Parameter estimation for fractional birth and fractional death processes

Dexter O. Cahoy · Federico Polito

Received: 5 August 2011 / Accepted: 26 October 2012 / Published online: 13 November 2012
© Springer Science+Business Media New York 2012

Abstract The fractional birth and the fractional death processes are more desirable in practice than their classical counterparts as they naturally provide greater flexibility in modeling growing and decreasing systems. In this paper, we propose formal parameter estimation procedures for the fractional Yule, the fractional linear death, and the fractional sublinear death processes. The methods use all available data possible, are computationally simple and asymptotically unbiased. The procedures exploited the natural structure of the random inter-birth and inter-death times that are known to be independent but are not identically distributed. We also showed how these methods can be applied to certain models with more general birth and death rates. The computational tests showed favorable results for our proposed methods even with relatively small sample sizes. The proposed methods are also illustrated using the branching times of the plethodontid salamanders data of (Syst. Zool. 28:579–599, 1979).

Keywords Birth process · Yule process · Yule–Furry process · Death process · Mittag–Leffler

1 Introduction

Recently, generalizations of the classical birth and death processes have been developed using the techniques of fractional calculus. These are called the fractional birth (Uchaikin et al. 2008; Orsingher and Polito 2010; Cahoy and Polito 2012) and the fractional death (Orsingher et al. 2010) processes, correspondingly. A major advantage of these models over their classical counterparts is that they can capture both Markovian and non-Markovian structures of a growing or decreasing system.

When the birth and death rates are both linear, they are then called the fractional linear birth or fractional Yule or Yule–Furry process (fYp) and fractional linear death process, respectively. The classical linear birth or Yule process has been widely used to model various stochastic systems such as cosmic showers in physics and epidemics in biology to name a few (see e.g., Nee et al. 1994; Aldous 2001; Nee 2001; Paradis 2012). Note also that the fractional linear birth process was partially investigated by Uchaikin et al. (2008) using the Riemann-Liouville derivative operator but was continued and generalized by Orsingher and Polito (2010) using the Caputo derivative. The inter-birth time distribution, which provided a way to simulate the fYp was derived in Cahoy and Polito (2012). With this, we adopt the fYp from Orsingher and Polito (2010). In addition, the definition of the fractional linear and fractional sublinear death processes are taken from Orsingher et al. (2010).

For completeness, we first enumerate some properties of the fractional Yule (with one progenitor) and the fractional linear death (with initial population size \( n_0 > 1 \)) processes, which will be used in the subsequent discussions. Table 1 below shows the probability \( \tilde{P}(t) \) of no event (no birth or no death) at time \( t \), the state probability mass function \( P_i(t) \) or the probability of having \( (i - 1) \) births or \( (n_0 - i) \) deaths by time \( t \), the probability density function \( f_i(t) \) of the independent but non-identically distributed random inter-event times, the mean, and the variance of the fractional Yule and the fractional linear death processes. Note that the fractional
is the Mittag–Leffler function.\footnote{Note: The entries with \((**\)} are new results and are derived in Sect. 2.

Yule and the fractional linear death processes have the parameters \(\lambda > 0\) and \(\mu > 0\) as the birth and death intensities, correspondingly.

Recall that

\[
E_{\delta, \beta}(x) = \sum_{j=0}^{\infty} \frac{x^j}{\Gamma(\delta j + \beta)}
\]

is the Mittag–Leffler function.\footnote{Note: The entries with \((**\)}

In this article, we propose regression-based procedures to estimate the parameters of the fractional linear birth, the fractional linear death, and the fractional sublinear death processes. The rest of the paper is organized as follows. In Sect. 2, the specific functional forms of the inter-death time distributions and the variances of the fractional linear and sublinear death processes are obtained. These results allowed us to apply our methods to these processes. Section 3 introduces the proposed method using the fractional Yule, the fractional linear death, and the fractional sublinear death processes as examples. The section also shows some extensions of the procedures to certain models. Section 4 contains the empirical test results and the real-data application of the proposed methods for the case of the fYp only as similar inference procedures can be applied to the fractional linear and fractional sublinear death processes. The summary and extensions of our study are given in Sect. 5.

We now derive some properties which will permit us to apply the proposed estimation procedures to the fractional linear death and the fractional sublinear death processes. More specifically, the theorems below showed that the inter-death times for both the fractional linear and sublinear death processes are Mittag-Leffler distributed. The variances of both processes are also derived.

\[\text{Theorem 2.1} \quad \text{The inter-death time } T^v_k \text{ of the fractional linear death process } \{M^v(t), t \geq 0\} \text{ with death rate intensity } \mu > 0, \text{ and } n_0 \in \mathbb{N} \text{ initial individuals are independent but are non-identically distributed with probability density function}
\]

\[\Pr\{T^v_k \in dt\}/dt = \mu(n_0 - k)t^{v-1}E_{v,v}(-\mu(n_0 - k)t^{v}).\]

where \(k = 0, 1, \ldots, n_0 - 1\), and \(T^v_k\) is the random time separating the \(k\)th and \((k + 1)\)th death.

\[\text{Proof}\] We prove the theorem by induction. When \(k = 0\) we obtain

\[\Pr\{T^v_0 \leq t\} = \Pr\{M^v(t) < n_0\} = 1 - \Pr\{M^v(t) = n_0\} = 1 - \tilde{P}(t) \quad \text{(see Table 1)}
\]

\[= 1 - E_{v,1}(-\mu n_0 t^{v}).\]  \hspace{1cm} (2.1)

Therefore

\[\Pr\{T^v_0 \in dt\}/dt = \frac{d}{dt} \Pr\{T^v_0 \leq t\} = \mu n_0 t^{v-1}E_{v,v}(-\mu n_0 t^{v}).\]  \hspace{1cm} (2.2)

For \(k = 1\) we observe

\[\Pr\{T^v_0 + T^v_1 \in dt\}/dt = \frac{d}{dt} \Pr\{T^v_0 + T^v_1 < t\}
\]

\[= \frac{d}{dt} \Pr\{M^v(t) < n_0 - 1\}
\]

\[= \frac{d}{dt} \left[1 - \Pr\{M^v(t) = n_0\} - \Pr\{M^v(t) = n_0 - 1\}\right].\]  \hspace{1cm} (2.3)

