Week 12

34 $M/M/1$ queue with finite capacity

Previous model assumes that there are no limit on the number of customers that could be in the system at the same time. However, in reality there is always a finite system capacity N, in the sense that there can be no more than N customers in the system at any time. Hence, in general, if an arriving customer finds that there are already N customers present, he does not enter the system.

In this case, as usual, the arrival process N_t is a Poisson process with rate λ and the service time $Y \sim Exp(\mu)$. Let L_t denote the number of customers (either waiting or being served) in the system.

- L_t is *the system length* at the time t.
- $\{L_t, t \geq 0\}$ is a continuous MC with finite state space $S = \{0, 1, ..., N\}$ and is a birth and death chain with birth rate $b_i = \lambda > 0, 0 \leq i \leq N-1$, $b_i = 0, i \geq N$, and death rate $\mu_0 = 0, \mu_i = \mu$, $1 \leq i \leq N$, i.e.,

$$
\begin{aligned}
&P(L_{t+h} = k \mid L_t = i) \\
= &\begin{cases}
1 - \lambda\, h + o(h) & \text{for } k = i = 0, \\
1 - (\lambda + \mu)h + o(h) & \text{for } k = i = 1, ..., N-1,, \\
1 - \mu h + o(h) & \text{for } k = i = N, \\
\lambda h + o(h) & \text{for } k = i + 1,\ 0 \leq i \leq N-1, \\
\mu h + o(h) & \text{for } k = i - 1,\ 1 \leq i \leq N, \\
o(h) & \text{for } |k - i| \geq 2\ .
\end{cases}
\end{aligned}
$$

The Q-matrix is given by

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & ... & 0 & 0 & 0 \\ \mu & -(\lambda+\mu) & \lambda & 0 & ... & 0 & 0 & 0 \\ 0 & \mu & -(\lambda+\mu) & \lambda & ... & 0 & 0 & 0 \\ . & . & . & & & & & \\ 0 & 0 & 0 & 0 & ... & \mu & -(\lambda+\mu) & \lambda \\ 0 & 0 & 0 & 0 & ... & 0 & \mu & -\mu \end{pmatrix}.$$

Now, by using Theorem 11.1, the $\pi = (\pi_0, \pi_1, ...\pi_N)$ is a stationary distribution iff

$$\pi\, Q = 0,$$

i.e.,

$$\begin{aligned} \lambda\pi_0 &= \mu\pi_1 \\ (\lambda+\mu)\pi_j &= \lambda\pi_{j-1} + \mu\pi_{j+1}, \quad 1 \le j \le N-1 \\ \mu\pi_N &= \lambda\pi_{N-1}. \end{aligned}$$

This yields that

$$\begin{aligned} \pi_1 &= \rho\,\pi_0 \\ \pi_{j+1} - \pi_j &= \rho(\pi_j - \pi_{j-1}) = ... = \rho^j(\pi_1 - \pi_0), \quad 1 \le j \le N-1 \\ \pi_N &= \rho\,\pi_{N-1} \end{aligned}$$

Hence, solving in terms of π_0, we obtain

$$\pi_k \;=\; \rho^k \pi_0, \quad 1 \le k \le N.$$

Since $\sum_{k=0}^{N} \pi_k = 1$, it follows that

$$\pi_k \;=\; \frac{1-\rho}{1-\rho^{N+1}}\rho^k, \quad k = 0, 1, 2, ..., N$$

Having found the stationary distribution of the process $\{L_t, t \geq 0\}$), which is the distribution of $L := L_\infty$, we may answer the questions regarding the system's behavior (in the long run).

- $P\,(\text{system is busy}) = P(L = N) = \frac{1-\rho}{1-\rho^{N+1}}\rho^N$;
- The expected number of customers (waiting and being served):

$$EL = \sum_{k=1}^{N} k\pi_k = \frac{1-\rho}{1-\rho^{N+1}} \sum_{k=1}^{N} k\rho^k.$$

- The expected the number of customers waiting to be served:

$$E\max\{L-1, 0\} = EL - \sum_{k=1}^{N} \pi_k.$$

- The expected number of customers being served:

$$E\min\{L, 1\} = \sum_{k=1}^{N} \pi_k = 1 - \pi_0 = \rho(1-\rho^N)/(1-\rho^{N+1}).$$

