

Article

A Phase-Type Distribution for the Sum of Two Concatenated Markov Processes Application to the Analysis Survival in Bladder Cancer

Belén García-Mora * , Cristina Santamaría  and Gregorio Rubio 

Instituto Universitario de Matemática Multidisciplinar, Universitat Politècnica de València, 46022 Valencia, Spain; crisanna@imm.upv.es (C.S.); grubio@imm.upv.es (G.R.)

* Correspondence: magarmo5@imm.upv.es

Received: 31 October 2020; Accepted: 20 November 2020; Published: 24 November 2020



Abstract: Stochastic processes are useful and important for modeling the evolution of processes that take different states over time, a situation frequently found in fields such as medical research and engineering. In a previous paper and within this framework, we developed the sum of two independent phase-type (PH)-distributed variables, each of them being associated with a Markovian process of one absorbing state. In that analysis, we computed the distribution function, and its associated survival function, of the sum of both variables, also PH-distributed. In this work, in one more step, we have developed a first approximation of that distribution function in order to avoid the calculation of an inverse matrix for the possibility of a bad conditioning of the matrix, involved in the expression of the distribution function in the previous paper. Next, in a second step, we improve this result, giving a second, more accurate approximation. Two numerical applications, one with simulated data and the other one with bladder cancer data, are used to illustrate the two proposed approaches to the distribution function. We compare and argue the accuracy and precision of each one of them by means of their error bound and the application to real data of bladder cancer.

Keywords: bladder cancer; Fréchet derivative; Kronecker product; Markov process; phase-type distribution; survival analysis

1. Introduction

Stochastic processes have proved to be useful in the evolution of processes that can take different states over time. In this regard, Markov models have been widely used for this purpose in fields like medical research and engineering as well as the phase-type distributions. A phase-type (PH) distribution is the distribution of the time until absorption in a finite-state absorbing Markov chain [1–4]. A PH distribution is represented by (α, T) where α is an initial probability vector and T is a squared matrix representing the rates between the transient states in the Markov process. Phase-type distributions are a powerful tool in stochastic models of real systems. Numerous applications have been reported in queueing theory models [5] and reliability in the context to model the failure of electrical components [6], in bladder cancer [7] in the context of survival analysis, and applications in shock and wear systems [8], among others. These distributions also arise in the evolution of some chronic diseases since the process goes through a series of states or phases [9,10] and in applications to the length of stay at hospitals [11–13]. See Chapter 1 in [14,15] for a sample of the diversity of contexts where the phase-type distributions are used (telecommunications, finance, teletraffic modeling, biostatistics, drug kinetics, and survival analysis).

On the other hand, one of the major interests in cancer research is to model its evolution from the beginning of the disease until death, going through a number of states before reaching the absorbing

state (death). However, sometimes, the real data do not include the complete evolution of the disease because this evolution is treated in different and independent units within the same department or hospital and consequently the real data are registered separately. For this reason, given the need to describe the complete evolution of the bladder cancer (from a primary tumor to the extirpation of the bladder), in a previous paper [16], we concatenated two Markov processes. Each one analyzed one part of the evolution of this chronic disease because in that study, we had two different and independent real databases belonging to different units of the La Fe University Hospital in Valencia (Spain). The first database described the disease progression from a primary tumor through several states to a more aggressive tumor. The second database described the evolution from the aggressive tumors to the bladder extirpation (death of the bladder), also through several states. Both bases were disconnected and our objective was to connect both and be able to study the evolution of the whole disease divided in two phases from the primary tumor (start state in the first database) to the removal of the bladder (absorbing state in the second database).

In the modeling of these two phases of the disease, we considered two consecutive homogeneous Markov processes with state spaces $\{1, 2, \dots, m+1\}$ and $\{1, 2, \dots, n+1\}$, respectively, and glued together to become a unique continuous-time Markov chain with $m+n+1$ states, identifying the state $m+1$ with the first state of the second Markov chain. The absorbing state of the resulting concatenated Markov process was the $m+n+1$ state. We have also considered two random continuous and independent variables representing two absorption times: the absorption time of the first process (the appearance of the first aggressive or progression tumor) and the absorption time of the second one (the bladder extirpation). Both variables were considered PH-distributed with representations (α, T) and (β, S) , respectively, with α and β initial probability vectors in each process. T and S were squared matrices of orders m and n , respectively, representing the rates between the transient states in each process. In order to study the survival function of the absorption time of this concatenated process ($m+n+1$ state), we obtained a new distribution function of the sum of these two variables, also PH-distributed.

Thinking about the practical application of our approach, for example, if we find large matrices because of dealing with many states, we aim to build approximations that facilitate the task. In Section 2, we present the distribution function of the previous paper [16]. In Section 3, we develop a first approximation of this distribution function to avoid the calculation of an inverse matrix in its expression due to the possibility of bad conditioning of the matrix. We start developing the expression of that inverse by a numeric series. We also compute an error bound for this first approximation. In a second step, we improve this result with a new and more accurate approximation. We have also developed an error bound for this second approach that allows us to compare them and improve the result of this study. In Section 4, we illustrate the results of both approaches with a numerical application of simulated data. We compare them with the exact distribution function in [16]. In Section 5, we apply this methodology to real data of bladder cancer, and we argue the improvement and the usefulness of the second approach compared to the first one. Finally, we conclude the study in Section 6 with some discussion.

2. The Survival Function of Two Concatenated Markov Processes

Throughout this paper, we use the Kronecker matrix form just like the Kronecker sum and product, denoted with \oplus and \otimes (mathematical notation), respectively. The main definitions about these two concepts as well as the properties are in Appendix A. We start with the following result proved in [16].

Theorem 1. *Let X_1 and X_2 be nonnegative random two independent variables representing the absorbing times in the Markov chains with state spaces $\{1, 2, \dots, m, m+1\}$ and $\{1, 2, \dots, n, n+1\}$, respectively. The groups of states $\{1, 2, \dots, m\}$ and $\{1, 2, \dots, n\}$ are the transient states in each Markov process and the states $m+1$ and $n+1$ are the absorbing ones. X_1 and X_2 are PH-distributed and the representations are (α, T) and (β, S) , respectively. Thus, the survival function $S(x) = P(X > x) = 1 - F(x)$ for the sum $X_1 + X_2$ is given by*

$$S(x) = \alpha e^{Tx} \mathbf{e}_m + \alpha_{m+1} \beta e^{Sx} \mathbf{e}_n + (\mathbf{e}_n' \otimes \alpha) \int_0^1 e^{sS'x \oplus (1-s)Tx} ds \text{vec}(T^0 \beta x) \tag{1}$$

where α_{m+1} is the initial probability of reaching the absorbing state $m + 1$, $\mathbf{e}_k = (1, 1, \dots, 1)' \in \mathbb{R}^{k \times 1}$, $T^0 = -T\mathbf{e}_m$, being T^0 a matrix of order $m \times 1$ with the entries $t_{i,m+1}$, $i = 1, \dots, m$, that represent the absorbing rates from the transient states and the function $\text{vec}()$ stacks the columns of a matrix into a column vector (defined in Appendix A). The expression for the calculated integral is

$$\int_0^1 e^{sS'x \oplus (1-s)Tx} ds = [S'x \oplus (-Tx)]^{-1} \left(e^{S'x} \otimes I_m - I_n \otimes e^{Tx} \right) \tag{2}$$

See Results (15)–(17) in [16].

3. The Approximated Survival Function for Two Concatenated Markov Processes

The calculation of the inverse of the previous matrix $S'x \oplus (-Tx)$ given in Equation (2) can present serious difficulties if this matrix is badly conditioned. To avoid a possible bad conditioning, we perform one approximation for the survival function $S(x)$.