Using Table 1 we get

\[\Pr\{T^v_0 + T^v_1 \in dt\}/dt = -\frac{d}{dt} E_{v,1}(-\mu n_0 t^{v}).\]
By exploiting again the Laplace transform and writing

\[ -\frac{d}{dt}[n_0 E_v,1(\mu t^\nu) - n_0 E_v,1(\mu t^\nu)] \]

\[ = \mu n_0 t^{\nu-1} E_v,v(-\mu t^\nu) \]

\[ + n_0(n_0 - 1)\mu t^{\nu-1} E_v,v(-\mu(n_0 - 1)t^\nu) \]

\[ - n_0^2 \mu t^{\nu-1} E_v,v(-\mu n_0 t^\nu) \]

\[ = n_0(n_0 - 1)\mu t^{\nu-1}[E_v,v(-\mu(n_0 - 1)t^\nu) \]

\[ - E_v,v(-\mu n_0 t^\nu)]. \]  \hfill (2.4)

To check the preceding results, we can obtain the Laplace transform as

\[ \int_0^\infty e^{-wt} \Pr\{T_0^\nu + T_1^\nu \in dt\} \]

\[ = \int_0^\infty e^{-ws} \Pr\{T_0^\nu \in ds\} \int_0^\infty e^{-wt} \Pr\{T_1^\nu \in ds\} \]

\[ = \int_0^\infty e^{-ws} \Pr\{T_0^\nu \in ds\} \int_0^\infty e^{-ut} \Pr\{T_1^\nu \in ds\} \]

\[ = \int_0^\infty \Pr\{T_0^\nu \in ds\} \int_s^\infty e^{-zt} \Pr\{T_1^\nu \in ds\}, \] \hfill (2.5)

which is just a convolution of two independent variables \( T_0^\nu \) and \( T_1^\nu \). For a general \( k \) it is sufficient to note that

\[ \Pr\{T_0^\nu + \cdots + T_k^\nu \in dt\} \]

\[ = \int_0^t \Pr\{T_k^\nu \in d(t - s)\} \Pr\{T_0^\nu + \cdots + T_{k-1}^\nu \in ds\}. \]  \hfill (2.6)

By exploiting again the Laplace transform and writing \( D_k^\nu = T_k^\nu + \cdots + T_k^\nu \), we have

\[ \int_0^\infty e^{-ws} \Pr\{D_k^\nu \in dt\} \]

\[ = \int_0^\infty e^{-ws} \int_0^t \Pr\{T_k^\nu \in d(t - s)\} \Pr\{D_{k-1}^\nu \in ds\} \]

\[ = \int_0^\infty \Pr\{D_{k-1}^\nu \in ds\} \int_s^\infty e^{-zt} \Pr\{T_k^\nu \in d(t - s)\} \]

\[ = \int_0^\infty e^{-ws} \Pr\{D_{k-1}^\nu \in ds\} \int_0^\infty e^{-ut} \Pr\{T_k^\nu \in dy\} \]

\[ = \prod_{j=0}^k \int_0^\infty e^{-ws} \Pr\{T_j^\nu \in ds\} \]

\[ = \prod_{j=0}^k \frac{\mu(n_0 - j)}{w^\nu + \mu(n_0 - j)}. \] \hfill (2.7)

We now determine the variance of the fractional linear death process \( \{M(t), t > 0\} \). Consider Eq. (1.6) of Orsingher et al. (2010). That is,

\[ \frac{d^\nu}{dt^\nu} p_k^\nu(t) = \mu(k + 1)p_{k+1}^\nu(t) - \mu k p_k^\nu(t), \quad 0 \leq k \leq n_0, \]

\[ p_k^\nu(0) = \begin{cases} 1, & k = n_0, \\ 0, & 0 \leq k < n_0. \end{cases} \] \hfill (2.8)

It is then straightforward to arrive at

\[ \int_0^\infty \frac{d^\nu}{dt^\nu} G^\nu(u, t) = -\mu(u - 1) \frac{d}{du} G^\nu(u, t), \]

\[ G^\nu(u, 0) = u^{n_0}, \] \hfill (2.9)

where \( G^\nu(u, t) = \sum_{k=0}^{n_0} u^k p_k^\nu(t) \) is the probability generating function of the fractional linear death process. This in turn leads to

\[ \int_0^\infty \frac{d^\nu}{dt^\nu} H(t) = -2\mu H(t), \]

\[ H(0) = n_0(n_0 - 1), \] \hfill (2.10)

where \( H(t) = \mathbb{E}(M^\nu(t)(M^\nu(t) - 1)) \) is the second factorial moment. The solution to (2.10) reads

\[ H(t) = n_0(n_0 - 1)e^{-2\mu t}, \] \hfill (2.11)

and the variance can be immediately obtained as

\[ \text{Var } M^\nu(t) = H(t) + \mathbb{E} M^\nu(t) - (\mathbb{E} M^\nu(t))^2 \]

\[ = n_0(n_0 - 1)E_{\nu,1}(-2\mu t^\nu) \]

\[ + n_0 E_{\nu,1}(-\mu t^\nu) - n_0^2 (E_{\nu,1}(-\mu t^\nu))^2. \] \hfill (2.12)

Note that the above expression, when \( \nu = 1 \), simplifies to the variance of the classical linear death process, i.e.

\[ \text{Var } M^1(t) = n_0 e^{-\mu t}(1 - e^{-\mu t}). \] \hfill (2.13)

Below is the algorithm to generate a typical sample path of a fractional linear death process in Fig. 1. Note that there are several sub-algorithms to generate the inter-death times \( T_j^\nu \)'s that are available in the literature (see e.g., Cahoy and Polito 2012).

**Algorithm**

Step 1. Let \( k = 0 \) and the population size equal \( n_0 \).

Step 2. Simulate \( T_k^\nu \), and let the \( k \)th death time be \( D_k^\nu = T_k^\nu + T_{k+1}^\nu + T_{k+2}^\nu + \cdots + T_k^\nu \).

Step 3. Set the population size \( n_0 - k \), and \( k = k + 1 \).

Step 4. Repeat Steps 2–3 for \( k = 1, \ldots, n_0 - 1 \).

It can be gleaned from Fig. 1 that the sample path of the fractional linear death process (bottom) seems to decay faster at small times but is slower for large times than its
Recall that the fractional sublinear death process can be determined by considering Eq. (3.45) of Orsingher et al. Theorem 2.2 preceding theorem as follows. The fractional linear death process can be easily deduced (whose proof corresponding to small times). The inter-death time distribution for the fractional sublinear death process is capable of producing death bursts especially at early stages (corresponding to small times).

