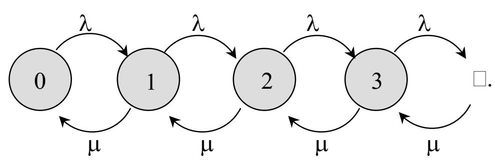
Section 7: Coupling

Contents

7.1 The M/M/1 Queue

The M/M/1 queue is a birth-death Markov jump process $X = (X(t) : t \geq 0)$ on the state space $\mathbb{Z}_+ (\triangleq \{0, 1, 2, \ldots\})$ having birth rates $\lambda_n \equiv \lambda > 0$ and death rates $\mu_n \equiv \mu > 0$. Its transition rate diagram is:



The process $X = (X(t) : t \geq 0)$ is the "number-in-system" process for a queue with infinite capacity waiting room, a single server, a Poisson arrival process (having rate λ), and iid $\text{Exp}(\mu)$ service times.

When $\lambda < \mu$, it is well known that

$$X(t) \Rightarrow X(\infty) \tag{7.1.1}$$

as $t \to \infty$ (where $\Rightarrow$ denotes "weak convergence"), where

$$P(X(\infty) = x) = (1 - \rho)\rho^x$$

for $x \geq 0$, where $\rho \triangleq \lambda/\mu$. The parameter ρ is called the "traffic-intensity" or "utilization" (note that $\rho = P(X(\infty) \geq 1) = P(\text{server is busy})$).

Let $L = \sum_{j=0}^{\infty} j\pi(j)$ be the steady-state mean number-in-system. For the M/M/1 queue,

$$L = \rho(1 - \rho)^{-1}.$$

Note the singularity at $\rho = 1$; this stylized model reflects the performance degradation that arises when a queue is in "heavy traffic".

7.2 The Heavy-Traffic Limit Theorem for the M/M/1 Queue

We are interested in the behavior of $X(\infty)$ as $\rho \nearrow 1$. Let $X_\rho(\infty)$ be a rv having the geometric distribution

$$P(X_\rho(\infty) = x) = (1 - \rho)\rho^x.$$

Theorem 7.2.1 $(1-\rho)X_\rho(\infty) \Rightarrow \mathrm{Exp}(1)$ as $\rho \nearrow 1$.

Remark 7.2.1 Theorem 7.2.1 suggests the approximation

$$P(X_\rho(\infty) \in \cdot) \approx P((1-\rho)^{-1}\mathrm{Exp}(1) \in \cdot)$$

for ρ close to 1.

To prove Theorem 7.2.1, we first offer a "sledge-hammer" proof based on use of transforms.

Definition 7.2.1 Given a rv X, the *characteristic function* of X is the function

$$c_X(\theta) = \mathrm{E}\exp(i\theta X)$$

Remark 7.2.2 The characteristic function is essentially the Fourier transform of the distribution of X:

$$c_X(\theta) = \int_{\mathbb{R}} \exp(i\theta x) P(X \in dx)$$

We need the following result on the relationship between characteristic functions and weak convergence.

Theorem 7.2.2 Given a sequence $(Z_n : 1 \le n \le \infty)$ of random variables, $Z_n \Rightarrow Z_\infty$ as $n \to \infty$ if and only if for each $\theta \in \mathbb{R}$,

$$c_{Z_n}(\theta) \to c_{Z_\infty}(\theta)$$

as $n \to \infty$.

For a proof, see R. Durrett, "Probability: Theory and Examples".

First Proof of Theorem 7.2.1 Note that

$$\begin{aligned}
\mathrm{E}\exp(i\theta(1-\rho)X_\rho(\infty)) &= \sum_{k=0}^{\infty} \exp(i\theta(1-\rho)k)(1-\rho)\rho^k \\
&= \frac{1-\rho}{1-\exp(i\theta(1-\rho))\rho} \\
&= \frac{1-\rho}{1-\rho\left(1+i\theta(1-\rho)-\frac{\theta^2}{2}(1-\rho)^2+\cdots\right)} \\
&= \frac{1-\rho}{1-\rho\left(1+i\theta(1-\rho)+o(1-\rho)\right)} \\
&\to \frac{1}{1-i\theta} \\
&= \int_0^\infty \exp(i\theta x)e^{-x}\,dx,
\end{aligned}$$

proving the result. □

We will soon offer a probabilistic proof of Theorem 7.2.1 based on the use of "coupling". The coupling argument will also provide an associated rate of convergence for Theorem 7.2.1.

