数理解析研究所講究録 1263 巻 2002 年 30-34

An Iterative Estimation Procedure for Mixed Poisson Processes via EM Algorithm and Its Application to Queueing Systems

岡村寛之，土肥正
Hiroyuki Okamura and Tadashi Dohi
Department of Information Engineering, Graduate School of Engineering, Hiroshima University, Japan

Abstract. This paper proposes an iterative estimation procedure for mixed Poisson processes. Based on the EM (expectation-maximization) algorithm, we develop an efficient iteration method to derive the MLE (maximum likelihood estimate). The result is applied to estimating some performance measures for *M/G/1* queueing systems.

1 Introduction

The mixed Poisson process [1] can be regarded as a natural extension of the homogeneous Poisson process. Let $\{N_P(t);\ t \geq 0\}$ be a homogeneous Poisson process with arrival intensity λ (> 0). The probability mass function at time t is expressed by the following Poisson distribution:

$$\Pr\{N_P(t)=n\} = \frac{(\lambda t)^n}{n!}\exp(-\lambda t). \tag{1}$$

Let $\{N_M(t);\ t \geq 0\}$ denote a mixed Poisson process with random arrival intensity Λ (> 0). The probability density function of Λ is given by $g(\lambda;\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ denotes a set of parameters. Then, the probability mass function of the mixed Poisson process is

$$\begin{aligned}\Pr\{N_M(t)=n\} &= \int_0^\infty \Pr\{N_P(t)=n|\Lambda=\lambda\}g(\lambda;\boldsymbol{\theta})d\lambda \\ &= \int_0^\infty \frac{(\lambda t)^n}{n!}\exp(-\lambda t)g(\lambda;\boldsymbol{\theta})d\lambda, \end{aligned} \tag{2}$$

which is the mixed Poisson distribution. When the probability density function $g(\cdot)$ is the Erlang distribution with shape parameter m and scale parameter β, *i.e.* for $\boldsymbol{\theta}=(m,\beta)$,

$$g(\lambda;m,\beta) = \frac{\beta^m}{(n-1)!}\lambda^{m-1}\exp(-\beta\lambda), \tag{3}$$

it can be verified that

$$\Pr\{N_M(t)=n\} = \binom{m+n-1}{n}\left(\frac{\beta}{\beta+t}\right)^m\left(\frac{t}{\beta+t}\right)^n \tag{4}$$

and that $N_M(t)$ obeys a negative binomial distribution. This special stochastic process is called the Pólya–Lundberg process [1]. For example, the density function $g(\cdot)$ is an inverse Gaussian distribution, a beta distribution, an uniform distribution, *etc.* Table 1 presents typical examples of $g(\cdot)$ and the corresponding mixed distributions [1].

2 An estimation procedure

In this section, we propose an estimation procedure for the mixed Poisson process. To give flexibility for the estimation algorithm, we widely extend the stochastic process under consideration not to the homogeneous Poisson process, but to the non–homogeneous Poisson process. That is, the arrival time distribution is assumed to be $F(\cdot;\lambda,\boldsymbol{\xi})$, where λ (> 0) and $\boldsymbol{\xi}$ denote an arrival intensity and a parameter set of the others. Then, the total number of arrivals before time t follows the non–homogeneous Poisson process having the probability mass function;

$$\Pr\{N_P(t)=n\} = \frac{(-\log\overline{F}(t;\lambda,\boldsymbol{\xi}))^n}{n!}\overline{F}(t;\lambda,\boldsymbol{\xi}), \tag{5}$$

Table 1: Typical examples of mixed Poisson distributions.

types	mixed Poisson distributions
Erlang distribution (gamma distribution)	negative binomial distribution
shifted gamma distribution	Delaporte distribution
generalized inverse Gaussian distribution	Sichel distribution
inverse Gaussian distribution	inverse Gaussian–Poisson distribution
beta distribution	beta–Poisson distribution
uniform distribution	uniform–Poisson distribution
truncated normal distribution	truncated normal–Poisson distribution
lognormal distribution	lognormal–Poisson distribution

where $\overline{F}(\cdot;\lambda,\boldsymbol{\xi}) = 1-F(\cdot;\lambda,\boldsymbol{\xi})$. In a fashion similar to the homogeneous Poisson process, if the stochastic process $\{N_M(t),\ t \geq 0\}$ is a mixed non-homogeneous Poisson process with random parameter Λ, we obtain

$$\Pr\{N_M(t) = n\} = \int_0^\infty \frac{(-\log \overline{F}(t;\lambda,\boldsymbol{\xi}))^n}{n!} \overline{F}(t;\lambda,\boldsymbol{\xi}) g(\lambda;\boldsymbol{\theta}) d\lambda. \tag{6}$$

If $\overline{F}(t;\lambda) = \exp\{-\lambda t\}$, the process in Equation (6) can be reduced into the mixed Poisson process.

