Lecture Notes in Management Science (2013) Vol. **5**: 46–56
5th International Conference on Applied Operational Research, Proceedings

ISSN 2008-0050 (Print), ISSN 1927-0097 (Online)

Modeling scheduled patient punctuality in an infusion center

Seunggyun Cheong [1], Robert R. Bitmead [2] and John Fontanesi [3]

[1] *Australian Centre for Field Robotics, School of Aerospace, Mechanical and Mechatronic Engineering, The University of Sydney, Australia*
s.cheong@acfr.usyd.edu.au

[2] *Department of Mechanical & Aerospace Engineering,*
University of California, San Diego, La Jolla CA, USA
rbitmead@ucsd.edu

[3] *Center for Management Science in Health,*
University of California, San Diego, La Jolla CA, USA
jfontanesi@ucsd.edu

Abstract. We introduce a new model structure for punctuality of scheduled patients. Each iteration of the model is a mixture of two exponential distributions, one for the punctuality of early-arriving patients and the other for late-arriving patients. Since patients' earliness and lateness are treated separately, the models can capture different characteristics of each while many other model structures such as normal distributions treat them symmetrically. The new model structure is tested on data collected in a hospital infusion room and demonstrates quantifiably better performance than normal distribution fitting. The approaches are compared using two goodness-of-fit measures. Further, patient punctuality is shown to vary throughout a day depending on patients' appointment times.

Keywords: scheduling; punctuality; exponential distribution

Introduction

The national goal to contain health care costs has generated multiple strategies to increase access, improve patient throughput, enhance patient satisfaction and provide high-caliber quality care. An example of the changes wrought is the dramatic expansion of ambulatory infusion treatment for a wide range of therapies including analgesics, narcotics, chemotherapy, and antibiotic or antiviral infusions. Each treatment is driven by a protocol designating the frequency, duration and staffing requirements for treatment. With over 750,000 individuals treated in ambulatory infusion centers 2005 and expected rate of increase in excess of 10%, understanding the behaviors of patients accessing their infusion treatment appointment is critical to matching demand to resources. Such strategies are contingent upon accurately anticipating patient appointment needs and dynamically matching those needs to health system resources as efficiently as possible. Doing so requires an understanding of variability in factors such as patient no-show, cancellation, punctuality, and treatment durations (Gupta and Denton, 2008). In this paper, we investigate punctuality behaviors of scheduled patients in a hospital infusion center with an appointment system in place.

Patient punctuality is often presumed to follow a normal distribution, e.g. Cayirli et al (2006), and, thus, its two-dimensional parameter, i.e. mean and variance, is computed to fit a probability density function (PDF) of a normal distribution to collected data. Other models have been proposed for punctuality, e.g. four-parameter Johnson distributions in Alexopoulos et al (2008), which also can capture asymmetric densities. By comparison, arrival and service data are often fitted with exponential distribution functions dependent on a single parameter. Naturally, if we fit a PDF with a higher-dimensional parameter to the same data, a better fit is expected. However, as a trade-off, the estimated parameters have a larger variance and the resulting estimated PDF captures potentially-embedded disturbance in data, which may not reappear in the future. Thus, to provide this balance, we employ the Akaike Information Criterion (AIC) and compare the estimate with a normal distribution to a new estimate with a three-dimensional distribution based on a two-sided mixture of exponential distributions. This new estimate is designed to treat early-arriving patients and late-arriving patients separately and shows the different behaviors of each category. The idea of modeling patient punctuality with a mixture of more than one distribution is introduced in Tai and Williams (2009) but this is focused on finding a model with a better fit to collected data rather than understanding and interpreting collected data using appropriate models, which is the focus of this paper.

Patients arrive early for different reasons than they arrive late. Some patients have earlier treatments in different departments in the same hospital so that they check in right after their earlier treatments regardless of their appointments. Other patients visit early for a potential early-admission and wait for any available time slot produced by no-show patients. On the other hand, aspects of late arriving patients depend on other factors such as unexpected traffic congestion and delayed same-day appointments.