We introduce the coupling idea by illustrating its use in a couple of more elementary settings.

Example 7.2.1 We will use coupling to prove that when $\mathrm{E}X^2 < \infty$, then the inequality $\mathrm{E}X^2 \geq (\mathrm{E}X)^2$ holds.

Despite the fact that this inequality involves only one X, the coupling proof introduces two iid copies of X, call them X_1 and X_2. Clearly, $Z = (X_1 - X_2)^2 \geq 0$ so $\mathrm{E}Z \geq 0$. But

$$\begin{aligned} \mathrm{E}Z &= \mathrm{E}(X_1 - X_2)^2 \\ &= \mathrm{E}X_1^2 - 2\mathrm{E}X_1X_2 + \mathrm{E}X_2^2 \\ &= 2\mathrm{E}X^2 - 2\mathrm{E}X \cdot \mathrm{E}X \\ &= 2(\mathrm{E}X^2 - (\mathrm{E}X)^2) \geq 0, \end{aligned}$$

proving the inequality. Note that we have constructed a particular joint distribution/probability space which serves to simplify the calculation.

Example 7.2.2 Let $X_1, X_2, \ldots, X_d$ be a collection of d independent rv's. We say that $f : \mathbb{R}^d \to \mathbb{R}$ is increasing if $f(x) \leq f(y)$ whenever $x \leq y$, where we write $x \leq y$ whenever all the components of $y - x$ are non-negative. Suppose that f_1 and f_2 are increasing functions in $\mathbb{R}^d$ such that $\mathrm{E}f_i(X_1, \ldots, X_d)^2 < \infty$ for $i = 1, 2$. It seems intuitively reasonable that

$$\mathrm{cov}(f_1(X_1, \ldots, X_d), f_2(X_1, \ldots, X_d)) \geq 0.$$

Under the conditions stated, the above inequality is valid. We will now prove this result for $d = 1$ by the use of coupling.

We again introduce two independent copies $\tilde{X}_1$ and $\tilde{X}_2$ of X_1. Note that because f_1 and f_2 are increasing, it is evident that $f_1(\tilde{X}_1) - f_1(\tilde{X}_2)$ and $f_2(\tilde{X}_1) - f_2(\tilde{X}_2)$ have the same sign. So

$$Z = (f_1(\tilde{X}_1) - f_1(\tilde{X}_2))(f_2(\tilde{X}_1) - f_2(\tilde{X}_2)) \geq 0,$$

Hence, $\mathrm{E}Z \geq 0$, so

$$\begin{aligned} \mathrm{E}Z &= 2(\mathrm{E}f_1(X_1)f_2(X_1) - \mathrm{E}f_1(X_1)\mathrm{E}f_2(X_1)) \\ &= 2\mathrm{cov}(f_1(X_1), f_2(X_1)) \geq 0. \end{aligned}$$

To provide a coupling-based proof of Theorem 7.2.1, let $Z \stackrel{\mathcal{D}}{=} \mathrm{Exp}(1)$.

Lemma 7.2.1 Put $Z_h = \lfloor Z/h \rfloor$. Then, $Z_h \stackrel{\mathcal{D}}{=} \mathrm{Geometric}(1 - e^{-h})$

Proof: Note that

$$P(Z_h \geq k) = P(Z \geq kh) = e^{-kh}. \quad \square$$

Observe that $X_\rho(\infty) \stackrel{\mathcal{D}}{=} \mathrm{Geometric}(1 - \rho)$. So,

$$X_\rho(\infty) \stackrel{\mathcal{D}}{=} Z_{\log(1/\rho)}.$$

For any $x \in \mathbb{R}$, $|h\lfloor x/h \rfloor - x| \leq h$, so evidently

$$|\log(1/\rho)Z_{\log(1/\rho)} - Z| \leq \log(1/\rho).$$

But

$$\log(1/\rho) = \log(1/(1-(1-\rho))) = 1-\rho + O((1-\rho)^2).$$

It follows that

$$(1-\rho)Z_{\log(1/\rho)} \to Z \quad \text{a.s.}$$

as $\rho \nearrow 1$, from which it is immediate that

$$(1-\rho)X_\rho(\infty) \Rightarrow Z = \text{Exp}(1)$$

as $\rho \nearrow 1$, thereby completing our coupling argument.

