Verjetnost 2
Sedmo poglavje
Čakalne vrste

Matjaž Omladič

Oktober 2010

Vsebina

Splošno ravnotežje

Generični primer je zdravnikova čakalnica. Prihajajoči pacienti tvorijo tok rojstev s svojimi intenzivnostmi $\{\lambda_i | i = 0, 1, 2, \ldots\}$. Odhajajoči pacienti tvorijo tok smrti s svojimi intenzivnostmi $\{\mu_i | i = 1, 2, \ldots\}$. Včasih je smiselno privzeti, da je $\lambda_{-i} = 0$ za $i = 1, 2, \ldots$ in $\mu_{-i} = 0$ za $i = 0, 1, 2, \ldots$. Zanima nas stacionarna porazdelitev dobljenega rojstno-smrtnega procesa, če obstaja.

Iskano porazdelitev označimo s $\pi = \{\pi_i | i = 0, 1, 2, \ldots\}$, kjer je spet smiselno privzeti, da je $\pi_{-i} = 0$ za $i = 1, 2, \ldots$. Pretok verjetnosti v k-to stanje je enak

$$\lambda_{k-1}\pi_{k-1} + \mu_{k+1}\pi_{k+1},$$

pretok verjetnosti iz k-tega stanja pa je enak

$$(\lambda_k + \mu_k)\pi_k.$$

Splošno ravnotežje

Generični primer je zdravnikova čakalnica. Prihajajoči pacienti tvorijo tok rojstev s svojimi intenzivnostmi $\{\lambda_i | i = 0, 1, 2, \ldots\}$. Odhajajoči pacienti tvorijo tok smrti s svojimi intenzivnostmi $\{\mu_i | i = 1, 2, \ldots\}$. Včasih je smiselno privzeti, da je $\lambda_{-i} = 0$ za $i = 1, 2, \ldots$ in $\mu_{-i} = 0$ za $i = 0, 1, 2, \ldots$. Zanima nas stacionarna porazdelitev dobljenega rojstno-smrtnega procesa, če obstaja.

Iskano porazdelitev označimo s $\pi = \{\pi_i | i = 0, 1, 2, \ldots\}$, kjer je spet smiselno privzeti, da je $\pi_{-i} = 0$ za $i = 1, 2, \ldots$. Pretok verjetnosti v k-to stanje je enak

$$\lambda_{k-1}\pi_{k-1} + \mu_{k+1}\pi_{k+1},$$

pretok verjetnosti iz k-tega stanja pa je enak

$$(\lambda_k + \mu_k)\pi_k.$$

V ravnotežju morata biti pretoka enaka, zato je:

Splošno ravnotežje

Generični primer je zdravnikova čakalnica. Prihajajoči pacienti tvorijo tok rojstev s svojimi intenzivnostmi $\{\lambda_i | i = 0, 1, 2, \ldots\}$. Odhajajoči pacienti tvorijo tok smrti s svojimi intenzivnostmi $\{\mu_i | i = 1, 2, \ldots\}$. Včasih je smiselno privzeti, da je $\lambda_{-i} = 0$ za $i = 1, 2, \ldots$ in $\mu_{-i} = 0$ za $i = 0, 1, 2, \ldots$. Zanima nas stacionarna porazdelitev dobljenega rojstno-smrtnega procesa, če obstaja.
Iskano porazdelitev označimo s $\pi = \{\pi_i | i = 0, 1, 2, \ldots\}$, kjer je spet smiselno privzeti, da je $\pi_{-i} = 0$ za $i = 1, 2, \ldots$. Pretok verjetnosti v k-to stanje je enak

$$\lambda_{k-1}\pi_{k-1} + \mu_{k+1}\pi_{k+1},$$

pretok verjetnosti iz k-tega stanja pa je enak

$$(\lambda_k + \mu_k)\pi_k.$$

V ravnotežju morata biti pretoka enaka, zato je:

Reševanje ravnotežnih enačb — 1

$$(\lambda_k + \mu_k)\pi_k = \lambda_{k-1}\pi_{k-1} + \mu_{k+1}\pi_{k+1}. \tag{1}$$

Z zgoraj vpeljanimi dodatnimi ničlami si zapis poenostavimo tako, da veljajo te enačbe za vsako celo število k. Dobljeni sistem enačb poimenujemo *ravnotežni sistem diferenčnih enačb*, njegovo rešitev, če je nenegativna in ima vsoto enako 1, pa *ravnotežna porazdelitev*.

Izrek (Stacionarna porazdelitev)

Ravnotežna porazdelitev, če obstaja, je element levega jedra generatorja rojstno smrtnega procesa, prirejenega čakajoči vrsti, in je zato enaka stacionarni porazdelitvi tega procesa.

Iz ničte ravnotežne enačbe $\lambda_0\pi_0 = \mu_1\pi_1$ dobimo najprej

$$\pi_1 = \pi_0\frac{\lambda_0}{\mu_1}.$$

Reševanje ravnotežnih enačb — 1

$$(\lambda_k + \mu_k)\pi_k = \lambda_{k-1}\pi_{k-1} + \mu_{k+1}\pi_{k+1}. \tag{1}$$

Z zgoraj vpeljanimi dodatnimi ničlami si zapis poenostavimo tako, da veljajo te enačbe za vsako celo število k. Dobljeni sistem enačb poimenujemo *ravnotežni sistem diferenčnih enačb*, njegovo rešitev, če je nenegativna in ima vsoto enako 1, pa *ravnotežna porazdelitev*.

Izrek (Stacionarna porazdelitev)

Ravnotežna porazdelitev, če obstaja, je element levega jedra generatorja rojstno smrtnega procesa, prirejenega čakajoči vrsti, in je zato enaka stacionarni porazdelitvi tega procesa.