The inter-death time distribution for the fractional sublinear death process can be easily deduced (whose proof follows from the previous result and is omitted) from the preceding theorem as follows.

**Theorem 2.2** The fractional sublinear death process \( \{\mathcal{M}^\nu(t), t > 0\} \), with death intensity rate \( \mu > 0 \), and \( n_0 \in \mathbb{N} \) initial individuals has the following probability density function of the inter-death times \( \tau^\nu_k \),

\[
Pr\{\tau^\nu_k \in dt\} = \mu(k + 1)\nu^{-1}E_{\nu,\nu}(-\mu(k + 1)t^\nu),
\]

with \( k = 0, 1, \ldots, n_0 - 1 \), where \( \tau^\nu_k \) is the random time separating the \( k \)th and \((k + 1)\)th death.

The variance of the fractional sublinear death process can be determined by considering Eq. (3.45) of Orsingher et al. (2010). Recall that

\[
\frac{\partial^2}{\partial u^2} \Phi^\nu(u, t) \bigg|_{u=1} = E[\mathcal{M}^\nu(t)(\mathcal{M}^\nu(t) - 1)] = H(t). \quad (2.14)
\]

Then

\[
\frac{d^\nu}{dt^\nu} H(t) = -2\mu(n_0 + 1)(E\mathcal{M}^\nu(t) + Pr\{\mathcal{M}^\nu(t) = 0\} - 1) + 2\mu H(t)
\]

\[
= -2\mu(n_0 + 1) \sum_{k=1}^{n_0} \binom{n_0}{k} (-1)^k E_{\nu,1}(-k\mu t^\nu)
+ \sum_{k=1}^{n_0} \binom{n_0 + 1}{k + 1} (-1)^{k+1} E_{\nu,1}(-\mu k t^\nu) + 2\mu H(t)
\]

\[
= -2\mu(n_0 + 1) \sum_{k=1}^{n_0} \left( \binom{n_0}{k} - \binom{n_0 + 1}{k + 1} \right)
\times (-1)^k E_{\nu,1}(-k\mu t^\nu) + 2\mu H(t)
\]

\[
= 2\mu(n_0 + 1) \sum_{k=1}^{n_0} \binom{n_0}{k + 1} (-1)^k E_{\nu,1}(-k\mu t^\nu)
+ 2\mu H(t). \quad (2.15)
\]

Using the initial condition \( H(0) = n_0(n_0 - 1) \) and letting \( \tilde{H}(w) \) be the Laplace transform of \( H(t) \), we write

\[
w^\nu \tilde{H}(w) - w^{\nu-1}n_0(n_0 - 1)
\]

\[
= 2\mu(n_0 + 1) \sum_{k=1}^{n_0} \binom{n_0}{k + 1}
\times (-1)^k \frac{w^{\nu-1}}{w^{\nu} + k\mu} + 2\mu \tilde{H}(w). \quad (2.16)
\]

Hence,

\[
\tilde{H}(w)
= n_0(n_0 - 1) \frac{w^{\nu-1}}{w^{\nu} - 2\mu} + 2\mu(n_0 + 1) \sum_{k=1}^{n_0} \binom{n_0}{k + 1}
\times (-1)^k \frac{w^{\nu-1}}{w^{\nu} + k\mu}(w^{\nu} - 2\mu)
\]

\[
= n_0(n_0 - 1) \frac{w^{\nu-1}}{w^{\nu} - 2\mu} + 2\mu(n_0 + 1)
\times \sum_{k=1}^{n_0} \binom{n_0}{k + 1} (-1)^k w^{\nu-1}
\times \left[ \frac{1}{w^{\nu} + k\mu} - \frac{1}{w^{\nu} - 2\mu} \right] \frac{1}{(w^{\nu} - 2\mu)}
\]

\[
= n_0(n_0 - 1) \frac{w^{\nu-1}}{w^{\nu} - 2\mu} + \frac{w^{\nu-1}}{w^{\nu} - 2\mu}(n_0 + 1)
\]
Tingly complex structure pose a computational challenge on
tributions. This observation and the Mittag-Leffler's seem-
=−H(t)
The second factorial moment can be easily shown as
Stat Comput (2014) 24:211–222 215
is obtained from each of the
ccess are given. This also insinuates that only a single datum
That is,
n
birth rate
3.1 Estimation for the fractional Yule or linear birth
process.
simulate sample trajectories of the fractional sublinear death
process.
Note that the algorithm above could be easily adopted to

3 Parameter estimation

3.1 Estimation for the fractional Yule or linear birth
process
We now illustrate our estimation approach for the fYp with
birth rate λi, i ≥ 1. Furthermore, assume that a sample tra-
jectory of n births corresponding to n random inter-birth
times Tj′s of the fractional linear birth process is observed.
That is, n independent but are not identically distributed
random inter-birth times of the fractional linear birth pro-
cess are given. This also insinuates that only a single datum
is obtained from each of the n different Mittag-Leffler

how to estimate the model parameters more efficiently espe-
cially for small population sizes. Recall the structural rep-
presentation of the Mittag-Leffler distributed random inter-
birth time Tj ≡ E1/νSν (see Cahoy and Polito 2012), where
E ≡ exp(λt) is independent of Sν which is a one-sided σ+-
able distributed random variable. Applying the logarithmic
transformation and taking the expectation on both sides, it
can be easily shown that the mean and variance (see details
in Cahoy et al. 2010) of the log-transformed i-th random
sojourn time T′i = ln(Ti) of the fYp are

\[
\mu_{T_i}' = \frac{-\ln(\lambda i)}{\nu} - \gamma, \tag{3.1}
\]
and

\[
\sigma_{T_i}'^2 = \pi^2 \left( \frac{1}{3 \nu^2} - \frac{1}{6} \right), \tag{3.2}
\]
respectively, where \( \gamma \approx 0.5772156649 \) is the Euler-
Mascheroni’s constant. The first two moments above therefore
suggest that the following simple linear regression model can
be fitted/formulated:

\[
T'_i = a_0 + a_1 \ln i + \varepsilon_i, \quad i = 1, \ldots, n, \tag{3.3}
\]
where

\[
a_0 = \frac{-\ln(\lambda_i)}{\nu} - \gamma, \quad a_1 = \frac{-1}{\nu}, \tag{3.4}
\]
and \( \varepsilon_i \overset{iid}{\sim} N(\mu_\varepsilon = 0, \sigma_\varepsilon^2 = \sigma_{T_i}'^2) \). The trick used here was to
factor out the non-identical means of the log-transformed
random inter-birth or sojourn times, which are linear func-
tions of the logarithm of the known fixed \( i \). Thus, this leads
to studying the widely used simple linear regression model
(see Montgomery et al. 2006).

