Classes of Equivalence and Identifiability of Age-Dependent Branching Processes

Ollivier Hyrien

University of Rochester

Workshop on Branching Processes and their Applications
Badajoz, April 7-10, 2015

Joint work with Rui Chen

Introduction

- Identifiability is central to most statistical procedures (e.g., asymptotic normality; construction of confidence intervals), and in interpreting models
- Studied for deterministic models (e.g., differential equation models) but literature on identifiability of stochastic processes remains limited
- Prakasa Rao, B.L.S. (1992). *Identifiability in Stochastic Models, Characterization of Probability Distributions.* Academic Press.
- Not studied for branching processes
- The property is often claimed without formal proof, partly out of convenience
- We investigated the issue for age-dependent branching processes (single-type, no immigration) and constructed a partition of the family of processes into classes of equivalence

A class of age-dependent branching processes

- The process begins with a single cell of age 0
- Every cell evolves independently of every other cell

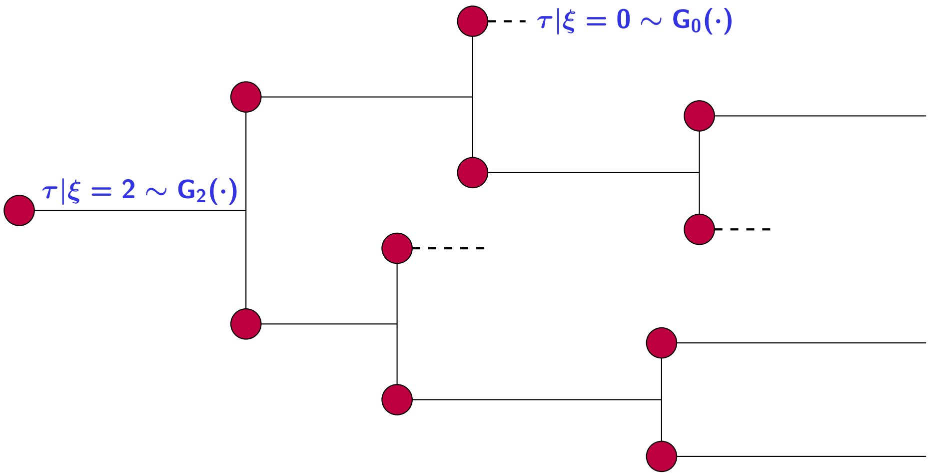
Offspring, ξ

- Upon completion of its lifespan, every cell produces a random number of offspring $\xi \in \mathcal{J} = \{0, 1, \cdots, J\}$
- $p := (p_0, \ldots, p_J)$, where $p_j := \mathbb{P}(\xi = j)$, $j \in \mathcal{J}$, for the offspring distribution
- In cell biology: $J = 2$ and $\xi = \begin{cases} 0 & \text{cell death} \\ 1 & \text{cell resting/differentiation} \\ 2 & \text{cell division} \end{cases}$
- Put $\mathcal{J}^*(p) := \{j \in \mathcal{J} : p_j > 0\}$

Lifespan, τ

- Put $G_j(t) := \mathbb{P}(\tau \le t | \xi = j)$, $t \ge 0$, for every $j \in \mathcal{J}^*(p)$
- Assume that $G_j \in \mathcal{D}$, $j \in \mathcal{J}^*(p)$, where $\mathcal{D}$ denotes the collection of all proper distributions

An example of tree generated by the process



$\sim$Sevastyanov process: $\mathcal{L}(\xi, \tau) = \mathcal{L}(\xi|\tau)\mathcal{L}(\tau)$

here: $\mathcal{L}(\xi, \tau) = \mathcal{L}(\tau|\xi)\mathcal{L}(\xi)$ (easier to interpret)

Bellman-Harris: $\mathcal{L}(\xi, \tau) = \mathcal{L}(\tau)\mathcal{L}(\xi)$

A non-identifiable formulation via competing risks

- The process begins with a single cell of age 0
- Every cell evolves independently of every other cell