3.1. An Approximated Survival Function

First, we consider the distribution function, $F(x) = 1 - S(x)$, using the expression given in Equation (1),

$$F(x) = 1 - \alpha e^{Tx} \mathbf{e}_m - \alpha_{m+1} \beta e^{Sx} \mathbf{e}_n - (\mathbf{e}_n' \otimes \alpha) \int_0^1 e^{sS'x \oplus (1-s)Tx} ds \text{vec}(T^0 \beta x), \tag{3}$$

and so on, let us calculate an approximation for the term $\int_0^1 e^{sS'x \oplus (1-s)Tx} ds$. For that purpose, we use the expansion of the Taylor series for an exponential function in the integrand of Equation (3). We denote this integrand by the function $f(S, T, s, x) = e^{sS'x \oplus (1-s)Tx}$

$$f(S, T, s, x) = I_{m \times n} + (S's \oplus T(1-s))x + \frac{1}{2!} (S's \oplus T(1-s))^2 x^2 + \dots \dots + \frac{1}{k!} (S's \oplus T(1-s))^k x^k + \dots \tag{4}$$

Let us call the k th term of Equation (4) by the function $f_k(S, T, s, x)$ and taking into account that

$$\|f_k(S, T, s, x)\|_2 = \frac{1}{k!} \|(S's \oplus T(1-s))^k\|_2 x^k \leq \frac{1}{k!} (\|S'\|_2 + \|T\|_2)^k x^k \tag{5}$$

then

$$\sum_{k=0}^{\infty} \frac{1}{k!} (\|S'\|_2 + \|T\|_2)^k x^k = e^{(\|S'\|_2 + \|T\|_2)x} \tag{6}$$

is a convergent series of constant terms. Then, applying the Weierstrass criterion of uniform convergence, the Taylor series expansion $f(S, T, s, x) = \sum_{k=0}^{\infty} f_k(S, T, s, x)$ converges uniformly in $s \in [0, 1]$. Thus, the integral in Equation (3) can be obtained by integrating Equation (4) term by term. For this, as $S' \otimes I$ and $I \otimes T$ commute (see the demonstration in Appendix A for $k = 2$), the binomial of Newton can be used in each term $f_k(S, T, s, x)$ and so one gets to

$$\begin{aligned} \sum_{k=0}^{\infty} f_k(S, T, s, x) &= I_{m \times n} + (S's \oplus T(1-s))x + \frac{x^2}{2!} \sum_{j=0}^2 \binom{2}{j} (S')^{2-j} \otimes T^j s^{2-j} (1-s)^j + \dots \\ &\dots + \frac{x^k}{k!} \sum_{j=0}^k \binom{k}{j} (S')^{k-j} \otimes T^j s^{k-j} (1-s)^j + \dots \end{aligned} \tag{7}$$

Integrating each term of Equation (7), it follows that

$$\begin{aligned}
 I(x) &= \int_0^1 e^{sS'x \oplus (1-s)Tx} ds \\
 &= I_{m \times n} + \left[S' \int_0^1 s ds \oplus T \int_0^1 (1-s) ds \right] x + \dots \\
 &\quad + \frac{x^k}{k!} \sum_{j=0}^k (S')^{k-j} \otimes T^j \binom{k}{j} \int_0^1 s^{k-j} (1-s)^j ds + \dots \\
 &= I_{m \times n} + \frac{1}{2} (S' \oplus T)x + \dots + \frac{x^k}{k!} \sum_{j=0}^k (S')^{k-j} \otimes T^j \binom{k}{j} B(j+1, k-j+1) + \dots \\
 &= I_{m \times n} + \frac{1}{2} (S' \oplus T)x + \dots + \frac{x^k}{k!} \sum_{j=0}^k (S')^{k-j} \otimes T^j \binom{k}{j} \frac{\Gamma(j+1)\Gamma(k-j+1)}{\Gamma(k+2)} + \dots \\
 &= I_{m \times n} + \frac{1}{2} (S' \oplus T)x + \dots + \frac{x^k}{k!} \sum_{j=0}^k (S')^{k-j} \otimes T^j \frac{k!}{j!(k-j)!} \frac{j!(k-j)!}{(k+1)!} + \dots \\
 &= I_{m \times n} + \sum_{k \geq 1} \frac{x^k}{(k+1)!} \sum_{j=0}^k (S')^{k-j} \otimes T^j
 \end{aligned} \tag{8}$$

Now, if we take in Equation (8), the following finite sum until the p th term, we have

$$\phi_p(x) = I_{m \times n} + \sum_{k=1}^p \frac{x^k}{(k+1)!} \sum_{j=0}^k (S')^{k-j} \otimes T^j \tag{9}$$

as an approximation to $I(x)$ and we can substitute Expression (9) by the integral of the distribution function Equation (3). Then, $\hat{F}_1(x)$ will represent an approximation of the distribution function $F(x)$,

$$\begin{aligned}
 \hat{F}_1(x) &= 1 - \alpha e^{Tx} \mathbf{e}_m - \alpha_{m+1} \beta e^{Sx} \mathbf{e}_n - \\
 &\quad \left(\mathbf{e}_n' \otimes \alpha \right) \left(I_{m \times n} + \sum_{k=1}^p \frac{x^k}{(k+1)!} \sum_{j=0}^k (S')^{k-j} \otimes T^j \right) \text{vec} (T^0 \beta x)
 \end{aligned} \tag{10}$$

and consequently, an approximation of the survival function would be

$$\begin{aligned}
 \hat{S}_1(x) &= \alpha e^{Tx} \mathbf{e}_m + \alpha_{m+1} \beta e^{Sx} \mathbf{e}_n + \\
 &\quad \left(\mathbf{e}_n' \otimes \alpha \right) \left(I_{m \times n} + \sum_{k=1}^p \frac{x^k}{(k+1)!} \sum_{j=0}^k (S')^{k-j} \otimes T^j \right) \text{vec} (T^0 \beta x)
 \end{aligned} \tag{11}$$

We are interested in obtaining an error bound for this first approach. For this, if $M = \max\{\|S\|_2, \|T\|_2\}$, it follows that

$$\begin{aligned}
 \|I(x) - \phi_p(x)\|_2 &= \left\| \sum_{k \geq p+1} \frac{x^k}{(k+1)!} \sum_{j=0}^k (S')^{k-j} \otimes T^j \right\|_2 \\
 &\leq \sum_{k \geq p+1} \frac{x^k}{(k+1)!} \sum_{j=0}^k \|S'\|_2^{k-j} \|T\|_2^j \\
 &\leq \sum_{k \geq p+1} \frac{x^k}{(k+1)!} (k+1) M^k \\
 &= \sum_{k \geq p+1} \frac{x^k}{k!} M^k = \frac{(Mx)^{p+1}}{(p+1)!} e^{M\theta x}
 \end{aligned} \tag{12}$$

where $0 < \theta < 1$, and so

$$\begin{aligned}
 \|S(x) - \hat{S}_1(x)\|_2 &\leq \|\mathbf{e}_n'\|_2 \|\alpha\|_2 \frac{M^{p+1} x^{p+1}}{(p+1)!} e^{M\theta x} \|\text{vec}(T^0 \beta x)\|_2 \\
 &\leq \sqrt{n} \|\alpha\|_2 \frac{M^{p+1} x^{p+2}}{(p+1)!} e^{Mx} \|T^0 \beta\|_F
 \end{aligned} \tag{13}$$

Therefore, the following result has been proved:

Theorem 2. Let X_1 and X_2 be nonnegative random independent variables representing the absorption times in two homogeneous Markov processes with state space $\{1, 2, \dots, m, m+1\}$ and $\{1, 2, \dots, n, n+1\}$, respectively, where $\{1, 2, \dots, m\}$ and $\{1, 2, \dots, n\}$ are the transient states in each process and $m+1$ and $n+1$ are the absorbing ones. Assume that both variables are PH-distributed with representation (α, T) and (β, S) , respectively. Then, an approximation of the survival function for the sum $X_1 + X_2$ is given by Equation (11), where α_{m+1} is the

initial probability of entering the absorbing state $m + 1$, $\mathbf{e}_k = (1, 1, \dots, 1)' \in \mathbb{R}^{k \times 1}$, and $T^0 = -T\mathbf{e}_m$. Moreover, an error bound of the approximation error is given in Expression (13), where $M = \max\{\|S\|_2, \|T\|_2\}$ and $\|\cdot\|_F$ is the Frobenius norm.