Consider an estimation problem under the assumption that all the occurrence times of events on m independent mixed Poisson processes are observed, namely, the following data set is available:

$$\mathcal{D}^i = (x_1^i, \ldots, x_{n_i}^i) \qquad \text{for } i = 1, \ldots, m.$$

The well-known method to estimate the model parameters $\boldsymbol{\xi}$ and $\boldsymbol{\theta}$ is the maximum likelihood (ML) estimation. For the mixed non-homogeneous Poisson process, the likelihood and the log-likelihood functions are given by

$$L(\boldsymbol{\xi},\boldsymbol{\theta}|\mathcal{D}^1,\ldots,\mathcal{D}^m) = \prod_{i=1}^m \left(\int_0^\infty \prod_{j=1}^{n_i} \frac{f(x_j^i;\lambda,\boldsymbol{\xi})}{\overline{F}(x_j^i;\lambda,\boldsymbol{\xi})} \overline{F}(x_{n_i}^i;\lambda,\boldsymbol{\xi}) g(\lambda;\boldsymbol{\theta}) d\lambda \right) \tag{7}$$

and

$$\log L(\boldsymbol{\xi},\boldsymbol{\theta}|\mathcal{D}^1,\ldots,\mathcal{D}^m) = \sum_{i=1}^m \left(\log \int_0^\infty \prod_{j=1}^{n_i} \frac{f(x_j^i;\lambda,\boldsymbol{\xi})}{\overline{F}(x_j^i;\lambda,\boldsymbol{\xi})} \overline{F}(x_{n_i}^i;\lambda,\boldsymbol{\xi}) g(\lambda;\boldsymbol{\theta}) d\lambda \right), \tag{8}$$

respectively. In general, the maximum likelihood estimate (MLE) can be found as the value maximizing Equation (8). However, in the case of the estimation for the mixed non-homogeneous Poisson process, since both the likelihood and the log-likelihood functions are not always simple forms, it is difficult to maximize them directly. This motivates us to develop an iterative estimation procedure based on the EM algorithm.

The EM algorithm is an iteration method for statistical estimation problems with incomplete data. Let X and $Y = u(X)$ be the unobserved random variable with probability density $f_X(\cdot;\boldsymbol{\alpha})$ and the observed random variable, respectively. Given the observed experiment $\boldsymbol{y}$, we estimate the parameter set $\boldsymbol{\alpha}$. The $(n+1)$-st step in the EM algorithm consists of finding $\boldsymbol{\alpha}_{n+1}$ which maximizes the expected log-likelihood function for the complete data, provided that the incomplete data is observed. That is,

$$\hat{\boldsymbol{\alpha}}_{n+1} = \underset{\boldsymbol{\alpha}}{\operatorname{argmax}} \left\{ \mathrm{E}[\log f_X(X;\boldsymbol{\alpha}) | u(X) = \boldsymbol{y}; \hat{\boldsymbol{\alpha}}_n] \right\}, \tag{9}$$

where $\hat{\boldsymbol{\alpha}}_n$ is the estimate for the n-th step in the EM algorithm and $\mathrm{E}[\cdot\ ;\boldsymbol{\alpha}]$ denotes the mathematical expectation operator with respect to the probability density $f_X(\cdot;\boldsymbol{\alpha})$.

Let us now return our concern to the estimation problem for the mixed non-homogeneous Poisson process. If the complete data like

$$(\mathcal{D}^i, \lambda_i) = (x_1^i, \ldots, x_{n_i}^i, \lambda_i) \qquad \text{for } i = 1, \ldots, m,$$

where λ_i is a sample for the density $g(\cdot;\boldsymbol{\theta})$ is given, the log-likelihood function is then derived as

$$\begin{aligned}\log L(\boldsymbol{\xi},\boldsymbol{\theta}|\mathcal{D}^1,\ldots,\mathcal{D}^m,\lambda_1,\ldots,\lambda_m) &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}\log f(x_j^i;\lambda_i,\boldsymbol{\xi}) - \sum_{i=1}^{m}\sum_{j=1}^{n_i-1}\log \overline{F}(x_j^i;\lambda_i,\boldsymbol{\xi}) \\ &\quad + \sum_{i=1}^{m}\log g(\lambda_i;\boldsymbol{\theta}). \end{aligned} \tag{10}$$

