Other properties of $M/M/1$

Přemysl Bejda

premysslbejda@gmail.com

2012
Contents

1 Reflected Lévy Process

2 Time dependent properties of $M/M/1$

3 Waiting times and queue disciplines in $M/M/1$
Contents

1. Reflected Lévy Process

2. Time dependent properties of $M/M/1$

3. Waiting times and queue disciplines in $M/M/1$
Contents

1. Reflected Lévy Process

2. Time dependent properties of $M/M/1$

3. Waiting times and queue disciplines in $M/M/1$
Definition Lévy process

Definition 1 (Lévy process).

A stochastic process \( \{ S_t \}_{t \geq 0} \) is said to be a Lévy process if

1. \( S_0 = 0 \) almost surely,
2. the increments are independent,
3. the increments are stationary (strong).

- We return to Lévy processes in IX.1.
- This is a natural way to define random walk for continuous time.
- Some examples:
  - linear deterministic drift \( (S_t = \theta t) \),
  - standard Brownian motion,
  - compound Poisson process.
Reflected version of Lévy process

- We need some similar tool for continuous time as was Lindley process for discrete case.
- The definition is not so direct.
- We use a proposition from previous presentation.

Definition 2 (Reflected version \( \{ V_t \}_{t \geq 0} \) of \( \{ S_t \}_{t \geq 0} \)).

\[
V_t = (V_0 + S_t) \vee \max_{0 \leq s \leq t} (S_t - S_s),
\]

when \( x = V_0 \) is of importance we will write \( V_t = V_t(x) \).

Further we define:
- \( M_T = \sup_{0 \leq t \leq T} S_t \),
- \( M = \sup_{0 \leq t \leq \infty} S_t \).
Propositions

If the mean of Lévy process $S_t$ is well defined then $E S_t = \mu t$.

**Proposition 3.**

\{ V_t \} is a strong Markov process.

**Corollary 4.**

$V_T \overset{\text{D}}{=} (V_0 + S_T) \lor M_T$. If $\mu < 0$, then $M < \infty$ and $V_T \rightarrow M$ in total variation.

**Proposition 5.**

Define $\omega = \inf\{ t > 0 : V_0 + S_t \leq 0 \}$. Then also

$\omega = \inf\{ t > 0 : V_t = 0 \}$, and $V_t = V_0 + S_t$ for $t < \omega$.

$V_t$ evolves as $S_t$ until the first hitting of 0.
Example

Consider a compound Poisson process of the form

\[ S_t = N_t^\beta - N_t^\delta. \]

\( N_t^\beta \) and \( N_t^\delta \) are independent Poisson processes with intensities \( \beta \) and \( \delta \) respectively.

The reflection means that jumps of \( N_t^\delta \) are ignored when \( V_t = 0 \).

\( \{ V_t \} \) is a Markov process on \( \mathbb{N}_0 \).

The only off-diagonal nonzero intensities are \( \lambda(i, i - 1) = \delta \) for \( i = 1, 2, \ldots \) and \( \lambda(i, i + 1) = \beta \) for \( i = 0, 1, \ldots \).

From previous results we can recognize \( \{ V_t \} \) as the \( M/M/1 \) queue length process.
Outline

1. Reflected Lévy Process

2. Time dependent properties of $M/M/1$

3. Waiting times and queue disciplines in $M/M/1$
The doubly infinite queue

- On the example before we will base the studium of $M/M/1$.
- The process $S_t$ from the example is called doubly infinite queue.
- From the corollary 4 and the example proceeds a following proposition.

**Proposition 6.**

The distribution of the $M/M/1$ queue length $X_t$ at time $t$ given $X_0 = i$ is that of $\max(i + S_t, M_t)$ where $S_t = B_t - D_t$ is the difference between two independent Poisson processes with intensities $\beta$ and $\delta$. 
The doubly infinite queue

Example 7 (Taxis and passengers before railway station).

- $B_t$ is a number of passengers arriving before $t$.
- $D_t$ is a number of taxis arriving before $t$.
- If $S_t > 0$ there is a queue of passengers in $t$ of length $S_t$.
- If $S_t < 0$ there is a queue of taxis in $t$ of length $-S_t$.

Proposition 8.