Notice that the error bound increases exponentially with time x , and consequently, for predictions, it would not be accurate. Thus, in the following section, we propose a second approximation method for the distribution function to improve this last obtained result.

3.2. Improving the Approximation to the Survival Function

The aim of this section is to improve $\hat{F}_1(x)$ in order to get a closer approximation to $F(x)$. For this, we use the expression of $F(x)$ obtained in our previous paper (see Equation (15) in [16]) given by the expression

$$F(x) = 1 - \alpha e^{Tx} \mathbf{e}_m - \left(\alpha_{m+1} \beta e^{Sx} + \alpha \int_0^1 e^{(1-s)Tx} T^0 \beta x e^{sSx} ds \right) \mathbf{e}_n \tag{14}$$

In [16], it is demonstrated, step by step, that Equation (14) is equivalent to $1 - S(x)$ in Equation (1) at the beginning of this paper (see Result (17) in [16]). For our purpose, we start working with the last term of Equation (14), specifically with

$$\alpha \int_0^1 e^{(1-s)Tx} T^0 \beta x e^{sSx} ds \mathbf{e}_n \tag{15}$$

We are going to reduce the previous bound Inequality (13). For this, let us consider the following expression with $k \geq 0$ positive integer

$$\int_0^1 e^{(1-s)\frac{Tx}{2^k}} \frac{T^0 \beta x}{2^k} e^{s\frac{Sx}{2^k}} ds, \tag{16}$$

and taking into account the Schur–Fréchet derivative and its function composition [17,18], we have the following expression

$$\int_0^1 e^{(1-s)\frac{Tx}{2^j}} \frac{T^0 \beta x}{2^j} e^{s\frac{Sx}{2^j}} ds e^{\frac{Sx}{2^j}} + e^{\frac{Tx}{2^j}} \int_0^1 e^{(1-s)\frac{Tx}{2^j}} \frac{T^0 \beta x}{2^j} e^{s\frac{Sx}{2^j}} ds \tag{17}$$

for $j = 1, 2, \dots, k$. See the demonstration of Expression (17) in Appendix B.

Now applying the function *vec* for both terms of this last expression, and using property 1 of Appendix A, we have for the first term

$$\text{vec} \left[\int_0^1 e^{(1-s)\frac{Tx}{2^{j-1}}} \frac{T^0 \beta x}{2^{j-1}} e^{s\frac{Sx}{2^{j-1}}} ds \right] = \left(\int_0^1 e^{s\frac{Sx}{2^{j-1}}} \otimes e^{(1-s)\frac{Tx}{2^{j-1}}} ds \right) \text{vec} \left(\frac{T^0 \beta x}{2^{j-1}} \right) \tag{18}$$

Now, in the second term of Expression (17), we have

$$\begin{aligned} & \text{vec} \left[\int_0^1 e^{(1-s)\frac{Tx}{2^j}} \frac{T^0 \beta x}{2^j} e^{s\frac{Sx}{2^j}} ds e^{\frac{Sx}{2^j}} + e^{\frac{Tx}{2^j}} \int_0^1 e^{(1-s)\frac{Tx}{2^j}} \frac{T^0 \beta x}{2^j} e^{s\frac{Sx}{2^j}} ds \right] \\ &= \text{vec} \left[\int_0^1 e^{(1-s)\frac{Tx}{2^j}} \frac{T^0 \beta x}{2^j} e^{(1+s)\frac{Sx}{2^j}} ds + \int_0^1 e^{(2-s)\frac{Tx}{2^j}} \frac{T^0 \beta x}{2^j} e^{s\frac{Sx}{2^j}} ds \right] \\ &= \left(\int_0^1 e^{(1+s)\frac{Sx}{2^j}} \otimes e^{(1-s)\frac{Tx}{2^j}} ds + \int_0^1 e^{s\frac{Sx}{2^j}} \otimes e^{(2-s)\frac{Tx}{2^j}} ds \right) \text{vec} \left(\frac{T^0 \beta x}{2^j} \right) \end{aligned} \tag{19}$$

Notice that in Equation (19), property 1 of Appendix A is used again, and in the first and second steps of Equation (20), properties 2 and 3, respectively. In the last two steps of Equation (20), properties 3 and 4 of Appendix A are used.

$$\begin{aligned}
 & \left[\left(e^{\frac{s'x}{2^j}} \otimes I_m \right) \int_0^1 e^{s\frac{s'x}{2^j}} \otimes e^{(1-s)\frac{T_x}{2^j}} ds + \int_0^1 e^{s\frac{s'x}{2^j}} \otimes e^{(1-s)\frac{T_x}{2^j}} ds \left(I_n \otimes e^{\frac{T_x}{2^j}} \right) \right] \text{vec} \left(\frac{T^0 \beta x}{2^j} \right) \\
 = & \left[\left(e^{\frac{s'x}{2^j}} \otimes I_m \right) \int_0^1 e^{s\frac{s'x}{2^j} \oplus (1-s)\frac{T_x}{2^j}} ds + \int_0^1 e^{s\frac{s'x}{2^j} \oplus (1-s)\frac{T_x}{2^j}} ds \left(I_n \otimes e^{\frac{T_x}{2^j}} \right) \right] \text{vec} \left(\frac{T^0 \beta x}{2^j} \right) \\
 = & \left[\left(e^{\frac{s'x}{2^j}} \otimes I_m \right) \int_0^1 e^{s\frac{s'x}{2^j} \oplus (1-s)\frac{T_x}{2^j}} ds + \left(I_n \otimes e^{\frac{T_x}{2^j}} \right) \int_0^1 e^{s\frac{s'x}{2^j} \oplus (1-s)\frac{T_x}{2^j}} ds \right] \text{vec} \left(\frac{T^0 \beta x}{2^j} \right) \tag{20} \\
 = & \left(e^{\frac{s'x}{2^j}} \otimes I_m + I_n \otimes e^{\frac{T_x}{2^j}} \right) \int_0^1 e^{s\frac{s'x}{2^j} \oplus (1-s)\frac{T_x}{2^j}} ds \text{vec} \left(\frac{T^0 \beta x}{2^j} \right) \\
 = & \left(e^{\frac{s'x}{2^j}} \otimes I_m + I_n \otimes e^{\frac{T_x}{2^j}} \right) \left(\int_0^1 e^{s\frac{s'x}{2^j}} \otimes I_m + I_n \otimes e^{(1-s)\frac{T_x}{2^j}} ds \right) \text{vec} \left(\frac{T^0 \beta x}{2^j} \right)
 \end{aligned}$$

Now, taking the equality Equations (18) and (20) into the above expression and commuting the integral $\int_0^1 e^{s\frac{s'x}{2^j}} \otimes I + I \otimes e^{(1-s)\frac{T_x}{2^j}} ds$, we have for $j = 1, 2, \dots, k$

$$\begin{aligned}
 & \left(\int_0^1 e^{s\frac{s'x}{2^{j-1}}} \otimes e^{(1-s)\frac{T_x}{2^{j-1}}} ds \right) \text{vec} \left(\frac{T^0 \beta x}{2^{j-1}} \right) = \\
 & \left(e^{\frac{s'x}{2^j}} \otimes I_m + I_n \otimes e^{\frac{T_x}{2^j}} \right) \left(\int_0^1 e^{s\frac{s'x}{2^j}} \otimes I_m + I_n \otimes e^{(1-s)\frac{T_x}{2^j}} ds \right) \text{vec} \left(\frac{T^0 \beta x}{2^j} \right) \tag{21}
 \end{aligned}$$

