = 0.75.$$

(b) Očekivani broj aviona koji čekaju na sletanje:

$$L_q = E\left(N_q\right) = \frac{\rho^2}{1-\rho} = \frac{(0.75)^2}{0.25} = 2.25.$$

(c) Očekivano vreme čekanja na sletanje:

$$W_q = E\left(T_q\right) = \frac{\lambda}{\mu(\mu-\lambda)} = \frac{15}{20(20-15)} = \frac{3}{20}\ časa = 9\ minuta.$$

(d) Verovatnoća da će čekanje biti više od 5 minuta? 10 minuta? Nema čekanja?

$$P\left\{T_q \le t\right\} = 1 - \rho e^{-(\mu-\lambda)t}$$

$$\Rightarrow \quad P\left\{T_q > t\right\} = \rho e^{-\mu(1-\rho)t},$$

$$P\left\{T_q > 5\ minuta\right\} = 0.75 e^{-20(1-0.75)\frac{5}{60}} =$$

$$= 0.75 e^{-\frac{25}{60}} = 0.4944,$$

$$P\left\{T_q > 10\ minuta\right\} = 0.75 e^{-20(1-0.75)\frac{10}{60}} =$$

$$= 0.75 e^{-\frac{50}{60}} = 0.3259,$$

$$P\{nema\ čekanja\} = P\left\{T_q = 0\right\} = 1 - \rho = 1 - 0.75 = 0.25.$$ □

2.3.2 Sistemi sa jednim serverom i konačnim kapacitetom ($M/M/1/c$ model)

U prethodnom modelu smo pretpostavili da nema granične vrednosti za broj klijenata u sistemu. Iako je $M/M/1$ model čekanja u redu veoma koristan za modeliranje različitih fenomena, ali pretpostavka o beskonačnom kapacitetu često ne odgovara stvarnosti i mnogo je realističnije pretpostaviti da je kapacitet sistema neki ceo broj $c < \infty$. U realnom svetu uvek postoji ograničen kapacitet sistema c, u smislu da ne može biti više od c klijenata u sistemu u bilo koje vreme. Ovim se misli na to da ako klijent koji dolazi zatekne stanje u sistemu gde je već prisutno c klijenata onda neće ući u sistem. Sva takva pristizanja u sistem, gde je prisutno tačno c klijenata, bivaju zaustavljena, a ti klijenti izgubljeni za sistem.

Prema tome i ovaj sistem može biti modeliran kao proces rađanja/umiranja, pri čemu su u ovom slučaju odgovarajući parametri:

$$\lambda_n = \lambda \quad za\ n = 0, 1, 2, \dots, c-1$$

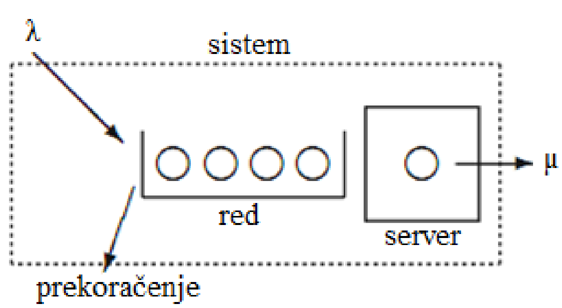
$$\lambda_c = 0$$

$$\mu_0 = 0$$

$$\mu_n = \mu \quad za\ n = 1, 2, \dots, c.$$

Budući da je $\lambda_c = 0$, stanje sistema nikad neće biti $c + 1$.

U slučaju kada je neophodno koristiti sistem sa konačnim kapacitetom , verovatnoće stanja i mere efikasnosti predstavljene u jednačinama (2.19), (2.20), (2.21), (2.23), (2.24), (2.26) i (2.28) su neodgovarajuće , i zato se moraju odrediti nove verovatnoće i mere efikasnosti za $M/M/1/c$ sistem.



Slika 2.4. Prikaz $M/M/1/5$ sistema sa punim kapacitetom.

Jednačine za određivanje verovatnoće stanja su identične kao i kod $M/M/1$ sistema, jedino se normirana jednačina razlikuje kao što ćemo i videti. Kada se sistem nalazi u stanju c, on iz njega samo može da izađe pošto korisnici koji su posluženi odlaze. Osim toga, u ovo stanje moguće je ući samo iz stanja $c-1$, po dolasku novog korisnika. Ako je c maksimalni mogući broj klijenata u sistemu, tada je *generator matrica* dimenzije $(c+1)\times(c+1)$ i ima sledeći oblik

$$A=\begin{bmatrix} -\lambda & \lambda & & & \\ \mu & -(\lambda+\mu) & \lambda & & \\ & \ddots & \ddots & \ddots & \\ & & \mu & -(\lambda+\mu) & \lambda \\ & & & \mu & -\mu \end{bmatrix}.$$

Sistem jednačina $\boldsymbol{p}A=\boldsymbol{0}$ nam daje *ravnotežne jednačine*:

$$\mu p_1=\lambda p_0$$

$$\mu p_{n+1}=(\lambda+\mu)p_n-\lambda p_{n-1} \quad za\ n=1,2,\dots,c-1$$

$$\mu p_c=\lambda p_{c-1}, \tag{2.30}$$

do kojih možemo doći i primenom principa jednakosti stopa, tj. *"stopa izlaska" ="stopa ulaska"*. Argument za stanje 0 je isti kao i kod $M/M/1$ sistema. Naime, kada je u stanju 0 sistem će napustiti to stanje samo po dolasku klijenta (koji se dešava po stopi λ), pa je stopa po kojoj sistem napušta stanje 0 jednaka λp_0. Drugim rečima, proces može ući u stanje 0 samo iz stanja 1 po odlasku klijenta; dakle stopa po kojoj proces ulazi u stanje 0 je μp_1. Jednačina za n-to stanje, gde je $1\le n<c$ je ista kao i ranije. Jednačina za stanje c je drugačija, jer se sada u stanje c može ući samo iz stanja $c-1$ po dolasku klijenta.

Sistem jednačina (2.30) se može rešiti sukcesivnom supstitucijom rešenja i izražavanjem svih promenljivih pomoću p_0, pa dobijamo

$$p_n=\rho^n p_0 \quad za\ n=0,1,\dots,c,$$

gde je $\rho=\frac{\lambda}{\mu}$. Poslednja jednačina se kod (2.30) može izostaviti, jer se kod konačnog nesvodljivog sistema Markova uvek javlja jedna suvišna jednačina. Koristeći činjenicu da je $\sum_{n=0}^{c} p_n=1$ dobijamo:

$$1=\sum_{n=0}^{c} p_n=p_0\sum_{n=0}^{c}\rho^n=\begin{cases} p_0\dfrac{1-\rho^{c+1}}{1-\rho} & za\ \rho\neq 1,\\ p_0(c+1) & za\ \rho=1,\end{cases}$$

odavde dobijamo da je

$$p_0 = \begin{cases} \dfrac{1-\rho}{1-\rho^{c+1}} & za\ \rho \neq 1, \\ \dfrac{1}{c+1} & za\ \rho = 1, \end{cases}$$

S obzirom da je gornja suma konačna, koristili smo konačnu geometrijsku progresiju koja dopušta da ρ bude veća od 1. Stoga, za $M/M/1/c$ sistem dobijamo

$$p_n = \begin{cases} \rho^n \dfrac{1-\rho}{1-\rho^{c+1}} & za\ \rho \neq 1, \\ \dfrac{1}{c+1} & za\ \rho = 1, \end{cases} \qquad za\ n = 0, 1, \dots, c.$$

Napomena 5.
(a) Ako je $\lambda = \mu$ (tj. $\rho = 1$), tada su sva moguća $c + 1$ stanja sistema koja je u ravnoteži jednako verovatna. Štaviše, kada c teži beskonačnosti , verovatnoća p_n teži u 0 za svaki konačan broj n. Ovo potvrđuje i činjenica da u $M/M/1/\infty$ modelu dužina reda se povećava beskonačno kada je $\lambda = \mu$, tako da nema stacionarnog stanja.

(b) Ako je $\lambda > \mu$, tada postoje granične verovatnoće. Međutim, što je količnik $\rho = \frac{\lambda}{\mu}$ veći, to će p_c više težiti ka 1 (a p_n težiti ka 0 za $n = 0, 1, \dots, c-1$), što je i logično .

(c) Iako u stvarnom životu kapacitet c ne može biti beskonačan, $M/M/1$ model predstavlja dobru aproksimaciju stvarnosti ako verovatnoća p_c, da je sistem pun, je veoma mala.

(d) Za $M/M/1/c$ redove čekanja ravnotežna stanja će postojati čak i ukoliko je $\rho \geq 1$. Dakle, u ovom slučaju nije potrebno postavljati uslov $\frac{\lambda}{\mu} < 1$, jer je veličina reda po definiciji konačna pa ne postoji mogućnost da se red povećava u beskonačnost (konačni kapacitet sistema kontroliše nagomilavanje klijenata prisutnih u sistemu).

Teorema 2.6. *U slučaju $M/M/1/c$ modela prosečan broj klijenata u sistemu koji je u ravnotežnom stanju je dato sa*

$$L = E(N) = \begin{cases} \dfrac{\rho}{1-\rho} - \dfrac{(c+1)\rho^{c+1}}{1-\rho^{c+1}} & ako\ je\ \rho \neq 1 \\ \dfrac{c}{2} & ako\ je\ \rho = 1 \end{cases} =$$

$$= \begin{cases} \rho \dfrac{1 + c\rho^{c+1} - (c+1)\rho^c}{(1-\rho)(1-\rho^{c+1})} & ako\ je\ \rho \neq 1 \\ \dfrac{c}{2} & ako\ je\ \rho = 1\,. \end{cases}$$

Dokaz: Kao i ranije L se može izraziti preko p_n da bismo dobili:

$$L = E(N) = \sum_{n=0}^{c} np_n = \sum_{n=1}^{c} np_n = p_0\rho \sum_{n=1}^{c} n\rho^{n-1}.$$

Prvo razmatramo slučaj kada je $\rho = 1$, tada imamo

$$L = \sum_{n=0}^{c} n\frac{1}{c+1} = \frac{1}{c+1}\sum_{n=1}^{c} n = \frac{1}{c+1}\,\frac{c(c+1)}{2} = \frac{c}{2},$$

gde smo koristili formulu za izračunavanje zbira prvih c članova aritmetičkog niza.

Kada je $\rho \neq 1$, moramo da odredimo konačnu sumu

$$L = \sum_{n=0}^{c} n\frac{\rho^n(1-\rho)}{1-\rho^{c+1}} = \frac{1}{1-\rho^{c+1}}\sum_{n=0}^{c} n\,\rho^n(1-\rho).$$

Definišimo $X = Z - 1$, gde slučajna promenljiva Z ima geometrijsku raspodelu sa parametrom $1-\rho$. Tada je raspodela verovatnoća za X dato sa

$$p_X(n) = P\{X = n\} = \rho^n(1-\rho) \quad za\; n = 0, 1, \dots\ .$$

Sledi

$$\sum_{n=0}^{\infty} p_X(n) = \sum_{n=0}^{\infty} \rho^n(1-\rho) = 1$$

i

$$\sum_{n=0}^{\infty} n\rho^n(1-\rho) = \rho(1-\rho)\sum_{n=0}^{\infty} n\rho^{n-1} = \rho(1-\rho)\left(\sum_{n=0}^{\infty}\rho^n\right)' = \rho(1-\rho)\left(\frac{1}{1-\rho}\right)' = \frac{\rho}{1-\rho}.$$

Koristeći ove formule, možemo pisati da je

$$\sum_{n=0}^{c} n\rho^n(1-\rho) = \frac{\rho}{1-\rho} - \sum_{n=c+1}^{\infty} n\rho^n(1-\rho)$$

i

$$\sum_{n=c+1}^{\infty} n\rho^n(1-\rho) = \sum_{n=c+1}^{\infty} [n-(c+1)+(c+1)]\,\rho^n(1-\rho) =$$

$$= \sum_{n=c+1}^{\infty} [n-(c+1)]\,\rho^n(1-\rho) + \sum_{n=c+1}^{\infty} (c+1)\rho^n(1-\rho) =$$

$$= \rho^{c+1}\sum_{m=0}^{\infty} m\rho^m(1-\rho) + (c+1)\rho^{c+1}\sum_{m=0}^{\infty} \rho^m(1-\rho) =$$

$$= \rho^{c+1}(1-\rho)\rho\sum_{m=0}^{\infty} m\rho^{m-1} + (c+1)\rho^{c+1}(1-\rho)\sum_{m=0}^{\infty} \rho^m =$$

$$= \rho^{c+2}(1-\rho)\left(\sum_{m=0}^{\infty} \rho^m\right)' + (c+1)\rho^{c+1}(1-\rho)\frac{1}{1-\rho} =$$

$$= \rho^{c+2}(1-\rho)\left(\frac{1}{1-\rho}\right)' + (c+1)\rho^{c+1} =$$

$$= \rho^{c+2}\frac{1}{1-\rho} + (c+1)\rho^{c+1} = \rho^{c+1}\frac{\rho}{1-\rho} + (c+1)\rho^{c+1}.$$

Imamo da je

$$L = \frac{1}{1-\rho^{c+1}}\left[\frac{\rho}{1-\rho} - \rho^{c+1}\frac{\rho}{1-\rho} - (c+1)\rho^{c+1}\right] =$$

$$= \frac{\rho - \rho^{c+1}\rho - (c+1)\rho^{c+1}(1-\rho)}{(1-\rho)(1-\rho^{c+1})} =$$

$$= \rho\,\frac{1 + c\rho^{c+1} - (c+1)\rho^c}{(1-\rho)(1-\rho^{c+1})}.$$

■

Prosečan broj klijenata u redu je

$$L_q = E(N_q) = \sum_{n=1}^{c-1} np_{n+1} = p_0\rho^2\sum_{n=1}^{c-1} n\rho^{n-1} =$$

$$= \begin{cases} L - \dfrac{\rho(1-\rho^c)}{1-\rho^{c+1}} & ako\ je\ \rho \neq 1 \\[2ex] \dfrac{c(c-1)}{2(c+1)} & ako\ je\ \rho = 1\,. \end{cases}$$

Napomena 6.

