Beweis der Erlang-C-Formel

Alexander Herzog

alexander.herzog@tu-clausthal.de
a-herzog.github.io/QueueCalc

1 Das Erlang-C-Modell

Das Erlang-C-Warteschlangenmodell beschreibt ein Bediensystem mit exponentiell verteilten Zwischenankunftszeiten mit einer Ankunftsrate von $\lambda>0$, exponentiell verteilten Bedienzeiten mit einer Bedienrate von $\mu>0$ und $c\in\mathbb{N}$ Bedienern. Die ankommenden Kunden bilden eine Warteschlange vor dem Bedienschalter. Jedes Mal, wenn ein Bediener frei wird, wird der Kunde, der am längsten gewartet hat, zu dem freien Bediener geleitet (FIFO Bedienreihenfolge; First-In-First-Out), siehe Abbildung 1.

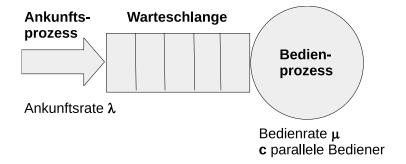


Abbildung 1: Erlang-C-Warteschlangenmodell

Die langfristige Auslastung des Systems beträgt

$$\rho := \frac{\lambda}{cu} \ .$$

Das System kann nur langfristig stabil arbeiten, wenn ho < 1 ist (die langfristige Ankunftsrate muss unterhalb der verfügbaren Bedienleistung liegen).

Üblicherweise interessiert man sich für die folgenden Kenngrößen des Systems:

- $P(W \le t)$: Wahrscheinlichkeit dafür, dass ein neu eintreffender Kunde nicht länger als $t \ge 0$ Sekunden warten muss (Erlang-C-Formel),
- **E**[W]: mittlere Wartezeit,
- $\mathbf{E}[V]$: mittlere Verweilzeit,
- $\mathbf{E}[N_O]$: mittlere Anzahl an Kunden in der Warteschlange,
- ullet $\mathbf{E}[N]$: mittlere Anzahl an Kunden im System.

2 Idee zur Herleitung von $P(W \le t)$

Die Idee zur Berechnung der Wartezeitverteilung der Kunden $P(W \leq t)$ (die Wahrscheinlichkeit, dass ein Kunde höchstens $t \geq 0$ Sekunden warten muss) besteht darin, die Wartezeitverteilung für alle möglichen Zustände, in denen sich das System befinden kann, zu bestimmen und diese Zeiten mit der Wahrscheinlichkeit zu gewichten, dass sich das System im jeweiligen Zustand befindet. Dies führt zu folgender Formel:

$$P(W \le t) = \sum_{n=0}^{\infty} P_{\mu, n - (c-1)}(W \le t) \cdot p_n , \qquad (1)$$

wobei p_n die Wahrscheinlichkeit dafür ist, dass sich n Kunden im System befinden, und $P_{\mu,m}(W \leq t)$ die Wahrscheinlichkeit ist, dass ein neuer Kunde t Sekunden oder weniger warten muss, wenn die Bedienrate $\mu>0$ ist und es m Kunden gibt, die zunächst zu Ende bedient werden müssen, bevor der Bedienprozess für den neuen Kunden beginnt. Für m=0 beginnt der neue Bedienprozess sofort $(P_{\mu,0}(W \leq t)=1$ für alle $t \geq 0)$ und aus Gründen der Konsistenz nehmen wir dasselbe für m<0 an

Fall $n \geq c$:

Wenn sich n < c Kunden im System befinden, kann ein neu eintreffender Kunde sofort bedient werden, da in diesem Fall nicht alle Bediener belegt sind. Wenn $n \ge c$ ist, müssen zunächst ein oder mehrere bereits laufende Bedienprozesse beendet werden, und dann müssen alle Kunden, die bereits vor dem neu angekommenen Kunden gewartet haben, bedient werden, bevor der neu angekommene Kunde bedient werden kann. Daher muss zwischen den Bediendauern von Kunden, deren Bedienprozess noch nicht begonnen hat, und den verbleibenden Bedienzeiten der Kunden, die sich bereits im Bedienprozess befinden, unterschieden werden. Da das Erlang-Modell jedoch die Exponentialverteilung für die Bedienzeiten verwendet, unterliegen diese beiden Zeitdauern derselben Wahrscheinlichkeitsverteilung mit demselben Parameter. Dies ist auf die Gedächtnislosigkeit der Exponentialverteilung zurückzuführen. (Diese Eigenschaft der Exponentialverteilung wird weiter unten in Abschnitt 3 bewiesen).

Fall n < c:

Für n < c Kunden im System ist die Wahrscheinlichkeit, dass ein neu eintreffender Kunde weniger als $t \geq 0$ Sekunden warten muss, für alle $t \geq 0$ immer 1 (da der neu eintreffende Kunde überhaupt nicht warten muss). Das bedeutet, dass $P_{\mu,n-(c-1)}(W \leq t) = 1$ für n < c und für alle $t \geq 0$ gilt und sich die Formel (1) vereinfacht zu:

$$P(W \le t) = \sum_{n=0}^{c-1} p_n + \sum_{n=c}^{\infty} P_{\mu, n-(c-1)}(W \le t) \cdot p_n .$$
 (2)

Da $P_{\mu,n-(c-1)}(W \leq t)$ die Hintereinandeausführung mehrerer Exponentialverteilungen bedeutet (und somit eine Erlang-Verteilung ist, siehe Abschnitt 5) und p_n die Wahrscheinlichkeit ist, dass sich die entsprechende Markov-Kette im Zustand n befindet, werden in den folgenden Abschnitten Ergebnisse zu Exponentialverteilungen (Abschnitt 3), Erlang-Verteilungen (Abschnitt 5) und Geburts- und Todesprozessen (Abschnitt 6) vorgestellt. Auf Basis dieser Vorüberlegungen wird es dann möglich sein, die Erlang-C-Formel (2) in Abschnitt 7 zu berechnen.

3 Exponentialverteilung

Da die Zwischenankunfts- und die Bedienzeiten in einem Erlang-Modell der Exponentialverteilung unterliegen, sollen im folgenden zunächst einige relevante Ergebnisse zu dieser Verteilung vorgestellt und bewiesen werden.

Definition 1 (Dichte der Exponentialverteilung). Die Funktion

$$f_{\lambda}(x) := \begin{cases} \lambda e^{-\lambda x}, & x \ge 0, \\ 0, & x < 0 \end{cases}$$

 $\min \lambda > 0$ wird die Dichte der Exponentialverteilung genannt.

Theorem 1 (Verteilungsfunktion der Exponentialverteilung). Die Funktion

$$F_{\lambda}(x) := \begin{cases} 1 - e^{-\lambda x}, & x \ge 0, \\ 0, & x < 0 \end{cases}$$

ist die Verteilungsfunktion einer Exponentialverteilung mit Dichte $f_{\lambda}(x)$, $\lambda > 0$.

