Noname manuscript No.
(will be inserted by the editor)

A Differential Equation for a Class of Discrete Lifetime Distributions with an Application in Reliability

A Demonstration of the Utility of Computer Algebra

Attila Csenki

Received: date / Accepted: date

Abstract It is shown that the probability generating function of a lifetime random variable T on a finite lattice with polynomial failure rate satisfies a certain differential equation. The interrelationship with Markov chain theory is highlighted. The differential equation gives rise to a *system* of differential equations which, when inverted, can be used in the limit to express the polynomial coefficients in terms of the factorial moments of T. This then can be used to estimate the polynomial coefficients. Some special cases are worked through symbolically using Computer Algebra. A simulation study is used to validate the approach and to explore its potential in the reliability context.

Keywords Polynomial failure rate · Probability generating function · Markov chain · Stirling numbers · Computer algebra · Point estimation · Reliability

Mathematics Subject Classification (2000) 60J22 · 90B25 · 62M05 · 65C50

1 Introduction

Parametric classes of distributions taken to model lifetime data are frequently chosen because they can be used to approximate closely the probability density, the cumulative distribution function or the failure rate function of the underlying distribution. Whereas traditionally the failure density is the main focus of attention, Berg has stressed in [2] the importance of choosing classes of lifetime distributions for reliability modelling with *failure rate* functions which are intuitively appealing and capable of taking a wide range of functional shapes. The question addressed in this paper is motivated by this viewpoint.

A. Csenki
School of Computing, Informatics and Media, University of Bradford, Bradford, UK Tel.: +44-1274-232323
E-mail: a.csenki@bradford.ac.uk

Polynomial failure rates are of interest for continuous lifetime models (e.g. [13]) since any 'smooth' function on a closed finite interval can be approximated uniformly by a polynomial. We are concerned here with *discrete* models, however, where even though this approximation argument is of a lesser importance, but the discrete failure rate function can be applied to approximately describe a continuous-time failure process. The object of interest here is the class of lifetime distributions with a finite lattice support and polynomial failure rate function.

After defining the framework in the next section, we present in Section 3.1 a differential equation for the probability generating function (pgf) of a lifetime distribution of the said kind. The differential equation is in terms of the polynomial coefficients and the number of lattice points of the failure rate function. (The differential equation is proved in the Appendix, Section 8.) In Section 3.2, the Markov framework is used to derive an explicit expression for the same pgf. In Section 4, a technique is described for expressing the polynomial coefficients in terms of factorial moments of the distribution. It is based on inverting (in the limit) a system of differential equations obtained from the original equation and an algebraic constraint. To obtain specific results, we turn to a computer algebra system. Some of the equations thus derived are then used in Section 5 in a simulation study.

The papers available on discrete lifetime random variables are legion; a small selection concerned with, for example, discrete versions of the Weibull distribution (in the reliability context) is [11], [12], [17] and [24]. [23] is of relevance if the interest is in *discrete failure rate* functions, in particular for comparing the corresponding distributions also in the multivariate case.

The area of discrete parameter distributions in the context of reliability and lifetime analysis is of current research interest as shown e.g. by [25].

[4] and [5] are remotely related to the subject matter of this paper in that they deal with discrete distributions where (generalizations of) the Stirling numbers appear; these numbers in their original (not generalized) form play a role in our results too.

Finally, the recent article [21] (and the references cited therein) show a continued interest in the related field of (continuous) linear failure rate models.

2 Modelling time to failure

The 'time' to failure T is assumed to take values in $\mathcal{T} = \{0, 1, 2, \ldots, n\}$, that is, failure can happen at any of the time instances in $\mathcal{T}$. In some applications T will be the actually elapsed time to failure; then T measures on a discretized time line the number of time units until failure. In some other applications T measures the number of cycles until failure. The *failure rate* is the conditional probability of failure in the kth time instant given survival until time $k-1$; it is defined by

$$r_k = P(T = k | T > k-1), \quad k \in \mathcal{T}. \tag{1}$$

Notice that (1) for $k = 0$ is the probability of instantaneous failure r_0 (as the conditioning event in (1) is now the full event). It is assumed that

$$r_n = 1, \tag{2}$$

i.e. failure is certain to occur at the latest after n usage periods. Finally, it will be assumed that

$$0 < r_k < 1, \quad k = 0, \ldots, n-1. \tag{3}$$

2.1 Polynomial failure rate

The failure rate function r in (1) will be assumed to be a *polynomial* of degree m, i.e.

$$r_k = a_0 + a_1 k + a_2 k^2 + \ldots + a_m k^m, \quad k \in \mathcal{T}. \tag{4}$$

The assumption of r being a polynomial is reasonable as it is well-known that any continuous function on the interval $[0, n]$ (of which $\mathcal{T}$ is a subset) can be arbitrarily closely approximated by a polynomial. From a practical point of view it is important to notice that in particular a (discretized) bathtub shaped failure rate curve can be obtained by choosing the parameters m and a_i in (4) appropriately.