(a) Lako nalazimo da je

$$\lim_{c\to\infty} L = \begin{cases} \dfrac{\rho}{1-\rho} & ako\ je\ \rho < 1 \\ \infty & ako\ je\ \rho \geq 1 \end{cases}$$

što odgovara rezultatima dobijenim u prethodnom delu.

(b) Kada je kapacitet c sistema veoma ograničen (na primer, kada je $c = 2, 3$ ili 4), tada ako jednom izračunamo verovatnoće p_n, iz definicije matematičkog očekivanja diskretne slučajne promenljive možemo direktno dobiti N:

$$L = E(N) = \sum_{n=0}^{c} n p_n.$$

Možemo izračunati i *disperziju slučajne promenljive* N:

$$D(N) := \sum_{n=0}^{c} \big(n - E(N)\big)^2 p_n = \sum_{n=0}^{c} n^2\, p_n - \big(E(N)\big)^2.$$

(c) Kada je $\rho = 1$, *disperziju slučajne promenljive* N nalazimo koristeći formulu:

$$\sum_{n=0}^{c} n^2 = \frac{c(c+1)(2c+1)}{6}.$$

Imamo da je

$$E(N^2) := \sum_{n=0}^{c} n^2 p_n = \sum_{n=0}^{c} n^2 \frac{1}{c+1} = \frac{c(2c+1)}{6},$$

tako da je

$$D(N) = \frac{c(2c+1)}{6} - \left(\frac{c}{2}\right)^2 = \frac{c(c+2)}{12}.$$

Dakle, *disperzije* za ove količine u slučaju $M/M/1/c$ sistem su

$$D(N) =$$

$$= \begin{cases} \dfrac{\rho}{(1-\rho^{c+1})(1-\rho)^2}\left[1+\rho-(c+1)^2\rho^c+(2c^2+2c-1)\rho^{c+1}-c^2\rho^{c+2}\right]-L^2 & za\ \rho \neq 1 \\ \dfrac{c(c+2)}{12} & za\ \rho = 1, \end{cases}$$

$$D(N_q) = D(N) - p_0(L + L_q).$$

Određivanje W i W_q se u ovom slučaju donekle razlikuje nego kod M/M/1 redova. Kod M/M/1 sistema λ, koja se pojavljuje u teoremi, predstavlja prosečan broj klijenata po jedinici vremena koje pristignu u sistem. Međutim, kod sistema sa konačnim kapacitetom klijenata, λ takođe predstavlja prosečan broj klijenata po jedinici vremena, ali λp_c od tih klijenata zatiče sistem popunjen i odlazi. Prema tome, prosečno $\lambda - \lambda p_c = \lambda(1 - p_c)$ klijenata po jedinici vremena će u stvari pristići u sistem.

Dakle, *prosečna stopa ulaska* (*efektivna stopa dolaska*) klijenata u sistem koji je u ravnoteži je data sa

$$\lambda_e = \lambda(1 - p_c),$$

jer klijenti uvek pristižu u skladu sa Poasonovim procesom sa stopom λ, ali mogu da uđu u sistem samo kada on nije pun, to je slučaj kada je sistem u jednom od sledećih stanja: $0, 1, \ldots, c-1$. Koristeći ove činjenice i Little-ovu formulu $L = \lambda_e W$, možemo napisati da je *prosečno vreme koje će klijent provesti u sistemu i u redu* jednaka

$$W = \frac{L}{\lambda(1 - p_c)}$$

i

$$W_q = W - \frac{1}{\mu}.$$

Izraz za W_q smo dobili iz Osobine 3. Iskušenje je da se koriste formule iz Osobine 2. za dobijanje izraza za disperziju vremena čekanja. Međutim, odnosi disperzija po principu "kao Little" zasnovani su na pretpostavci da su dolasci u sistem Poasonove raspodele. Zbog ograničenja u kapacitetu isključeni su Poasonovi dolasci u sistem, pa se generalizacija za disperziju iz Little-ovog zakona ne može koristiti.

Primer 2.3. Jedna korporacija treba da održava jedan veliki vozni park traktora. Ima i jednog mehaničara koji popravlja pokvarene traktore po redosledu njihovog dolaska u radionicu. Dolazak traktora u radionicu procenjuje se na osnovu Poasonove raspodele, i to u proseku 3 traktora nedeljno. Vreme trajanja popravki varira prema eksponencijalnoj raspodeli, sa tim da je prosek trajanja popravke pola nedelje po traktoru. Korporacija je odlučila da u slučajevima kada u njenoj radionici stoji više od dva traktora unajmi jednu spoljašnju radionicu, čime je broj traktora na čekanju sveden na jedan. Svaka nedelja koju traktor provede u radionici kompaniju košta $100\,dolara$. Unajmljivanje spoljašnje radionice košta $500\,dolara$ po traktoru (u šta je uračunato i izgubljeno vreme). Želimo da proverimo poslovanje korporacije i odredimo optimalnu graničnu tačku na kojoj će se unajmiti jedna spoljašnja radionica. Drugim rečima, želimo da odredimo maksimalan broj traktora u radionici kompanije pre nego što se oni pošalju u neku spoljašnju radionicu. Ukupni nedeljni troškovi su

$$Trošak_c = 100L + 500\lambda p_c.$$

Dalje, imamo da je iskorišćenost radionice

$$\begin{aligned}
&\lambda = stopa\, pristizanja = 3/nedelja,\\
&\mu = stopa\, usluživanja = 2/nedelja,\\
&\rho = stopa\, iskorišćenosti = \tfrac{\lambda}{\mu} = \tfrac{3}{2} = 1.5.
\end{aligned}$$

Za ovakvo poslovanje, što predstavlja jedan $M/M/1/2$ sistem, imamo da je

$$L = E(N) = \rho\frac{1 + c\rho^{c+1} - (c+1)\rho^c}{(1-\rho)(1-\rho^{c+1})} = 1.5\frac{1 + 2\cdot 1.5^3 - 3\cdot 1.5^2}{-0.5\cdot(1-1.5^3)} = 1.26$$

i

$$p_2 = \rho^2\frac{1-\rho}{1-\rho^3} = 1.5^2\frac{-0.5}{1-3.375} = 0.474,$$

pa sledi

$$Trošak_{c=2} = 100\cdot 1.26 + 500\cdot 3\cdot 0.474 = 837\, dolara/nedelja.$$

U slučaju da kompanija dozvoli da u sistemu budu tri traktora, tada imamo da je $L = 1.98$ i $p_3 = 0.415$, što daje da je

$$Trošak_{c=3} = 100\cdot 1.98 + 500\cdot 3\cdot 0.415 = 820\, dolara/nedelja.$$

U slučaju da su dozvoljena četiri traktora u sistemu, imamo da je $L = 2.76$ i $p_4 = 0.384$, što daje da je

$$Trošak_{c=4} = 100\cdot 2.76 + 500\cdot 3\cdot 0.384 = 852\, dolara/nedelja.$$

Dakle, preporuka je da se traktor pošalje u spoljašnju radionicu tek kada je više od tri traktora u sistemu.

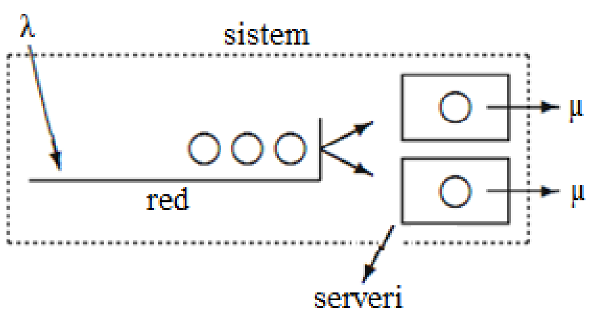
□

2.4 Sistemi masovnog opsluživanja sa više servera

2.4.1 Sistemi sa s serverom i beskonačnim kapacitetom ($M/M/s/\infty$ model)

Do važne generalizacije $M/M/1$ modela dolazimo kada pretpostavimo da u sistemu postoji *s servera* (koji usluge pružaju nezavisno jedni od drugih) i da svi oni poslužuju *eksponencijalnom stopom* μ. Klijenti stižu u sistem u skladu sa *Poasonovim procesom* sa stopom λ. *Kapacitet sistema* je *beskonačan* i usluživanje se vrši po principu *"ko prvi dođe, prvi ode"*.

Višeserverski $M/M/s$ sistem masovnog posluživanja predstavlja model koji se najviše koristi u analizi uslužnih stanica sa više od jednog servera, kao što su banke, kase u marketima, mesta čekiranja karata na aerodromima i slično. Pretpostavljamo da klijenti koji pristižu formiraju *jedan red čekanja* i da korisnik koji je sledeći na redu prilazi serveru koji prvi postane slobodan. Dok ima korisnika, ni jedan server ne stoji besposlen. Jasno je da je sistem sa ovakvom disciplinom čekanja efikasniji nego sistem u kojem ispred svakog servera stoji red, pošto u tom slučaju neki serveri će možda da stoje besposleno, dok ispred drugih korisnici stoje u redu.



Slika 2.5. Prikaz jednog M/M/2 sistema sa jednim redom čekanja i dva servera.

Napomena 7. Ne moraju se formirati pravi redovi čekanja u kojima korisnici stoje jedni iza drugih. Dovoljno je da korisnici koji pristižu u sistem uzmu broj, ili da se zadaci koje serveri obavljaju numerišu prema njihovom prispeću u sistem.

S obzirom da klijenti dolaze jedan po jedan i bivaju usluženi jedan po jedan, proces $\{X(t), t > 0\}$, gde $X(t)$ predstavlja *broj klijenata u sistemu u trenutku* t, je jedan *proces rađanja i umiranja* sa sledećim stopama za rođenje i umiranje:

$$\begin{aligned} \lambda_n &= \lambda \quad za \;\; n = 0, 1, \dots, \\ \mu_n &= \begin{cases} n\mu & za \;\; n = 1, 2, \dots, s-1 \\ s\mu & za \;\; n = s, s+1, \dots . \end{cases} \end{aligned} \tag{2.30}$$

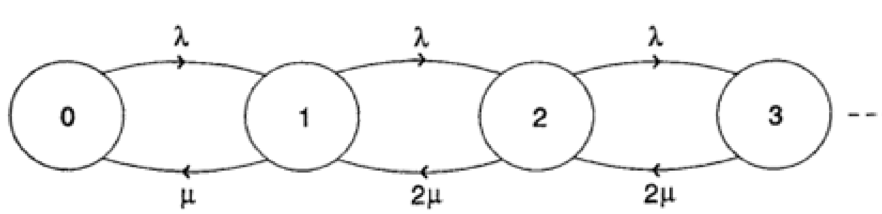
Razlog zbog kojeg je $\mu_n = n\mu$ za $n = 1, \dots, s-1$ je taj da kada u sistemu postoji manje od s klijenata, svaki klijent biva poslužen, pa će tako stopa posluživanja biti jednaka broju korisnika, pošto će serveri na kojima nema korisnika biti besposleni (tj. slobodni serveri ne pomažu zauzetim serverima). Ako postoji više od s klijenata u sistemu, tada je tačno s servera zauzet i stoga stopa usluživanja mora bit $s\mu$. Napomenimo da dva proizvoljna klijenta ne mogu napustiti sistem u istom vremenskom trenutku, jer su vremena opsluživanja neprekidne slučajne promenljive.