Beweis. Gemäß der Definition der Verteilungsfunktion $F(x) := \int_{-\infty}^x f(t) \, \mathrm{d}t$ gilt:

$$F_{\lambda}(x) = \int_{0}^{x} \lambda e^{-\lambda t} dt = \left[-e^{-\lambda t} \right]_{t=0}^{t=x} = -e^{-\lambda x} - (-1) = 1 - e^{-\lambda x}$$

für $x \ge 0$ und $F_{\lambda}(x) = 0$ sonst.

Theorem 2. Die Funktion $f_{\lambda}(x)$, $\lambda > 0$, ist die Dichte einer Wahrscheinlichkeitsverteilung, d.h. es gelten $f_{\lambda}(x) \geq 0$ und $\int_{-\infty}^{\infty} f_{\lambda}(x) \, dx = 1$.

Beweis. Die Behauptung $f_{\lambda}(x) \geq 0$ ergibt sich unmittelbar aus der Definition von $f_{\lambda}(x)$. Mit dem vorherigen Satz gilt:

$$\int_{-\infty}^{\infty} f_{\lambda}(t) dt = \lim_{x \to \infty} F_{\lambda}(x) = \lim_{x \to \infty} 1 - e^{-\lambda x} = 1.$$

Theorem 3 (Kenngrößen der Exponentialverteilung). Für eine Exponentialverteilung mit Parameter $\lambda > 0$ gelten:

1. $\mathbf{E}[X] = \frac{1}{\lambda}$,

2. $\mathbf{Std}[X] = \frac{1}{\lambda}$ und

3. CV[X] = SCV[X] = 1.

Beweis. 1. Mit der Definition von $\mathbf{E}[X]$ folgt:

$$\begin{split} \mathbf{E}[X] &= \int_{-\infty}^{\infty} x \cdot f_{\lambda}(x) \, \mathrm{d}x = \int_{0}^{\infty} \lambda x \mathrm{e}^{-\lambda x} \, \mathrm{d}x \\ &= \left[\lambda x \cdot \left(-\frac{1}{\lambda} \right) \mathrm{e}^{-\lambda x} \right]_{x=0}^{x=\infty} - \int_{0}^{\infty} \lambda \left(-\frac{1}{\lambda} \right) \mathrm{e}^{-\lambda x} \, \mathrm{d}x = 0 + \int_{0}^{\infty} \mathrm{e}^{-\lambda x} \, \mathrm{d}x \\ &= \left[\left(-\frac{1}{\lambda} \right) \mathrm{e}^{-\lambda x} \right]_{x=0}^{x=\infty} = 0 - \left(-\frac{1}{\lambda} \right) = \frac{1}{\lambda} \, . \end{split}$$

2. Mit der Definition von $\mathbf{E}[X^2]$ folgt zunächst:

$$\begin{split} \mathbf{E}[X^2] &= \int_{-\infty}^{\infty} x^2 \cdot f_{\lambda}(x) \, \mathrm{d}x = \int_{0}^{\infty} \lambda x^2 \mathrm{e}^{-\lambda x} \, \mathrm{d}x \\ &= \left[\lambda x^2 \left(-\frac{1}{\lambda} \right) \mathrm{e}^{-\lambda x} \right]_{x=0}^{x=\infty} - \int_{0}^{\infty} 2\lambda x \left(-\frac{1}{\lambda} \right) \mathrm{e}^{-\lambda x} \, \mathrm{d}x = 0 + \int_{0}^{\infty} 2x \mathrm{e}^{-\lambda x} \, \mathrm{d}x \\ &= \left[2x \left(-\frac{1}{\lambda} \right) \mathrm{e}^{-\lambda x} \right]_{x=0}^{x=\infty} - \int_{0}^{\infty} 2 \left(-\frac{1}{\lambda} \right) \mathrm{e}^{-\lambda x} \, \mathrm{d}x = 0 + \frac{2}{\lambda} \left[\left(-\frac{1}{\lambda} \right) \mathrm{e}^{-\lambda x} \right]_{x=0}^{x=\infty} \\ &= 0 - \frac{2}{\lambda} \left(-\frac{1}{\lambda} \right) = \frac{2}{\lambda^2} \, . \end{split}$$

Mit dem Verschiebungssatz gilt:

$$\mathbf{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}.$$

Und damit gilt schließlich für die Standardabweichung:

$$\mathbf{Std}[X] = \sqrt{\mathbf{Var}[X]} = \frac{1}{\lambda}.$$

3. Aus 1 und 2 folgt unmittelbar:

$$\mathbf{CV}[X] = rac{\mathbf{Std}[X]}{|\mathbf{E}[X]|} = rac{rac{1}{\lambda}}{rac{1}{\lambda}} = 1$$
 .

Des Weiteren ist $\mathbf{SCV}[X] = (\mathbf{CV}[X])^2 = 1.$

3.1 Gedächtnislosigkeit der Exponentialverteilung

Theorem 4 (Gedächtnislosigkeit der Exponentialverteilung). Ist eine Zeitdauer exponentiell verteilt und ist das Ereignis bis zum Zeitpunkt $s \geq 0$ noch nicht eingetreten, so ist die Verteilungsfunktion der Wahrscheinlichkeitsverteilung, dass das Ereignis innerhalb der nächsten $t \geq 0$ Zeiteinheiten eintritt (d. h. bis zu dem Zeitpunkt s+t), wieder exponentiell verteilt mit demselben Parameter $\lambda > 0$, d. h. es gilt:

$$P(X \le t + s | X \ge s) = P(X \le t), \tag{3}$$

wobei $P(X \leq a | X \geq b)$ die bedingte Wahrscheinlichkeit für das Ereignis " $X \leq a$ " unter dem Vorauswissen, dass $X \geq b$ ist, sei.

Beweis. Mit den Rechenregeln für bedingte Wahrscheinlichkeiten gilt:

$$P(X \le t + s | X \ge s) = \frac{P(\{X \le t + s\} \cap \{X \ge s\})}{P(X \ge s)} = \frac{F_{\lambda}(t + s) - F_{\lambda}(s)}{1 - F_{\lambda}(s)}$$

$$= \frac{1 - e^{-\lambda(t + s)} - 1 + e^{-\lambda s}}{1 - (1 - e^{-\lambda s})} = \frac{e^{-\lambda s} - e^{-\lambda(t + s)}}{e^{-\lambda s}} = 1 - e^{-\lambda t}$$

$$= F_{\lambda}(t) = P(X \le t).$$

Konsequenz:

Ist die Bediendauer eines Kunden exponentiell verteilt mit Parameter $\lambda>0$ und befindet sich der Kunde bereits in Bedienung, so ist seine Restbediendauer wiederum exponentiell verteilt mit demselben Parameter λ . Soll die Verteilungsfunktion der Wartezeit eines neu eintreffenden Kunden bestimmt werden, so muss also nicht zwischen den Restbedienzeiten der momentan in Bedienung befindlichen Kunden und den Bedienzeiten der vor dem neu eingetroffenen Kunden wartenden anderen Kunden unterschieden werden.