Let W and V denote the waiting time and the service time of a customer in the queue $M/M/1$, respectively. Note that

$$W = \begin{cases} 0 & \text{if } L = 0, \\ Y_1^* + Y_2 + ... + Y_L & \text{if } L \geq 1, \end{cases}$$

where Y_1^* is the residual service time of the customer being served and Y_j is the service time for j-th customer in the queue. By the fact that Y_j are iid $Exp(\mu)$ and Y_1^* is also exponential with parameter μ, due to the lack of memory property of the exponential distribution, it is easy to get:

- The expected waiting time of a customer before service starts:

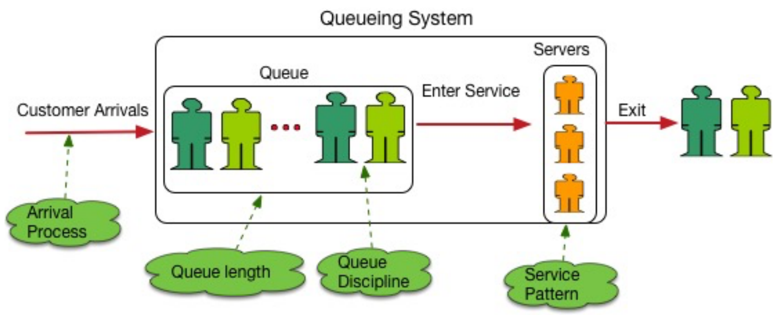
$$E\,(W) \;=\; E\Big(\sum_{j=1}^{L} Y_j I(L \geq 1)\Big) \;=\; EL\, EY_1 = EL/\mu.$$

- Total time of a customer spent in the system:

$$E\,(W + V) \;=\; (EL + 1)/\mu.$$

35 The queue $M/M/k$, $1 < k < \infty$

Consider there are k servers, working in parallel independent of each other, in the queue system. As usual, the arrival process N_t is a Poisson process with rate λ and the service time $Y \sim Exp(\mu)$.



Suppose that the customers are served on a "first come, first served" basis and the arriving customers form a single queue and go to the counter which will become free first.

Let L_t denote the number of customers (either waiting or being served) in the $M/M/k$.

- L_t is *the system length* at the time t.
- As in the case of a $M/M/1$ queue, $\{L_t, t \geq 0\}$ is a birth and death chain with birth rate $b_i = \lambda > 0, i \geq 0$, and death rate $\mu_0 = 0$ and

$$\mu_i = \begin{cases} i\mu & \text{if } 1 \leq i < k, \\ k\mu & \text{if } i \geq k. \end{cases} ,$$

i.e.,

$$P(L_{t+h} = j \mid L_t = i) = \begin{cases} 1 - (\lambda_i + \mu_i)\, h + o(h) & \text{for } j = i, \\ \lambda_i h + o(h) & \text{for } j = i+1, \\ \mu_i h + o(h) & \text{for } j = i-1, \\ o(h) & \text{for } |j-i| \geq 2 \ . \end{cases}$$

The Q-matrix is given by

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & 0 & ... & 0 & ... & ... \\ \mu_1 & -(\lambda+\mu_1) & \lambda & 0 & ... & ... & & & \\ 0 & \mu_2 & -(\lambda+\mu_2) & \lambda & 0 & ... & ... & & \\ . & . & \mu_3 & ... & & & & & \\ . & . & . & .... & & & & & \\ 0 & & 0 & 0 & 0 & ... & \lambda & ... & ... \\ 0 & & 0 & 0 & 0 & ... & -(\lambda+\mu_k) & \lambda & ... \\ 0 & & 0 & 0 & 0 & ... & \mu_k & -(\lambda+\mu_k) & ... \\ 0 & & 0 & 0 & 0 & ... & 0 & \mu_k & ... \\ . & . & . & .... & & & & & \end{pmatrix}.$$

Write $\rho_1 = \rho/k$ (*traffic intensity*), where $\rho = \lambda/\mu$, and

$$p_{ij}(t) = P(L_{t+s} = j \mid L_s = i).$$

Th12.1. Suppose $\rho_1 < 1$ ($\rho < k$). Then, for all $i, j \geq 0$,

$$\begin{aligned} \pi_j &:= \lim_{t\to\infty} P(L_t = j) = \lim_{t\to\infty} p_{ij}(t) \\ &= \begin{cases} \pi_0 \rho^j/j! & \text{if } 0 \leq j < k, \\ \pi_0 k^k (\rho/k)^j/k! & \text{if } j \geq k, \end{cases} \\ &= \begin{cases} \pi_0\, \rho^j/j! & \text{if } 0 \leq j \leq k, \\ \pi_k\, (\rho/k)^{j-k} & \text{if } j > k, \end{cases} \end{aligned}$$

where

$$\pi_0 = \Big[\sum_{j=0}^{k-1} \frac{\rho^j}{j!} + \frac{\rho^k}{k!}(1-\rho/k)^{-1}\Big]^{-1}.$$

Note: $\pi = (\pi_0, \pi_1, ...)$ satisfies that $\pi Q = 0$, i.e., π is the stationary distribution $M/M/k$ queue if $\rho/k < 1$.