Generator matrica A za proces je data sa

$$A = \begin{array}{c} 0 \\ 1 \\ \vdots \\ s \\ s+1 \\ \vdots \end{array} \begin{bmatrix} -\lambda & \lambda & & & & & \\ \mu & -(\lambda+\mu) & \lambda & & & & \\ & & \ddots & & & & \\ & & & s\mu & -(\lambda+s\mu) & \lambda & \\ & & & & s\mu & -(\lambda+s\mu) & \lambda \\ & & & & & \ddots & \end{bmatrix}. \tag{2.31}$$

Ravnotežne jednačine sistema su sledeće:

stanje i	*stopa po kojoj proces napušta stanje i* ("stopa izlaska")	=	*stopa po kojoj proces ulazi u stanje i* ("stopa ulaska")
0	λp_0	$=$	μp_1
$0 < n < s$	$(\lambda + n\mu)p_n$	$=$	$\lambda p_{n-1} + (n+1)\mu p_{n+1}$ (2.32)
$s \le n < \infty$	$(\lambda + s\mu)p_n$	$=$	$\lambda p_{n-1} + s\mu p_{n+1}$.



Slika 2.6. Dijagram prelaza između stanja za $M/M/2$ model.

Slično kao i u slučaju M/M/1 sistema pomoću rekurzije dobijamo sledeće:

$$n\mu p_n = \lambda p_{n-1} \quad za\ n = 1, 2, \dots, s,$$

$$s\mu p_n = \lambda p_{n-1} \quad za\ n = s+1, s+2, \dots .$$

Dakle, imamo da je

$$p_n = \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n p_0, \quad 0 \leq n \leq s,$$

$$p_{s+r} = \left(\frac{\lambda}{s\mu}\right)^r p_s, \quad r = 0, 1, 2, \dots, \tag{2.33}$$

$$p_n = \left(\frac{\lambda}{s\mu}\right)^{n-s} p_s, \quad n = s, s+1, \dots .$$

Neka je $\rho = \frac{\lambda}{s\mu}$, tada imamo da je

$$p_n = \frac{1}{n!} s^n \left(\frac{\lambda}{s\mu}\right)^n p_0 = \frac{1}{n!} s^n \rho^n p_0, \quad 0 \leq n < s$$

$$p_n = \frac{1}{s!}\left(\frac{\lambda}{\mu}\right)^s \left(\frac{\lambda}{s\mu}\right)^{n-s} p_0 = \frac{1}{s!} s^s \left(\frac{\lambda}{s\mu}\right)^s \left(\frac{\lambda}{s\mu}\right)^{n-s} p_0, \quad s \leq n < \infty.$$

Odnosno,

$$p_n = \begin{cases} \frac{1}{n!}(s\rho)^n p_0, & 0 \leq n \leq s, \\ \frac{1}{s!}(s\rho)^s \rho^{n-s} p_0, & s \leq n < \infty. \end{cases} \tag{2.34}$$

Koristeći uslov $\sum_{n=0}^{\infty} p_n = 1$, iz (2.34) sledi

$$p_0 = \left(1 + \sum_{n=1}^{s} \frac{(s\rho)^n}{n!} + \sum_{n=s+1}^{\infty} \frac{1}{s!}(s\rho)^s \rho^{n-s}\right)^{-1} =$$

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

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

$$= \left(\sum_{n=0}^{s-1} \frac{(s\rho)^n}{n!} + \frac{1}{s!}(s\rho)^s \sum_{r=0}^{\infty} \rho^{r} \right)^{-1} =$$

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

$$p_n = \begin{cases} \dfrac{(s\rho)^n}{n!} p_0, & 0 \le n \le s, \\ \dfrac{s^s \rho^n}{s!} p_0, & s \le n < \infty. \end{cases} \tag{2.35}$$

pod uslovom da je $\rho = \frac{\lambda}{s\mu} < 1$. Pošto je $s\mu$ maksimalna stopa usluživanja, kako smo i gore definisali ρ možemo posmatrati kao i*ntenzitet dolazaka* (*saobraćaja*) za sistem.

Ako uvrstimo $\rho = \frac{\lambda}{s\mu}$ u (2.33) dobijamo

$$p_n = \rho^{n-s} p_s, \quad n \ge s,$$

i možemo reći da kada j e broj korisnika u sistemu veći ili jednak s, sistem se ponaša kao $M/M/1$ sistem sa stopom usluživanja $s\mu$. Možemo pisati da je $\alpha = \frac{\lambda}{\mu}$, tada je $\frac{\alpha}{s} = \rho$. Alternativni oblik za (2.35) koristeći α je dato sa

$$p_0 = \left[\sum_{n=0}^{s-1} \frac{\alpha^n}{n!} + \frac{\alpha^s}{s!} \left(1 - \frac{\alpha}{s}\right)^{-1} \right]^{-1},$$

$$p_n = \begin{cases} \dfrac{\alpha^n}{n!} p_0, & 0 \le n \le s, \\ \dfrac{\alpha^s}{s!} \left(\dfrac{\alpha}{s}\right)^{n-s} p_0, & s \le n < \infty. \end{cases} \tag{2.36}$$

Klijenti će morati da sačekaju na usluživanje ako je broj klijenata u sistemu veći ili jednak s. Verovatnoća ovog događaja je data sa $\sum_{n=s}^{\infty} p_n$, stoga

$$P\{zadržavanje\ klijenta\} = P\{N \geq s\} = C(s,\alpha) =$$

$$= \sum_{n=s}^{\infty} p_n = \sum_{n=s}^{\infty} \frac{\alpha^s}{s!}\left(\frac{\alpha}{s}\right)^{n-s} p_0 = \frac{\alpha^s}{s!} p_0 \left(\frac{\alpha}{s}\right)^{-s} \sum_{n=s}^{\infty} \left(\frac{\alpha}{s}\right)^{n} =$$

$$= \frac{\alpha^s}{s!} p_0 \left(\frac{\alpha}{s}\right)^{-s} \sum_{m=0}^{\infty} \left(\frac{\alpha}{s}\right)^{m+s} = \frac{\alpha^s}{s!} p_0 \frac{1}{1-\frac{\alpha}{s}} =$$

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

odnosno

$$C(s,\alpha) = \frac{\frac{\alpha^s}{s!}}{(1-\rho)\sum_{n=0}^{s-1} \frac{\alpha^n}{n!} + \frac{\alpha^s}{s!}}.$$

Može se lako pokazati da je

$$p_0 = \frac{s!\,(1-\rho)C(s,\alpha)}{\alpha^s}.$$

Formula za $C(s,\alpha)$ u literaturi je poznata kao *Erlangova formula zadržavanja* ili kao *Druga Erlangova formula*[5], i često se označava sa $E_{2,s}(\alpha)$.

Ako sa L i L_q označimo *prosečan broj klijenata u sistemu i u redu*, respektivno, tada možemo da ih izvedemo na sledeći način : koristeći izraz iz (2.34) i ako znamo da je $s\rho = \alpha$, dobijamo

$$\begin{aligned} L = E(N) = \sum_{n=1}^{\infty} np_n \quad &= p_0 \left[\sum_{n=1}^{s} n\frac{\alpha^n}{n!} + \sum_{n=s+1}^{\infty} n\rho^{n-s}\frac{\alpha^s}{s!}\right] = \\ &= p_0 \left[\alpha \sum_{n=1}^{s} \frac{\alpha^{n-1}}{(n-1)!} + \frac{\alpha^s}{s!} \sum_{n-s=1}^{\infty} n\rho^{n-s}\right] = \\ &= p_0 \left[\alpha \sum_{m=0}^{s-1} \frac{\alpha^m}{m!} + \frac{\alpha^s}{s!} \sum_{m=1}^{\infty} (m+s)\rho^m\right] = \\ &= p_0 \left[\alpha \sum_{m=0}^{s-1} \frac{\alpha^m}{m!} + \frac{\alpha^s}{s!}\left(\rho \sum_{m=0}^{\infty} m\rho^{m-1} + s\rho \sum_{r=0}^{\infty} \rho^r\right)\right] = \end{aligned}$$

[5] Ovaj rezultat je prvi put objavio Erlang 1917. godine. Pre nego što su došli računari, u telefonskoj industriji su korišćene $C(s,\alpha)$ kartice sa različitim kombinacijama s i α.

$$= p_0 \left[\alpha \sum_{m=0}^{s-1} \frac{\alpha^m}{m!} + \frac{\alpha^s}{s!} \left(\rho \left(\frac{1}{1-\rho} \right)' + s\rho \frac{1}{1-\rho} \right) \right] =$$

$$= p_0 \left[\alpha \sum_{m=0}^{s-1} \frac{\alpha^m}{m!} + \frac{\alpha^s}{s!} \left(\rho \frac{1}{(1-\rho)^2} + \alpha \frac{1}{1-\rho} \right) \right] =$$

$$= \frac{\rho \alpha^s p_0}{s!\,(1-\rho)^2} + \alpha p_0 \left[\sum_{m=0}^{s-1} \frac{\alpha^m}{m!} + \frac{\alpha^s}{s!\,(1-\rho)} \right].$$

Primetimo da je izraz između [ ] poslednjeg reda jednaka p_0^{-1}, pa imamo da je

$$L = \alpha + \frac{\rho \alpha^s p_0}{s!\,(1-\rho)^2},$$

što možemo zapisati kao

$$L = \alpha + \frac{\rho p_s}{(1-\rho)^2}.$$

Sada ćemo izraziti L_q, pišemo

$$L_q = E(N_q) = \sum_{n=s+1}^{\infty} (n-s) p_n = \sum_{n=s+1}^{\infty} (n-s) \frac{\alpha^s}{s!} \rho^{n-s} p_0 =$$

$$= \frac{\alpha^s}{s!} p_0 \sum_{n-s=1}^{\infty} (n-s)\, \rho^{n-s} = \frac{\alpha^s}{s!} p_0 \rho \sum_{m=0}^{\infty} m \rho^{m-1} =$$

$$= \frac{\alpha^s}{s!} p_0 \rho \left(\sum_{m=0}^{\infty} \rho^m \right)' = \frac{\alpha^s}{s!} p_0 \rho \left(\frac{1}{1-\rho} \right)' =$$

$$= \frac{\alpha^s}{s!} p_0 \rho \frac{1}{(1-\rho)^2} = \frac{\rho \alpha^s p_0}{s!\,(1-\rho)^2}, \tag{2.37}$$

što možemo zapisati na sledeći način

$$L_q = \frac{\rho p_s}{(1-\rho)^2}. \tag{2.38}$$

Ako u poslednjem koraku (2.37) uvrstimo izraz za p_0, dobijamo da je

$$L_q = \frac{\rho C(s, \alpha)}{1-\rho}.$$

Poređenjem izraza za L i L_q, možemo zaključiti da $s\rho(=\alpha)$ predstavlja *očekivani broj zauzetih servera*.

Za raspravu o vremenu čekanja klijenata pretpostavimo da se oni služe po principu $FCFS$. Pristigli klijent treba da čeka u redu za opsluživanje samo kada je broj korisnika u sistemu veći ili jednak s, tj. $N = n \geq s$. Svih s servera su zauzeti, vremena između odlazaka imaju eksponencijalnu raspodelu sa stopom $s\mu$. s klijenata su na serverima i $n - s$ klijenata čeka u redu. Stoga, novi klijent treba da sačeka prethodnih $n - s + 1$ klijenata da oni budu usluženi i tek nakon toga dolazi on na red (ako je $n = s$, što znači da nikoga nema u redu, novi klijent koji pristigne mora da sačeka da se jedno usluživanje završi; ako je $n = s + 1$, tada treba se čekati da se dva usluživanja završe; ...).

Dakle, vreme čekanja u redu je zbir $n - s + 1$ nezavisnih eksponencijalnih slučajnih promenljivih, svaki sa srednjom vrednošću $1/(s\mu)$, što predstavlja *gama raspodelu* sa parametrima $n - s + 1$ i $s\mu$. Ako je t$>$ 0 i kako je $\Gamma(n - s + 1) = (n - s)!$, možemo pisati:

$$T_q : \Gamma(n - s + 1),$$

gde smo sa T_q označili *vreme čekanja klijenata*.