Let us call $I_j = \int_0^1 e^{s\frac{s'x}{2^j}} \otimes I_m + I_n \otimes e^{(1-s)\frac{T_x}{2^j}} ds$, $j = 0, 1, \dots, k$; thus, Expression (21) is

$$I_{j-1} \text{vec} \left(\frac{T^0 \beta x}{2^{j-1}} \right) = \left(e^{\frac{s'x}{2^j}} \otimes I_m + I_n \otimes e^{\frac{T_x}{2^j}} \right) I_j \text{vec} \left(\frac{T^0 \beta x}{2^j} \right) \quad \text{for } j = 1, 2, \dots, k \tag{22}$$

If now we denote $\phi_{k,p}$ the approximated function Equation (9) for the integral

$$I_k = \int_0^1 e^{s\frac{s'x}{2^k}} \otimes I_m + I_n \otimes e^{(1-s)\frac{T_x}{2^k}} ds,$$

it follows that

$$\phi_{k,p} \text{vec} \left(\frac{T^0 \beta x}{2^k} \right) \simeq I_k \text{vec} \left(\frac{T^0 \beta x}{2^k} \right)$$

Now, we define

$$\phi_{j-1,p} = \frac{e^{\frac{s'x}{2^j}} \otimes I_m + I_n \otimes e^{\frac{T_x}{2^j}}}{2} \phi_{j,p} \quad \text{for } j = k, k-1, \dots, 1$$

Then, for $j = k$, and taking Equation (22) into account, we have

$$\begin{aligned}
 \phi_{k-1,p} \text{vec} \left(\frac{T^0 \beta x}{2^{k-1}} \right) &= \frac{e^{\frac{s'x}{2^k}} \otimes I_m + I_n \otimes e^{\frac{T_x}{2^k}}}{2} \phi_{k,p} \text{vec} \left(\frac{T^0 \beta x}{2^k} \right) \simeq \\
 & \left(e^{\frac{s'x}{2^k}} \otimes I_m + I_n \otimes e^{\frac{T_x}{2^k}} \right) I_k \text{vec} \left(\frac{T^0 \beta x}{2^k} \right) = I_{k-1} \text{vec} \left(\frac{T^0 \beta x}{2^{k-1}} \right) \tag{23}
 \end{aligned}$$

Similarly, for $j = k - 1$, one gets

$$\begin{aligned} \phi_{k-2,p} \operatorname{vec} \left(\frac{T^0 \beta x}{2^{k-2}} \right) &= \frac{e^{\frac{s'x}{2^{k-1}} \otimes I_m + I_n} \otimes e^{\frac{Tx}{2^{k-1}}}}{2} \phi_{k-1,p} \operatorname{vec} \left(\frac{T^0 \beta x}{2^{k-2}} \right) \simeq \\ &\left(e^{\frac{s'x}{2^{k-1}} \otimes I_m + I_n} \otimes e^{\frac{Tx}{2^{k-1}}} \right) I_{k-1} \operatorname{vec} \left(\frac{T^0 \beta x}{2^{k-1}} \right) = I_{k-2} \operatorname{vec} \left(\frac{T^0 \beta x}{2^{k-2}} \right) \end{aligned} \tag{24}$$

and finally, for $j = 1$, we have

$$\begin{aligned} \phi_{0,p} \operatorname{vec} (T^0 \beta x) &= \frac{e^{\frac{s'x}{2} \otimes I_m + I_n} \otimes e^{\frac{Tx}{2}}}{2} \phi_{1,p} \operatorname{vec} (T^0 \beta x) \simeq \\ &\left(e^{\frac{s'x}{2} \otimes I_m + I_n} \otimes e^{\frac{Tx}{2}} \right) I_1 \operatorname{vec} \left(\frac{T^0 \beta x}{2} \right) = I_0 \operatorname{vec} (T^0 \beta x) \end{aligned} \tag{25}$$

and so

$$\phi_{0,p} \operatorname{vec} (T^0 \beta x) \approx I_0 \operatorname{vec} (T^0 \beta x) \tag{26}$$

where $I_0 = \int_0^1 e^{sS'x} \otimes I_m + I_n \otimes e^{(1-s)Tx} ds = \int_0^1 e^{sS'x} \oplus e^{(1-s)Tx} ds$. Thus, the approximation for $I_0 \operatorname{vec} (T^0 \beta x)$ is

$$\begin{aligned} \phi_{0,p} \operatorname{vec} (T^0 \beta x) &= \left(e^{\frac{s'x}{2} \otimes I_m + I_n} \otimes e^{\frac{Tx}{2}} \right) \phi_{1,p} \operatorname{vec} \left(\frac{T^0 \beta x}{2} \right) \\ &= \left(e^{\frac{s'x}{2} \otimes I_m + I_n} \otimes e^{\frac{Tx}{2}} \right) \left(e^{\frac{s'x}{2^2} \otimes I_m + I_n} \otimes e^{\frac{Tx}{2^2}} \right) \phi_{2,p} \operatorname{vec} \left(\frac{T^0 \beta x}{2^2} \right) \\ &\quad \vdots \\ &= \left(e^{\frac{s'x}{2} \otimes I_m + I_n} \otimes e^{\frac{Tx}{2}} \right) \left(e^{\frac{s'x}{2^2} \otimes I_m + I_n} \otimes e^{\frac{Tx}{2^2}} \right) \dots \\ &\quad \dots \left(e^{\frac{s'x}{2^k} \otimes I_m + I_n} \otimes e^{\frac{Tx}{2^k}} \right) \phi_{k,p} \operatorname{vec} \left(\frac{T^0 \beta x}{2^k} \right) \\ &= \prod_{j=1}^k \left(e^{\frac{s'x}{2^j} \otimes I_m + I_n} \oplus e^{\frac{Tx}{2^j}} \right) \phi_{k,p} \operatorname{vec} \left(\frac{T^0 \beta x}{2^k} \right) \end{aligned} \tag{27}$$

In this way, the second approximation for $F(x)$ is

$$\hat{F}_2(x) = 1 - \alpha e^{Tx} \mathbf{e}_m - \alpha_{m+1} \beta e^{Sx} \mathbf{e}_n - \left(\mathbf{e}_n' \otimes \alpha \right) \prod_{j=1}^k \left(e^{\frac{s'x}{2^j} \otimes I_m + I_n} \oplus e^{\frac{Tx}{2^j}} \right) \phi_{k,p} \operatorname{vec} \left(\frac{T^0 \beta x}{2^k} \right) \tag{28}$$

and consequently, the survival function is

$$\hat{S}_2(x) = \alpha e^{Tx} \mathbf{e}_m + \alpha_{m+1} \beta e^{Sx} \mathbf{e}_n + \left(\mathbf{e}_n' \otimes \alpha \right) \prod_{j=1}^k \left(e^{\frac{s'x}{2^j} \otimes I_m + I_n} \oplus e^{\frac{Tx}{2^j}} \right) \phi_{k,p} \operatorname{vec} \left(\frac{T^0 \beta x}{2^k} \right) \tag{29}$$