Latent failure times, $\mathbf{T} = (T_0, \ldots, T_J)$

- $\mathbf{T} \sim F(t_0, \ldots, t_J)$
- Interpretation: T_j is the hypothetical time needed by some intra-cellular processes to complete their tasks so the cell would generate j daughters
- Special case: independent competing risks (Waugh 1955)

$$F(t_0, \ldots, t_J) = F_1(t_0) \ldots F_J(t_J)$$

Offspring and lifespan, (ξ, τ)

- If $T_{j_0} = \min_{j=0,\ldots,J} T_j$ then $\begin{cases} \xi = j_0 \\ \tau = T_{j_0} \end{cases}$

A non-identifiable formulation via competing risks

- Competing risks branching processes used to gain insights into intra-cellular processes involved in cell fate determination

$\Rightarrow$ Independent latent failure times/competing risks describe these intra-cellular processes

- This conclusion may not provide much insights because non-identifiability occurs in competing risks models if the latent failure times are not independent (Cox, 1959; Tsiatis 1975; Peterson 1976, and many others)
- Specifically, for every distribution $F(t_0, \ldots, t_J)$ for dependent failure times $\mathbf{T}$, there exists a distribution

$$F^{(*)}(t_0, \ldots, t_J) = F_1^{(*)}(t_0) \ldots F_1^{(*)}(t_J)$$

  such that $\mathcal{L}(\tau, \xi | F) = \mathcal{L}(\tau, \xi | F^{(*)})$
- Our class of processes may be equivalent to the class of competing risks models, but each process is a representative member of one equivalence class, such that the above type of nonidentifiability is not a concern

Formulation of the the problem

Put

- $G = \{G_j, j \in \mathcal{J}^*(p)\}$
- $C = (p, G)$: *characteristics* of the process
- $Z(t)$: the size of the population at time $t \geq 0$

Define

- $\mathcal{P}$ denotes the set of all processes that satisfy the above assumptions
- $\mathcal{P}_0 \subset \mathcal{P}$ is the subset of processes with characteristics (p, G) satisfying $p_1 = 0$

Question:

Are there distinct characteristics (p, G) under which the distribution of $Z(t)$ is identical for all $t \geq 0$?

Answering this question will inform us about what can or cannot be estimated by only observing $Z(t)$

Equivalence

Definition (equivalence)

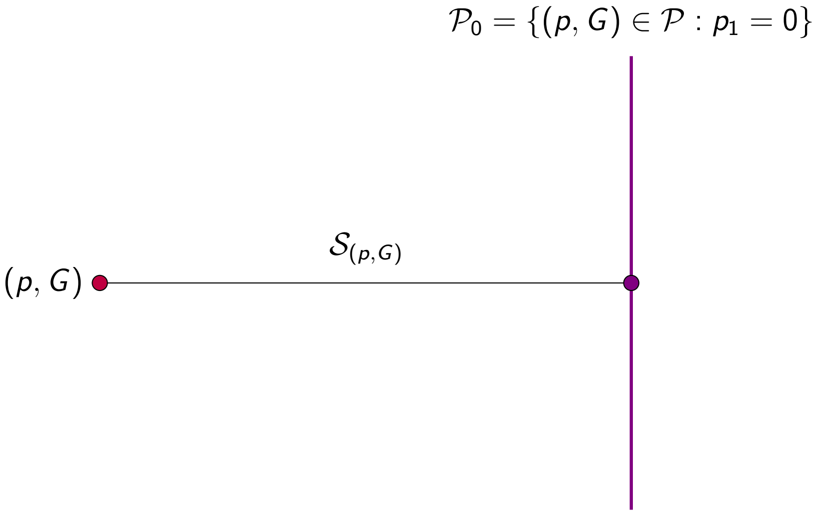
We shall say that two processes with characteristics (p, G) *and* $(\hat{p}, \hat{G})$ *are equivalent if, for all* $t \geq 0$*, the distribution of* $Z(t)$ *is the same under either characteristics.*

