Simulating Coxian Phase-type Distributions for Patient Survival

Published in:
International Transactions in Operational Research

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Simulating Coxian phase-type distributions for patient survival

Adele H. Marshall[a] and Mariangela Zenga[b]

[a]*Department of Applied Mathematics and Theoretical Physics, Queen's University of Belfast, Belfast, Northern Ireland,*
[b]*Dipartimento di Metodi Quantitativi per le Scienze Economiche ed Aziendali, Università di Milano-Bicocca, Milan, Italy*
E-mails: a.h.marshall@qub.ac.uk, mariangela.zenga@unimib.it

Abstract

Coxian phase-type distributions are a special type of Markov model that can be used to represent survival times in terms of phases through which an individual may progress until they eventually leave the system completely. Previous research has considered the Coxian phase-type distribution to be ideal in representing patient survival in hospital. However, problems exist in fitting the distributions. This paper investigates the problems that arise with the fitting process by simulating various Coxian phase-type models for the representation of patient survival and examining the estimated parameter values and eigenvalues obtained. The results indicate that numerical methods previously used for fitting the model parameters do not always converge. An alternative technique is therefore considered. All methods are influenced by the choice of initial parameter values. The investigation uses a data set of 1439 elderly patients and models their survival time, the length of time they spend in a UK hospital.

Keywords: Markov processes; stochastic models; survival analysis; statistical distributions; health services; elderly healthcare

1. Introduction

Coxian phase-type distributions (Cox, 1955) are a special type of Markov model that can be used to represent duration of time as a stochastic process consisting of phases through which elements in the model progress until they leave the system completely at any stage into a final absorbing phase.

Previous research has used Coxian phase-type distributions to represent survival times as the length of time until a certain event occurs, where the phases are considered to be stages in the survival and the absorbing, final stage, the event that occurs causing the individual or element to leave the system completely. For instance, this event could be a patient recovering from an illness,

a patient having a relapse, an individual leaving a certain type of employment, a piece of equipment failing, or a patient dying. Faddy (1994) illustrates how useful the Coxian phase-type distributions are in representing survival times for various applications such as the length of treatment spell of control patients in a suicide study, the time prisoners spend on remand and the lifetime of rats used as controls in a study of ageing.

In particular, Faddy and McClean (1999) used the Coxian phase-type distribution to find a suitable distribution for the duration of stay of a group of male geriatric patients in hospital. They found that the phase-type distributions were ideal for measuring the lengths of stay of patients in hospital and showed how it was also possible to consider other variables that may influence duration. More recently, Marshall and McClean (2003) have demonstrated how the Coxian phase-type distribution can, unlike alternative approaches, adequately model the survival of various groups of elderly patients in hospital uniquely capturing the typical skewed nature of such survival data in the form of a Conditional phase-type model (C-Ph) which incorporates a Bayesian network of inter-related variables.

The Coxian phase-type distribution is in fact a sub-class of the more general phase-type distribution, very similar in character except that the transient states (or phases) of the model are ordered in the Coxian phase-type distribution. The benefit of such a feature in the Coxian phase-type distribution is the reduction in the number of parameters required to represent the distribution therefore easing the process of parameter estimation. However, the fitting of appropriate Coxian phase-type distributions can also lead to problems.

The focus of this paper is to investigate the fitting process, how this has been carried out in the past and the potential problems that may arise. The process is examined by simulating various Coxian phase-type models, and considering the fitted distributions for representation of patient survival.

The investigation uses a data set of 1439 elderly patients and models their survival time, the length of time they spend in a Geriatric department in a UK hospital.

2. Coxian phase-type distributions

A Coxian phase-type distribution $\{X(t); t \geq 0\}$ may be defined as a (latent) Markov chain in continuous time with states $\{1, 2, \ldots n, n+1\}$, $X(0) = 1$, and for $i = 1, 2, \ldots n-1$

$$\text{prob}\{X(t + \delta t) = i + 1 | X(t) = i\} = \lambda_i \delta t + o(\delta t) \tag{1}$$

and for $i = 1, 2, \ldots n$

$$\text{prob}\{X(t + \delta t) = n + 1 | X(t) = i\} = \mu_i \delta t + o(\delta t). \tag{2}$$

Here, states $\{1, 2, \ldots n\}$ are latent (transient) states of the process and state $(n+1)$ is the (absorbing) state. λ_i represents the transition from state i to state $(i+1)$ and μ_i the transition from state i to the absorbing state $(n+1)$ (Fig. 1).

The process begins in the first phase and may either progress through the phases sequentially or enter into the absorbing state (the terminating event – phase $n+1$). Such phases may then be used to describe stages of a process which terminates at some stage. For example, in the case of the process of the duration of stay of elderly patients in hospital, transitions through the ordered
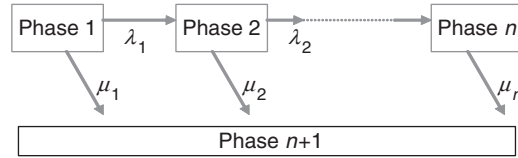
Fig. 1. An illustration of phase-type distributions.

transient states could correspond to various stages in the patient's stay such as stages of diagnosis, assessment, rehabilitation and long-stay care where patients eventually are absorbed when they leave the hospital completely due to discharge, transfer or death (Faddy, 1994).

