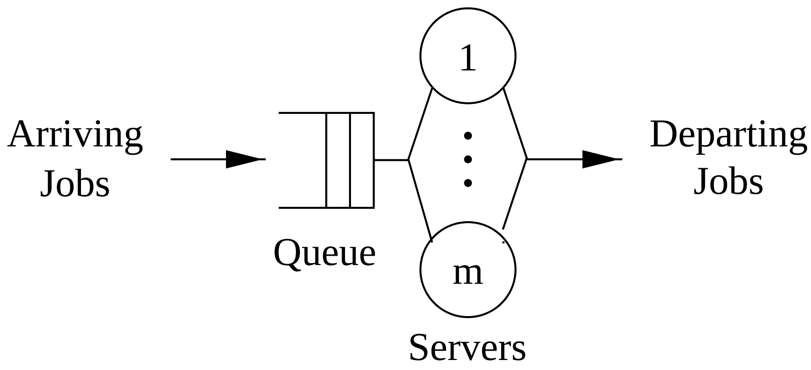
D Einfache Wartesysteme

D.1 Beschreibung (Kendall'sche Notation)



- **Parameter**:
  - Ankunftsrate: $\lambda = 1/\overline{T}_A$ $\quad \overline{T}_A$ = Zwischenankunftszeit
  - Bedienrate: $\mu = 1/\overline{T}_B$ $\quad \overline{T}_B$ = Bedienzeit
  - Anzahl der Bedieneinheiten: *m*

Kendall’sche Notation:

A/B/m - Warteschlangendisziplin

A Verteilung der Zwischenankunftszeiten

B Verteilung der Bedienzeiten

m Anzahl der identischen Bedieneinheiten

- Für ***A*** und ***B*** werden traditionell folgende Symbole verwendet:
  - M Exponentialverteilung (Markov-Eigenschaft)
  - E_k Erlangverteilung mit k Phasen
  - H_k Hyperexponentialverteilung mit k Phasen
  - D Deterministische Verteilung (T_A bzw. T_B sind konstant)
  - G Allgemeine Verteilung
  - GI Allgemeine unabhängige Verteilung

- **Warteschlangendisziplin** (Bedienstrategie) legt fest, welcher Auftrag aus der Warteschlange (WS) als nächster bedient wird:
  - *FCFS* (First-Come-First-Served):
    Die Aufträge werden in der Reihenfolge der Ankunft bedient
    Ist keine Disziplin angegeben so gilt FCFS.
  - *LCFS* (Last-Come-First-Served):
    Der zuletzt angekommene Auftrag wird als nächster bedient.
  - *SIRO* (Service-In-Random-Order):
    Die Auswahl erfolgt zufällig.
  - *RR* (Round Robin):
    Ist die Bedienung eines Auftrages nach einer fest vorgegebenen Zeitscheibe noch nicht beendet, so wird er verdrängt und wieder in die WS eingereiht, die nach FCFS abgearbeitet wird. Dies wiederholt sich so oft bis der Auftrag vollständig bedient ist.
  - *PS* (Processor Sharing):
    Entspricht RR mit infinitesimal kleiner Zeitscheibe. Dadurch entsteht der Eindruck, als ob alle Aufträge gleichzeitig bedient würden mit entsprechend längerer Bedienzeit.

- *IS* (Infinite Server):
  Es sind immer mindestens so viele Bedieneinheiten vorhanden wie Aufträge im System, so daß keine WS entsteht.
- *Statische Prioritäten:*
  Die Auswahl erfolgt nach fest vorgegebenen Prioritäten der Aufträge. Haben mehrere Aufträge dieselbe Priorität so erfolgt die Auswahl unter diesen nach FCFS.
- *Dynamische (zeitabhängige) Prioritäten:*
  Die Auswahl erfolgt nach dynamische Prioritäten, die sich in Abhängigkeit von der Zeit ändern.
- *Verdrängung (preemption):*
  Bei den Prioritätsdisziplinen und LCFS kann eine Verdrängung des gerade in Bedienung befindlichen Auftrags erfolgen, wenn ein neu in der WS ankommender Auftrag höhere Priorität hat.