- Let $\mathcal{C}_{p,G}$ denote the equivalence class defined as the collection of all processes included in $\mathcal{P}$ that are equivalent to the process with characteristics (p, G).
- Objective: construct the class $\mathcal{C}_{p,G}$ for any admissible characteristics (p, G).
- Application to study identifiability of families of parametric models
  - A process with exponential lifespan (not always identifiable)
  - A process with (non-exponential) gamma distributed lifespan (always identifiable)
  - A Smith-Martin process (not always identifiable)

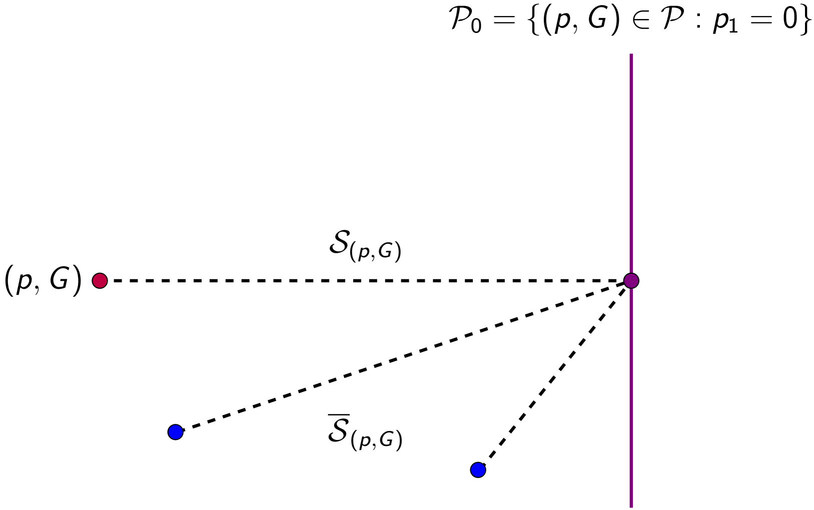
Step 1

(p, G)

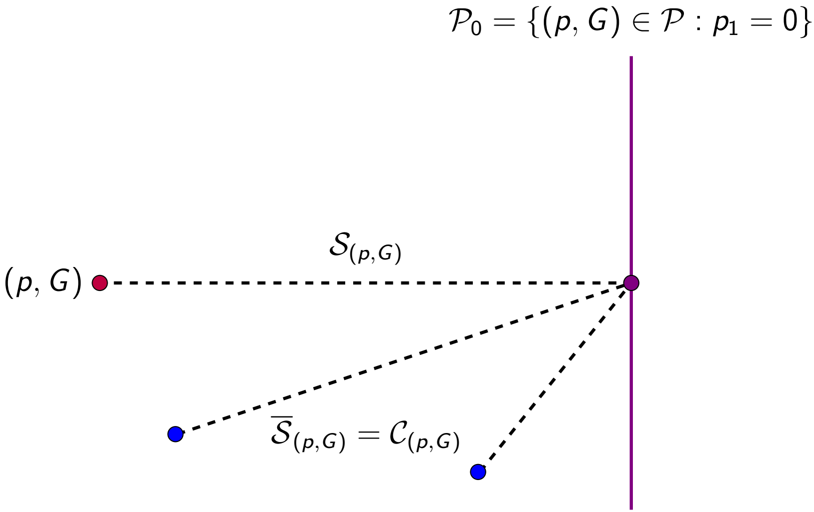
Step 1



Step 2


$\mathcal{P}_0 = \{(p, G) \in \mathcal{P} : p_1 = 0\}$
$\mathcal{S}_{(p,G)}$
(p, G)
$\overline{\mathcal{S}}_{(p,G)}$

Step 3


$\mathcal{P}_0 = \{(p, G) \in \mathcal{P} : p_1 = 0\}$
$\mathcal{S}_{(p,G)}$
(p, G)
$\overline{\mathcal{S}}_{(p,G)} = \mathcal{C}_{(p,G)}$