In addition, these early and late patterns vary throughout a day. Patients with early appointments are unlikely to have earlier treatments in the same hospital and their earliness may have a similar pattern to their lateness. Patients with later appointments may demonstrate a tendency for early arrival.

In Section 2, we describe how the data of patient punctuality, to which our models are fitted, were collected. In Section 3, we introduce a new three-dimensional distribution for modeling the patients' punctuality and study evaluation measures for comparing distributions. In Section 4, a quantitative comparison is made between the new distribution and a normal distribution based on the empirical goodness-of-fit measures. Conclusions and future work follow in Section 5.

Data collection

The Infusion Center at the University of California, San Diego, Moores Cancer Center provides patients with chemotherapy infusion treatments. 3162 patients were treated in the center in November of 2012 alone and the demand for the treatments is increasing by 16% annually. The center is operated 7 days a week from 07:30 to 18:30. The center manages appointment time slots every 30 minutes between 07:30 and 18:30, i.e. at 07:30, 08:00, …, and 18:30, and each patient makes an appointment at one of the time slots before visiting the center so that there are no walk-in patients.

Patients' appointment times and arrival times were observed between January 1st and June 30th in 2012. Since our focus is their punctuality, we ignore the records of patients who did not turn up and the total number of observations of patient visits is 16410. There were multiple visits by the same patients but, for simplicity, we treat, in this paper, each visit as an independent visit by a different patient.

We analyze the data of the patient punctuality for each appointment time slot separately in order to investigate the dependency of the patient punctuality on the time slots. Given a time slot, denote by N the number of patients who had appointments at the time slot and turned up. The punctuality of the patients is represented by

$$x_n = (\text{Arrival time of patient } n) - (\text{Appointment time of patient } n)$$
$$\forall n \in \{1, \dots, N\}.$$

Since the center opens at 07:00, the earliest possible arrival time of patients is 07:00. With these data, an empirical cumulative distribution function (CDF) for the patients' punctuality corresponding to each time slot can be obtained as

$$F_N(x) = \frac{1}{N}\sum_{n=1}^{N} I_{x_n \le x} \quad (1)$$

where indicator function $I_{x_n \le x}$ is 1 if $x_n \le x$ and 0 otherwise. In Section 4, these CDFs are compared to models developed in Section 3.

Models for the patient punctuality

A model structure

For modeling the patient punctuality, we consider a model structure, M_e, whose probability density function (PDF) is parametrized as

$$f(x, \theta) = \begin{cases} a\lambda_1 e^{\lambda_1 x}, & x \le 0 \\ (1-a)\lambda_2 e^{-\lambda_2 x}, & x > 0 \end{cases} \quad (2)$$

For $\theta \in \Theta_e$ where $\Theta_e = \{\theta = [a \;\; \lambda_1 \;\; \lambda_2]^T \mid 0 \leq a \leq 1,\; \lambda_1, \lambda_2 \geq 0\}$. Unlike other model structures such as normal distributions, this model structure represents early-arriving patient punctuality and late-arriving patient punctuality with a mixture of two different exponential distributions: λ_1 is the exponential parameter of early arrival, λ_2 is the exponential parameter of late arrival, and $a \in [0,1]$ is the mixture parameter capturing the proportion of each distribution. Thus,

$$\int_{-\infty}^{0} f(x,\theta)dx = a\lambda_1 \left[\frac{1}{\lambda_1} e^{\lambda_1 x}\right]_{-\infty}^{0} = a \quad \text{and} \quad \int_{0}^{\infty} f(x,\theta)dx$$
$$= (1-a)\lambda_2 \left[\frac{1}{-\lambda_2} e^{-\lambda_2 x}\right]_{0}^{\infty} = 1 - a.$$

The three-parameter cumulative distribution function (CDF) of $f(x,\theta)$ is

$$F(x,\theta) = \begin{cases} a e^{\lambda_1 x}, & x \leq 0 \\ 1 - (1-a)e^{-\lambda_2 x}, & x > 0 \end{cases}.$$

