Outline

Source: Stewart, W.J. (2009): Probability, Markov Chains, Queues, and Simulation

11.2 Birth-Death Processes: The $M/M/1$ Queue

Birth-death processes as continuous-time Markov chains with a special structure:

- States indexed by the integers 0,1,2,...
- Transitions permitted only from state $i > 0$ to states $i - 1$ and $i + 1$
- Arrival to the queueing system $\rightarrow$ **birth**
- Departure $\rightarrow$ **death**

Source: Stewart, W.J. (2009): Probability, Markov Chains, Queues, and Simulation

11.2 Birth-Death Processes: The $M/M/1$ Queue

The simplest of all queueing systems $\rightarrow$ $M/M/1$ queue:

- FCFS scheduling discipline
- Poisson process for arrivals
- Service time exponentially distributed

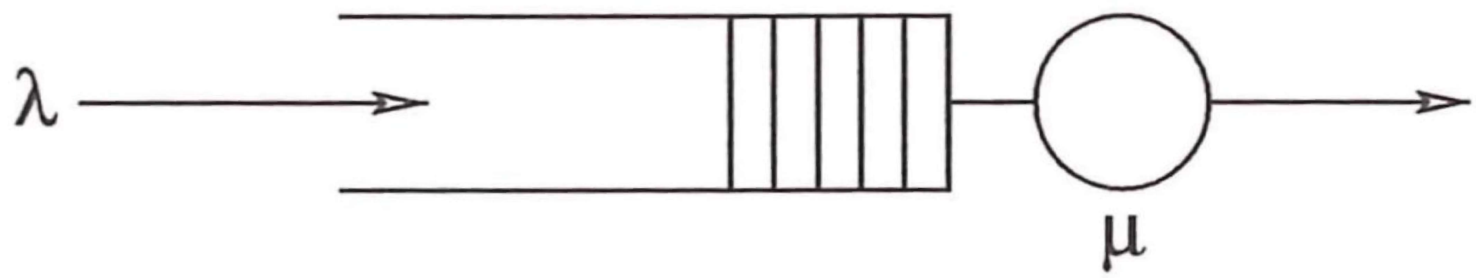


Figure 11.10. The *M/M/1* queue.

Source: Stewart, W.J. (2009): Probability, Markov Chains, Queues, and Simulation

$M/M/1$ Queue: Description and Steady-State Solution

State of the $M/M/1$ queue:

- State completely characterized by specifying the number of customers present
- n: State with n customers in the system (including the one in service)

Object of investigation:

- **State probabilities**:

  $\rightarrow$ Probability the system is in any given state n at any time t:

$$p_n(t) = \text{Prob}\{n \text{ in system at time } t\}$$

Source: Stewart, W.J. (2009): Probability, Markov Chains, Queues, and Simulation

$M/M/1$ Queue: Description and Steady-State Solution

Steady-state probabilities:

$$p_n = \lim_{t \to \infty} p_n(t)$$

If the limit exists, the probability of finding n customers present does not change over time.

Source: Stewart, W.J. (2009): Probability, Markov Chains, Queues, and Simulation

Probability the System is in State n at Time $t + \Delta t$

Goal:

$\rightarrow$ Computing the respective state probabilities

The system will be in state n at time $t + \Delta t$ if one of the following events occur:

1. The system is in state n at time t and no change occurs in $(t, t + \Delta t]$
. The system is in state $n - 1$ at time t and an arrival occurs in $(t, t + \Delta t]$
3. The system is in state $n + 1$ at time t and a departure occurs in $(t, t + \Delta t]$

Source: Stewart, W.J. (2009): Probability, Markov Chains, Queues, and Simulation

Probability the System is in State n at Time $t + \Delta t$

For sufficiently small values of Δt:

$$\begin{aligned}
\text{Prob}\{1 \text{ arrival in } (t, t + \Delta t]\} &= \lambda \Delta t + o(\Delta t), \\
\text{Prob}\{1 \text{ departure in } (t, t + \Delta t]\} &= \mu \Delta t + o(\Delta t), \\
\text{Prob}\{0 \text{ arrivals in } (t, t + \Delta t]\} &= 1 - \lambda \Delta t + o(\Delta t), \\
\text{Prob}\{0 \text{ departures in } (t, t + \Delta t]\} &= 1 - \mu \Delta t + o(\Delta t).
\end{aligned}$$

$\rightarrow$ Probability for multiple arrivals/departures in small interval sufficiently small

Source: Stewart, W.J. (2009): Probability, Markov Chains, Queues, and Simulation

Probability the System is in State n at Time $t + \Delta t$

State transition probability diagram:

