The XIII International Conference
"Applied Stochastic Models and Data Analysis"
(ASMDA-2009)
June 30-July 3, 2009, Vilnius, LITHUANIA

ISBN 978-9955-28-463-5
L. Sakalauskas, C. Skiadas and
E. K. Zavadskas (Eds.): ASMDA-2009
Selected papers. Vilnius, 2009, pp. 482–485

RECENT DEVELOPMENTS IN FITTING COXIAN PHASE-TYPE DISTRIBUTIONS IN HEALTHCARE

Adele H. Marshall[1], Mariangela Zenga[2]

[1]*Centre for Statistical Science and Operational Research (CenSSOR), Queen's University of Belfast, Northern Ireland*
[2]*Dipartimento di Metodi Quantitativi per le Scienze Economiche ed Aziendali, Università Milano-Bicocca, Milano*
E-mail: [1]a.h.marshall@qub.ac.uk; [2]mariangela.zenga@unimib.it

Abstract: In the past few decades, Coxian phase-type distributions have become increasingly more popular as a means of representing survival times in healthcare models. In particular, they are considered suitable for modelling the length of stay of patients in hospital and more recently for modelling the patient waiting times in Accident and Emergency Departments. However, problems have arisen in how to accurately estimate the parameters of the model from healthcare data. This paper examines the various approaches and considers the recent developments in fitting the Coxian phase-type distribution.

Keywords: Coxian phase-type distribution, fitting, simulation, algorithm.

1. Introduction

The Coxian phase-type distribution is a special type of stochastic model that represents the time to absorption of a finite Markov chain in continuous time where there is a single absorbing state and the stochastic process starts in a transient state (Neuts, 1981). Such models describe duration until an event occurs in terms of a process consisting of a sequence of ordered latent phases. It is this special feature of ordered phases that distinguishes the Coxian phase-type from the more general phase-type distribution and in doing so reduces the number of paramaters requiring estimation.

There is, however, a general issue with regard to fitting phase-type distributions in terms of the estimation of parameters. Even with the reduced number of parameters required for the Coxian phase-type distribution, estimation can still be problematic. This is considered a major drawback for their current use in everyday applications. In the last few decades, Coxian phase-type distributions have become increasingly more popular as a means of representing survival times in healthcare models. If there was a robust method for fitting the distribution, they have the potential of being a versatile approach in survvial modelling.

This paper examines the various approaches for parameter estimation and considers the most recent developments in fitting the Coxian phase-type distribution.

2. Coxian phase-type distribution

The Coxian phase-type distribution is defined as having a transition matrix **Q** of the following form,

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

where the λ_i's and μ_i's are its parameters. The λ_i's describe transitions through the ordered transient states {1, 2, ... n} and the μ_i's transitions into the (single) absorbing state n+1. The probability density function (*p.d.f.*) of T is given by

$$f(\mathrm{t}) = \mathbf{p}\exp\{\mathbf{Q}\mathrm{t}\}\mathbf{q} \tag{2}$$

where

$$\mathbf{p} = (1\ 0\ 0\ ...0\ 0\ ), \mathbf{q} = -\mathbf{Q1} = (\mu_1\ \mu_2 ... \mu_n)^{\mathrm{T}}, \tag{3}$$

and **1** is a column vector of 1's and **Q** as above (Faddy, 1994).

Parameter estimation for the Coxian phase-type distribution is nontrivial. The literature reports that there are several known problems with attempting to fit real data or a simulated distribution to the general phase-type distribution (PH). The difficulties are caused by the non-linear problem of fitting, the number of parameters to be simultaneously optimized and the non unique representations of PH distributions (Lang et al, 1997). As such it is not possible to find an exact solution so a numerical algorithm is required (Neuts, 1981). The following are well known accepted methods of estimation; the methods of maximum likelihood, methods of moments (moment matching) and the least squares approach. Many authors have considered these three techniques and applied them to the phase-type distribution.

A group of algorithms based on the maximum likelihood method has evolved over the years. These perform the fit of the Coxian phase-type distribution by maximizing (or minimizing the (-)) log-likelihood of (4). Faddy, (1994, 1998, 2002), and Faddy and McClean (1999) utilise the Nelder Mead algorithm to perform the minimization in MATLAB software (2001). Whereas Asmussen, Nerman and Olson (1996) developed the EMPht program in C (programming language) using a fitting procedure based on the expectation-maximization (EM) algorithm for the complete class of phase-type distributions. Prior to both of these approaches, Bobbio and Cumani (1992) had developed an algorithm in a FORTRAN program to maximize the log-likelihood function by combining a linear program with a line search at each iteration.