Let $\rho = \beta / \delta$. Then a.s.:

1. $S_t \to -\infty$, $M < \infty$ when $\rho < 1$.
2. $S_t \to \infty$, $M = \infty$ when $\rho > 1$.
3. $\lim_{t \to \infty} S_t = \infty$, $\lim_{t \to \infty} S_t = -\infty$ when $\rho = 1$. 
The transition probabilities

Definition 9 (Modified Bessel function of order $n \in \mathbb{N}$).

$$I_n(x) = \sum_{k=0}^{\infty} \frac{(x/2)^{n+2k}}{k!(n+k)!}, \quad I_{-n}(x) = I_n(x), \quad n \in \mathbb{N}.$$ 

- Let $\mu = \sqrt{\beta \delta}$ and $\rho = \beta / \delta$.
- In the following we will always employ $x = 2\mu t$ and $I_n$ will denote $I_n(2\mu t)$.
- Let $\iota_n = e^{-(\beta+\delta)t} \rho^{n/2} I_n$.
- From previous we get $\iota_{-n} = \rho^{-n} \iota_n$, $n \in \mathbb{N}$. 
The transition probabilities

Proposition 10. 

In the $M/M/1$ queue with $0 < \rho < \infty$, $p^t_{ij} = P(X_t = j | X_0 = i)$

$$p^t_{ij} = \frac{\rho^{-i} \nu_{i+j+1} + (1 - \rho) \rho^j \sum_{n=-\infty}^{j-i-2} \nu_n}{\sum_{n=-\infty}^{j-i} \nu_n}.$$  

- Derivation of this formula is rather technical.
- One of the ideas employed in the proof is using a process with $\beta, \delta$ both replaced by $\mu$.
- Let us denote this process as $P_0$-process. Such a process has traffic intensity 1.
- Since the $P_0$-process is symmetric we can employ the reflection principle.
The transition probabilities

- There are also derived alternative formulas for $p_{ij}^t$.
- But all these formulas are for such a simple case as $M/M/1$ quite complicated.
- This suggests that time dependent explicit solutions are in general not tractable.
- The most feasible solution may be to apply numerical integration of appropriate formulas.
- The same is true for the busy period distribution formulas. See next slides.
The busy period distribution

- **Busy period** $G$ is the time from when a customer enters an empty system until the system is empty again.
- When the system is empty this is so called *idle period*.
- $G + H$ constitute for *busy cycle*. 
The busy period distribution

- In the $M/M/1$ case are $G$ and $H$ independent.
- $H$ is exponentially distributed with intensity $\beta$.
- The following proposition can be derived from the proposition about the first passage time $\tau = \inf\{ t > 0 : S_t = 1 \}$ from 0 to 1, by a symmetry argument.
- Since we may identify $G$ with the time of passage from 0 to -1.

**Proposition 11.**

The busy period distribution of the $M/M/1$ queue is given by the density

$$g(t) = \delta e^{-(\beta+\delta)t} (I_0 - I_2) = \frac{\rho^{-1/2}}{t} e^{-(\beta+\delta)t} I_1.$$  

But $g$ is defective for $\rho > 1$. LIFO case
Transform methods

- In many cases there is no way to derive $p_{ij}^t$, but even in these cases it may be still possible to find explicit expression for transforms.
- The main drawback of this method is that the expressions are little transparent.
- Their derivation may be complicated.

Lemma 12.

For any $\alpha$, \( \{ Y_t \} = \{ \exp(\alpha S_t - t\kappa(\alpha)) \} \) is a continuous time martingale and we have for $\rho \leq 1$ and $\kappa'(\alpha)$ (i.e. $\alpha \leq -\ln \rho/2$) that

\[
1 = E Y_0 = E Y_G = \exp(-\alpha) E \exp(-G\kappa(\alpha)),
\]

where $\kappa(\alpha) = \ln E \exp(\alpha S_1) = \ln (E \exp(\alpha B_1) E \exp(-\alpha D_1)) = \beta(\exp(\alpha) - 1) + \delta(\exp(-\alpha) - 1)$.
Proposition 13.

For $\rho < 1$, the Laplace transform of the $M/M/1$ busy period is given by

$$E e^{-\theta G} = \xi(\theta) = \frac{1}{2\beta} \left( \beta + \delta + \theta - \sqrt{(\beta + \delta + \theta)^2 - 4\beta\delta} \right)$$

for $\theta \geq 0$. Then

$$EG = -\xi'(0) = \frac{1}{\delta(1 - \rho)}, \quad \text{var}G = \xi''(0) - \xi'(0)^2 = \frac{1 + \rho}{\delta^2(1 - \rho)^3}.$$ 

From this follows also distributional properties of busy cycle. E.g.

$$EG + EH = \frac{1}{\delta(1 - \rho)} + \frac{1}{\beta} = \frac{1}{\beta(1 - \rho)}.$$
The relaxation time

**Theorem 14.**

If $\rho < 1$, then

$$ p_{ij}^t = (1 - \rho) \rho^j + \frac{e^{-rt}}{4\sqrt{\pi(\beta\delta)^{3/2}}} t^{-3/2} C_2(i, j) + o \left( \frac{e^{-rt}}{t^{3/2}} \right) $$

as $t \to \infty$, where $C_2(i, j)$ is a constant depending on $\rho$, $i$ and $j$ and $r = (\sqrt{\delta} - \sqrt{\beta})^2$.