Step 1: a collection of equivalent processes

- Let $(p, G) \in \mathcal{P}$, assume that $p_1 \in [0, 1)$
- For every $a \in [0, p_1]$, define

$$p^{(a)} = (p_0^{(a)}, \dots, p_J^{(a)}) \text{ where } p_j^{(a)} := \begin{cases} \frac{p_j}{1-a} & j \in \mathcal{J} \setminus \{1\} \\ \frac{p_1 - a}{1-a} & j = 1 \end{cases}$$

and, for every $t \geq 0$ and $j \in \mathcal{J}^*(p)$, put

$$G_j^{(a)}(t) := (1-a) G_j * \sum_{k=0}^{\infty} a^k G_1^{*k}(t)$$

where $G_j * G_1(t) := \int_0^t G_j(t-x) dG_1(x)$

- Let $\mathcal{S}_{p,G}$ denote the collection of processes with characteristics $(p^{(a)}, G^{(a)})$, $a \in [0, p_1]$

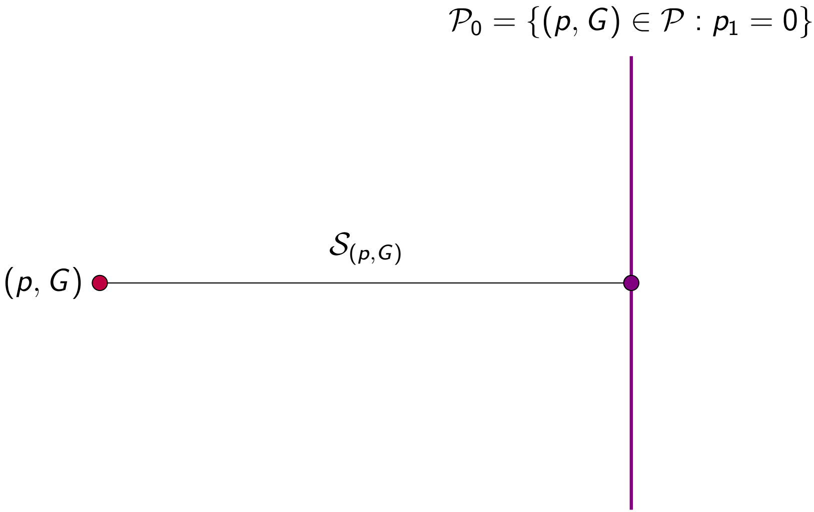
- $\mathcal{S}_{p,G}$ includes
  - the process (p, G) since $(p^{(0)}, G^{(0)}) = (p, G)$
  - one process from $\mathcal{P}_0$ since $p_1^{(p_1)} = 0$

Step 1: A collection of equivalent processes

- Theorem

  For all $t \geq 0$, the distribution of the population size process $Z(t)$ is identical under all processes included in $\mathcal{S}_{p,G}$; that is, $\mathcal{S}_{p,G} \subseteq \mathcal{C}_{p,G}$.

Step 1



Step 2: A larger collection of equivalent processes

- Setting $a = p_1$ yields

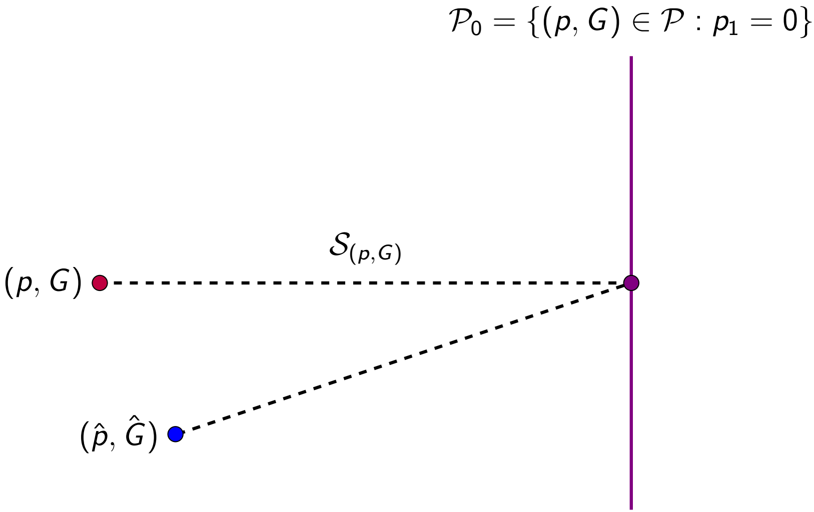
$$
\begin{aligned}
p_j^{(p_1)} &= \begin{cases} \frac{p_j}{1-p_1} & j \in \mathcal{J}\backslash\{1\} \\ 0 & j = 1, \end{cases} \\
\mathcal{L}_{g_j^{(p_1)}}(s) &= \frac{(1-p_1)\mathcal{L}_{g_j}(s)}{1 - p_1\mathcal{L}_{g_1}(s)}, \quad j \in \mathcal{J}^*(p)\backslash\{1\},
\end{aligned}
$$