Remark 7.2.3 The above argument makes clear that

$$|(1-\rho)Z_{\log(1/\rho)} - Z| = O(1-\rho)$$

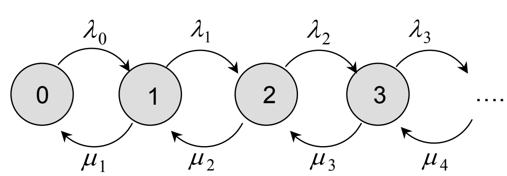
as $\rho \nearrow 1$. From this, we can show that (we leave this as an exercise)

$$|P((1-\rho)X_\rho(\infty) \le x) - P(Z \le x)| = O(1-\rho)$$

as $\rho \nearrow 1$, yielding our rate of convergence.

7.3 Additional Applications of Coupling

Suppose that $X = (X(t) : t \ge 0)$ is a birth-death process on $\mathbb{Z}_+$ with birth rates $(\lambda_n : n \ge 0)$ and death rates $(\mu_n : n \ge 1)$, so that its transition rate diagram is:



It seems intuitively reasonable that the larger the starting state of X, the higher X should be. To make this precise, we need a definition.

Definition 7.3.1 Suppose that X and Y are real-valued rv's. We say that X is *stochastically larger* than Y (and write $X \overset{\text{st}}{\ge} Y$) if

$$P(X > x) \ge P(Y > x)$$

for each $x \in \mathbb{R}$.

Here is the theorem we would like to prove.

Theorem 7.3.1 Suppose $x \ge y$. Then, for each $t \ge 0$,

$$P_x(X(t) \in \cdot) \overset{\text{st}}{\ge} P_y(X(t) \in \cdot)$$

Before proving this result, we wish to offer a useful re-formulation of "stochastic ordering".

Proposition 7.3.1 The following are equivalent:

i.) $X \overset{\text{st}}{\geq} Y$;

ii.) $P(X > x) \geq P(Y > x)$ for all $x \in \mathbb{R}$;

iii.) $P(X \leq x) \leq P(Y \leq x)$ for all $x \in \mathbb{R}$;

iv.) For each non-decreasing non-negative $f : \mathbb{R} \to \mathbb{R}$,

$$\mathrm{E}f(X) \geq \mathrm{E}f(Y).$$

The key to proving the above result is the following lemma.

Lemma 7.3.1 Let X be a real-valued rv and set

$$F_X^{-1}(x) = \sup\{y : P(X \leq y) \leq x\}$$

If U is uniform on $[0,1]$, then $F_X^{-1}(U) \overset{\mathcal{D}}{=} X$, where $\overset{\mathcal{D}}{=}$ denotes "equality in distribution".

Proof of Proposition 7.3.1 The only non-trivial implication is that ii.) implies iv.); we prove this via "coupling". Observe that ii.) implies that $F_X^{-1}(U) \geq F_Y^{-1}(U)$ a.s., from which it follows that $f(F_X^{-1}(U)) \geq f(F_Y^{-1}(U))$ a.s.. This clearly implies that

$$\mathrm{E}f(X) = \mathrm{E}f(F_X^{-1}(U)) \geq \mathrm{E}f(F_Y^{-1}(U)) = \mathrm{E}f(Y). \quad \square$$

We now turn to the proof of Theorem 7.3.1.

Proof of Theorem 7.3.1 Let $X_x = (X_x(t) : t \geq 0)$ be a version of the birth-death process starting from x and let $X_y = (X_y(t) : t \geq 0)$ be an independent version starting from y. Since X_x and X_y are independent, X_x and X_y will never jump simultaneously (since the jump times are independent continuous rv's).