Iz ničte ravnotežne enačbe $\lambda_0\pi_0 = \mu_1\pi_1$ dobimo najprej

$$\pi_1 = \pi_0\frac{\lambda_0}{\mu_1}.$$

Reševanje ravnotežnih enačb – 2

Reševanje nadaljujemo rekurzivno. Končno rešitev seveda preverimo z indukcijo.

Izrek (Rešitev ravnotežnih enačb)

Rešitev sistema ravnotežnih enačb je enaka

$$\pi_k = \pi_0 \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}} \quad za \quad k = 0, 1, 2, \ldots \tag{2}$$

Za obstoj stacionarne porazdelitve moramo privzeti dodatni pogoj, da konvergira vrsta

$$S_1 = \sum_{k=1}^{\infty} \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}}. \tag{3}$$

Reševanje ravnotežnih enačb – 2

Reševanje nadaljujemo rekurzivno. Končno rešitev seveda preverimo z indukcijo.

Izrek (Rešitev ravnotežnih enačb)

Rešitev sistema ravnotežnih enačb je enaka

$$\pi_k = \pi_0 \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}} \quad \textit{za} \quad k = 0, 1, 2, \ldots \tag{2}$$

Za obstoj stacionarne porazdelitve moramo privzeti dodatni pogoj, da konvergira vrsta

$$S_1 = \sum_{k=1}^{\infty} \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}}. \tag{3}$$

Reševanje ravnotežnih enačb — 3

Izrek (Stacionarna porazdelitev)

Stacionarna (oziroma tavnotežna) porazdelitev obstaja takrat in natanko takrat, kadar je konvergentna vrsta S_1, definirana s (3). V tem primeru je stacionarna porazdelitev podana s

$$\pi_0 = \frac{1}{1+S_1}$$

in enačbami (2).

Navadno privzamemo, da so vse intenzivnosti na intervalu $(0, \infty)$. Zato je veriga nerazcepna. Kadar je konvergentna vrsta S_1, so vsa stanja rojstno-smrtnega procesa povrnljiva neničelna in veriga je ergodijska. V primeru, da ta pogoj ni izpolnjen, se izkaže, da moramo ločiti dva primera, v odvisnosti od tega, ali je konvergentna še ena pomembna vrsta.

Reševanje ravnotežnih enačb — 3

Izrek (Stacionarna porazdelitev)

Stacionarna (oziroma tavnotežna) porazdelitev obstaja takrat in natanko takrat, kadar je konvergentna vrsta S_1, definirana s (3). V tem primeru je stacionarna porazdelitev podana s

$$\pi_0 = \frac{1}{1 + S_1}$$

in enačbami (2).

Navadno privzamemo, da so vse intenzivnosti na intervalu $(0, \infty)$. Zato je veriga nerazcepna. Kadar je konvergentna vrsta S_1, so vsa stanja rojstno-smrtnega procesa povrnljiva neničelna in veriga je ergodijska. V primeru, da ta pogoj ni izpolnjen, se izkaže, da moramo ločiti dva primera, v odvisnosti od tega, ali je konvergentna še ena pomembna vrsta.

Reševanje ravnotežnih enačb — 4

$$S_2 = \sum_{k=1}^{\infty} \left[\lambda_k \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}} \right]^{-1} . \tag{4}$$

Sledenjo vrsto sta prva študirala Karlin in McGregor v povezavi s povrnljivostjo posameznih stanj. Dokazala sta izrek:

Izrek (Minljiva stanja)

Kadar veriga ni ergodijska (oziroma ekvivalentno, je vrsta (3) divergentna), potem so bodisi

- *vsa njena stanja povrnljiva ničelna, kar se zgodi natanko takrat, kadar je vrsta (4) divergentna, bodisi*
- *so vsa njena stanja minljiva, kar se zgodi natanko takrat, kadar je vrsta (4) konvergentna.*

Reševanje ravnotežnih enačb — 4

$$S_2 = \sum_{k=1}^{\infty} \left[\lambda_k \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}} \right]^{-1} . \tag{4}$$

Slednjo vrsto sta prva študirala Karlin in McGregor v povezavi s povrnljivostjo posameznih stanj. Dokazala sta izrek:

Izrek (Minljiva stanja)

Kadar veriga ni ergodijska (oziroma ekvivalentno, je vrsta (3) divergentna), potem so bodisi

- *vsa njena stanja povrnljiva ničelna, kar se zgodi natanko takrat, kadar je vrsta (4) divergentna, bodisi*
- *so vsa njena stanja minljiva, kar se zgodi natanko takrat, kadar je vrsta (4) konvergentna.*

Ravnotežje sistema $M/M/1$

Klasični primer čakalnega sistema. Prihodi v sistem tvorijo Poissonov tok s paramterom λ. Sistem ima enega samega strežnika (v generičnem primeru zdravnika), ki obdeluje čakajoče zahteve drugo za drugo eksponentno s parametrom μ. Tako rojstne kot smrtne intenzivnosti so konstantne:

$$\lambda_k = \lambda \quad \text{za} \quad k = 0, 1, 2, \ldots$$

$$\mu_k = \mu \quad \text{za} \quad k = 1, 2, 3, \ldots$$

Od tod izračunamo potencialno ravnotežno porazdelitev:

$$\pi_k = \pi_0 \left(\frac{\lambda}{\mu}\right)^k \quad \text{za} \quad k = 0, 1, 2, \ldots.$$

Stacionarna porazdelitev bo potem res obstajala natanko tedaj, kadar bo konvergentna vrsta S_1, to pa bo natanko tedaj, kadar bo $\lambda < \mu$. Stacionarno porazdelitev v tem primeru izračunamo. Dobimo geometrijsko porazdelitev s parametrom $q = \dfrac{\lambda}{\mu}$.