Neka je $t > 0$ i $F_{T_q}(t) = P\{T_q \leq t\}$. Tada je

$$F_{T_q}(0) = P\{T_q = 0\} = P\{N < s\} = \sum_{n=0}^{s-1} p_n = p_0 \sum_{n=0}^{s-1} \frac{\alpha^n}{n!}.$$

Iz prve jednačine kod (2.36) imamo da je

$$p_0 \left(\sum_{n=0}^{s-1} \frac{\alpha^n}{n!} + \frac{\alpha^s}{s!} \left(1 - \frac{\alpha}{s}\right)^{-1} \right) = 1,$$

odnosno

$$\sum_{n=0}^{s-1} \frac{\alpha^n}{n!} + \frac{\alpha^s}{s!} \left(1 - \frac{\alpha}{s}\right)^{-1} = \frac{1}{p_0}.$$

Odavde sledi da je

$$\sum_{n=0}^{s-1} \frac{\alpha^n}{n!} = \frac{1}{p_0} - \frac{\alpha^s}{s!} (1 - \rho)^{-1}.$$

Sada imamo da je

$$F_{T_q}(0) = p_0 \left(\frac{1}{p_0} - \frac{\alpha^s}{s!} (1 - \rho)^{-1} \right) = 1 - \frac{\alpha^s p_0}{s!\,(1 - \rho)}. \tag{2.39}$$

Kod sistema sa više servera, slično kao i kod $M/M/1$ sistema, imamo da je

$$\begin{aligned} dF_{T_q}(t) &= \sum_{n=s}^{\infty} p_n e^{-s\mu t} \frac{(s\mu t)^{n-s}}{(n-s)!} s\mu dt = p_s e^{-s\mu t} \sum_{n=s}^{\infty} \rho^{n-s} \frac{(s\mu t)^{n-s}}{(n-s)!} s\mu dt = \\ &= s\mu p_s e^{-s\mu t} \sum_{n-s=0}^{\infty} \frac{(s\mu t\rho)^{n-s}}{(n-s)!} dt = s\mu p_s e^{-s\mu t} \sum_{m=0}^{\infty} \frac{(s\mu t\rho)^{m}}{m!} dt = \\ &= s\mu p_s e^{-s\mu t} e^{s\mu t\rho} dt = s\mu p_s e^{-s\mu t(1-\rho)} dt, \end{aligned} \tag{2.40}$$

što takođe možemo da zapišemo kao

$$dF_{T_q}(t) = \frac{s\mu\alpha^s}{s!} p_0 e^{-s\mu(1-\rho)t} dt. \tag{2.41}$$

Iz (2.40) imamo da je

$$\begin{aligned} W_q &= E(T_q) = \int_0^{\infty} t dF_{T_q}(t) = \int_0^{\infty} ts\mu p_s e^{-s\mu(1-\rho)t} dt = \\ &= \frac{p_s}{s\mu(1-\rho)^2}. \end{aligned} \tag{2.42}$$

Korišćenjem p_0 umesto p_s, možemo da pišemo sledeće

$$W_q = \frac{\alpha^s p_0}{s!\, s\mu(1-\rho)^2}, \tag{2.43}$$

što je dalje jednako

$$W_q = \frac{C(s,\alpha) W_s}{s(1-\rho)},$$

pošto je $W_s = \frac{1}{\mu}$.

Ako uporedimo (2.38) sa (2.42) (ili (2.37) sa (2.43)), ponovo možemo da potvrdimo Little-ovu formulu $L_q = \lambda W_q$.

Funkciju raspodele $F_{T_q}(t)$ *vremena čekanja* sada možemo da odredimo iz (2.39) i (2.41):

$$F_{T_q}(t) = F_{T_q}(0) + \int_0^t \frac{s\mu\alpha^s}{s!} p_0 e^{-s\mu(1-\rho)x} dx =$$

$$= 1 - \frac{\alpha^s p_0}{s!\,(1-\rho)} + \frac{\alpha^s p_0}{s!\,(1-\rho)} \int_0^t s\mu(1-\rho) e^{-s\mu(1-\rho)x} dx =$$

$$= 1 - \frac{\alpha^s p_0}{s!\,(1-\rho)} e^{-s\mu(1-\rho)t}. \tag{2.44}$$

Odnosno, ako uvrstimo izraz za p_0 dobijamo

$$F_{T_q}(t) = 1 - C(s,\alpha) e^{-\mu t(s-\alpha)}.$$

Za određivanje očekivane vrednosti za T_q i T mogli smo jednostavno samo primeniti Little-ovu formulu

$$W_q = \frac{L_q}{\lambda},$$

i Osobinu 3.

$$W = W_q + \frac{1}{\mu},$$

i konačno ponovo primenjujemo Little-ovu formulu

$$L = L_q + \alpha.$$

Razlog zbog kojeg dopuštamo da je $\alpha = \frac{\lambda}{\mu}$ je zbog toga, jer je ρ rezervisano za stopu pristizanja podeljeno sa najvećom stopom usluživanja, tj. $\rho = \frac{\lambda}{s\mu}$. Može se pokazati da ρ daje iskorišćenost servera, tj. ρ je deo vremena za koji je jedan proizvoljno izabran server zauzet.
U nastavku navešćemo formule koje su potrebne za izvođenje disperzija za M/M/s sistem, ali zbog složenosti nećemo ih odrediti.

$$E[N_q(N_q - 1)] = \frac{2p_0 n^s \rho^2}{s!\,(1-\rho)^3},$$

$$E(T_q^2) = \frac{2p_0 n^s}{\mu^2 s^2 s!\,(1-\rho)^3},$$

$$D(T) = D(T_q) + \frac{1}{\mu^2},$$

$$D(N) = \lambda^2 D(T) + L.$$

Primer 2.4. Korporacija, iz prethodnog primera, je implementirala proceduru prema kojoj se nikada ne dozvoljava da u njihovoj radionici boravi više od tri traktora. Za $600\ dolara$ nedeljno mogu da angažuju još jednog mehaničara. Postavlja se pitanje, da li se isplati da se to učini ako se kao kriterijum uzima očekivani trošak? Da bismo mogli dati odgovor na ovo pitanje, uporedićemo troškove $M/M/1/3$ sistema sa troškovima predloženog $M/M/2/3$ sistema. Jednačine rođenja i smrti (2.31) koristimo zajedno sa

$$\lambda_n = \begin{cases} \lambda & za\ n = 0, 1, 2, \\ 0 & za\ n = 3, 4, \dots, \end{cases}$$

$$\mu_n = \begin{cases} \mu & za\ n = 1, \\ 2\mu & za\ n = 2, 3, \end{cases}$$

gde je $\lambda = 3/nedelja$ i $\mu = 2/nedelja$. To nam daje

$$\rho = \frac{\lambda}{s\mu} = \frac{3}{2 \cdot 2} = 0.75$$

$$p_1 = 1.5p_0$$

$$p_2 = 1.125p_0$$

$$p_3 = 0.84375p_0$$

$$p_0 = \frac{1}{1 + 1.5 + 1.125 + 0.84375} = 0.224\ .$$

Očekivani broj klijenata u sistemu je

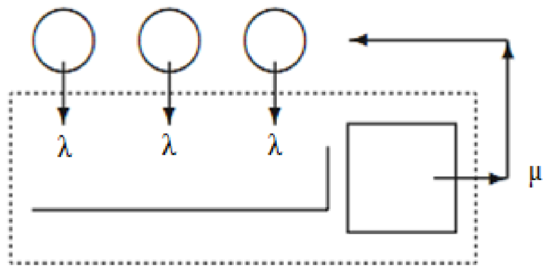
$$L = E(N) = 0.224 \cdot (1 \cdot 1.5 + 2 \cdot 1.125 + 3 \cdot 0.84375) = 1.407\ .$$

Troškovi predloženog sistema su

$$Trošak_{s=2} = 100 \cdot 1.407 + 500 \cdot 3 \cdot 0.189 + 600 = 824.20\ \ dolara/nedelja.$$

Dakle, korporaciji se ne isplati da angažuje još jednog mehaničara, jer će troškovi u tom slučaju biti veći od $820\ dolara$ kolko iznose kada je angažovan samo jedan mehaničar. □

Primer 2.5. Jedna mala korporacija ima tri stare mašine koje se stalno kvare. Svaka mašina se pokvari u proseku jednom nedeljno. Korporacija ima jednog mehaničara kome je potrebno u proseku jedna polovina nedelje za popravku mašine.



Slika 2.7. Šematski prikaz sistema čekanja u redu za popravku mašine (Primer 2.5.).

Pod pretpostavkom da su kvarovi i popravci slučajne promenljive sa eksponencijalnom raspodelom, jednačine rođenja i smrti se mogu koristiti na sledeći način :

$$\lambda_n = \begin{cases} (3-n)\lambda & za\ n = 0, 1, 2, \\ 0 & za\ n = 3, 4, \dots , \end{cases}$$

$$\mu_n = \mu \qquad za\ n = 1, 2, 3,$$

gde je $\lambda = 1/nedelja$ i $\mu = 2/nedelja$. Iz ovoga sledi

$$p_1 = 1.5p_0$$

$$p_2 = 1.5p_0$$

$$p_3 = 0.75p_0$$

$$p_0 = \frac{1}{1 + 1.5 + 1.5 + 0.75} = 0.21 .$$

$$L = E(N) = 0.21 \cdot (1 \cdot 1.5 + 2 \cdot 1.5 + 3 \cdot 0.75) = 1.42 .$$

Dalje, pretpostavimo da za svaki sat koji mašina provede u servisu, korporacija gubi $25\ dolara$. Trošak ovog sistema na sat, zbog nedostupnosti mašina, se izračunava kao

$$Trošak_{s=1} = 25 \cdot L = 25 \cdot 1.42 = 35.50\ dolara/č. \qquad \square$$

Primer 2.6. Kako bi izgledale performanse aerodroma iz Primera 2.2. da na njemu postoje dve piste, uz istu stopu pristizanja i posluživanja?

(a) Iskorišćenost piste:

$$\lambda = stopa\ pristizanja = 15/č,$$

$$\mu = stopa\ usluživanja = \tfrac{60}{3}/č = 20/č,$$

$$s = broj\ servera = 2,$$

$$\rho = stopa\ iskori\check{s}\acute{c}enosti\ svake\ piste = \frac{\lambda}{s\mu} = \frac{3}{8},$$

$$\alpha = s\rho = \frac{3}{4}.$$

(b) Očekivani broj aviona koji čekaju na sletanje:

$$L_q = E(N_q) = \frac{\rho\alpha^s p_0}{s!\,(1-\rho)^2},$$

$$p_0 = \left[\sum_{n=0}^{s-1}\frac{\alpha^n}{n!} + \frac{\alpha^s}{s!}(1-\rho)^{-1}\right]^{-1} = \left[1 + \frac{3}{4} + \frac{\left(\frac{3}{4}\right)^2}{2}\left(1-\frac{3}{8}\right)^{-1}\right]^{-1} = 0.4545\,,$$

$$L_q = \frac{\frac{3}{8}\cdot\left(\frac{3}{4}\right)^2\cdot 0.4545}{2\cdot\left(\frac{5}{8}\right)^2} = 0.1227\,.$$

(c) Očekivano vreme čekanja na sletanje:

$$W_q = E(T_q) = \frac{\alpha^s p_0}{s!\,s\mu(1-\rho)^2} =$$

$$= \frac{\left(\frac{3}{4}\right)^2\cdot 0.4545}{2\cdot 2\cdot 20\cdot\left(1-\frac{3}{8}\right)^2} =$$

$$= 0.00818\ \check{c}asa = 0.49\ minuta.$$

(d) Verovatnoća da će čekanje biti više od 5 minuta? 10 minuta? Nema čekanja?

$$P\{T_q \le t\} = 1 - \frac{\alpha^s p_0}{s!\,(1-\rho)}e^{-s\mu(1-\rho)t}$$

$$\Rightarrow \qquad P\{T_q > t\} = \frac{\alpha^s p_0}{s!\,(1-\rho)}e^{-s\mu(1-\rho)t},$$

$$P\{T_q > 5\ minuta\} = \frac{\left(\frac{3}{4}\right)^2\cdot 0.4545}{2\cdot\left(1-\frac{3}{8}\right)}\ e^{-2\cdot\frac{1}{3}\cdot\frac{5}{8}\cdot 5} = 0.1245\,,$$

$$P\{T_q > 10\ minuta\} = 0.0155\,.$$

$$P\{nema\ \check{c}ekanja\} = F_{T_q}(0) = P\{T_q = 0\} = 1 - \frac{\alpha^s p_0}{s!(1-\rho)} = 0.7955. \qquad \square$$

Primer 2.7. Banka je ustanovila dve blagajne – jednu za komercijalne i jednu za privatne korisnike. Stope pristizanja i usluživanja kod blagajne za komercijalne klijente su 6 i 12 klijenata na sat , respektivno. Odgovarajuće stope kod blagajne za privatne korisnike su 12 i 24 klijenata na sat, respektivno. Pretpostavljamo da su dolasci u skladu sa Poasonovim procesom i da vremena opsluživanja imaju eksponencijalnu raspodelu.

(a) Pod pretpostavkom da ova dva šaltera rade nezavisno jedan od drugog, odredićemo očekivani broj klijenata koji čekaju u redu i njihova srednja vremena čekanja na svakom šalteru. Rezultati su navedeni u Tabeli 2.1.

