Kapitel 3

Markov-Modelle II

3.1 Ankunft oder Abfertigung in Gruppen

3.1.1 Das $M^{[X]}/M/1$ Modell

Wir nehmen an, dass die Ankünfte einem Poisson-Prozess mit Rate λ genügen und zu jedem Ankunftszeitpunkt T eine Anzahl X_T Kunden gleichzeitig eintrifft, wobei die X_T unabhängige und identisch verteilte Zufallsvariablen sind. Dieser Ankunftsprozess heißt *Poisson-Prozess mit Mehrfachpunkten* bzw *Compound Poisson-Prozess*. Da die Bedienungszeiten exponentialverteilt sind, ist die Anzahl der Kunden im System eine Markov-Kette. Mit der Bezeichnung $c_x = \Pr[X_T = x]$ gilt für ihren infinitesimalen Generator

$$\mathbf{Q} = \begin{pmatrix} -\lambda & \lambda c_1 & \lambda c_2 & \lambda c_3 & \dots \\ \mu & -\lambda-\mu & \lambda c_1 & \lambda c_2 & \dots \\ 0 & \mu & -\lambda-\mu & \lambda c_1 & \dots \\ 0 & 0 & \mu & -\lambda-\mu & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}.$$

Damit muss die Gleichgewichtsverteilung folgendes Gleichungssystem erfüllen ($n \geq 1$)

$$\begin{aligned} -\lambda p_0 + \mu p_1 &= 0 \\ \lambda \sum_{k=1}^{n} c_k p_{n-k} - (\lambda+\mu) p_n + \mu p_{n+1} &= 0. \end{aligned}$$

Wir werden das System mittels erzeugender Funktionen lösen. Seien $P(z) = \sum_{n=0}^{\infty} p_n z^n$ und $C(z) = \sum_{n=1}^{\infty} c_n z^n$, so erhalten wir

$$\begin{aligned} \lambda \sum_{n=1}^{\infty} \sum_{k=1}^{n} c_k z^k p_{n-k} z^{n-k} - (\lambda+\mu) \sum_{n=1}^{\infty} p_n z^n + \frac{\mu}{z} \sum_{n=1}^{\infty} p_{n+1} z^{n+1} &= 0 \\ \lambda \sum_{k=1}^{\infty} c_k z^k \sum_{n=k}^{\infty} p_{n-k} z^{n-k} - (\lambda+\mu) \sum_{n=1}^{\infty} p_n z^n + \frac{\mu}{z} \sum_{n=2}^{\infty} p_n z^n &= 0 \\ \lambda C(z) P(z) - (\lambda+\mu)(P(z) - p_0) + \frac{\mu}{z}(P(z) - p_0 - p_1 z) &= 0. \end{aligned}$$

Setzen wir nun die erste obige Gleichung ein, so ergibt sich

$$\lambda C(z) P(z) - (\lambda + \mu) P(z) + \mu \left(1 - \frac{1}{z}\right) p_0 + \frac{\mu}{z} P(z) = 0,$$

also

$$P(z) = \frac{\mu (1-z) p_0}{z\lambda (C(z) - 1) + \mu (1-z)}.$$

Da $P(1) = 1$ und $C(1) = 1$ gelten muss, folgt mit der Regel von De l'Hospital

$$1 = \frac{-\mu p_0}{\lambda C'(1) - \mu} = \frac{\mu p_0}{-\lambda \mathbb{E}[X] + \mu}.$$

Damit die Gleichgewichtsverteilung existiert, muss $\mathbb{E}[X] < \mu/\lambda$ gelten. Dann gilt

$$p_0 = 1 - \frac{\lambda}{\mu} \mathbb{E}[X] = 1 - \rho.$$

Um die p_n berechnen zu können, ist die Reihenentwicklung von $P(z)$ erforderlich, dazu muss aber $C(z)$ bekannt sein. Die erwartete Anzahl von Kunden im System $L = P'(1)$ kann aber auch mit weniger Information berechnet werden. Es gilt

$$L = \frac{\lambda \left(\mathbb{E}[X] + \mathbb{E}\left[X^2\right]\right)}{2 (\mu - \lambda \mathbb{E}[X])}.$$

Die Reihenentwicklung von $P(z)$ ist möglich für

- Konstante Gruppengröße, dh $C(z) = z^c$ mit c klein
- Geometrisch verteilte Gruppengröße, dh

  $$C(z) = \sum_{n=1}^{\infty} (1-\alpha) \alpha^{n-1} z^n = \frac{(1-\alpha) z}{1 - \alpha z}$$

  Es ergibt sich $p_0 = 1 - \frac{\lambda}{\mu(1-\alpha)}$ und

  $$p_n = \left(1 - \frac{\lambda}{\mu (1-\alpha)}\right) \frac{\lambda}{\mu} \left(\alpha + \frac{\lambda}{\mu}\right)^{n-1}.$$

Beispiel in [1] Seite 120.

3.1.2 Das $M/M^{[K]}/1$ Modell

Wir nehmen an, dass die Ankünfte einem Poisson-Prozess mit Rate λ genügen. Es werden jeweils höchstens K Kunden gleichzeitig bedient, wobei eine Bedienung eine exponential verteilte Zeitspanne mit Rate μ dauert, an deren Ende alle Kunden in der Bedienungsanlage das System verlassen. Solange weniger als K Kunden im System sind, treten neu eintreffende Kunden sofort in die Bedienungsanlage ein und werden bedient. Sind mehr als K Kunden im System, so werden K Kunden bedient und die anderen warten.

Die Gleichgewichtsverteilung muss folgendes Gleichungssystem erfüllen ($n \geq 1$):

$$\begin{aligned} -\lambda p_0 + \mu p_1 + \mu p_2 + \ldots + \mu p_{K-1} + \mu p_K &= 0 \\ \lambda p_{n-1} - (\lambda + \mu) p_n + \mu p_{n+K} &= 0. \end{aligned}$$

Dessen Lösungen lassen sich mit Hilfe der Nullstellen r_i des Operator-Polynoms, dh Lösungen der charakteristischen Gleichung

$$\mu r^{K+1} - (\lambda + \mu) r + \lambda = 0$$

schreiben. Wir erhalten also $p_n = \sum_{i=0}^{K} c_i r_i^n$. Weil die Summe der Wahrscheinlichkeiten gleich 1 sein muss, kommen nur Wurzeln $r_i \in]0,1[$ in Frage. Wie man sich leicht überlegt, hat das Operatorpolynom für $K\mu \leq \lambda$ in $]0,1[$ keine Nullstelle, dh es existiert keine Gleichgewichtsverteilung.
Für $K\mu > \lambda$ hat das Operatorpolynom in $]0,1[$ genau eine Nullstelle, die wir mit r_0 bezeichnen. Damit erhalten wir für $K\mu > \lambda$ als Gleichgewichtsverteilung

$$p_n = (1 - r_0) r_0^n \text{ für } n \geq 0.$$

Weil die Anzahl der Kunden im System also geometrisch verteilt ist, können wir die Formeln für L, L_q, W, W_q unmittelbar vom $M/M/1$ Modell übernehmen, wenn wir ρ durch r_0 ersetzen.
Analog kann auch ein Modell betrachtet werden, in dem eintreffende Kunden immer solange warten müssen, bis K gemeinsam abgefertigt werden können. Hier lautet das Gleichungssystem für die Gleichgewichtsverteilung

$$\begin{array}{ll} 0 = -\lambda p_0 + \mu p_K & \\ 0 = -\lambda p_n + \mu p_{n+K} + \lambda p_{n-1} & 1 \leq n < K \\ 0 = -(\lambda + \mu) p_n + \mu p_{n+K} + \lambda p_{n-1} & n \geq K \end{array} .$$

Benützt man die letzte Zeile, so erhält man wie oben für $n \geq K$

$$p_n = c r_0^n,$$

wobei c eine Konstante und r_0 die Wurzel der charakteristischen Gleichung

$$\mu r^{K+1} - (\lambda + \mu) r + \lambda = 0$$

in $]0,1[$. Verwenden wir nun zusätzlich die erste Zeile, so erhalten wir

$$p_K = \frac{\lambda}{\mu} p_0 = c r_0^K, \text{ also } c = \frac{\lambda p_0}{\mu r_0^K}.$$

Durch Verwendung der mittleren Zeile ergibt sich

$$\begin{aligned} \lambda (p_n - p_{n-1}) &= \mu p_{n+K} = \mu \frac{\lambda p_0}{\mu r_0^K} r_0^{n+K} = \lambda p_0 r_0^n \\ p_n - p_{n-1} &= p_0 r_0^n \end{aligned}$$

und daraus durch sukzessives Einsetzen für $1 \leq n < K$

$$\begin{aligned} p_1 &= p_0 \left(1 + r_0\right) \\ p_2 &= p_1 + p_0 r_0^2 = p_0 \left(1 + r_0 + r_0^2\right) \\ & \dots \end{aligned}$$

Wir erhalten also

$$p_n = p_0 \begin{cases} \frac{1-r_0^{n+1}}{1-r_0} & \text{falls} \quad 1 \leq n < K \\ \frac{\lambda}{\mu} r_0^{n-K} & \text{falls} \quad n \geq K \end{cases},$$

wobei sich

$$p_0 = \left(1 + \frac{K-1}{1-r_0} - r_0^2 \frac{1-r_0^{K-1}}{\left(1-r_0\right)^2} + \frac{\lambda}{\mu} \frac{1}{1-r_0}\right)^{-1}$$

durch Normierung ergibt.