Ravnotežje sistema $M/M/1$

Klasični primer čakalnega sistema. Prihodi v sistem tvorijo Poissonov tok s paramterom λ. Sistem ima enega samega strežnika (v generičnem primeru zdravnika), ki obdeluje čakajoče zahteve drugo za drugo eksponentno s parametrom μ. Tako rojstne kot smrtne intenzivnosti so konstantne:

$$\lambda_k = \lambda \quad \text{za} \quad k = 0, 1, 2, \ldots$$

$$\mu_k = \mu \quad \text{za} \quad k = 1, 2, 3, \ldots$$

Od tod izračunamo potencialno ravnotežno porazdelitev:

$$\pi_k = \pi_0 \left(\frac{\lambda}{\mu}\right)^k \quad \text{za} \quad k = 0, 1, 2, \ldots .$$

Stacionarna porazdelitev bo potem res obstajala natanko tedaj, kadar bo konvergentna vrsta S_1, to pa bo natanko tedaj, kadar bo $\lambda < \mu$. Stacionarno porazdelitev v tem primeru izračunamo. Dobimo geometrijsko porazdelitev s parametrom $q = \dfrac{\lambda}{\mu}$.

Stacionarna porazdelitev

Preverimo še drugi pogoj

$$S_2 = \sum_{k=1}^{\infty} \frac{\mu^k}{\lambda^{k+1}}.$$

Ko je $\mu < \lambda$, je vrsta S_2 konvergentna, vsa stanja so minljiva.
V mejnem primeru $\lambda = \mu$ so vsa stanja povrnljiva ničelna.
Najprej izračunamo povprečno število strank v sistemu

$$E(N) = \frac{q}{1-q},$$

disperzija pa je

$$D(N) = \frac{q}{(1-q)^2}.$$

Littlovo pravilo za povprečni časa bivanja stranke v sistemu:

$$T = \frac{E(N)}{\lambda} = \frac{1}{\mu - \lambda}.$$

Verjetnost dogodka, da bo v sistemu vsaj k strank je q^k.

Stacionarna porazdelitev

Preverimo še drugi pogoj

$$S_2 = \sum_{k=1}^{\infty} \frac{\mu^k}{\lambda^{k+1}}.$$

Ko je $\mu < \lambda$, je vrsta S_2 konvergentna, vsa stanja so minljiva. V mejnem primeru $\lambda = \mu$ so vsa stanja povrnljiva ničelna. Najprej izračunamo povprečno število strank v sistemu

$$E(N) = \frac{q}{1-q},$$

disperzija pa je

$$D(N) = \frac{q}{(1-q)^2}.$$

Littlovo pravilo za povprečni časa bivanja stranke v sistemu:

$$T = \frac{E(N)}{\lambda} = \frac{1}{\mu - \lambda}.$$

Verjetnost dogodka, da bo v sistemu vsaj k strank je q^k.

Nekaj količin

Oglejmo si nekaj količin, ki zanimajo uporabnike, upravljalce, organizatorje in lastnike strežnih sistemov:

- $\alpha(t)$ naj bo število strank, ki so prišle v sistem v časovnem intervalu $(0, t)$.
- $\delta(t)$ naj bo število strank, ki so odšle iz sistema v časovnem intervalu $(0, t)$.
- Privzamemo, da na začetku v sistemu ni bilo strank, torej je skupno število strank v sistemu v času t

$$N(t) = \alpha(t) - \delta(t).$$

- $\gamma(t)$ je skupni čas, ki so ga vse stranke prebile v sistemu v intervalu $(0, t)$, merjen v enoti stranka$\times$sekunda. Dobimo ga po formuli

$$\gamma(t) = \int_0^t N(s)\, ds.$$

Nekaj količin

Oglejmo si nekaj količin, ki zanimajo uporabnike, upravljalce, organizatorje in lastnike strežnih sistemov:

- $\alpha(t)$ naj bo število strank, ki so prišle v sistem v časovnem intervalu $(0, t)$.
- $\delta(t)$ naj bo število strank, ki so odšle iz sistema v časovnem intervalu $(0, t)$.
- Privzamemo, da na začetku v sistemu ni bilo strank, torej je skupno število strank v sistemu v času t

$$N(t) = \alpha(t) - \delta(t).$$

- $\gamma(t)$ je skupni čas, ki so ga vse stranke prebile v sistemu v intervalu $(0, t)$, merjen v enoti stranka$\times$sekunda. Dobimo ga po formuli

$$\gamma(t) = \int_0^t N(s)\, ds.$$

Nekaj količin — nadaljevanje

- λ_t, tj povprečno intenzivnost prihodov strank v časovnem intervalu $(0, t)$ dobimo kot $\alpha(t)/t$.
- $\overline{N}_t$, povprečno število strank v sistemu na intervalu $(0, t)$, pa je potem enako
$$\overline{N}_t = \frac{\gamma(t)}{t}.$$
- T_t je povprečni čas, ki ga stranka prebije v sistemu v intervalu $(0, t)$. Dobimo ga po formuli
$$T_t = \frac{\gamma(t)}{\alpha(t)} = \frac{\overline{N}_t}{\lambda_t}.$$

Slednjo izpeljavo iz definicij prepišemo v ti *Littlovo pravilo*:
$$\overline{N}_t = \lambda_t T_t, \tag{5}$$
Opazimo, da so vsa časovna *povprečja* označena po principu $\cdot_t$; pri teh pričakujemo v ravnotežnih pogojih obstoj limite, ko gre t proti neskončnosti, v oznaki za limito pa potem t izpustimo.