Theorem 5. Die Exponentialverteilung $F_{\lambda}(x)$ ist die einzige stetige Verteilungsfunktion mit F(x)=0 für $x\leq 0$, die die Eigenschaft der

Beweis. Der Beweis erfolgt in drei Schritten:

4

Schritt 1: Funktionalgleichung zur Beschreibung der Gedächtnislosigkeit

Um den Beweis formal möglichst einfach darstellen zu können, wird zunächst die Funktion G(t):=1-F(t) definiert, die die Wahrscheinlichkeit angibt, dass das Ereignis bis zum Zeitpunkt $t\geq 0$ noch nicht eingetreten ist. (Diese Definition stammt aus der Zuverlässigkeitstheorie; dort wird G(t) auch die Überlebenswahrscheinlichkeit genannt.) Aus Gleichung (3) wird damit mit Hilfe der Rechenregeln für bedingte Wahrscheinlichkeiten:

$$\begin{split} P(X \leq t + s | X \geq s) &= P(X \leq t) \iff \frac{F(t + s) - F(s)}{1 - F(s)} = F(t) \\ \iff \frac{1 - G(t + s) - (1 - G(s))}{G(s)} &= 1 - G(t) \\ \iff -G(t + s) + G(s) &= G(s) - G(s)G(t) \iff G(s + t) = G(s)G(t) \,. \end{split}$$

Es ist jetzt also zu zeigen, dass nur die Überlebenswahrscheinlichkeit unter Annahme einer Exponentialverteilung, d.h. $G_{\lambda}(t)=\mathrm{e}^{-\lambda t}$, die Gleichung G(s+t)=G(s)G(t) erfüllt.

Schritt 2: Explizite Darstellung für G(t)

Aus G(s+t)=G(s)G(t) folgt für $a\in\mathbb{N}$ und $b\in\mathbb{N}$:

$$\begin{split} G\left(\frac{a}{b}\right) &= G\left(\frac{1}{b} + \frac{a-1}{b}\right) = G\left(\frac{1}{b}\right) \cdot G\left(\frac{a-1}{b}\right) = \ldots = \underbrace{G\left(\frac{1}{b}\right) \cdots G\left(\frac{1}{b}\right)}_{a \text{ Faktoren}} \\ &= \left[G\left(\frac{1}{b}\right)\right]^{a}. \end{split}$$

 $\text{Mit } 1 = \tfrac{b}{b} \text{ gilt außerdem } G(1) = \left[G\left(\tfrac{1}{b}\right)\right]^b \text{ bzw. } \left[G(1)\right]^{\frac{1}{b}} = G\left(\tfrac{1}{b}\right). \text{ Damit folgt weiter: } G(1) = \left[G\left(\tfrac{1}{b}\right)\right]^b \text{ bzw. } \left[G(1)\right]^{\frac{1}{b}} = G\left(\tfrac{1}{b}\right).$

$$G\left(\frac{a}{b}\right) = \left[G\left(\frac{1}{b}\right)\right]^a = \left[G(1)\right]^{\frac{a}{b}}.$$

Das bedeutet, dass der Zusammenhang $G(q)=[G(1)]^q$ für alle positiven, rationalen Zahlen q gilt. Aufgrund der Tatsache, dass F als stetig angenommen wurde und somit auch G stetig ist, gilt $G(t)=[G(1)]^t$ für alle reellen $t\geq 0$. Weiter gilt mit den Rechenregeln für die Exponentialfunktion und Logarithmen:

$$G(t) = e^{\ln(G(1)^t)} = e^{t \ln(G(1))}$$
.

Mit der Wahl von $\lambda:=-\ln(G(1))$ folgt $G(t)=\mathrm{e}^{-\lambda t}$ bzw. $F(t)=1-\mathrm{e}^{-\lambda t}$. Zu zeigen bleibt, dass $0<\lambda<\infty$ bzw. 0< G(1)<1 gilt.

Schritt 3: Nachweis, dass 0 < G(1) < 1 gilt

Es ist G(1) = 1 - F(1). Damit gilt sofort $0 \le G(1) \le 1$. Zu zeigen bleibt $0 \ne G(1) \ne 1$.

- Wenn G(1)=0 wäre, würde $G(x)=\left[G(1)\right]^x=0^x=0$ für alle x>0 gelten. Damit wäre F(x)=1 für alle x>0. Da aber F als stetig vorausgesetzt wurde und außerdem F(0)=0 nach Voraussetzung galt, kann dies nicht der Fall sein.
- Wenn G(1)=1 wäre, würde $G(x)=[G(1)]^x=1^x=1$ für alle $x\geq 0$ gelten. Damit wäre F(x)=0 für alle $x\geq 0$, was im Widerspruch dazu steht, dass F(x) die Verteilungsfunktion einer Wahrscheinlichkeitsverteilung ist und somit $\lim_{x\to\infty}F(x)=1$ gelten muss.

4 Faltung von Wahrscheinlichkeitsverteilungen

Werden zwei Prozesse, die durch Wahrscheinlichkeitsverteilungen beschrieben werden können, hintereinander ausgeführt, so spricht man davon, dass die Dichten der beiden Wahrscheinlichkeitsverteilungen miteinander gefaltet werden. Das Ergebnis der Faltung ist die Dichte der Wahrscheinlichkeitsverteilung, die den Gesamtprozess beschreibt.

Definition 2 (Faltung). Sind f(x) und g(x) die Dichten von zwei Wahrscheinlichkeitsverteilungen, so nennt man

$$f * g(x) := \int_{-\infty}^{\infty} f(t)g(x-t) dt$$

die Faltung von f und g.

Anschaulich werden durch obige Formel alle denkbaren Fälle, wie die Gesamtzeit x auf die beiden Teilprozesse f und g aufgeteilt werden kann, aufintegriert.

Da die Faltung durch ein Integral über das Produkt von zwei Dichten definiert ist, gilt auch hier wieder, dass diese für alle Wahrscheinlichkeitsverteilungen, bei denen die Bestimmung der Stammfunktion der Dichte schwierig ist – was für fast alle stetigen Wahrscheinlichkeitsverteilungen außer der Exponentialverteilung gilt –, nur sehr schwierig oder überhaupt nicht explizit berechenbar ist. Daher wird in den Erlang-Formeln grundsätzlich die Exponentialverteilung für die Bedienzeiten verwendet.

5 Erlang-Verteilung

Wie sich im Folgenden zeigen wird, beschreibt die Erlang-Verteilung einen Prozess, der aus der Hintereinanderausführung von $n \in \mathbb{N}$ Teilprozessen besteht, deren Zeitdauern jeweils mit demselben Parameter $\lambda > 0$ exponentiell verteilt sind.