3.2 Erlang-Modelle

Sei $k \in \mathbb{N}$ und seien $X_1, \dots, X_k$ unabhängige und exponential-verteilte Zufallsvariablen mit Parameter μ, so ist ihre Summe $X = X_1 + \ldots + X_k$ Erlang-verteilt mit den Parametern k und μ, also $\mathrm{E}\left(k, \mu\right)$. Es gilt

$$\Pr\left[X < x\right] = \int_0^x \frac{\left(\mu k\right)^k}{(k-1)!} t^{k-1} e^{-k\mu t} dt,$$

sowie $\mathbb{E}[X] = 1/\mu$ und $\mathrm{Var}[X] = 1/\left(k\mu^2\right)$. Für $k = 1$ ergibt sich die Exponentialverteilung, die die größte Varianz aufweist. Allgemein nennt man Verteilungen von Absorptionszeiten[1] endlicher Markov-Ketten *Phasentypverteilungen*. Beispiele dafür sind die *Erlangverteilung*, die eine Faltung von Exponentialverteilungen mit demselben Parameter ist, aber auch die Faltung von Exponentialverteilungen mit verschiedenen Parametern, oder auch die *Hyperexponentialverteilung*, die durch Mischung von Exponentialverteilungen entsteht.
Setzt sich nun etwa eine Bedienung aus mehreren Phasen zusammen, so kann ihre Dauer im einfachsten Fall durch eine Erlangverteilung angepasst werden. Aber auch wenn die Phasen nicht argumentiert werden können, so kann doch durch Einbeziehung von Phasentypverteilungen die Wirklichkeit genauer modelliert werden. Analog kann man auch die Zwischenankunftszeiten (wenn etwa die Varianz nicht so groß ist wie bei einer Exponentialverteilung) durch passende Erlangverteilungen oder Phasentypverteilungen modellieren.

[1] Ein Zustand i einer Markovkette heißt absorbierend, wenn $q_{ii} = 0$ gilt. Dann kann die Markov-Kette, die in i angekommen ist, von dort nicht mehr entkommen. Die zufällige Zeit, die eine Markov-Kette braucht, um in einen absorbierenden Zustand zu gelangen, nennt man ihre Absorptionszeit. Stellt man sich nun die Verweilzeiten der Markov-Kette in den verschiedenen Zuständen als Phasen vor, so sind diese Phasen unabhängig und exponentialverteilt mit verschiedenen Parametern und die Absorptionszeit ist eine zufällige Summe von solchen Phasen. Alle Verteilungen, die als Verteilung einer Absorptionszeit auftreten können, nennt man Phasentypverteilungen.

3.2.1 Das $M/E_k/1$ Modell

Es ist charakterisiert durch exponentialverteilte Zwischenankunftszeiten, einen unendlichen Warteraum, einen Server, Bedienung in k Phasen, beginnend mit Phase k und endend mit Phase 1. Wir bezeichnen mit $p_{n,i}$ die Wahrscheinlichkeit, dass n Kunden im System sind und die Bedienung in Phase i ist. Die Gleichgewichtsverteilung erfüllt folgendes Gleichungssystem:

$$\begin{array}{ll} 0 = -\lambda p_0 + k\mu p_{1,1} & \\ 0 = -(\lambda + k\mu)\, p_{1,k} + k\mu p_{2,1} + \lambda p_0 & \\ 0 = -(\lambda + k\mu)\, p_{1,i} + k\mu p_{1,i+1} & 1 \le i < k-1 \\ 0 = -(\lambda + k\mu)\, p_{n,k} + k\mu p_{n+1,1} + \lambda p_{n-1,k} & n \ge 2 \\ 0 = -(\lambda + k\mu)\, p_{n,i} + k\mu p_{n,i+1} + \lambda p_{n-1,i} & n \ge 2, 1 \le i \le k-1 \end{array}.$$

Die Transformation $(n,i) \to (n-1)\,k + i$, wobei man anstatt der Kunden im System die Phasen zählt, führt auf

$$\begin{array}{ll} 0 = -\lambda p_0 + k\mu p_1 & \\ 0 = -(\lambda + k\mu)\, p_{(n-1)k+i} + k\mu p_{(n-1)k+i+1} + \lambda p_{(n-2)k+i} & n \ge 1, 1 \le i \le k \end{array},$$

wobei p_j mit negativem j stets gleich 0 gesetzt wird. Bezeichnet man $m = (n-1)\,k + i$, so ergibt sich

$$\begin{array}{ll} 0 = -\lambda p_0 + k\mu p_1 & \\ 0 = -(\lambda + k\mu)\, p_m + k\mu p_{m+1} + \lambda p_{m-k} & m \ge 1 \end{array}.$$

Dieses Gleichungssystem ist uns aber bei der Ankunft in Gruppen schon begegnet, wenn wir dort konstante Gruppengröße k (Anzahl der Phasen eines Kunden) und Bedienungsrate $k\mu$ (für eine Phase) einsetzen.
Wie dort erhalten wir die Existenz der Gleichgewichtsverteilung für $k < k\mu/\lambda$, dh $\lambda < \mu$. Dann gilt $p_0 = 1 - \lambda/\mu$. Die erwartete Anzahl von Phasen im System und die mittlere Verweilzeit für eine Phase sind gleich

$$L = \frac{k+1}{2}\frac{\lambda}{\mu - \lambda} \text{ und } W = \frac{k+1}{2k\,(\mu - \lambda)}.$$

Dei Ergebnisse lassen sich aber auf Kunden nicht unmittelbar umrechnen, da jede Phase im System nicht mehr vorkommt, sobald sie fertig abgearbeitet ist, der Kunde aber noch im System verweilt, bis auch seine letzte Phase abgearbeitet ist.

3.2.2 Das $E_k/M/1$ Modell

Durch analoge Überlegungen kann dieses Modell mit einem Modell mit Bedienung in Gruppen in Verbindung gebracht werden. Ebenso können auch $E_j/E_k/1$ Modelle analysiert werden. Das Schöne daran ist, dass durch Einführung der Phasen die Markov-Eigenschaft gerettet werden kann.
Beispiel 3.6 Seite 139 [1]

3.3 Prioritäten

Bis jetzt haben wir nur FCFS-Modelle behandelt. In der Praxis treten aber häufig auch andere Prioritätsregeln auf. Man muss zwischen absoluter Priorität und relativer Priorität unterscheiden, wobei für einen Kunden mit absoluter Priorität sogar der Server geräumt wird, wenn dort gerade ein Kunde mit niederer Priorität bedient wird. Bei diesem ist dann zu entscheiden, ob bei seiner neuerlichen Bedienung die vorherige fortgesetzt wird oder ob die Bedienung von vorne beginnen muss. Kunden mit relativer Priorität werden nur nach Freiwerden eines Servers zuerst bedient.
Modelle mit Prioritäten sind naturgemäß sehr viel schwieriger zu behandeln als die FCFS-Modelle. Geht man von exponentialverteilten Bedienungszeiten aus, so spielt es wegen ihrer Nichtalterungseigenschaft keine Rolle, ob absolute oder relative Priorität herrscht. Sind die Bedienungszeiten der verschiedenen Prioritätstypen gleich, so ändert sich der stochastische Prozess, der die Anzahl der Kunden im System beschreibt, nicht. Ebenso gelten die Formeln von Little. Die Verteilung der Wartezeit allerdings ändert sich schon. Man kann zeigen, dass unter allen Strategien, die die Bedienungszeit nicht berücksichtigen, die FCFS-Strategie die stochastisch kürzeste Wartezeit aufweist. Außerdem ist ja gerade die Wartezeit für die verschiedenen Typen interessant. Diese geht aber aus diesem Prozess gar nicht hervor.
Die Bestimmung der stationären Verteilung ist schon für ein $M/M/1$ Modell mit zwei Typen von Kunden und relativer Priorität sehr kompliziert. Die Erwartungswerte der Anzahlen von Kunden des jeweiligen Typs können über zweidimensionale erzeugende Funktionen berechnet werden (siehe [1] S144 ff). Man kann aber auch die Methode der direkten Mittelwertberechnung verwenden (siehe [1] S151 ff).

Kapitel 4

Netzwerke

In der Praxis kommen fast ausschließlich Netzwerke von Bedienungsanlagen vor. Sie bestehen aus einzelnen Knoten, die aus verschieden vielen Servern bestehen und die von den Kunden in gewisser Reihenfolge durchlaufen werden müssen. Man unterscheidet *offene und geschlossene Netzwerke*, je nachdem, ob Kunden in das System eintreten und es wieder verlassen können oder immer dieselben Kunden im System vorhanden sind. Es kann sein, dass Kunden das Netzwerk in verschiedenen Routen durchlaufen, dass sie über verschiedene Knoten ins System eintreten und es über verschiedene (andere) wieder verlassen, es kann Wartemöglichkeiten geben oder nicht, Abhängigkeiten, verschiedene Verteilungen, $\ldots \rightarrow$ Arena-Praktikum

4.1 Jackson-Netzwerke

Ein Netzwerk mit $k+1$ Knoten heißt *Jackson*[1]*-Netzwerk*, wenn folgende 3 Bedingungen erfüllt sind:

- Die **Ankünfte**, die von außen im Knoten i ($i = 1, \ldots, k$) eintreffen, gehorchen einem Poisson-Prozess mit Rate γ_i.
- Die **Bedienungszeiten** sind unabhängig voneinander und in jedem Knoten i ($i = 1, \ldots, k$) ist die Bedienungszeit exponentialverteilt, wobei der Parameter μ_i von der Anzahl der Kunden im Knoten i abhängen darf.
- Die **Verzweigungswahrscheinlichkeiten** sind unabhängig vom Systemzustand, wobei r_{ij} die Wahrscheinlichkeit bezeichnet, dass ein Kunde nach Bedienung in i ($i = 1, \ldots, k$) nach j ($j = 0, 1, \ldots, k$) wechselt (routing probability). Wechselt ein Kunde in den Knoten 0, so heißt das, dass er das System verlassen hat. Von dort ist keine Rückkehr ins System möglich, dh $r_{0j} = 0$.