- The remainder term decreases exponentially at rate $r$.
- For this reason $r^{-1}$ is denoted as *relaxation time*.
- It determines the time needed to relax in the steady state, when the initial condition $X_0 = i$ becomes unimportant.
Outline

1. Reflected Lévy Process
2. Time dependent properties of $M/M/1$
3. Waiting times and queue disciplines in $M/M/1$
Waiting times and workload in the FIFO case

- So far we studied the length of the queue.
- Now we will study the waiting times, with respect to the disciplines, which are harder to handle with.
- $P_e$ denotes the probability in the steady state.
- Let us define $M(k) = T_1 + \cdots + T_{k-1}$ and $W_k = V_{M(k)-}$.

Theorem 15 (The $M/M/1$ in the steady state).

The waiting time and the workload have a common distribution that is a mixture with weights $1 - \rho$, $\rho$ of an atom at 0 and an exponential distribution with intensity $\gamma = \delta - \beta$,

$$P_e(W_n \leq y) = P_e(V_t \leq y) = 1 - \rho + \rho(1 - e^{-\gamma y}) = 1 - \rho e^{-\gamma y}.$$
Customer vs. time stationarity

- It is necessary to distinguish between time and customer stationarity.
- Consider the time stationary version $\{ V_t^* \}_{t \geq 0}$ of the workload process.
- Then the waiting time of the first customer is $V_{T_1^-}^*$, where $T_1$ is the first arrival time.
- Obviously $V_{T_1^-}^* = (V_0^* - T_1)^+$ is smaller than the representative $V_0^*$.
- The thing is that we work with the information about the order of customers.
The LIFO case

- The distribution of the length of the queue in the fixed time $t$ is the same as in the FIFO case.
- It is same also before an arrival of other customer.
- This among others means that for $W_t = 0$ there is a weight $1 - \rho$.
- $W_k = 0$ if $n = 0$, where $n$ is the length of the queue.
- If $n > 0$, customer $k$ must wait for the server to finish the customer presently in front of him and to clear customers arriving later than customer $k$. 
The LIFO case

- His service can start at time
  \[ M(k) + W_k = \inf\{ t \geq M(k) : X_t = n - 1 \}. \]
- But this show that independently of \( n \) \( W_k \) is distributed as the time of doubly infinite queue from 0 to -1.

**Proposition 16.**

*Consider the LIFO \( M/M/1 \) queue in the steady state. Then*

\[
P_e(W_n \leq y) = 1 - \rho + \rho^{1/2} \int_0^y \frac{1}{t} e^{-(\beta+\delta)t} l_1 \, dt.
\]
The SIRO case

Once more the distribution of the length of the queue in the fixed time $t$ is the same as in the FIFO case.

Customer of type $n$ is a customer who meets $n$ other customers in the system upon arrival.

Let us denote $H_n(y)$ the probability that the waiting time of a customer of type $n$ strictly exceeds $y$.

With use of Taylor expansion it is possible to derive following theorem.
The SIRO case

Theorem 17 (The steady state $M/M/1$ SIRO waiting time distribution).

\[ H_n(y) = \sum_{k=0}^{\infty} h_n^{(k)} \frac{y^k}{k!}, \quad P_e(W_n > y) = \sum_{k=0}^{\infty} h_n^{(k)} \frac{y^k}{k!}, \]

where the $h_n^{(k)}$ is recursively determined by $h_0^{(k)} = 0$, $h_1^{(k)} = 1$ and

\[ h_n^{(k)} = H_n^{(k)}(0) = \delta h_{n-1}^{(k-1)} \frac{n-1}{n} - (\beta + \delta) h_n^{(k-1)} + \beta h_{n+1}^{(k-1)}, \]

further $h^{(k)} = \sum_{\infty} \pi_n h_n^{(k)}$. 

The PS case

- Because a service starts immediately after the customer comes, it does not make sense to investigate waiting time.
- Instead, we shall be interested in sojourn time $W_k^*$, it means how long the customer $k$ spends in the system.
- We always have $W_k^* \geq U_k$, where $U_k$ is a service time.
- $W_k = W_k^* - U_k$ may be interpreted as the delay caused by the possible presence of other customers.
- Finding of probability $P_e(W_k^* > y)$ may be transformed to the case of SIRO and the result is also similar.
S. Asmussen.
Applied probability and queues.

Z. Prášková, P. Lachout
Základy náhodných procesů.

P. Lachout.
Diskrétní martingaly.
Time for your questions
Thank you for your attention