Nekaj količin — nadaljevanje

- λ_t, tj povprečno intenzivnost prihodov strank v časovnem intervalu $(0, t)$ dobimo kot $\alpha(t)/t$.
- $\overline{N}_t$, povprečno število strank v sistemu na intervalu $(0, t)$, pa je potem enako
$$\overline{N}_t = \frac{\gamma(t)}{t}.$$
- T_t je povprečni čas, ki ga stranka prebije v sistemu v intervalu $(0, t)$. Dobimo ga po formuli
$$T_t = \frac{\gamma(t)}{\alpha(t)} = \frac{\overline{N}_t}{\lambda_t}.$$

Slednjo izpeljavo iz definicij prepišemo v ti *Littlovo pravilo*:

$$\overline{N}_t = \lambda_t T_t, \tag{5}$$

Opazimo, da so vsa časovna *povprečja* označena po principu $\cdot_t$; pri teh pričakujemo v ravnotežnih pogojih obstoj limite, ko gre t proti neskončnosti, v oznaki za limito pa potem t izpustimo.

Uporabnostni faktor ρ

Opazimo, da so vsa povprečja mišljena kot povprečja glede na čas. Vse količine pa so v resnici tudi slučajne spremenljivke. Pojem uporabnostnega faktorja prihaja iz teorije *(pretočnih) sistemov*. V splošnem gre za kvocient $\rho = R/C$, kjer je R vstopni tok dobrin, ki prihajajo v sistem in C ti kapaciteta sistema, torej količina dobrin, ki jih sistem uspe obdelati. Obe količini sta izraženi v enoti, v kateri merimo dobrino, na časovno enoto, kvocient pa je potem brez enot. Kadar je $\rho > 1$, ostaja dobrina neobdelana na vhodu v sitem, iz sistema pa prihaja C dobrin; kadar pa je $\rho \leqslant 1$, potem je izhod iz sstema enak vhodu R.

Uporabnostni faktor ρ

Opazimo, da so vsa povprečja mišljena kot povprečja glede na čas. Vse količine pa so v resnici tudi slučajne spremenljivke. Pojem uporabnostnega faktorja prihaja iz teorije *(pretočnih) sistemov*. V splošnem gre za kvocient $\rho = R/C$, kjer je R vstopni tok dobrin, ki prihajajo v sistem in C ti kapaciteta sistema, torej količina dobrin, ki jih sistem uspe obdelati. Obe količini sta izraženi v enoti, v kateri merimo dobrino, na časovno enoto, kvocient pa je potem brez enot. Kadar je $\rho > 1$, ostaja dobrina neobdelana na vhodu v sitem, iz sistema pa prihaja C dobrin; kadar pa je $\rho \leqslant 1$, potem je izhod iz sstema enak vhodu R.

V primeru strežnih sistemov lahko izračunamo ρ iz podatkov λ (povprečna intenzivnost prihodov strank v časovnem intervalu $(0, t)$ v limiti, ko pošljemo t proti neskončnosti) in $\overline{x}$ (povprečen čas, ki ga stranka prebije v strežbi) po formuli $\rho = \lambda \overline{x}$. V primeru sistema $M/M/1$ je $\rho = \lambda/\mu = q$.

Uporabnostni faktor ρ

Opazimo, da so vsa povprečja mišljena kot povprečja glede na čas. Vse količine pa so v resnici tudi slučajne spremenljivke. Pojem uporabnostnega faktorja prihaja iz teorije *(pretočnih) sistemov*. V splošnem gre za kvocient $\rho = R/C$, kjer je R vstopni tok dobrin, ki prihajajo v sistem in C ti kapaciteta sistema, torej količina dobrin, ki jih sistem uspe obdelati. Obe količini sta izraženi v enoti, v kateri merimo dobrino, na časovno enoto, kvocient pa je potem brez enot. Kadar je $\rho > 1$, ostaja dobrina neobdelana na vhodu v sitem, iz sistema pa prihaja C dobrin; kadar pa je $\rho \leqslant 1$, potem je izhod iz sstema enak vhodu R.
V primeru strežnih sistemov lahko izračunamo ρ iz podatkov λ (povprečna intenzivnost prihodov strank v časovnem intervalu $(0, t)$ v limiti, ko pošljemo t proti neskončnosti) in $\overline{x}$ (povprečen čas, ki ga stranka prebije v strežbi) po formuli $\rho = \lambda\overline{x}$. V primeru sistema $M/M/1$ je $\rho = \lambda/\mu = q$.

Odvračani prihodi

Nekoliko bolj realno od modela $M/M/1$ je privzeti, da stranke dolge čakajoče vrste odvračajo od čakanja. Najpreprostejše je *harmonično odvračanje*, ki ga modeliramo z intenzivnostmi

$$\lambda_k = \frac{\alpha}{k+1} \quad \text{za} \quad k = 0, 1, 2, \ldots$$

$$\mu_k = \mu \quad \text{za} \quad k = 1, 2, 3, \ldots$$

Potencialna ravnotežna porazdelitev je:

$$\pi_k = \pi_0 \frac{1}{k!} \left(\frac{\alpha}{\mu}\right)^k \quad \text{za} \quad k = 0, 1, 2, \ldots.$$