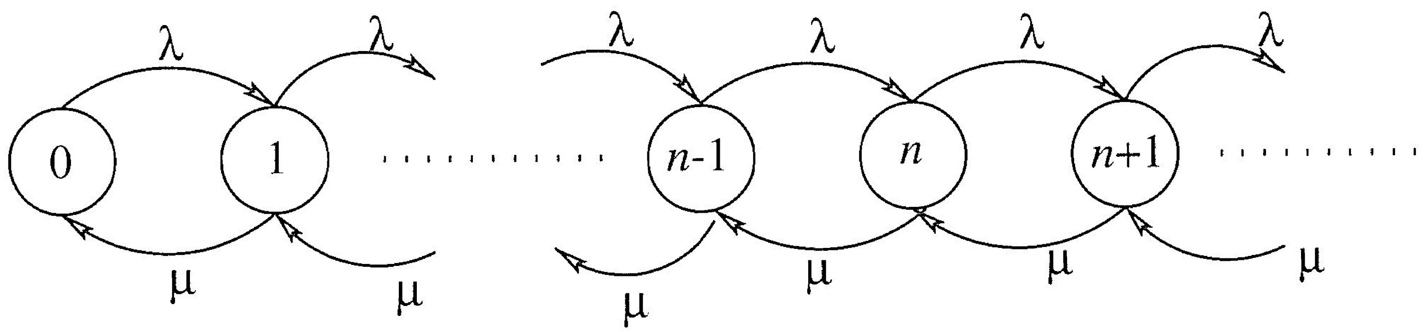


Figure 11.11. State transitions in the $M/M/1$ queue.

Source: Stewart, W.J. (2009): Probability, Markov Chains, Queues, and Simulation

Probability the System is in State n at Time $t + \Delta t$

For $n \geq 1$:

$$\begin{aligned} p_n(t+\Delta t) = & p_n(t)[1-\lambda\Delta t + o(\Delta t)][1-\mu\Delta t + o(\Delta t)] \\ & + p_{n-1}(t)[\lambda\Delta t + o(\Delta t)] \\ & + p_{n+1}(t)[\mu\Delta t + o(\Delta t)] \end{aligned}$$

When $n = 0$:

$$\begin{aligned} p_0(t+\Delta t) = & p_0(t)[1-\lambda\Delta t + o(\Delta t)] \\ & + p_1(t)[\mu\Delta t + o(\Delta t)] \end{aligned}$$

Source: Stewart, W.J. (2009): Probability, Markov Chains, Queues, and Simulation

Probability the System is in State n at Time $t + \Delta t$

Expanding the right-hand side:

$$\begin{aligned} p_n(t+\Delta t) =& p_n(t) - (\lambda+\mu)\Delta t p_n(t) + \lambda \Delta t p_{n-1}(t) + \mu \Delta t p_{n+1}(t) + o(\Delta t), \\ & n \geq 1, \\ p_0(t+\Delta t) =& p_0(t) - \lambda \Delta t p_0(t) + \mu \Delta t p_1(t) + o(\Delta t) \end{aligned}$$

Subtracting $p_n(t)$ from each side and dividing by Δt:

$$\begin{aligned} \frac{p_n(t+\Delta t) - p_n(t)}{\Delta t} =& -(\lambda+\mu)p_n(t) + \lambda p_{n-1}(t) + \mu p_{n+1}(t) + \frac{o(\Delta t)}{\Delta t}, \\ & n \geq 1 \\ \frac{p_0(t+\Delta t) - p_0(t)}{\Delta t} =& -\lambda p_0(t) + \mu p_1(t) + \frac{o(\Delta t)}{\Delta t} \end{aligned}$$

Source: Stewart, W.J. (2009): Probability, Markov Chains, Queues, and Simulation

Probability the System is in State n at Time $t + \Delta t$

Limit $\Delta t \to 0$:

$$\frac{dp_n(t)}{dt} = -(\lambda + \mu)p_n(t) + \lambda p_{n-1}(t) + \mu p_{n+1}(t), \; n \geq 1,$$

$$\frac{dp_0(t)}{dt} = -\lambda p_0(t) + \mu p_1(t)$$

$\rightarrow$ Difference-Differential equation!