In einem M/M/1-System lässt sich die Wartezeitverteilung eines neu eintreffenden Kunden durch die Bediendauern der vor ihm wartenden Kunden und der Restbediendauer des gerade in Bedienung befindlichen Kunden darstellen. Aufgrund der Gedächtnislosigkeit der Exponentialverteilung ist die Restbediendauer des momentan in Bedienung befindlichen Kunden aber genauso verteilt, wie seine gesamte Bediendauer (siehe Gleichung (3)). D.h. befinden sich vor dem neu eintreffenden Kunden bereits $k \geq 0$ Kunden in der Warteschlange und sind die Bediendauern mit Parameter $\mu \geq 0$ exponentiell verteilt, so ist die Wartezeit des neuen Kunden Erlang-verteilt mit den Parametern μ und k+1.

Definition 3 (Dichte der Erlang-Verteilung). Die Funktion

$$f_{\lambda,n}(x) := \begin{cases} \frac{\lambda^n x^{n-1}}{(n-1)!} e^{-\lambda x}, & x \ge 0, \\ 0, & x < 0 \end{cases}$$

 $\mbox{mit } \lambda > 0 \mbox{ und } n \in \mathbb{N} \mbox{ wird die Dichte der Erlang-Verteilung genannt.}$

Theorem 6 (Zusammenhang mit der Exponentialverteilung). Die Funktion $f_{\lambda,n}(x)$, $\lambda > 0$, $n \in \mathbb{N}$ ist die n-fache Faltung von $f_{\lambda}(x)$.

Beweis. Der Beweis wird per Induktion geführt. Offensichtlich gilt $f_{\lambda,1}(x) = \lambda e^{-\lambda x} = f_{\lambda}(x)$. Für die Faltung von zwei exponentiell verteilten Zeitdauern gilt:

$$f_{\lambda,1} * f_{\lambda}(x) = f_{\lambda} * f_{\lambda}(x) = \int_0^x \lambda e^{-\lambda t} \cdot \lambda e^{-\lambda(x-t)} dt = \left[\lambda^2 e^{-\lambda x} t\right]_{t=0}^{t=x} = \lambda^2 x e^{-\lambda x} = f_{\lambda,2}(x)$$

für $x \geq 0$ und $f_{\lambda,2}(x) = 0$ sonst. Es wird nun angenommen, dass

$$f_{\lambda,n}(x) = \underbrace{f_{\lambda} * \cdots * f_{\lambda}}_{n \text{ -fach}}(x)$$

für ein $n \in \mathbb{N}$ gelte. Dann gilt:

$$f_{\lambda,n} * f_{\lambda}(x) = \int_{0}^{x} \frac{\lambda^{n} t^{n-1}}{(n-1)!} e^{-\lambda t} \cdot \lambda e^{-\lambda(x-t)} dt = \frac{\lambda^{n+1}}{(n-1)!} e^{-\lambda x} \int_{0}^{x} t^{n-1} dt$$
$$= \frac{\lambda^{n+1}}{(n-1)!} e^{-\lambda x} \left[\frac{1}{n} t^{n} \right]_{t=0}^{t=x} = \frac{\lambda^{n+1} x^{n}}{n!} e^{-\lambda x} = f_{\lambda,n+1}(x)$$

für $x \ge 0$ und $f_{\lambda, n+1}(x) = 0$ sonst.

Theorem 7 (Verteilungsfunktion der Erlang-Verteilung). Die Funktion

$$F_{\lambda,n}(x) := \begin{cases} 1 - e^{-\lambda x} \sum_{i=0}^{n-1} \frac{(\lambda x)^i}{i!}, & x \ge 0, \\ 0, & x < 0 \end{cases}$$

ist die Verteilungsfunktion einer Erlang-Verteilung mit Dichte $f_{\lambda,n}(x)$, $\lambda > 0$, $n \in \mathbb{N}$.

Beweis. Der Beweis wird ebenfalls per Induktion geführt. Für den Fall n=1 gilt:

$$F_{\lambda,1}(x) = 1 - e^{-\lambda x} \sum_{i=0}^{1-1} \frac{(\lambda x)^i}{i!} = F_{\lambda}(x) .$$

Da $F_{\lambda}(x)$ die Verteilungsfunktion der Exponentialverteilung ist und $f_{\lambda,1}(x)=f_{\lambda}(x)$ galt, ist damit der Induktionsanfang gezeigt. Es wird nun angenommen, dass bereits bekannt sei, dass

$$F_{\lambda,n-1}(x) = 1 - e^{-\lambda x} \sum_{i=0}^{n-2} \frac{(\lambda x)^i}{i!}$$

für ein $n \in \mathbb{N}$, $n \ge 2$, für $x \ge 0$ gelte. Dann ist:

$$\begin{split} F_{\lambda,n}(x) &= \int_0^x f_{\lambda,n}(t) \, \mathrm{d}t = \int_0^x \frac{\lambda^n t^{n-1}}{(n-1)!} \mathrm{e}^{-\lambda t} \, \mathrm{d}t \\ &= \left[\frac{\lambda^n t^{n-1}}{(n-1)!} \left(-\frac{1}{\lambda} \mathrm{e}^{-\lambda t} \right) \right]_{t=0}^{t=x} - \int_0^x \frac{\lambda^n t^{n-2}}{(n-2)!} \left(-\frac{1}{\lambda} \mathrm{e}^{-\lambda t} \right) \, \mathrm{d}t \\ &= -\frac{(\lambda x)^{n-1}}{(n-1)!} \mathrm{e}^{-\lambda x} + \int_0^x f_{\lambda,n-1}(t) \, \mathrm{d}t = -\frac{(\lambda x)^{n-1}}{(n-1)!} \mathrm{e}^{-\lambda x} + F_{\lambda,n-1}(t) \\ &= -\frac{(\lambda x)^{n-1}}{(n-1)!} \mathrm{e}^{-\lambda x} + 1 - \mathrm{e}^{-\lambda x} \sum_{i=0}^{n-2} \frac{(\lambda x)^i}{i!} = 1 - \mathrm{e}^{-\lambda x} \sum_{i=0}^{n-1} \frac{(\lambda x)^i}{i!} \end{split}$$

für $x \ge 0$ und $F_{\lambda,n}(x) = 0$ sonst.