  which identifies a process in $\mathcal{P}_0$
- Any process with characteristics $(\hat{p}, \hat{G})$ that satisfy

$$
\begin{cases} \hat{p}_j^{(\hat{p}_1)} = p_j^{(p_1)} & j \in \mathcal{J}\backslash\{1\} \\ \mathcal{L}_{\hat{g}_j^{(\hat{p}_1)}}(s) = \mathcal{L}_{g_j^{(p_1)}}(s) & j \in \mathcal{J}^*(p)\backslash\{1\} \end{cases} \tag{1}
$$

  belongs to $\mathcal{C}_{p,G}$ (here, $\mathcal{L}_h(s)$ is the Laplace transform of any H)

Step 2



Step 2: A larger collection of equivalent processes

- By solving eqns. (1) we find that $(\hat{p}, \hat{G})$ satisfies

$$\begin{cases} \hat{p}_j = p_j^{(p_1)}(1-\hat{p}_1) & j \in \mathcal{J}\backslash\{1\} \\ \mathcal{L}_{\hat{g}_j}(s) = \mathcal{L}_{g_j}^{(p_1)}(s)\{1-\hat{p}_1\mathcal{L}_{\hat{g}_1}(s)\}/(1-\hat{p}_1) & j \in \mathcal{J}^*(p)\backslash\{1\}, \end{cases} \tag{2}$$

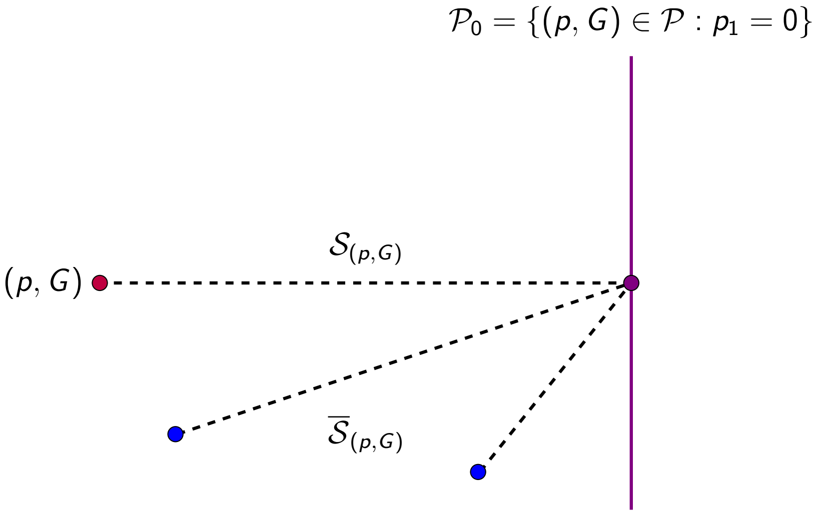
  where $\hat{p}_1 \in [0,1)$ and $\hat{G}_1 \in \mathcal{D}_{p,G}$, and where $\mathcal{D}_{p,G} \subseteq \mathcal{D}$
- Write $(p_{\hat{p}_1,\hat{G}_1}, G_{\hat{p}_1,\hat{G}_1})$ for any characteristics that satisfy eqns. (2).
- Then, the collection of processes