Ein Beispiel für ein geschlossenes Jackson-Netzwerk ($\gamma_i = 0$ und $r_{i0} = 0$) haben wir schon kennengelernt, das Modell für die Wartung von Maschinen, die in diesem Fall die Kunden sind. Hier gibt es 2 Knoten, nämlich einen, in dem die Maschinen repariert werden,

[1] Jackson, James Russell, Arbeiten aus 1957 und 1963.

und einen, in dem die Maschinen selbst arbeiten. Jede Maschine wechselt deterministisch zwischen diesen beiden Knoten. Der "Reparaturknoten" hat beschränkte Kapazität, die möglicherweise ein Warten erfordert. Ebenso hat der "Arbeitsknoten" beschränkte Kapazität, seine Warteschlange sind die Reservemaschinen. Die "Bedienungszeiten" in den einzelnen Knoten hängen von der Anzahl der dort anwesenden "Kunden" ab. Da es sich bei geschlossenen Jackson-Netzwerken um endliche Markov-Ketten handelt, ist eine Berechnung der Zustands- und Gleichgewichtsverteilung zumindest numerisch möglich, solange die Netzwerke nicht zu allzu groß werden.

4.1.1 Bedienungsanlagen in Serie

Häufig kommt es vor, dass ein Kunde mehrere Bedienungsstellen hintereinander besuchen muss, wobei immer dieselbe Reihenfolge einzuhalten ist. Wir wählen daher die Anzahl der Bedienungsanlagen $k \in \mathbb{N}$ fix und wählen ein Markov-Modell: Ein Poissonstrom von Kunden tritt in den Knoten 1 ein, der c_1 Server mit Bedienungsrate μ_1 umfasst, jeder dort fertig bediente Kunde wechselt in den Knoten 2 mit c_2 Servern mit Bedienungsrate μ_2, ... aus dem Knoten k wechselt ein fertig bedienter Kunde in den Knoten 0, dh er verläßt das System. Dabei handelt es sich um ein Jackson-Netzwerk mit den Parametern

$$\lambda_i = \begin{cases} \lambda & \text{falls } i = 1 \\ 0 & \text{sonst} \end{cases} \quad \text{und } r_{ij} = \begin{cases} 1 & \text{falls } j = (i+1)_{k+1}\,,\ i = 1, \ldots, k \\ 0 & \text{sonst} \end{cases} \quad .$$

Man muss nun unterscheiden, ob es bei den Knoten einen Warteraum gibt oder nicht. Angenommen, es gibt bei allen Knoten einen unbegrenzten Warteraum.
Dann ist der erste Knoten ein $M/M/c_1/\infty$ Modell. Der Output des ersten Knotens ist der Input für den zweiten. Der zweite Knoten ist ein $?/M/c_2/\infty$ Modell. Wobei noch zu überlegen bleibt, welcher Input vorliegt. Wir müssen also die Verteilung des Outputs eines $M/M/c_1/\infty$ Modells berechnen.
Sei $(X_t)_{t\in\mathbb{R}}$ die Markov-Kette für das $M/M/c_1/\infty$ Modell im Gleichgewicht. Nun gilt aber für alle i, j wie man sich durch Einsetzen der konkreten Werte leicht überlegt

$$p_i q_{ij} = p_j q_{ji},$$

dh die Markov-Kette ist reversibel. Deswegen sind

$$\Pr\left[X(t) = i, X(t+\Delta t) = j\right] = p_i p_{ij}(t, t+\Delta t) = p_i\left(q_{ij}\Delta t + o(\Delta t)\right) = p_i q_{ij}\Delta t + o(\Delta t)$$

und

$$\Pr\left[X(t) = i, X(t-\Delta t) = j\right] = p_j p_{ji}(t, t+\Delta t) = p_j\left(q_{ji}\Delta t + o(\Delta t)\right) = p_j q_{ji}\Delta t + o(\Delta t)$$

gleich und der Prozess ist derselbe, egal ob man die Zeit vorwärts oder rückwärts laufen läßt. Dreht man aber die Zeitachse um, so wird aus dem Inputstrom ein Outputstrom und umgekehrt. Der Output der urprünglichen Markov-Kette entspricht also dem Input der zeitumgekehrten und weil die beiden wegen der Reversibilität übereinstimmen, ist der Outputstrom im $M/M/c_1/\infty$ Modell im Gleichgewicht ebenso wie der Inputstrom ein Poissonstrom mit Rate λ.

Für ein Seriensystem von Knoten mit unbeschränktem Warteraum existiert also genau dann ein Gleichgewicht, wenn für jeden Knoten eines existiert. Im Gleichgewicht verhält sich dann jeder Knoten im Seriensystem wie eine einzelne Bedienungsanlage mit einem Poisson-Input mit Rate λ. Man kann zeigen, dass im Gleichgewicht

$$\Pr[X_1 = i_1, \ldots, X_k = i_k] = \Pr[X_1 = i_1] \cdot \ldots \cdot \Pr[X_k = i_k]$$

gilt (wobei X_i die Anzahl der Kunden beim Server i beschreibt), dh die gemeinsame Wahrscheinlichkeit mit dem Produkt der stationären Wahrscheinlichkeiten übereinstimmt (vgl [1] Aufgabe 4.4).

Die Analyse von Seriensystemen wird sehr viel komplizierter, wenn einige beschränkte Warteräume vorhanden sind und das System dadurch teilweise blockiert werden kann. Sind jedoch alle Warteräume beschränkt, so handelt es sich um endliche Markov-Ketten, deren Analyse zwar mühsam aber zumindest numerisch möglich ist.

4.1.2 Offene Jackson-Netzwerke

Diese sind eine Verallgemeinerung des vorangegangenen Modells. Sie zeichnen sich dadurch aus, dass nicht alle γ_i gleich 0 sind. Es können also Kunden von außen ins Netzwerk kommen. Der Zustand des Systems wird durch die Anzahl von Kunden in jedem Knoten beschrieben, dh durch einen k-dimensionalen Vektor $\bar{n} = (n_1, n_2, \ldots, n_k)$. Folgende Bezeichnungsweise hat sich bewährt

$$\begin{array}{ll} (n_1, \ldots, n_i, \ldots, n_j, \ldots, n_k) & \bar{n} \\ (n_1, n_2, \ldots, n_i + 1, \ldots, n_j, \ldots, n_k) & \bar{n}; i^+ \\ (n_1, n_2, \ldots, n_i, \ldots, n_j - 1, \ldots, n_k) & \bar{n}; j^- \\ (n_1, n_2, \ldots, n_i + 1, \ldots, n_j - 1, \ldots, n_k) & \bar{n}; i^+ j^- \end{array}.$$

Der Einfachheit halber nehmen wir an, das alle Knoten nur aus einem Server bestehen. Die Gleichgewichtsverteilung muss nun die stochastischen Bilanzgleichungen

$$\sum_{i=1}^{k} \gamma_i p_{\bar{n};i^-} + \sum_{\substack{i,j=1 \\ i \neq j}}^{k} \mu_i r_{ij} p_{\bar{n};i^+ j^-} + \sum_{i=1}^{k} \mu_i r_{i0} p_{\bar{n};i^+} = \sum_{i=1}^{k} \mu_i \left(1 - r_{ii}\right) p_{\bar{n}} + \sum_{i=1}^{k} \gamma_i p_{\bar{n}}$$

erfüllen (für $n_i \geq 0$ wenn man p mit negativem Index als 0 setzt). Jackson hat in seinen Arbeiten (1957, 1963) gezeigt, dass die Lösung dieses Systems Produktgestalt

$$p_{\bar{n}} = c \cdot c_1^{n_1} \cdot c_2^{n_2} \cdot \ldots \cdot c_k^{n_k}$$

hat. Wir werden die Lösung präsentieren und durch Einsetzen erkennen, dass die Bilanzgleichungen erfüllt sind.
Zuerst bestimmen wir durch heuristische Überlegung die Raten λ_i für den *Input in den Knoten i*. Es muss ein *Flussgleichgewicht* herrschen, daher wählen wir λ_i $(i = 1, \ldots, k)$ so, dass die *Bilanzgleichungen*

$$\lambda_i = \gamma_i + \sum_{j=1}^{k} \lambda_j r_{ji} \text{ bzw } \bar{\lambda} = \bar{\gamma} + \bar{\lambda} \cdot R$$

gelten. Ebenso muss der Input und der Output in bezug auf die Außenwelt übereinstimmen, das erfordert für die λ_i eine zusätzliche *Bilanzgleichung*

$$\sum_{i=1}^{k} \lambda_i r_{i0} = \sum_{i=1}^{k} \gamma_i.$$

Bezeichnen wir nun $\rho_i = \lambda_i/\mu_i$, so werden wir zeigen, dass

$$p_{\bar{n}} = (1-\rho_1)\cdot\rho_1^{n_1}\cdot(1-\rho_2)\cdot\rho_2^{n_2}\cdot\ldots\cdot(1-\rho_k)\cdot\rho_k^{n_k}$$

die Bilanzgleichungen erfüllt.
Setzen wir $p_{\bar{n}}$ in die Bilanzgleichungen ein, so ergibt sich

$$\begin{aligned}\sum_{i=1}^{k}\gamma_i p_{\bar{n}}\frac{1}{\rho_i} + \sum_{\substack{i,j=1\\i\neq j}}^{k}\mu_i r_{ij} p_{\bar{n}}\frac{\rho_i}{\rho_j} + \sum_{i=1}^{k}\mu_i r_{i0} p_{\bar{n}}\rho_i &\stackrel{?}{=} \sum_{i=1}^{k}\mu_i\left(1-r_{ii}\right)p_{\bar{n}} + \sum_{i=1}^{k}\gamma_i p_{\bar{n}}\\ \sum_{i=1}^{k}\gamma_i\frac{\mu_i}{\lambda_i} + \sum_{j=1}^{k}\frac{\mu_j}{\lambda_j}\sum_{\substack{i=1\\i\neq j}}^{k} r_{ij}\lambda_i + \sum_{i=1}^{k}\lambda_i r_{i0} &\stackrel{?}{=} \sum_{i=1}^{k}\mu_i\left(1-r_{ii}\right) + \sum_{i=1}^{k}\gamma_i\end{aligned}$$

Wegen des Flussgleichgewichts gilt

$$\sum_{\substack{i=1\\i\neq j}}^{k} r_{ij}\lambda_i = \lambda_j - \gamma_j - r_{jj}\lambda_j \text{ und } \sum_{i=1}^{k}\lambda_i r_{i0} = \sum_{i=1}^{k}\gamma_i,$$

was auf

$$\sum_{i=1}^{k}\gamma_i\frac{\mu_i}{\lambda_i} + \sum_{j=1}^{k}\frac{\mu_j}{\lambda_j}\left(\lambda_j - \gamma_j - r_{jj}\lambda_j\right) \stackrel{?}{=} \sum_{i=1}^{k}\mu_i\left(1-r_{ii}\right)$$

führt. Wählen wir nun überall als Summationsindex i, so folgt

$$\sum_{i=1}^{k}\left(\gamma_i\frac{\mu_i}{\lambda_i} + \checkmark\,\mu_i - \gamma_i\frac{\mu_i}{\lambda_i} - \checkmark\checkmark\, r_{ii}\mu_i\right) \stackrel{?}{=} \sum_{i=1}^{k}\left(\checkmark\,\mu_i - \checkmark\checkmark\,\mu_i r_{ii}\right),$$

was offensichtlich richtig ist.