Cox and Miller (1965) develop the theory of Markov chains such as those defined by (1) and (2). The Coxian phase-type distribution is defined as having a transition matrix $\mathbf{Q}$ of the following form,

$$\mathbf{Q} = \begin{pmatrix} -(\lambda_1 + \mu_1) & \lambda_1 & 0 & \ldots & 0 & 0 \\ 0 & -(\lambda_2 + \mu_2) & \lambda_2 \ldots & 0 & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & .. & . & . \\ 0 & 0 & 0 & \ldots & -(\lambda_{n-1} + \mu_{n-1}) & \lambda_{n-1} \\ 0 & 0 & 0 & \ldots & 0 & -\mu_n \end{pmatrix}, \tag{3}$$

where the λ_i's and μ_i's are from Cox and Miller's (1965) theory of Markov chains defined by (1) and (2).

The *probability density function* (pdf) of the Coxian phase-type distribution for time t is given by

$$f(t) = \mathbf{p} \exp\{\mathbf{Q}t\}\mathbf{q}, \tag{4}$$

where $\mathbf{p}$ is a $1 \times n$ vector of probabilities defining the initial transient phases and $\mathbf{q}$ is a $n \times 1$ vector of rates from the transient states to the absorbing state, as shown in (5) and (6), respectively.

$$\mathbf{p} = (1\,0\,0\ldots0\,0) \tag{5}$$

and

$$\mathbf{q} = -\mathbf{Q}\mathbf{1} = (\mu_1\,\mu_2\ldots\mu_n)^{\mathrm{T}}. \tag{6}$$

Previous research has fitted the pdf in equation (4) to various data sets of patient survival times in hospital and found that it can be used to consider patients' survival as a continuous variable notionally partitioned into stages or phases of time. A patient will enter hospital at time zero thus commencing a stay in hospital in phase 1. The patient spends a number of days in hospital until he is discharged or dies upon which stage he leaves the system completely at a rate of μ_1, or he continues his stay in hospital and moves into the next stage or phase of stay, phase 2. The movement from phase 1 to phase 2 of duration of stay is represented by the rate, λ_1. The final transition is when the patient leaves the hospital completely thus reaching the absorbing state of the system. Phase 1 is generally considered as a state containing more of the acute shorter stay cases with long stay patients more likely to continue their stay into the second phase, comprising

more of the those elderly patients who eventually transfer to a nursing home (Marshall and McClean, 2003).

3. Fitting Coxian phase-type distributions

As previously noted, phase-type distributions have not always been considered a flexible means of data representation due to the difficulties faced with estimating all the parameters of the distribution. As it is impossible to find an exact solution, a numeric algorithm is required (Neuts, 1981). The implementation of such efficient numerical procedures for estimating the distribution parameters remains an open problem that sometimes limits the use of this class of distribution in applications. In order to overcome this difficulty, many studies have defined restrictions of the Coxian representation; see for example Asmussen et al. (1996), Augustin and Buscher (1982), Bux and Herzog (1977) and Faddy (1994, 1998, 2002).

The interesting nature of the phase-type representation of the Coxian distribution has inspired many authors to continue to seek methods to solve the parameter estimation problem. Over the years, several ways have been developed in pursuit of an ideal solution, of which three key techniques have evolved.

The first of such methods is based on the maximization of the log-likelihood, or the minimization of (−)log-likelihood:

$$-\sum_{i=1}^{n} \log[\mathbf{p}\exp\{\mathbf{Q}t_i\}\mathbf{q}] \tag{7}$$

with respect to the parameters. Bobbio and Cumani (1992) suggested solving (7) by an iterative linearization method of estimation. In particular, they proposed a technique based on the canonical representation of acyclic phase-type (ACP), which is suited for the interpretation of censored data samples. Asmussen et al. (1996) proposed a fitting procedure for the complete class of phase-type distribution, based on the EM algorithm of Dempster et al. (1977) for incomplete data developing the EMPht program to estimate the distribution parameters. Faddy (1994, 1998) utilised the Nelder and Mead algorithm (1965) in the MATLAB (2001) software to find the solution of (7), and remedied the problem of the non-convergence of the algorithm by introducing penalized likelihoods (Faddy, 2002).

The second group of algorithms are based on moment matching. In a series of papers, Johnson and Taaffe (1989, 1990a, b, 1991) show how the first k (finite) moments of any non-degenerate distribution can be matched by a mixture of Erlang distributions of common order. The package MEDA, developed by Schmickler (1992), works on a mixture of Erlang distributions and implements a criterion based on the matching of the first moment combined with the minimization of a deviation measure.

The third method used to fit the phase-type distributions is based on minimum distance (MD). Parr and Schucany (1980) proposed a MD estimation procedure on the restricted class of ACP distributions. The chosen distance to minimize was the distance of Kolmogorov–Smirnov goodness-of-fit test. However, the most used method out of all of these techniques was based on the least squares.

In general, some authors compared the previously discussed methods concluding that, when the original distribution is unknown, the method of the moments is the least effective (Riska et al., 2002).