3.1.1 Point estimation

Inverting the least squares (LS) estimators

\[
\hat{a}_1 = \frac{\sum_{j=1}^{n} T'_j (\ln j - \ln i)}{\sum_{j=1}^{n} (\ln j - \ln i)^2} \tag{3.5}
\]
and \( \hat{a}_0 = \bar{T}'_i - \hat{a}_1 \cdot \ln i \) gives the LS-based point estimators
of \( \nu \) and \( \lambda \) as

\[
\hat{\nu}_{LS} = \frac{1}{\hat{a}_1} \tag{3.6}
\]
and

\[
\hat{\lambda}_{LS} = \exp \left( (\hat{a}_0 + \gamma) / \hat{a}_1 \right). \tag{3.7}
\]
respectively, where
\[
\ln i = \sum_{j=1}^{n} \ln j/n, \quad \text{and} \quad T' = \sum_{j=1}^{n} T'_j/n.
\]
Equating \( \sigma_i^2 \) or \( \sigma_i^2/T'_i \) in (3.1) with its unbiased estimator
\[
\hat{\sigma}_u^2 = \sum_{j=1}^{n} \hat{\varepsilon}_j^2/(n-2),
\]
we get the residual-based point estimators
\[
\hat{\nu}_{res} = \frac{1}{\sqrt{3(\hat{\sigma}_u^2/\pi^2 + \frac{1}{6})}} \quad \text{(see Cahoy et al. 2010) (3.9)}
\]
and
\[
\hat{\lambda}_{res} = \exp(-\hat{\nu}_{res}(\hat{a}_0 + \gamma)) \quad \text{(3.10)}
\]
of the model parameters \( \nu \) and \( \lambda \), correspondingly where \( \hat{\nu}_i = T'_i - T'_i \) and \( T'_i = \hat{a}_0 + \hat{a}_1 \ln i \). Note that the residual-based estimators exploit the residuals to estimate \( \nu \) rather than the negative inverse of the LS estimate of the slope \( a_1 \).

### 3.1.2 Interval estimation

We now develop interval estimators using the large-sample properties of the least squares estimators \( \hat{a}_0 \) and \( \hat{a}_1 \) above. The following result shows the joint asymptotic behavior of the proposed point estimators of \( \nu \) and \( \lambda \) for the fYp.

**Theorem 3.1** Let \( 0 < \nu \leq 1 \) and \( \lambda > 0 \). Then
\[
\sqrt{n} \left( \frac{\hat{\nu}_{ls} - \nu}{\hat{\lambda}_{ls} - \lambda} \right) \xrightarrow{d} N[0, n\sigma_{\nu}^2 C]
\]
where “\( \xrightarrow{d} \)” denotes convergence in distribution,
\[
C = \begin{pmatrix}
C_1 & C_{12} \\
C_{21} & C_2
\end{pmatrix},
\]
\[
C_1 = \nu^4 s^{-1},
\]
\[
C_{12} = C_{21} = \lambda \nu^3 ((\ln i + \ln(\lambda))/s),
\]
\[
C_2 = (\nu \lambda)^2 (1/n + (\ln i^2 + 2 \ln(\lambda) / (\ln(\lambda))^2)/s),
\]
and \( s = \sum_{j=1}^{n} (\ln j - \ln i)^2 \).

**Proof** Recall the large-sample normality of the least squares estimators \( \hat{a}_0 \) and \( \hat{a}_1 \), i.e.,
\[
\sqrt{n} \left( \frac{\hat{a}_0 - a_0}{\hat{a}_1 - a_1} \right) \xrightarrow{d} N[0, \Sigma]
\]
where the covariance matrix \( \Sigma \) is defined as
\[
\Sigma = n\sigma_{\nu}^2 \begin{pmatrix}
(1/n + \ln i^2/s) & -\ln i/s \\
-\ln i/s & s^{-1}
\end{pmatrix}.
\]
Recall the multivariate delta method (Ferguson 1996): If
\[
\sqrt{n}(\hat{\beta}_n - \beta) \xrightarrow{d} N(0, \Sigma)
\]
then
\[
\sqrt{n}(g(\hat{\beta}_n) - g(\beta)) \xrightarrow{d} N(0, \hat{g}(\beta)\Sigma\hat{g}(\beta)^\top).
\]
Hence, using the delta method above, \( \hat{\beta}_n = (\hat{a}_0, \hat{a}_1)^\top \),
\[
\hat{g}(\beta) = \begin{pmatrix}
0 \\
1/a_1
\end{pmatrix} - \exp((a_0 + \gamma)/a_1)/a_1
\]
we obtain the final expression of the covariance matrix by simply substituting back \( a_0 = -\ln(\lambda)/\nu - \gamma \) and \( a_1 = -1/\nu \).

**Corollary 3.1** Approximate \( (1 - \alpha)100 \% \) confidence intervals for \( \nu \) and \( \lambda \) can be deduced as
\[
\hat{\nu}_{ls} \pm z_{\alpha/2} \hat{\sigma}_{\nu ls} \hat{\nu}_{ls} \sqrt{s^{-1}},
\]
and
\[
\hat{\lambda}_{ls} \pm z_{\alpha/2} \hat{\sigma}_{\lambda ls} \hat{\lambda}_{ls} (1/n + (\ln i^2 + 2 \ln(\hat{\lambda}_{ls}) / (\ln(\lambda))^2)/s)^{1/2},
\]
respectively, where \( z_{\alpha/2} \) is the \((1 - \alpha/2)\)th quantile of the standard normal distribution and \( 0 < \alpha < 1 \).