Um die stationäre Wahrscheinlichkeit für ein offenes Jackson-Netzwerk zu bestimmen, müssen zuerst $\lambda_1,\ldots,\lambda_k$ aus dem Flussgleichgewicht bestimmt werden, dh

$$\bar{\lambda}\cdot(I-R) = \bar{\gamma}$$

muss gelöst werden. Eine eindeutige Lösung existiert, wenn R so beschaffen ist, dass jeder Kunde, der das Netzwerk betritt, dieses auch wieder verlassen kann. Gilt nun für alle i, dass $\rho_i = \lambda_i/\mu_i < 1$ gilt, so ergibt sich für $p_{\bar{n}}$ tatsächlich eine Wahrscheinlichkeit, und eine Gleichgewichtsverteilung existiert. Jeder Knoten hat dann als Gleichgewichtsverteilung gerade die Gleichgewichtsverteilung eines $M/M/1$ Modells, die Zwischenankunftszeiten sind aber im allgemeinen keineswegs exponentialverteilt. Natürlich gilt aber die Formel

von Little sowohl für jeden einzelnen Knoten als auch für das gesamte Netzwerk. Die Wartezeiten in den einzelnen Knoten sind im allgemeinen nicht unabhängig.
Offene Jackson-Netzwerke mit mehr als einem Server in den Knoten werden analog behandelt. Für offene Jackson-Netzwerke mit verschiedenen Kundentypen (verschiedene r_{ij}) geht man folgendermaßen vor: Man berechnet zuerst das Flussgleichgewicht für jeden einzelnen Kundentyp. Die Inputraten für alle Kunden ergeben sich dann als Summe der Raten der einzelnen Typen. Mit diesen Raten ist schließlich die Gleichgewichtsverteilung zu berechnen. Will man nun die Verteilung der Anzahl von Kunden vom Typ A bestimmen, so hat man zu beachten, dass jeder Kunde in jedem Knoten i mit Wahrscheinlichkeit λ_i^A/λ_i vom Typ A ist.

4.1.3 Geschlossene Jackson-Netzwerke

Gilt $\gamma_i = 0$ und $r_{i0} = 0$ für alle Knoten i, so handelt es sich um ein geschlossenes Jackson-Netzwerk mit N Kunden. Das ist das bereits bekannte Markov-Modell mit endlicher Population. Die soeben gewählte Vorgangsweise liefert uns aber nun auch für dieses Modell eine sehr elegante Methode zur Berechnung der Gleichgewichtsverteilung.
Folgendes Gleichungssystem ist zu lösen (Bilanzgleichungen)

$$\sum_{\substack{i,j=1\\ i\neq j}}^{k} \mu_i r_{ij} p_{\bar{n};i^+j^-} = \sum_{i=1}^{k} \mu_i \left(1 - r_{ii}\right) p_{\bar{n}}.$$

Analog zum vorigen Abschnitt hat die Lösung die Gestalt

$$p_{\bar{n}} = c \cdot \rho_1^{n_1} \cdot \rho_2^{n_2} \cdot \ldots \cdot \rho_k^{n_k},$$

wobei die $\rho_i = \lambda_i/\mu_i$ aus dem Flussgleichgewicht

$$\lambda_i = \sum_{j=1}^{k} \lambda_j r_{ji}$$

zu bestimmen sind. Man beachte, dass $\bar{\lambda}$ ein Gleichgewichtsmaß (stationäres Maß) für R ist. Wir nehmen an, dass die Matrix R irreduzibel ist. Damit existiert $\bar{\lambda}$ und ist bis auf einen konstanten Faktor eindeutig. Wir können daher ein λ_i beliebig, zB gleich μ_i, wählen. Zum Abschluss muss $p_{\bar{n}}$ noch normiert werden und es folgt

$$c = c\left(N\right) = \left(\sum_{n_1+n_2+\ldots+n_k=N} \rho_1^{n_1} \cdot \rho_2^{n_2} \cdot \ldots \cdot \rho_k^{n_k}\right)^{-1}.$$

Falls N groß ist, kann gerade das numerisch aufwendig werden. Ein möglicher Algorithmus ist der von Buzen (1973), vgl [1] Seite 186.

Mittelwertanalyse

Die Mittelwertanalyse ist ein Verfahren, mit dem Mittelwerte (L_i, W_i) für das System im Gleichgewicht direkt, dh ohne Berechnung der $p_{\bar{n}}$ und damit ohne Berechnung von $c\left(N\right)$,

berechnet werden. Ebenso werden damit die Randverteilungen $p_i(n, N) := \Pr_N [N_i = n]$ $(i = 1, \ldots, k,\ n = 1, \ldots, N)$ direkt, dh ohne Berechnung der $p_{\bar{n}}$, bestimmt, wobei $\Pr_N$ Wahrscheinlichkeiten für ein System mit N Kunden bezeichnet. Hat jeder Knoten nur einen Server, so ist die Mittelwertanalyse im Vergleich zum obigen Verfahren einschließlich Buzen-Algorithmus effizienter.
Wir beschränken uns auf den Fall, wo jeder Knoten einen Server aufweist, und begründen das Verfahren nur heuristisch. Genaueres findet sich in [1].
Die Mittelwertanalyse beruht auf 3 Prinzipien:

1. In einem geschlossenen Jackson-Netzwerk mit N Kunden ist die Warteschlange in irgendeinem Knoten, die von einem dort ankommenden Kunden beobachtet wird, dieselbe, wie die Warteschlange in diesem Knoten in einem geschlossenen Jackson-Netzwerk mit $N - 1$ Kunden, dh
$$W_i(N) = \frac{L_i(N-1) + 1}{\mu_i}$$
2. Die Formel von Little gilt im gesamten Netzwerk, dh
$$L_i(N) = \lambda_i(N)\, W_i(N)\,.$$
3. Wegen der Reversibilität der Markov-Kette (Beweis [1] Seite 197) gilt für $n = 1, \ldots, N$ und $N \geq 1$
$$\mu_i p_i(n, N) = \lambda_i(n)\, p_i(n-1, N-1)$$
und $p_i(0, 0) = 1$.

Zur Berechnung der $\lambda_i(N)$ geht man nun folgendermaßen vor: Man bestimmt ein stationäres Maß $\bar{\nu}$ von R mit oBdA $\nu_1 = 1$. Aus der Theorie der Markov-Ketten folgt, dass dann ν_i die erwartete Anzahl von Aufenthalten in i pro Aufenthalt im Knoten 1 ist. Die erwartete Rückkehrzeit eines Kunden vom Knoten 1 in den Knoten 1 ist damit gleich $\sum_{i=1}^{k} \nu_i W_i(N)$. Diese interpretieren wir als mittlere Systemzeit für den Knoten 1. Weil immer N Kunden im System sind, gilt wegen der Formel von Little
$$N = \lambda_1(N) \sum_{i=1}^{k} \nu_i W_i(N)\,.$$
Daraus erhalten wir
$$\lambda_1(N) = \frac{N}{\sum_{i=1}^{k} \nu_i W_i(N)} \text{ und } \lambda_i(N) = \nu_i \lambda_1(N)\,.$$
Der Algorithmus Mittelwertanalyse liegt nun auf der Hand:

- Berechne $\bar{\nu}$ aus $\bar{\nu} = \bar{\nu} R$ mit $\nu_1 = 1$.
- Setze $L_i(0) = 0$ und $p_i(0, 0) = 1$, $i = 1, \ldots, k$.

- Für $n = 1, \ldots, N$ berechne

$$\begin{aligned} W_i(n) &= \tfrac{L_i(n-1)+1}{\mu_i}, \\ \lambda_1(n) &= \tfrac{n}{\sum_{i=1}^{k} \nu_i W_i(n)}, \\ \lambda_i(n) &= \nu_i \lambda_1(n), \\ L_i(n) &= \lambda_i(n) W_i(n). \end{aligned}$$

  - Für $m = 1, \ldots, n$ berechne

$$\begin{aligned} p_i(m,n) &= \tfrac{\lambda_i(m)}{\mu_i} p_i(m-1, n-1) \\ p_i(0,n) &= 1 - \sum_{m=1}^{n} p_i(m,n). \end{aligned}$$

Aus den $\lambda_i(N)$ könnten mit Hilfe der Produktformel die $p_{\bar{n}}$ berechnet werden, und auch daraus könnten die $p_i(n, N)$ bestimmt werden, was aber im Vergleich zur iterativen Methode einen höheren Rechenaufwand erfordern würde.

Beispiel: Zyklische Anordnung von Knoten ([1] Seite 199).