Tabela 2.1. Rezultati iz Primera 2.7.(a).

	Blagajna za komercijalne klijente	Blagajna za privatne klijente
λ	$6/čas$	$12/čas$
μ	$12/čas$	$24/čas$
$\rho = \frac{\lambda}{\mu}$	0.5	0.5
$L_q = \frac{\rho^2}{1-\rho}$	0.5	0.5
$W_q = \frac{\rho}{\mu(1-\rho)}$	$5\ minuta$	$2.5\ minuta$

(b) Šta je efekat kod čekanja u dva reda koji funkcioniše kao jedan dvoserverni red sa stopom pristizanja $18/čas$ i stopom posluživanja $18/čas$?

Tabela 2.2. Rezultati iz Primera 2.7.(b).

	Red sa dva servera
λ	$18/čas$
μ	$18/čas$
Broj servera (s)	2
$\rho = \frac{\lambda}{s\mu}$	0.5
$\alpha = \frac{\lambda}{\mu}$	1
$p_0 = \left[\sum_{n=0}^{s-1} \frac{\alpha^n}{n!} + \frac{\alpha^s}{s!\,(1-\rho)}\right]^{-1}$	0.4

$L_q = \dfrac{\rho \alpha^s p_0}{s!\,(1-\rho)^2}$	0.4
$W_q = \dfrac{\alpha^s p_0}{s!\, s\mu(1-\rho)^2}$	1.33 $minuta$

Možemo zaključiti da je red sa dva servera efikasniji nego kad se serveri koriste pojedinačno. □

Primer 2.8. *Čekanje u redu u banci.*

Uvod: Ljudi po pravilu ne vole redove u bankama, a da bi poboljšao nivo usluga menadžer banke želi da sazna:

1. Koliko je prosečan broj klijenata u banci (tj. broj klijenata u sistemu);
2. Koliko je prosečan broj klijenata koji čekaju u banci (tj. broj klijenata u redu);
3. Koliko će prosečno vremena klijent provesti u redu i koliko u sistemu;
4. Kolika je dužina vremena kad su blagajne slobodne?

U zavisnosti od toga koliko automata primenjuje tokom pauze za ručak, spreman je da zaposli najviše pet blagajnika, ali ne manje od jednog.

Pretpostavke i detalji modela:

- Raspodela dužina vremena koje je blagajniku potrebno za obavljanje zadatka je eksponencijalna sa srednjom vrednošću od $2\, minuta$ i standardnom devijacijom od $5/4\, minuta$.
- Gotovo da ne postoji ograničenje na dužinu reda, jer banka ima veliku čekaonicu.
- Kupci stižu u skladu sa Poasonovim raspodelom, sa srednjom vrednošću od $25\, klijenata\, na\, sat$.
- Usluživanje se vrši po principu $FCFS$.
- Ovo je jedan $M/M/s$ sistem, gde je $1 \leq s \leq 5$.

Oznake:

λ	prosečna stopa pristizanja ($25/čas$)
μ	1/prosečno vreme usluživanja klijenta= $30/čas$
s	broj blagajnika (servera)
ρ	prosečna količina posla za svakog servera po času
L	prosečan broj klijenata u sistemu
L_q	prosečna dužina reda
W	prosečno vreme koje će klijent provesti u sistemu
W_q	prosečno vreme koje će klijent provesti u redu

Računanja:

Na osnovu izloženog lako možemo da izračunamo sledeće vrednosti koje ćemo koristiti za određivanje L, L_q, W i W_q:

$$\alpha = \frac{\lambda}{\mu} = \frac{25}{30} = \frac{5}{6},$$

$$\rho = \frac{\lambda}{s\mu} = \frac{\alpha}{s} = \frac{5}{6s}.$$

p_0, L, L_q, W, W_q i dužinu vremena kada je server slobodan možemo izračunati korišćenjem sledećih jednačina:

$$p_0 = \left[\sum_{n=0}^{s-1} \frac{\alpha^n}{n!} + \frac{\alpha^s}{s!\,(1-\rho)} \right]^{-1},$$

$$L = \frac{\rho \alpha^s p_0}{s!\,(1-\rho)^2} + \alpha = \frac{5^{s+1} p_0}{6^{s+1} s!\, s \left(1 - \frac{5}{6s}\right)^2} + \frac{5}{6},$$

$$L_q = \frac{\rho \alpha^s p_0}{s!\,(1-\rho)^2} = \frac{5^{s+1} p_0}{6^{s+1} s!\, s \left(1 - \frac{5}{6s}\right)^2},$$

$$W_q = \frac{L_q}{\lambda} = \frac{L_q}{25},$$

$$W = W_q + \frac{1}{\mu} = W_q + \frac{1}{30},$$

$$Slobodno\ vreme = 1 - \rho = 1 - \frac{5}{6s}.$$

Rezultati ovih jednačina za 1 do 5 blagajnika su dati u sledećoj tabeli:

Blagajnici (s)	ρ	$\frac{\alpha^n}{n!}$	p_0	L_q	L	W	W_q	Slobodno vreme
1	0.833	0.833	0.167	4.167	5.000333	0.2	0.16668	0.167
2	0.417	0.347	0.412	0.175	1.008333	0.04	0.007	0.583
3	0.278	0.096	0.432	0.022	0.855333	0.034	0.00088	0.722
4	0.208	0.02	0.434	0.003	0.836333	0.033	0.00012	0.792
5	0.167	0.003	0.435	0.000	0.833333	0.033	0	0.833

Iz tabele se vidi da je dovoljno zaposliti jednog ili dva blagajnika, pošto se time skraćuje vreme čekanja, a blagajnici najveći deo radnog vremena nisu besposleni. □

2.4.2 Sistemi sa s serverom i konačnim kapacitetom c ($M/M/s/c$ model)

Kada je kapacitet čekaonice u sistemu masovnog posluživanja ograničen, govorimo o *ograničenom sistemu* masovnog posluživanja. U većini situacija , javljaju se konačni redovi što je i mnogo prirodnije nego čekaonice sa beskonačnim kapacitetom. Međutim, sa pomeranjem granice kapaciteta, sistem se ponaša kao da se radi o sistemu beskonačnog kapaciteta, pa u tom slučaju možemo da ignorišemo ograničenja u pogledu veličine. Dobar primer za ograničeni sistem masovnog posluživanja predstavlja komunikacioni sistem sa ograničenim baferima i nekoliko uslužnih kanala.

Posmatrajmo jedan sistem masovnog posluživanja gde je *broj servera* s, *pristizanja* su prema *Poasonovoj raspodeli*, *posluživanje* je *eksponencijalno*, a *granični kapacitet broja korisnika u sistemu* je c. Jasno je da je $c \geq s$. Pretpostavimo da su λ i μ *stope pristizanja i usluživanja*, respektivno. Tada imamo sledeće infinitezimalne stope prelaza kod generalizovanog *modela rađanja i smrti*:

$$\lambda_n = \lambda \quad za\ n = 0, 1, 2, \dots , c-1,$$

$$\mu_n = \begin{cases} n\mu & za\ n = 1, 2, \dots, s-1, \\ s\mu & za\ n = s, s+1, \dots, c. \end{cases}$$

Pretpostavljamo da se korisnicima, kada njihov broj u sistemu dostigne c, više ne dozvoljava ulazak u sistem (ili se proces dolazaka prekine). *Generator matrica* A je

$$A = \begin{bmatrix} -\lambda & \lambda & & & \\ \mu & -(\lambda+\mu) & \lambda & & \\ & \ddots & \ddots & \ddots & \\ & & s\mu & -(\lambda+s\mu) & \lambda \\ & & & s\mu & -s\mu \end{bmatrix}.$$

Za *granične verovatnoće* $\{p_n, n = 0, 1, 2, \dots, c\}$, jednačine ravnoteže se mogu odrediti na sličan način kao kod $M/M/s/\infty$ modela. Rešenje koje odgovara (2.34) je dato kao

$$p_n = \begin{cases} \dfrac{1}{n!}\left(\dfrac{\lambda}{\mu}\right)^n p_0, & 0 \leq n \leq s, \\ \dfrac{1}{s!}\left(\dfrac{\lambda}{\mu}\right)^s \left(\dfrac{\lambda}{s\mu}\right)^{n-s} p_0, & s \leq n \leq c. \end{cases}$$

Ako označimo da je $\frac{\lambda}{s\mu} = \rho$ i $\frac{\lambda}{\mu} = \alpha$, tada p_0 možemo odrediti pomoću uslova $\sum_{n=0}^{c} p_n = 1$:

$$p_0 = \left[\sum_{n=0}^{s-1} \frac{\alpha^n}{n!} + \frac{\alpha^s}{s!}\sum_{n=s}^{c} \rho^{n-s}\right]^{-1}.$$

S obzirom da je druga suma na desnoj strani kod izraza za p_0 konačna, ne moramo nametnuti uslov $\rho < 1$ za rešenje gde je $p_0 > 0$. Imamo da je

$$p_0 = \begin{cases} \left[\sum_{n=0}^{s-1} \frac{\alpha^n}{n!} + \frac{\alpha^s}{s!}\frac{1-\rho^{c-s+1}}{1-\rho}\right]^{-1}, & \rho \neq 1, \\ \left[\sum_{n=0}^{s-1} \frac{\alpha^n}{n!} + \frac{\alpha^s}{s!}(c-s+1)\right]^{-1}, & \rho = 1, \end{cases}$$

$$p_n = \begin{cases} \frac{\alpha^n}{n!} p_0, & 0 \leq n \leq s, \\ \frac{\alpha^s}{s!}\rho^{n-s} p_0, & s \leq n \leq c. \end{cases}$$

Kada raspravljamo o karakteristikama vremena čekanja korisnika u ograničenom sistemu masovnog posluživanja treba da dozvolimo mogućnost da se neki korisnici koji pristignu neće priključiti sistemu. Kada je sistem u ravnoteži, *verovatnoća da klijent koji stigne neće se priključiti sistemu* je p_c. Stoga kada u sistemu ima n $(n < c)$ klijenata, *verovatnoća da će klijent koji dođe ući u sistem* je dato sa $p_n/(1-p_c)$. Koristeći notaciju kao i ranije za *raspodelu vremena čekanja*, imamo da je

$$F_{T_q}(t) = P\{T_q \leq t\} = F_{T_q}(0) + P\{0 < T_q \leq t\},$$

gde je

$$F_{T_q}(0) = \sum_{n=0}^{s-1} \frac{p_n}{1-p_c}.$$

Takođe, imamo da je

$$dF_{T_q}(t) = \sum_{n=s}^{c-1} \frac{p_n}{1-p_c} e^{-s\mu t} \frac{(s\mu t)^{n-s}}{(n-s)!} s\mu dt,$$

$$F_{T_q}(t) = F_{T_q}(0) + \frac{1}{1-p_c}\sum_{n=s}^{c-1} p_n \int_0^t e^{-s\mu t} \frac{(s\mu t)^{n-s}}{(n-s)!} s\mu dt =$$

$$= F_{T_q}(0) + \frac{1}{1-p_c}\sum_{n=s}^{c-1} p_n \left(1 - \int_t^{\infty} e^{-s\mu t}\,\frac{(s\mu t)^{n-s}}{(n-s)!}\,s\mu dt\right).$$

Prilikom pojednostavljivanja prethodnog izraza, treba da imamo u vidu sledeće:

$$F_{T_q}(0) + \frac{1}{1-p_c}\sum_{n=s}^{c-1} p_n = 1$$

i

$$\int_t^{\infty} e^{-s\mu t}\,\frac{(s\mu t)^{n-s}}{(n-s)!}\,s\mu dt = \sum_{r=0}^{n-s} e^{-s\mu t}\,\frac{(s\mu t)^r}{r!}.$$

Tada imamo

$$F_{T_q}(t) = 1 - \frac{1}{1-p_c}\sum_{n=s}^{c-1} p_n \sum_{r=0}^{n-s} e^{-s\mu t}\,\frac{(s\mu t)^r}{r!}.$$

Očekivano vreme čekanja u redu je dato sa

$$W_q = E(T_q) = \int_0^{\infty} t dF_{T_q}(t) = \sum_{n=s}^{c-1} \frac{p_n}{1-p_c}\int_0^{\infty} e^{-s\mu t}\,\frac{(s\mu t)^{n-s}}{(n-s)!}\,s\mu t dt =$$

$$= \frac{1}{s\mu(1-p_c)}\sum_{n=s}^{c-1}(n-s+1)p_n.$$

Očekivano vreme čekanja u sistemu se može odrediti na sledeći način

$$W = W_q + \frac{1}{\mu}.$$

Očekivani broj klijenata u redu i u sistemu se može dobiti pomoću *Little-ove formule* i primenjivanjem *efektivne stope dolazaka* λ_e koja je jednaka $\lambda(1-p_c)$, dobijamo:

$$L = \lambda(1-p_c)W,$$

$$L_q = \lambda(1-p_c)W_q.$$

Drugi izraz za L je

$$L = L_q + \frac{\lambda_e}{\mu} = L_q + \frac{\lambda(1-p_c)}{\mu} = L_q + \alpha(1-p_c).$$

2.4.3 Sistem gubitka($M/M/s/s$ model)

Sistem $M/M/s/s$ u kojem se pristiglim korisnicima ne dozvoljava ulazak ako su svi serveri zauzeti (ne formira se red čekanja) jedan je od prvih sistema koje je A.K.Erlang (1917.) razmatrao. Taj sistem se zove *sistem bez čekanja* ili *sistem gubitka*. Pre uvođenja bafera za čekanje poziva, telefonski sistemi su funkcionisali isključivo kao sistemi gubitka.