The prime objective is to interrelate the parameters $a_0, \ldots, a_n$ by a differential equation for the pgf of the distribution of T. The differential equation will be presented in Section 3.1. Subsequently, the differential equation will be used to estimate the parameters $a_0, \ldots, a_n$ based on ℓ independent realizations $t_1, \ldots, t_\ell$ of T. Together with the boundary condition (2), now in the form,

$$\sum_{i=0}^{m} a_i n^i = 1, \tag{5}$$

this gives initially two equations for the $m + 1$ parameters $a_0, \ldots, a_m$. The missing $m - 1$ conditions for the $m + 1$ parameters will be obtained by differentiating the differential equation $(m - 1)$ times.

2.2 Continuous time analogue

Noteworthy is the continuous time analogue of the class of lifetime distributions considered here. If r in (4) is a continuos time failure rate function, i.e. if

$$r(t) = \sum_{j=0}^{m} a_j t^j, \quad , t \geq 0,$$

then the reliability function of the corresponding lifetime distribution on $[0, +\infty)$ is

$$R(t) = \exp(-H(t)), \tag{6}$$

with the cumulative hazard

$$H(t) = \int_0^t r(s)ds = \sum_{j=0}^{m} \frac{a_j}{j+1} t^{j+1}. \tag{7}$$

Equations (6) and (7) show that R is then the reliability function of a lifetime variable distributed like the *minimum* of m independent Weibull random variables where the jth of them has scale parameter α_j and shape parameter β_j with

$$\alpha_j = \sqrt[j+1]{\frac{j+1}{a_j}}, \qquad \beta_j = j+1.$$

Because of the continuous nature of the model and because the support is the entire non-negative axis, there is now no constraint interrelating the parameters a, i.e. there is no continuous time analogue of (5).

A differential equation has been derived recently in [7] for the Laplace transform of continuously distributed T with a polynomial failure rate. The result in there is an analogue of that in Section 3.1. The continuous case and its proof are, however, much simpler than their counterpart presented here in Section 3.1.

3 Probability generating function

The pgf G of T is defined by (e.g. [9], [20])

$$G(z) = E\left(z^T\right) = \sum_{i=0}^{n} z^i P(T=i), \quad |z| \le 1.$$

The pgf is the transform usually considered for probability distributions on the set of non-negative integers.

3.1 Differential equation

The following differential equation holds for G.

Theorem 1 *G satisfies the differential equation*

$$G(z) = \sum_{k=0}^{m} u_k z^k \frac{d^k}{dz^k}\left(\frac{1-z^{n+1}}{1-z}\right) - \sum_{k=0}^{m} v_k z^{k+1} \frac{d^k}{dz^k}\left(\frac{G(z)-z^n}{1-z}\right), \tag{8}$$

where the coefficients $\mathbf{u}$ *and* $\mathbf{v}$ *in (8) are respectively given by*

$$u_j = \begin{cases} a_0 & \text{for } j = 0, \\ \sum_{s=j}^{m} a_s \sigma_s^{(j)} & \text{for } j = 1, \dots, m, \end{cases} \tag{9}$$

$$v_j = \sum_{s=j+1}^{m+1} a_{s-1} \sigma_s^{(j+1)}, \ \text{for } j = 0, \dots, m, \tag{10}$$

with $\sigma_s^{(j)}$ *standing for Stirling numbers of the second kind ([1], [3] and [20, p. 137]); they are determined by the relation*

$$x^s = \sum_{j=1}^{s} \sigma_s^{(j)} x(x-1)\dots(x-j+1). \tag{11}$$

Furthermore, the following holds for the coefficients u and v.

Proposition 1 *If* $\mathbf{u}$ *and* $\mathbf{v}$ *are respectively defined by (9) and (10) then they satisfy for all non-negative integers* k *and real* z *the equations*

$$\left(\sum_{i=0}^{m} a_i k^i \right) z^k = \sum_{i=0}^{m} u_i z^i \frac{d^i}{dz^i}(z^k), \tag{12}$$

and

$$\left(\sum_{i=0}^{m} a_i k^i \right) z^k = \sum_{i=0}^{m} v_i z^{i+1} \frac{d^i}{dz^i}(z^{k-1}). \tag{13}$$

$\mathbf{u}$ *and* $\mathbf{v}$ *are unique solutions of (12) and (13) respectively.*

The proofs can be found in the Appendix (Section 8).

Remark 1 Proposition 1 in conjunction with a computer algebra system offers an easy alternative to (9) and (10) for evaluating the coefficients $\mathbf{u}$ and $\mathbf{v}$. The MAXIMA implementation referred to in Section 5 is indeed based on coefficient-matching in Proposition 1 for finding $\mathbf{u}$ and $\mathbf{v}$.