Theorem 8. Die Funktion $f_{\lambda,n}(x)$, $\lambda > 0$, $n \in \mathbb{N}$, ist die Dichte einer Wahrscheinlichkeitsverteilung, d. h. es gelten $f_{\lambda,n}(x) \geq 0$ und

$$\int_{-\infty}^{\infty} f_{\lambda,n}(x) \, dx = 1 \; .$$

Beweis. Die Behauptung $f_{\lambda,n}(x) \geq 0$ ergibt sich unmittelbar aus der Definition von $f_{\lambda,n}(x)$. Mit der Reihendarstellung der Exponentialfunktion liefert die n-fache Anwendung der Regel von l'Hospital:

$$F_{\lambda,n}(x) = 1 - e^{-\lambda x} \sum_{i=0}^{n-1} \frac{(\lambda x)^i}{i!} = 1 - \frac{\sum_{i=0}^{n-1} \frac{(\lambda x)^i}{i!}}{\sum_{i=0}^{\infty} \frac{(\lambda x)^i}{i!}} \xrightarrow{x \to \infty} 1.$$

Damit gilt:

$$\int_{-\infty}^{\infty} f_{\lambda,n}(t) dt = \lim_{x \to \infty} F_{\lambda,n}(x) = 1.$$

Theorem 9 (Kenngrößen der Erlang-Verteilung). Für eine Erlang-Verteilung mit Parametern $\lambda > 0$ und $n \in \mathbb{N}$ gelten:

1. $\mathbf{E}[X] = \frac{n}{\lambda}$,

2.
$$\mathbf{Std}[X] = \frac{\sqrt{n}}{\lambda}$$
 und

3.
$$\mathbf{CV}[X] = \frac{1}{\sqrt{n}} \text{ and } \mathbf{SCV}[X] = \frac{1}{n}.$$

Beweis. Die Erlang-Verteilung mit den Parametern $\lambda>0$ und $n\in\mathbb{N}$ stellt die Hintereinanderausführung von n unabhängigen Exponentialverteilungen mit Parameter λ dar. Aufgrund der Unabhängigkeit der einzelnen Teilprozesse dürfen nicht nur die Erwartungswerte addiert werden, sondern auch die Varianzen. Für den Erwartungswert der Erlang-Verteilung folgt damit sofort

$$\mathbf{E}[X] = n \cdot \frac{1}{\lambda} = \frac{n}{\lambda} .$$

Für die Standardabweichung gilt:

$$\mathbf{Std}[X] = \sqrt{\mathbf{Var}[X]} = \sqrt{n \cdot \frac{1}{\lambda^2}} = \frac{\sqrt{n}}{\lambda}.$$

Damit gilt unmittelbar für den Variationskoeffizienten:

$$\mathbf{CV}[X] = \frac{\mathbf{Std}[X]}{|\mathbf{E}[X]|} = \frac{\frac{\sqrt{n}}{\lambda}}{\frac{n}{\lambda}} = \frac{1}{\sqrt{n}}.$$

Des Weiteren ist damit $\mathbf{SCV}[X] = (\mathbf{CV}[X])^2 = \frac{1}{n}$.

6 Geburts- und Todesprozesse

Markov-Prozesse, bei denen nur Übergänge zu dem jeweils nächst höheren Zustand oder dem nächst niedrigeren Zustand möglich sind (z.B. Ankünfte und Abgänge von einzelnen Kunden), werden Geburts- und Todesprozesse genannt. Eine "Geburt" stellt die Ankunft eines Kunden dar, ein "Tod" den Abgang eines zu Ende bedienten Kunden, siehe Abbildung 2.

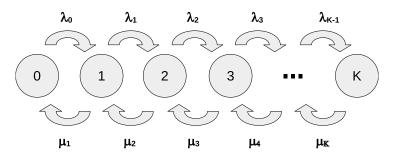


Abbildung 2: Ein Geburts- und Todesprozess dargestellt als Markov-Kette

6.1 Ankunfts- und Bedienrate

Für die zustandsabhängige Ankunftsrate am System in einem Bedienprozess mit $K \in \mathbb{N}$ Warte- und Bedienplätzen gilt für den Zustand $n \in \mathbb{N}_0$:

$$\lambda_n^{\text{Sys}} := \begin{cases} \lambda, & n < K \\ 0, & n = K \end{cases} \tag{4}$$

wobei $\lambda>0$ die Ankunftsrate bezogen auf die Kunden ist. Befinden sich bereits n=K Kunden im System, so können keine weiteren Kunden am System eintreffen, d.h. es ist in diesem Fall $\lambda^{\mathrm{Sys}}=0$.

Für die zustandsabhängige Bedienrate bezogen auf das gesamte System gilt:

$$\mu_n^{\text{Sys}} := \begin{cases} n\mu, & n \le c, \\ c\mu, & n > c, \end{cases} \tag{5}$$

wobei $\mu > 0$ die Bedienrate eines einzelnen Bedieners ist.

6.2 Transiente Zustandswahrscheinlichkeiten

Es wird nun angenommen, dass die Wahrscheinlichkeit $p_n(t)$, mit der sich zu einem Zeitpunkt $t \geq 0$ insgesamt $n \in \mathbb{N}_0$ Kunden im System befinden, zu einem bestimmten vorgegebenen Zeitpunkt $t \geq 0$ bekannt sei. Auf dieser Basis wird bestimmt, mit welcher Wahrscheinlichkeit zu einem Zeitpunkt $t + \Delta t$ sich $n \in \mathbb{N}_0$ Kunden im System befinden. Dabei kann man sich $\Delta t > 0$ als einen kleinen Augenblick vorstellen, der seit dem Zeitpunkt t vergangen ist. Im Fall n=0 setzt sich $p_0(t+\Delta t)$ aus der Wahrscheinlichkeit, dass zum Zeitpunkt t auch bereits 0 Personen im System waren und kein Kunde eingetroffen ist, und der Wahrscheinlichkeit, dass zum Zeitpunkt t ein Kunde im System war und dieser in der Zeit Δt zu Ende bedient wurde, zusammen. Des Weiteren könnten in der Zeitdauer Δt auch ein Kunde eingetroffen und bereits wieder aus dem System ausgeschieden sein usw. Im Vergleich zu den zuerst genannten beiden Wahrscheinlichkeiten sind die Wahrscheinlichkeiten für solche Ereignisse jedoch verschwindend gering – mathematisch spricht man davon, dass sie für $\Delta t \to 0$ schneller gegen 0 konvergieren als Δt . Dies drückt man durch das Landau-Symbol $o(\Delta t)$ aus. Damit ergibt sich für n=0:

$$p_0(t + \Delta t) = p_0(t)(1 - \Delta t \lambda_0^{\text{Sys}}) + p_1(t)\Delta t \mu_1^{\text{Sys}} + o(\Delta t)$$
.