$$\overline{\mathcal{S}}_{p,G} := \bigcup_{\hat{p}_1\in[0,1)} \bigcup_{\hat{G}_1\in\mathcal{D}_{p,G}} \{\text{process with characteristics } (p_{\hat{p}_1,\hat{G}_1}, G_{\hat{p}_1,\hat{G}_1})\}$$

  is included in $\mathcal{C}_{p,G}$. It is also clear that $\mathcal{S}_{p,G} \subset \overline{\mathcal{S}}_{p,G}$.
- So far, we have shown that

$$\mathcal{S}_{p,G} \subset \overline{\mathcal{S}}_{p,G} \subseteq \mathcal{C}_{p,G}$$

Step 2



$\mathcal{P}_0 = \{(p, G) \in \mathcal{P} : p_1 = 0\}$
$\mathcal{S}_{(p,G)}$
(p, G)
$\overline{\mathcal{S}}_{(p,G)}$

Step 3: Exhaustivity of $\overline{\mathcal{S}}_{p,G}$ (when $J = 2$ and $J = 3$)

- Since $\overline{\mathcal{S}}_{p,G} \subseteq \mathcal{C}_{p,G}$, it is sufficient to prove that $\mathcal{C}_{p,G} \subseteq \overline{\mathcal{S}}_{p,G}$
- Let $(\hat{p}, \hat{G})$ be the characteristics of any process included in $\mathcal{C}_{p,G}$. Then, by construction, the process with characteristics $(\hat{p}^{(\hat{p}_1)}, \hat{G}^{(\hat{p}_1)})$ belongs to $\mathcal{P}_0$

Lemma

Suppose that $J = 2$ or $J = 3$. For every admissible (p, G), the equivalence class $\mathcal{C}_{p,G}$ includes a single process in $\mathcal{P}_0$.

- Hence $(\hat{p}^{(\hat{p}_1)}, \hat{G}^{(\hat{p}_1)}) = (p^{(p_1)}, G^{(p_1)})$ and the process with characteristics $(\hat{p}, \hat{G})$ belongs to $\overline{\mathcal{S}}_{p,G}$, which implies that $\mathcal{C}_{p,G} \subseteq \overline{\mathcal{S}}_{p,G}$

Theorem

We have $\mathcal{C}_{p,G} = \overline{\mathcal{S}}_{p,G}$ for every admissible (p, G) when $J = 2$ and $J = 3$.

Characterization of $\mathcal{C}_{p,G}$ using moments when $J = 2$

- In applications, model parameters are sometimes estimated using moments of the process rather than its distribution.
- Then, a relevant question is which moments are sufficient to fully characterize the equivalence class $\mathcal{C}_{p,G}$?
- Answer when $J = 2$: expectation and variance

Theorem

Assume that $J = 2$ and that the marginal distribution of $\{Z(t), t \geq 0\}$, is determined by its moments. Then, $\mathcal{C}_{p,G} = \{$processes with characteristics $(\hat{p}, \hat{G}) : \hat{m}(t) = m(t), \hat{m}_2(t) = m_2(t), t \geq 0\}$.

Application to model identifiability

- Suppose that G belongs to a family of distributions (e.g., exponential, gamma...)
- Let $\mathcal{M} \subset \mathcal{P}$ denote the resulting family of processes
- Let $(p, G) \in \mathcal{M}$
- We want to know whether (p, G) is identifiable within the family of models $\mathcal{M}$
- It is sufficient to show that

$$\mathcal{C}_{p,G}^{\mathcal{M}} := \mathcal{C}_{p,G} \cap \mathcal{M} = \{(p, G)\}$$