◆ **Beispiel** zur Kendall’schen Notation:

M/G/1 - LCFS Preemptive Resume (PR)

- Zwischenankunftszeit ist **exponentiell** verteilt
- Bedienzeit hat eine **beliebige** Verteilung (gegeben durch eine Verteilungsfunktion oder durch Mittelwert und Varianz)
- Anzahl der Bedieneinheiten ist ***m = 1***.
- Die Warteschlangendisziplin ist **LCFS**
- **Preemptive** bedeutet, daß ein neu ankommender Auftrag den in Bedienung befindlichen Auftrag verdrängt.
- **Resume** bedeutet, daß der verdrängte Auftrag ab der Stelle weiter-bearbeitet wird an der er verdrängt wurde, wenn er wieder die Bedieneinheit erhält.
  (Repeat würde bedeuten, daß der Auftrag nach einer Verdrängung wieder von Anfang an beabeitet werden muß).

D.2 Verteilungsfunktionen

Exponentialverteilung

- Verteilungsfunktion:

$$F_X(x) = \begin{cases} 1 - \exp\left(-\frac{x}{\overline{X}}\right), & 0 \le x < \infty, \\ 0, & \text{otherwise.} \end{cases}$$

$$\text{with } \overline{X} = \begin{cases} \frac{1}{\lambda}, & \text{if } X \text{ represents interarrival times}, \\ \frac{1}{\mu}, & \text{if } X \text{ represents service times.} \end{cases}$$

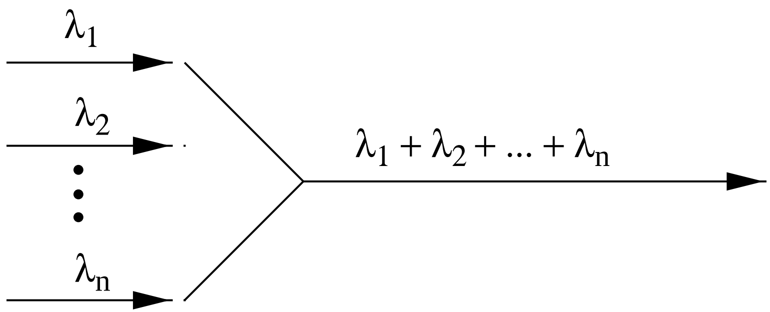
◆ Kenngrößen:

$$
\begin{array}{ll}
\text{Dichte:} & f_X(x) = \lambda e^{-\lambda x}\,, \\
\text{Mittelwert:} & \overline{X} = \dfrac{1}{\lambda}\,, \\
\text{Varianz:} & \mathrm{var}(X) = \dfrac{1}{\lambda^2}\,, \\
\text{Variationskoeffizient:} & c_X = 1\,.
\end{array}
$$

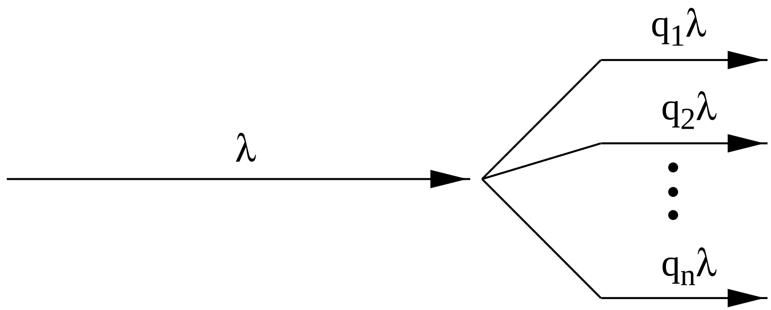
◆ Eigenschaft der Gedächtnislosigkeit
(Markoveigenschaft, memoryless property)

$$P(X \leq u + t \mid X > u) = 1 - \exp\left(-\frac{t}{\overline{X}}\right) = P(X \leq t)$$

◆ Verschmelzung:

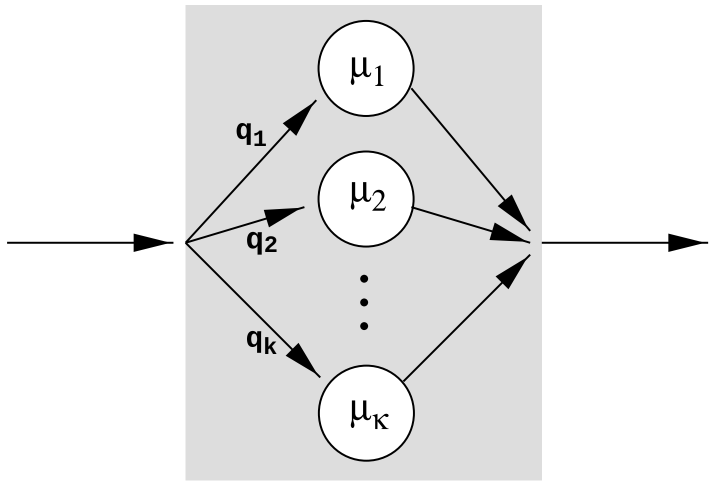


◆ Aufspaltung:



Hyperexponentialverteilung, H_k

◆ Modellvorstellung:



◆ Verteilungsfunktion:

$$F_X(x) = \sum_{j=1}^{k} q_j (1 - e^{-\mu_j x}), \qquad x \geq 0$$

◆ Kenngrößen:

Dichte: $$f_X(x) = \sum_{j=1}^{k} q_j \mu_j e^{-\mu_j x}, \quad x > 0,$$

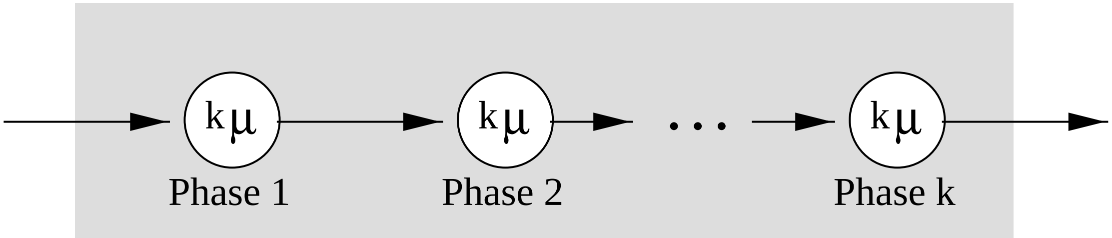
Mittelwert: $$\overline{X} = \sum_{j=1}^{k} \frac{q_j}{\mu_j} = \frac{1}{\mu}, \quad x > 0,$$

Varianz: $$\mathrm{var}(X) = 2\sum_{j=1}^{k} \frac{q_j}{\mu_j^2} - \frac{1}{\mu^2},$$

Variationskoeffizient: $$c_X = \sqrt{2\mu^2 \sum_{j=1}^{k} \frac{q_j}{\mu_j^2} - 1} \quad \geq 1.$$

Erlang-k-Verteilung, E_k

◆ Modellvorstellung:



◆ Verteilungsfunktion:

$$F_X(x) = 1 - e^{-k\mu x} \cdot \sum_{j=0}^{k-1} \frac{(k\mu x)^j}{j!}, \quad x \geq 0,\ k = 1, 2 \ldots .$$

◆ Kenngrößen:

Dichte: $f_X(x) = \frac{k\mu(k\mu x)^{k-1}}{(k-1)!} e^{-k\mu x},\ x > 0,\ k = 1, 2, \ldots\ ,$

Mittelwert: $\overline{X} = \frac{1}{\mu}\,,$

Varianz: $\mathrm{var}(X) = \frac{1}{k\mu^2}\,,$

Variationskoeffizient: $c_X = \frac{1}{\sqrt{k}} \leq 1\,.$

Hypoexponentialverteilung:

◆ Verteilungsfunktion:

$$F_X(x) = 1 - \frac{\mu_2}{\mu_2 - \mu_1} e^{-\mu_1 x} + \frac{\mu_1}{\mu_2 - \mu_1} e^{-\mu_2 x}, \qquad x \geq 0.$$

◆ Kenngrößen:

Dichte: $$f_X(x) = \frac{\mu_1 \mu_2}{\mu_1 - \mu_2}(e^{-\mu_2 x} - e^{-\mu_1 x}), \quad x > 0\,,$$

Mittelwert: $$\overline{X} = \frac{1}{\mu_1} + \frac{1}{\mu_2}\,,$$

Varianz: $$\mathrm{var}(X) = \frac{1}{\mu_1^2} + \frac{1}{\mu_2^2}\,,$$

Variationskoeffizient: $$c_X = \frac{\sqrt{\mu_1^2 + \mu_2^2}}{\mu_1 + \mu_2} < 1\ .$$

Gammaverteilung:

◆ Verteilungsfunktion:

$$F_X(x) = \int_0^x \frac{\alpha\mu \cdot (\alpha\mu u)^{\alpha-1}}{\Gamma(\alpha)} \cdot e^{-\alpha\mu u} du, \quad x \geq 0, \ \alpha > 0,$$

$$\text{with } \Gamma(\alpha) = \int_0^\infty u^{\alpha-1} \cdot e^{-u} du, \quad \alpha > 0.$$

◆ Kenngrößen:

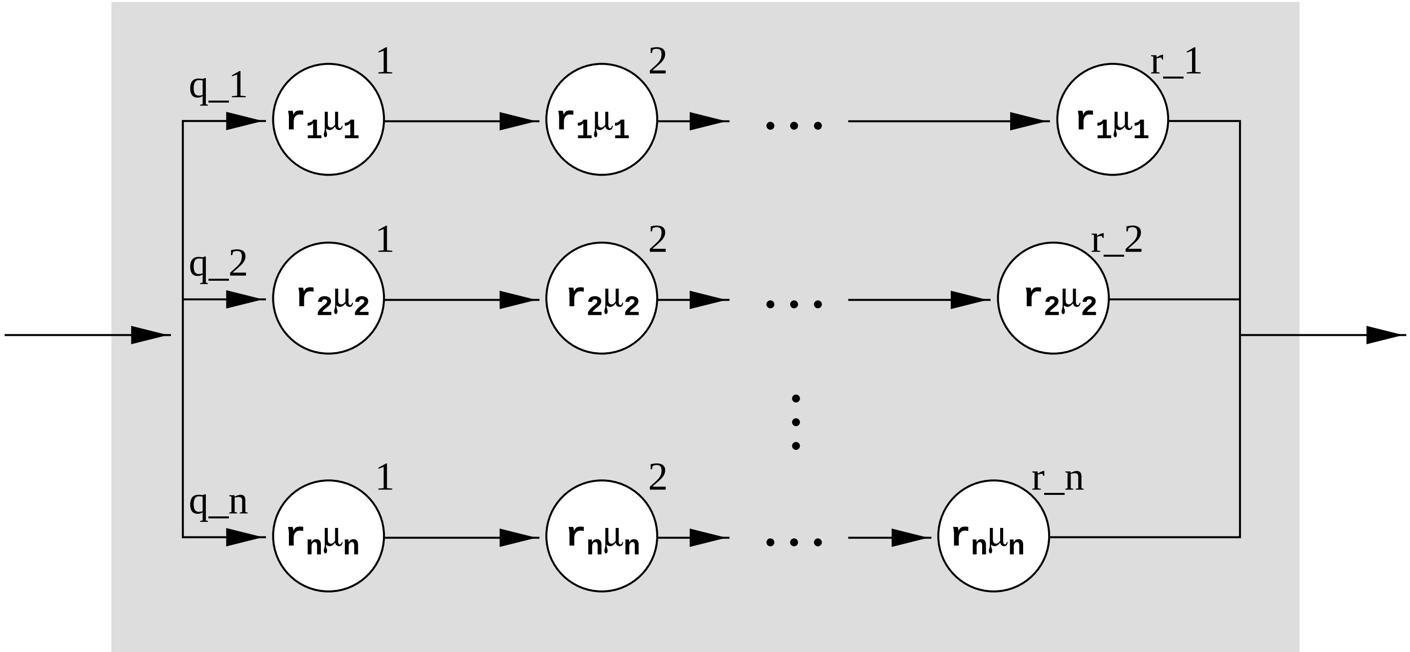
Dichte: $$f_X(x) = \frac{\alpha\mu\cdot(\alpha\mu x)^{\alpha-1}}{\Gamma(\alpha)} e^{-\alpha\mu x}, \quad x > 0,\ \alpha > 0\,,$$