4. Medical data set

This paper considers a data set containing information collected during 2002–2003 to assist in the management of patients in a UK department of geriatric medicine. There are 1439 patient records which include the patient duration of stay (in days) in hospital. Patients were admitted to the Geriatric Department where they stayed, receiving treatment for a period of time before leaving the hospital either to return home, leaving to transfer to another hospital for care or leaving due to death. Length of stay in hospital is measured from the patient's time of admission to the time of departure. It is a continuous variable with mean 24 days, median 13 days and mode 8 days. The distribution of patient duration of stay in hospital, as illustrated in Fig. 2, is highly skewed consisting of a large peak at the start of the distribution that gradually tails off as duration increases. This skewness is also confirmed by the mean, median and mode which are not similar in value.

The data were fitted to a Coxian phase-type distribution using the maximum likelihood function and the Nelder-Mead simplex algorithm to perform parameter estimation. The sequential procedure adopted by Faddy (1998), starts with a 1-phase distribution, $n = 1$
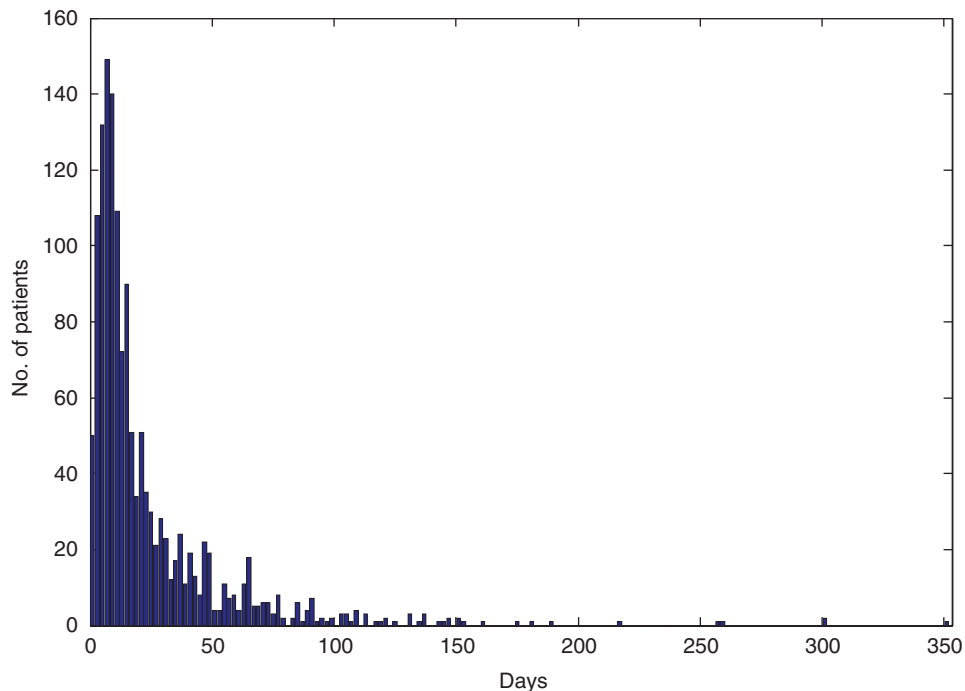

Fig. 2. Patient length of stay in hospital.

(corresponding to the exponential distribution) and considers such a representation for the patient duration of stay. It then continues the fitting process trying to find a better fit for the data by increasing the number of n phases until little improvement in the fit to the data can be obtained by adding a new phase.

When fitting the Coxian phase-type distribution to this data set of patient duration of stay, difficulties arose when estimating the parameters for more than three phases. In particular, choosing the initial values in order to guarantee convergence appeared to be problematic. This inspired the following investigation.

5. Simulating fits for the parameters of the Coxian phase-type distribution

The fitting procedure for the Coxian phase-type distribution is investigated by generating 1300 samples of initial values of the unknown parameters and using these as initial estimates for fitting the 4-phase and 5-phase models of patient duration of stay in hospital. For the 4-phase model, the samples will consist of seven units, the initial values for λ_1, λ_2, λ_3 and μ_1, μ_2, μ_3, μ_4, while for the 5-phase model, every sample is made of nine units, the initial values for λ_1, λ_2, λ_3, λ_4 and μ_1, μ_2, μ_3, μ_4, μ_5. The initial values were generated from uniform distributions and the length of stay distribution of elderly people in hospital fitted to a Coxian phase-type distribution using the Maximum Likelihood method. In particular, it is wished to minimize equation (7) subject to the following:

$$\begin{cases} 0 \leqslant \mu_j \leqslant 1, & j = 1, \ldots, n \\ 0 \leqslant \lambda_j \leqslant 1, & j = 1, \ldots, (n-1) \end{cases} \tag{8}$$

using a Sequential Quadratic Programming (SQP) method. In this method, a Quadratic Programming (QP) subproblem is solved at each iteration along with an update on the estimate of the Hessian of the Lagrangian. A full description of this algorithm is found in the "Constrained Optimization" section of the Introduction to Algorithms chapter of the MATLAB (2001) toolbox.

The choice of this particular algorithm, in comparison with the simplex method used by other authors (Faddy and McClean, 1999), is its ability to always converge, even when results are enforced.