In fact, due to $\pi Q = 0$, we have

$$\begin{aligned}
\lambda\pi_0 &= \mu\pi_1 \\
(\lambda + j\mu)\pi_j &= \lambda\pi_{j-1} + (j+1)\mu\,\pi_{j+1}, \quad 1 \le j \le k-1 \\
(\lambda + k\mu)\pi_j &= \lambda\pi_{j-1} + k\,\mu\,\pi_{j+1}, \quad j \ge k,
\end{aligned}$$

indicating

$$\begin{aligned}
\pi_{j+1} &= \frac{\rho}{j+1}\pi_j = ... = \frac{\rho^{j+1}}{(j+1)!}\pi_0, \quad 1 \le j \le k-1, \\
\pi_{j+1} &= \frac{\rho}{k}\pi_j = ... = (\rho/k)^{j+1-k}\pi_k, \quad j \ge k-1
\end{aligned}$$

Having found the distribution $\{\pi_j\}$ of $L = L_\infty$ (the stationary distribution of the process $\{L_t, t \geq 0\}$), we may answer the questions regarding the system's behavior (in the long run).

- $$\begin{aligned} P\,(\text{all servers busy}) &= P(L \geq k) = \pi_k \sum_{j=k}^{\infty} (\rho/k)^{j-k} \\ &= (1-\rho/k)^{-1}\pi_k; \end{aligned}$$

- The expected number of customers (waiting and being served):

$$\begin{aligned} EL &= \sum_{n=1}^{k-1} \frac{n\rho^n}{n!}\pi_0 + \sum_{n=k}^{\infty} n\left(\frac{\rho}{k}\right)^{n-k}\pi_k \\ &= \rho + \pi_k \frac{\rho}{k}(1-\rho/k)^{-2}. \end{aligned}$$

- The expected number of customers waiting to be served:

$$\begin{aligned} E\max\{L-k, 0\} &= \sum_{j=k+1}^{\infty} (j-k)\,P(L=j) \\ &= \pi_k \sum_{j=1}^{\infty} j\,(\rho/k)^j = \pi_k\,(\rho/k)\,(1-\rho/k)^{-2}. \end{aligned}$$

- The expected number of customers being served: $E\min\{L, k\} = \rho$.

- The expected waiting time of a customer before service starts:

$$E\,(W) \;=\; \sum_{j=1}^{\infty} j\pi(j+k)E\min\{V_1, V_2, ..., V_k\} = \frac{\pi_k}{k\mu}(1-\rho/k)^{-2},$$

where V_j denotes the service time at the j-th counter, and W the waiting time.

Note: Let R_i denote the time of i-th person departure from the system. In terms of the non-memory property of exponential distribution, we have $r_i = R_i - R_{i-1}$ are independent and

$$r_i =_d \min\{V_1, V_2, ..., V_k\}.$$

It follows that

$$W = \begin{cases} 0, & \text{if } L < k, \\ R_{L-k+1}, & \text{if } L \geq k, \end{cases}$$

and

$$\begin{aligned} EW &= \sum_{j=k}^{\infty} P(L=j)\,(j-k+1)E\min\{V_1, V_2, ..., V_k\} \\ &= \sum_{j=0}^{\infty} P(L=j+k)\,(j+1)\,\frac{1}{\mu k} \\ &= \frac{\pi_k}{\mu k}\sum_{j=0}^{\infty}(j+1)\,(\rho/k)^j \\ &= \frac{\pi_k}{\mu k}\,(1-\rho/k)^{-2}. \end{aligned}$$

Note: As in the queue $M/M/1$, it is easy to check **Little's law**:

$$\begin{aligned} E(\#\text{ waiting to be served }) &= \lambda\, E(\text{waiting time before service starts}),\\ E(\#\text{ being served }) &= \lambda\, E(\text{service time}),\\ E(\#\text{ in the system }) &= \lambda\, E(\text{time in the system}). \end{aligned}$$

That is

$$\begin{aligned} E\max\{L-k,0\} &= \lambda\, EW; \quad E\min\{L,k\} = \lambda\, EV;\\ EL &= \lambda\,(EW+EV) \end{aligned}$$

The queue $M/M/\infty$

In this case, there is no queue: each customer, upon arrival, is served immediately. The system length L_t becomes the number of busy servers at time t.