We now propose another interval estimators which utilize the residual-based estimate of \( \nu \), and a bootstrap technique. It can be inferred from Cahoy et al. (2010) that a residual-based \((1 - \alpha)100 \% \) confidence interval for \( \nu \) can be
\[
\hat{\nu}_{res} \pm z_{\alpha/2} \sqrt{\hat{\nu}_{res}^2 (32 - 20 \hat{\nu}_{res}^2 - \hat{\sigma}_{\nu res}^4)/40n},
\]
where \( z_{\alpha/2} \) is defined above. A residual-based \((1 - \alpha)100 \% \) interval estimate for \( \lambda \) can also be
\[
\hat{\lambda}_{res} \pm z_{\alpha/2} \left[ e^{-2\hat{\nu}_{res}(\hat{a}_0 + \gamma)} \left( \hat{\nu}_{res}^2 (32 - 20 \hat{\nu}_{res}^2 - \hat{\sigma}_{\nu res}^4)/40n \right) \right]^{1/2}.
\]
Since the small-sample performance of \( \hat{\nu}_{res} \) and the residual-based interval estimator in (3.13) have been shown to perform well already (see, e.g., Cahoy et al. 2010), we apply a non-parametric percentile bootstrap technique to \( \hat{\lambda}_{res} \) using...
the fixed-regressor approach to obtain a small-sample interval estimator of $\lambda$. This well-known procedure is slightly modified by first dividing each residual $\hat{e}$ by $\sqrt{1 - h_i}$, where $h_i$ is the $i$th leverage or the $i$th diagonal entry in the hat matrix before sampling from the transformed residuals. Note that the division of $\sqrt{1 - h_i}$ is simply for correction as the true variance of the residual $\hat{e}_i$ is $\text{Var}(\hat{e}_i) = \sigma^2(1 - h_i)$ (see Montgomery et al. 2006). Hence, the bootstrap counterpart of $\hat{\lambda}_{res}$ is calculated as $\hat{\lambda}_{res}^* = \exp(-\hat{\lambda}_{res}^*(\hat{G}^*_0 + \gamma))$ where $\hat{G}^*_0$ used the bootstrapped transformed or weighted residuals. A clear advantage of the asymptotic-based procedures over the re-sampling-based ones is that they are faster to calculate especially for large sample sizes.

3.2 Estimation for the fractional linear and the fractional sublinear death processes

Assuming that a sample trajectory of $n_0$ deaths corresponding to $n_0$ random inter-death times $T^{\nu}_k$’s of a fractional death process is observed. Following the procedure for the fractional linear birth process in the preceding subsection, we can estimate the parameters $\nu$ and $\mu$ by regressing $\ln(T^{\nu}_k)$ with $\ln(n_0 - k)$. That is, we fit the following simple linear regression model:

$$\ln(T^{\nu}_k) = b_0 + b_1 \ln(n_0 - k) + \epsilon_k,$$  

(3.15)

where $k = 0, \ldots, n_0 - 1$, $b_0 = -\ln(\mu)/\nu - \gamma$, $b_1$ is given in (3.4) of Sect. 3.1, and $\epsilon_k \overset{iid}{\sim} N(\mu, \sigma^2_{\ln(T^{\nu}_k)})$. Following the methodology in the preceding subsection, we can straightforwardly obtain the corresponding LS-based point estimates of $\nu$ and $\mu$ from (3.6) and (3.7) as

$$\hat{\nu}_{ls} = -\frac{1}{b_1}$$  

(3.16)

and

$$\hat{\mu}_{ls} = \exp(\hat{b}_0 + \gamma)/\hat{b}_1,$$  

(3.17)

respectively, where $\ln(n_0 - k) = \sum_{j=0}^{n_0-1} \ln(n_0 - j)$, $\ln(T^{\nu}_k) = \sum_{j=0}^{n_0-1} \ln(T^{\nu}_j)(\ln(n_0 - j) - \ln(n_0 - k))$, $\hat{b}_0 = \ln(T^{\nu}_k) - \hat{b}_1 \ln(n_0 - k)$, and

$$\hat{b}_1 = \frac{\sum_{j=0}^{n_0-1} \ln(T^{\nu}_j)(\ln(n_0 - j) - \ln(n_0 - k))}{\sum_{j=0}^{n_0-1} \ln(n_0 - j) - \ln(n_0 - k)^2}.$$  

(3.18)

Furthermore, the LS-based interval estimates for $\nu$ and $\mu$ directly follow from (3.11) and (3.12) of Corollary 3.1 in Sect. 3.1.2, correspondingly. Hence, the approximate $(1 - \alpha)100\%$ for $\nu$ and $\mu$ can be explicitly written as

$$\hat{\nu}_{ls} \pm z_{\alpha/2}^* \hat{\sigma}_{\nu\hat{\nu}_{ls}}^2 \left( \sum_{j=0}^{n_0-1} \frac{1}{\ln(n_0 - j) - \ln(n_0 - k)}^2 \right)^{-1},$$

and

$$\hat{\mu}_{ls} \pm z_{\alpha/2}^* \hat{\sigma}_{\mu\hat{\mu}_{ls}} \left( \sum_{j=0}^{n_0-1} (\ln(n_0 - j) - \ln(n_0 - k))^2 \right)^{-1} \left( \frac{1}{\ln(n_0 - k)} + \frac{2}{\ln(n_0 - k)^2} \right),$$

(3.19)

correspondingly. On the other hand, the residual-based point and interval estimators of $\nu$ and $\mu$ immediately follow from Sect. 3.1 as well, where $\hat{d}_0$ is replaced by $\hat{b}_0$ in (3.10), $\hat{\epsilon}_k = \ln(T^{\nu}_k) - \ln(T^{\nu}_0)$, and $\ln(T^{\nu}_0) = \hat{b}_0 + \hat{b}_1 \ln(n_0 - k)$.

A similar approach can be done to obtain estimates for the fractional sublinear death process. That is, we regress $\ln(T^{\nu}_k)$ with $\ln(k + 1), k = 0, 1, \ldots, n_0 - 1$, or fit the model

$$\ln(T^{\nu}_k) = c_0 + c_1 \ln(k + 1) + \epsilon_k,$$  

(3.20)

and follow the procedures used for fractional Yule and the fractional linear death processes. In general, we simply replace $\lambda$, $\ln(i)$, $\ln(t)$ by $\mu$, $\ln(n_0 - k)$ or $\ln(k + 1)$, and $\ln(n_0 - k)$ or $\ln(k + 1)$, accordingly in the methods of Sect. 3.1 to obtain the parameter estimators for the fractional linear death and the fractional sublinear death processes.