Akaike information criterion (AIC)

Given the punctuality data $x_1, \dots, x_N$, we seek to fit the best model in the model set of interest in the previous section. For an evaluation measure, we seek to minimize the AIC

$$V(M,\theta) = \frac{1}{N}\left\{\sum_{i=1}^{N} -\ln f(x_n,\theta) + \dim\theta\right\} \tag{3}$$

(Ljung, 1999) where M is a set of PDFs whose elements are described by $f(x,\theta)$, $\theta \in \Theta$ with an appropriate parameter set Θ. Thus, M represents a model structure, i.e. a structure of a PDF, Θ is a set of all possible parameters in M, and $\dim\theta$ is the dimension of θ. The AIC is based on the Kullback-Leibler information (See, for example, deLeeuw, 1992) and can act both as an estimation criterion and as a model selection criterion.

We seek a PDF that minimizes the AIC in (3). This criterion for density estimation is the same as maximum likelihood estimation (MLE) for a model structure comprised of the same dimensional parameters. The density estimate associated with (3) and the model structure M_e in Section 3.1 leads to

$$\hat{\theta}(M_e) = \begin{bmatrix}\hat{a} & \hat{\lambda}_1 & \hat{\lambda}_2\end{bmatrix}^T = \begin{bmatrix} \frac{N^-}{N} & \frac{N^-}{\sum_{\substack{i=1 \\ x_n \leq 0}}^{N} (-x_n)} & \frac{N^+}{\sum_{\substack{i=1 \\ x_n > 0}}^{N} x_n} \end{bmatrix}^T \tag{4}$$

where N^- and N^+ are the numbers of non-positive and positive x_n's, respectively, and, hence, satisfy $N^- + N^+ = N$. The proof is presented in the Appendix. Clearly, the estimated parameters $\hat{\alpha}$, $\hat{\lambda}_1$, and $\hat{\lambda}_2$ signify the proportion of the early-arriving patients among all arriving patients, the reciprocal of the average earliness, and the reciprocal of the average lateness, respectively.

Model evaluation

The density estimation in the previous section can be performed over different model structures, which leads to different estimates of the patient punctuality density function. This brings a need for goodness-of-fit measure across model structures. There are many evaluation measures of model structures and, in this paper, we consider two, the AIC and the Kolmogorov-Smirnov (K-S) statistic, and apply these measures to the fitted two-sided exponential distributions in Section 3.1 and the fitted normal distributions.

Note, the AIC in (3) can be used as a tool to find an optimal parameter in a given model structure M, as in the previous section, and, also, can be used to compare different model structures by comparing the AIC values of the best models from each model structure.

The K-S statistic is given by

$$D_N = \sup_x \left| F_N(x) - F\left(x, \hat{\theta}\right) \right|$$

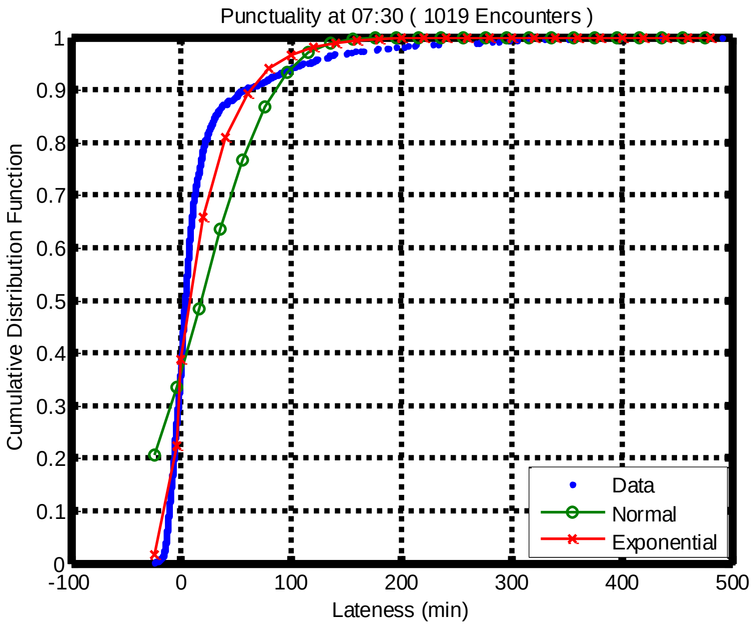
where $F_N(x)$ is the empirical CDF in (1) and $F\left(x, \hat{\theta}\right)$ is the CDF of a model corresponding to a parameter estimate $\hat{\theta}$. It follows that