Exponentially distributed lifespan

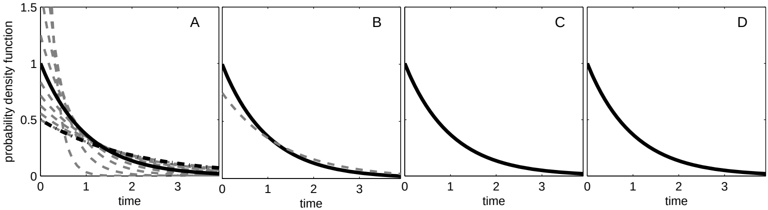
Corollary

Suppose that $J = 2$ and, for every $j \in \mathcal{J}^*(p)$, that $G_j(t) = 1 - e^{-\psi_j t}$, $t \geq 0$, for some $\psi_j \in \mathbb{R}_+^*$. Then, (p, G) is uniquely identified by the process $\{Z(t), t \geq 0\}$, except in the following cases:

Case 1: *If $\psi_j = \psi$, $j \in \mathcal{J}^*(p)$, then $\mathcal{C}_{p,G}^{\mathcal{M}}$ includes the processes with characteristics $(\hat{p}, \hat{G})$ such that $\hat{p}_1 \in [0,1), \hat{p}_j = p_j(1-\hat{p}_1)/(1-p_1), j \in \mathcal{J}^*(p)\backslash\{1\}, \hat{\psi}_j = \psi(1-p_1)/(1-\hat{p}_1), j \in \mathcal{J}^*(\hat{p})$.*

Case 2: *If $\psi_j = \psi$, $j \in \mathcal{J}^*(p)\backslash\{1\}$, $p_1 \in (0,1)$, $\psi < \psi_1$, and $p_1 \neq 1 - \psi/\psi_1$, then $\mathcal{C}_{p,G}^{\mathcal{M}}$ consists of the processes with characteristics (p, G) and $(\hat{p}, \hat{G})$ where $\hat{p}_1 = 1 - \psi/\psi_1$, $\hat{p}_j = p_j\psi/\{(1-p_1)\psi_1\}$, $\hat{\psi}_1 = \psi_1$, $\hat{\psi}_0 = \hat{\psi}_2 = (1-p_1)\psi_1$, $j \in \mathcal{J}^*(p)\backslash\{1\}$.*

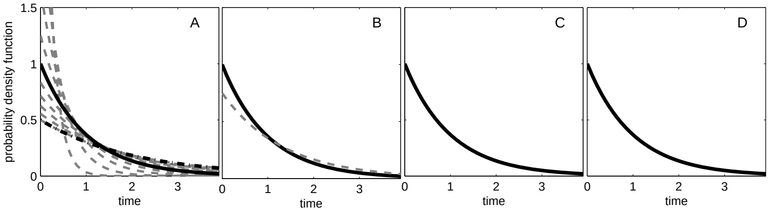
An illustrative example



- Assume that $\xi \sim p = (\frac{1}{5}, \frac{1}{2}, \frac{3}{10})$ and $\tau|\xi \sim$ exponential with parameters $\psi_0 = \psi_1 = \psi_2 = 1$ (Case 1)
- Equivalence class $\mathcal{C}_{p,G}^{\mathcal{M}} = \{(\hat{p}, \hat{G}) \in \mathcal{M}$:

$$\begin{cases} \hat{p} = \left(\frac{2}{5}(1-\hat{p}_1), \hat{p}_1, \frac{3}{5}(1-\hat{p}_1)\right) \\ \hat{\psi}_0 = \hat{\psi}_1 = \hat{\psi}_2 = \frac{1}{2(1-\hat{p}_1)} \end{cases} \quad \text{where } \hat{p}_1 \in [0,1)\}$$

- Setting $\hat{p}_1 = 0$ yields $\hat{p} = (\frac{2}{5}, 0, \frac{3}{5})$ and $\hat{\psi}_0 = \hat{\psi}_2 = 0.5$, which belongs to $\mathcal{P}_0$
- Fig. 1.A displays examples of probability density functions $\hat{g}_2$ for a sample of processes that belong to $\mathcal{C}_{p,G}$.