Mittelwert: $$\overline{X} = \frac{1}{\mu}\,,$$

Varianz: $$\mathbf{var}(X) = \frac{1}{\alpha\mu^2}\,,$$

Variationskoeffizient: $$c_X = \frac{1}{\sqrt{\alpha}}\,.$$

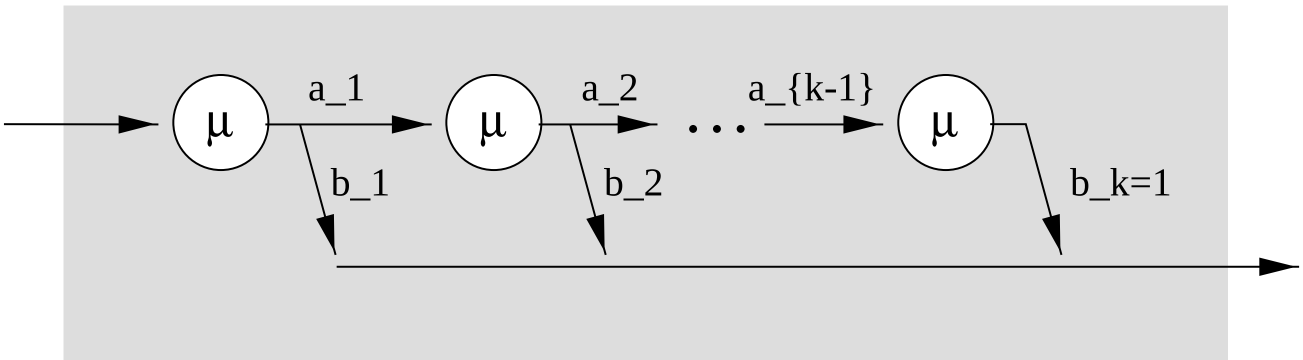
Verallgemeinerte Erlangverteilung:



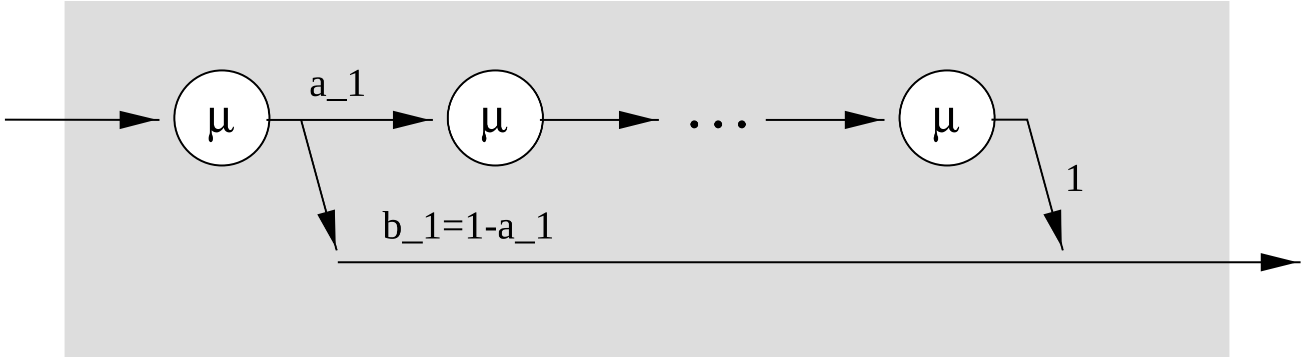
Dichte: $$f_X(x) = \sum_{j=1}^{n} q_j \cdot \frac{r_j\mu_j(r_j\mu_j x)^{r_j-1}}{(r_j-1)!} \cdot e^{-r_j\mu_j x}, \qquad x \geq 0.$$