3.3 Some extensions

Assume that a fractional birth or death process exists with rates $\theta_j$, $j = 1, 2, \ldots, n$, where the $j$th inter-event time $X_j$ is Mittag-Leffler distributed with parameter $\theta_j$. Then the mean of $X'_j = \ln(X_j)$ is

$$\mu_{X'_j} = -\frac{\ln(\theta_j)}{\nu} - \gamma.$$  

(3.21)

Based on the above mean formulation, we use the model

$$X'_j = d_0 + d_1 \cdot q(j) + \epsilon_j$$  

(3.22)

to estimate more forms of the parameters or rates under the two cases below.

Case 1. When $\ln(\theta_j) = m(\theta) + q(j)$ for some appropriate known functions $m(\theta)$ and $q(j)$ of the parameter $\theta$ and $j \in \mathbb{N}$, correspondingly.

In this case, the general form of the regression model that could be used for estimation is

$$X'_j = -\left( \gamma + \frac{m(\theta)}{\nu} \right) - \frac{1}{\nu} \cdot q(j) + \epsilon_j.$$  

(3.23)

Clearly, $d_0 = -(\gamma + m(\theta)/\nu)$, $d_1 = -1/\nu$, and $q(j)$ is the regressor variable. Using $\hat{\nu}_{res}$ or $-1/\hat{d}_1$ and inverting the least squares estimate $\hat{b}_0$, we can compute $m(\theta)$ and $\hat{\theta}$ sequentially. Note that the explicitness of $\hat{\theta}$ depends on the form of $m$. 

© Springer
Table 2  Mean point estimates of and dispersions from the true parameters $\nu$ and $\lambda$

| $(\nu, \lambda)$ | Estimator | $n = 100$ | $n = 500$ | $n = 1000$ |
|------------------|-----------|-----------|-----------|------------|
|                  | Mean      | MAD       | Mean      | MAD        | Mean      | MAD        |
| $(0.1, 1)$       | $\hat{\nu}_{ls}$ | 0.104 | 0.020 | 0.101 | 0.008 | 0.100 | 0.006 |
|                  | $\hat{\nu}_{res}$ | 0.103 | 0.008 | 0.100 | 0.004 | 0.100 | 0.003 |
|                  | $\hat{\lambda}_{ls}$ | 3.190 | 0.665 | 1.151 | 0.407 | 1.077 | 0.318 |
|                  | $\hat{\lambda}_{res}$ | 1.318 | 0.725 | 1.091 | 0.408 | 1.051 | 0.322 |
| $(0.25, 0.1)$    | $\hat{\nu}_{ls}$ | 0.261 | 0.048 | 0.252 | 0.021 | 0.251 | 0.014 |
|                  | $\hat{\nu}_{res}$ | 0.252 | 0.022 | 0.251 | 0.010 | 0.250 | 0.007 |
|                  | $\hat{\lambda}_{ls}$ | 0.119 | 0.031 | 0.106 | 0.025 | 0.103 | 0.021 |
|                  | $\hat{\lambda}_{res}$ | 0.131 | 0.071 | 0.109 | 0.044 | 0.106 | 0.033 |
| $(0.5, 0.5)$     | $\hat{\nu}_{ls}$ | 0.521 | 0.100 | 0.505 | 0.040 | 0.501 | 0.028 |
|                  | $\hat{\nu}_{res}$ | 0.506 | 0.041 | 0.501 | 0.018 | 0.500 | 0.014 |
|                  | $\hat{\lambda}_{ls}$ | 0.825 | 0.280 | 0.565 | 0.190 | 0.532 | 0.142 |
|                  | $\hat{\lambda}_{res}$ | 0.640 | 0.350 | 0.555 | 0.216 | 0.528 | 0.164 |
| $(0.75, 0.25)$   | $\hat{\nu}_{ls}$ | 0.774 | 0.121 | 0.755 | 0.052 | 0.751 | 0.036 |
|                  | $\hat{\nu}_{res}$ | 0.755 | 0.056 | 0.752 | 0.023 | 0.750 | 0.016 |
|                  | $\hat{\lambda}_{ls}$ | 0.313 | 0.093 | 0.266 | 0.069 | 0.259 | 0.056 |
|                  | $\hat{\lambda}_{res}$ | 0.300 | 0.146 | 0.268 | 0.094 | 0.260 | 0.072 |
| $(0.95, 5)$      | $\hat{\nu}_{ls}$ | 0.969 | 0.131 | 0.953 | 0.058 | 0.952 | 0.042 |
|                  | $\hat{\nu}_{res}$ | 0.955 | 0.055 | 0.950 | 0.024 | 0.950 | 0.018 |
|                  | $\hat{\lambda}_{ls}$ | 11.251 | 3.492 | 5.836 | 2.104 | 5.397 | 1.654 |
|                  | $\hat{\lambda}_{res}$ | 5.978 | 2.544 | 5.375 | 1.635 | 5.206 | 1.317 |

Example 1.1 When $\theta_j$ is linear, i.e., $\theta_j = \theta$ then $\ln(\theta_j) = m(\theta) + \ln(j)$, where $m(\theta) = \ln(\theta)$ and $q(j) = \ln(j)$, respectively. Note that this parametrization corresponds to the fractional Yule, the fractional linear death, and the fractional sublinear death processes.

Example 1.2 If $\theta_j = e^{\theta + j}$ then $\ln(\theta_j) = \theta + j$, where $m(\theta) = \theta$ and $q(j) = j$, correspondingly. This suggests that $d_0 = -(\gamma + \theta / \nu)$.

Case 2 When $\ln(\theta_j) = m(\theta) \cdot q(j)$ for some appropriate known functions $m(\theta)$ and $q(j)$ of the parameter $\theta$ and $j \in \mathbb{N}$, correspondingly.

The general form of the regression model in this case is

$$X_j' = -\gamma - \frac{m(\theta)}{\nu} \cdot q(j) + \varepsilon_j.$$  \hspace{1cm} (3.23)

4 Method testing and application

4.1 Empirical test

For the sake of reproducibility, we now test our procedures using the fYp as a particular example as similar approach can be carried out for both the fractional linear death and
The fractional sublinear death processes. In point estimation testing, we evaluated the finite-sample properties (unbiasedness and homogeneity) by computing the average and the median absolute deviation (MAD) of the estimates using 1000 simulations for sample sizes 100, 500, and 1000. These values are shown in Table 2. The relative fluctuation (RF = 100% × MAD/mean) of $\hat{\nu}$ decreases from 19.23% (corresponds to $\nu = 0.1, n = 100$) to as little as 4.41% (with $\nu = 0.95$ and $n = 1000$). On the other hand, the residual-based $\hat{\nu}$ ranges from 19.23% (with $\nu = 0.95$ and $n = 1000$) to 1.89% (corresponds to $\nu = 0.95, n = 1000$). While $\hat{\lambda}$'s RF improves from 33.94% (corresponds to $\lambda = 0.5, n = 100$) to 30.48% (with $\lambda = 5$ and $n = 1000$), $\hat{\lambda}^{\text{res}}$'s RF decays faster from 55% ($\lambda = 1, n = 100$) to 25.29% ($\lambda = 5$ and $n = 1000$). In general, the relative fluctuations of the residual-based estimators tend to decay faster than the LS-based estimates. They are also less bias than the LS-based estimators especially for $n \leq 100$. Nonetheless, both the residual- and LS-based point estimators are asymptotically unbiased as expected.

