Phase-type Distributions in Healthcare Modelling II

Mark Fackrell

Department of Mathematics and Statistics
University of Melbourne
Victoria 3010
AUSTRALIA

email: mfackrel@ms.unimelb.eu.au

In the second of these three articles on phase-type (PH) distributions in healthcare modelling I will introduce Coxian distributions, a particularly useful subclass, and explain briefly how they have been used in healthcare modelling. In the first article (see Fackrell [3]) the case for using stochastic models in healthcare was made, and PH distributions were introduced. Fackrell [2] contains a more comprehensive treatment and bibliography.

An order p Coxian distribution has a representation of the form

$$\boldsymbol{\alpha} = \begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_p \end{pmatrix} \tag{1}$$

$$\boldsymbol{T} = \begin{pmatrix} -\lambda_1 & \lambda_1 & 0 & \dots & 0 \\ 0 & -\lambda_2 & \lambda_2 & \dots & 0 \\ 0 & 0 & -\lambda_3 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -\lambda_p \end{pmatrix}, \tag{2}$$

where $0 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_p$. Figure 1 shows the state transition diagram for an order p Coxian distribution.

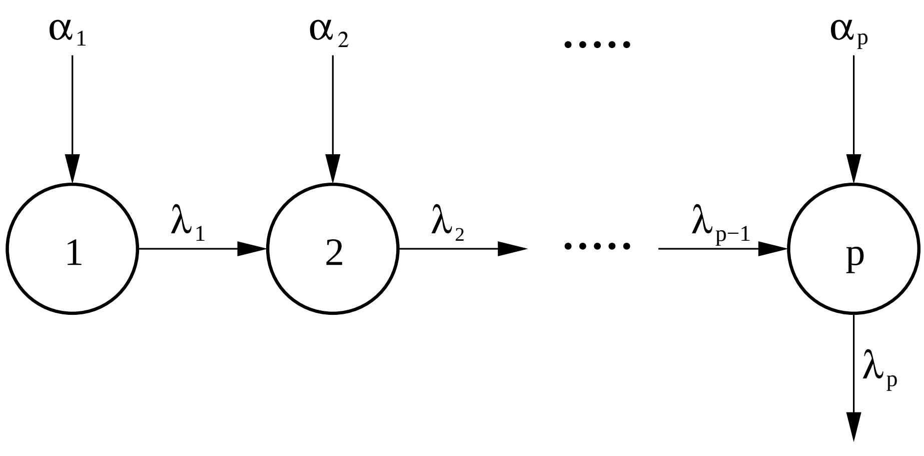

Figure 1: State transition diagram for an order p Coxian distribution.

The first thing we notice is that the representation (1)–(2) depends on only $2p - 1$ parameters. General PH representations require $p^2 + p - 1$ parameters, but only $2p - 1$ parameters are needed to define the distribution uniquely. Thus, general PH distributions are considerably overparameterized, whereas Coxian distributions are not. When it comes to fitting data with PH distributions, Coxian distributions have been the prefered option for practitioners in healthcare modelling and elsewhere, partly because of this reason - there are a lot less parameters to estimate.

Coxian distributions exhibit quite a lot of flexibility. Figure 2 shows the density functions for three order 4 Coxian distributions with the same generator $\boldsymbol{T}$, but different vectors $\boldsymbol{\alpha}$. We can see by their shapes that they exhibit much more flexibility than the exponential distribution, which should come as no surprise - they rely on 7 free parameters instead of just one! Indeed, Coxian distributions are more versatile than both hyperexponential distributions and generalized Erlang distributions, which also depend on $2p-1$ parameters, see Fackrell [2].