Source: Stewart, W.J. (2009): Probability, Markov Chains, Queues, and Simulation

Probability the System is in State n at Time $t + \Delta t$

Assumption $\rightarrow$ steady state exists:

$$\frac{dp_n(t)}{dt} = 0, \qquad n = 0, 1, \ldots$$

$$0 = -(\lambda + \mu)p_n + \mu p_{n+1} + \lambda p_{n-1}, \qquad n \geq 1,$$

$$0 = -\lambda p_0 + \mu p_1 \implies p_1 = \frac{\lambda}{\mu} p_0$$

Rearranging the set of difference equations:

$$p_{n+1} = \frac{\lambda + \mu}{\mu} p_n - \frac{\lambda}{\mu} p_{n-1}, \qquad n \geq 1$$

Source: Stewart, W.J. (2009): Probability, Markov Chains, Queues, and Simulation

Probability the System is in State n at Time $t + \Delta t$

Substituting $n = 1$:

$$\begin{aligned} p_2 &= \frac{\lambda + \mu}{\mu}\left(\frac{\lambda}{\mu}\right)p_0 - \frac{\lambda}{\mu}p_0 \\ &= \frac{\lambda}{\mu}\left(\frac{\lambda+\mu}{\mu} - 1\right)p_0 \\ &= \frac{\lambda}{\mu}\left(\frac{\lambda}{\mu}\right)p_0 = \frac{\lambda^2}{\mu^2}p_0 \end{aligned}$$

As well as:

$$p_3 = \left(\frac{\lambda}{\mu}\right)^3 p_0$$

We conjecture:

$$p_n = \left(\frac{\lambda}{\mu}\right)^n p_0$$

Source: Stewart, W.J. (2009): Probability, Markov Chains, Queues, and Simulation

Probability the System is in State n at Time $t + \Delta t$

Proof by induction for $n + 1$:

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

Determine p_0 for $\rho = \lambda/\mu$:

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

$$p_0 = \frac{1}{\sum_{n=0}^{\infty} \rho^n}$$

Source: Stewart, W.J. (2009): Probability, Markov Chains, Queues, and Simulation

Probability the System is in State n at Time $t + \Delta t$

$\sum_{n=0}^{\infty} \rho^n$: geometric series $\rightarrow$ converges if and only if $|\rho| < 1$
For $\rho < 1$:

$$\sum_{n=0}^{\infty} \rho^n = \frac{1}{1-\rho}$$

Hence: $p_0 = 1 - \rho = 1 - \lambda/\mu$ or alternatively $\rho = 1 - p_0$

Steady-state solution for the $M/M/1$ queue:

$$p_n = \rho^n(1-\rho) \text{ for } \rho = \lambda/\mu < 1$$

Note:

$\rightarrow$ (discrete) probability mass function of the *geometrically distributed* random variable N

Source: Stewart, W.J. (2009): Probability, Markov Chains, Queues, and Simulation

Performance Measures for the $M/M/1$ Queue

Mean Number in System

Random variable N:

$\rightarrow$ Number of customers at steady state: $L = E[N]$

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

Assumption: system is stable $\rightarrow \rho < 1$:

$$\sum_{n=0}^{\infty} n\rho^{n-1} = \frac{\delta}{\delta\rho}\left[\sum_{n=0}^{\infty}\rho^n\right] = \frac{\delta}{\delta\rho}\left[\frac{1}{1-\rho}\right] = \frac{1}{(1-\rho)^2}$$

$$L = (1-\rho)\cdot\rho\cdot\frac{1}{(1-\rho)^2} = \frac{\rho}{1-\rho} = \frac{\lambda}{\mu-\lambda}$$

Source: Stewart, W.J. (2009): Probability, Markov Chains, Queues, and Simulation

Performance Measures for the $M/M/1$ Queue

Mean Queue Length

Random variable N_q:

$\rightarrow$ Number of customers waiting in the queue at steady state: $L_q = \mathrm{E}[N_q]$

$$L_q = \mathrm{E}[N_q] = 0 \times p_0 + \sum_{n=1}^{\infty}(n-1)p_n$$

$$L_q = \sum_{n=1}^{\infty} np_n - \sum_{n=1}^{\infty} p_n = L - (1 - p_0) = \frac{\rho}{1-\rho} - \rho = \frac{\rho^2}{1-\rho} = \rho L$$

$$L_q = \rho L = L - \rho$$

Source: Stewart, W.J. (2009): Probability, Markov Chains, Queues, and Simulation

Performance Measures for the $M/M/1$ Queue

Average Response Time

Little's law $L = \lambda W$ states:

$$\begin{aligned} \mathrm{E}[N] &= \lambda \mathrm{E}[R] \\ \mathrm{E}[R] = \frac{1}{\lambda}\mathrm{E}[N] &= \frac{1}{\lambda}\frac{\rho}{1-\rho} = \frac{1/\mu}{1-\rho} = \frac{1}{\mu-\lambda} \end{aligned}$$