Für $n=1,\ldots,K-1$ ergibt sich zusätzlich die Möglichkeit, dass der Zustand n aus dem Zustand n-1 heraus erreicht wurde:

$$p_n(t + \Delta t) = p_n(t)(1 - \Delta t\lambda_n^{\text{Sys}} - \Delta t\mu_n^{\text{Sys}}) + p_{n-1}(t)\Delta t\lambda_{n-1}^{\text{Sys}} + p_{n+1}(t)\Delta t\mu_{n+1}^{\text{Sys}} + o(\Delta t).$$

Für n=K schließlich kann der Zustand nur aus den Vorgängerzuständen n und n-1 heraus erreicht werden:

$$p_K(t+\Delta t) = p_K(t)(1-\Delta t \mu_K^{\mathrm{Sys}}) + p_{K-1}(t)\Delta t \lambda_{K-1}^{\mathrm{Sys}} + o(\Delta t) \ .$$

6.3 Stationäre Zustandswahrscheinlichkeiten

Termumformung in den drei obigen Gleichungen ergibt zunächst (für $n=1,\dots,K-1$ in der mittleren Gleichung):

$$\frac{p_{0}(t + \Delta t) - p_{0}(t)}{\Delta t} = -p_{0}(t)\lambda_{0}^{\text{Sys}} + p_{1}(t)\mu_{1}^{\text{Sys}} + \frac{o(\Delta t)}{\Delta t}, \qquad (6)$$

$$\frac{p_{n}(t + \Delta t) - p_{n}(t)}{\Delta t} = -p_{n}(t)(\lambda_{n}^{\text{Sys}} + \mu_{n}^{\text{Sys}}) + p_{n+1}(t)\mu_{n+1}^{\text{Sys}} + p_{n-1}(t)\lambda_{n-1}^{\text{Sys}} + \frac{o(\Delta t)}{\Delta t},
\frac{p_{K}(t + \Delta t) - p_{K}(t)}{\Delta t} = -p_{K}(t)\mu_{K}^{\text{Sys}} + p_{K-1}(t)\lambda_{K-1}^{\text{Sys}} + \frac{o(\Delta t)}{\Delta t}.$$

Wählt man nun für Δt eine immer kürzere Zeitspanne, d.h. führt man den Grenzübergang $\Delta t \to 0$ durch, so ergibt sich für $n=0,\ldots,K$ auf den linken Seiten der Gleichungen (6):

$$\lim_{\Delta t \to 0} \frac{p_n(t + \Delta t) - p_n(t)}{\Delta t} = p'_n(t) .$$

Des Weiteren gilt $\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0$ nach Definition von $o(\Delta t)$. Damit wird für $\Delta t \to 0$ aus (6):

$$p'_{0}(t) = -p_{0}(t)\lambda_{0}^{\text{Sys}} + p_{1}(t)\mu_{1}^{\text{Sys}},$$

$$p'_{n}(t) = -p_{n}(t)(\lambda_{n}^{\text{Sys}} + \mu_{n}^{\text{Sys}}) + p_{n+1}(t)\mu_{n+1}^{\text{Sys}} + p_{n-1}(t)\lambda_{n-1}^{\text{Sys}},$$

$$p'_{K}(t) = -p_{K}(t)\mu_{K}^{\text{Sys}} + p_{K-1}(t)\lambda_{K-1}^{\text{Sys}}.$$
(7)

Die Betrachtung eines Warteschlangensystems im stationären Zustand bedeutet, dass man das System nach sehr langer Laufzeit, also für $t \to \infty$ betrachtet. Der stationäre Zustand zeichnet sich in einem System, welches diesen erreicht, also in dem z.B. nicht kontinuierlich mehr Kunden eintreffen, als bedient werden können, dadurch aus, dass sich die Zustandswahrscheinlichkeiten stabilisieren, d.h. es gilt $p_n'(t) \to 0$ für alle $n = 0, \ldots, K$ für $t \to \infty$. Setzt man in (7) $p_n'(t) = 0$ so ergibt sich das folgende lineare Gleichungssystem für die stationären, d.h. zeitunabhängigen, Zustandswahrscheinlichkeiten $p_n := \lim_{t \to \infty} p_n(t)$:

$$p_{1} = \frac{\lambda_{0}^{\text{Sys}}}{\mu_{1}^{\text{Sys}}} p_{0} ,$$

$$p_{n+1} = \frac{\lambda_{n}^{\text{Sys}}}{\mu_{n+1}^{\text{Sys}}} p_{n} + \frac{\mu_{n}^{\text{Sys}}}{\mu_{n+1}^{\text{Sys}}} p_{n} - \frac{\lambda_{n-1}^{\text{Sys}}}{\mu_{n+1}^{\text{Sys}}} p_{n-1} ,$$

$$p_{K} = \frac{\lambda_{K-1}^{\text{Sys}}}{\mu_{K}^{\text{Sys}}} p_{K-1} .$$
(8)

Theorem 10. Mit den obigen Bezeichnungen gelten:

$$p_{n} = \prod_{i=1}^{n} \frac{\lambda_{i-1}^{\text{Sys}}}{\mu_{i}^{\text{Sys}}} p_{0} \quad \text{für } n = 1, \dots, K \quad und \quad p_{0} = \left[\sum_{n=1}^{K} \prod_{i=1}^{n} \frac{\lambda_{i-1}^{\text{Sys}}}{\mu_{i}^{\text{Sys}}} + 1 \right]^{-1} . \tag{9}$$

Beweis. Der Beweis für p_n , $n=1,\ldots,K$, erfolgt per Induktion. Der Induktionsanfang ergibt sich unmittelbar aus der ersten Gleichung von (8). Es sei nun angenommen, dass die Gleichung (9) für ein $n \in \mathbb{N}$ gelte. Dann ergibt sich aus der zweiten Gleichung von (8) für n+1:

$$p_{n+1} = \frac{\lambda_{n}^{\text{Sys}}}{\mu_{n+1}^{\text{Sys}}} p_{n} + \frac{\mu_{n}^{\text{Sys}}}{\mu_{n+1}^{\text{Sys}}} p_{n} - \frac{\lambda_{n-1}^{\text{Sys}}}{\mu_{n+1}^{\text{Sys}}} p_{n-1}$$

$$= \frac{\lambda_{n}^{\text{Sys}}}{\mu_{n+1}^{\text{Sys}}} \prod_{i=1}^{n} \frac{\lambda_{i-1}^{\text{Sys}}}{\mu_{i}^{\text{Sys}}} p_{0} + \frac{\mu_{n}^{\text{Sys}}}{\mu_{n+1}^{\text{Sys}}} \prod_{i=1}^{n} \frac{\lambda_{i-1}^{\text{Sys}}}{\mu_{i}^{\text{Sys}}} p_{0} - \frac{\lambda_{n-1}^{\text{Sys}}}{\mu_{n+1}^{\text{Sys}}} \prod_{i=1}^{n-1} \frac{\lambda_{i-1}^{\text{Sys}}}{\mu_{i}^{\text{Sys}}} p_{0}$$

$$= \prod_{i=1}^{n+1} \frac{\lambda_{i-1}^{\text{Sys}}}{\mu_{i}^{\text{Sys}}} p_{0} + \underbrace{\left[\frac{\mu_{n}^{\text{Sys}} \lambda_{n-1}^{\text{Sys}}}{\mu_{n+1}^{\text{Sys}}} - \frac{\lambda_{n-1}^{\text{Sys}}}{\mu_{n+1}^{\text{Sys}}}\right]}_{i=1} \prod_{i=1}^{n-1} \frac{\lambda_{i-1}^{\text{Sys}}}{\mu_{i}^{\text{Sys}}} p_{0}.$$

Das Warteschlangensystem muss sich stets in einem der Zustände $0, \ldots, K$ befinden, d.h. die Summe aller Wahrscheinlichkeiten muss 1 ergeben ($\sum_{n=0}^{K} p_n = 1$). Mit der Formel p_n ergibt sich:

$$p_0 + \sum_{n=1}^{K} \prod_{i=1}^{n} \frac{\lambda_{i-1}^{\text{Sys}}}{\mu_i^{\text{Sys}}} p_0 = 1.$$

Löst man diese Gleichung nach p_0 auf, so ergibt sich sofort die Aussage für p_0 .