Finally, we construct an error bound for the difference between $F(x)$ and this second approximation $\hat{F}_2(x)$. As by Equation (22), we have

$$I_0 \operatorname{vec} (T^0 \beta x) = \left(e^{\frac{s'x}{2} \otimes I_m + I_n} \otimes e^{\frac{Tx}{2}} \right) I_1 \operatorname{vec} \left(\frac{T^0 \beta x}{2} \right) \tag{30}$$

and given that $M = \max\{\|S'\|_2, \|T\|_2\}$, we establish

$$\begin{aligned}
 & \|I_0 \text{vec}(T^0 \beta x) - \phi_{0,p} \text{vec}(T^0 \beta x)\|_2 \leq \\
 & \leq 2 e^{\frac{Mx}{2}} \left\| I_1 \text{vec}\left(\frac{T^0 \beta x}{2}\right) - \phi_{1,p} \text{vec}\left(\frac{T^0 \beta x}{2}\right) \right\|_2 \\
 & \leq 2^2 e^{\left(\frac{Mx}{2} + \frac{Mx}{4}\right)} \left\| I_2 \text{vec}\left(\frac{T^0 \beta x}{2^2}\right) - \phi_{2,p} \text{vec}\left(\frac{T^0 \beta x}{2^2}\right) \right\|_2 \\
 & \vdots \\
 & \leq 2^k e^{\frac{Mx}{2} \left(\frac{1}{2} + \frac{1}{2^2} + \dots + \frac{1}{2^k}\right)} \left\| I_k \text{vec}\left(\frac{T^0 \beta x}{2^k}\right) - \phi_{k,p} \text{vec}\left(\frac{T^0 \beta x}{2^k}\right) \right\|_2 \\
 & = e^{Mx \left(1 - \frac{1}{2^k}\right)} \left\| I_k - \phi_{k,p} \right\|_2 \|T^0 \beta x\|_F \\
 & \leq e^{Mx \left(1 - \frac{1}{2^k}\right)} x^{p+2} \left(\frac{M}{2^k}\right)^{p+1} \frac{1}{(p+1)!} \|T^0 \beta\|_F
 \end{aligned} \tag{31}$$

where the difference $\|I_k - \phi_{k,p}\|_2$ has been calculated previously with Equation (12). Thus, the error bound for the survival function and this second approximation is

$$\|S(x) - \hat{S}_2(x)\|_2 \leq \sqrt{n} \|\alpha\|_2 e^{Mx \left(1 - \frac{1}{2^k}\right)} x^{p+2} \left(\frac{M}{2^k}\right)^{p+1} \frac{1}{(p+1)!} \|T^0 \beta\|_F \tag{32}$$

We illustrate the result in the following theorem

Theorem 3. Let X_1 and X_2 be nonnegative random independent variables representing the absorption times in two homogeneous Markov processes with state space $\{1, 2, \dots, m, m + 1\}$ and $\{1, 2, \dots, n, n + 1\}$, respectively, where $\{1, 2, \dots, m\}$ and $\{1, 2, \dots, n\}$ are the transient states in each process and $m + 1$ and $n + 1$ are the absorbing ones. Assume that both variables are PH-distributed with representation (α, T) and (β, S) , respectively. Then an approximation of the survival function for the sum $X_1 + X_2$ is given in Equation (29), where α_{m+1} is the initial probability of entering the absorbing state $m + 1$, $\mathbf{e}_k = (1, 1, \dots, 1)' \in \mathbb{R}^{k \times 1}$, $T^0 = -T\mathbf{e}_m$, and $\phi_{k,p}$ is given by Equation (9). Moreover, an error bound of the approximation error is given in Expression (32), where $M = \max\{\|S\|_2, \|T\|_2\}$ and $\|\cdot\|_F$ is the Frobenius norm.

Comparing Expressions (13) and (32), it is clear that this last one is more accurate, given that $\frac{M}{2^k} < M$ and $e^{Mx \left(1 - \frac{1}{2^k}\right)} < e^{Mk}$ where k is big. Moreover, this second error bound does not increase with time as fast as the previous one, Equation (13), and so the convergence is assured, a subject that is interesting in survival or reliability analysis for long-term predictions. This suggests that Equation (29) is better than Equation (11); however, Equation (29) requires the computation of more matrix exponentials. Thus, both approximations may be useful in applications, and we apply them to a numerical example of simulated data and real bladder cancer data in the next sections.

4. Numerical Application with Simulated Data

In this section, we compute the two approximations obtained above with simulated data. In order to clarify the two proposed approaches, first, we considered two homogeneous Markov processes and the corresponding concatenated process (Figure 1). We have considered the T and S transition matrices with $m = 10$ and $n = 11$ states, respectively, where each entry different from zero has been uniformly distributed in the range $[0.01, 0.05]$. We have chosen that range for the rates between transitional states by similarity with our real data; then, the absorbing rate is much greater (see expression $T^0 = -T\mathbf{e}_m$ in [16]). All simulated rates ($N = 100$) of the uniform distribution for each transition have been generated randomly with the Mathematica[®] software in the above range 90% of the simulations according to this uniform distribution led to badly conditioned matrices for the survival function $S(x)$, which corroborates the need to apply the two developed approximations, $\hat{S}_1(x)$ and $\hat{S}_2(x)$ of this study.

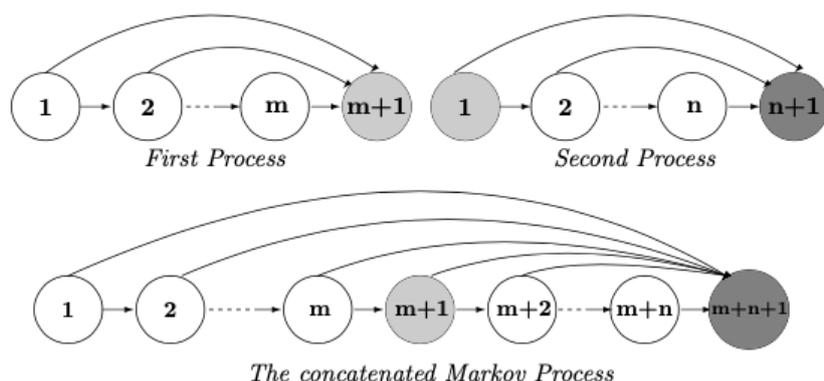


Figure 1. Two concatenated Markov processes. The resulting Markov chain with the final state $m + n + 1$ from any transitional state.

The first approximated survival function Equation (11) has been compared with the theoretical model Equation (1) previously developed in [16]. In Figure 2, we present the results of this fit between both functions, $S(x)$ and $\hat{S}_1(x)$, for $k = 5$ and $p = 7$ (9). We can observe a slight mismatch between both survival functions from 60 units of time onward, and in 80 units of time, the probability is over 60% and 50%. This is because the rates between the transitional states are small while these rates have more weight from each transitional state to the absorbing state.

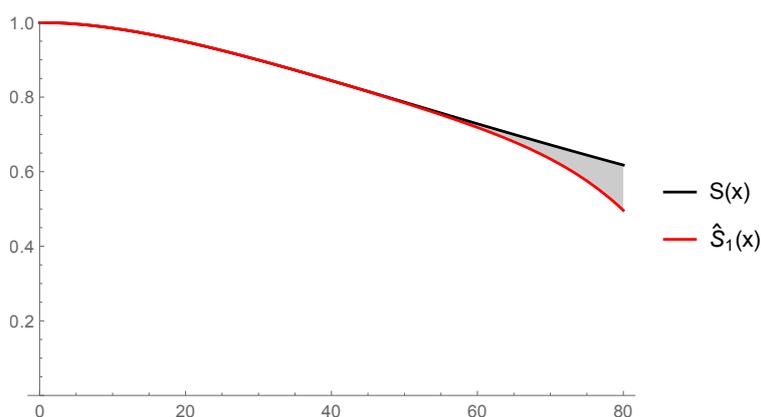


Figure 2. Survival function, $S(x)$, and the first approximation, $\hat{S}_1(x)$, in the concatenated Markov process for $p = 7$ with simulated data from a uniform distribution.

In Table 1, the error bound Inequality (13), calculated for the difference between $\hat{S}_1(x)$ and $S(x)$, increases from $x = 30$ until $x = 80$, reaching a large difference and leading to poor predictions in our calculations. In these cases, to get a lower bound of this difference, about 10^{-15} , for example, leads to an increase in the number of p terms in Equation (9), at least up to $p = 20$ in $x = 30$ units of time. Notice that the obtained error bound for this first approximation Equation (9) increases exponentially with time x , giving inaccurate predictions and a high error level.

In order to improve this first approximated distribution function, we proposed a second approximation to get a better convergence (Equation (29)). In Figure 3, we can check the accuracy between both survival functions, $S(x)$ and $\hat{S}_2(x)$, where absence of mismatches between them can be observed for high units of time. We can observe that the lines for both functions are superposed. In this case, the error bound Inequality (32) is low enough (about 10^{-14}) with only $p = 7$ and $k = 5$ as in Equation (9). The fit between both functions is almost perfect, even after one hundred units of time (error bound about 10^{-8}).