6. Results

The 1300 samples of initial parameter values for the 4-phase Coxian distribution, as expected, yielded different fitted results. In fact, it was found that the representation was not unique, that it was possible for the estimation procedure to obtain equivalent solutions. The 1300 samples of initial parameter values used for fitting the 5-phase Coxian distribution also lacked an agreed fit. The best model for the 4-phase Coxian distribution is illustrated in Fig. 3. The $-$loglikelihood value is 5888.98, the smallest of those obtained from the 1300 samples of initial parameter values. The mean and median values for expression (7) are 5906.4 and 5889.0994, respectively.

Table 1 displays the average values and confidence intervals for each of the seven parameters.
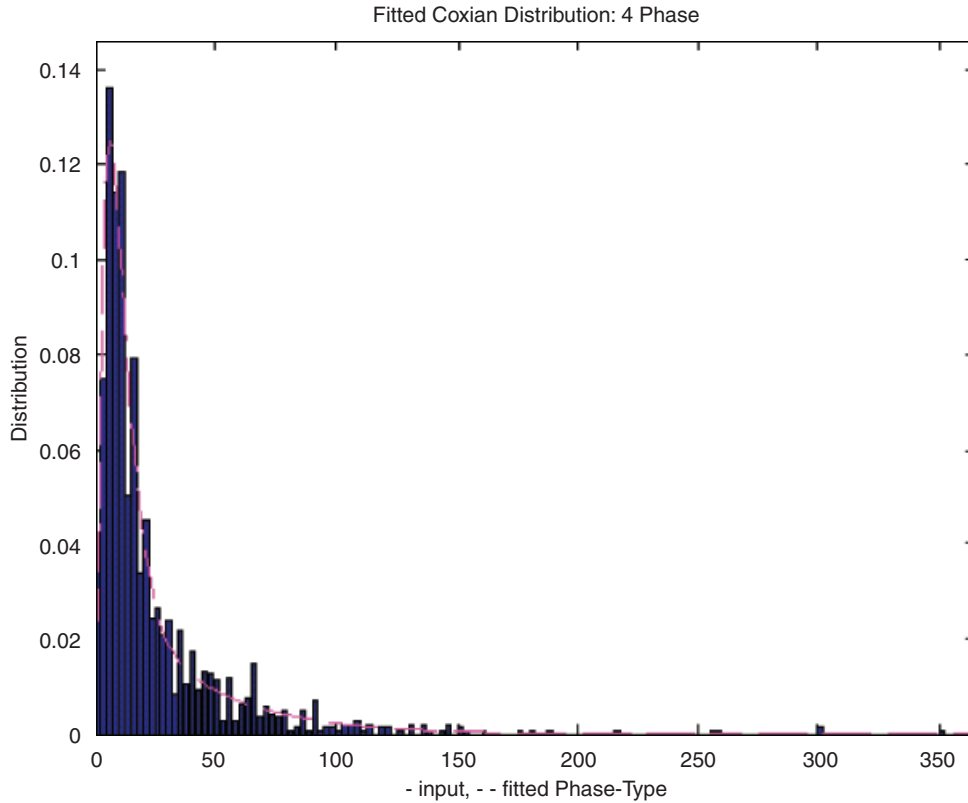
Fig. 3. The best fit for the 4-phase Coxian model.

Table 1
Average values and confidence intervals for parameters (four-phases model)

Parameter	Mean	Confidence interval $(1 - \alpha) = 0.99$
μ_1	0.00107	(0.00057, 0.00156)
μ_2	0.02678	(0.02181, 0.03176)
μ_3	0.33290	(0.31223, 0.35357)
μ_4	0.13114	(0.11149, 0.15079)
λ_1	0.34901	(0.32860, 0.36942)
λ_2	0.34133	(0.32021, 0.36245)
λ_3	0.20926	(0.19467, 0.22384)