Table 3 shows the averaged lower and upper 95% confidence bounds using the formulae in Sect. 3. These bounds used 1000 simulation runs for each of the sample sizes $n = 15, 30, 100,$ and 500. Note that the residual-based interval estimator $\hat{\lambda}^{\text{res}}$ utilized 500 bootstrap samples and $\hat{\sigma}_{\epsilon}^2 = \pi^2 (1/(3\hat{\nu}^2) - 1/6)$ is used to estimate the error variance in our LS-based procedures. Observe that some of the interval estimates for sample sizes $n = 15,$ and $n = 30$ are omitted as they are unreliable due to the multiple error warnings that showed up during the computation process. Moreover, the coverage probabilities to their true levels for the LS-based

| $(\nu, \lambda)$ | Estimator | $n = 15$ | $n = 30$ | $n = 100$ | $n = 500$ |
|------------------|-----------|---------|---------|-----------|-----------|
| (0.1, 1)         | $\hat{\nu}$ | (0.060, 0.158) | (0.071, 0.137) | (0.083, 0.118) | (0.084, 0.117) |
|                  | $\hat{\nu}^{\text{res}}$ | ($-35.8067, 58.864$) | ($-0.56, 2.885$) | ($0.297, 5.560$) | ($0.46, 2.652$) |
|                  | $\hat{\lambda}$ | (0.015, 2.183) | (0.020, 1.271) | (0.028, 0.551) | (0.044, 0.265) |
| (0.25, 0.1)      | $\hat{\nu}$ | (0.151, 0.389) | (0.211, 0.298) | (0.209, 0.296) | (0.231, 0.269) |
|                  | $\hat{\nu}^{\text{res}}$ | (0.034, 0.201) | (0.018, 0.291) | (0.018, 0.202) | (0.044, 0.265) |
|                  | $\hat{\lambda}$ | (0.099, 9.154) | (0.113, 5.261) | (0.164, 2.542) | (0.241, 1.246) |
| (0.5, 0.5)       | $\hat{\nu}$ | (0.319, 0.749) | (0.366, 0.664) | (0.424, 0.586) | (0.464, 0.536) |
|                  | $\hat{\nu}^{\text{res}}$ | (0.151, 0.389) | (0.211, 0.298) | (0.209, 0.296) | (0.231, 0.269) |
|                  | $\hat{\lambda}$ | (0.034, 0.704) | (0.020, 1.271) | (0.044, 0.265) | (0.074, 0.798) |
| (0.75, 0.25)     | $\hat{\nu}$ | (0.527, 1.057) | (0.573, 0.953) | (0.648, 0.857) | (0.704, 0.798) |
|                  | $\hat{\nu}^{\text{res}}$ | (0.008, 0.618) | (0.076, 0.453) | (0.087, 0.555) | (0.127, 0.573) |
| (0.95, 5)        | $\hat{\nu}$ | (0.719, 1.221) | (0.779, 1.150) | (0.848, 1.059) | (0.904, 0.999) |
|                  | $\hat{\nu}^{\text{res}}$ | (0.103, 1.389) | (0.261, 1.881) | (0.121, 0.978) | (0.151, 0.968) |
|                  | $\hat{\lambda}$ | (0.099, 9.154) | (0.113, 5.261) | (0.164, 2.542) | (0.241, 1.246) |
method is made faster by using the error variance estimate \( \hat{\sigma}^2 = \pi^2(1/(3\hat{\nu}^2) - 1/6) \). From Table 3, it is apparent that the residual-based interval estimates of \( \nu \) are narrower and are better centered around the true parameter values than the least-squares’ even when the sample size is as large as 500. In addition, our simulations showed that the asymptotic or non-bootstrapped residual-based interval estimator of \( \lambda \) gives more sensible results than the LS-based procedure for small samples. Nevertheless, the LS-based interval estimates for \( \lambda \) are more accurately centered than the bootstrapped residual-based estimates especially for large samples.

The corresponding coverage probabilities and the widths of the interval estimates above with a confidence level of 95 % are displayed in Table 4. When the sample size \( n = 15 \), the residual-based interval estimators have minimum coverage of 90.1 % and 91.1 % for \( \nu = 0.95 \) and \( \lambda = 0.1 \), respectively. When \( n = 500 \), the bootstrap interval estimator of \( \lambda \) has coverage probabilities which are closer to the true confidence level than the LS-based procedure for large values of \( \lambda \). However, the LS-based estimator of \( \lambda \) has a better coverage than the bootstrapped residual-based interval estimator for small \( \lambda \) values. The residual-based interval estimator for \( \lambda \) seemed to have slower convergence than the LS-based method. Furthermore, the residual-based estimator of \( \nu \) outperformed the LS-based method as its coverage probabilities are closer to 95 %, and has narrower intervals. Overall, the coverage probabilities and interval widths still provide good merits for our estimators even when the sample size is as small as \( n = 15 \).