Since the samples, $\lambda_1,\ldots,\lambda_m$, cannot be observed, we consider the expected log-likelihood function under the incomplete data $\mathcal{D}^i$ $(i=1,\ldots,m)$ and find the parameter set maximizing the expected log-likelihood function. Therefore, at the $(n+1)$-st step in the EM algorithm, the estimates of parameters can be computed as follows.

$$\begin{aligned}\boldsymbol{\xi}^{(n+1)} &= \underset{\boldsymbol{\xi}}{\operatorname{argmax}}\Bigg\{\sum_{i=1}^{m}\sum_{j=1}^{n_i}\mathrm{E}[\log f(x_j^i;\Lambda_i,\boldsymbol{\xi})|\mathcal{D}^i;\boldsymbol{\xi}^{(n)},\boldsymbol{\theta}^{(n)}] \\ &\quad - \sum_{i=1}^{m}\sum_{j=1}^{n_i-1}\mathrm{E}[\log \overline{F}(x_j^i;\Lambda_i,\boldsymbol{\xi})|\mathcal{D}^i;\boldsymbol{\xi}^{(n)},\boldsymbol{\theta}^{(n)}]\Bigg\}, \end{aligned}\tag{11}$$

$$\boldsymbol{\theta}^{(n+1)} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\Bigg\{\sum_{i=1}^{m}\mathrm{E}[\log g(\Lambda_i;\boldsymbol{\theta})|\mathcal{D}^i;\boldsymbol{\xi}^{(n)},\boldsymbol{\theta}^{(n)}]\Bigg\}, \tag{12}$$

where $\boldsymbol{\xi}^{(n)}$ and $\boldsymbol{\theta}^{(n)}$ are the estimated parameter sets at the n-th step in the EM algorithm. We derive the useful formula to calculate the expected log-likelihood function. For any measurable function $h(\cdot)$, the following equation holds:

$$\mathrm{E}[h(\Lambda_i)|\mathcal{D}^i;\boldsymbol{\xi},\boldsymbol{\theta}] = \frac{\int_0^\infty h(\lambda)\big(\prod_{j=1}^{n_i} f(x_j^i;\lambda,\boldsymbol{\xi})/\overline{F}(x_j^i;\lambda,\boldsymbol{\xi})\big)\overline{F}(x_{n_i}^i;\lambda,\boldsymbol{\xi})g(\lambda;\boldsymbol{\theta})d\lambda}{\int_0^\infty \big(\prod_{j=1}^{n_i} f(x_j^i;\lambda,\boldsymbol{\xi})/\overline{F}(x_j^i;\lambda,\boldsymbol{\xi})\big)\overline{F}(x_{n_i}^i;\lambda,\boldsymbol{\xi})g(\lambda;\boldsymbol{\theta})d\lambda}. \tag{13}$$

The similar argument to the above can be also applied to the case of the counting data observed. Consider the case where the number of arrivals during a constant period $[0,t_i)$ $(i=1,\ldots,m)$ is observed, *i.e.*

$$\mathcal{D}^i = (n_i,t_i), \qquad \text{for } i=1,\ldots,m.$$

If the complete data $(\mathcal{D}^i,\lambda_i)$ is observed, the log-likelihood function is obtained as

$$\begin{aligned}\log L(\boldsymbol{\xi},\boldsymbol{\theta}|\mathcal{D}^1,\ldots,\mathcal{D}^m,\lambda_1,\ldots,\lambda_m) &= \sum_{i=1}^{m} n_i\log(-\log \overline{F}(t_i;\lambda_i,\boldsymbol{\xi})) - \sum_{i=1}^{m}\log(n_i!) \\ &\quad + \sum_{i=1}^{m}\log \overline{F}(x_j^i;\lambda_i,\boldsymbol{\xi}) + \sum_{i=1}^{m}\log g(\lambda_i;\boldsymbol{\theta}). \end{aligned}\tag{14}$$