Let $T = \inf\{t \geq 0 : X_x(t) \leq X_y(t)\}$. Because a birth-death process can only jump to a neighboring state and X_x and X_y never jump simultaneously, it follows that $X_x(T) = X_y(T)$ whenever $T < \infty$. Put $\tilde{X}_x = X_x$ and

$$\tilde{X}_y(t) = \begin{cases} X_y(t); & t < T \\ X_x(t); & t \geq T. \end{cases}$$

Then, $\tilde{X}_x \overset{\mathcal{D}}{=} X_x$ and $\tilde{X}_y \overset{\mathcal{D}}{=} X_y$. But

$$\tilde{X}_x \geq \tilde{X}_y \quad \text{a.s.}$$

for all $t \geq 0$, from which it follows that

$$P(X_x(t) > z) = P(\tilde{X}_x(t) > z) \geq P(\tilde{X}_y(t) > z) = P(X_y(t) > z)$$

for $z \in \mathbb{R}$. $\square$

We can also use coupling to prove that whenever a birth-death Markov jump process has a stationary distribution π (so that if $P(X(0) = x) = \pi(x)$ for $x \geq 0$ then $P(X(t) = x) = \pi(x)$ for $x \geq 0$ and $t \geq 0$), then

$$X(t) \Rightarrow X(\infty)$$

as $t \to \infty$, where $P(X(\infty) = x) = \pi(x)$ for $x \geq 0$. (In other words, we will be proving (7.1.1) for general birth-death processes.)

Remark 7.3.1 It is well known that such a stationary distribution π must satisfy $\pi Q = 0$, where Q is the rate matrix of X.

Theorem 7.3.2 Let $X = (X(t) : t \geq 0)$ be a birth-death process on $\mathbb{Z}_+$ with a stationary distribution π. Suppose that $P_x(T_0 < \infty) = 1$ for each $x \in \mathbb{Z}_+$, where

$$T_0 = \inf\{t \geq 0 : X(t) = 0\}.$$

Then,

$$|P_x(X(t) = y) - \pi(y)| \leq P_x(T_0 > t) + P_\pi(T_0 > t)$$

and hence

$$P_x(X(t) = y) \to \pi(y)$$

as $t \to \infty$.

Proof: As in the proof of Theorem 7.3.1, let $X_x = (X_x(t) : t \geq 0)$ be a version of the birth-death process starting from x and let $X_\pi = (X_\pi(t) : t \geq 0)$ be an independent version in which $X(0)$ is distributed according to π.

Let $T = \inf\{t \geq 0 : X_x(t) = X_\pi(t)\}$ be the "coupling time" of the two processes. Because of the nearest neighbor transitions of X_x and X_π, it is evident that

$$T \leq \max\{T_0^x, T_0^\pi\},$$

where T_0^x and T_0^π are the first times that X_x and X_π hit 0. It follows that

$$P(T > t) \leq P_x(T_0 > t) + P_\pi(T_0 > t).$$

Put $\tilde{X}_\pi = X_\pi$ and

$$\tilde{X}_x(t) = \begin{cases} X_x(t); & t < T \\ X_\pi(t); & t \geq T. \end{cases}$$

Once again, $\tilde{X}_\pi \stackrel{\mathcal{D}}{=} X_\pi$ and $\tilde{X}_x \stackrel{\mathcal{D}}{=} X_x$. Note that

$$\begin{aligned} P_x(X(t) = y) &= P(\tilde{X}_x(t) = y) \\ &= P(\tilde{X}_x(t) = y, T \leq t) + P(\tilde{X}_x(t) = y, T > t) \\ &= P(\tilde{X}_\pi(t) = y, T \leq t) + P(\tilde{X}_x(t) = y, T > t) \\ &= P(\tilde{X}_\pi(t) = y) + P(\tilde{X}_x(t) = y, T > t) - P(\tilde{X}_\pi(t) = y, T > t) \\ &= P(X_\pi(t) = y) + P(\tilde{X}_x(t) = y, T > t) - P(\tilde{X}_\pi(t) = y, T > t) \\ &= \pi(y) + P(\tilde{X}_x(t) = y, T > t) - P(\tilde{X}_\pi(t) = y, T > t) \end{aligned}$$

So,

$$P_x(X(t) = y) - \pi(y) \leq P(\tilde{X}_x(t) = y, T > t) \leq P(T > t)$$

and

$$\pi(y) - P_x(X(t) = y) \le P(\tilde{X}_\pi(t) = y, T > t) \le P(T > t).$$

We conclude that

$$|P_x(X(t) = y) - \pi(y)| \le P(T > t),$$

proving the theorem. □