### Table 4 Coverage probabilities and mean widths of 95 % interval estimates for different values of \( \nu \) and \( \lambda \)

| \((\nu, \lambda)\) | Estimator | \( n = 15 \) | \( n = 30 \) | \( n = 100 \) | \( n = 500 \) |
|----------------|----------|-------------|-------------|-------------|-------------|
|                | Coverage | Width       | Coverage    | Width       | Coverage    | Width       | Coverage    | Width       |
| (0.1, 1)       | \( \hat{\nu}_n \) | 0.956 | 0.098 | 0.958 | 0.066 | 0.949 | 0.078 | 0.952 | 0.033 |
|                | \( \hat{\nu}_{res} \) | 0.956 | 0.035 | 0.957 | 0.016 | 0.956 | 0.035 | 0.957 | 0.016 |
| (0.25, 0.1)    | \( \hat{\nu}_n \) | 0.947 | 0.238 | 0.953 | 0.162 | 0.940 | 0.167 | 0.949 | 0.107 |
|                | \( \hat{\nu}_{res} \) | 0.895 | 0.322 | 0.925 | 0.184 | 0.895 | 0.322 | 0.925 | 0.184 |
| (0.5, 0.5)     | \( \hat{\nu}_n \) | 0.952 | 0.430 | 0.953 | 0.298 | 0.954 | 0.161 | 0.949 | 0.072 |
|                | \( \hat{\nu}_{res} \) | 0.931 | 2.141 | 0.923 | 0.817 | 0.931 | 2.141 | 0.923 | 0.817 |
| (0.75, 0.25)   | \( \hat{\nu}_n \) | 0.921 | 0.529 | 0.945 | 0.379 | 0.941 | 0.481 | 0.950 | 0.291 |
|                | \( \hat{\nu}_{res} \) | 0.903 | 0.627 | 0.934 | 0.377 | 0.903 | 0.627 | 0.934 | 0.377 |
| (0.95, 5)      | \( \hat{\nu}_n \) | 0.901 | 0.502 | 0.927 | 0.317 | 0.941 | 0.210 | 0.947 | 0.095 |
|                | \( \hat{\nu}_{res} \) | 0.899 | 20.221 | 0.939 | 9.371 | 0.899 | 20.221 | 0.939 | 9.371 |

 Springer
Collectively, Tables 2–4 strongly indicate that the proposed point and interval estimators performed well in our computational tests. We emphasize that the point estimates could also be regarded as reasonable starting values for better iterative estimation algorithms.

4.2 Application

We now apply our proposed methods to a real dataset. In particular, we estimate the parameters of the fractional Yule model using the branching times for the plethodontid salamander dataset from Highton and Larson (1979) (see also Nee et al. 1994; Nee 2001). The 25 data points are the times measured from each node to the present of a phylogenetic tree, and can be downloaded from the package laser of the R software. The summary statistics of the inter-branching times of the plethodontid dataset are given in Table 5.

The point and the 95 % confidence interval estimates are given in Table 6. The LS-based point estimate (0.749) of the fractional parameter \( \nu \) seemed to suggest that the plethodontid salamander branching process is not a standard Yule process while the residual-based point estimate (1.119) appeared to suggest otherwise. Moreover, both the LS- and residual-based interval estimates indicated that \( \nu \) could be strictly less than one, which implies that a non-standard Yule process could model the plethodontid salamander dataset with a confidence level of 95 %. The residual-based point estimate (0.049) of \( \lambda \) is more conservative than the bootstrap- and LS-based estimate (0.049). A similar observation can be gleaned from the 95 % interval estimates, i.e., the residual-based 95 % interval estimate is narrower than the bootstrap- and LS-based interval estimates.

We also tested the residuals for normality using the Shapiro-Wilk, Anderson-Darling, Cramer-von Mises, Lilliefors, Pearson chi-square, and the Shapiro-Francia tests, which gave the \( p \)-values 0.811, 0.651, 0.619, 0.609, 0.849, and 0.461, correspondingly. Hence, these \( p \)-values indicated good fit of the fractional Yule process to the plethodontid salamander data.

5 Concluding remarks

We have proposed closed-form expressions of the estimators of the parameters \( \nu \) and \( \lambda \) for the fractional linear birth or Yule, the fractional linear death, and the fractional sublinear death processes. The estimators were derived by taking advantage of the known structural form of the logarithm of the random inter-event times and the well-studied least squares regression procedure. The explicit formulas led to computationally simple and fast parameter estimation procedures. The inter-death time distributions and variances of the fractional linear and sublinear death processes were also obtained. These statistical properties were necessary for generating sample trajectories and for our estimation procedures to be applicable in these processes. It has also been shown that the proposed procedure can be easily extended to certain models that have different model parameterizations than the linear ones. The proposed methods were used to model a real physical process. Generally, the extensive computational tests showed favorable results for the proposed estimators.

We cite some extensions which would be worth pursuing in the future. For instance, improving the small-sample performance of the least squares-based estimators and developing other estimators using the likelihood approach or a re-sampling technique would be valuable pursuits. The application of these methods in practice, and the characterization of the appropriate functions \( m(\theta) \) and \( q(j) \) would also be of interest.

Acknowledgements The authors are highly grateful to the editor, the associate editor, and the reviewers for their valuable comments and suggestions that improved the article. This research was supported under the Louisiana Board of Regents Research Competitiveness Subprogram grant LEQSF (2011-14)-RD-A-15.

References

Aldous, D.J.: Stochastic models and descriptive statistics for phylogenetic trees, from Yule to today. Stat. Sci. 16(1), 23–34 (2001)
Cahoy, D.O., Polito, F.: Simulation and estimation for the fractional Yule process. Methodol. Comput. Appl. Probab. 14(2), 383–403 (2012)
Cahoy, D.O., Uchaikin, V.V., Woyczynski, W.A.: Parameter estimation for fractional Poisson processes. J. Stat. Plan. Inference 140, 3106–3120 (2010)
Ferguson, T.: A Course in Large Sample Theory. Chapman & Hall, London (1996)

Highton, R., Larson, L.A.: The genetic relationships of the salamanders of the genus Plethodon. Syst. Zool. 28, 579–599 (1979)

Montgomery, D.C., Peck, E.A., Vining, G.G.: Introduction to Linear Regression Analysis 4 ed. Wiley, New York (2006)

Nee, S., Holmes, E.C., May, R.M., Harvey, P.H.: Extinction rates can be estimated from molecular phylogenies. Philos. Trans. R. Soc. Lond. B, Biol. Sci. 344, 77–82 (1994)

Nee, S.: Inferring speciation rates from phylogenies. Evolution 55, 661–668 (2001)

Orsingher, E., Polito, F.: Fractional pure birth processes. Bernoulli 16, 858–881 (2010)

Orsingher, E., Polito, F., Sakhno, L.: Fractional non-linear, linear and sublinear death processes. J. Stat. Phys. 141, 68–93 (2010)

Paradis, E.: Analysis of Phylogenetics and Evolution with R, 2nd edn. Springer, New York (2012)

Uchaikin, V.V., Cahoy, D.O., Sibatov, R.T.: Fractional processes: from Poisson to branching one. Int. J. Bifurc. Chaos 18, 2717–2725 (2008)