- In Fig. 1.B, we set $\begin{cases} p = \left(\frac{1}{5}, \frac{1}{2}, \frac{3}{10}\right) \\ \psi_0 = \psi_2 = 1 \\ \psi_1 = 1.5 \end{cases} \Longleftrightarrow \begin{cases} \hat{p} = \left(\frac{4}{15}, \frac{1}{3}, \frac{2}{5}\right) \\ \hat{\psi}_0 = \hat{\psi}_2 = 0.75 \\ \hat{\psi}_1 = 1.5 \end{cases}$

- In Fig. 1.C, we set $\begin{cases} p = \left(\frac{1}{5}, \frac{1}{2}, \frac{3}{10}\right) \\ \psi_0 = \psi_2 = 1 \\ \psi_1 = 0.75 \end{cases}$ no equivalent process

- In Fig. 1.D, we set $\begin{cases} p = \left(\frac{1}{5}, \frac{1}{2}, \frac{3}{10}\right) \\ \psi_0 = \psi_1 = 1 \\ \psi_2 = 2 \end{cases}$ no equivalent process

Gamma distributed lifespan

Corollary

Suppose that $J = 2$ *and, for every* $j \in \mathcal{J}^*(p)$*, that* G_j *is a gamma distribution with parameters* $\kappa_j > 0$*,* $\omega_j > 0$ *and* $\omega_j \neq 1$*; that is,*

$$G_j(t) := \int_0^t \frac{\kappa_j^{\omega_j}}{\Gamma(\omega_j)} x^{\omega_j - 1} e^{-\kappa_j x} dx$$

for some $\psi_j := (\omega_j, \kappa_j) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$*,* $j \in \mathcal{J}^*(p)$*. Then,* (p, G) *is uniquely identified by the process* $\{Z(t), t \geq 0\}$*.*

Identifiability of a Smith-Martin process

Corollary

Suppose that $J = 2$ and, for every $j \in \mathcal{J}^*(p)$, that

$$G_j(t) = 1 - e^{-\psi_j(t-\delta_j)} \quad (t \geq \delta_j).$$

Then, (p, G) is uniquely identified by $\{Z(t), t \geq 0\}$ except:

Case 1: *If $\psi_j = \psi$, $j \in \mathcal{J}^*(p)$, and $\delta_1 = 0$ when $p_1 = 0$, $\mathcal{C}^{\mathcal{SM}}_{p,G}$ includes the S.M. processes with characteristics $(\hat{p}, \hat{G}) \in \{\hat{p}_1 \in (0,1), \hat{p}_j = p_j(1-\hat{p}_1)/(1-p_1), \hat{\delta}_1 = 0, \hat{\delta}_j = \delta_j, \hat{\psi}_1 = \hat{\psi}_j = \psi(1-p_1)/(1-\hat{p}_1), j \in \mathcal{J}^*(p)\backslash\{1\}\} \cup \{\hat{p}_1 = 0, \hat{p}_j = p_j/(1-p_1), \hat{\delta}_j = \delta_j, \hat{\psi}_j = \psi(1-p_1), j \in \mathcal{J}^*(p)\backslash\{1\}\}$.*

Case 2: *If $p_1 \in (0,1)$, $p_1 \neq 1 - \psi/\psi_1$, $\delta_1 = 0$, $\psi_j = \psi$, $j \in \mathcal{J}^*(p)\backslash\{1\}$, and $\psi < \psi_1$, $\mathcal{C}^{\mathcal{SM}}_{p,G}$ consists of the S.M. processes with characteristics (p, G) and $(\hat{p}, \hat{G})$ where $\hat{p}_1 = 1 - \psi/\psi_1$, $\hat{p}_j = p_j\psi/\{(1-p_1)\psi_1\}$, $\hat{\delta}_1 = 0, \hat{\delta}_j = \delta_j, \hat{\psi}_1 = \psi_1$, $\hat{\psi}_j = (1-p_1)\psi_1$, $j \in \{0, 2\}$.*

Conclusion

- We have partitioned a family of age-dependent branching processes into classes of equivalence
- These equivalence classes are easy to construct by solving a set of equations
- This result is useful to study identifiability of parametric age-dependent branching processes

- Chen, R., and Hyrien, O. (2014). On classes of equivalence and identifiability of age-dependent branching processes. *Advances in Applied Probability*, 46, 704-718.

Thank you!