Ergodijski pogoj nam da tokrat le $\alpha/\mu < \infty$. Pripelje nas do $\pi_0 = e^{-\alpha/\mu}$ in do Poissonove porazdelitve. Povprečno število strank v sistemu je $E(N) = \overline{N} = \frac{\alpha}{\mu}$. Littlovo pravilo nam da:

$$T = \frac{E(N)}{\lambda} = \frac{\alpha}{\mu\lambda}.$$

Odvračani prihodi

Nekoliko bolj realno od modela $M/M/1$ je privzeti, da stranke dolge čakajoče vrste odvračajo od čakanja. Najpreprostejše je *harmonično odvračanje*, ki ga modeliramo z intenzivnostmi

$$\lambda_k = \frac{\alpha}{k+1} \quad \text{za} \quad k = 0, 1, 2, \dots$$

$$\mu_k = \mu \quad \text{za} \quad k = 1, 2, 3, \dots$$

Potencialna ravnotežna porazdelitev je:

$$\pi_k = \pi_0 \frac{1}{k!} \left(\frac{\alpha}{\mu}\right)^k \quad \text{za} \quad k = 0, 1, 2, \dots .$$

Ergodijski pogoj nam da tokrat le $\alpha/\mu < \infty$. Pripelje nas do $\pi_0 = e^{-\alpha/\mu}$ in do Poissonove porazdelitve. Povprečno število strank v sistemu je $E(N) = \overline{N} = \frac{\alpha}{\mu}$. Littlovo pravilo nam da:

$$T = \frac{E(N)}{\lambda} = \frac{\alpha}{\mu\lambda}.$$

Uporabnostni faktor v sistemih $G/G/1$

Toda kaj vzeti v tem primeru za povprečno intenzivnost prihodov λ. V sistemih tipa $G/G/1$ jo lahko izračunamo iz uporabnostnega faktorje ρ. V nekem zelo dolgem časovnem intervalu dolžin τ lahko po ZVŠ z verjetnostjo 1 pričakujemo $\lambda\tau$ strank. Ker je verjetnost, da je (edini) strežnik brezdelen enaka p_0, je v tem časovnem intervalu strežnik zaseden $\tau - \tau p_0$ časa in lahko obdela $(\tau - \tau p_0)/\overline{x}$ zahtev. V ravnotežnem stanju mora veljati

$$\lambda\tau = \frac{\tau - \tau p_0}{\overline{x}}.$$

Ta izraz nam pomaga izračunati faktor ρ iz ravnotežne porazdelitve $\rho = \lambda\overline{x} = 1 - p_0$, od tod pa lahko tudi intenzivnost λ izrazimo iz ravnotežne porazdelitve $\lambda = \rho/\overline{x} = (1 - p_0)/\overline{x}$. V primeru odvračanih prihodov dobimo

$$T = \frac{\alpha}{\mu^2(1 - e^{-\alpha/\mu})}.$$

Uporabnostni faktor v sistemih $G/G/1$

Toda kaj vzeti v tem primeru za povprečno intenzivnost prihodov λ. V sistemih tipa $G/G/1$ jo lahko izračunamo iz uporabnostnega faktorje ρ. V nekem zelo dolgem časovnem intervalu dolžin τ lahko po ZVŠ z verjetnostjo 1 pričakujemo $\lambda\tau$ strank. Ker je verjetnost, da je (edini) strežnik brezdelen enaka p_0, je v tem časovnem intervalu strežnik zaseden $\tau - \tau p_0$ časa in lahko obdela $(\tau - \tau p_0)/\overline{x}$ zahtev. V ravnotežnem stanju mora veljati

$$\lambda\tau = \frac{\tau - \tau p_0}{\overline{x}}.$$

Ta izraz nam pomaga izračunati faktor ρ iz ravnotežne porazdelitve $\rho = \lambda\overline{x} = 1 - p_0$, od tod pa lahko tudi intenzivnost λ izrazimo iz ravnotežne porazdelitve $\lambda = \rho/\overline{x} = (1 - p_0)/\overline{x}$. V primeru odvračanih prihodov dobimo

$$T = \frac{\alpha}{\mu^2(1 - e^{-\alpha/\mu})}.$$

Odzivni strežniki

Eden od ekstremnih primerov strežnega sistema je tudi *sistem z odzivnimi strežniki*. V tem sistemu privzamemo, da ima sistem na voljo poljubno mnogo strežnikov, ki se vključujejo v strežbo v odvisnosti od prihajanja strank. Modeliramo ga z intenzivnostmi

$$\lambda_k = \lambda \quad \text{za} \quad k = 0, 1, 2, \ldots$$

$$\mu_k = k\mu \quad \text{za} \quad k = 1, 2, 3, \ldots$$

Potencialna ravnotežna porazdelitev je:

$$\pi_k = \pi_0 \frac{1}{k!} \left(\frac{\lambda}{\mu}\right)^k \quad \text{za} \quad k = 0, 1, 2, \ldots .$$

Ergodijski pogoj nam da $\lambda/\mu < \infty$. Pripelje nas do $\pi_0 = e^{-\lambda/\mu}$ in do Poissonove porazdelitve. Povprečno število strank v sistemu je $E(N) = \overline{N} = \frac{\lambda}{\mu}$, vendar pa je tokrat $T = \frac{1}{\mu}$.

Odzivni strežniki

Eden od ekstremnih primerov strežnega sistema je tudi *sistem z odzivnimi strežniki*. V tem sistemu privzamemo, da ima sistem na voljo poljubno mnogo strežnikov, ki se vključujejo v strežbo v odvisnosti od prihajanja strank. Modeliramo ga z intenzivnostmi

$$\lambda_k = \lambda \quad \text{za} \quad k = 0, 1, 2, \ldots$$

$$\mu_k = k\mu \quad \text{za} \quad k = 1, 2, 3, \ldots$$

Potencialna ravnotežna porazdelitev je:

$$\pi_k = \pi_0 \frac{1}{k!} \left(\frac{\lambda}{\mu}\right)^k \quad \text{za} \quad k = 0, 1, 2, \ldots.$$

Ergodijski pogoj nam da $\lambda/\mu < \infty$. Pripelje nas do $\pi_0 = e^{-\lambda/\mu}$ in do Poissonove porazdelitve. Povprečno število strank v sistemu je $E(N) = \overline{N} = \frac{\lambda}{\mu}$, vendar pa je tokrat $T = \frac{1}{\mu}$.