4.2 Verallgemeinerungen

Jackson selbst hat schon gewisse Verallgemeinerungen in den Modellen zugelassen: Er hat etwa als Ankunftsprozess einen Poisson-Prozess zugelassen, dessen Intensität von der Gesamtanzahl der Kunden im System abhängt, und er hat Bedienungsraten zugelassen, die von der Anzahl der Kunden im Knoten abhängen. Auch hier hat die Gleichgewichtsverteilung die Produktform. Die Normierungskonstante ist allerdings ähnlich aufwendig zu bestimmen, wie im Fall der geschlossenen Jackson-Netzwerke.

Auch die r_{ij} können als zustandsabhängig angenommen werden. Es können zwischen den Knoten zufällige Transferzeiten zugelassen werden. Ebenso können auch mehrere Typen von Kunden zugelassen werden.

Die Gleichungssysteme für die Gleichgewichtsverteilung bei Netzwerken mit endlich vielen Zuständen werden sehr schnell riesig groß, hat man etwa 50 Kunden und 10 Knoten, so sind theoretisch

$$\binom{50+10-1}{50} = 1.257 \cdot 10^{10}$$

verschiedene Zustände möglich.

Kapitel 5

Modelle mit allgemeiner Ankunfts- oder Bedienungszeit

Durch die Abkehr von der Exponentialverteilung geht die Markov-Eigenschaft verloren. Diese stochastischen Prozesse sind also keine Markov-Prozesse. Betrachtet man allerdings günstige Zeitpunkte, so kann man eingebettete Markov-Ketten finden.

5.1 M/G/1 Modelle

Wir nehmen an, der Ankunftsprozess ist ein Poisson-Prozess mit Rate λ und die Bedienungszeiten $T_s^{(k)}$ aller Kunden k sind unabhängig voneinander und vom Ankunftsprozess und haben die Verteilungsfunktion F_{T_s}, dh

$$\Pr\left[T_s^{(k)} < t\right] = F_{T_s}(t).$$

Der Erwartungswert von $T_s^{(k)}$ sei $1/\mu$ und die Varianz sei σ_S^2. Bezeichnet A_k die Anzahl der Kunden, die während der Bedienung des k-ten Kunden eintreffen, so gilt

$$\begin{aligned}\Pr[A_k = a] &= \int \Pr\left[A_k = a \mid T_s^{(k)} = t\right] dF_{T_s}(t) \\ &= \int e^{-\lambda t} \frac{(\lambda t)^a}{a!} dF_{T_s}(t).\end{aligned}$$

Ist $(\tau_k)_{k\in\mathbb{N}_0}$ die Folge derjenigen Zeitpunkte, zu denen Kunden fertig bedient sind, so definieren wir N_k als die Anzahl der Kunden, die der k-te Kunde im System zurückläßt. Weil N_k durch

$$N_{k+1} = \begin{cases} A_{k+1} & \text{falls } N_k = 0 \\ N_k - 1 + A_{k+1} & \text{falls } N_k \neq 0 \end{cases}$$

beschrieben wird, hängt bei vorgegebenem $N_k = n$ die nächste Zustand N_{k+1} lediglich von A_{k+1} ab und ist damit unabhängig von der Vergangenheit $N_1, \ldots, N_{k-1}$. $(N_k)_{k\in\mathbb{N}_0}$ ist daher eine Markov-Kette.

5.1.1 Erwartungswerte, Pollaczek-Khintchine Formel

Für den Fall, dass der Prozess im Gleichgewicht ist, bezeichne $L^{(D)}$ die erwartete Anzahl von Kunden im System, wenn ein Kunde gerade fertig bedient ist und das System verläßt. Diese wollen wir nun berechnen.
Wir berechnen zuerst

$$\begin{aligned}\mathbb{E}\left[A_{k+1}\right] &= \sum_{a=1}^{\infty} a \int e^{-\lambda t}\frac{(\lambda t)^a}{a!}dF_{T_s}(t) \\ &= \int (\lambda t)\sum_{a=1}^{\infty} e^{-\lambda t}\frac{(\lambda t)^{a-1}}{(a-1)!}dF_{T_s}(t) = \lambda\mathbb{E}\left[T_s\right] = \frac{\lambda}{\mu}\end{aligned}$$

und

$$\begin{aligned}\mathbb{E}\left[A_{k+1}^2\right] &= \sum_{a=1}^{\infty} a^2 \int e^{-\lambda t}\frac{(\lambda t)^a}{a!}dF_{T_s}(t) \\ &= \int (\lambda t)\sum_{a=1}^{\infty} a\ e^{-\lambda t}\frac{(\lambda t)^{a-1}}{(a-1)!}dF_{T_s}(t) \\ &= \int (\lambda t)\left(\sum_{a=1}^{\infty}(a-1)\,e^{-\lambda t}\frac{(\lambda t)^{a-1}}{(a-1)!} + \sum_{a=1}^{\infty} e^{-\lambda t}\frac{(\lambda t)^{a-1}}{(a-1)!}\right)dF_{T_s}(t) \\ &= \int (\lambda t)\left((\lambda t)\sum_{a=2}^{\infty} e^{-\lambda t}\frac{(\lambda t)^{a-2}}{(a-2)!}+1\right)dF_{T_s}(t) \\ &= \lambda^2\mathbb{E}\left[T_s^2\right] + \lambda\mathbb{E}\left[T_s\right] \\ &= \lambda^2\left(\sigma_S^2 + \mathbb{E}\left[T_s\right]^2\right) + \lambda\mathbb{E}\left[T_s\right] \\ &= \lambda^2\sigma_S^2 + \left(\frac{\lambda}{\mu}\right)^2 + \frac{\lambda}{\mu}.\end{aligned}$$

Es gilt

$$N_{k+1} = N_k - \mathbf{1}_{\{N_k>0\}} + A_{k+1},$$

was für die Erwartungswerte auf

$$L^{(D)} = L^{(D)} - \Pr\left[N_k > 0\right] + \mathbb{E}\left[A_{k+1}\right]$$

führt. Daraus folgt nun

$$\Pr\left[N_k > 0\right] = \mathbb{E}\left[A_{k+1}\right] = \frac{\lambda}{\mu}.$$

Quadriert man die Gleichung für die Zufallsvariablen, bildet dann wieder die Erwartungswerte und setzt die obige Gleichheiten ein, so erhält man

$$\begin{aligned}N_{k+1}^2 &= N_k^2 + \mathbf{1}_{\{N_k>0\}} + A_{k+1}^2 - 2N_k - 2A_{k+1}\mathbf{1}_{\{N_k>0\}} + 2N_kA_{k+1} \\ 0 &= \frac{\lambda}{\mu} + \left(\lambda^2\sigma_S^2 + \left(\frac{\lambda}{\mu}\right)^2 + \frac{\lambda}{\mu}\right) - 2L^{(D)} - 2\frac{\lambda}{\mu}\frac{\lambda}{\mu} + 2L^{(D)}\frac{\lambda}{\mu},\end{aligned}$$

was auf

$$L^{(D)} = \frac{\lambda}{\mu} + \frac{\lambda^2\sigma_S^2 + \left(\frac{\lambda}{\mu}\right)^2}{2\left(1-\frac{\lambda}{\mu}\right)}$$

führt. Wir werden später zeigen, dass ebenso

$$L = \frac{\lambda}{\mu} + \frac{\left(\frac{\lambda}{\mu}\right)^2 + \lambda^2\sigma_S^2}{2\left(1-\frac{\lambda}{\mu}\right)}$$

gilt. Das ist nun die gesuchte Pollaczek-Khintchine Formel. Daraus wird nun mit der Formel von Little W hergeleitet, daraus wie üblich $W_q = W - 1/\mu$ und $L_q = L - \lambda/\mu$.
Sind die Bedienungszeiten exponentialverteilt, dh $\sigma_S^2 = 1/\mu^2$, so ergibt sich die entsprechende $M/M/1$-Formel.
Beispiel 5.1 in [1] Seite 213.

5.1.2 Gleichgewichtsverteilung

Für die Markov-Kette $(N_k)_{k\in\mathbb{N}_0}$gilt wegen

$$N_{k+1} = \begin{cases} A_{k+1} & \text{falls} \quad N_k = 0 \\ N_k - 1 + A_{k+1} & \text{falls} \quad N_k \neq 0 \end{cases}$$

für $n \geq 0, m > 0$

$$\Pr\left[N_{k+1} = m + n \mid N_k = m\right] = \Pr\left[A_{k+1} = n\right] =: a_n.$$

Sie hat daher die Übergangsmatrix

$$\mathbf{P} = \begin{pmatrix} a_0 & a_1 & a_2 & \dots & \dots \\ a_0 & a_1 & a_2 & \dots & \dots \\ 0 & a_0 & a_1 & a_2 & \dots \\ 0 & 0 & a_0 & a_1 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix},$$

deren stationäre Verteilungen sich aus

$$\pi\mathbf{P} = \pi$$

ergeben. Zu lösen ist also das Gleichungssystem für $n \geq 0$

$$\pi_n = \pi_0 a_n + \sum_{k=1}^{n+1} a_{n-k+1}\pi_k$$

und $\sum_{n=0}^{\infty} \pi_n = 1$.
Mittels der erzeugenden Funktionen $\Pi(z)$ und $A(z)$ für die (π_n) und (a_n) erhalten wir

$$\Pi(z) = \pi_0 \frac{(1-z)\,A(z)}{A(z) - z},$$

woraus unter der Voraussetzung $1 - \frac{\lambda}{\mu} > 0$ wegen $\Pi(1) = A(1) = 1$ und $A'(1) = \lambda/\mu$ zuerst $\pi_0 = 1 - \frac{\lambda}{\mu}$ und dann

$$\Pi(z) = \left(1 - \frac{\lambda}{\mu}\right) \frac{(1-z)\, A(z)}{A(z) - z}$$

folgt. Die erwartete Anzahl von Kunden im System kann als $\Pi'(1)$ auch hier berechnet werden.
Beispiel mit konstanter Bedienungszeit: [1] Seite 219, Beispiel 5.3