$$D_N = \max_{n \in \{1,\ldots,N\}} \max \left\{ \left| \frac{1}{N} \sum_{i=1}^{N} I_{x_i < x_n} - F\left(x_n, \hat{\theta}\right) \right|, \left| \frac{1}{N} \sum_{i=1}^{N} I_{x_i \leq x_n} - F\left(x_n, \hat{\theta}\right) \right| \right\},$$

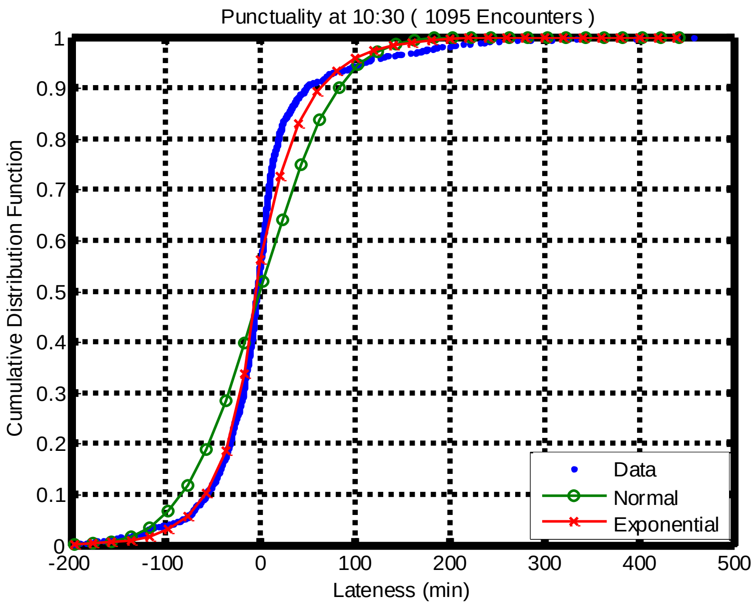
which can be computed efficiently.

Results of density estimation and model comparison

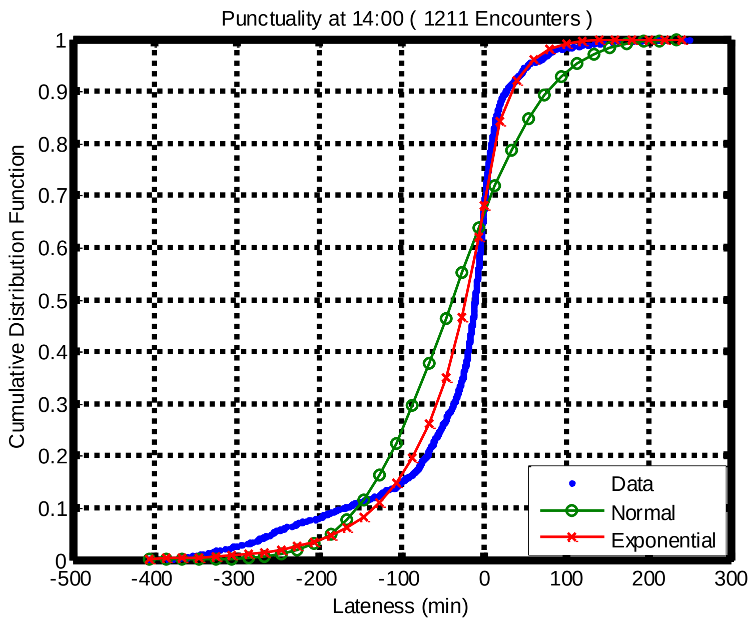
In this section, we apply the density estimation presented in the previous section to the punctuality data collected in the infusion center. Figure 1 shows empirical CDFs and estimated CDFs corresponding to normal distributions and two-sided exponential functions at four time slots.