Neka su *dolasci* klijenata u skladu sa *Poasonovim procesom* sa stopom λ i neka su *vremena usluživanja eksponencijalna* sa srednjom vrednošću $1/\mu$. Postoji s servera, i svi korisnici koji dolaze kada su svi serveri zauzeti predstavljaju *gubitak* za sistem. *Skup stanja* za ovakav sistem je $\{0, 1, 2, \ldots, s\}$. *Generator matrica* je sledeća:

$$A = \begin{bmatrix} -\lambda & \lambda & & & \\ \mu & -(\lambda+\mu) & \lambda & & \\ & \ddots & \ddots & \ddots & \\ & & (s-1)\mu & -[\lambda+(s-1)\mu] & \lambda \\ & & & s\mu & -s\mu \end{bmatrix}.$$

Shodno tome, granične verovatnoće ćemo dobiti iz *ravnotežnih jednačina*:

stanje i	*stopa po kojoj proces napušta stanje i* ("stopa izlaska")	=	*stopa po kojoj proces ulazi u stanje i* ("stopa ulaska")	
0	λp_0	=	μp_1	
$1 \le n < s$	$(\lambda + n\mu)p_n$	=	$\lambda p_{n-1} + (n+1)\mu p_{n+1}$	(2.45)
s	$s\mu p_s$	=	λp_{s-1}.	

Označimo da je $\frac{\lambda}{\mu} = \alpha$. (2.45) možemo da rešimo pomoću rekurzije i dobijamo

$$p_0 = \left[1 + \alpha + \frac{\alpha^2}{2!} + \ldots + \frac{\alpha^s}{s!}\right]^{-1},$$

$$p_n = \frac{\alpha^n}{n!} p_0, \qquad n = 0, 1, \ldots, s.$$

Dalje sledi

$$p_s = \frac{\frac{\alpha^s}{s!}}{1 + \alpha + \frac{\alpha^2}{2!} + \ldots + \frac{\alpha^s}{s!}}, \tag{2.46}$$

što predstavlja *verovatnoću da će klijent biti blokiran prilikom ulaska u sistem*. (Izgubljeni telefonski pozivi.) Takođe, λp_s daje *očekivani broj klijenata koji će biti blokirani prilikom ulaska u sistem u jedinici vremena*. Izraz (2.46) je poznat kao E*rlangova formula gubitka* ili *Prva Erlangova formula* i označava se kao $E_{1,s}(\alpha)$ ili $B(s,\alpha)$. Ova formula se intenzivno koristi u projektovanju telefonskih sistema.

Vrednosti za p_s date su za različite vrednosti s, kao i za različite vrednosti *ponuđenog prometa* α. U literaturi o telefonskom saobraćaju, ponuđeni promet (odnos stope dolazaka i posluživanja) se meri u Erlangu. Napomenimo i to da je u žargonu telefonske industrije obavljeni promet dato sa $\alpha(1-p_s)$, pošto se deo p_s za dolazeće korisnike gubi u sistemu.

Izraz na desnoj strani formule (2.46) je konveksna funkcija od s u intervalu $[0,\infty)$ za $\alpha > 0$. Još jedna od karakteristika ove formule je njeno važenje čak i kada vremena usluživanja imaju opšte raspodele.

2.4.4 Sistem masovnog posluživanja sa beskonačnim brojem servera ($M/M/\infty$ model)

Nazivanje nekog sistema sistemom masovnog posluživanja sa beskonačnim brojem servera (beskonačan broj servera i samim tim nema čekanja u redu) je pogrešno, $M/M/\infty$ sistem se zbog svoje strukture ipak tako naziva. Klijenti dolaze u skladu sa *Poasonovim procesom* i vremena usluživanja imaju *eksponencijalnu raspodelu*. Neka su λ i µ *stope pristizanja i usluživanja*. Pretpostavimo da je sistem u stanju da pruži uslugu čim kupac stigne. Jednostavan primer za ovo predstavlja veliki supermarket u kojem korisnici sami sebe poslužuju dok biraju robu. Kase na izlazu se tada modeliraju kao jedan $M/M/s$ sistem.

Kada u sistemu ima n klijenata, stopa usluživanja je $n\mu$ $(n = 1, 2, \ldots)$. *Parametri modela rađanja i smrti* tada su

$$\begin{aligned} \lambda_n &= \lambda, \qquad & n &= 0, 1, 2, \ldots, \\ \mu_n &= n\mu, \qquad & n &= 1, 2, 3, \ldots. \end{aligned} \tag{2.47}$$

Generator matrica se dobija proširivanjem prvog dela matrice (2.31) kod $M/M/s$ modela sa $n = s+1, s+2, \ldots$. Dobijamo

$$A = \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \end{array} \begin{bmatrix} -\lambda & \lambda & & & \\ \mu & -(\lambda+\mu) & \lambda & & \\ & 2\mu & -(\lambda+2\mu) & \lambda & \\ & & & \ddots & \end{bmatrix}.$$

Ravnotežne jednačine stanja za granične verovatnoće $\{ p_n, n = 0, 1, 2, \ ...\}$ su sledećeg oblika

$$\begin{aligned} \lambda p_0 &= \mu p_1, \\ (\lambda + n\mu)p_n &= \lambda p_{n-1} + (n+1)\mu p_{n+1}, \quad n = 1, 2, \ldots \ . \end{aligned}$$

Iz uslova $\sum_{n=0}^{\infty} p_n = 1$ sledi

$$p_0 = e^{-\frac{\lambda}{\mu}},$$

$$p_n = \frac{e^{-\lambda/\mu}\left(\frac{\lambda}{\mu}\right)^n}{n!}, \quad n = 0, 1, 2, \ldots,$$

što je *Poasonov proces* sa parametrom λ/μ.

Zbog strukture parametara rađanja i umiranja (2.47), sistem se takođe može posmatrati kao sistem masovnog posluživanja gde su dolasci u skladu sa Poasonovim procesom i vremena usluživanja su eksponencijalna sa linearno zavisnom stopom usluživanja $n\mu$ kada postoji $n\mu$ klijenata u sistemu.

Napomena 8. $M/M/s\,(= M/M/s/\infty)$ i $M/M/s/s$ sistemi masovnog posluživanja predstavljaju dva *ekstremna slučaja* koje treba razmotriti. $M/M/s/c$ sa $s < c < \infty$ predstavlja model koji najčešće dobro reprezentuje stvarnost, pošto obično postoji prostor u kojem potencijalni korisnici mogu da čekaju dok ne budu posluženi, ali taj prostor nije beskonačan.

Osim toga, zaključili smo da za $M/M/s$ sistem, da bi postigao stacionarni režim rada, mora da bude zadovoljen uslov da je $\lambda < s\mu$, to jest, stopa pristizanja korisnika u sistem mora da bude manja od stope posluživanja korisnika kada je broj zauzetih servera s, ili će se dužina čekanja u redu otegnuti u beskonačnost. U praksi, međutim, neki dolazeći klijenti neće ući u sistem ako procene da je vreme čekanja predugo. Neki će pak da napuste sistem pre nego što budu posluženi ukoliko procene da su previše vremena utrošili čekajući u redu.

Glava 3

3. Simulacija

Analitičko rešenje nekog problema, odnosno tačne informacije o tom problemu korišćenjem standardnih matematičkih metoda možemo dobiti samo ako su relacije na kojima je izgrađen matematički model dovoljno jednostavne. Međutim, većina sistema iz stvarnog sveta su veoma komplikovana da bi se njihovi modeli mogli rešiti analitički, zbog toga se ti modeli rešavaju pomoću simulacija.

Simulacije predstavljaju tehniku koja koristeći računar oponaša funkcionisanje različitih sistema u realnom svetu. U cilju proučavanja, objašnjavanja i opisivanja takvih sistema potrebno je odrediti skup pretpostavki o tome kako oni funkcionišu. Pretpostavke simulacija se formulišu u zavisnosti od naših saznanja i prikupljenih podataka o sistemu.

U ovom delu rada izvršićemo simulaciju na konkretnim primerima za izračunavanje ocene *prosečnog broja klijenata u sistemu, u redu i na serveru, prosečnog vremena zadržavanja u sistemu, u redu i na serveru, kao i iskorišćenost servera* za $M/M/s/c$ modele redova čekanja za različite vrednosti s i c.

3.1 Simulacija čekanja u redu u pošti

Kao primer za simulaciju, razmatraćemo usluživanje građana/klijenata u pošti od strane *jednog* ili *više blagajnika*. Pretpostavimo da se pristizanje klijenata u poštu odvija saglasno *Poasonovom procesu* sa srednjom stopom λ. Na osnovu iskustva, procenjuje se da je vreme usluživanja eksponencijalno raspodeljena, odnosno usluživanja klijenata mogu se modelirati slučajnim promenljivama sa *eksponencijalnom raspodelom* sa očekivanom vrednošću $1/\mu$. Znamo, λ predstavlja prosečan broj klijenata po jedinici vremena koje pristižu u poštu, dok je μ prosečan broj usluženih klijenata po jedinici vremena. Dalje, pretpostavićemo slučaj kada je *kapacitet pošte beskonačan*, odnosno *konačan*, a usluživanje se kod svakog modela vrši po principu *"ko prvi dolazi, prvi odlazi"*. Prema tome, svaki klijet koji pri dolasku u pošti zatekne slobodnu blagajnu odmah biva uslužena, dok u situaciji u kojoj su sve blagajne zauzete staje u red za pružanje usluge. Nakon pružene usluge klijent napušta poštu.

3.1.1 Simulacija M/M/s modela čekanja u pošti

Sistemi u kojima figurišu $M/M/1$ redovi čekanja su veoma jednostavni u odnosu na sisteme iz realnog sveta, ali primer ove simulacije je reprezentativan i za mnogo komplikovanije sisteme.

Posmatraćemo tri modela. U sva tri slučaja pretpostavićemo da je kapacitet pošte beskonačan i da je prosečna stopa pristizanja u poštu 25 klijenata na sat, dok je prosečna stopa usluživanja 30 klijenata na sat. Odnosno, vreme između dolazaka klijenata je u proseku $2{,}4$ minuta, a prosečno vreme usluživanja je 2 minuta. U prvom slučaju imamo $M/M/1$ model, tj. u pošti postoji jedan red u kojem klijenti čekaju da budu usluženi od strane jednog blagajnika. U drugom i trećem slučaju pretpostavićemo da su poštu proširili sa još jednim, odnosno dva šaltera, respektivno. Dakle, u poslednja dva slučaja imaćemo $M/M/2$ i $M/M/3$ sisteme.

Rezultati matematičkog modela za posmatrane M/M/1, M/M/2 i M/M/3 modele date su u tabelama 3.1., 3.3. i 3.5., redom, a rezultati simulacije u tabelama 3.2., 3.4. i 3.6., respektivno.

U drugoj koloni u tabelama 3.2., 3.4. i 3.6. prikazani su rezultati dobijeni simulacijom, iz čega vidimo da je prosečan broj klijenata koji čekaju u redu (L_q) u pošti sa jednim, dva i tri šaltera $4{,}45$, $0{,}17$ i $0{,}03$, redom, a prosečan broj klijenata na šalterima (L_s) je $0{,}84$, $0{,}85$ i $0{,}84$, respektivno. Dakle, prosečan broj klijenata u pošti (L) je $5{,}29$, $1{,}02$ i $0{,}87$ redom. Odgovor na pitanje prosečno koliko vremena će klijent provesti u pošti, daje prosečno vreme čekanja u pošti (W) što kod posmatranih sistema iznosi $0{,}2$, $0{,}4$ i $0{,}04$, odnosno 12, 24 i $2{,}4\ minuta$, redom. Prosečno vreme provedeno u pošti (W) je zbir vremena provedeno u redu (W_q) i vremena provedeno na blagajni (W_s). Stopa efikasnosti (iskorišćenosti, ρ) ukazuje na to da su blagajnici zauzeti u proseku 83%, 43% i 28% vremena, redom.