3.2 Markov chain model

The distribution of $N = T + 1$ is the number of transient states visited until absorption in state $n+1$ of the Markov chain shown in Fig. 1 if it is started in state 0. (N is called the 'length of *sojourn* in the set of transient states'.) The

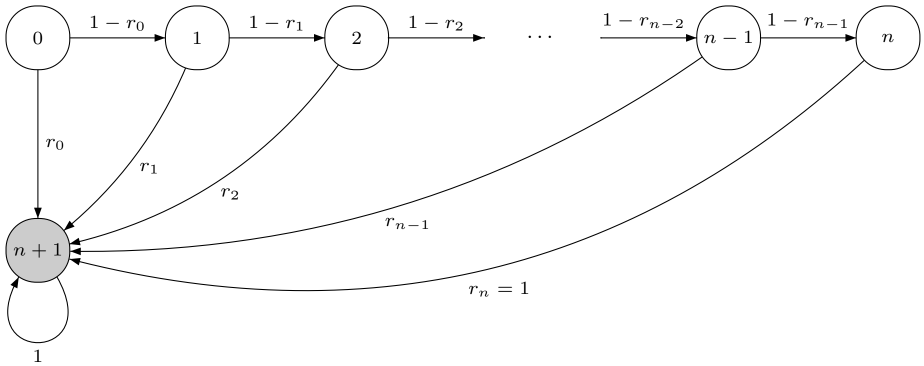

Fig. 1 Absorbing Markov chain

transition probability matrix of the chain in Fig. 1 is, in a partitioned form,

$$\mathbf{P} = \left[\begin{array}{c|c} \mathbf{Q} & \mathbf{r} \\ \hline \mathbf{0} & 1 \end{array}\right],$$

with the $(n+1) \times (n+1)$ square matrix

$$\mathbf{Q} = \begin{bmatrix} 0 & (1-r_0) & 0 & 0 & \cdots & 0 \\ 0 & 0 & (1-r_1) & 0 & \cdots & 0 \\ 0 & 0 & 0 & (1-r_1) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & (1-r_{n-1}) \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \tag{14}$$

and the column vector

$$\mathbf{r} = \begin{bmatrix} r_0 \\ r_1 \\ \vdots \\ r_{n-1} \\ 1 \end{bmatrix},$$

which is of length $(n+1)$. It is well known (e.g. [6]) that N has probability mass function

$$P(N=k) = \begin{cases} 0 & \text{for } k = 0, \\ \boldsymbol{\alpha}^t \mathbf{Q}^{k-1}(\mathbf{I} - \mathbf{Q})\mathbf{1} & \text{for } k \geq 1, \end{cases}$$

where $\boldsymbol{\alpha}$ is the initial probability (column) vector on the set of transient states (of length $(n+1)$), and $\mathbf{1}$ is the column vector of all ones; as the chain is started in state 0, it is $\boldsymbol{\alpha}^t = (1, 0, \ldots, 0)$. Therefore, the pgf of N is

$$\begin{aligned} E\left(z^N\right) &= \sum_{k=1}^{\infty} z^k \boldsymbol{\alpha}^t \mathbf{Q}^{k-1} (\mathbf{I} - \mathbf{Q}) \mathbf{1} \\ &= z \boldsymbol{\alpha}^t \left(\mathbf{I} - z\mathbf{Q}\right)^{-1} \left(\mathbf{I} - \mathbf{Q}\right) \mathbf{1}, \end{aligned}$$

from which it follows by

$$T = N - 1, \tag{15}$$

that

$$G(z) = \boldsymbol{\alpha}^t \left(\mathbf{I} - z\mathbf{Q}\right)^{-1} \left(\mathbf{I} - \mathbf{Q}\right) \mathbf{1}. \tag{16}$$

Equation (16) records the pgf of T *explicitly*. The main goal of the paper is the recognition that G satisfies the differential equation (8), which in turn gives access to the coefficients $a_0, \ldots, a_m$, as will be shown in Section 4 below.

The explicit expression in (16) for the pgf *could* be taken (theoretically!) to confirm that the differential equation (8) holds. However, this verification would probably be very tedious and it has not been attempted.

4 Polynomial coefficients and factorial moments

In this section it will be demonstrated that the differential equation (8) can be used to obtain the polynomial coefficients $a_0, \ldots, a_m$ in terms of the factorial moments of T. Such a representation is of practical interest as the factorial moments are readily estimated from samples of T.

In Sections 4.1-4.3, the cases $m = 1, 2, 3$ will be considered in turn by using the computer algebra system Maxima ([10], [22], [19]).