In this case, since the arrival intensities, $\lambda_1,\ldots,\lambda_m$, are unobservable, we derive the expected log-likelihood function and find the parameter set maximizing it. Then we have the following iteration algorithm;

$$\begin{aligned}\boldsymbol{\xi}^{(n+1)} &= \underset{\boldsymbol{\xi}}{\operatorname{argmax}}\Bigg\{\sum_{i=1}^{m} n_i\mathrm{E}[\log(-\log \overline{F}(t_i;\Lambda_i,\boldsymbol{\xi}))|\mathcal{D}^i;\boldsymbol{\xi}^{(n)},\boldsymbol{\theta}^{(n)}] \\ &\quad + \sum_{i=1}^{m}\mathrm{E}[\log \overline{F}(x_j^i;\Lambda_i,\boldsymbol{\xi})|\mathcal{D}^i;\boldsymbol{\xi}^{(n)},\boldsymbol{\theta}^{(n)}]\Bigg\}, \end{aligned}\tag{15}$$

$$\boldsymbol{\theta}^{(n+1)} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\Bigg\{\sum_{i=1}^{m}\mathrm{E}[\log g(\Lambda_i;\boldsymbol{\theta})|\mathcal{D}^i;\boldsymbol{\xi}^{(n)},\boldsymbol{\theta}^{(n)}]\Bigg\}. \tag{16}$$

Further, the expected value in Equations (15) and (16) can be derived by

$$\mathrm{E}[h(\Lambda_i)|\mathcal{D}^i;\boldsymbol{\xi},\boldsymbol{\theta}] = \frac{\int_0^\infty h(\lambda)(-\log \overline{F}(t_i;\lambda,\boldsymbol{\xi}))^{n_i}/n_i!\overline{F}(t_i;\lambda,\boldsymbol{\xi})g(\lambda;\boldsymbol{\theta})d\lambda}{\int_0^\infty (-\log \overline{F}(t_i;\lambda,\boldsymbol{\xi}))^{n_i}/n_i!\overline{F}(t_i;\lambda,\boldsymbol{\xi})g(\lambda;\boldsymbol{\theta})d\lambda}. \tag{17}$$

For example, it is assumed that

$$f(t;\lambda) = \lambda \exp\{-\lambda t\}, \tag{18}$$

$$g(\lambda;\beta) = \beta \exp\{-\beta\lambda\}. \tag{19}$$

From Equation (4), the probability mass function is the following geometric distribution;

$$\Pr\{N_M(t) = n\} = \left(\frac{\beta}{\beta+t}\right)\left(\frac{t}{\beta+t}\right)^n. \tag{20}$$

Find the iteration algorithm for estimating the parameter β under two kinds of observed data;

$$\mathcal{D}_1^i = (x_1,\ldots,x_{n_i}) \qquad \text{and} \qquad \mathcal{D}_2^i = (n_i,t_i) \qquad \text{for } i=1,\ldots,m.$$

For the first data set, at the $(n+1)$-st step in the EM algorithm, we have

$$\hat{\beta}^{(n+1)} = \frac{m}{\sum_{i=1}^m \mathrm{E}[\Lambda_i|\mathcal{D}_1^i;\hat{\beta}^{(n)}]}. \tag{21}$$

From Equation (13), the expected value can be derived as

$$\begin{aligned}\mathrm{E}[\Lambda_i|\mathcal{D}_1^i;\beta] &= \frac{\int_0^\infty \lambda^{n_i+1}\exp\{-\lambda x_{n_i}\}\beta\exp\{-\beta\lambda\}d\lambda}{\int_0^\infty \lambda^{n_i}\exp\{-\lambda x_{n_i}\}\beta\exp\{-\beta\lambda\}d\lambda} \\ &= \frac{n_i+1}{x_{n_i}+\beta}.\end{aligned} \tag{22}$$

Consequently, the estimate can be computed by the following iteration scheme;

$$\hat{\beta}^{(n+1)} = \frac{m}{\displaystyle\sum_{i=1}^m \frac{n_i+1}{x_{n_i}+\hat{\beta}^{(n)}}}. \tag{23}$$

Next, consider the iteration algorithm under the data set $\mathcal{D}_2^i$ $(i=1,\ldots,m)$. In this case, the algorithm for estimating the parameter β is represented as

$$\hat{\beta}^{(n+1)} = \frac{m}{\sum_{i=1}^m \mathrm{E}[\Lambda_i|\mathcal{D}_2^i;\hat{\beta}^{(n)}]}. \tag{24}$$

Then, Equation (17) yields

$$\mathrm{E}[\Lambda_i|\mathcal{D}_2^i;\beta] = \frac{n_i+1}{t_i+\beta}. \tag{25}$$