6.4 Stationäre Zustandswahrscheinlichkeiten für die konkreten Ankunfts- und Bedienraten

Bisher wurden in den Formeln für die Zustandswahrscheinlichkeiten die Ankunftsrate bezogen auf das Gesamtsystem λ^{Sys} und die Bedienrate des Gesamtsystems μ^{Sys} verwendet. Im Folgenden sollen nun die Ankunftsraten der einzelnen Kunden λ bzw. die Bedienraten der einzelnen Bedieners μ verwendet werden. Die Größen λ^{Sys} und μ^{Sys} sind für drei Bereiche unterschiedlich definiert: für n=0, für $n=1,\ldots,c$ und für $n=c+1,\ldots,K$. Aus diesem Grund wird p_n im Folgenden auch für diese Bereiche getrennt berechnet:

Bestimmung von p_0

Setzt man die Zustandswahrscheinlichkeiten aus (4) und (5) für p_0 in (9) ein, so ergibt sich:

$$p_{0} = \left[\sum_{n=1}^{K} \prod_{i=1}^{n} \frac{\lambda_{i-1}^{\text{Sys}}}{\mu_{i}^{\text{Sys}}} + 1\right]^{-1} = \left[\sum_{n=1}^{K} \left(\lambda^{n} \cdot \frac{1}{\prod_{i=1}^{\min(n,c)} i\mu \cdot \prod_{i=\min(n,c)+1}^{n} c\mu}\right) + 1\right]^{-1}$$

$$= \left[\sum_{n=1}^{c} \frac{\lambda^{n}}{\mu^{n} n!} + \sum_{n=c+1}^{K} \frac{\lambda^{n}}{\mu^{c} c! \cdot (c\mu^{n-c})} + 1\right]^{-1} = \left[\sum_{n=1}^{c} \frac{\lambda^{n}}{\mu^{n} n!} + \sum_{n=c+1}^{K} \frac{\lambda^{n}}{\mu^{n} c! c^{n-c}} + 1\right]^{-1}.$$

Mit $a := \frac{\lambda}{\mu}$ gilt weiter:

$$p_0 = \left[\sum_{n=1}^c \frac{a^n}{n!} + \sum_{n=c+1}^K \frac{a^n}{c!c^{n-c}} + 1\right]^{-1}.$$

Mit der Wahl von

$$C_n := \begin{cases} \frac{a^n}{n!} & \text{für } n \le c, \\ \frac{a^n}{c!c^{n-c}} & \text{für } c < n \le K \end{cases}$$
 (10)

gilt schließlich

$$p_0 = \left[\sum_{n=0}^K C_n\right]^{-1} .$$

Bestimmung von p_n für $n = 1, \ldots, c$

Setzt man die Zustandswahrscheinlichkeiten aus (4) und (5) für p_n in (9) ein, so ergibt sich mit der Definition von C_n in (10) für $n = 1, \ldots, c$:

$$p_n = \prod_{i=1}^n \frac{\lambda_{i-1}^{\text{Sys}}}{\mu_i^{\text{Sys}}} \cdot p_0 = \prod_{i=1}^n \frac{\lambda}{n\mu} \cdot p_0 = \frac{a^n}{n!} \cdot p_0 = C_n p_0$$
.

Bestimmung von p_n für $n = c + 1, \dots, K$

Setzt man die Zustandswahrscheinlichkeiten aus (4) und (5) für p_n in (9) ein, so ergibt sich mit der Definition von C_n in (10) für $n=c+1,\ldots,K$:

$$p_{n} = \prod_{i=1}^{n} \frac{\lambda_{i-1}^{\text{Sys}}}{\mu_{i}^{\text{Sys}}} \cdot p_{0} = \prod_{i=1}^{c} \frac{\lambda}{n\mu} \cdot \prod_{i=c+1}^{n} \frac{\lambda}{c\mu} \cdot p_{0} = \frac{a^{n}}{c!c^{n-c}} \cdot p_{0} = C_{n}p_{0} .$$

Die obigen Überlegungen für p_n lassen sich zu folgendem Satz zusammenfassen:

Theorem 11. Mit der Wahl von C_n gemäß (10) gilt für die Zustandswahrscheinlichkeiten p_n in einem Geburts- und Todesprozess mit Ankunfts- und Bedienraten gemäß (4) und (5):

$$p_n = \begin{cases} \left[\sum_{i=0}^K C_i \right]^{-1} & \text{für } n = 0, \\ C_n p_0 & \text{für } n > 0. \end{cases}$$

$$(11)$$

7 Die Erlang-C-Formel

Mit der Wahl von $K:=\infty$ folgt zunächst für C_n (vgl. Formel (10)):

$$C_n := \begin{cases} \frac{a^n}{n!} & \text{für } n < c ,\\ \frac{a^n}{c!c^{n-c}} & \text{für } n \ge c . \end{cases}$$

Die Wahrscheinlichkeit dafür, dass keine Kunden im System sind, p_0 berechnet sich in diesem Fall aus

$$1 = p_0 + \sum_{n=1}^{\infty} C_n p_0 = p_0 \left(1 + \sum_{n=1}^{\infty} C_n \right)$$

zu:

$$p_0 = \left[1 + \sum_{n=1}^{\infty} C_n\right]^{-1} = \left[\sum_{n=0}^{c-1} \frac{a^n}{n!} + \sum_{n=c}^{\infty} \frac{a^n}{c!c^{n-c}}\right]^{-1} = \left[\sum_{n=0}^{c-1} \frac{a^n}{n!} + \frac{a^c}{c!} \sum_{n=0}^{\infty} \frac{a^n}{c^n}\right]^{-1}$$

$$= \left[\sum_{n=0}^{c-1} \frac{a^n}{n!} + \frac{a^c}{c!} \cdot \frac{1}{1 - \frac{a}{c}}\right]^{-1} = \left[\sum_{n=0}^{c-1} \frac{a^n}{n!} + \frac{a^c \cdot c}{c!(c-a)}\right]^{-1}.$$

Bei der Umformung

$$\sum_{n=0}^{\infty} \left(\frac{a}{c}\right)^n = \frac{1}{1 - \frac{a}{c}}$$

im vorletzten Rechenschritt wurde dabei die *geometrische Reihe* verwendet und die Voraussetzung ausgenutzt, dass im Fall eines Warteschlangenmodells ohne Abbrecher $\lambda < c\mu$ gelten muss, also $\frac{a}{c} < 1$ sein muss, was für die Konvergenz der Reihe notwendig ist.