It is well known that the ith derivative of the pgf of T *at* $z = 1$ equals the ith factorial moment of T,

$$g_i \underset{def}{=} \left.\frac{d^i}{dz^i}(G(z))\right|_{z=1} = E\left(T(T-1)\ldots(T-i+1)\right). \tag{17}$$

4.1 r is a first order polynomial

It is

$$r_k = a_0 + a_1 k, \quad k \in \mathcal{T},$$

and (8) becomes

$$\begin{aligned} G(z) = a_0 \frac{1 - z^{n+1}}{1 - z} &+ a_1 z \frac{d}{dz}\left(\frac{1 - z^{n+1}}{1 - z}\right) \\ &- (a_0 + a_1) z \frac{G(z) - z^n}{1 - z} - a_1 z^2 \frac{d}{dz}\left(\frac{G(z) - z^n}{1 - z}\right). \end{aligned} \tag{18}$$

Solving for a_0 and a_1 the system comprising (18) and the equation $a_0+a_1n=1$, we get with Maxima,

$$a_0 \equiv \frac{\left(z^3-z^2\right)\left(\frac{d}{d\,z}\,G\left(z\right)\right)+\left(-n\,z^2+\left(2\,n-1\right)\,z-n\right)\,G\left(z\right)+z}{\left(z^3-z^2\right)\left(\frac{d}{d\,z}\,G\left(z\right)\right)+\left(\left(n-1\right)\,z-n\,z^2\right)\,G\left(z\right)+\left(n+1\right)\,z-n}, \quad (19)$$

$$a_1 \equiv -\frac{\left(z-1\right)\,G\left(z\right)-z+1}{\left(z^3-z^2\right)\left(\frac{d}{d\,z}\,G\left(z\right)\right)+\left(\left(n-1\right)\,z-n\,z^2\right)\,G\left(z\right)+\left(n+1\right)\,z-n}. \quad (20)$$

Applying l'Hospital's rule to (19) and (20) twice each with $z\to 1$, we get with Maxima

$$a_i = \frac{\gamma_{i,0}+\gamma_{i,1}n}{\kappa_0+\kappa_1 n}, \quad i=0,1, \tag{21}$$

where

$$\begin{aligned} \gamma_{0,0} &= g_2+2g_1, & \gamma_{0,1} &= -2,\\ \gamma_{1,0} &= -2g_1, & \gamma_{1,1} &= 0, \end{aligned}$$

and

$$\kappa_0 = g_2+2g_1, \qquad \kappa_1 = -2g_1-2.$$

4.2 r is a quadratic polynomial

It is now

$$r_k = a_0+a_1k+a_2k^2, \quad k\in\mathcal{T},$$

and (8) reads

$$\begin{aligned} G(z) &= a_0\frac{1-z^{n+1}}{1-z}\\ &\quad+(a_1+a_2)z\frac{d}{dz}\left(\frac{1-z^{n+1}}{1-z}\right)+a_2z^2\frac{d^2}{dz^2}\left(\frac{1-z^{n+1}}{1-z}\right)\\ &\quad-(a_0+a_1+a_2)z\frac{G(z)-z^n}{1-z}\\ &\quad-(a_1+3a_2)z^2\frac{d}{dz}\left(\frac{G(z)-z^n}{1-z}\right)-a_2z^3\frac{d^2}{dz^2}\left(\frac{G(z)-z^n}{1-z}\right). \end{aligned} \tag{22}$$

The task is to solve for a_0, a_1 and a_2 the system comprising the three equations: (22), the first derivative of (22), and, the equation $a_0+a_1n+a_2n^2=1$. This we accomplish with Maxima as before, though the resulting formulae are now somewhat more complicated. We get

$$a_i = \frac{\gamma_{i,0}+\gamma_{i,1}n+\gamma_{i,2}n^2}{\kappa_0+\kappa_1n+\kappa_2n^2}, \quad i=0,1,2, \tag{23}$$

where

$$\begin{aligned}
\gamma_{0,0} &= -90\, g_2\, g_4 - 180\, g_1\, g_4 + 80\, g_3^2 - 960\, g_1\, g_3 + 360\, g_2^2 - 720\, g_1\, g_2,\\
\gamma_{0,1} &= 180\, g_4 - 240\, g_1\, g_3 + 1440\, g_3 - 1080\, g_1\, g_2 + 2520\, g_2 - 720\, g_1^2 + 720\, g_1,\\
\gamma_{0,2} &= -240\, g_3 + 360\, g_1\, g_2 - 1080\, g_2 + 720\, g_1^2 - 720\, g_1,\\
\gamma_{1,0} &= 180\, g_1\, g_4 - 120\, g_2\, g_3 + 1440\, g_1\, g_3 - 540\, g_2^2 + 2160\, g_1\, g_2 + 720\, g_1^2,\\
\gamma_{1,1} &= 0,\\
\gamma_{1,2} &= 360\, g_2 - 720\, g_1^2,\\
\gamma_{2,0} &= -240\, g_1\, g_3 + 180\, g_2^2 - 720\, g_1\, g_2 - 720\, g_1^2,\\
\gamma_{2,1} &= 720\, g_1^2 - 360\, g_2,\\
\gamma_{2,2} &= 0,
\end{aligned}$$

