STOCHASTIC COMPARISONS AMONG CONDITIONAL PROCESSES DERIVED FROM A SEMIEXPLOSIVE GALTON-WATSON BRANCHING PROCESS

Michael D. Perlman
Department of Statistics
University of Washington
Seattle WA 98105 USA
michael@stat.washington.edu

1 March 2012 (revised 26 May 2013)

Abstract

Let $\mathcal{X} \equiv \{X_n \mid n \geq 0\}$ be a semiexplosive Galton-Watson branching process, where ultimate extinction ($X_n \to 0$) and ultimate explosion ($X_n \to \infty$) both occur with positive probability. Stochastic orderings among $\mathcal{X}$ and four conditional processes derived from it are studied: $\tilde{\mathcal{X}} \equiv \mathcal{X} \mid$ extinction, $\check{\mathcal{X}} \equiv \mathcal{X} \mid$ explosion, $\check{\mathcal{X}} \equiv \mathcal{X} \mid$ no individual ever dies without offspring, and $\check{\mathcal{X}} \equiv \{X_n\} \mid$ explosion, where $\check{X}_n$ is the number of individuals in the $n$th generation whose lines of descent never die out. It might be expected that $\check{\mathcal{X}} \prec (\mathcal{X}, \tilde{\mathcal{X}}) \prec (\check{\mathcal{X}}, \tilde{\mathcal{X}})$ and $\check{\mathcal{X}} \prec \check{\mathcal{X}}$, but only four of these seven stochastic orderings hold in general. More refined results are given for special cases, including geometric and Poisson offspring distributions. An application to the problem of predicting extinction or explosion is noted.
1. The Galton-Watson branching process

The classical Galton-Watson (GW) branching process is a discrete-time Markov chain that describes the growth or decline of a population that reproduces by simple branching, or splitting. Applications range from population growth models to nuclear chain reactions. The classic reference is Harris [5] Ch. I; also see Feller [2], Karlin [7], Athreya and Ney [1], Jagers [6], Taylor and Karlin [9], and Guttrop [3] among many others.

For each time \( n = 0, 1, 2 \ldots \) let \( X_n \) denote the population size at time \( n \); assume that \( X_0 = 1 \). At time \( n = 0 \) this single individual splits into a random number \( \xi_1^{(0)} \sim \xi \) of first-generation offspring, where the random variable (rv) \( \xi \) has probability distribution \( (p_0, p_1, p_2, \ldots) \) on \( \{0, 1, 2, \ldots \} \). The \( i \)-th individual in generation \( n \geq 1 \) similarly splits into a random number \( \xi_i^{(n)} \sim \xi \) of \((n+1)\)-th generation offspring independently of its siblings. Thus the total population size in the \((n+1)\)-th generation satisfies

\[
X_{n+1} = \xi_1^{(n)} + \cdots + \xi_{X_n}^{(n)}, \quad n \geq 0,
\]

where \( \xi_1^{(n)}, \ldots, \xi_{X_n}^{(n)} \) are iid rv's each \( \sim \xi \). If \( p_0 = 1 \) then \( X_n \equiv 0 \) for all \( n \geq 1 \); if \( p_1 = 1 \) then \( X_n \equiv 1 \) for all \( n \geq 0 \). Thus to avoid these trivial cases we assume that \( p_0 < 1 \) and \( p_1 < 1 \).

Let \( \phi \equiv \phi_\xi \) denote the probability generating function (pgf) of \( \xi \):

\[
\phi(s) = E(s^\xi) = p_0 + p_1 s + p_2 s^2 + p_3 s^3 + \cdots,
\]

defined for \( 0 \leq s \leq 1 \). Then \( \phi(0) = p_0, \phi(1) = 1, 0 \leq \phi(s) \leq 1 \) for \( 0 \leq s \leq 1 \), \( \phi'(1) = E(\xi) \), and \( \phi \) is strictly increasing and convex on \([0, 1]\); it is strictly convex unless \( p_0 + p_1 = 1 \), in which case \( \phi(s) = p_0 + p_1 s \) is linear and no growth is possible. Then (1.1) yields the following fundamental results:

Proposition 1.1. (i) For the GW process \( \mathcal{X} \equiv \{X_n \mid n \geq 0\} \), the pgf of \( X_n \) (\( n \geq 1 \)) is the \( n \)th functional iterate \( \phi_n \) of \( \phi \), that is,

\[
\phi_{X_n}(s) \equiv \phi_n(s) = \underbrace{\phi \cdots \phi(\phi(s)) \cdots}_{n \text{ times}}.
\]