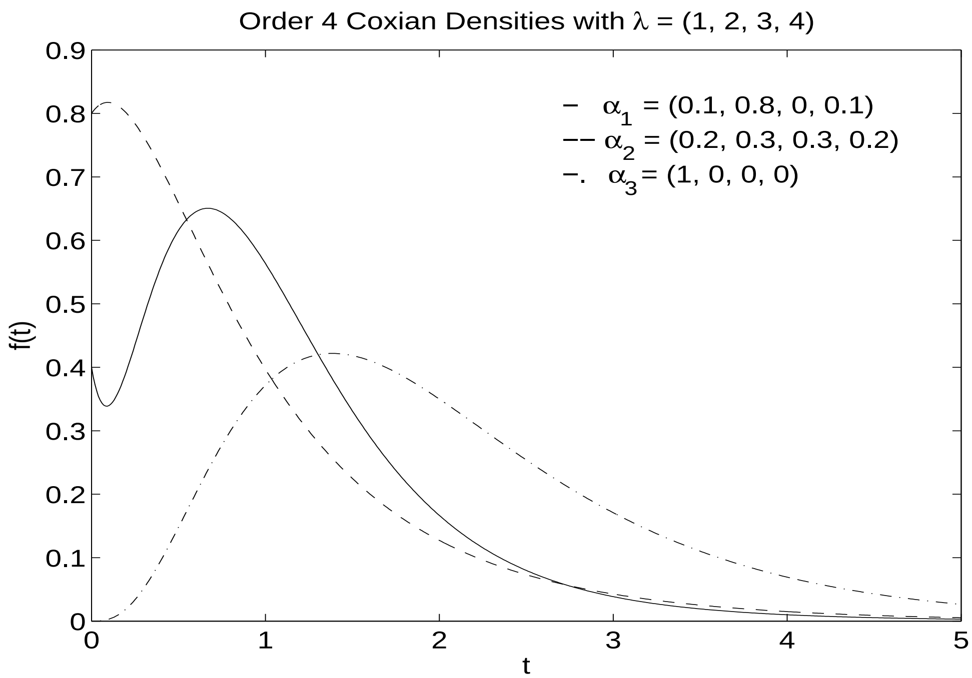

Figure 2: Density functions for three different order 4 Coxian distributions with $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$, $\lambda_4 = 4$.

A curious mathematical fact is that any PH distribution whose generator $\boldsymbol{T}$ is an upper triangular matrix, has a Coxian representation of the same or lower order, see Cumani [1] or O'Cinneide [9]. For example, the hyperexponential distribution with representation

$$\boldsymbol{\alpha} = \begin{pmatrix} \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \end{pmatrix} \tag{3}$$

$$\boldsymbol{T} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{4}$$

has an equivalent Coxian representation

$$\boldsymbol{\beta} = \begin{pmatrix} \frac{1}{9} & \frac{1}{6} & \frac{13}{18} \end{pmatrix} \tag{5}$$

$$\boldsymbol{S} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -2 & 2 \\ 0 & 0 & -3 \end{pmatrix}, \tag{6}$$

In fact, any *PH* distribution whose generator $\boldsymbol{T}$ has only real eigenvalues has a Coxian representation of *some* order. The million dollar question is, of course, how big is the order? It is widely believed that there are examples of *PH* distributions of relatively low order that have Coxian representations of high order. Nevertheless, Coxian distributions are a very important subclass of *PH* distributions and need to be studied.

Coxian distributions have been popular with healthcare modellers because the states (or groups of states) can sometimes be given a physical interpretation. For example, Xie, Chaussalet, and Millard [10] modelled the length of stay (LOS) of geriatric patients in residential and nursing home care with two, 2-state Coxian distributions. Figure 3 shows the state transition diagram for their model. Here, patients enter the system via the