In order to evaluate both functions, $\hat{S}_1(x)$ and $\hat{S}_2(x)$, we can conclude that the second approach is more accurate than the first one compared to the exact model Equation (1) and with the same

computational cost. It can also be concluded that $\hat{S}_1(x)$ is always calculated with more terms (higher p) and $\hat{S}_2(x)$ with p' and k where $p' < p$ and with a error bound much smaller than with $\hat{S}_1(x)$. For this reason, we will only consider the second approximation in our application with the real bladder cancer data of the next section, given that we want to assure an error bound by 10^{-10} with only $k = 5$ and $p = 7$ terms.

Table 1. Error bound calculated for both approximations, $\hat{S}_1(x)$ and $\hat{S}_2(x)$, compared to the exact model $S(x)$ in the concatenated Markov process with simulated data from a uniform distribution for the transition rates.

Time Unit	Error Bound $\ S(x) - \hat{S}_1(x)\ _2$			Error Bound $\ S(x) - \hat{S}_2(x)\ _2$
	$p = 7$	$p = 10$	$p = 20$	$p = 7, k = 5$
$x = 30$	0.0143	42.07×10^{-6}	1.154×10^{-15}	1.2459×10^{-14}
$x = 50$	3.6811	0.05006	2.270×10^{-10}	3.1081×10^{-12}
$x = 80$	1054.3	58.7284	29.29×10^{-6}	8.5134×10^{-10}
$x = 100$	20343.3	2213.28	0.01028	1.5945×10^{-8}

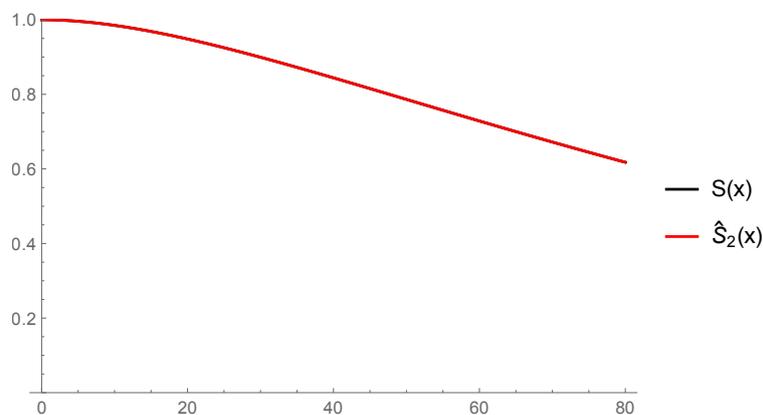


Figure 3. Survival function, $S(x)$, and the second approximation, $\hat{S}_2(x)$, in the concatenated Markov process for $k = 5$ and $p = 7$ with simulated data from a uniform distribution.

5. Illustration with a Real Data Set in Bladder Cancer

Bladder cancer can be classified into two well-differentiated types: non-muscle-invasive bladder cancer (NMIBC) and muscle-invasive tumor (MIT). Each type of tumor has a different follow-up protocol and treatment. Between 30–80% of the patients with a non-muscle-invasive primary tumor (the first tumor) have a recurrence of the disease (same type of tumor) and between 1–45% of these patients progress to muscle-invasive tumor (more aggressive). Patients with a muscle-invasive primary tumor may have a progression of the disease (with the possibility of the extirpation of the bladder) after some recurrences or directly. The first aim of the present work is to model both processes of the disease and, after the concatenated process, from the first NMBIC to the bladder extirpation (absorbing state $m + n + 1$).

For this, we have considered two Markov chains of an absorbing state, each one of them: (1) a first process, from a NMIBC primary tumor to a muscle-invasive tumor (first absorbing state), and (2) a second process starting in a muscle invasive primary tumor until the arrival to a progression with the final of the bladder (extirpation, the second absorbing state). In this context, we have worked with two continuous variables, X_1 and X_2 , where each one of them represented the two absorbing times mentioned above, each one of them PH-distributed. Two independent databases (with a different follow-up protocol and treatment) have been collected in this application, both from the Urology Department of at *La Fe University Hospital* in Valencia (Spain), each one with for a Markov process. Both databases cover between January 1995 and January 2010.

In the *first stage* of the disease, five transient states and one absorbing state were considered: the *primary non-muscle invasive tumor (NMIBC)*, a *first recurrence* and until a *fourth recurrence* (reappearance of a new *NMIBC* with similar characteristics to the first initial tumor). The absorbing state is the appearance of a progression to a *muscle invasive tumor (MIT)*, a much more aggressive tumor. The first database collected the information of 800 patients with mean follow-up period of 22.74 months.

The *second stage* showed three transient states and the absorbing state: three *muscle invasive tumors (MIT)* and the end of the *bladder* (extirpation). This database consisted of 160 patients with a *MIT* and 14.53 months of mean follow-up. The progression in a *MIT* is much faster than when the patient presents a *NMIBC*.

We have concatenated both independent processes to study the Survival function of a patient who had a primary non-muscle invasive tumor *NMIBC* (start of the illness) until the extirpation of its bladder (the worst episode in this disease) going through all the states (transient and absorbent ones) in each process (see Figure 1) with $m = 5$ and $n = 3$. For this, we have obtained the distribution function of the sum of the two variables (also *PH*-distributed) for each absorption state with the aim to obtain the survival function, $S(x)$, and the two approaches, $\hat{S}_1(x)$ and $\hat{S}_2(x)$ developed in this work.

The squared matrices T and S from each real database, referred to the rates between the transient states in each process, were estimated using the maximum likelihood method by the `msm()` function in the multistate modeling with R software and the `msm` package [19]. This function `msm` models transition rates with hidden Markov chains and data with observations with censored states. The dimensions of both matrices are five and three, respectively, corresponding to the number of transient states considered in the real data of each process as it has been mentioned above. The transient rates for T and S resulted in a range of $[0.001, 0.0001]$, smaller quantities than in the simulated data of the previous section.

We have also calculated the mean absolute percentage error (MAPE) for both approximations, $\hat{S}_1(x)$ and $\hat{S}_2(x)$, compared to the exact function $S(x)$: 0.0200759 and 1.1701×10^{-13} , respectively. The value of the MAPE for $\hat{S}_1(x)$ and $\hat{S}_2(x)$ shows the obvious mismatch between both functions mentioned in simulated data. We have also calculated the difference between both functions with $S(x)$ to conclude the same result that with simulated data: more precision with the second approach (Table 2). Finally, we have represented in the Figure 4 the exact function $S(x)$ and the second approximation, $\hat{S}_2(x)$, where we can see the good match between these two functions developed for the survival function of the concatenated process.

Table 2. Error bound calculated for both approximations, $\hat{S}_1(x)$ and $\hat{S}_2(x)$, compared to the exact model $S(x)$ in the concatenated Markov process with simulated real bladder cancer data.

time	Error Bound $\ S(x) - \hat{S}_1(x)\ _2$			Error Bound $\ S(x) - \hat{S}_2(x)\ _2$
	$p = 7$	$p = 10$	$p = 20$	$p = 7, k = 5$
$x = 30$	1.917×10^{-22}	9.252×10^{-29}	6.160×10^{-52}	1.739×10^{-34}
$x = 50$	2.004×10^{-20}	4.477×10^{-26}	4.931×10^{-47}	1.815×10^{-32}
$x = 80$	1.489×10^{-18}	1.362×10^{-23}	1.650×10^{-42}	1.345×10^{-30}
$x = 100$	1.169×10^{-17}	2.089×10^{-22}	2.356×10^{-40}	1.054×10^{-29}

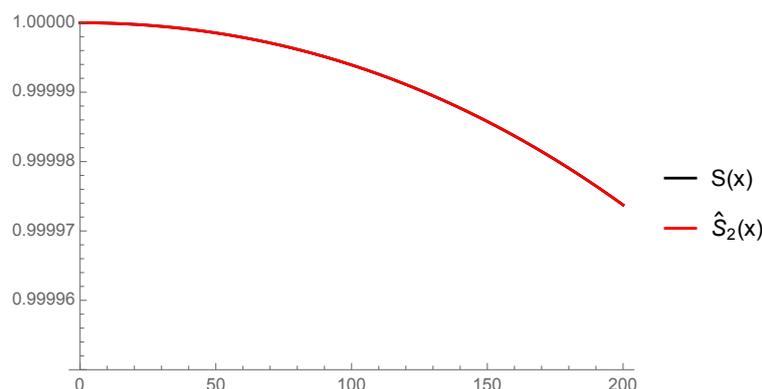


Figure 4. Survival function, $S(x)$, and the second approximation, $\hat{S}_2(x)$, for $k=5$ and $p=7$ of the concatenated Markov process for real bladder cancer data.