(ii) The joint pgf \( \phi_{X_1, \ldots, X_n}(s_1, \ldots, s_n) \equiv E(s_1^{X_1} \cdots s_n^{X_n}) \) of \((X_1, \ldots, X_n)\) is

\[
\phi_{X_1, \ldots, X_n}(s_1, \ldots, s_n) = \phi(s_1 \phi(s_2 \phi(\cdots s_{n-1} \phi(s_n) \cdots))).
\]
Set $\mu = \mathbb{E}(\xi) = \mathbb{E}(X_1) > 0$; for simplicity assume $\mu < \infty$. From (1.1) or (1.3), $\mathbb{E}(X_n) = \mu^n$. The GW process is called subcritical (critical) if $\mu < 1$ (1), in which case extinction always occurs: \( \Pr[X_n \to 0] = 1 \). The subcritical and critical cases together constitute the subexplosive case. The process is explosive if $\mu > 1$ and $p_0 = 0$ (so death cannot occur); here explosion always occurs: \( \Pr[X_n \to \infty] = 1 \). The process is called semiexplosive if $\mu > 1$ and $p_0 > 0$ (so death can occur); here extinction and explosion occur with positive probabilities $u_p$ and $1 - u_p$ respectively, where $0 < u_p < 1$ is the unique solution in $(0, 1)$ of the equation

\[
(1.5) \quad \phi(s) = s.
\]

**Figure 1 (subexplosive):**

$\mu \equiv \phi'(1) \leq 1$.

**Figure 2 (semiexplosive):**

$\mu \equiv \phi'(1) > 1, \ p_0 > 0$.

**Figure 3 (explosive):** $\mu \equiv \phi'(1) > 1, \ p_0 = 0$. 

3
In this paper we shall consider a semiexplosive GW process. Because both extinction and explosion can occur, it is of interest to examine the conditional behavior of the process given either ultimate extinction or ultimate explosion. Four conditional processes will be defined and compared stochastically to each other and to the original GW process. The results are not entirely as expected.

2. Four conditional processes based on a semiexplosive GW process

Consider a semiexplosive GW process \( \mathcal{X} \equiv \{X_n | n \geq 0\} \) \( (X_0 = 1) \) with pgf \( \phi \). Referring to Figure 2, consider the functions \( \bar{\phi} \) and \( \tilde{\phi} \) determined by those portions of the graph of \( \phi \) that lie in the squares \([0, u]^2\) and \([u, 1]^2\), respectively, but rescaled by the factors \(1/u\) and \(1/(1-u)\) to transform each of these two smaller squares to the unit square \([0, 1]^2\). Thus for \(0 \leq s \leq 1\),

\[
(2.1) \quad \bar{\phi}(s) = \frac{1}{u} \phi(us), \\
(2.2) \quad \tilde{\phi}(s) = \frac{\phi((1-u)s + u) - u}{1-u}. 
\]

First consider \( \bar{\phi} \). From (2.1),

\[
(2.3) \quad \bar{\phi}(s) = \sum_{k=0}^{\infty} p_k u^{k-1} s^k,
\]

so \( \bar{\phi} \) is itself a pgf corresponding to the probability distribution \( (\bar{p}_0, \bar{p}_1, \ldots) \), where \( \bar{p}_k \equiv p_k u^{k-1} \) satisfies \( \sum_{k=0}^{\infty} \bar{p}_k = 1 \) by (1.5). We now show that the conditional process

\[
(2.4) \quad \bar{\mathcal{X}} \equiv \mathcal{X} | \text{extinction},
\]

is itself a subexplosive GW process with \( \bar{\phi} \) as its generating pgf. In fact \( \bar{\phi} \) is subcritical since \( \bar{\phi}'(1) = \bar{\phi}'(u) < 1 \) (see Figure 2).

**Proposition 2.1.** (Athreya and Ney [1] §I.12, Theorem 3)). For a semiexplosive GW process \( \mathcal{X} \) with pgf \( \phi \), the distribution of \( \mathcal{X} | \text{extinction} \) is the same as the distribution of a subcritical GW process with generating pgf \( \bar{\phi} \).
Short Proof. Conditional on ultimate extinction the family line of each individual eventually terminates, so the conditional process \( X \) behaves like a subcritical GW process. [See the Long Proof for a rigorous proof of this fact.] It remains to show that \( \tilde{\phi} \) is the pgf for this process. But for this, simply note that for each \( k = 0, 1, 2, \ldots \), Bayes formula yields

\[
\Pr[\bar{X}_1 = k] \equiv \Pr[X_1 = k | \text{extinction}] = \frac{\Pr[\text{extinction} | X_1 = k] \Pr[X_1 = k]}{\Pr[\text{extinction}]} = \frac{u^k p_k}{u} \equiv \tilde{p}_k,
\]

which is the coefficient of \( s^k \) in \( \tilde{\phi}(s) \), as required.

Long Proof. By Proposition 1.1(ii), it suffices to show that for each \( n \geq 1 \), the joint pgf \( \phi_{\bar{X}_1, \ldots, \bar{X}_n} \) of \( (\bar{X}_1, \ldots, \bar{X}_n) \) satisfies

\[
\phi_{\bar{X}_1, \ldots, \bar{X}_n}(s_1, \ldots, s_n) = \tilde{\phi}(s_1 \tilde{\phi}(s_2 \tilde{\phi}(\ldots s_{n-1} \tilde{\phi}(s_n) \ldots)))).
\]

The proof proceeds by induction on \( n \).