In the case of the 5-phase Coxian model, the best representation has a ($-$loglikelihood) value equal to 5884.2, significantly lower than that of the 4-phase Coxian fit as shown by the χ^2 goodness-of-fit test at the 90% level. The Akaike Information Criterion (AIC) penalizes for the addition parameters, and thus selects a model that fits well but has a minimum number of parameters based on simplicity and parsimony (Akaike, 1974). Using the AIC criterion, it appears
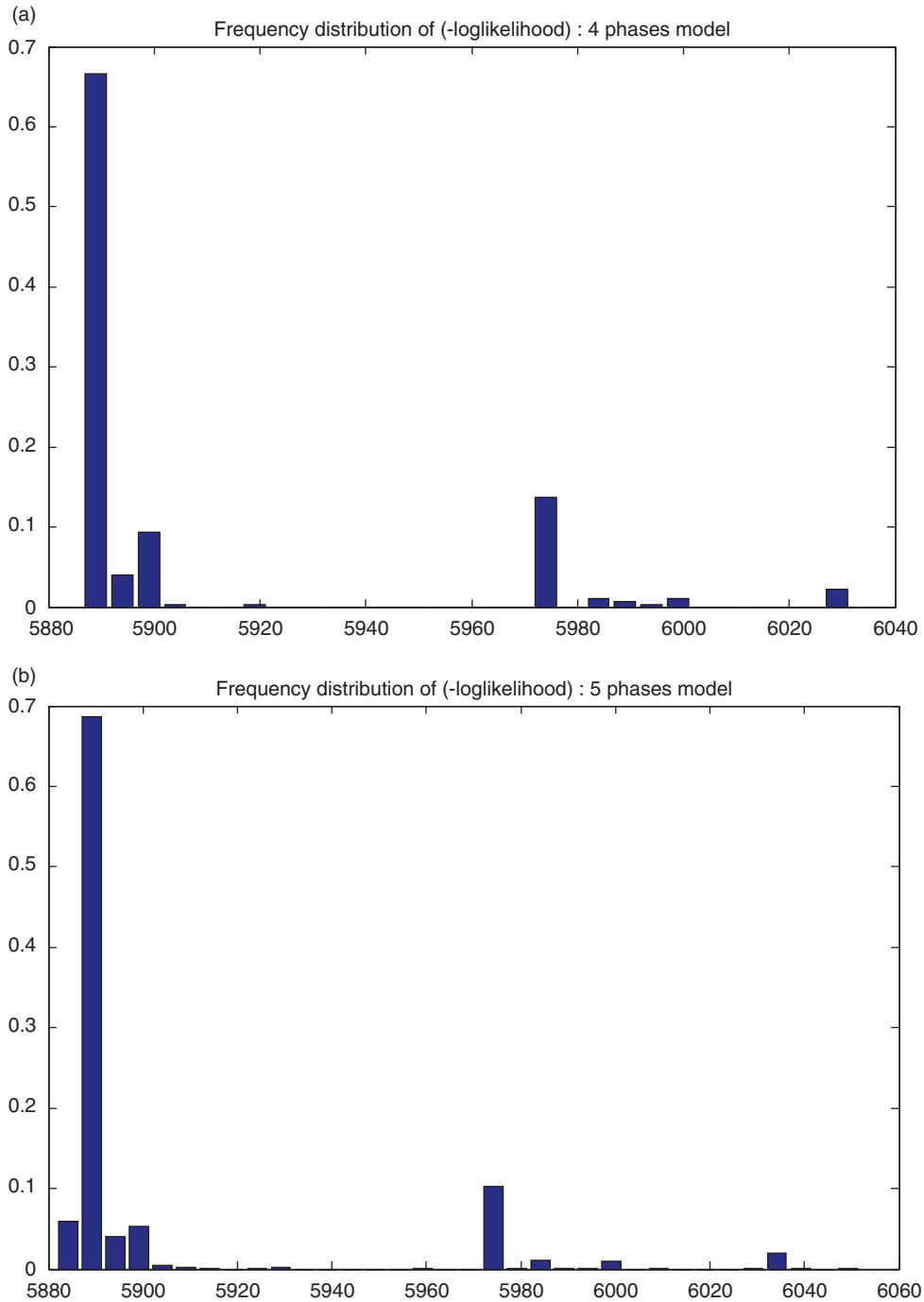
(a)

Frequency distribution of (-loglikelihood) : 4 phases model

(b)

Frequency distribution of (-loglikelihood) : 5 phases model

Fig. 4. The distributions of (− loglikelihood) for the (a) 4-phase and (b) 5-phase model.

that the 5-phase distribution having an Akaike weight of 0.94, better represents the data than the 4-phase distribution. Thus, this 5-phase model is 15.67 times more likely to be the best model than the 4-phase model with the lower (− loglikelihood). Expression (7) has mean and median values of 5904.61 and 5889.09, respectively for the 5-phase model. These values are almost the same as those obtained for the 4-phase model.

Figure 4 displays the distributions for expression (7) for both the 4-phase and 5-phase model, respectively. The 5-phase model appears to be better due to its lowest value of (7), however, the median and the modal classes are the same for both models. The largest value for expression (7) is almost identical in both models.

Table 2 displays the average values and confidence intervals for each of the nine parameters of the 5-phase distribution. In comparison with the parameter values for the 4-phase distribution, it is apparent that the parameters are not so different for both models.

Figure 5 illustrates the distributions for μ_2 and λ_4. From inspection of the values, it is possible to see for μ_2 that the results are very close to one specified value. However, on inspection of the values of parameter λ_4, it is apparent that there is a problem with identifying one specific value as being prevalent. There is a considerable amount of variability in the parameter value. This demonstrates that the initial values used for the distribution parameters appears to influence the resulting fit thus confirming what was initially expected that upon simulating and fitting various different distributions, the selection of initial values is vitally important as they can heavily influence the fitted results and the final model.

It is also noteworthy, as in the case of the research carried out by Faddy (1998) to consider some interesting features discovered for the fitted parameters in this paper. For instance, in many cases during the fitting process, at least one of the parameter values was actually estimated to be zero. In the case of the 4-phase fits, 643 out of the 1300 (nearly 50%) had one parameter value equal to zero and 535 (41%) out of the remaining 657 fits had two parameters equal to 0. This would strongly suggest over-fitting. Likewise, in the case of the 5-phase fits, 519 (nearly 40%) out of the 1300 had one parameter value equal to zero, 600 (46%) out of the remaining 781 fits had two parameters equal to 0, and 149 fits (11%) had three parameters equal to 0. Such results indicate parameter redundancy in the models which agrees exactly with what Faddy (1998) has previously reported.