Therefore, for the counting data set, the iteration algorithm is given by

$$\hat{\beta}^{(n+1)} = \frac{m}{\displaystyle\sum_{i=1}^m \frac{n_i+1}{t_i+\hat{\beta}^{(n)}}}. \tag{26}$$

3 Application

Consider an *M/G/1* queueing system with arrival rate λ and service density function $g(\cdot;\boldsymbol{\theta})$. Let $\{X_n; n \geq 0\}$ denote the queue length at departures. For any $n = 1,2,\ldots$, it can be seen that

$$\Pr\{X_n - X_{n-1} + 1 = k | X_{n-1} > 0\} = \int_0^\infty \frac{(\lambda t)^n}{n!}\exp\{-\lambda t\}g(t;\boldsymbol{\theta})dt, \tag{27}$$

and

$$\Pr\{X_n = k | X_{n-1} = 0\} = \int_0^\infty \frac{(\lambda t)^n}{n!}\exp\{-\lambda t\}g(t;\boldsymbol{\theta})dt. \tag{28}$$

In this section, we treat the estimation problem for the service parameter set $\boldsymbol{\theta}$ when both the arrival rate and the queue length at departures are observed. Given that

$$\mathcal{D} = (X_0 = x_0, X_1 = x_1, \ldots, X_m = x_m),$$

the estimation problem to obtain $\hat{\boldsymbol{\theta}}$ can be reduced to that for mixed Poisson processes under the data set;

$$\mathcal{D}^i = (x_i - x_{i-1} + \mathbf{1}_{\{x_{i-1}>0\}}) \qquad \text{for } i = 1, \dots, m,$$

where

$$\mathbf{1}_{\{x_i>0\}} = \begin{cases} 0 & \text{for } x_i = 0 \\ 1 & \text{for } x_i > 0. \end{cases}$$

Therefore, from the result in Section 2, the estimates can be calculated as follows.

E-Step:

$$\mathrm{E}[h(X_i)|\mathcal{D}^i; \boldsymbol{\theta}] = \frac{\int_0^\infty h(t)(\lambda t)^{n_i} \exp\{-\lambda t\} g(t; \boldsymbol{\theta}) dt}{\int_0^\infty (\lambda t)^{n_i} \exp\{-\lambda t\} g(t; \boldsymbol{\theta}) dt}. \tag{29}$$

M-Step:

$$\boldsymbol{\theta}^{(n+1)} = \arg\max_{\boldsymbol{\theta}} \left\{ \sum_{i=1}^{m} \mathrm{E}[\log g(X_i; \boldsymbol{\theta})|\mathcal{D}^i; \boldsymbol{\theta}^{(n)}] \right\}, \tag{30}$$

where $n_i = x_i - x_{i-1} + \mathbf{1}_{\{x_{i-1}>0\}}$.

For example, if $g(\cdot; \mu)$ is an exponential distribution with parameter μ (> 0), the underlying queueing system is an *M/M/1* system. Without any loss of generality, when $\lambda = 1$, we have the following iterative estimation algorithm for the data $\mathcal{D}^i$:

$$\hat{\mu}^{(n+1)} = \frac{m}{\displaystyle\sum_{i=1}^{m} \frac{n_i + 1}{1 + \hat{\mu}^{(n)}}}. \tag{31}$$

In this case, it is easy to verify that the above estimate converges to

$$\hat{\mu}^{(\infty)} = \frac{m}{\displaystyle\sum_{i=1}^{m} n_i}. \tag{32}$$

Since the arrival rate is fixed as 1, the first moment of service time is equivalent to the traffic intensity. Therefore, the estimate in Equation (32) satisfies our intuition.

4 Conclusion

In this paper, we have developed the iterative estimation procedure for mixed Poisson processes. The iteration method has been derived based on the EM algorithm. Since we have considered a mixture of non–homogeneous Poisson processes, our method proposed here can be widely applied to the estimation problem for the mixed Poisson process. In addition, for the parameter estimation in an *M/G/1* queueing system, we have presented the iteration algorithm for estimating the service parameters on the queue length data at departures.

References

[1] J. Grandell, *Mixed Poisson Processes*, Chapman & Hall, London, 1997.

[2] A. P. Dempster, N. M. Laird and D. B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, *J. Roy. Statist. Soc. Ser. B*, 39, pp. 1–38, 1977.