$\{L_t, t \geq 0\}$ is a birth and death chain with birth rate $\lambda_i = \lambda > 0, i \geq 0$, and death rate $\mu_0 = 0$ and $\mu_i = i\mu, i \geq 1$.

The Q-matrix is given by

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & 0 & \ldots \\ \mu_1 & -(\lambda+\mu_1) & \lambda & 0 & .. & \\ 0 & \mu_2 & -(\lambda+\mu_2) & \lambda & 0 & .. \\ . & . & . & & & \end{pmatrix}.$$

It can be shown by using $\pi\, Q = 0$, i.e.,

$$\begin{aligned} \lambda\pi_0 &= \mu\pi_1 \\ (\lambda + j\mu)\pi_j &= \lambda\pi_{j-1} + (j+1)\mu\,\pi_{j+1}, \quad j \geq 1, \end{aligned}$$

that, for all $0 < \lambda, \mu < \infty$, the stationary distribution of $\{L_t, t \geq 0\}$ exists and is given by

$$\pi_j = \lim_{t\to\infty} p_{ij}(t) = \big(\lambda/\mu\big)^j \, \frac{e^{-\lambda/\mu}}{j!}, \quad j \geq 0,$$

i.e., $\{\pi_j\}$ is a Poisson distribution with parameter λ/μ.

36 Introduction to Brownian motion and martingale

A random process $\{X_t\}_{t\in T}$ may be considered as a well-defined process once its state space S, index set T and the joint distribution family are prescribed. Depending on the nature of the state space S and the index set T, we may classify random processes into four classes:

1. Both S and T discrete: discrete valued random processes. The random walk and Markov chain are in this class.

2. S discrete and T continuous: discrete valued continuous parametric random processes. The Poisson process and point processes are in this class.

3. S continuous and T discrete: continuous random process with discrete parameter. Many financial random processes, such as the process in Ex 1.3, in this class.

4. Both S and T continuous: continuous random processes. Brownian motion, etc, are in this class.

Examples of continuous stochastic processes

1. **Gaussian processes**

   A stochastic process $\{X_t, t \geq 0\}$ is called a **Gaussian process** if all its finite dimensional distribution are multivariate normal.

2. **Markov processes**

   A continuous-time process $\{X_t, t \geq 0\}$ is called a **Markov process** if for each $t \geq 0$, each $A \in \sigma(X_s, s > t)$ (the "future") and $B \in \sigma(X_s, s < t)$ (the "past"),

   $$P(A|X_t, B) = P(A|X_t).$$

   That is, if you know where you are (at time t), how you got there doesn't matter so far as predicting the future is concerned – equivalently, past and future are conditionally independent given the present.

3. **Martingales**

   A continuous-time process $\{X_t, t \geq 0\}$ is called a **martingale** if

   (ii) $E|X_t| < \infty$ for all $t < \infty$,
   (iii) $E[X_t|X_s] = X_s$ for all $0 \leq s \leq t$.

   Note: more general definition exists. One can also define upper-martingales, sub-martingales, semi-martingales, local martingales, etc.

4. **Diffusion, Levy processes, Fractional processes, etc. ..**

5. **Brownian motion**

A stochastic process $\{B_t, t \geq 0\}$ is a **Brownian motion** if

(a) It starts at zero: $B_0 = 0$ a.s.;

(b) It has stationary, independent increments;

(c) For every $t > 0$, $B_t \sim N(0, t)$;

(d) The path $t \to B_t$ is continuous almost surely.

(It has continuous sample paths: no jumps.)

An **equivalent definition** of Brownian motion:
A stochastic process $\{B_t, t \geq 0\}$ is a Brownian motion if and only if

(a') B_t is a Gaussian process (i.e., all its finite dimensional distributions (f.d.d.s.) are multivariate normal),

(b') $EB_s = 0$ and $E(B_s B_t) = s \wedge t = \min\{s, t\}$.

(c') With probability one, $t \to B_t$ is continuous.