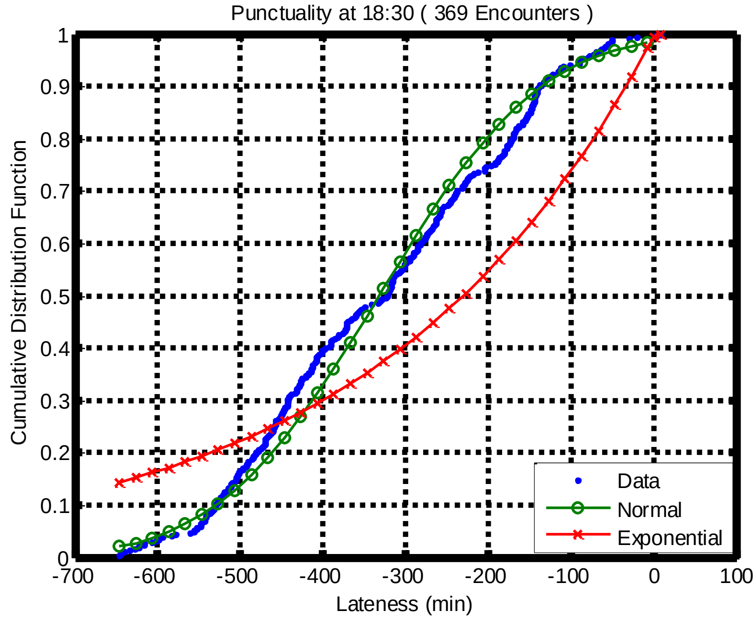
(a) At 07:30



(b) At 10:30



(c) At 14:00



(d) At 18:30

Fig. 1. CDF estimates for the patient punctuality at four appointment time slots.

For time slots where the data richly contain both patient earliness and lateness such as (a), (b), and (c) in Figure 1, the two-sided exponential distributions are lucidly better fits than the normal distributions and this can also be confirmed by comparing their corresponding PDFs. For the 11:30 time slot, PDF estimates are shown in Figure 2 with the data described by a histogram.

At 18:30, almost all data correspond to patients arriving early to which the normal distribution assigns 2 parameters and the two-sided exponential distribution assigns a single parameter. Thus, it is natural that the normal distribution fits the data at 18:30 better.