For \( n = 1 \), (2.6) follows by Bayes formula as in (2.5). Next, assume that (2.6) holds for \( n \) and consider the case \( n + 1 \). Then

\[
\phi_{\bar{X}_1, \ldots, \bar{X}_{n+1}}(s_1, \ldots, s_{n+1}) = \mathbb{E} \left[ \prod_{i=1}^{n+1} s_i^{X_i} \mid \text{extinction} \right]
\]

\[
= \mathbb{E} \left\{ \prod_{i=1}^{n} s_i^{X_i} \cdot \mathbb{E} \left[ \xi_1^{(n)} + \cdots + \xi_{X_n}^{(n)} \mid X_1, \ldots, X_n, \text{extinction} \right] \mid \text{extinction} \right\}
\]

\[
* = \mathbb{E} \left\{ \prod_{i=1}^{n} s_i^{X_i} \cdot (\tilde{\phi}(s_{n+1}))^{X_n} \mid \text{extinction} \right\}
\]

\[
= \mathbb{E} \left\{ \prod_{i=1}^{n-1} s_i^{X_i} \cdot (s_n \tilde{\phi}(s_{n+1}))^{X_n} \mid \text{extinction} \right\}
\]

\[
\equiv \phi_{\bar{X}_1, \ldots, \bar{X}_n}(s_1, \ldots, s_{n-1}, s_n \tilde{\phi}(s_{n+1}))
\]

\[
** \equiv \tilde{\phi}(s_1 \tilde{\phi}(s_2 \tilde{\phi}(\ldots s_{n-1} \tilde{\phi}(s_n \tilde{\phi}(s_{n+1}) \ldots))))),
\]
which confirms (2.6) for the case \( n + 1 \). The equality (*) follows because conditional on ultimate extinction, \( \xi_1^{(n)}, \ldots, \xi_{X_n}^{(n)} \) are iid rvs with pgf \( \tilde{\phi} \) rather than \( \phi \), since each individual in generation \( n \) must give rise to a family line that becomes extinct. The equality (**) follows from the induction hypothesis (2.6). \( \square \)

From (2.1), the 2nd iterate of \( \tilde{\phi} \) is
\[
\tilde{\phi}_2(s) = \frac{1}{u} \phi(u \frac{1}{u} \phi(us)) = \frac{1}{u} \phi(\phi(us)) = \frac{1}{u} \phi_2(us),
\]
so by induction, the \( n \)th iterate of \( \tilde{\phi} \) and its first derivative are
\[
\tilde{\phi}_n(s) = \frac{1}{u} \phi_n(us),
\]
\[
\tilde{\phi}_n'(s) = \phi_n'(us).
\]
Using the relation \( \phi(u) = u \) it can be shown that \( \phi_n'(u) = [\phi'(u)]^n \), so
\[
E(\tilde{X}_n) = \tilde{\phi}_n(1) = \phi_n'(u) = [\phi'(u)]^n \equiv \tilde{\mu}^n,
\]
where \( \tilde{\mu} = \phi'(u) = E(\tilde{X}_1) \). Note that \( \tilde{\mu} < 1 < \mu \) (see Figure 2).

Next consider \( \tilde{\phi} \). Note that \( \tilde{\phi}(0) = 0 \) (since \( \phi(u) = u \)) and \( \tilde{\phi}(1) = 1 \), both being properties of an explosive pgf. That \( \tilde{\phi} \) is in fact an explosive pgf will be shown in Proposition 2.2, and by analogy with \( \tilde{\phi} \) it is tempting to think that \( \tilde{\phi} \) is the generating pgf for the conditional process \( \mathcal{X} \mid \text{explosion} \). However, this is not true: unlike the case of extinction, where the family line of every individual dies out, in the case of explosion only some family lines survive forever while the rest die out. This cannot occur in an explosive GW process, where no deaths without reproduction are possible.\(^1\)

Instead, \( \tilde{\phi} \) is the pgf of the conditional partial process
\[
\tilde{\mathcal{X}} \equiv \{ \tilde{X}_n \mid n \geq 0 \} \mid \text{explosion},
\]
where \( \tilde{X}_n \) is the number of individuals at time \( n \) whose family lines never end. Unconditionally, the events \( \{ \tilde{X}_0 = 0 \} \equiv \{ \text{extinction} \} \) and \( \{ \tilde{X}_0 = 1 \} \equiv \{ \text{explosion} \} \) occur with probabilities \( u \) and \( 1 - u \) respectively, and

\(^1\) In fact the conditional process \( \tilde{\mathcal{X}} \equiv \mathcal{X} \mid \text{explosion} \) is not a GW process; see (7) in Section 3.
(a) \( \check{X}_n \leq X_n \) for \( n = 0, 1, 2, \ldots \);
(b) \( \check{X}_0 = 0 \Rightarrow \check{X}_1 = \check{X}_2 = \cdots = 0 \);
(c) \( \check{X}_0 = 1 \Rightarrow 1 \leq \check{X}_1 \leq \check{X}_2 \leq \cdots \).

Following Karlin [7] Problem 27, p.335 and Athreya and Ney [1] §I.12, we now show that the behavior of the conditional process \( \check{X} \) is the same as the behavior of an explosive GW process with generating pgf \( \check{\phi} \). When considering \( \check{X} \) hereafter, the condition of explosion will be assumed unless explicitly noted to the contrary.

**Proposition 2.2.** For a semiexplosive GW process \( X \) with pgf \( \phi \), the distribution of the partial process \( \check{X} \) is the same as the distribution of an explosive GW process with generating pgf \( \check{\phi} \).