Coxverteilung, C_k

◆ Modellvorstellung:



$C_X < 1$:



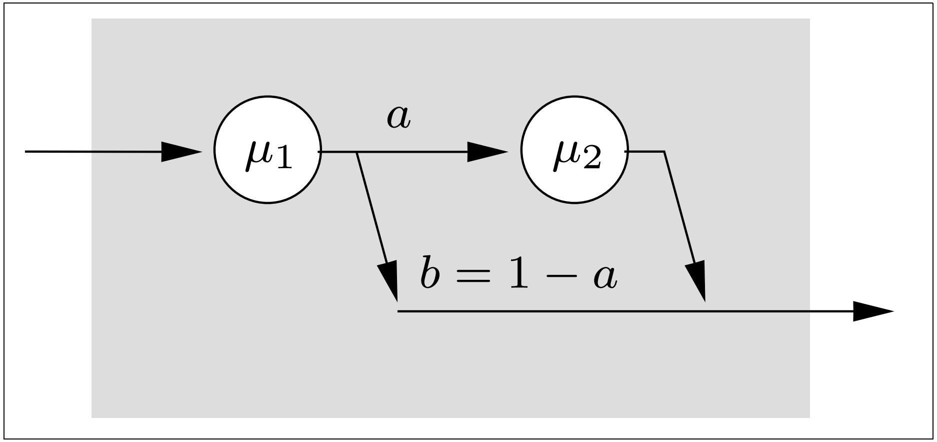
◆ Kenngrößen:

Mittelwert: $$\overline{X} = \frac{b_1 + k(1 - b_1)}{\mu},$$

Varianz: $$\operatorname{var}(X) = \frac{k + b_1(k-1)\,(b_1(1-k) + k - 2)}{\mu^2},$$

Quadrierter Variationskoeffizient: $$c_X^2 = \frac{k + b_1(k-1)\,(b_1(1-k) + k - 2)}{[b_1 + k(1 - b_1)]^2}.$$

$C_X > 1$:



- Kenngrößen:

Mittelwert: $$\overline{X} = \frac{1}{\mu_1} + \frac{a}{\mu_2},$$

Varianz: $$\mathrm{var}\,(X) = \frac{\mu_2^2 + a\mu_1^2(2-a)}{\mu_1^2 \cdot \mu_2^2},$$

Quadrierter Variationskoeffizient: $$c_X^2 = \frac{\mu_2^2 + a\mu_1^2(2-a)}{(\mu_2 + a\mu_1)^2}.$$

Weibull-Verteilung:

$$F_X(x) = 1 - \exp(-(\lambda x)^\alpha)\,, \quad x \ge 0\,.$$

- Kenngrößen:

Dichte: $$f_X(x) = \alpha\lambda(\lambda x)^{\alpha-1}\exp(-(\lambda x)^\alpha)\,,\ \ \lambda > 0\,,$$

Mittelwert: $$\overline{X} = \frac{1}{\lambda}\Gamma\left(1 + \frac{1}{\alpha}\right)\,,$$

Quadr. Variationskoeffizient: $$c_X^2 = \frac{\Gamma(1+2/\alpha)}{\{\Gamma(1+1/\alpha)\}^2} - 1\,.$$

- Mittelwert, Varianz und Variationskoeffizient wichtiger Verteilungen:

Distribution	Parameter	$E[X]$	$\mathrm{var}(X)$	c_X
Exponential	μ	$\frac{1}{\mu}$	$\frac{1}{\mu^2}$	1
Erlang	μ, k $k=1,2,\ldots$	$\frac{1}{\mu}$	$\frac{1}{k\mu^2}$	$\frac{1}{\sqrt{k}} \leq 1$
Gamma	μ, α $(0<\alpha<\infty)$	$\frac{\alpha}{\mu}$	$\frac{\alpha}{\mu^2}$	$0 < \frac{1}{\sqrt{\alpha}} < \infty$
Hypoexponential	μ_1, μ_2	$\frac{1}{\mu_1} + \frac{1}{\mu_2}$	$\frac{1}{\mu_1^2} + \frac{1}{\mu_2^2}$	$\frac{\sqrt{\mu_1^2 + \mu_2^2}}{\mu_1 + \mu_2} < 1$
Hyperexponential	k, μ_i, q_i	$\sum_{i=1}^{k} \frac{q_i}{\mu_i} = \frac{1}{\mu}$	$2\sum_{i=1}^{k} \frac{q_i}{\mu_i^2} - \frac{1}{\mu^2}$	$\sqrt{2\mu^2 \sum_{i=1}^{k} \frac{q_i}{\mu_i^2} - 1} > 1$

Formeln zur Bestimmung der Parameter wichtiger Verteilungen

Distribution	Parameter	Calculation of the Parameters
Exponential	μ	$\mu = 1/\overline{X}$
Erlang	μ, k $k=1,2,\ldots$	$k = \text{ceil}(1/c_X^2)$ $\mu = 1/(c_X^2 \cdot k\overline{X})$
Gamma	μ, α $0<\alpha<\infty$	$\alpha = 1/c_X^2$ $\mu = 1/\overline{X}$
Hypoexponential	μ_1, μ_2	$\mu_{1/2} = \frac{2}{\overline{X}}\left[1 \pm \sqrt{1+2(c_X^2-1)}\right]^{-1}$
Hyperexponential (H_2)	μ_1, μ_2, q_1, q_2	$\mu_1 = \frac{1}{\overline{X}}\left[1 - \sqrt{\frac{q_2}{q_1}\frac{c_X^2-1}{2}}\right]^{-1}$ $\mu_2 = \frac{1}{\overline{X}}\left[1 + \sqrt{\frac{q_1}{q_2}\frac{c_X^2-1}{2}}\right]^{-1}$ $q_1 + q_2 = 1, \mu_2 > 0$

Distribution	Parameter	Calculation of the Parameters
Cox $(c_X \leq 1)$	k, b_i, μ_i	$k = \mathrm{ceil}(1/c_X^2)$ $b_1 = \dfrac{2kc_X^2 + k - 2 - \sqrt{k^2 + 4 - 4kc_X^2}}{2(c_X^2 + 1)(k-1)}$ $b_2 = b_3 = \ldots = b_{k-1} = 0, \quad b_k = 1$ $\mu_1 = \mu_2 = \ldots = \mu_k = \dfrac{k - b_1 \cdot (k-1)}{\overline{X}}$
Cox $(c_X > 1)$	k, b, μ_1, μ_2	$k = 2$ $b = c_X^2 \left[1 - \sqrt{1 - \dfrac{2}{1 + c_X^2}}\right]$ $\mu_{1/2} = \dfrac{1}{\overline{X}} \left[1 \pm \sqrt{1 - \dfrac{2}{1 + c_X^2}}\right]$

◆ Abschätzung von Mittelwert und Variationskoeffizient aus Stichproben X_i:

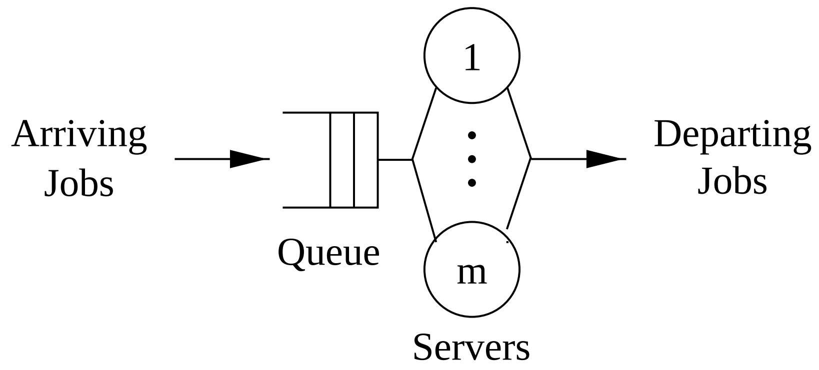
Mittelwert: $$\overline{X} = \frac{1}{k}\sum_{i=1}^{k} X_i$$