Table 2
Average values and the confidence intervals for parameters (5-phase model)

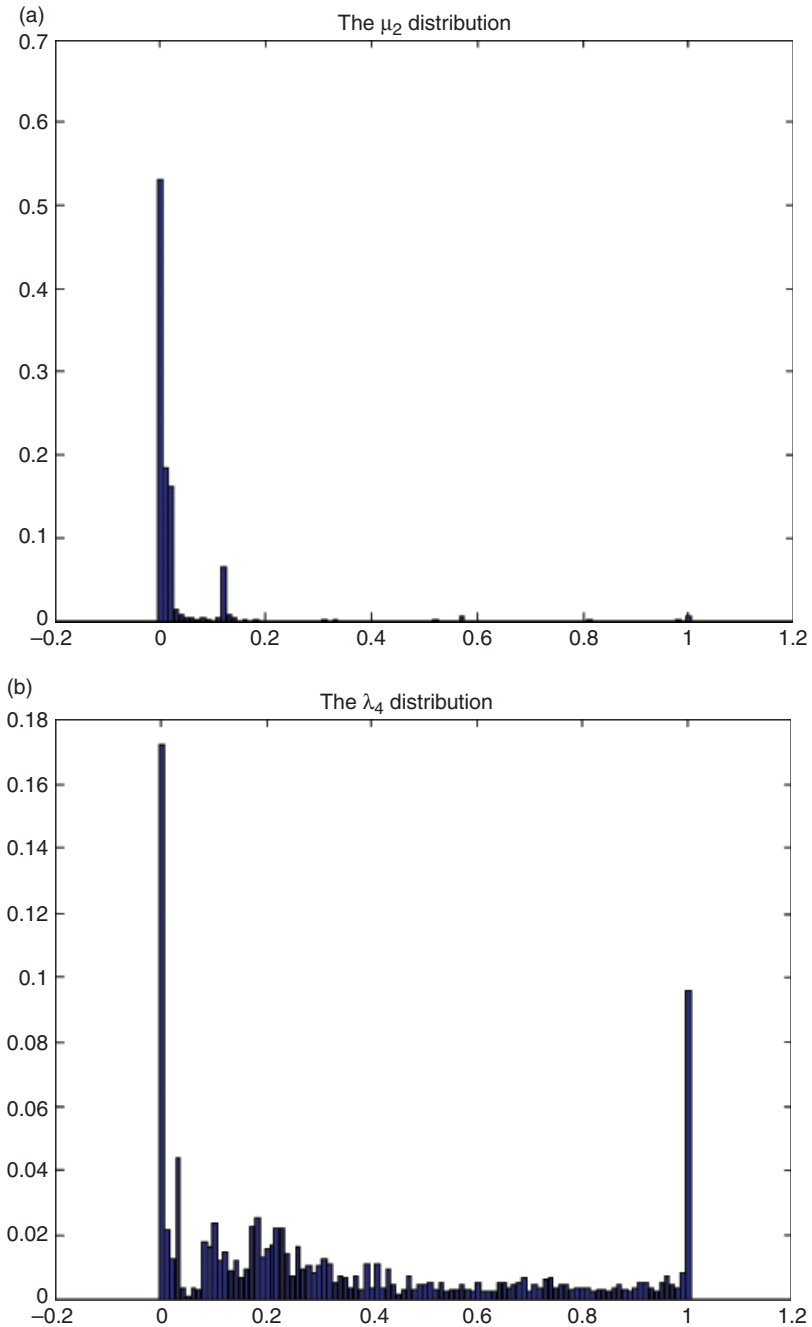| Parameter	Mean	Confidence interval ($\alpha = 0.99$)
μ_1	0.00140	(0.00086, 0.00194)
μ_2	0.02621	(0.01971, 0.03271)
μ_3	0.30781	(0.28833, 0.32730)
μ_4	0.18402	(0.16337, 0.20467)
μ_5	0.14310	(0.12088, 0.15974)
λ_1	0.36493	(0.34448, 0.38504)
λ_2	0.38575	(0.36385, 0.40764)
λ_3	0.31964	(0.29730, 0.34198)
λ_4	0.35212	(0.32826, 0.37598)

Fig. 5. The distributions of (a) μ_2 and (b) λ_4 for the 5-phase model.

Another interesting observation is the study of the eigenvalues, $\lambda_j + \mu_j$, of the matrix **Q** (3) (Faddy, 1994). In the case of the 4-phase fits, a total of 10 of the 1300 simulated fits present an eigenvalue ($\lambda_j + \mu_j$) equal to zero while a total of 99 of the 1300 simulated fits have two equal
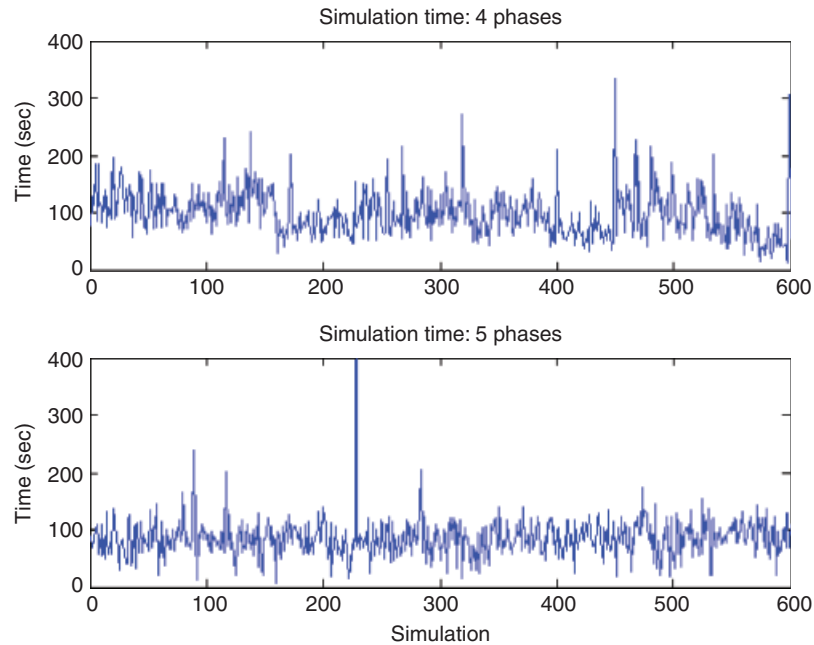
Fig. 6. The computational times for simulations of 4- and 5-phase Coxian distributions.