Sistem $M/M/m$

V tem sistemu spet privzamemo konstantno intenzivnost prihodov in neskončno čakalnico. Imamo pa na voljo več strežnikov, a največ m. Sistem spet modeliramo z rojstno-smrtnim procesom z intenzivnostmi

$$\lambda_k = \lambda \quad \text{za} \quad k = 0, 1, 2, \ldots$$

$$\mu_k = \begin{cases} k\mu, & \text{če } 0 \leqslant k \leqslant m; \\ m\mu, & \text{če } m \leqslant k. \end{cases} \qquad \text{za} \quad k = 1, 2, 3, \ldots$$

Označimo $q = \frac{\lambda}{m\mu}$; potencialna ravnotežna porazdelitev je:

$$\pi_k = \begin{cases} \pi_0 \frac{(mq)^k}{k!}, & \text{če } k \leqslant m; \\ \pi_0 \frac{q^k m^m}{m!}, & \text{če } k \geqslant m. \end{cases} \qquad \text{za} \quad k = 0, 1, 2, \ldots .$$

Ergodijski pogoj nam da $\lambda < m\mu$ in nas pripelje do

$$\pi_0 = \left[\sum_{k=0}^{m-1} \frac{(mq)^k}{k!} + \frac{(mq)^m}{m!} \frac{1}{1-q} \right]^{-1} .$$

Sistem $M/M/m$

V tem sistemu spet privzamemo konstantno intenzivnost prihodov in neskončno čakalnico. Imamo pa na voljo več strežnikov, a največ m. Sistem spet modeliramo z rojstno-smrtnim procesom z intenzivnostmi

$$\lambda_k = \lambda \quad \text{za} \quad k = 0, 1, 2, \dots$$

$$\mu_k = \begin{cases} k\mu, & \text{če } 0 \leqslant k \leqslant m; \\ m\mu, & \text{če } m \leqslant k. \end{cases} \qquad \text{za} \quad k = 1, 2, 3, \dots$$

Označimo $q = \frac{\lambda}{m\mu}$; potencialna ravnotežna porazdelitev je:

$$\pi_k = \begin{cases} \pi_0 \frac{(mq)^k}{k!}, & \text{če } k \leqslant m; \\ \pi_0 \frac{q^k m^m}{m!}, & \text{če } k \geqslant m. \end{cases} \qquad \text{za} \quad k = 0, 1, 2, \dots.$$

Ergodijski pogoj nam da $\lambda < m\mu$ in nas pripelje do

$$\pi_0 = \left[\sum_{k=0}^{m-1} \frac{(mq)^k}{k!} + \frac{(mq)^m}{m!} \frac{1}{1-q} \right]^{-1}.$$

Sistem $M/M/1/K$ s končno čakalnico

Od tod dobimo *Erlangovo C formulo* iz telefonije za verjetnost dogodka, da bo stranka v sistemu z m strežniki čakala

$$C(m,\lambda/\mu) = \frac{(mq)^m}{m!}\frac{1}{1-q}\left[\sum_{k=0}^{m-1}\frac{(mq)^k}{k!} + \frac{(mq)^m}{m!}\frac{1}{1-q}\right]^{-1}.$$

Tudi problem končne čakalnice, kjer pomeni K skupno število čakalnih in strežnih mest, lahko še zmerom modeliramo z rojstno-smrtnim procesom. Dobimo

$$\lambda_k = \begin{cases} \lambda, & \text{če } k < K; \\ 0, & \text{sicer;} \end{cases} \quad \text{in} \quad \mu_k = \mu \quad \text{za} \quad k = 1, 2, \ldots, K.$$

Sistem $M/M/1/K$ s končno čakalnico

Od tod dobimo *Erlangovo C formulo* iz telefonije za verjetnost dogodka, da bo stranka v sistemu z m strežniki čakala

$$C(m, \lambda/\mu) = \frac{(mq)^m}{m!}\frac{1}{1-q}\left[\sum_{k=0}^{m-1}\frac{(mq)^k}{k!} + \frac{(mq)^m}{m!}\frac{1}{1-q}\right]^{-1}.$$

Tudi problem končne čakalnice, kjer pomeni K skupno število čakalnih in strežnih mest, lahko še zmerom modeliramo z rojstno-smrtnim procesom. Dobimo

$$\lambda_k = \begin{cases} \lambda, & \text{če } k < K; \\ 0, & \text{sicer;} \end{cases} \quad \text{in} \quad \mu_k = \mu \quad \text{za} \quad k = 1, 2, \ldots, K.$$

To je proces s končno mnogo stanji, zmerom ergodijski. Vpeljemo $q = \lambda/\mu$, da dobimo $\pi_k = \pi_0 q^k$ za $k = 1, 2, \ldots, K$ in $\pi_k = 0$ sicer. Iz normalizacijskega pogoja dobimo

$$\pi_0 = \frac{1-q}{1-q^{K+1}}.$$

Sistem $M/M/1/K$ s končno čakalnico

Od tod dobimo *Erlangovo C formulo* iz telefonije za verjetnost dogodka, da bo stranka v sistemu z m strežniki čakala

$$C(m, \lambda/\mu) = \frac{(mq)^m}{m!}\frac{1}{1-q}\left[\sum_{k=0}^{m-1}\frac{(mq)^k}{k!} + \frac{(mq)^m}{m!}\frac{1}{1-q}\right]^{-1}.$$