and

$$\begin{aligned}
\kappa_0 &= -90\, g_2\, g_4 - 180\, g_1\, g_4 + 80\, g_3^2 - 960\, g_1\, g_3 + 360\, g_2^2 - 720\, g_1\, g_2,\\
\kappa_1 &= 180\, g_1\, g_4 + 180\, g_4 - 120\, g_2\, g_3 + 1200\, g_1\, g_3 + 1440\, g_3 - 540\, g_2^2\\
&\quad + 1080\, g_1\, g_2 + 2520\, g_2 + 720\, g_1,\\
\kappa_2 &= -240\, g_1\, g_3 - 240\, g_3 + 180\, g_2^2 - 360\, g_1\, g_2 - 1080\, g_2 - 720\, g_1,
\end{aligned}$$

l'Hospital's rule was applied six times each when calculating (23).

4.3 **r** is a cubic polynomial

To obtain the parameters a_0, a_1, a_2, a_3, l'Hospital's rule has been applied 12 times. The resulting formulae are too long to be shown here in full. Suffice it to say that the structure of the expressions obtained for the a_i is analogous to (21) and (23), i.e. it is

$$a_i = \frac{\gamma_{i,0} + \gamma_{i,1} n + \gamma_{i,2} n^2 + \gamma_{i,3} n^3}{\kappa_0 + \kappa_1 n + \kappa_2 n^2 + \kappa_3 n^3}, \quad i = 0, 1, 2, 3. \tag{24}$$

As an example, we show the coefficient of the third power of n in the denominator of (24),

$$\begin{aligned}
\kappa_3 = {} & 31933440\, g_1\, g_3\, g_5 + 31933440\, g_3\, g_5 - 23950080\, g_2^2\, g_5 + 47900160\, g_1\, g_2\, g_5\\
& + 143700480\, g_2\, g_5 + 95800320\, g_1\, g_5 - 29937600\, g_1\, g_4^2 - 29937600\, g_4^2\\
& + 39916800\, g_2\, g_3\, g_4 - 79833600\, g_3\, g_4 - 119750400\, g_2^2\, g_4\\
& + 239500800\, g_1\, g_2\, g_4 + 958003200\, g_2\, g_4 + 958003200\, g_1\, g_4 - 17740800\, g_3^3\\
& + 79833600\, g_2\, g_3^2 - 106444800\, g_1\, g_3^2 - 585446400\, g_3^2 - 79833600\, g_2^2\, g_3\\
& + 159667200\, g_1\, g_2\, g_3 + 479001600\, g_2\, g_3 + 2235340800\, g_1\, g_3 - 479001600\, g_2^2\\
& + 958003200\, g_1\, g_2.
\end{aligned}$$

5 Simulation study

Assume that $t_1, \dots, t_\ell$ are ℓ independent samples of T. Our objective here is to estimate the parameters $a_0, \dots, a_m$ based on this sample. We want to discuss results for a simulation study with $m = 2$.

5.1 The experiment

Let us assume that the random lifetime T of some technical equipment is measured in the number of *whole* years until it fails. Assume also that the item's maximum lifetime is $n = 20$ years, and that the failure rate function is quadratic with

$$r_k = \frac{2}{17} - \frac{k}{34} + \frac{k^2}{272} = \frac{(k-4)^2 + 16}{272}, \quad k = 0, \dots, 20. \tag{25}$$

From (25) it is immediately seen that (2) and (3) hold with $n = 20$ and r takes its minimum at $k = 4$. The motivation for definig r by (25) is that it resembles a 'bathtub' shaped failure rate curve frequently considered in Reliability Theory.

By (17), the ith sample factorial moment

$$\hat{\phi}_i = \frac{1}{\ell} \sum_{j=1}^{\ell} t_j(t_j - 1) \dots (t_j - i + 1)$$

is an estimator of g_i.

The results of the simulation study are shown in Table 1; the estimates of a_0, a_1, a_2 in Table 1 are based on Section 4.2. (We note in passing that the simulation study was implemented in the functional programming language Haskell.) It is seen from Table 1 that the estimated coefficients of the fail-

Table 1 Estimating a_0, a_1, a_2 and their actual values

Sample size ℓ	*Estimates of*		
	a_0	a_1	a_2
10	0.1025480	−0.03447523	0.003967391
50	0.1350796	−0.03251291	0.003787946
100	0.1329796	−0.03060992	0.003698047
500	0.1229448	−0.02844011	0.003614644
1,000	0.1208637	−0.02777423	0.003586552
5,000	0.1142492	−0.02707575	0.003568164
10,000	0.1150938	−0.02773551	0.003599041
'Exact' values	0.1176471	−0.02941176	0.003676471

ure rate function tend to their exact counterpart very rapidly. (This is hardly surprising as the sample factorial moment is an asymptotically consistent estimator of the corresponding population factorial moment and it is known from in Sect. 4.2 that the coefficients a_i are *continuous* functions of the factorial moments.) At this stage there is no assessment of the speed of the convergence available.