eigenvalues, where $\lambda_2 + \mu_2 = \lambda_3 + \mu_3$ occurred 11 times out of the 99. Likewise, in the case of the 5-phase fits, 105 of the 1300 simulated fits show at least two equal eigenvalues where $\lambda_3 + \mu_3 = \lambda_4 + \mu_4$ occurred in five cases. Further examination of the 4-phase fits reveals that nine of them have the estimated matrix $\mathbf{Q}$ (3) with $\lambda_3 + \mu_3 = \mu_4$, indicating that there is less redundancy of the parameters than first highlighted. Also for the 5-phase fits, only five of them have $\lambda_4 + \mu_4 = \mu_5$. Although this may imply a conflict with what was reported earlier in the paper, it is possible that the reason for such results are due to there being parameter redundancy in the earlier phases and not in fact in the final phase of the model.

All simulations were executed on an Intel Pentium 4 (3.2 GHz) computer equipped with 1 GB RAM. The mean computational time (standard deviation in brackets) is 97.6699 (39.2589) s for the simulations of Coxian distribution with 4 phases and 94.4795 (27.8307) s for the simulations of Coxian distribution with 5 phases. Figure 6 illustrates the computational times for 600 simulations of fitting the Coxian phase-type distribution with four and five estimated phases.

From inspection of the plot, it is possible to identify an outlier in the computational times for the 5-phase distribution fits, its time being 4498.89 s. Although in general it seems that the algorithm reaches the convergence sooner for fitting the 5-phase Coxian distribution.

Figure 7 shows the distribution of the simulation time of computation; for the 5 phases distributions, the outlier was removed to aid comparison of the two distributions.

As previously observed, the 5-phase distribution tends to be fitted faster than that of the 4-phase also suggesting that there was no over-fitting.
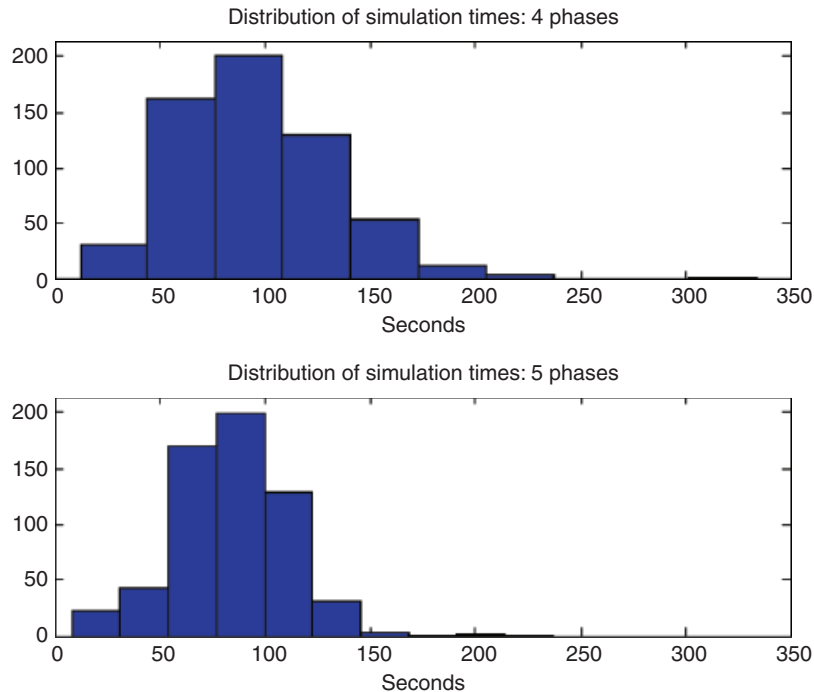
Fig. 7. The distributions of computational times taken to fit the 4- and 5-phase simulated Coxian distributions.

7. Conclusion and future work

This paper considers the Coxian phase-type distribution and how it may be fitted to a data set to represent survival times, in particular the length of time a group of elderly patients spend in hospital. In doing so, the paper investigated the fitting process of the parameter estimation, the various techniques that have been used to perform this in the past and the potential problems that may arise. On fitting the data, one problem that emerged is the inability to find a unique representation for the data. In fact, it was possible for the estimation procedure to obtain equivalent solutions.