Tudi problem končne čakalnice, kjer pomeni K skupno število čakalnih in strežnih mest, lahko še zmerom modeliramo z rojstno-smrtnim procesom. Dobimo

$$\lambda_k = \begin{cases} \lambda, & \text{če } k < K; \\ 0, & \text{sicer}; \end{cases} \quad \text{in} \quad \mu_k = \mu \quad \text{za} \quad k = 1, 2, \ldots, K.$$

To je proces s končno mnogo stanji, zmerom ergodijski. Vpeljemo $q = \lambda/\mu$, da dobimo $\pi_k = \pi_0 q^k$ za $k = 1, 2, \ldots, K$ in $\pi_k = 0$ sicer. Iz normalizacijskega pogoja dobimo

$$\pi_0 = \frac{1-q}{1-q^{K+1}}.$$

Sistem m strežnikov z izgubo strank $M/M/m/m$

Imamo m strežnikov brez dodatnih čakalnih mest. Stranka, ki se pojavi, ko so vsi strežniki zasedeni, je izgubljena (blocked calls cleared). Sistem modeliramo z intenzivnostmi:

$$\lambda_k = \begin{cases} \lambda, & \text{če } k < m; \\ 0, & \text{sicer;} \end{cases} \quad \text{in} \quad \mu_k = k\mu \quad \text{za} \quad k = 1, 2, \ldots, m.$$

Vpeljemo $q = \lambda/\mu$, da dobimo $\pi_k = \pi_0 \frac{q^k}{k!}$ za $k = 0, 1, 2, \ldots, m$ in $\pi_k = 0$ sicer. Iz normalizacijskega pogoja dobimo

Sistem m strežnikov z izgubo strank $M/M/m/m$

Imamo m strežnikov brez dodatnih čakalnih mest. Stranka, ki se pojavi, ko so vsi strežniki zasedeni, je izgubljena (blocked calls cleared). Sistem modeliramo z intenzivnostmi:

$$\lambda_k = \begin{cases} \lambda, & \text{če } k < m; \\ 0, & \text{sicer;} \end{cases} \quad \text{in} \quad \mu_k = k\mu \quad \text{za} \quad k = 1, 2, \ldots, m.$$

Vpeljemo $q = \lambda/\mu$, da dobimo $\pi_k = \pi_0 \frac{q^k}{k!}$ za $k = 0, 1, 2, \ldots, m$ in $\pi_k = 0$ sicer. Iz normalizacijskega pogoja dobimo

$$\pi_0 = \left[\sum_{k=0}^{m} \frac{q^k}{k!} \right]^{-1}.$$

Od tod dobimo delež časa, ko so vsi strežniki zasedeni kot

$$B(m, \lambda/\mu) = \frac{1}{m!} \left(\frac{\lambda}{\mu} \right)^m \left[\sum_{k=0}^{m} \frac{1}{k!} \left(\frac{\lambda}{\mu} \right)^k \right]^{-1}.$$

Erlangova B formula = verjetnost dogodka, da bo klic zgubljen.

Sistem m strežnikov z izgubo strank $M/M/m/m$

Imamo m strežnikov brez dodatnih čakalnih mest. Stranka, ki se pojavi, ko so vsi strežniki zasedeni, je izgubljena (blocked calls cleared). Sistem modeliramo z intenzivnostmi:

$$\lambda_k = \begin{cases} \lambda, & \text{če } k < m; \\ 0, & \text{sicer}; \end{cases} \quad \text{in} \quad \mu_k = k\mu \quad \text{za} \quad k = 1, 2, \ldots, m.$$

Vpeljemo $q = \lambda/\mu$, da dobimo $\pi_k = \pi_0 \frac{q^k}{k!}$ za $k = 0, 1, 2, \ldots, m$ in $\pi_k = 0$ sicer. Iz normalizacijskega pogoja dobimo

$$\pi_0 = \left[\sum_{k=0}^{m} \frac{q^k}{k!} \right]^{-1}.$$

Od tod dobimo delež časa, ko so vsi strežniki zasedeni kot

$$B(m, \lambda/\mu) = \frac{1}{m!} \left(\frac{\lambda}{\mu} \right)^m \left[\sum_{k=0}^{m} \frac{1}{k!} \left(\frac{\lambda}{\mu} \right)^k \right]^{-1}.$$

Erlangova B formula = verjetnost dogodka, da bo klic zgubljen.

$M/M/1//M$ — en strežnik, končno mnogo strank

Skupno M uporabnikov sistema prihaja v sistem eksponentno s parametrom λ. Z naraščajočim številom strank v sistemu se zmanjšuje število strank izven sistema in s tem pritisk na nadaljnje prihajanje in obratno, zato deluje sistem samoregulativno. Sistem modeliramo z intenzivnostmi:

$$\lambda_k = \begin{cases} \lambda(M-k), & \text{če } 0 \leqslant k \leqslant M; \\ 0, & \text{sicer;} \end{cases} \quad \text{in} \quad \mu_k = \mu \quad \text{za} \quad k = 1, 2, \ldots$$

$M/M/1//M$ – en strežnik, končno mnogo strank

Skupno M uporabnikov sistema prihaja v sistem eksponentno s parametrom λ. Z naraščajočim številom strank v sistemu se zmanjšuje število strank izven sistema in s tem pritisk na nadaljnje prihajanje in obratno, zato deluje sistem samoregulativno. Sistem modeliramo z intenzivnostmi:

$$\lambda_k = \begin{cases} \lambda(M-k), & \text{če } 0 \leqslant k \leqslant M; \\ 0, & \text{sicer;} \end{cases} \quad \text{in} \quad \mu_k = \mu \quad \text{za} \quad k = 1, 2, \ldots$$