6. Concluding Remarks

In this study, we developed two approaches to a distribution function, proposed in a previous paper, to avoid the calculation of the inverse of a matrix (due to the possibility of a badly conditioned matrix) in the expression of that distribution function. Of the two developed approaches, we find the second one is much more accurate for performing predictions, corroborated with the calculation of the error bound for both approaches.

Regarding the computations of both approaches, small values of p and k terms in the expressions were used to get the desired accuracy. There is no problem in increasing the number of terms p and the parameter k to improve the precision with the second approximation. The mathematical expressions and the calculations were presented in a closed form that allowed algebraic treatment and the corresponding computational implementation. Moreover, this is easily interpretable and has a relatively low computational cost. Calculations were performed with the Mathematica[®] software, and all codes are available from the authors on request.

The aim of this work arose from the need to develop an approximation to the survival function for the disease when a database from the start of the illness to the bladder extirpation is not available, but two disconnected bases are available. The two real databases in this work and the previous paper were from different units at the La Fe University Hospital (hence independent of each other), and we were interested in examining the process until bladder extirpation from the beginning of the disease (the primary NMIBC). However, the presented approach is general, and this analysis can be applied to other similar data in chronic diseases or in a reliability context.

Summarizing, the results obtained in this paper with their applications allow to compute the survival or reliability function when the matrix in Equation (2.2) of the resulting process is poorly conditioned.

Author Contributions: Design and development of the Mathematical model, B.G.-M. and C.S.; Formal analysis, B.G.-M. and C.S.; computational framework, B.G.-M. and C.S.; implementation of the simulations, B.G.-M., C.S. and G.R.; performance of the calculations, B.G.-M., C.S. and G.R.; application to the bladder cancer, B.G.-M., C.S. and G.R.; writing—original draft, B.G.-M., C.S. and G.R.; writing—review and editing, B.G.-M., C.S. and G.R. All authors have read and agreed to the published version of the manuscript.

Funding: This paper has been supported by the Generalitat Valenciana grant AICO/2020/114.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Kronecker Sum and Kronecker Product

We are going to describe the most relevant properties to operate the PH distributions.

Definition A1. Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$ be matrices; then, the Kronecker product $A \otimes B$ is the $mp \times nq$ block matrix $(a_{ij}B)$ with a_{ij} the ij th element in the matrix A .

Definition A2. Let $A \in \mathbb{R}^{m \times m}$ and $B \in \mathbb{R}^{n \times n}$ be matrices and I_k denotes the identity matrix of order k ; then, the Kronecker sum is given by the following expression $A \oplus B = A \otimes I_n + I_m \otimes B$.

These properties are used through the paper. Let $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{p \times q}$, $C \in \mathbb{R}^{n \times o}$, and $D \in \mathbb{R}^{q \times r}$ be matrices; then,

1. $\text{vec}(ADC) = (C' \otimes A) \text{vec}(D)$ where C' denotes the transpose of the C matrix. The function $\text{vec}(\cdot)$ stacks the columns of a matrix into a column vector.
2. $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$
3. $e^A \otimes e^B = e^{A \oplus B}$
4. $\sigma(A \otimes B) = \{\mu\rho \mid \mu \in \sigma(A), \rho \in \sigma(B)\}$ where $\sigma(A)$ and $\sigma(B)$ are the spectrums of the A and B matrices respectively.
5. $f(A \otimes I) = f(A) \otimes I$; $f(I \otimes B) = I \otimes f(B)$ where f is an analytic function.
6. $\sigma(A \oplus B) = \{\mu + \rho \mid \mu \in \sigma(A), \rho \in \sigma(B)\}$

For more details of these properties, see [20,21].

As $S' \otimes I_n$ and $I_m \otimes T$ commute, $(S' \otimes I_n)(I_m \otimes T) = S' \otimes T = (I_m \otimes T)(S' \otimes I_n)$, it can be developed the following expression

$$\begin{aligned} (S's \oplus (1-s)T)^2 &= (S's \otimes I_n + I_m \otimes (1-s)T)^2 \\ &= (S')^2 s^2 \otimes I_n + I_m \otimes (1-s)^2 T^2 + 2(S's \otimes (1-s)T) \\ &= \sum_{j=0}^2 \binom{2}{j} (S')^{2-j} \otimes T^j s^{2-j} (1-s)^j \end{aligned} \tag{A1}$$

Appendix B. Proof of the Expression (16)

Definition A3 (The Fréchet derivative). The Fréchet derivative $L_F(Z, A)$ of F at A in the matrix direction Z is the limit of the Newton quotient for F

$$L_F(Z, A) \equiv \lim_{\delta \rightarrow 0} \frac{F(A + \delta Z) - F(A)}{\delta} \tag{A2}$$

If $F(x) = x^2$, the Newton quotient for the squaring function is

$$L_F(Z, X) \equiv \lim_{\delta \rightarrow 0} \frac{F(X + \delta Z) - X^2}{\delta} = \lim_{\delta \rightarrow 0} (XZ + ZX + \delta Z^2) = XZ + ZX \tag{A3}$$

This result is reviewed in the following definition

Definition A4. The Fréchet derivative $L_F(Z, A)$ of F at A in the matrix direction Z is the limit of the Newton quotient for F

$$L_F(Z, A) \equiv \lim_{\delta \rightarrow 0} \frac{F(A + \delta Z) - F(A)}{\delta} \tag{A4}$$

Definition A5. The standard integral formula of the Fréchet derivative of the exponential map at X in the direction L is

$$L_F(L, X) = \int_0^1 e^{(1-s)X} L e^{sX} ds \tag{A5}$$

Definition A6. The Fréchet derivative of the composition of functions is the composition of the Fréchet derivatives

$$L_{(G \circ F)(x)}(N, X) = L_{G(F(x))}(N, X) = L_G[L_F(N, X), F(x)] \tag{A6}$$

This last result is used in the composition of functions $X \rightarrow \frac{X}{2^k}$ followed by exponentiation and then squaring $e^X = (e^{\frac{X}{2^k}})^{2^k}$

Firstly, the standard integral formula of the Fréchet derivative of the exponential map at X in the direction N is applied to the function $e^{\frac{X}{2^{k-1}}}$ (A5)

$$L_{e^{\frac{X}{2^{k-1}}}}(N, X) = L_{e^X} \left(\frac{N}{2^{k-1}}, \frac{X}{2^{k-1}} \right) = \int_0^1 e^{(1-s)\frac{X}{2^{k-1}}} \frac{N}{2^{k-1}} e^{s\frac{X}{2^{k-1}}} ds \tag{A7}$$

However, as $L_{e^{\frac{X}{2^{k-1}}}}(N, X) = L_{e^{\left(\frac{X}{2^k}\right)^2}}$, the composition of functions (A6) is applied to the last expression

$$L_{e^{\left(\frac{X}{2^k}\right)^2}} = L_2 \left(L_{e^{\frac{X}{2^k}}}(N, X), e^{\frac{X}{2^k}} \right) \tag{A8}$$