5.1.3 Beweis von $\pi_n = p_n$

Wir zeigen, dass die stationären Wahrscheinlichkeiten π_n zu den Zeitpunkten τ_k mit den über die ganze Zeit gesehen stationären Wahrscheinlichkeiten p_n übereinstimmen. Für $t \geq 0$ und $n \in \mathbb{N}_0$ bezeichnen wir mit $N_n^+(t)$ die Anzahl der Übergänge von n auf $n+1$ im Zeitintervall $]0,t[$ und mit $N_n^-(t)$ die Anzahl der Übergänge von $n+1$ auf n im Zeitintervall $]0,t[$. Da immer nur Sprünge der Höhe 1 auftreten, gilt

$$\left|N_n^+(t) - N_n^-(t)\right| \leq 1.$$

Bezeichnen $N^+(t) := \sum_{n=0}^{\infty} N_n^+(t)$ und $N^-(t) := \sum_{n=0}^{\infty} N_n^-(t)$ die Summe der Ankünfte bzw Abgänge im Intervall $]0,t[$, so gilt

$$N^-(t) = N^+(t) + X(0) - X(t),$$

wobei $X(t)$ die Anzahl der Kunden zum Zeitpunkt t bezeichnet. Existiert nun eine Grenzverteilung π für $X(\tau_k)$, so gilt fast sicher

$$\begin{aligned}
\pi_n &= \lim_{t\to\infty} \frac{N_n^-(t)}{N^-(t)} \\
&= \lim_{t\to\infty} \frac{N_n^+(t) - N_n^+(t) + N_n^-(t)}{N^+(t) + X(0) - X(t)} \\
&= \lim_{t\to\infty} \frac{N_n^+(t)}{N^+(t) + X(0) - X(t)} - \lim_{t\to\infty} \frac{N_n^+(t) - N_n^-(t)}{N^+(t) + X(0) - X(t)} \\
&= \lim_{t\to\infty} \frac{N_n^+(t)}{N^+(t)}.
\end{aligned}$$

Die letzte Gleichheit folgt, weil $N^+(t)$ fast sicher gegen ∞ geht und sowohl $X(0)$ als auch $X(t)$ wegen der Existenz der Grenzverteilung mit Sicherheit endlich sind. Weil aber der Ankunftsprozess ein Poisson-Prozess ist, findet ein nach langer Zeit ankommender Kunde mit Wahrscheinlichkeit p_n gerade n Kunden im System vor, sodass fast sicher

$$\lim_{t\to\infty} \frac{N_n^+(t)}{N^+(t)} = p_n$$

gilt.

5.1.4 Existenz der Grenzverteilung

Unter der Voraussetzung $1 - \frac{\lambda}{\mu} > 0$ ist die eingebettete Markov-Kette irreduzibel und hat die stationäre Verteilung π, wegen Satz **??** ist sie damit positiv rekurrent und π ist die einzige stationäre Verteilung. Die eingebettete Markov-Kette ist darüber hinaus auch aperiodisch, es gilt sogar $p_{nn} > 0$ für alle n, wegen Satz **??** ist daher die stationäre Verteilung auch Grenzverteilung.
Gilt auch die Umkehrung? Folgt aus der Existenz der Grenzverteilung, dass $1 - \frac{\lambda}{\mu} > 0$ gilt? Die Grenzverteilung ist auch stationäre Verteilung und ihre erzeugende Funktion ist $\Pi(z)$. Damit aber $\Pi(1) = 1$ gilt, muss

$$\begin{aligned} 1 &= \Pi(1) = \pi_0 \left.\frac{(1-z)\,A(z)}{A(z) - z}\right|_{z=1} \\ &= \pi_0 \frac{-A(1) + (1-1)\,A'(1)}{A'(1) - 1} \\ &= \pi_0 \frac{-1}{\frac{\lambda}{\mu} - 1} \end{aligned}$$

gelten, was wegen $\pi_0 > 0$ auch $1 - \frac{\lambda}{\mu} > 0$ zur Folge hat.
Es existiert also genau dann eine Grenzverteilung, wenn $1 - \frac{\lambda}{\mu} > 0$ gilt.

5.1.5 Wartezeiten

Wegen der Formel von Little gilt für den Erwartungswert der Systemzeit $W = L/\lambda$. Formeln für die Momente der Systemzeit $\mathrm{E}T^k$ können durch folgende Überlegungen gewonnen werden.
Angenommen, das System ist im Gleichgewicht und zum Zeitpunkt 0 trifft ein Kunde ein. Bezeichne T die Systemzeit dieses Kunden. Weil der Systemzustand zum Zeitpunkt, in dem der Kunde das System verläßt, gerade durch p bzw π beschrieben wird und die Anzahl der während eines fixen Intervalls eintreffenden Kunden Poisson-verteilt ist, gilt für $n \geq 0$

$$\begin{aligned} p_n &= \pi_n = \Pr[X(T) = n] \\ &= \int \Pr[X(T) = n \mid T = t]\, dF_T(t) \\ &= \int \frac{1}{n!} (\lambda t)^n e^{-\lambda t} dF_T(t)\,. \end{aligned}$$

Einführung der erzeugenden Funktionen ergibt

$$\begin{aligned} P(z) &= \sum_{n=0}^{\infty} p_n z^n = \int \sum_{n=0}^{\infty} \frac{1}{n!} (\lambda t z)^n e^{-\lambda t} dF_T(t) \\ &= \int e^{-\lambda t(1-z)} dF_T(t) = F_T^*(\lambda(1-z))\,, \end{aligned}$$

wobei F_T^* die Laplace-Stielties-Transformierte[1] von F_T bezeichnet. Mittels Kettenregel gilt daher

$$\begin{aligned}\frac{d^k P(z)}{dz^k} &= \sum_{n=k}^{\infty} \frac{n!}{k!} p_n z^{n-k} = \int (\lambda t)^k e^{-\lambda t(1-z)} dF_T(t) \\ &= \lambda^k \mathrm{E}\left[T^k e^{-\lambda T(1-z)}\right].\end{aligned}$$

Setzen wir nun $z = 1$ ein, so ergibt sich eine Verallgemeinerung der Formel von Little

$$\begin{aligned}\mathrm{E}\left[\frac{X!}{k!}\mathbf{1}_{\{X\geq k\}}\right] &= \lambda^k \mathrm{E}\left[T^k\right], \text{ oder kurz} \\ L_{(k)} &= \lambda^k W_k\end{aligned}$$

wobei links das k-te faktorielle Moment der Kundenanzahl $L_{(k)}$ steht und rechts λ^k-mal das k-te Moment der Systemzeit W_k. Für die Laplace-Stielties-Transformierten der Bedienungs- und der System- bzw der Wartezeit gilt

$$\begin{aligned}F_T^*(s) &= \frac{\left(1-\frac{\lambda}{\mu}\right) s F_{T_s}^*(s)}{s - \lambda\left(1 - F_{T_s}^*(s)\right)} \\ F_{T_q}^*(s) &= \frac{\left(1-\frac{\lambda}{\mu}\right) s}{s - \lambda\left(1 - F_{T_s}^*(s)\right)}.\end{aligned}$$

Beweis Es gilt

$$F_T^*(\lambda(1-z)) = P(z) = \Pi(z) = \left(1 - \frac{\lambda}{\mu}\right)\frac{(1-z)A(z)}{A(z)-z}$$

und

$$\begin{aligned}A(z) &= \sum_{n=0}^{\infty} \Pr[A_k = n] z^n \\ &= \sum_{n=0}^{\infty} \int \Pr\left[A_k = n \mid T_s^{(k)} = t\right] dF_{T_s}(t) z^n \\ &= \int \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t z)^n}{n!} dF_{T_s}(t) \\ &= \int e^{-\lambda t(1-z)} dF_{T_s}(t) \\ &= F_{T_s}^*(\lambda(1-z)).\end{aligned}$$

Daraus folgt die erste Behauptung, wenn man $s = \lambda(1-z)$ setzt. Beachtet man, dass wegen $T = T_q + T_s$ für die Laplace-Stielties-Transformierten $F_T^*(s) = F_{T_q}^*(s) F_{T_s}^*(s)$ gilt, so folgt die zweite Behauptung. □

[1] $F^*(s) = \int e^{-st} dF(t)$

Durch Ableiten der Beziehung

$$F_{T_q}^*(s)\left(s-\lambda\left(1-F_{T_s}^*(s)\right)\right)=\left(1-\frac{\lambda}{\mu}\right)s$$

lassen sich die Momente der Wartezeit aus Momenten der Bedienungszeit berechnen. Es gilt

$$\mathrm{E}\left[T_q^k\right]=\frac{\lambda^2}{\mu-\lambda}\sum_{j=1}^{k}\binom{k}{j}\mathrm{E}\left[T_q^{k-j}\right]\frac{\mathrm{E}\left[T_s^{j+1}\right]}{j+1}.$$

5.1.6 Endliche M/G/1 Modelle

Sei $(X(t))_{t\geq 0}$ die Anzahl von Kunden in einem M/G/1/K Modell.
Hier gilt die Pollaczek-Khintchine Formel nicht, weil die erwartete Anzahl von Kunden während der Bedienung des k-ten Kunden von der Systemgröße abhängt.
Die stationäre Verteilung π für die eingebettete Markov-Kette läßt sich analog zum unendlichen M/M/1 Modell berechnen. Sie ist auf die Zustände $\{0,1,\ldots,K-1\}$ beschränkt und kann daher nicht mit der stationären Verteilung für den ursprünglichen Prozess übereinstimmen, weil bei dieser auch der Zustand K eine positive Wahrscheinlichkeit hat. Wohl aber kann man sich analog zum unendlichen Modell überlegen, dass, falls $(X(t))_{t\geq 0}$ stationär ist,

$$\pi_i=\Pr\left[X(t)=i\mid X(t)<K\right]$$

gilt. Angenommen, wir würden $p_K=\Pr\left[X(t)=K\right]$ kennen, so könnten wir auch für $i<K$

$$\begin{aligned}p_i &= \Pr\left[X(t)=i\right]\\ &= \Pr\left[X(t)=i\wedge X(t)<K\right]\\ &= \Pr\left[X(t)=i\mid X(t)<K\right]\Pr\left[X(t)<K\right]\\ &= \pi_i\left(1-p_K\right)\end{aligned}$$

berechnen. Die stationäre Wahrscheinlichkeit p_K können wir nun berechnen, indem wir den effektiven Eingangsstrom mit dem effektiven Ausgangsstrom vergleichen und beachten, dass diese übereinstimmen müssen. So erhalten wir

$$\lambda\left(1-p_K\right)=\mu\left(1-p_0\right).$$

Mit $p_0=\pi_0\left(1-p_K\right)$ ergibt sich

$$\begin{aligned}\lambda\left(1-p_K\right) &= \mu\left(1-\pi_0\left(1-p_K\right)\right)\\ \left(\frac{\lambda}{\mu}+\pi_0\right)\left(1-p_K\right) &= 1\\ 1-p_K &= \frac{1}{\frac{\lambda}{\mu}+\pi_0}\end{aligned}$$

und daraus schließlich

$$p_i=\begin{cases}\frac{\pi_i}{\frac{\lambda}{\mu}+\pi_0} & \text{falls } 0\leq i<K\\ 1-\frac{1}{\frac{\lambda}{\mu}+\pi_0} & \text{falls } i=K\end{cases}.$$