2. Moment: $$\overline{X^2} = \frac{1}{k}\sum_{i=1}^{k} X_i^2$$

Varianz: $$\sigma_X^2 = \mathrm{var}(X) = \overline{(X-\overline{X})^2} = \overline{X^2} - \overline{X}^2 ,$$

Variationskoeffizient: $$c_X = \frac{\sigma_X}{\overline{X}}$$

D.3 Leistungsgrößen



- Zustandswahrscheinlichkeit π_k

$$\pi_k = P[\text{there are } k \text{ jobs in the system}]\,.$$

Auslastung ρ:

- Eine Bedieneinheit (single server) *m = 1*:

$$\rho = \frac{\textbf{mittlere Bedienzeit}}{\textbf{mittlere Zwischenankunftszeit}} = \frac{\textbf{Ankunftsrate}}{\textbf{Bedienrate}} = \frac{\lambda}{\mu}\,.$$

- Mehrere Bedieneinheiten (multiple server) *m > 1*:

  mit **m**μ = Gesamtbedienrate der *m* Bedieneinheiten

$$\rho = \frac{\lambda}{m\mu}\,,$$

- Stabilitätsbedingung:

$$\rho < 1 \quad \longrightarrow \quad \lambda < m\mu$$

- Durchsatz λ:
  - ➤ Anzahl der pro Zeiteinheit bedienten Aufträge (Abgangsrate)
  - ➤ Im stabilen Fall ($\rho < 1$) gilt: Zugangsrate = Abgangsrate und damit:

$$\boldsymbol{\lambda = m \cdot \rho \cdot \mu}$$

- Antwortzeit T:
  - ➤ Zeit die ein Auftrag im Gesamtsystem (WS und Bedieneinheiten) verbringt
  - ➤ auch: Verweilzeit oder Systemzeit

- Wartezeit W:

  - ➤ Zeit die ein Auftrag in der WS verbringt
  - ➤ Es gilt:

    Antwortzeit = Wartezeit + Bedienzeit

  - ➤ Und für die Mittelwerte:

$$\overline{T} = \overline{W} + \frac{1}{\mu}$$

- Warteschlangenlänge Q:

  - ➤ Anzahl der Aufträge in der WS

- Anzahl der Aufträge K:
  - Gesamtanzahl der Aufträge im Wartesystem (WS und Bedieneinheiten)
  - Mittlere Anzahl der Aufträge:

$$\overline{K} = \sum_{k=1}^{\infty} k \cdot \pi_k \,.$$

- Little's Gesetz:
  - Grundlegendes Gesetz der Warteschlangentheorie
  - Zusammenhang zwischen der mittleren Aufenthaltszeit in einem System und der mittleren Anzahl der Aufträge in dem System, z.B.:

$$\overline{K} = \lambda \overline{T} \,,$$

$$\overline{Q} = \lambda \overline{W} \,.$$

  - Wenn eine der Größen $\overline{K}$, $\overline{Q}$, $\overline{T}$ und $\overline{W}$ bekannt ist, können die anderen berechnet werden (Bei bekanntem λ)

Zusammenstellung wichtiger Formeln:

Auslastung:

$$\rho = \frac{\lambda}{m\mu} ;$$

Little's Gesetz:

$$\overline{K} = \lambda \overline{T} ,$$

$$\overline{Q} = \lambda \overline{W}$$

Mittl. Antwortzeit:

$$\overline{T} = \overline{W} + \frac{1}{\mu}$$

Mittl. Anzahl der Jobs:

$$\overline{K} = \overline{Q} + m\rho$$

Mittl. Anzahl der Jobs:

$$\overline{K} = \sum_{k=1}^{\infty} k \cdot \pi_k$$