5.2 Evaluation and comments

The experiment simulates the situation where the failure rate function is known to be well described by a quadratic polynomial. The estimates of the coefficients are seen to approximate very well their respective exact values even for small sample sizes. (Notice that the coefficients of the failure rate based on simulation for $\ell \geq 500$ are very close to their respective exact counterpart.)

It should be pointed out that the symbolic formulae for the parameters a in Section 4 will have to be worked out *once* only. Therefore, a database of them can be prepared by running MAXIMA once only in advance of any subsequent numerical work. The separation of symbolic and numerical computations will thus make the procedure described here more acceptable for practical purposes.

6 Conclusions and further work

A class of discrete lifetime distributions on a finite lattice with polynomial failure rate was considered. A differential equation for the pgf was shown to hold which then allowed the polynomial coefficients to be expressed in terms of the pgf and its derivatives. Using computer algebra, symbolic expressions were derived in some special cases for the polynomial coefficients in terms of the factorial moments of the distribution. A simulation study with an assumed second order failure rate function resembling a bathtub curve was used to explore the question to what extent the technique can be used to estimate the polynomial coefficients.

The estimation technique described here resembles the maximum likelihood method in that the estimates are solutions of (systems of) equations. The technique is, however, also reminiscent of the method of moments in that the (unknown) coefficients in the equations are estimated by sample factorial moments.

The system of equations is laborious to solve symbolically even for low degree polynomials. Further work is needed for devising simplified (approximate) methods and also for exploring properties of the proposed estimators.

The *main* purpose of this paper was to establish the differential equation (1) for the pgf G and then to illustrate its use by obtaining estimators for the coefficients a. A possible avenue for follow-up work is, for instance, by exploring properties of the estimators themselves. By the *Delta Method* (e.g. [14]), it should be possible to show that they are asymptotically normal. (For recent uses of the *Delta Method*, see, for example, [8], [16] and [26].) This result then could be used for obtaining asymptotic tests and confidence regions for the coefficients a. (A pertinent practical problem for an assumed quadratic failure rate function is, for example, testing $H_0 : a_2 = 0$ versus $H_1 : a_2 \neq 0$.)

The recent paper [15] describes a general, semi-Markov framework where the present ideas may well be applicable.

Future work may also involve exploring the connection of the subject matter of the present paper with [4] and [5].

7 Addendum and Acknowledgment

In this paper attention is focused on the differential equation for the pgf of the time to failure distribution; and, the equation and its consequences are claimed to be novel. The observations on the Markov framework in Section 3.2 are incidental and are not essential for the main line of argument. Nevertheless, as has been kindly pointed out by the referee to the author, the Markov approach (involving phase-type distributions) allows the factorial moments of T in (15) to be written in terms of the matrix $\mathbf{Q}$ in (14) (see [18, Chapter 2]). The two methods may be compared in subsequent work.

8 Appendix: Proofs

We start by proving Proposition 1 as it will be needed in the proof of Theorem 1.

Proof of Proposition 1. Let the coefficients $u_0, \ldots, u_m$ be *defined by* (9). Then, because of (11), we have

$$\begin{aligned}
\sum_{i=0}^{m} u_i z^i \frac{d^i}{dz^i}(z^k) &= u_0 z^k + \sum_{i=1}^{m} u_i z^i \frac{d^i}{dz^i}(z^k) \\
&= a_0 z^k + \sum_{i=1}^{m} \left(\sum_{s=i}^{m} a_s \sigma_s^{(i)} \right) z^i k(k-1)\ldots(k-i+1) z^{k-i} \\
&= a_0 z^k + \sum_{s=1}^{m} a_s \left(\sum_{i=1}^{s} \sigma_s^{(i)} k(k-1)\ldots(k-i+1) \right) z^k \\
&= \sum_{s=0}^{m} a_s k^s z^k,
\end{aligned}$$

i.e. the $\mathbf{u}$ satisfy (12). Likewise, if the coefficients $v_0, \ldots, v_m$ are *defined by* (10), then they satisfy (13) since

$$\begin{aligned}
k\sum_{i=0}^{m} v_i z^{i+1}\frac{d^i}{dz^i}(z^{k-1}) &= \sum_{i=0}^{m} v_i k(k-1)(k-2)\ldots(k-i)z^k \\
&= \sum_{j=1}^{m+1} v_{j-1}k(k-1)\ldots(k-j+1)z^k \\
&= \sum_{j=1}^{m+1}\left(\sum_{s=j}^{m+1} a_{s-1}\sigma_s^{(j)}\right)k(k-1)\ldots(k-j+1)z^k \\
&= \sum_{s=1}^{m+1} a_{s-1}\sum_{j=1}^{s}\sigma_s^{(j)}k(k-1)\ldots(k-j+1)z^k \\
&= \sum_{s=1}^{m+1} a_{s-1}k^s z^k = k(a_0 + a_1 k + \ldots + a_m k^m)z^k.
\end{aligned}$$