Vpeljemo $q = \lambda/\mu$, da dobimo

$$\pi_k = \pi_0 q^k \frac{M!}{(M-k)!}$$

za $k = 0, 1, 2, \ldots, M$ in $\pi_k = 0$ sicer. Iz normalizacijskega pogoja dobimo

$$\pi_0 = \left[\sum_{k=0}^{m} q^k \frac{M!}{(M-k)!}\right]^{-1}.$$

$M/M/1//M$ — en strežnik, končno mnogo strank

Skupno M uporabnikov sistema prihaja v sistem eksponentno s parametrom λ. Z naraščajočim številom strank v sistemu se zmanjšuje število strank izven sistema in s tem pritisk na nadaljnje prihajanje in obratno, zato deluje sistem samoregulativno. Sistem modeliramo z intenzivnostmi:

$$\lambda_k = \begin{cases} \lambda(M-k), & \text{če } 0 \leqslant k \leqslant M; \\ 0, & \text{sicer;} \end{cases} \quad \text{in} \quad \mu_k = \mu \quad \text{za} \quad k = 1, 2, \ldots$$

Vpeljemo $q = \lambda/\mu$, da dobimo

$$\pi_k = \pi_0 q^k \frac{M!}{(M-k)!}$$

za $k = 0, 1, 2, \ldots, M$ in $\pi_k = 0$ sicer. Iz normalizacijskega pogoja dobimo

$$\pi_0 = \left[\sum_{k=0}^{m} q^k \frac{M!}{(M-k)!} \right]^{-1}.$$

$M/M/\infty//M$ – poljubno strežnikov na končno mnogo strank

Tokrat imamo za vsakega od M uporabnikov svoj strežnik, tako da dobimo intenzivnosti:

$$\lambda_k = \begin{cases} \lambda(M-k), & \text{če } 0 \leqslant k \leqslant M; \\ 0, & \text{sicer}; \end{cases} \quad \text{in} \quad \mu_k = k\mu \quad \text{za} \quad k = 1, 2, \ldots$$

Vpeljemo $q = \lambda/\mu$, da dobimo

$$\pi_k = \pi_0 q^k \binom{M}{k}$$

za $k = 0, 1, 2, \ldots, M$ in $\pi_k = 0$ sicer. Iz normalizacijskega pogoja dobimo

$$\pi_0 = (1+q)^{-M}.$$

Lahko je izračunati še nekatere druge količine, npr

$$\overline{N} = \frac{Mq}{1+q}.$$

$M/M/\infty//M$ – poljubno strežnikov na končno mnogo strank

Tokrat imamo za vsakega od M uporabnikov svoj strežnik, tako da dobimo intenzivnosti:

$$\lambda_k = \begin{cases} \lambda(M-k), & \text{če } 0 \leqslant k \leqslant M; \\ 0, & \text{sicer}; \end{cases} \quad \text{in} \quad \mu_k = k\mu \quad \text{za} \quad k = 1, 2, \ldots$$

Vpeljemo $q = \lambda/\mu$, da dobimo

$$\pi_k = \pi_0 q^k \binom{M}{k}$$

za $k = 0, 1, 2, \ldots, M$ in $\pi_k = 0$ sicer. Iz normalizacijskega pogoja dobimo

$$\pi_0 = (1+q)^{-M}.$$

Lahko je izračunati še nekatere druge količine, npr

$$\overline{N} = \frac{Mq}{1+q}.$$

Pa še kombinacija $M/M/m/K/M$

Smiselno je privzeti, da je $m \leqslant K \leqslant M$, kar nas pripelje do intenzivnosti:

$$\lambda_k = \begin{cases} \lambda(M-k), & \text{če } 0 \leqslant k \leqslant K-1; \\ 0, & \text{sicer}; \end{cases} \quad \text{in} \quad \mu_k = \begin{cases} k\mu, & \text{če } 0 \ldots \\ m\mu, & \text{sicer.} \end{cases}$$

Vpeljemo $q = \lambda/\mu$, da dobimo najprej pri $0 \leqslant k \leqslant m-1$

$$\pi_k = \pi_0 q^k \binom{M}{k}.$$

Pri $m \leqslant k \leqslant K$ pa dobimo še

$$\pi_k = \pi_0 q^k \binom{M}{k} \frac{k!}{m!} m^{m-k}.$$

Normalizacijski pogoj nas spet pripelje do izraza za π_0, ki je še nekoliko bolj zapleten, kot vsi doslej.

Pa še kombinacija $M/M/m/K/M$

Smiselno je privzeti, da je $m \leqslant K \leqslant M$, kar nas pripelje do intenzivnosti:

$$\lambda_k = \begin{cases} \lambda(M-k), & \text{če } 0 \leqslant k \leqslant K-1; \\ 0, & \text{sicer;} \end{cases} \quad \text{in} \quad \mu_k = \begin{cases} k\mu, & \text{če } 0 \\ m\mu, & \text{sicer.} \end{cases}$$

Vpeljemo $q = \lambda/\mu$, da dobimo najprej pri $0 \leqslant k \leqslant m-1$

$$\pi_k = \pi_0 q^k \binom{M}{k}.$$

Pri $m \leqslant k \leqslant K$ pa dobimo še

$$\pi_k = \pi_0 q^k \binom{M}{k} \frac{k!}{m!} m^{m-k}.$$

Normalizacijski pogoj nas spet pripelje do izraza za π_0, ki je še nekoliko bolj zapleten, kot vsi doslej.