**Proof.** Because \( X \) is a GW process, the partial process \( \check{X} \) is also a GW process: at each time \( n \) we may simply ignore those individuals whose family lines ultimately die out, because neither they nor their descendants contribute to subsequent generations \( \check{X}_{n+1}, \check{X}_{n+2}, \ldots \). It remains to show that the generating pgf for \( \check{X} \) is given by \( \check{\phi} \) in (2.2).

For this it suffices to show that

\[
E[s^{\check{X}_1} | \check{X}_0 = 1] = \frac{\phi((1 - u)s + u) - u}{1 - u} \equiv \check{\phi}(s).
\]

(2.12)

First note that \( \check{p}_0 \equiv \Pr[\check{X}_1 = 0] = 0 \), while for \( k \geq 1 \),

\[
\check{p}_k = \Pr[\check{X}_1 = k]
= \frac{1}{1 - u} \sum_{l \geq k} \Pr^*[X_1 = l] \Pr^*[\check{X}_1 = k | X_1 = l] \quad \text{[by (a)]}
= \frac{1}{1 - u} \sum_{l \geq k} p_l \Pr^*[k \text{ of } l \text{ live forever, } l - k \text{ go extinct} | X_1 = l]
\]

(2.13)

\[
= \frac{1}{1 - u} \sum_{l \geq k} p_l \binom{l}{k} (1 - u)^k u^{l - k},
\]

where \( \Pr^* \) denotes unconditional probability and \( p_l = \Pr^*[X_1 = l] \). Thus
\[ E[s^{\tilde{X}_1} | \tilde{X}_0 = 1] \equiv \sum_{k \geq 0} \tilde{p}_k s^k \]
\[ = \frac{1}{1-u} \sum_{k \geq 1} \left[ \sum_{l \geq k} p_l \binom{l}{k} (1-u)^k u^{l-k} \right] s^k \]
\[ = \frac{1}{1-u} \sum_{l \geq 1} p_l \sum_{k=1}^{l} \binom{l}{k} [(1-u)s]^k u^{l-k} \]
\[ = \frac{1}{1-u} \sum_{l \geq 1} p_l \left[ \sum_{k=0}^{l} \binom{l}{k} [(1-u)s]^k u^{l-k} - u^l \right] \]
\[ = \frac{1}{1-u} \sum_{l \geq 1} p_l \left[ ((1-u)s + u)^l - u^l \right] \]
\[ = \frac{\phi((1-u)s + u) - \phi(u)}{1-u}, \]

which confirms (2.12) since \( \phi(u) = u \).

By Proposition 1.1(i) the distribution of \( \tilde{X}_n \) is determined by the \( n \)th iterate \( \tilde{\phi}_n \) of \( \tilde{\phi} \). The 2nd iterate is
\[
\tilde{\phi}_2(s) \equiv \tilde{\phi}(\tilde{\phi}(s)) = \frac{\phi \left\{(1-u) \left[ \frac{\phi((1-u)s + u) - u}{1-u} \right] + u \right\} - u}{1-u}
\]
\[ = \frac{\phi(\phi((1-u)s + u)) - u}{1-u}
\]
\[ = \frac{\phi_2((1-u)s + u) - u}{1-u}. \]

By induction, the \( n \)th iterate of \( \tilde{\phi} \) is
\[ \tilde{\phi}_n(s) = \frac{\phi_n((1-u)s + u) - u}{1-u}, \]
(cf. Karlin [7] Problem 27 p.335), from which it follows that
\[ \mathbb{E}(\tilde{X}_n) = \tilde{\phi}'_n(1) = \phi'_n(1) = \mathbb{E}(X_n) = \mu^n. \]
Next consider the conditional process

(2.16) \[ \tilde{\mathcal{X}} \equiv \mathcal{X} \mid \text{explosion}. \]

Unlike \( \tilde{\mathcal{X}} \equiv \mathcal{X} \mid \text{extinction} \), \( \tilde{\mathcal{X}} \) is not a GW process (see (7) in Section 3), but the pgf \( \tilde{\phi}_{[n]} \) of \( \tilde{X}_n \) and thence its moments can be obtained from the following mixture representation for \( \mathcal{X} \):

(2.17) \[ \mathcal{X} = \begin{cases} \tilde{\mathcal{X}} & \text{with probability } u, \\ \hat{\mathcal{X}} & \text{with probability } 1 - u. \end{cases} \]

This implies that

\[
\phi_n(s) \equiv \mathbb{E}(s^{X_n}) = \mathbb{E}(s^{X_n} \mid \text{extinction}) \Pr[\text{extinction}] \\
+ \mathbb{E}(s^{X_n} \mid \text{explosion}) \Pr[\text{explosion}]
\]

\[
= \mathbb{E}(s^{\tilde{X}_n}) u + \mathbb{E}(s^{\hat{X}_n}) (1 - u)
\]