The *uniqueness* of the values $\mathbf{u}$ solving (12) is assured, because of the linearity of (12) (in $\mathbf{a}$ and $\mathbf{u}$), if it can be shown that $a_0 = \ldots = a_m = 0$ implies that all the $\mathbf{u}$ are zero. The inference is by induction on $k = 0, 1, \ldots$.

It is seen from (12) by $k = 0$ that $u_0 = 0$. (As usual, the 0th power of any number is unity.)

Let us assume that $0 = u_0 = \ldots = u_k$ for $k \leq m-1$. Then, $u_{k+1} = 0$ since

$$0 \equiv \sum_{i=0}^{m} u_i z^i \frac{d^i}{dz^i}(z^{k+1}) = \sum_{i=k+1}^{m} u_i z^i \frac{d^i}{dz^i}(z^{k+1}) = (k+1)!\; u_{k+1}z^{k+1}.$$

A similar reasoning also shows that the coefficients $\mathbf{v}$ are *unique* solutions of (13).

Proof of Theorem 1. By the definition of the conditional probability in (1), the probability mass function of T for $k \in \mathcal{T}$ is given by

$$P(T=k) = r_k(1 - P(T \leq k-1)) = \left(\sum_{i=0}^{m} a_i k^i\right)(1 - P(T \leq k-1)). \quad (26)$$

Now, we get (8) by the following sequence of equations:

$$
\begin{aligned}
G(z) &= \sum_{k=0}^{n} z^k P(T=k) && (27)\\
&= \sum_{k=0}^{n} \left(\sum_{i=0}^{m} a_i k^i \right) z^k - \sum_{k=0}^{n} \left(\sum_{i=0}^{m} a_i k^i \right) z^k P(T \le k-1) && (28)\\
&= \sum_{k=0}^{n} \left(\sum_{i=0}^{m} a_i k^i \right) z^k - \sum_{k=1}^{n} \left(\sum_{i=0}^{m} a_i k^i \right) z^k P(T \le k-1) && (29)\\
&= \sum_{k=0}^{n} \sum_{i=0}^{m} u_i z^i \frac{d^i}{dz^i}(z^k) - \sum_{k=1}^{n} \sum_{i=0}^{m} v_i z^{i+1} \frac{d^i}{dz^i}(z^{k-1}) P(T \le k-1) && (30)\\
&= \sum_{i=0}^{m} u_i z^i \frac{d^i}{dz^i} \left(\sum_{k=0}^{n} z^k \right) - \sum_{i=0}^{m} v_i z^{i+1} \frac{d^i}{dz^i} \left(\sum_{k=1}^{n} z^{k-1} P(T \le k-1) \right) && (31)\\
&= \sum_{i=0}^{m} u_i z^i \frac{d^i}{dz^i} \left(\frac{1-z^{n+1}}{1-z} \right) - \sum_{i=0}^{m} v_i z^{i+1} \frac{d^i}{dz^i} \left(\frac{G(z)-z^n}{1-z} \right). && (32)
\end{aligned}
$$

The justifications of the steps leading from (27) through to (32) are indicated in Table 2.

Table 2 Justifying (27) – (32)

From Eqn Nr.	*To Eqn Nr.*	*Justification*
(27)	(28)	(26)
(28)	(29)	$P(T \le -1) = 0$
(29)	(30)	(12) and (13) in Proposition 1
(30)	(31)	Interchange summation and differentiation
(31)	(32)	Geometric summation and Lemma 1

The last step in the above reasoning is by the following Lemma 1.

Lemma 1 *The pgf G of any discrete random variable T on the integers $\{0, \ldots, n\}$ satisfies for $|z| < 1$ the equation*

$$\sum_{\ell=0}^{n-1} z^\ell P(T \le \ell) = \frac{G(z) - z^n}{1-z}. \tag{33}$$

Proof. We have

$$
\begin{aligned}
\sum_{\ell=0}^{\infty} z^\ell P(T \le \ell) &= \sum_{\ell=0}^{\infty} z^\ell \sum_{i=0}^{\ell} P(T=i) = \sum_{0 \le i \le \ell < \infty} z^\ell P(T=i)\\
&= \sum_{i=0}^{\infty} P(T=i) \sum_{\ell=i}^{\infty} z^\ell = \sum_{i=0}^{\infty} P(T=i) \frac{z^i}{1-z}. && (34)
\end{aligned}
$$