Previous research has also been performed on fitting the Coxian phase-type using the second group of algorithms for moment matching. Schmickler (1992) developed the MEDA package (in PASCAL programming language) using the method of moments to match a mixture of Erlang distributions with the first moment combined with the minimization of a deviation measure. Alternatively, Johnson (1993) wrote the MEFIT algorithm in FORTRAN programming language, to match the first three moments of any non-degenerate distribution to the first three moments of a mixture of Erlang distributions.

Lang et al. (1997) compared the moment-matching parameter method with the maximum likelihood method. They asserted that satisfactory approximations using the Johnson's method are obtained only at the price of high orders of the approximated Erlang mixture. Instead the maximum likelihood method seems to be most effective when based on the acyclic phase-type class of distribution. However the numerical implementation is very intensive due to the initial starting points of the algorithm. In fact, different starting points can actually lead to different final solutions.

The third group of algorithms is based on least squares estimation. Faddy (1993) used the least squares method (by Quasi Newton algorithm in MATLAB) to estimate the parameters for compartmental models for drug kinetics.

In summary, all of these approaches unfortunately have limitations. In general, the method of the moments is considered the least effective when the original distribution is unknown (Riska et al., 2002).

3. Simulation

This section addresses the parameter estimation problem for Coxian phase-type distribution using the maximum likelihood (MLE) method. In particular, it is wished to minimize the (-)log-likelihood of the Coxian phase-type distribution:

$$-\sum_{i=1}^{n}\log\left[\mathbf{p}\cdot\exp\{\mathbf{Q}\cdot t_i\}\mathbf{q}\right] \tag{4}$$

subject to

$$\begin{cases} 0 \le \mu_j \le 1 & j = 1,...,n \\ 0 < \lambda_j \le 1 & j = 1,...,(n-1) \end{cases} \tag{5}$$

with respect to the parameters.

The Quasi-Newton (QN) algorithm and the Nelder Mead (NM) algorithm are both used in the parameter estimation. This facilitates a comparison of the results in terms of evaluating both their global/local convergence properties and convergence speed. Two performance measures are introduced to evaluate the global/local convergence and the convergence speed for the two algorithms (Okumara, Dohi 2006). The first is the rate of convergence (ROC):

$$ROC = \frac{\text{number of successful estimations}}{\text{total number of estimation executions}} \times 100 \tag{6}$$

where "successful estimations" means that each estimation procedure terminates without any exceptions such as overflow and underflow errors. The second performance measure is the mean relative distance to MLEs (MRD):

$$MRD = \frac{\sum \frac{|\text{MLEs - estimates}|}{\text{MLEs}}}{\text{number of successful estimations}} \tag{7}$$

If the ROC is higher than the MRD, then the algorithm has the global convergence property, and therefore is stable. However, a smaller MRD means the convergence speed of the algorithm is faster.

The numerical experiment is performed for both algorithms by generating a discrete sample from a Coxian phase-type distribution with 4 phases, n=1500, and parameter values μ_i = [0.024, 0.065, 0.016, 0.0023] and λ_i= [0.067, 0.026, 0.0011] (Faddy, McClean, 1999). The starting values (103 initial values) in each algorithm are generated in advance randomly in the interval [E(t)± 0.5]. The criteria for stopping the algorithm is ε = 1.0 × 10-6 and the maximum number of iterations is 1000. All simulations were executed using MATLAB on an Intel Pentium 4 (3.2 GHz) computer equipped with 1GB RAM.

4. Results

The results for both the Quasi-Newton (QN) and Nelder Mead (NM) algorithms are reported in Tables 1 and 2. The rate of convergence (ROC) for the QN algorithm is always 100%, that is, the QN algorithms never cause any computation errors such as overflow and underflow (Table 1). The NM algorithm performs poorly with only 295 simulations (29.5%) converging and only 10 with acceptable results.

Table 1. Comparison of the iteration schemes based on the Quasi-Newton and Nelder Mead algorithms.