By means of the definitions (A3) and (A5), the expression (A8) follows as

$$\begin{aligned} L_2 \left(L_{e^{\frac{X}{2^k}}}(N, X), e^{\frac{X}{2^k}} \right) &= e^{\frac{X}{2^k}} L_{e^{\frac{X}{2^k}}}(N, X) + L_{e^{\frac{X}{2^k}}}(N, X) e^{\frac{X}{2^k}} \\ &= L_{e^X} \left(\frac{N}{2^k}, \frac{X}{2^k} \right) e^{\frac{X}{2^k}} + e^{\frac{X}{2^k}} L_{e^X} \left(\frac{N}{2^k}, \frac{X}{2^k} \right) \\ &= \int_0^1 e^{(1-s)\frac{X}{2^k}} \frac{N}{2^k} e^{s\frac{X}{2^k}} ds e^{\frac{X}{2^k}} + e^{\frac{X}{2^k}} \int_0^1 e^{(1-s)\frac{X}{2^k}} \frac{N}{2^k} e^{s\frac{X}{2^k}} ds \end{aligned} \tag{A9}$$

Then, the expressions (A7) and (A9) are equal

$$\int_0^1 e^{(1-s)\frac{X}{2^{k-1}}} \frac{N}{2^{k-1}} e^{s\frac{X}{2^{k-1}}} ds = \int_0^1 e^{(1-s)\frac{X}{2^k}} \frac{N}{2^k} e^{s\frac{X}{2^k}} ds e^{\frac{X}{2^k}} + e^{\frac{X}{2^k}} \int_0^1 e^{(1-s)\frac{X}{2^k}} \frac{N}{2^k} e^{s\frac{X}{2^k}} ds \tag{A10}$$

Now, the result (A10) is applied to the following upper triangular matrices by blocks

$$X = \begin{pmatrix} Tx & 0 \\ 0 & Sx \end{pmatrix} \quad N = \begin{pmatrix} 0 & T^0\beta x \\ 0 & 0 \end{pmatrix} \tag{A11}$$

where T and S are nonsingular matrices of order m and n , respectively. T_0 is a $m \times 1$ matrix [16].

$$L_{e^{\frac{X}{2^{k-1}}}} \left[\begin{pmatrix} 0 & T^0\beta x \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} Tx & 0 \\ 0 & Sx \end{pmatrix} \right] = L_{e^{\frac{X}{2^{k-1}}}} \left[\begin{pmatrix} 0 & T^0\beta x \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} Tx & 0 \\ 0 & Sx \end{pmatrix} \right] e^{\frac{X}{2^k}} + e^{\frac{X}{2^k}} L_{e^{\frac{X}{2^{k-1}}}} \left[\begin{pmatrix} 0 & T^0\beta x \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} Tx & 0 \\ 0 & Sx \end{pmatrix} \right] \tag{A12}$$

and after operating the standard integral formula of the Fréchet derivative of the exponential map at X in the direction N (A7), the expression (A12) will result

$$\int_0^1 e^{(1-s)\frac{Tx}{2^j}} \frac{T^0\beta x}{2^j} e^{s\frac{Sx}{2^j}} ds e^{\frac{Sx}{2^j}} + e^{\frac{Tx}{2^j}} \int_0^1 e^{(1-s)\frac{Tx}{2^j}} \frac{T^0\beta x}{2^j} e^{s\frac{Sx}{2^j}} ds \tag{A13}$$

References

1. Neuts, M.F. *Matrix Geometric Solutions in Stochastic Models. An Algorithmic Approach*; Johns Hopkins University Press: Baltimore, MA, USA, 1981.
2. O’Cinneide, C. Characterization of phase-type distributions. *Commun. Stat. Stoch. Model.* **1990**, *6*, 1–57. [CrossRef]
3. Buchholz, P.; Kriege, J.; Felko, I. *Input Modeling with Phase-Type Distributions and Markov Models. Theory and Applications*; Springer: Cham, Switzerland, 2014.
4. Lee, S.; Bain, P.; Musa, A. A Markov chain model for analysis of physician workflow in primary care clinics. *Health Care Manag. Sci.* **2020**. [CrossRef] [PubMed]
5. Asmussen, S. *Applied Probability and Queues*; Springer: New York, NY, USA, 2003.

6. Rodríguez, J.; Lillo, R.E.; Ramírez-Cobo, P. Failure modeling of an electrical N -component framework by the non-stationary Markovian arrival process. *Reliab. Eng. Syst. Saf.* **2015**, *134*, 126–133. [[CrossRef](#)]
7. García-Mora, B.; Santamaría, C.; Rubio, G. Markovian modeling for dependent interrecurrence times in bladder cancer. *Math. Methods Appl. Sci.* **2020**, *43*, 8302–8310. [[CrossRef](#)]
8. Montoro-Cazorla, D.; Pérez-Ocón, R. Matrix stochastic analysis of the maintainability of a machine under shocks. *Reliab. Eng. Syst. Saf.* **2014**, *121*, 11–17. [[CrossRef](#)]
9. Aalen, O.O. On phase type distributions in survival analysis. *Scand. J. Stat.* **1995**, *22*, 447–463.
10. Fackrell, M. Modelling healthcare systems with phase-type distributions. *Health Care Manag. Sci.* **2009**, *12*, 11–26. [[CrossRef](#)] [[PubMed](#)]
11. Garg, L.; McClean, S.; Meenan, B.; Millard, P. Phase-Type Survival Trees and Mixed Distribution Survival Trees for Clustering Patients' Hospital Length of Stay. *Informatica* **2011**, *22*, 57–72. [[CrossRef](#)]
12. Marshall, A.H.; McClean, S.I. Conditional phase-type Distributions for modelling patient length of stay in hospital. *Int. Trans. Oper. Res.* **2003**, *10*, 565–576. [[CrossRef](#)]
13. Marshall, A.H.; McClean, S.I. Using Coxian Phase-Type Distributions to Identify Patient Characteristics for Duration of Stay in Hospital. *Health Care Manag. Sci.* **2004**, *7*, 285–289. [[CrossRef](#)] [[PubMed](#)]
14. Fackrell, M. Characterization of Matrix-Exponential Distributions. Ph.D. Thesis, School of Applied Mathematics, University of Adelaide, Adelaide, South Australia, 2003.
15. Fackrell, M. A semi-infinite programming approach to identifying matrix-exponential distributions. *Int. J. Syst. Sci.* **2012**, *9*, 1623–1631. [[CrossRef](#)]
16. García-Mora, B.; Santamaría, C.; Rubio, G.; Pontones, J.L. Computing survival functions of the sum of two independent Markov processes: an application to bladder carcinoma treatment. *Int. J. Comput. Math.* **2014**, *91*, 209–220. [[CrossRef](#)]
17. Kenney, C.; Laub, A.J. Condition Estimates for Matrix Functions. *SIAM J. Matrix Anal. Appl.* **1989**, *10*, 191–209. [[CrossRef](#)]
18. Kandolf, P.; Koskela, A.; Relton, S.; Schweitzer, M. Computing low-rank approximations of the Fréchet derivative of a matrix function using Krylov subspace methods. *arXiv* **2008**, arXiv:2008.12926.
19. Jackson, C.H. Multi-State Models for Panel Data: The msm Package for R. *J. Stat. Softw.* **2011**, *38*, 1–29. [[CrossRef](#)]
20. Graham, A. *Kronecker Products and Matrix Calculus with Applications*; Ellis Horwood Series in Mathematics and Its Applications; Halsted Press: Canberra, Australia, 1981.
21. L, L.M.; Reynolds, J. Scalable, Portable, Verifiable Kronecker Products on Multi-scale Computers. In *Constraint Programming and Decision Making. Studies in Computational Intelligence*; Ceberio, M., Kreinovich, V., Eds.; Springer: Cham, Switzerland, 2014; Volume 539. [[CrossRef](#)]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).