5.2 M/G/c Modelle

Für M/G/c Modelle im allgemeinen existiert diese eingebettete Markov-Kette nicht, weil für die Beschreibung des Systemzustandes die Anzahl der Kunden im System nicht ausreicht. Man müßte zusätzlich angeben, in welchem Stadium der Bedienung die Kunden in den verschiedenen Servern sind. Es gibt jedoch einzelne spezielle Modelle, die dennoch eingebettete Markov-Ketten besitzen, etwa M/D/c. Wir wollen einige allgemeine Resultate für M/G/c/∞, M/G/∞ und M/G/c/c Modelle herleiten.

5.2.1 M/G/c/∞ Modelle

Der stochastische Prozess $(X_t)_{t\geq 0}$ sei im Gleichgewicht. Sei T_q die Wartezeit eines Kunden in der Warteschlange, $W_{q\,(k)}$ das k-te Moment, $X_q^{(D)}$ die Anzahl von Kunden in der Warteschlange zu einem Zeitpunkt, in dem ein Kunde das System verläßt, und $L_q^{(D)}{}_{(k)}$ das k-te faktorielle Moment davon, so gilt eine verallgemeinerte Formel von Little

$$L_q^{(D)}{}_{(k)} = \lambda^k W_{q\ k}.$$

Beweis Die Rechnung erfolgt wie in Abschnitt 5.1.5, indem man X durch X_q und T durch T_q ersetzt. □

5.2.2 M/G/∞ Modell

Sei $X(t)$ die Anzahl der Kunden im System zur Zeit t mit $X(0) = 0$. Es gilt

$$X(t) = N^+(t) - N^-(t),$$

wobei $N^+(t)$ die Anzahl der Ankünfte bis t und $N^-(t)$ die Anzahl der Abgänge bis t bezeichnet. Als Vorüberlegung berechnen wir

$$q(t) := \Pr[\text{in } [0,t] \text{ ankommender Kunde wird nach } t \text{ fertig}]$$

indem wir beachten, dass die Ankunftszeit eines in $[0,t]$ ankommenden Kunden im Intervall $[0,t]$ gleichverteilt ist. Damit gilt

$$\begin{aligned} q(t) &= \int \Pr[T_s > t-\tau \mid \text{Ankunft} \ldots \text{zur Zeit } \tau]\, d\Pr[\text{Ankunft} \ldots \text{zur Zeit } \tau] \\ &= \int_0^t (1 - F_{T_s}(t-\tau))\frac{1}{t} d\tau \\ &= \frac{1}{t}\int_0^t (1 - F_{T_s}(\tau))\, d\tau. \end{aligned}$$

Für die Verteilung von $X(t)$ gilt damit

$$\begin{aligned}
\Pr[X(t)=n] &= \sum_{i=n}^{\infty} \Pr\left[X(t)=n \mid N^{+}(t)=i\right] \Pr\left[N^{+}(t)=i\right] \\
&= \sum_{i=n}^{\infty} \Pr\left[N^{-}(t)=i-n \mid N^{+}(t)=i\right] \frac{(\lambda t)^{i}}{i!} e^{-\lambda t} \\
&= \sum_{i=n}^{\infty} \binom{i}{n} q(t)^{n} (1-q(t))^{i-n} \frac{(\lambda t)^{i}}{i!} e^{-\lambda t} \\
&= \frac{e^{-\lambda t} (\lambda t q(t))^{n}}{n!} \sum_{i=n}^{\infty} (1-q(t))^{i-n} \frac{(\lambda t)^{i-n}}{(i-n)!} \\
&= \frac{e^{-\lambda t} (\lambda t q(t))^{n}}{n!} e^{\lambda t(1-q(t))} \\
&= \frac{(\lambda t\, q(t))^{n}}{n!} e^{-\lambda t\, q(t)}.
\end{aligned}$$

Zur Zeit t ist die Anzahl der Kunden im System also Poisson-verteilt mit Parameter $\lambda t\, q(t)$. Die stationäre Verteilung ergibt sich durch Grenzübergang $t \to \infty$ und ist daher eine Poissonverteilung mit Parameter

$$\begin{aligned}
\lim_{t\to\infty} \lambda t\, q(t) &= \lambda \int_{0}^{\infty} \left(1-F_{T_s}(\tau)\right) d\tau \\
&= \lambda \int_{0}^{\infty} \tau\, f_{T_s}(\tau)\, d\tau = \frac{\lambda}{\mu}
\end{aligned}$$

wie beim $M/M/\infty$ Modell. Die Verteilung des Outputstroms $N^{-}(t)$ überlegt man sich analog und erhält

$$\Pr\left[N^{-}(t)=n\right] = \frac{(\lambda t(1-q(t)))^{n}}{n!} e^{-\lambda t(1-q(t))}.$$

5.2.3 M/G/c/c Modell

Wir haben schon früher behauptet, dass im Fall M/G/c/c wie im Fall M/M/c/c für die stationäre Verteilung die Erlangsche Verlustformel

$$p_n = \frac{\lambda^n}{\mu^n n!} \left(\sum_{k=0}^{c} \frac{\lambda^k}{\mu^k k!} \right)^{-1}$$

gilt, dass also die Verteilung der Bedienungszeit für die stationäre Verteilung keine Rolle spielt.

Beweisidee:

Der stochastische Prozess $(X(t), T_1(t), \ldots, T_c(t))_{t\geq 0}$, wobei die $T_i(t)$ die bis zur Zeit t bereits vergangenen der Größe nach sortierten Bedienungszeiten bezeichnen, ist eine reversible Markov-Kette. Man beachte, dass für $X(t)=n$ gilt $T_1(t) = \ldots = T_{c-n}(t) = 0$.

Zuerst zeigt man, dass die stationäre Verteilung dieser Markov-Kette Produktgestalt hat, dh

$$\begin{aligned} p_n(u_1,\ldots,u_c) \quad &:= \lim_{t\to\infty} \Pr\left[X(t)=n, T_1(t)=u_1,\ldots,T_c(t)=u_c\right] \\ &= \frac{p_0(0,\ldots,0)}{n!} a_n \left(1-F_{T_s}(u_1)\right)\ldots\left(1-F_{T_s}(u_c)\right), \end{aligned}$$

wobei $u_1 = \ldots = u_{c-n} = 0$ gilt, und a_n zur Wahrscheinlichkeit proportional ist, dass n Server arbeiten, die wiederum zu λ^n proportional ist. Der Term $n!$ steht für die $n!$ verschiedenen Anordnungen der Server, die den sortierten Bedienungszeiten entsprechen können. Daraus folgt wegen

$$\int \left(1-F_{T_s}(u_i)\right) du_i = 1/\mu$$

die Beziehung

$$p_n = \int \ldots \int p_n(u_1,\ldots,u_c)\, du_1 \ldots du_n = p_0 \frac{\lambda^n}{\mu^n n!}$$

□

Weil nun das M/G/c/c Modell dieselbe stationäre Verteilung hat, wie das M/M/c/c Modell gilt auch im stationären M/G/c/c System

$$\lambda p_n = (n+1)\,\mu p_{n+1}.$$

Man kann zeigen, dass man im obigen Gleichgewicht statt λ auch eine zustandsabhängige Rate λ_n einsetzen kann.

5.2.4 M/D/c Modell

Siehe [1] Seite 286.