	Quasi-Newton (QN)	Nelder Mead (NM)
Rate of Convergence (ROC)	100%	29.5%
Mean Relative Distance (MRD)	0.00285	0.00214
Mean Convergence (seconds)	62.077	31.383

The MRDs for QN and NM reach almost the same values. Using these convergence measures, it can be concluded that the QN algorithms are more effective to compute the MLEs. However, the mean convergence time for the QN algorithm is 62 seconds taking double the mean convergence time of 31 seconds for the NM algorithm.

Table 2. Average value for the parameter estimates with standard deviation (in brackets) and (–)log-likelihood value.

	Parameter Value	Quasi-Newton	Nelder Mead		Parameter Value	Quasi-Newton	Nelder Mead
μ_1	0.024	0.0187 (0.00058)	0.0190 (0.00037)	λ_1	0.067	0.1003 (0.01298)	0.0931 (0.02094)
μ_2	0.065	0.0596 (0.01034)	0.0619 (0.01028)	λ_2	0.026	0.0631 (0.01358)	0.0486 (0.02180)
μ_3	0.016	0.0453 (0.01011)	0.0346 (0.01937)	λ_3	0.0011	0.0911 (0.08932)	0.0547 (0.02540)
μ_4	0.0023	0.0394 (0.08430)	0.0328 (0.00502)	-log likelihood	-	5823.10	5823.40

Table 2 displays the average values for each of the 7 parameters fitted using the Quasi-Newton method and Nelder Mead algorithm. By comparing the estimates of both methods with their real values, the Nelder Mead algorithm appears to be closer in every case and thus seems to provide the better values, even if the estimations have more dispersion. The smallest negative log-likelihood value obtained from the 1000 samples of initial parameter values is almost identical for both methods of minimization.

5. Conclusion and future work

This paper has performed a comparison on two algorithms using the *maximum likelihood* approach to estimate the parameters of the Coxian phase-type distribution. In a numerical experiment, 1000 samples of data have been simulated from a known 4 phase Coxian distribution. The simulated samples have then been fitted to a Coxian phase-type distribution and the resulting parameters and true parameters compared along with other performance measures. In terms of the rate of convergence of the algorithm the Quasi-Newton algorithm shows the best performance while the Nelder Mead algorithm performs better in terms of parameter estimation giving values closer to the real parameters.

As the maximization of the log-likelihood of the Coxian phase-type distribution is a numeric algorithm, it is apparent that the initial values heavily influence the fitted results and the final model. This represents a problem for the parameter estimation of this kind of distribution. To overcome the problem, it could be necessary to perform the optimization using the eigenvalues of the matrix **Q**.

References

Asmussen, S., Nerman, O., Olsson, M. 1996. Fitting Phase-Type Distributions via the EM Algorithm. *Scandinavian Journal of Statistics*, 23, 419–441.

Bobbio A., Cumani A. 1992. ML estimation of the parameters of a PH distribution in triangular canonical form. In G. Balbo and G. Serazzi, editors, Computer Performance Evaluation, 22, 33–46.

Bobbio, A.,. Telek, M., 1994. A Benchmark for PH Estimation Algorithm: Results for Acyclic PH, *Stochastic Models*, 10, 661–667.

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

Faddy, M., 1990. Compartmental models with residence-time distribution. *Applied stochastic models and data analysis*, 6, 121–127..

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. and Taylor, P., editors, *Matrix-analytic Methods. Theory and Applications*, pages 107–114. World Scientific, New Jersey.

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. 1993. Selecting parameters of phase distribution: combining non linear programming, heuristics and Erlang distribution. *ORSA Journal on Computing*, 5, 69–83.

Lang A., Arthur J.L. 1997. Parameter approximation for phase-type distributions in Chakravarthy, S.R. and Alfa A.S. editors *Matrix-analytic methods in stochastic models*, volume 189, 151–206 *Lecture notes in pure and applied mathematics*, Marcel Dekkar, New York 1997

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.

Marshall A.H., Zenga MA. 2009. Simulating Coxian Phase-type Distributions for Patient Survival. To be appearing on *International Transactions in Operational Research*

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

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

Okumara H., Dohi T. 2006. Building Phase-Type Software Reliability, *ISSRE '06*, 289–298.

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.