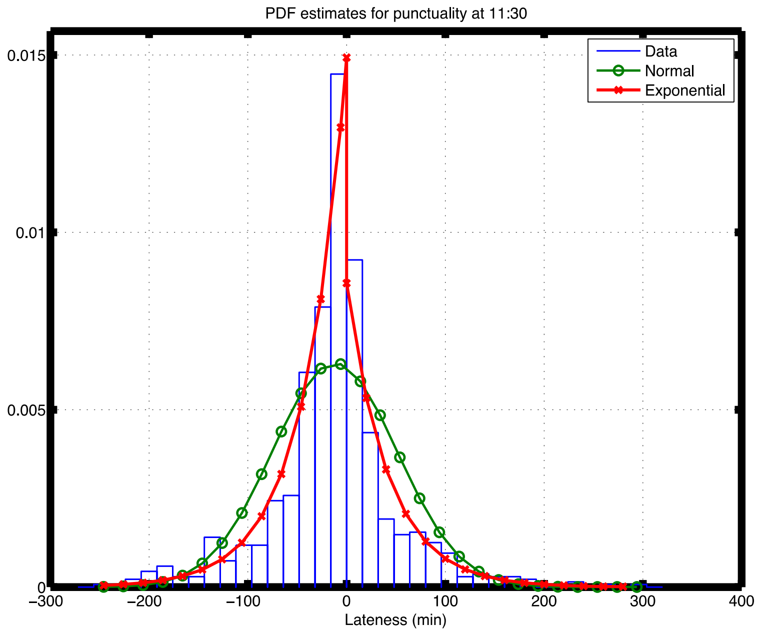
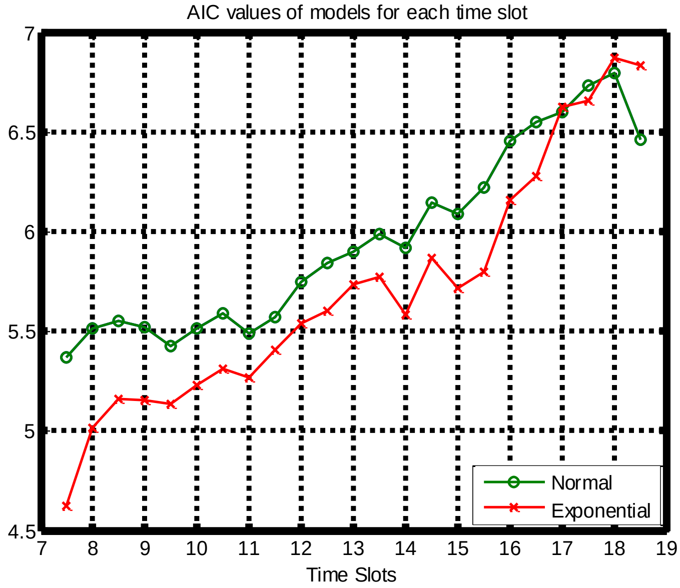
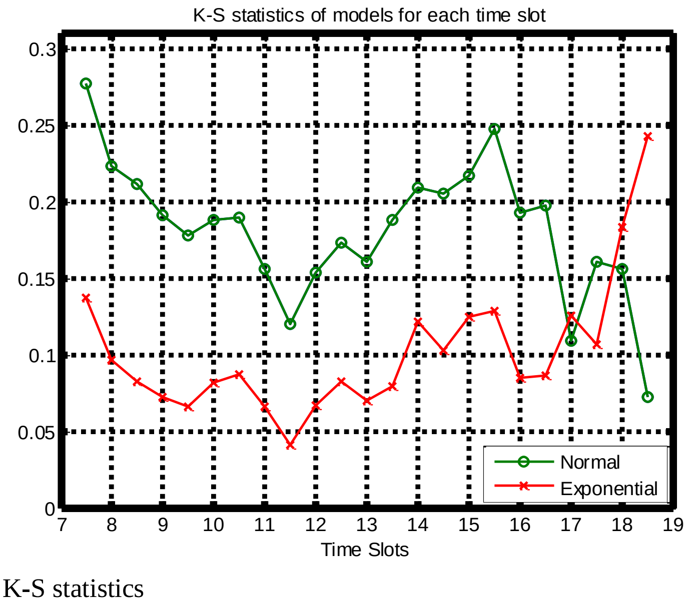


Fig. 2. PDF estimates for the patients' punctuality at 11:30

For quantified comparisons, the model structures for each time slot are compared using the AIC values and the K-S statistics in Figure 3. This also confirms that the two-sided exponential distribution provides better fits except at 17:00, 18:00, and 18:30. Also, the AIC values for both distributions become larger for later appointment time slots, which may imply that the patient punctuality deviates from both model structures and a better-suited model structure is needed to describe this punctuality better.



(a) AIC values



(b) K-S statistics

Fig. 3. Evaluation values of models at each time slot

From the density estimation with a structure of two-sided exponential distributions, the estimated parameter $\hat{a}$ increases, which means that the percentage of early-arriving patients increases and, at 18:30, almost every patient arrives earlier than their appointment time. The corresponding changes over time slot of $\hat{\lambda}_1$ and $\hat{\lambda}_2$, respectively, indicate that patients arrive relatively earlier for later appointments.