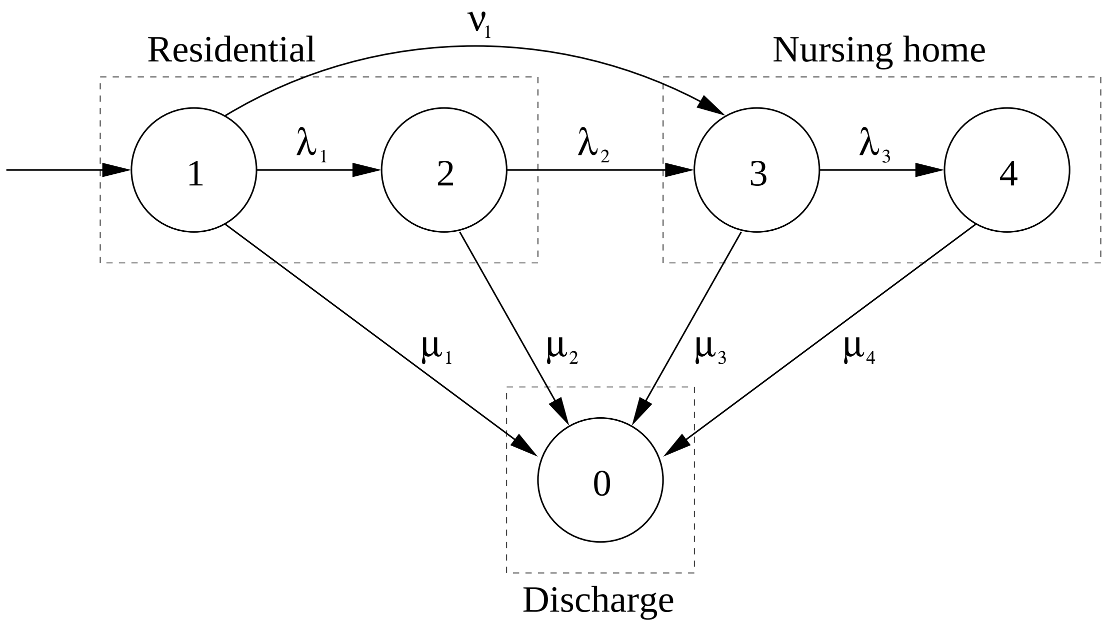

Figure 3: State transition diagram to model the length of stay in residential and nursing home care.

residential home care block where they can spend a short time (state 1 only), or a long time (state 1 followed by state 2). They can be discharged from either state, or progress to nursing home care, where again they can spend a short time (state 3), or a long time (states 3 and 4) before being discharged or dying.

The corresponding *PH* representation for the model is

$$\boldsymbol{\alpha} = \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix} \tag{7}$$

$$\boldsymbol{T} = \begin{pmatrix} -(\lambda_1+\mu_1+\nu_1) & \lambda_1 & \nu_1 & 0 \\ 0 & -(\lambda_2+\mu_2) & \lambda_2 & 0 \\ 0 & 0 & -(\lambda_3+\mu_3) & \lambda_3 \\ 0 & 0 & 0 & -\mu_4 \end{pmatrix}. \tag{8}$$

We remark here that because $\boldsymbol{T}$ is an upper triangular matrix, the *PH* distribution with representation $(\boldsymbol{\alpha}, \boldsymbol{T})$ is a Coxian distribution.

The authors fitted the Coxian distribution to four years data from the social services department of a London borough using maximum likelihood estimation. They reported

$\lambda_1 = \lambda_2 = \mu_2 = 0$, $\nu_1 = 0.000228$, $\mu_1 = 0.000855$, $\lambda_3 = 0.010874$, $\mu_3 = 0.006138$, and $\mu_4 = 0.001275$. The LOS in residential care was modelled by an exponential distribution (state 2 was unneccessary) with an average LOS of 923 days, with 21% of patients moving on to nursing home care and the rest getting discharged. The LOS for short stay patients in nursing home care (36 %) was modelled with an exponential distribution (average LOS 59 days), and the LOS for long stay patients (64 %) was modelled with a 2-state generalized Erlang distribution (average LOS 843 days).

Other notable papers where Coxian distributions have been used to model systems in healthcare include Faddy and McClean [4] and [5], Faddy and Taylor [6], McClean, Faddy, and Millard [8], and Marshall and McClean [7].

In the next issue's article I will propose some new ways in which *PH* distributions could be used in healthcare modelling.

References

[1] Cumani, A. On the canonical representation of homogeneous Markov processes modelling failure-time distributions. Microelectronics and Reliability **1982**, *22*, 583–602.

[2] Fackrell, M. Using phase-type distributions in health and social care modelling. Submitted, 2007.

[3] Fackrell, M. Phase-type distributions in health care modelling I. Nosokinetic News, Issue 4.4, August, 2007.

[4] Faddy, M. J.; McClean, S. I. Analysing data on lengths of stay of hospital patients using phase-type distributions. Applied Stochastic Models in Business and Industry **1999**, *15*, 311–317.

[5] Faddy, M. J.; McClean, S. I. Markov chain modelling for geriatric patient care. Methods of Information in Medicine **2005**, *44*, 369–373.

[6] Faddy, M. J.; Taylor, G. T. Stochastic modelling of the onset of bronchiolitis obliterans syndrome following lung transplantations: An analysis of risk factors. Mathematical and Computer Modelling **2003**, *38*, 1185–1189.

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

[8] McClean, S. I.; Faddy, M. J.; Millard, P. H. Markov model-based clustering for efficient patient care. Proceedings of the 18th IEEE Symposium on Computer-Based Medical Systems **2005**, 467–472.

[9] O'Cinneide, C. A. On non-uniqueness of representations of phase-type distributions. Communications in Statistics – Stochastic Models **1989**, *5*, 247–259.

[10] Xie, H.; Chaussalet, T. J.; Millard. P. H. A continuous time Markov model for the length of stay of elderly people in institutional long-term care. Journal of the Royal Statistical Society. Series A **2005**, *168*, 51–61.