(2.18)

\[
= \tilde{\phi}_n(s) u + \hat{\phi}_{[n]}(s) (1 - u),
\]

hence

(2.19) \[ \hat{\phi}_{[n]}(s) = \frac{\phi_n(s) - \tilde{\phi}_n(s) u}{1 - u} = \frac{\phi_n(s) - \phi_n(us)}{1 - u} \]

by (2.8). Similarly by (2.10) and (2.15),

(2.20) \[ \mathbb{E}(\tilde{X}_n) = \frac{\mathbb{E}(X_n) - \mathbb{E}(\hat{X}_n) u}{1 - u} = \frac{\mu^n - \tilde{\mu}_n u}{1 - u}. \]

Finally, we might wish consider the conditional process

(2.21) \[ \check{\mathcal{X}} \equiv \mathcal{X} \mid \text{no individual dies without offspring}. \]

Because \( p_0 > 0 \) for the underlying process \( \mathcal{X} \), however, the event that no individual dies without offspring is null, so strictly speaking cannot be considered as a conditioning event. Instead we define the process \( \check{\mathcal{X}} \) to be the modified GW process with offspring distribution \( (\check{p}_0, \check{p}_1, \check{p}_2, \ldots) \), where \( \check{p}_0 = 0 \) and

(2.22) \[ \check{p}_k = \frac{p_k}{1 - p_0} \quad \text{for } k \geq 1. \]
That is, the modified offspring rv $\tilde{\xi}$ has the conditional distribution of the original $\xi$ given that $\xi \geq 1$, i.e., death without offspring is not allowed. Thus $\tilde{\mathcal{X}}$ is an explosive GW process with pgf

$$
(2.23) \quad \tilde{\phi}(s) = \frac{\phi(s) - p_0}{1 - p_0}.
$$

This implies that

$$
(2.24) \quad \mathbb{E}(\tilde{\mathcal{X}}_n) = \tilde{\phi}'(1) = [\phi'(1)]^n = \frac{[\phi'(1)]^n}{(1 - p_0)^n} \equiv \frac{\mu^n}{(1 - p_0)^n},
$$

which differs from $\mathbb{E}(\tilde{\mathcal{X}}_n)$ in (2.20). That $\tilde{\mathcal{X}}$ behaves differently than $\mathcal{X}$ is to be expected because, as already mentioned, some individuals in the original process may die without offspring even though explosion occurs.

3. Stochastic comparisons among $\mathcal{X}$, $\tilde{\mathcal{X}}$, $\tilde{\mathcal{X}}$, $\check{\mathcal{X}}$, $\check{\mathcal{X}}$

A nonnegative integer-valued rv $Y$ is stochastically smaller than a nonnegative rv $Z$, denoted by $Y \preceq Z$, if

$$
(3.1) \quad \Pr[Y \geq x] \leq \Pr[Z \geq x] \quad \text{for all } x = 0, 1, \ldots.
$$

If strict inequality holds for at least one $x$ then $Y$ is strictly stochastically smaller than $Z$, denoted by $Y < Z$.

Equivalently, $Y \preceq Z$ iff $\mathbb{E}[g(Y)] \leq \mathbb{E}[g(Z)]$ for all nondecreasing integrable functions $g$ on $\mathbb{R}^1$, and $Y < Z$ iff $Y \preceq Z$ and strict inequality holds for at least one $g$. Clearly

$$
Y \preceq (\prec) Z \Rightarrow \mathbb{E}(Y) \leq (\prec) \mathbb{E}(Z)
$$

(provided the expectations are finite). It is also straightforward to show that if $U, V, Y, Z$ are independent, then

$$
(3.2) \quad U \preceq V \text{ and } Y \preceq (\prec) Z \Rightarrow U + Y \preceq (\prec) V + Z.
$$

A random vector $\mathbf{Y}_n \equiv (Y_1, \ldots, Y_n)$ is stochastically smaller than a random vector $\mathbf{Z}_n \equiv (Z_1, \ldots, Z_n)$, denoted by $\mathbf{Y}_n \preceq \mathbf{Z}_n$, if

$$
(3.3) \quad \mathbb{E}[g(\mathbf{Y}_n)] \leq \mathbb{E}[g(\mathbf{Z}_n)]
$$
for all nondecreasing integrable functions \( g \) on \( \mathbb{R}^n \): \( Y_n \) is \textit{strictly stochastically smaller} than \( Z_n \), denoted by \( Y_n \prec Z_n \), if \( Y_n \preceq Z_n \) and strict inequality holds in (3.3) for at least one \( g \).

A stochastic process \( \mathcal{Y} \equiv \{Y_n\} \) is (\textit{strictly}) \textit{stochastically smaller} than a stochastic process \( \mathcal{Z} \equiv \{Z_n\} \), denoted by \( \mathcal{Y} (\prec) \preceq \mathcal{Z} \), if \( Y_n (\prec) \preceq Z_n \) for all \( n \geq 1 \), where \( Y_n = (Y_1, \ldots, Y_n) \) and \( Z_n = (Z_1, \ldots, Z_n) \).

Now return to our explosive GW process \( \mathcal{X} \). From the definitions of the four conditional processes \( \tilde{X}, \tilde{X}, \hat{X}, \), and \( \tilde{X}, \) it might be conjectured that
\[
\tilde{X} \prec (\mathcal{X}, \tilde{X}) \prec (\hat{X}, \tilde{X}) \text{ and } \hat{X} \prec \tilde{X}.
\]

We shall investigate these seven possible stochastic orderings.