Density estimation with a normal distribution indicates several properties. The mean value is positive at 07:30, which implies tendency of lateness, and becomes smaller for later appointment time slots. The tendency of earliness is clear for the afternoon appointments from the negative mean values. The variance steadily becomes larger for a later appointment times.

Conclusion

This paper investigates patient punctuality by looking into patient earliness and lateness separately. The patterns of arriving early and late are modeled by a mixture of two one-sided exponential distributions and these models are shown to provide better fits to the collected data than normal distributions except where the effect of one of the parameters in the two-sided exponential distributions is diminished. Two evaluation measures are used for quantified comparison. Further, the patterns of the patient punctuality are shown to vary throughout a day. More accurate models of patient punctuality should engender more effective scheduling strategies.

For future work, other distribution models will be compared to the model in this paper and the patient punctuality will be investigated on a long-term basis rather than looking at the variations only in one day so that we can identify variations over a week, e.g. different patterns in weekdays and weekends potentially affected by traffic. Also, a no-show pattern can be included in the models by adding 1 more parameter.

References

Alexopoulos, C, Goldsman, D, Fontanesi, J, Kopald, D and Wilson JR (2008). Modeling patient arrivals in community clinics. Omega 36: 33-43

Cayirli, T, Veral, E and Rosen, H (2006). Designing appointment scheduling systems for ambulatory care services. Health Care Manage Sci 9: 47-58

deLeeuw, J (1992). Introduction to Akaike (1973) information theory and an extension of the maximum likelihood principle. In ‘Breakthroughs in Statistics. Vol.1’. pp 599-609. Springer-Verlag, New York

Gupta, D and Denton, B (2008). Appointment scheduling in health care: challenges and opportunities. IIE Transactions 40: 800-819

Ljung, L (1999). System Identification: Theory for the User. 2nd edition. Prentice-Hall, Upper Saddle River, New Jersey, USA

Tai, G and Williams, P (2009). A novel characterization of patient arrival process in healthcare facilities. The 35th Conference on Operational Research Applied to Health Services, Leuven, Belgium, July, 2009.

Appendix

A simple proof of (4) is given below. From (2) and (3), it follows that

$$\hat{\theta}(M_e) = \arg\min_{\theta\in\Theta_e} V(M_e,\theta) = \arg\min_{\theta\in\Theta_e} \frac{1}{N}\left\{\sum_{i=1}^{N} -\ln f(x_n,\theta) + \dim\theta\right\}$$

$$= \arg\min_{\theta\in\Theta_e} \sum_{i=1}^{N} \begin{cases} -\ln a - \ln\lambda_1 - \lambda_1 x_n, & x_n \le 0 \\ -\ln(1-a) - \ln\lambda_2 + \lambda_2 x_n, & x_n > 0 \end{cases}$$

$$= \arg\min_{\theta\in\Theta_e} -N^- \ln a - N^+ \ln(1-a) - N^- \ln\lambda_1 + \lambda_1 \sum_{\substack{i=1 \\ x_n\le 0}}^{N} (-x_n) - N^+ \ln\lambda_2 + \lambda_2 \sum_{\substack{i=1 \\ x_n>0}}^{N} x_n$$

where: $\Theta_e = \{\theta = [a \;\; \lambda_1 \;\; \lambda_2]^T \mid 0 \le a \le 1,\ \lambda_1,\lambda_2 \ge 0\}$, which represents the model structure M_e, and N^- and N^+ are the numbers of non-positive and positive x_n's, respectively. The third equality comes from the fact that $\dim\theta$, which is a constant 3, and $1/N$ have no effect on the minimization. This makes the estimate $\hat{\theta}(M_e)$ the same as the MLE. Then, it is straightforward to show that the estimate $\hat{\theta}(M_e)$ is given by (4).