Auf dieser Basis lässt sich nun auch für den Fall eines Erlang-C-Modells die Wartezeitverteilung $P(W \le t)$ bestimmen:

$$\begin{split} P(W \leq t) &= \sum_{n=0}^{c-1} C_n p_0 + \sum_{n=c}^{\infty} F_{c\mu,n-c+1}(t) C_n p_0 \\ &= \sum_{n=0}^{c-1} C_n p_0 + \sum_{n=c}^{\infty} \int_0^t \frac{(c\mu)^{n-c+1} x^{n-c}}{(n-c)!} \mathrm{e}^{-c\mu x} \, \mathrm{d}x C_n p_0 \\ &= \sum_{n=0}^{c-1} C_n p_0 + \int_0^t \sum_{n=c}^{\infty} \frac{(c\mu)^{n-c+1} x^{n-c}}{(n-c)!} \mathrm{e}^{-c\mu x} \, \mathrm{d}x C_n p_0 \\ &= \sum_{n=0}^{c-1} C_n p_0 + p_0 c\mu \int_0^t \mathrm{e}^{-c\mu x} \sum_{n=0}^{\infty} \frac{(c\mu x)^n}{n!} \cdot \frac{a^{n+c}}{c!c^n} \, \mathrm{d}x \\ &= \sum_{n=0}^{c-1} C_n p_0 + p_0 \frac{a^c c\mu}{c!} \int_0^t \mathrm{e}^{-c\mu x} \sum_{n=0}^{\infty} \frac{(\mu x a)^n}{n!} \, \mathrm{d}x \end{split}$$

$$= \sum_{n=0}^{c-1} C_n p_0 + p_0 \frac{a^c c \mu}{c!} \int_0^t e^{-c\mu x} \cdot e^{a\mu x} dx$$

$$= \sum_{n=0}^{c-1} C_n p_0 + p_0 \frac{a^c c \mu}{c!} \left[-\frac{1}{(c-a)\mu} e^{-(c-a)\mu x} \right]_{x=0}^{x=t}$$

$$= \sum_{n=0}^{c-1} C_n p_0 - p_0 \frac{a^c c}{c!(c-a)} \left(e^{-(c-a)\mu t} - 1 \right)$$

$$= p_0 \left(\sum_{n=0}^{c-1} \frac{a^n}{n!} + \frac{a^c \cdot c}{c!(c-a)} \right) - p_0 \frac{a^c c}{c!(c-a)} e^{-(c-a)\mu t}$$

$$= 1 - p_0 \frac{a^c c}{c!(c-a)} e^{-(c-a)\mu t} .$$

Mit der Definition

$$P_1 := p_0 \frac{a^c c}{c!(c-a)} = \frac{\frac{a^c c}{c!(c-a)}}{\sum_{n=0}^{c-1} \frac{a^n}{n!} + \frac{a^c \cdot c}{c!(c-a)}}$$
(12)

ergibt sich daraus schließlich die übliche Erlang-C-Formel:

$$P(W \le t) = 1 - P_1 e^{-(c-a)\mu t}$$

7.1 Weitere Kenngrößen

Im Falle eines M/M/c-Systems lassen sich die weiteren Kenngrößen unmittelbar aus den bereits bestimmten Größen p_n und C_n ableiten:

• Mittlere Warteschlangenlänge: Mit der Definition von P_1 aus (12) gilt zunächst für n>c:

$$p_n = C_n p_0 = \frac{a^n}{c!c^{n-c}} p_0 = P_1 \rho^{n-c} (1-\rho).$$

Damit ergibt sich für die mittlere Warteschlangenlänge unmittelbar mit der Definition des Erwartungswerts:

$$\mathbf{E}[N_Q] = \sum_{n=c+1}^{\infty} (n-c)p_n = P_1(1-\rho) \sum_{n=c+1}^{\infty} (n-c)\rho^{n-c}$$

$$= P_1(1-\rho) \sum_{n=1}^{\infty} n\rho^n = P_1(1-\rho) \sum_{n=0}^{\infty} n\rho^n = P_1(1-\rho)\rho \sum_{n=1}^{\infty} n\rho^{n-1}.$$

Da für $f(\rho) := \rho^n$ gilt $f'(\rho) = n\rho^{n-1}$, gilt zunächst weiter:

$$\mathbf{E}[N_Q] = P_1(1-\rho)\rho \sum_{n=1}^{\infty} [\rho^n]'$$
.

Aufgrund der absoluten Konvergenz der Reihe bedingt durch die Tatsache, dass $\rho < 1$ ist, können Grenzwertbildung und Differentiation vertauscht werden. Mit der Anwendung der geometrischen Reihe $(\sum_{n=0}^{\infty} a^n = \frac{1}{1-a}$ für |a| < 1) gilt:

$$\mathbf{E}[N_Q] = P_1(1-\rho)\rho \left[\sum_{n=1}^{\infty} \rho^n\right]' = P_1(1-\rho)\rho \left[\frac{1}{1-\rho}\right]' = P_1(1-\rho)\rho \frac{0-(-1)}{(1-\rho)^2}$$
$$= P_1\frac{\rho}{1-\rho} = P_1\frac{a}{c-a}.$$

• Mittlere Anzahl an Kunden im System:

Die mittlere Anzahl an Kunden im System ergibt sich als Summe aus der mittleren Warteschlangenlänge $\mathbf{E}[N_Q]$ und der mittleren Anzahl an Kunden in Bedienung a:

$$\mathbf{E}[N] = P_1 \frac{a}{c-a} + a .$$

Die Tatsache, dass die Arbeitslast a gerade der mittleren Anzahl an Kunden in Bedienung entspricht, ergibt sich dabei als Konsequenz aus der Formel von Little.

• Mittlere Wartezeit:

Mit Hilfe der Formel von Little ergibt sich aus $\mathbf{E}[N_Q]$ direkt die mittlere Wartezeit:

$$\mathbf{E}[W] = \frac{1}{\lambda} \mathbf{E}[N_Q] = P_1 \frac{1}{\lambda} \frac{a}{c-a} = P_1 \frac{\frac{1}{\mu}}{c-a} = P_1 \frac{1}{c\mu - \lambda}.$$

• Mittlere Verweilzeit:

Die mittlere Verweilzeit setzt sich aus der mittleren Wartezeit $\mathbf{E}[W]$ und der mittleren Bediendauer $\mathbf{E}[S] = \frac{1}{\mu}$ zusammen:

$$\mathbf{E}[V] = P_1 \frac{1}{c\mu - \lambda} + \frac{1}{\mu} .$$

Zusammenfassend:

$$\begin{split} \mathbf{E}[N_Q] &= P_1 \frac{a}{c - a} \,, \\ \mathbf{E}[N] &= P_1 \frac{a}{c - a} + a \,, \\ \mathbf{E}[W] &= P_1 \frac{1}{c\mu - \lambda} \,, \\ \mathbf{E}[V] &= P_1 \frac{1}{c\mu - \lambda} + \frac{1}{\mu} \,. \end{split}$$