To begin, the following comparisons among the expected values are obtained from (2.10), (2.15), (2.20), and (2.24):

\[
(3.4) \quad E(\tilde{X}_n) < E(X_n) = E(\tilde{X}_n) \left\{ \begin{array}{c} < E(\tilde{X}_n) \\ < E(\tilde{X}_n) \end{array} \right.,
\]

\[
(3.5) \quad E(\tilde{X}_n) < E(\hat{X}_n) \text{ for sufficiently large } n,
\]

in particular for \( n > \frac{\log(1-u)}{\log(1-p_0)} \). This suggests that the following six stochastic orderings may hold among the five processes:

\[
(3.6) \quad \tilde{X} \left\{ \begin{array}{c} (1) \prec \mathcal{X} \prec \{ (2) \hat{X} \} \\ (3) \prec \{ (4) \tilde{X} \} \prec \hat{X} \end{array} \right.,
\]

and possibly a seventh for sufficiently large \( n \):

\[
(3.7) \quad \hat{X} \prec (7) \tilde{X}.
\]

We now examine the seven possible stochastic orderings (1) – (7).

\textbf{Lemma 3.1.} Let \( \mathcal{Y} \) and \( \mathcal{Z} \) be GW branching processes with offspring distributions \( \eta \) and \( \zeta \) respectively. Then \( \eta \prec \zeta \Rightarrow \mathcal{Y} \prec \mathcal{Z} \).
Proof. This follows directly from the fundamental reproductive property (1.1) of a GW process, (3.2), and induction. □

(1) $\tilde{\mathcal{X}} \prec \mathcal{X}$ is valid:
This result is expected since $\tilde{X}_n$ is defined to be $X_n$ conditioned on extinction, that is, on $X_n \to 0$. The processes $\tilde{\mathcal{X}}$ and $\mathcal{X}$ are GW processes with offspring distributions $\xi \sim (\tilde{p}_0, \tilde{p}_1, \ldots)$ and $\xi \sim (p_0, p_1, \ldots)$ respectively, so by Lemma 3.1 it suffices to show that $\tilde{\xi} \prec \tilde{\xi}$. But by (2.5) the pair $(\tilde{\xi}, \xi)$ has a strictly increasing likelihood ratio, that is, $p_k / \tilde{p}_k = 1 / u^{k-1}$ is strictly increasing in $k$ for those $k$ s.t. $p_k > 0$, since $0 < u < 1$. Therefore $\tilde{\xi} \prec \xi$ by a standard result of monotone likelihood ratio (MLR) theory (cf. Lehmann and Romano [8] Problem 3.39(iii)). □

(2) $\mathcal{X} \prec \tilde{\mathcal{X}}$ is valid:
This result is also expected since $\tilde{X}_n$ is defined to be $X_n$ conditioned on explosion, that is, on $X_n \to \infty$. We might (incorrectly!) argue as in (1) as follows: $\mathcal{X}$ and $\tilde{\mathcal{X}}$ are GW processes with offspring distributions $\xi \sim (p_0, p_1, \ldots)$ and $\hat{\xi} \sim (\hat{p}_0, \hat{p}_1, \ldots)$ respectively, so by Lemma 3.1 it suffices to show that $\xi \prec \hat{\xi}$. But $\hat{p}_0 = \text{Pr}[\hat{X}_1 = 0] = 0$, while for $k \geq 1$, (2.19) with $n = 1$ implies that

$$\hat{p}_k = p_k \left( \frac{1 - u^k}{1 - u} \right).$$

Thus, here too the LR $\hat{p}_k / p_k$ is strictly increasing in $k$ for those $k$ s.t. $p_k > 0$, hence as with (1), $\xi \prec \hat{\xi}$ by MLR theory.

However, this argument is incorrect, or at least incomplete, because it is not obvious that $\tilde{\mathcal{X}} \equiv \mathcal{X} | \text{explosion}$ is a GW process: even though the population grows indefinitely the family lines of some individuals may become extinct, so these individuals would not reproduce according to the distribution of $\hat{\xi}$. Therefore we cannot appeal to the fundamental reproductive property (1.1).

Instead, (2) can be established directly by means of the mixture representation (2.17). For any $n \geq 1$ and any nondecreasing integrable function $g$ on $\mathbb{R}^n$,

$$E[g(X_n)] = E[g(\tilde{X}_n)] u + E[g(\tilde{X}_n)] (1 - u),$$

(3.9)
where \( X_n = (X_1, \ldots, X_n), \tilde{X}_n = (\tilde{X}_1, \ldots, \tilde{X}_n), \) and \( \hat{X}_n = (\hat{X}_1, \ldots, \hat{X}_n). \)

Because \( \mathcal{X} \prec \mathcal{X} \) by (1), \( E[g(\tilde{X}_n)] \leq E[g(\hat{X}_n)] \) with strict inequality for at least one \( g \). Thus, since \( 0 < u < 1 \), \( E[g(X_n)] \leq E[g(\hat{X}_n)] \) by (3.9), with strict inequality for at least one \( g \). Therefore \( X_n \prec \hat{X}_n \) for all \( n \geq 1 \), hence \( \mathcal{X} \prec \mathcal{X} \).