This result states: any continuous real-valued process $\{X_t, t \geq 0\}$ that is zero-mean Gaussian process with covariance $Cov(X_s, X_t) = s \wedge t$ is a Brownian motion. This simple fact turns out to be an extremely efficient means of checking when a process is a Brownian motion, the following few simple but extremely important examples serve to illustrate this.

Let $\{B_t, t \geq 0\}$ be a Brownian motion. So are the following

1. $\{B_{t+t_0} - B_{t_0}, t \geq 0\}, \forall t_0 \geq 0$;
2. $\{-B_t, t \geq 0\}$;
3. $\{cB_{t/c^2}, t \geq 0\}, \forall c \neq 0$;
4. $\{tB_{1/t}, t \geq 0\}$, where $tB_{1/t}$ is taken to be zero when $t = 0$.

Remark

Brownian motion is a continuous-time and continuous-state space Markov process which is originally used to explain the physical phenomenon behind the rapid ceaseless irregular motion of a small particle suspended in a fluid. In past decades, it has become very popular in financial mathematics. Indeed, stock markets, the foreign exchange markets, commodity markets and bond markets are all assumed to follow Brownian Motion where assets are changing continually over very small intervals of time and the position, change of state, is being altered by random amounts. More importantly, the mathematical models used to describe Brownian Motion are the fundamental tools on which all financial asset pricing and derivatives pricing models are based. These models are the key to so much of the work that is being done on market models and risk analysis.

Martingales in related to Brownian motion

Let $\{B_t, t \geq 0\}$ be a Brownian motion. Then,

- $B_t, t \geq 0$ is a martingale,
- $B_t^2 - t, t \geq 0$ is a martingale,
- for any u, $e^{uB_t - u^2 t/2}, t \geq 0$ is a martingale.

Remark 1. If a process X_t is a continuous martingale such that $X^2(t) - t$ is also a martingale, then $X(t)$ is a Bronian motion.
Remark 2. Brownian motion possesses Markov property.

Hitting times, maximum variable

Let $T_a = \inf\{t \geq 0, B_t = a\}$ denote the hitting time of a by a Brownian motion. If $a = 0$, then $T_a = 0$. Furthermore, we have

$$\begin{aligned} P\left(\max_{0\leq s\leq t} B_s \geq a\right) &= P(T_a \leq t) \\ &= 2P(B_t \geq a) = \frac{2}{\sqrt{2\pi t}} \int_a^\infty \exp(-x^2/2t)dx. \end{aligned}$$

The above formulas show the importance of a Brownian motion. Usually, for arbitrary stochastic process X_t, it is hard to compute

$$P(\sup_{0\leq s\leq t} X_s > a).$$

Such expression have fundamental importance in queueing theory, insurance etc.

Processes derived from Brownian motion

Brownian bridge

Consider the process

$$X_t = B_t - tB_1, \qquad 0 \le t \le 1.$$

Obviously,

$$X_0 = B_0 - 0 \times B_1 = 0, \qquad X_1 = B_1 - 1 \times B_1 = 0.$$

For this simple reason, X_t is called **Brownian bridge** or **tied down Brownian motion**. Brownian bridge appears as the limit of the normalized **empirical d.f.** of a sample of i.i.d. $U(0,1)$ r.v.'s, and has found important applications in nonparametric statistics.

It can be easily shown that X_t is a Gaussian process with mean and covariance functions given by

$$EX_t = 0, \qquad Cov(X_s, X_t) = \min(s,t) - st, \quad s,t \in [0,1].$$

Brownian motion with (linear) drift

Consider the process

$$X_t = \mu t + \sigma B_t, \qquad t \ge 0.$$

It is called **Brownian motion with (linear) drift** μ.

It can be easily shown that X_t is a Gaussian process with mean and covariance functions given by

$$EX_t = \mu t, \qquad Cov(X_s, X_t) = \sigma^2 \min(s,t), \quad s,t \ge 0.$$

Geometric Brownian motion

Bachelier (1900) was the first to use Brownian motion with drift to model stock prices, which are non-negative. However, Brownian motion can be negative. One way to overcome this is to use **geometric Brownian motion**:

$$X_t = X_0 e^{\mu t + \sigma B_t}, \qquad t \geq 0.$$

It can be easily shown that X_t is NOT a Gaussian process (Why? This is because X_t is no longer Gaussian; in fact, X_t follows a log-normal distribution.)