Povećavanjem broja šaltera (servera) smanjuje se prosečan broj klijenata u pošti (sistemu) i u redu, a smanjuje se i vreme provedeno u redu. Proširivanjem pošte jednim šalterom $(M/M/2)$, znatno se smanji broj klijenata u sistemu, ali i stopa efikasnosti. Svako dalje proširivanje se čini bespotrebnim, s obzirom da će šalteri najveći deo vremena biti besposleni, a broj klijenata u redu težiti nuli, što se može videti iz tabela.

Tabela 3.1.

Queue Station	Pošta1
Arrival Rate (λ)	25
Service Rate/Channel (μ)	30
Number of Servers (s)	1
Max. Number in System (c)	***
Number in Population (p)	***
Type	M/M/1
Mean Number at Station (L)	5.0000
Mean Time at Station (W)	0.2000
Mean Number in Queue (L_q)	4.1667
Mean Time in Queue (W_q)	0.1667
Mean Number in Service (L_s)	0.8333
Mean Time in Service (W_s)	0.0333
Efficiency (ρ)	0.8333
Probability All Servers Idle	0.1667
Prob. All Servers Busy	0.8333
Prob. System Full	0.0000
P(0)	0.1667
P(1)	0.1389
P(2)	0.1157
P(3)	0.0965
P(4)	0.0804
P(5)	0.0670
P(6)	0.0558
P(7)	0.0465
P(8)	0.0388
P(9)	0.0323
P(10)	0.0269

Tabela 3.2.

Queueing Simulaton	
Entity-Not Dynamic	
Queue Station	Pošta1
Arrival Rate (λ)	25
Service Rate (μ)	30
Number of Servers (s)	1
Max. Number in System (c)	***
Type	M/M/1
Arrival Seed	***
Service Seed	***
Number in Simulation	100
Start Data Time	0
Stop Data Time	3.7534
Mean Number at Station (L)	5.2833
Mean Time at Station (W)	0.2003
Mean Number in Queue (L_q)	4.4475
Mean Time in Queue (W_q)	0.1686
Mean Number in Service (L_s)	0.8359
Mean Time in Service (W_s)	0.0317
Efficiency (ρ)	0.8359

Tabela 3.3.

Queue Station	Pošta2
Arrival Rate (λ)	25
Service Rate/Channel (μ)	30
Number of Servers (s)	2
Max. Number in System (c)	***
Number in Population (p)	***
Type	M/M/2
Mean Number at Station (L)	1.0084
Mean Time at Station (W)	0.0403
Mean Number in Queue (L_q)	0.1751
Mean Time in Queue (W_q)	0.0070
Mean Number in Service (L_s)	0.8333
Mean Time in Service (W_s)	0.0333
Efficiency (ρ)	0.4167
Probability All Servers Idle	0.4118
Prob. All Servers Busy	0.2451
Prob. System Full	0.0000
P(0)	0.4118
P(1)	0.3431
P(2)	0.1430
P(3)	0.0596
P(4)	0.0248
P(5)	0.0103
P(6)	0.0043
P(7)	0.0018
P(8)	0.0007
P(9)	0.0003
P(10)	0.0001

Tabela 3.4.

Queueing Simulaton	
Entity-Not Dynamic	
Queue Station	Pošta2
Arrival Rate (λ)	25
Service Rate (μ)	30
Number of Servers (s)	2
Max. Number in System (c)	***
Type	M/M/2
Arrival Seed	***
Service Seed	***
Number in Simulation	100
Start Data Time	0
Stop Data Time	3.8954
Mean Number at Station (L)	1.0244
Mean Time at Station (W)	0.0403
Mean Number in Queue (L_q)	0.1704
Mean Time in Queue (W_q)	0.0067
Mean Number in Service (L_s)	0.8540
Mean Time in Service (W_s)	0.0336
Efficiency (ρ)	0.4270

Tabela 3.5.

Queue Station	Pošta3
Arrival Rate (λ)	25
Service Rate/Channel (μ)	30
Number of Servers (s)	3
Max. Number in System (c)	***
Number in Population (p)	***
Type	M/M/3
Mean Number at Station (L)	0.8555
Mean Time at Station (W)	0.0342
Mean Number in Queue (L_q)	0.0222
Mean Time in Queue (W_q)	0.0009
Mean Number in Service (L_s)	0.8333
Mean Time in Service (W_s)	0.0333
Efficiency (ρ)	0.2778
Probability All Servers Idle	0.4321
Prob. All Servers Busy	0.0577
Prob. System Full	0.0000
P(0)	0.4321
P(1)	0.3601
P(2)	0.1500
P(3)	0.0417
P(4)	0.0116
P(5)	0.0032
P(6)	0.0009
P(7)	0.0002
P(8)	0.0001
P(9)	0.0000
P(10)	0.0000

Tabela 3.6.

Queueing Simulaton	
Entity-Not Dynamic	
Queue Station	Pošta3
Arrival Rate (λ)	25
Service Rate (μ)	30
Number of Servers (s)	3
Max. Number in System (c)	***
Type	M/M/3
Arrival Seed	***
Service Seed	***
Number in Simulation	100
Start Data Time	0
Stop Data Time	4.2757
Mean Number at Station (L)	0.8686
Mean Time at Station (W)	0.0375
Mean Number in Queue (L_q)	0.0300
Mean Time in Queue (W_q)	0.0013
Mean Number in Service (L_s)	0.8386
Mean Time in Service (W_s)	0.0362
Efficiency (ρ)	0.2795

Odgovarajući rezultati dobijeni za matematički model zajedno sa prvih 11 verovatnoća stanja prikazani su u tabelama 3.1., 3.3. i 3.5., redom. Verovatnoće stanja u prvom slučaju ukazuju na to da će šalter slobodan biti 17% vremena ($P(0)$), a zauzet 83% $vremena$ $(1 - P(0))$. U drugom slučaju oba šaltera će slobodna biti 41% $vremena$, a zauzeta 24% $vremena$ $(1 - P(0) - P(1))$. U slučaju tri šaltera, sva tri će biti slobodna 43% $vremena$, a zauzeta 6% $vremena$ $(1 - P(0) - P(1) - P(2))$. U ovim slučajevima pošta nije nikad pun, s obzirom da joj je kapacitet beskonačan.

3.1.2 Simulacija M/M/s/c modela čekanja u pošti

U većini si stema iz stvarnog sveta javljaju se konačni redovi što je i mnogo prirodnije nego čekaonice sa beskonačnim kapacitetom.

Posmatraćemo četiri modela. U sva četiri slučaja pretpostavićemo da je kapacitet pošte 10, prosečna stopa pristizanja u poštu je 30 klijenata na sat, dok je prosečna stopa usluživanja 25 klijenata na sat. Odnosno, vreme između dolazaka klijenata je u proseku 2 minuta, a prosečno vreme usluživanja je 2,4 minuta. U prvom slučaju imamo $M/M/1/10$ model, tj. u pošti postoji jedan red u kojem klijenti čekaju da budu usluženi od strane jednog blagajnika, dok je kapacitet čekaonice ograničen na 10. U ostalim slučajevima pretpostavićemo da su poštu proširili sa još jednim, dva, odnosno tri šaltera, respektivno. Dakle, u poslednja tri slučaja imaćemo $M/M/2/10$, $M/M/3/10$ i $M/M/4/10$ sisteme.

Rezultati matematičkog modela za posmatrane $M/M/1/10$, $M/M/2/10$, $M/M/3/10$ i $M/M/4/10$ modele date su u tabelama 3.7., 3.9., 3.11. i 3.13, redom, a rezultati simulacije u tabelama 3.8., 3.10., 3.12. i 3.14., respektivno.

Iz rezultata dobijene simulacijom vidimo da je prosečan broj klijenata koji čekaju u redu (L_q) u pošti sa jednim, dva, tri i četiri šaltera 5,67, 0,77, 0,11 i 0,01, redom, a prosečan broj klijenata na šalterima (L_s) je 0,94, 1,18, 1,15 i 1,31 respektivno. Prema tome, prosečan broj klijenata u pošti (L) je 6,61, 1,95, 1,26 i 1,32 redom. Prosečno vreme čekanja u pošti (W) kod ova četiri sistema iznosi 0,32, 0,07, 0,05 i 0,05, odnosno 19,2, 4,2, 3 i 3 $minuta$, redom. Stopa efikasnosti (ρ) ukazuje na to da su šalteri zauzet u proseku 94%, 59%, 38% i 33% vremena, redom.

Posmatranjem rezultata dobijene simulacijom, uočavamo da povećanjem broja šaltera na dva, očekivani broj klijenata u pošti se znatno smanjilo, kao što i vreme provedeno u redu, ali svakim novim uključivanjem šaltera smanjuje se iskorišćenost šaltera.

Tabela 3.7.

Queue Station	Pošta1
Arrival Rate (λ)	30
Service Rate/Channel (μ)	25
Number of Servers (s)	1
Max. Num. in System (c)	10
Number in Population (p)	***
Type	M/M/1/10
Mean Num. at Station (L)	6.7107
Mean Time at Station (W)	0.2770
Mean Num. in Queue (L_q)	5.7418
Mean Time in Queue (W_q)	0.2370
Mean Num. in Service (L_s)	0.9689
Mean Time in Service (W_s)	0.0400
Efficiency (ρ)	0.9689
Probability All Servers Idle	0.0311
Prob. All Servers Busy	0.9689
Prob. System Full	0.1926
P(0)	0.0311
P(1)	0.0373
P(2)	0.0448
P(3)	0.0537
P(4)	0.0645
P(5)	0.0774
P(6)	0.0929
P(7)	0.1115
P(8)	0.1337
P(9)	0.1605
P(10)	0.1926

Tabela 3.8.

Queueing Simulaton	
Entity-Not Dynamic	
Queue Station	Pošta1
Arrival Rate (λ)	30
Service Rate (μ)	25
Number of Servers (s)	1
Max. Num. in System (c)	10
Type	M/M/1/10
Arrival Seed	***
Service Seed	***
Number in Simulation	100
Start Data Time	0
Stop Data Time	3.5837
Mean Num. at Station (L)	6.6103
Mean Time at Station (W)	0.3201
Mean Num. in Queue (L_q)	5.6667
Mean Time in Queue (W_q)	0.2744
Mean Num. in Service(L_s)	0.9437
Mean Time in Service(W_s)	0.0457
Efficiency (ρ)	0.9437

Tabela 3.9.

Queue Station	Pošta2
Arrival Rate (λ)	30
Service Rate/Channel (μ)	25
Number of Servers (s)	2
Max. Num. in System (c)	10
Number in Population (p)	***
Type	M/M/2/10
Mean Num. at Station (L)	1.8266
Mean Time at Station (W)	0.0611
Mean Num. in Queue (L_q)	0.6302
Mean Time in Queue (W_q)	0.0211
Mean Num. in Service (L_s)	1.1964
Mean Time in Service (W_s)	0.0400
Efficiency (ρ)	0.5982
Probability All Servers Idle	0.2511
Prob. All Servers Busy	0.4475
Prob. System Full	0.0030
P(0)	0.2511
P(1)	0.3014
P(2)	0.1808
P(3)	0.1085
P(4)	0.0651
P(5)	0.0391
P(6)	0.0234
P(7)	0.0141
P(8)	0.0084
P(9)	0.0051
P(10)	0.0030

Tabela 3.10.

Queueing Simulaton	
Entity-Not Dynamic	
Queue Station	Pošta2
Arrival Rate (λ)	30
Service Rate (μ)	25
Number of Servers (s)	2
Max. Num. in System (c)	10
Type	M/M/2/10
Arrival Seed	***
Service Seed	***
Number in Simulation	100
Start Data Time	0
Stop Data Time	3.2966
Mean Num. at Station (L)	1.9521
Mean Time at Station (W)	0.0663
Mean Num. in Queue (L_q)	0.7747
Mean Time in Queue (W_q)	0.0263
Mean Num. in Service(L_s)	1.1774
Mean Time in Service(W_s)	0.0400
Efficiency (ρ)	0.5887

Tabela 3.11.

Queue Station	Pošta3
Arrival Rate (λ)	30
Service Rate/Channel (μ)	25
Number of Servers (s)	3
Max. Num. in System (c)	10
Number in Population (p)	***
Type	M/M/3/10
Mean Num. at Station (L)	1.2932
Mean Time at Station (W)	0.0431
Mean Num. in Queue (L_q)	0.0933
Mean Time in Queue (W_q)	0.0031
Mean Num. in Service (L_s)	1.1998
Mean Time in Service (W_s)	0.0400
Efficiency (ρ)	0.3999
Probability All Servers Idle	0.2941
Prob. All Servers Busy	0.1411
Prob. System Full	0.0001
P(0)	0.2941
P(1)	0.3530
P(2)	0.2118
P(3)	0.0847
P(4)	0.0339
P(5)	0.0136
P(6)	0.0054
P(7)	0.0022
P(8)	0.0009
P(9)	0.0003
P(10)	0.0001

Tabela 3.12.