(3) \( \mathcal{X} \prec \tilde{\mathcal{X}} \) is valid:

The processes \( \mathcal{X} \) and \( \tilde{\mathcal{X}} \) are GW processes with offspring distributions \( \xi \sim (p_0, p_1, \ldots) \) and \( \tilde{\xi} \sim (\tilde{p}_0, \tilde{p}_1, \ldots) \) respectively. Thus by Lemma 3.1 it suffices to show that \( \xi \prec \tilde{\xi} \). But this follows directly from (2.22) since \( p_0 > 0 \).

(4) \( \tilde{\mathcal{X}} \prec \tilde{\mathcal{X}} \) holds for \( u \leq \frac{1}{2} \) but not necessarily for \( u > \frac{1}{2} \):

Because \( \tilde{\xi} \overset{d}{=} \tilde{X}_1 \) and \( \xi \overset{d}{=} \hat{X}_1 \), it follows from (2.5) with \( k = 0 \) that

\[
(3.10) \quad Pr[\xi \geq 1] = 1 - \frac{p_0}{u} < 1 = Pr[\tilde{\xi} \geq 1]
\]

for all values of \( u \). Also, it follows from (2.13) and (2.5) that for \( r \geq 2 \),

\[
Pr[\xi \geq r] = \sum_{k \geq r} \tilde{p}_k
\]

\[
= \frac{1}{1 - u} \sum_{k \geq r} \sum_{l \geq k} p_l \binom{l}{k} (1 - u)^k u^{l-k}
\]

\[
= \frac{1}{1 - u} \sum_{l \geq r} p_l u^l \sum_{k = r}^l \binom{l}{k} \left( \frac{1 - u}{u} \right)^k
\]

\[
= \frac{1}{1 - u} \left( \frac{1 - u}{u} \right)^r \sum_{l \geq r} p_l u^l \sum_{m = 0}^{l-r} \binom{l}{m+r} \left( \frac{1 - u}{u} \right)^m
\]

\[
\geq \frac{1}{u} \sum_{l \geq r} p_l u^l = \sum_{l \geq r} \tilde{p}_l = Pr[\tilde{\xi} \geq r],
\]

(3.11)

where the inequality (3.11) holds provided that \( u \leq \frac{1}{2} \). Thus, in this case \( \xi \prec \tilde{\xi} \) and therefore \( \mathcal{X} \prec \tilde{\mathcal{X}} \) by Lemma 3.1, since both \( \mathcal{X} \) and \( \tilde{\mathcal{X}} \) are GW processes.
In the geometric case (Example 4.1), in fact $\tilde{X} \prec \tilde{X}$ for all $n \geq 1$ when $u \leq \frac{\sqrt{5} - 1}{2} \approx 0.618$, and $\tilde{X}_n \prec \tilde{X}_n$ for sufficiently large $n$ (depending on $u$) when $u > \frac{\sqrt{5} - 1}{2}$. For a quadratic pgf (Example 3.1), however, $\tilde{X} \not\prec \tilde{X}$ for all $n \geq 1$ when $u > \frac{1}{2}$, so (4) fails in this case. This is somewhat surprising in view of the facts that in general:

- the conditioning events for $\tilde{X}$ and $\tilde{X}$ are $\{X_n \to 0\}$ and $\{X_n \to \infty\}$ respectively;
- $\tilde{X}_n \to 0$ (because $\tilde{X}$ is a subcritical GW process) while $\tilde{X}_n \to \infty$ (because $\tilde{X}$ is an explosive GW process);
- $\mathbb{E}(\tilde{X}_n) = \tilde{\mu}^n \to 0$ and $\mathbb{E}(\tilde{X}_n) = \mu^n \to \infty$ (since $\tilde{\mu} < 1 < \mu$).

(5) $\tilde{X} \prec \tilde{X}$ is valid:
Clearly $\tilde{X} \prec \tilde{X}$ because $\tilde{X}_n \leq X_n = \tilde{X}_n$ on the common conditioning event $\{\text{extinction}\}$ for $\tilde{X}$ and $\tilde{X}$. But for each $n \geq 1$,

$$\Pr[\tilde{X}_n < \tilde{X}_n] = \mathbb{E}\{\Pr[\tilde{X}_n < \tilde{X}_n | \tilde{X}_n]\}$$

$$= \mathbb{E}\{1 - \Pr[\tilde{X}_n = \tilde{X}_n | \tilde{X}_n]\}$$

$$= \mathbb{E}\{1 - (1 - u)\tilde{X}_n\} > 0$$

(3.12)

because $u > 0$ and $\tilde{X}_n \geq 1$, hence $\tilde{X}_n < \tilde{X}_n$, so $\tilde{X} \prec \tilde{X}$.

(6) $\tilde{X} \prec \tilde{X}$ does not hold in general:
Examples 3.1, 4.1, and 4.2 show that $\tilde{X} \prec \tilde{X}$ in the cases of quadratic, geometric, and Poisson offspring distributions. The cubic Example 3.2 shows, however, that $\tilde{X} \prec \tilde{X}$ does not hold in general, despite the fact that

- $\mathbb{E}(\tilde{X}_n) \ll \mathbb{E}(\tilde{X}_n)$ for large $n$ by (2.15) and (2.24).