This fitting process, along with the choice of initial parameters was further examined by simulating 1300 samples of initial parameter estimates used for 4-phase and another 1300 samples of initial parameter estimates for the 5-phase Coxian representations of patient survival. The resulting fitted distributions and associated parameter estimates were considered for the simulations for both 4-phase and 5-phase Coxian distributions. From inspection of the results, it was apparent that the initial values heavily influence the fitted results and the final model.

One very encouraging aspect of the analysis was the performance of the SPQ method used in the estimation process. Firstly, the execution of resolution performed much faster than other well-known programs. In particular, the constrained minimization for a set of initial values spent on average less time (at least 30% less) than a minimization done by the EMPht program (Asmussen et al., 1996). Secondly SPQ performs appropriately if a parameter is estimated as zero. When the

EMPht estimates a parameter as zero, it remarkably keeps this parameter as zero for all future iterations. Additionally, the SPQ always converges. This is not the case with other algorithms such as the Nelder-Mead simplex algorithm which faces the problem of not always converging to a suitable distribution.

As in iterative estimation procedures, the methods described in this article have encountered difficulties regarding the choice of the initial parameter values. The choice of such values influences the final resulting fits. In fact, in some cases a considerable amount of variability was apparent in the final parameter values. This drawback could possibly be overcome by using the considerations regarding the eigenvalues of the matrix **Q** (3) (Faddy, 1994) to estimate the parameters of the Coxian phase-type distribution, certainly an area worthy of further investigation in future research.

Another aspect not considered in this paper is the further development of a general model suitable for representing censored data in any survival analysis. Such an extension is the focus of future work.

Acknowledgements

References

Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19, 6, 716–723.

Asmussen, S., Nerman, O., Olsson, M., 1996. Fitting phase-type distributions via the EM algorithm. *Scandinavian Journal of Statistics* 23, 419–441.

Augustin, R., Buscher, K.-J., 1982. Characteristics of the Cox-distribution. *ACM Sigmetrics Performance Evaluation Review* 12, 1, 22–32.

Bobbio, A., Cumani, A., 1992. ML estimation of the parameters of a PH distribution in triangular canonical form. In Balbo, G., Serazzi, G. (eds) *Computer Performance Evaluation*, Volume 22, Elsevier Science, Amsterdam, pp. 33–46.

Bux, W., Herzog, U., 1977. The phase concept: approximation of measured data and performance analysis. In Chandy, K.M., Reiser, M. (eds.) *Computer Performance*. North-Holland, Amsterdam, pp. 23–38.

Cox, D.R., 1955. A use of complex probabilities in the theory of stochastic processes. *Cambridge Philosophical Society* 51, 313–319.

Cox, D.R., Miller, H.D., 1965. *The Theory of Stochastic Processes*. Methuen, London.

Dempster, A., Laird, D., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B* 39, 1–38.

Faddy, M., 1994. Examples of fitting structured phase-type distributions. *Applied Stochastic Models and Data Analysis* 10, 247–255.

Faddy, M., 1998. On inferring the number of phases in a Coxian phase-type distribution. *Communication in Statistics-Stochastic Models* 14, 407–417.

Faddy, M.J., 2002. Penalized maximum likelihood estimation of the parameters in a Coxian phase-type distribution. In Latouche, G., Taylor, P. (eds) *Matrix-Analytic Methods. Theory and Applications.* World Scientific, New Jersey, pp. 107–114.

Faddy, M., McClean, S.I., 1999. Analysing data on lengths of stay of hospital patients using phase-type distribution. *Applied Stochastic Models in Business and Industry* 311–317.

Johnson, M.A., Taaffe, M.R., 1989. Matching moments to phase distributions: mixtures of Erlang distributions of common order. *Communication in Statistics-Stochastic Models* 5, 711–743.

Johnson, M.A., Taaffe, M.R., 1990a. Matching moments to phase distributions: nonlinear programming approaches. *Communication in Statistics-Stochastic Models* 6, 259–281.

Johnson, M.A., Taaffe, M.R., 1990b. Matching moments to phase distributions: density function shapes. *Communication in Statistics-Stochastic Models* 6, 283–306.

Johnson, M.A., Taaffe, M.R., 1991. An investigation of phase-type-distribution moment-matching algorithms for use in queueing models. *Queueing System* 8, 129–148.

Marshall, A.H., McClean, S.I., 2003. Conditional phase-type distributions for modelling patient length of stay in hospital. *International Transactions in Operational Research* 10, 565–576.

MATLAB 2001. Optimization Toolbox For Use with Matlab, *User's Guide, Version 2*, The MathWorks Inc., Natick, MA.

Nelder, J.A., Mead, R., 1965. A simplex method for function minimization. *Computer Journal* 7, 308–313.

Neuts, M.F., 1981. *Matrix-Geometric Solutions in Stochastic Models – An Algorithmic Approach.* Johns Hopkins University Press, Baltimore.

Parr, W.C., Schucany, W.R., 1980. Minimum distance and robust estimation. *Journal of the Marican Statistical Association* 75, 616–624.

Riska, A., Diev, V., Smirni, E., 2002. Efficient fitting of long tailed data sets into PH distribution. *Internet Performance Symposium, IEEE GlobeCom 2002,* Taipei, Taiwan.

Schmickler, L., 1992. MEDA: Mixed Erlang distributions as phase-type. Representations of empirical distribution functions. *Communication in Statistics-Stochastic Models* 8, 131–156.