Queueing Simulaton	
Entity-Not Dynamic	
Queue Station	Pošta3
Arrival Rate (λ)	30
Service Rate (μ)	25
Number of Servers (s)	3
Max. Num. in System (c)	10
Type	M/M/3/10
Arrival Seed	***
Service Seed	***
Number in Simulation	100
Start Data Time	0
Stop Data Time	3.5856
Mean Num. at Station (L)	1.2616
Mean Time at Station (W)	0.0457
Mean Num. in Queue (L_q)	0.1071
Mean Time in Queue (W_q)	0.0039
Mean Num. in Service(L_s)	1.1545
Mean Time in Service(W_s)	0.0418
Efficiency (ρ)	0.3848

Tabela 3.13.

Queue Station	Pošta4
Arrival Rate (λ)	30
Service Rate/Channel (μ)	25
Number of Servers (s)	4
Max. Num. in System (c)	10
Number in Population (p)	***
Type	M/M/4/10
Mean Num. at Station (L)	1.2158
Mean Time at Station (W)	0.0405
Mean Num. in Queue (L_q)	0.0158
Mean Time in Queue (W_q)	0.0005
Mean Num. in Service (L_s)	1.2000
Mean Time in Service (W_s)	0.0400
Efficiency (ρ)	0.3000
Probability All Servers Idle	0.3002
Prob. All Servers Busy	0.0370
Prob. System Full	0.0000
P(0)	0.3002
P(1)	0.3602
P(2)	0.2161
P(3)	0.0865
P(4)	0.0259
P(5)	0.0078
P(6)	0.0023
P(7)	0.0007
P(8)	0.0002
P(9)	0.0001
P(10)	0.0000

Tabela 3.14.

Queueing Simulaton	
Entity-Not Dynamic	
Queue Station	Pošta4
Arrival Rate (λ)	30
Service Rate (μ)	25
Number of Servers (s)	4
Max. Num. in System (c)	10
Type	M/M/4/10
Arrival Seed	***
Service Seed	***
Number in Simulation	100
Start Data Time	0
Stop Data Time	3.5819
Mean Num. at Station (L)	1.3227
Mean Time at Station (W)	0.0479
Mean Num. in Queue (L_q)	0.0138
Mean Time in Queue (W_q)	0.0005
Mean Num. in Service(L_s)	1.3089
Mean Time in Service(W_s)	0.0474
Efficiency (ρ)	0.3272

Za simulaciju koristili smo Microsoft Office Excel, a u okviru toga Add-Ins OR_MM.

Zaključak

U ovom radu smo se uglavnom bavili $M/M/s/c$ sistemima za različite vrednosti s i c, koji predstavljaju broj servera i kapacitet sistema, redom. Poasonova pristizanja i eksponencijalne usluge omogućavaju primenu Markovljevih modela masovnog usluživanja koji se lako mogu analizirati i čiji se rezultati mogu praktično primeniti. Najjednostavniji sistemi čekanja za analizu su oni koji uključuju jedan eksponencijali server i gde dolasci odgovaraju Poasonovom procesu. Ove sisteme relativno je lako proučavati, a mogu da posluže i za demonstriranje nekih opštih tehnika analize masovnog posluživanja koje se mogu proširiti i na složenije sisteme.

Do važne generalizacije $M/M/1$ modela dolazimo kada pretpostavimo da u sistemu postoji s *servera* (koji usluge pružaju nezavisno jedni od drugih) i da svi oni poslužuju *eksponencijalnom stopom* μ. Višeserverski $M/M/s$ sistem masovnog posluživanja predstavlja model koji se najviše koristi u analizi uslužnih stanica sa više od jednog servera. $M/M/s$ i $M/M/s/s$ sistemi masovnog posluživanja predstavljaju dva *ekstremna slučaja*. $M/M/s/c$ sa $s < c < \infty$ predstavlja model koji najčešće dobro reprezentuje stvarnost, pošto obično postoji prostor u kojem potencijalni korisnici mogu da čekaju dok ne budu posluženi, ali taj prostor je konačan. Zaključili smo da za $M/M/s$ sistem, da bi postigao stacionarni režim rada, mora da bude zadovoljen uslov da je stopa pristizanja korisnika u sistem manja od stope posluživanja korisnika kada je broj zauzetih servera s, ili će se dužina čekanja u redu otegnuti u beskonačnost. U praksi, međutim, neki dolazeći klijenti neće ući u sistem ako procene da je vreme čekanja predugo. Neki će pak da napuste sistem pre nego što budu posluženi ukoliko procene da su previše vremena utrošili čekajući u redu.

U poslednjem delu rada na konkretnom primeru smo izvršili simulaciju, čije rezultate smo uporedili sa rezultatima dobijenim pomoću matematičkog modela. Za opisivanje efikasnosti sistema redova čekanja kao kriterijum uzeli smo očekivani broj klijenata u sistemu i očekivani broj klijenata u redu, odnosno prosečno vreme čekanja u sistemu i u redu.

Literatura

[1] U. Narayan Bhat, *An Introduction to Queueing Theory*, Birkhäuser, 2008.

[2] Richard M. Feldman, Ciriaco Valdez-Flores, *Applied Probability and Stochastic, Processes*, Springer, 2010.

[3] Mario Lefebvre, *Applied Stochastic Processes*, Springer, 2007.

[4] D. Gross, J. F. Shortle, J. M. Thompson, C. M. Harris, *Fundamentals of Queueing Theory*, Wiley, 2011.

[5] Danijela Rajter-Ćirić, *Verovatnoća*, Univerzitet u Novom Sadu, Prirodno-matematički fakultet u Novom Sadu, 2008.

[6] Danijela Rajter-Ćirić, *Stohastička analiz*a, beleške sa predavanja i skripta, 2009.

[7] http://en.wikipedia.org

[8] http://mathworld.wolfram.com

[9] http://www.me.utexas.edu/~jensen/ORMM/computation/unit/que_add/index.html

Kratka biografija

Rođena sam 16. januara 1987. godine u Novom Sad. Završila sam Osnovnu školu „Petefi Šandor“ u Novom Sadu 2002. godine kao nosilac diplome „Vuk Karadžić“.

Opšti smer Gimnazije „Svetozar Marković“ u Novom Sadu završila sam 2006. godine, takođe kao nosilac Vukove diplome.

Osnovne studije na Prirodno-matematičkom fakultetu, Univerziteta u Novom Sadu, na Departmanu za matematiku i informatiku, na smeru Finansijska matematika upisala sam 2006. godine, a završila 2010. godine sa prosečnom ocenom 9.66.

Master akademske studije - Primenjena matematika, modul Matematika finansija upisala sam 2010. godine na Prirodno–matematičkom fakultetu u Novom Sadu, na Departmanu za matematiku i informatiku, gde sam položila sve ispite predviđene nastavnim planom i programom, kao i grupu pedagoško-psihološko-metodičkih predmeta, zaključno sa septembrom 2011. godine sa prosečnom ocenom 9.63.

Novi Sad, oktobar 2012. Elvira Klebečko

UNIVERZITET U NOVOM SADU
PRIRODNO MATEMATIČKI FAKULTET
KLJUČNA DOKUMENTACIJSKA INFORMACIJA

Redni broj: **RBR**	
Identifikacioni broj: **IBR**	
Tip dokumentacije: **TD**	Monografska dokumentacija
Tip zapisa: **TZ**	Tekstualni štampani materijal
Vrsta rada: **VR**	Master rad
Autor: **AU**	Elvira Klebečko
Mentor: **MN**	dr Danijela Rajter-Ćirić
Naslov rada: **NR**	Redovi čekanja i njihove primene
Jezik publikacije: **JP**	srpski (latinica)
Jezik izvoda: **JI**	s/en
Zamlja publikovanja: **ZP**	Srbija
Uže geografsko područje: **UGP**	Vojvodina
Godina: **GO**	2012.

Izdavač: Autorski reprint
IZ

Mesto i adresa: Novi Sad, Trg Dositeja Obradovića 4
MA

Fizički opis rada: (3/93/0/18/7/0/0)
FO (broj poglavlja, strana, lit.citata, tabela, slika, grafika, priloga)

Naučna oblast: Matematika
NO

Naučna disciplina: Stohastička analiza
ND

Predmetne odrednica, ključne reči: Stohastički proces, Poasonov proces, eksponencijalna raspodela, Little-ova teorema, čekanje u redu, sistemi masovnog usluživanja, M/M/s/c sistemi, broj servera, kapacitet sistema, princip rađanja/umiranja
PO
UDK

Čuva se: U biblioteci Departmana za matematiku i informatiku
ČU

Važna napomena: nema
VN

Izvod:
IZ

Rad se bavi redovima čekanja i njihovim primenama u sistemima masovnog usluživanja. U uvodnom delu rada obrađene su osnove teorije verovatnoće i stohastičke analize. U nastavku predstavljeni su sistemi masovnog usluživanja sa jednim serverom, tj. $M/M/1$ i $M/M/1/c$ model, kao i sistemi sa više servera, odnosno $M/M/s/\infty$, $M/M/s/c$, $M/M/s/s$ i M/M/∞ model, gde s i c predstavljaju broj servera i kapacitet sistema, redom. Za svaki od ovih modela predstavljena su osnovna svojstva i osobine. Pored toga, većina je ilustrovana primerima. Pretpostavljeno je da su procesi pristizanja u sisteme masovnog usluživanja Poasonovi procesi, kao i da su vremena usluživanja eksponencijalno raspodeljena. Poasonova pristizanja i eksponencijalne usluge omogućavaju primenu Markovljevih modela masovnog usluživanja koji se lako mogu analizirati i čiji se rezultati mogu praktično primeniti. Detaljno je prikazan opšti model masovnog usluživanja po principu rađanja i umiranja.

Datum prihvatanja teme od strane NN veća: **DP**	01.02.2012.
Datum odbrane: **DO**	
Članovi komisije: **KO**	
Predsednik:	dr Dora Seleši, docent Prirodno-matematičkog fakulteta u Novom Sadu
Član:	dr Sanja Rapajić, vanredni profesor Prirodno-matematičkog fakulteta u Novom Sadu
Mentor:	dr Danijela Rajter-Ćirić, redovni profesor Prirodno-matematičkog fakulteta u Novom Sadu

UNIVERSITY OF NOVI SAD
FACULTY OF SCIENCE
KEY WORDS DOCUMENTATION

Serial number: **SNO**	
Identification number: **INO**	
Document type: **DT**	Monograph type
Type of record: **TR**	Printed text
Content code: **CC**	Master's thesis
Author: **AU**	Elvira Klebečko
Menthor: **MN**	dr Danijela Rajter-Ćirić
Title: **TI**	Queueing Theory and its applications
Language of text: **LT**	Serbian (Latin)
Language of abstract: **LA**	s/en
Country of publication: **CP**	Serbia
Locality of publication: **LP**	Vojvodina
Publication year: **PY**	2012.

Publisher: **PU**	Author's reprint
Publication place: **PP**	Novi Sad, Trg Dositeja Obradovića 4
Physical description: **PD**	(3/93/0/18/7/0/0) (chapters/pages/literature/tables/pictures/ graphs/add.lists)
Science field: **SF**	Mathematics
Scientific discipline: **SD**	Stochastic Analysis
Subject, Key word: **SKW**	Stochastic process, Poisson process, Exponential distribution, the principle of birth and death., Little's law, M/M/s/c queueing models, the number of servers, the capacity of the system
Holding data: **HD**	In library of Department of Mathematics and Informatics
Note: **N**	no

Abstract:
AB

This thesis is about queues and their applications in mass service systems. The first part covered the basics of probability theory and stochastic analysis. Below are presented the mass service systems with a single server, namely $M/M/1/c$ and $M/M/1$ model, as well as systems with multiple servers, these are $M/M/s/\infty$, $M/M/s/c$, $M/M/s/s$ and $M/M/\infty$ model, where s and c represent the capacity of servers and systems, respectively. For each of these models are presented the basic properties and characteristics. In addition, most of these models are illustrated with examples. It is assumed that the processes of arriving to the mass service systems are Poisson processes, and the service interarrival times are exponentially distributed. Poisson arrivals and exponential services enable to apply the Markov models of mass services that can be easily analyzed and which results can be practically applied. The common mass service model based on the principle of birth and death are presented in detail.

Accepted on Scientific Board on: **AS**	01.02.2012.
Defended: **DE**	
Thesis defended board: **DB**	
President:	dr Dora Seleši, assistant professor, Faculty of Sciences, Novi Sad
Member:	dr Sanja Rapajić, associate professor, Faculty of Science, Novi Sad
Mentor:	dr Danijela Rajter-Ćirić, full professor, Faculty of Science, Novi Sad