Modeling, Simulating, and Measuring the Performance of Queues

Anthony Hung

2019-03-01

Prerequisites

This vignette continues from concepts covered in the vignette: "Introduction to Queueing Theory and Queueing Models".

Introduction

In addition to simply being able to represent waiting lines mathematically, queueing theory allows for the evaluation of the behavior and performance of queues. Being able to measure the performance of queues also allows us to determine the effects of altering components of the queue on performance.

Simulating a M/M/1 queue

In simulating a M/M/1 queue, we want to keep track of four values of the queue over time.

1. Arrival times of customers
2. Departure times of customers
3. The number of customers in the system at the moment of arrivals or departures
4. The service time of each customer

```
lambda <- 4
mu <- 5
```

Introduction to performance measures for queues

There are several formal quantities used to measure the performance of a queueing system (with c servers).

1. p_j := The stationary probability that there are j customers in the system
2. a := Offered load. The mean number of requests per service time.
3. ρ := Traffic intensity. Offered load per server (a/c).
4. a' := Carried load. Mean number of busy servers.
5. ρ' := Server occupancy. Carried load per server (a'/c).
6. W_s := Mean length of time between a customer's arrival and the customer's departure from the system.
7. W_q := Mean length of time between a customer's arrival and when the customer's service starts.
8. L_s := Mean number of customers in the system, including those in the buffer and at servers.

9. L_q := Mean number of customers waiting in the buffer.

Perform an analysis of the simulated M/M/1 queueing process including a steady state analysis of the system (stationary distribution) and performance measures (loss probability, average waiting time, idle time).

Multiple servers: the M/M/c queue