5.3 G/M/1 Modelle

Kunden kommen einzeln und die Zeitabstände zwischen aufeinanderfolgenden Kunden sind unabhängig und identisch verteilt. Bezeichnet $(X(t))_{t\geq 0}$ die Anzahl von Kunden im System, so werden wir auch hier eine eingebettete Markov-Kette finden, indem wir den Prozess zu allen Zeitpunkten unmittelbar vor der Ankunft eines Kunden betrachten.
Den Zeitabstand zwischen der Ankunft des k-ten Kunden und des $(k+1)$-ten Kunden bezeichnen wir mit $T_0^{(k)}$. Weiters bezeichnen wir mit

$$X_k := X\left(T_0^{(1)} + \ldots + T_0^{(k-1)}\right) - 1$$

die Anzahl der Kunden, die der k-te Kunde im System vorfindet, und mit B_k die Anzahl von Kunden, die während $T_0^{(k)}$ bedient werden. Für $k \geq 0$ gilt $B_k \leq X_k + 1$ und

$$X_{k+1} = X_k + 1 - B_k.$$

Für die Verteilung von B_k gilt

$$\begin{aligned} b_n := \quad & \Pr[B_k = n] = \int \Pr\left[B_k = n \mid T_0^{(k)} = t\right] dF_{T_0}(t) \\ = \quad & \int e^{-\mu t} \frac{(\mu t)^n}{n!} dF_{T_0}(t). \end{aligned}$$

Da sich X_{k+1} aus X_k durch einen von allen anderen Zuständen unabhängigen Zuwachs ergibt, ist $(X_k)_{k\in\mathbb{N}}$ eine Markov-Kette. Sie hat die Übergangswahrscheinlichkeiten

$$\begin{aligned} p_{ij} &= \Pr[X_{k+1} = j \mid X_k = i] \\ &= \begin{cases} \Pr[B_k = i - j + 1] & \text{falls } 1 \le j \le i+1 \\ 0 & \text{falls } i+1 < j \end{cases} \\ &= \begin{cases} b_{i-j+1} & \text{falls } 1 \le j \le i+1 \\ 0 & \text{falls } i+1 < j \end{cases} \end{aligned}$$

Man beachte, dass die Formel für den Fall $j = 0$ nicht stimmt, weil der Server dann eine Zeit leerstehen kann, bevor der nächste Kunde eintrifft. Es gilt

$$p_{i0} = 1 - \sum_{j=1}^{i+1} p_{ij}.$$

Die Übergangsmatrix hat daher die Gestalt

$$\mathbf{P} = \begin{pmatrix} 1 - b_0 & b_0 & 0 & \dots & \dots \\ 1 - \sum_{i=0}^{1} b_i & b_1 & b_0 & 0 & \dots \\ 1 - \sum_{i=0}^{2} b_i & b_2 & b_1 & b_0 & 0 \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}.$$

Die stationäre Verteilung q für diesen Prozess erfüllt

$$q\mathbf{P} = q \text{ und } q\mathbf{1} = 1,$$

was mit

$$q_j = \sum_{n=0}^{\infty} q_{j-1+n} b_n \text{ für } i > 0 \text{ und } q_0 = \sum_{n=0}^{\infty} q_n \left(1 - \sum_{i=0}^{n} b_i\right)$$

gleichbedeutend ist. Man beachte, dass unendliche Summen auftreten. Mit der Wahrscheinlichkeit q_j findet ein Kunde, der ins stationäre System eintritt, gerade j Kunden im System vor. Man kann zeigen, dass q_j nur dann mit der stationären Wahrscheinlichkeit p_j, die über die gesamte Zeit mittelt, übereinstimmt, wenn die Zwischenankunftszeiten exponentialverteilt sind.
Die Berechnung der stationären Verteilung q verläuft über Operatoren. In Operatorschreibweise mit dem Shiftoperator D schreibt sich die obige Gleichung als

$$\left(D - \sum_{n=0}^{\infty} b_n D^n\right) q_{j-1} = 0,$$

was auf die charakteristische Gleichung

$$z - \beta(z) = 0$$

führt, wobei für

$$\begin{aligned}
\beta(z) &:= \sum_{n=0}^{\infty} b_n z^n \\
&= \int e^{-\mu t} \sum_{n=0}^{\infty} \frac{(\mu t z)^n}{n!} dF_{T_0}(t) \\
&= \int e^{-\mu t}\, e^{\mu t z}\, dF_{T_0}(t) \\
&= F_{T_0}^{*}(\mu(1-z))
\end{aligned}$$

gilt. Weil β eine monoton wachsende und konvexe Funktion ist mit $\beta(0) = b_0$, $\beta(1) = 1$ und $\beta'(1) = \frac{\mu}{\lambda}$ ist, hat die charakteristische Gleichung im Fall $\frac{\mu}{\lambda} \leq 1$ keine Lösung für $z \in]0, 1[$ und sonst genau eine Lösung, die wir mit r_0 bezeichnen.
Für $\mu > \lambda$ existiert also die stationäre Verteilung und hat die Gestalt

$$q_j = (1 - r_0)\, r_0^j.$$

Das Modell G/M/1 hat also, wenn man nur die Ankunftszeitpunkte betrachtet, formal dieselbe stationäre Verteilung wie das Modell M/M/1, wobei $\rho = \frac{\lambda}{\mu}$ durch r_0 ersetzt wird. Damit gehen alle Formeln des M/M/1 Modells durch Ersetzen von ρ durch r_0 in Formeln für das M/G/1 Modell über, wenn man nur die Ankunftszeitpunkte betrachtet. Es gilt also

$$L^{(A)} = \frac{r_0}{1 - r_0} \text{ und } L_q^{(A)} = \frac{r_0^2}{1 - r_0}.$$

Die erwarteten Verweilzeiten ergeben sich daraus, indem man sich überlegt, dass jeder eintreffende Kunde im Mittel $\frac{1}{\mu}$-mal $L^{(A)}$ Zeiteinheiten in der Warteschlange verbringt, bis er bedient wird, und im Mittel noch weitere $\frac{1}{\mu}$ Zeiteinheiten, bevor er das System verläßt. Es gilt also

$$W = \frac{1}{\mu(1 - r_0)} \text{ und } W_q = \frac{r_0}{\mu(1 - r_0)}.$$

Auch die zeitabhängigen Formeln für die Verweilzeiten gelten analog. Mit Hilfe der Formel von Little können nun L und L_q für den ursprünglichen Prozess berechnet werden, nämlich

$$L = \frac{\lambda}{\mu(1 - r_0)} \text{ und } L_q = \frac{\lambda r_0}{\mu(1 - r_0)}.$$

Für die Berechnung von r_0 müssen numerische Verfahren verwendet werden, zB sukzessive Substitution.
Beispiel[1] Seite 253.

5.4 G/M/c Modelle

Siehe [1] Seite 255.

Kapitel 6

Allgemeinere Modelle

6.1 G/G/1 Modelle

Die Gleichung von Lindley ist eine Integralgleichung für die Wartezeit eines Kunden, wenn der Prozess im Gleichgewicht ist. Bezeichne $T_0^{(k)}$ die Zwischenankunftszeit zwischen dem k-ten und dem $(k+1)$-ten Kunden, $T_q^{(k)}$ die Wartezeit des k-ten Kunden und $T_s^{(k)}$ die Bedienungszeit des k-ten Kunden, dann gilt

$$\begin{aligned} T_q^{(k+1)} &= \begin{cases} T_q^{(k)} + T_s^{(k)} - T_0^{(k)} & \text{falls } T_q^{(k)} + T_s^{(k)} - T_0^{(k)} > 0 \\ 0 & \text{sonst} \end{cases} \\ &= \max\left(0, T_q^{(k)} + T_s^{(k)} - T_0^{(k)}\right). \end{aligned}$$

Der Prozess $\left(T_q^{(k)}\right)_{k\in\mathbb{N}}$ ist also ein Markov-Prozess mit diskreter Zeit. Wegen der obigen Beziehung gilt für die (linksseitig stetige) Verteilungsfunktion von $T_q^{(k+1)}$

$$\begin{aligned} F_{T_q^{(k+1)}}(t) &= \Pr\left[T_q^{(k+1)} \le t\right] \\ &= \Pr\left[T_q^{(k)} + T_s^{(k)} - T_0^{(k)} \le t\right]. \end{aligned}$$

Wenn wir nun eine Zufallvariable $U^{(k)} := T_s^{(k)} - T_0^{(k)}$ mit der Verteilungsfunktion $F_{U^{(k)}}$ einführen, so gilt für $0 \le t < \infty$

$$F_{T_q^{(k+1)}}(t) = \int_{-\infty}^{t} F_{T_q^{(k)}}(t-x)\, dF_{U^{(k)}}(x).$$

Für das Gleichgewicht, wenn $\lambda/\mu < 1$ gilt, folgt die *Gleichung von Lindley*

$$\begin{aligned} F_{T_q}(t) &= \begin{cases} \int_{-\infty}^{t} F_{T_q}(t-x)\, dF_U(x) & \text{falls } 0 \le t < \infty \\ 0 & \text{falls } t < 0 \end{cases} \\ &= \begin{cases} -\int_{0}^{\infty} F_{T_q}(y)\, dF_U(t-y) & \text{falls } 0 \le t < \infty \\ 0 & \text{falls } t < 0 \end{cases}. \end{aligned}$$

Dabei ist $U = T_s - T_0$ und F_U die Verteilungsfunktion von U, dh die Faltung von F_{T_s} und F_{-T_0}.
Wie man mit einer solchen Gleichung umgeht und einfache Beispiele dazu findet man in [1] ab Seite 274.

6.2 Semi-Markov Modelle

Hier bildet die Folge der Zustände eine Markov-Kette, die Übergangszeiten sind aber im allgemeinen nicht exponentialverteilt und hängen von Ausgangs- und vom Endzustand ab. Kurze Einleitung siehe [1] Seite 288.

6.3 Design und Kontrolle

Probleme des Design von Bedienungssystemen sind uns im Rahmen von Vorlesung und Übung schon häufig begegnet, es geht darum, Modell und Parameter so zu wählen, dass gewisse Zielgrößen, etwa die Wartezeit, die mittlere Systemgröße oder erwartete Kosten, minimiert werden. Für eine weitere Diskussion siehe [1] Seite 302. Bei Kontrollproblemen geht es darum, während des Betriebs eines Bedienungssystems so einzugreifen, zB Reserveserver auf- und auch wieder zuzumachen, dass gewisse Optimalitätskriterien erfüllt sind, viele Referenzen siehe [1] Seite 308. Neben den Kontrollproblemen für die Gleichgewichtssituation gibt es auch Systeme, die von einem "vollen" Systemzustand ausgehen, und wo jene Strategie gesucht ist, die das System am günstigsten leert, wenn keine neuen Kunden eintreffen.

6.4 Statistische Methoden

In der Bedienungstheorie werden statistische Methoden benötigt, um Parameter zu schätzen oder zu testen, ob gewisse Parameter und Verteilungen vorliegen. In [1] Seite 313 sind zB Maximum-Likelihood Schätzer für die Parameter eines M/M/1 Modells durchgerechnet, das über einen gewissen Zeitraum im Gleichgewicht vollständig beobachtet wurde.

Literaturverzeichnis

[1] Gross, D., Harris, C.M., *Fundamentals of Queueing Theory* , 1998, Wiley Series in Probability and Statistics