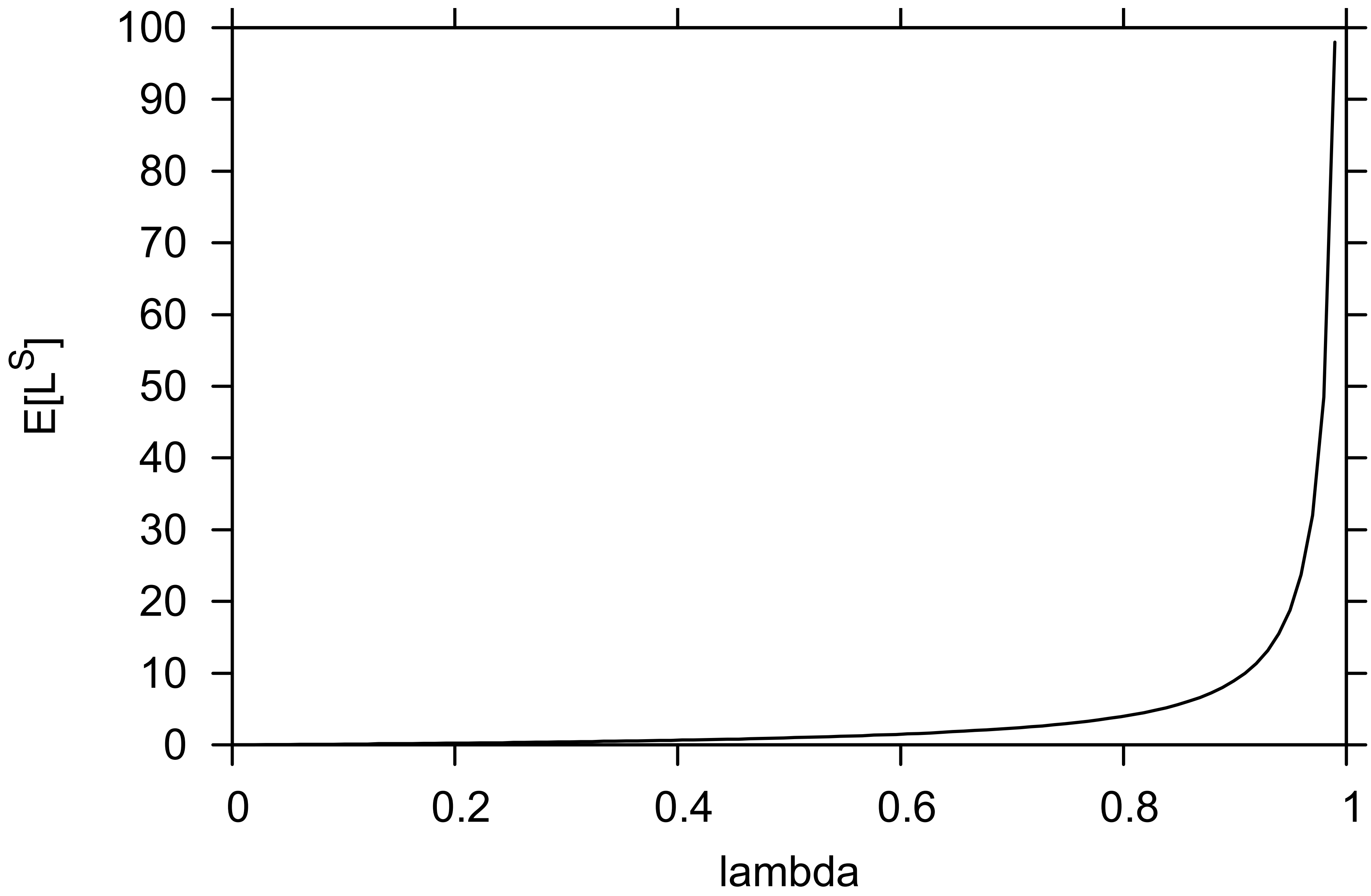
Average Waiting Time

From Little's law:

$$\begin{aligned} L_q &= \lambda W_q \\ W_q = \frac{L_q}{\lambda} &= \frac{\rho}{\lambda} \cdot L = \frac{\lambda}{\mu(\mu-\lambda)} = \frac{\rho}{\mu-\lambda} \end{aligned}$$

Source: Stewart, W.J. (2009): Probability, Markov Chains, Queues, and Simulation

Expected Number of Customers in the System ($\mu = 1$)



Source: Stewart, W.J. (2009): Probability, Markov Chains, Queues, and Simulation