Equation (33) follows from (34) in conjunction with $P(T \leq \ell) = 1, \ell \geq n$, thus

$$\sum_{\ell=0}^{n-1} z^\ell P(T \leq \ell) = \sum_{\ell=0}^{\infty} z^\ell P(T \leq \ell) - \sum_{\ell=n}^{\infty} z^\ell P(T \leq \ell)$$
$$= \frac{G(z)}{1-z} - \sum_{\ell=n}^{\infty} z^\ell P(T \leq \ell) = \frac{G(z)}{1-z} - \frac{z^n}{1-z} = \frac{G(z) - z^n}{1-z}.$$

References

1. Abramowitz M, Stegun A, Handbook of Mathematical Functions, Dover Publications, New York (1972)
2. Berg M P (1996). Towards rational age-based failure modelling, *Reliability and Maintenance of Complex Systems, Proceedings of the NATO Advanced Study Institute on Current Issues and Challenges in the Reliability and Maintenance of Complex Systems*, Kemer-Antalya, Turkey, June 12-22, 1995, Özekici S (Ed.), Springer-Verlag, NATO ASI Series F, Vol. 154, 107–113
3. Biggs N L, Discrete Mathematics, Clarendon Press, Oxford (1989)
4. Charalambides C A, Moments of a class of discrete q-distributions, Journal of Statistical Planning and Inference, 135, 64–76 (2005)
5. Crippa D, Simon K, q-distributions and Markov processes, Discrete Mathematics, 170, 81–98 (1997)
6. Csenki A, Dependability for Systems with a Partitioned State Space — Markov and Semi-Markov Theory and Computational Implementation, Lecture Notes in Statistics, Vol. 90, Springer, New York (1994)
7. Csenki A, On continuous lifetime distributions with polynomial failure rate with an application in reliability, Reliability Engineering and System Safety, 96, 1587–1590 (2011)
8. Csenki A, Asymptotics for continuous lifetime distributions with polynomial failure rate with an application in reliability, Reliability Engineering and System Safety, 102, 1–4 (2012)
9. Grimmett G, Welsh D, Probability — An Introduction. Clarendon Press, Oxford (1986)
10. Heller B, MACSYMA for Statisticians, Wiley–Interscience, New York (1991)
11. Jazi M A, Lai C D, Alamatsaz M H, A discrete inverse Weibull distribution and estimation of its parameters, Statistical Methodology, 7, 121–132 (2010)
12. Khan M S A, Kalique A, Abouammoh A M, On estimating parameters in a discrete Weibull distribution, IEEE Transactions on Reliability, 38, 348–348 (1989)
13. Lawless J F, Statistical Models and Methods for Lifetime Data, Wiley, New York (1982)
14. Lehmann E L, Elements of Large-Sample Theory, Springer, New York (1999)
15. Limnios N, Reliability measures of semi-Markov systems with general state space, Methodology and Computing in Applied Probability, Online First, 19 January 2011 (2011) doi: 10.1007/s11009-011-9211-5
16. Ma Y, Genton M G, Parzen E, Asymptotic properties of sample quantiles of discrete distributions, Annals of the Institute of Statistical Mathematics, 63, 227–243 (2011)
17. Nakagawa T A, Osaki S (1975) The discrete Weibull distribution, IEEE Transactions on Reliability, 24, 300–301 (1975)
18. Neuts M, Matrix-geometric solutions in stochastic models: An algorithmic approach. The Johns Hopkins University Press, Baltimore (1981)
19. Rand R H, Introduction to Maxima, Dept. of Theoretical and Applied Mechanics, Cornell University (2010) `http://maxima.sourceforge.net/docs/intromax/intromax.html`
20. Rényi A (1970), Probability Theory, North-Holland, Amsterdam, London (1970)
21. Sarhan A M, Hamilton D C, Smith B, Kundu D, The bivariate generalized linear failure rate distribution and its multivariate extension, Computational Statistics and Data Analysis, 55, 644–654 (2011)

22. Schelter W F, Maxima Manual, version 5.9.3, (2006)
`http://maxima.sourceforge.net/docs/manual/en/maxima.html`
23. Shaked M, Shantikumar G J, Valdez-Torres J B, Discrete hazard rate functions, Computers and Operations Research, 22, 391–402 (1995)
24. Stein W E, Dattero R, A new discrete Weibull Distribution, IEEE Transactions on Reliability, 33, 196–197 (1984)
25. Wang Z, One mixed negative binomial distribution with application, Journal of Statistical Planning and Inference, 141, 1153–1160 (2011)
26. Withers C, Nadarajah S, Stabilizing the asymptotic covariance of an estimate, Electronic Journal of Statistics, 4, 161–171 (2010)