(7) $\tilde{X} \prec \tilde{X}$ does not hold in general:
In the geometric case (Example 4.1), $\tilde{X}_n \prec \tilde{X}_n$ for sufficiently large $n$. For a quadratic pgf (Example 3.1), however, $\tilde{X}_n \not\prec \tilde{X}_n$ for all $n$, so (7) does not hold in general. This may be somewhat surprising in view of the facts that

- $\mathbb{E}(\tilde{X}_n) \ll \mathbb{E}(\tilde{X}_n)$ for large $n$ (by (2.20) and (2.24));
- no deaths are allowed for the process $\tilde{X}$;
• $\xi < \tilde{\xi}$ always,\(^2\) seen as follows: Clearly $\tilde{p}_0 = \tilde{\rho}_0 = 0$, while for $k \geq 1$, (2.22) and (3.8) combine to yield

$$\frac{\tilde{p}_k}{\bar{p}_k} = \left( \frac{1 - u^k}{1 - u} \right) (1 - p_0),$$

which is strictly increasing in $k$. Thus $(\tilde{\xi}, \tilde{\xi})$ satisfies the strict MLR property, so $\xi < \tilde{\xi}$. This confirms the fact that $\tilde{\mathcal{X}}$ does not evolve according to a GW branching process, for otherwise Lemma 3.1 would imply that $\mathcal{X} \prec \tilde{\mathcal{X}}$, contradicting the fact that $E(\tilde{X}_n) < E(\tilde{X}_n)$ for large $n$. \hfill \square

Here is a summary of these results:

(3.13) \[
\tilde{\mathcal{X}} \begin{cases} \prec \mathcal{X} \\ (u \leq \frac{1}{2}) \prec \hat{\mathcal{X}} \end{cases} \begin{cases} \prec \tilde{\mathcal{X}} \\ \prec \tilde{\hat{\mathcal{X}}} \end{cases},
\]

(3.14) \[
\tilde{X}_1 \prec \tilde{X}_1, \quad \tilde{X}_n \prec \tilde{X}_n \text{ for } n \geq 2,
\]

where "(-)" indicates that no general result holds.

**Remark 3.1.** Because $\tilde{X}_1 \overset{d}{=} \tilde{\xi} \prec \xi \overset{d}{=} \hat{X}_1$, it follows from (2.24) and (2.20) with $n = 1$ that

$$\frac{\mu}{1 - p_0} = E(\tilde{X}_1) \prec E(\hat{X}_1) = \frac{\mu - \tilde{\mu} u}{1 - u},$$

from which a lower bound for the extinction probability $u$ is obtained:

(3.15) \[
u > \frac{\mu p_0}{\mu - \tilde{\mu} (1 - p_0)} \]

**Example 3.1 (The general quadratic pgf).**

\(^2\) Unless $p_k > 0$ for only one $k \geq 1$; see Example 3.1.
Suppose that each individual is replaced by either 0, 1, or 2 new individuals. Then \( \phi(s) = \phi_{p_0,p_2}(s) = p_0 + p_1 s + p_2 s^2 \), where \( p_0 + p_1 + p_2 = 1 \) and \( p_0, p_2 > 0 \). Thus \( \mu = \phi'(1) = 1 + p_2 - p_0 \), so

\[
\mu \begin{cases} 
\leq 1 & \text{if } p_0 \geq p_2 \quad \text{(subexplosive);} \\
> 1 & \text{if } 0 < p_0 < p_2 \quad \text{(semiexplosive);} \\
> 1 & \text{if } 0 = p_0 < p_2 \quad \text{(explosive).}
\end{cases}
\]  

(3.16)

Assume that the process is semiexplosive. The extinction probability is \( u = p_0 / p_2 \), the unique solution in \((0, 1)\) of the quadratic equation \( \phi_{p_0,p_2}(s) = s \). From (2.1), (2.2), (2.19) \((n = 1)\), and (2.23) it follows that

\[
\tilde{\phi}_{p_0,p_2}(s) = p_2 + (1 - p_0 - p_2) s + p_0 s^2 \equiv \phi_{p_2,p_0}(s), 
\]  

(3.17)

\[
\phi_{p_0,p_2}(s) = (1 + p_0 - p_2) s + (p_2 - p_0) s^2 \equiv \phi_{0,p_2-p_0}(s), 
\]  

(3.18)

\[
\phi_{p_0,p_2;1}(s) = (1 - p_0 - p_2) s + (p_0 + p_2) s^2 \equiv \phi_{0,p_0+p_2}(s), 
\]  

(3.19)

\[
\phi_{p_0,p_2}(s) = \frac{1 - p_0 - p_2}{1 - p_0} s + \frac{p_2}{1 - p_0} s^2 \equiv \phi_{0,p_2/(1-p_0)}(s). 
\]  

(3.20)

Each of these are quadratic pgfs, the first subcritical and the rest explosive. Since \( p_0 < p_2 < \frac{p_2}{1-p_0} \leq p_0 + p_2 \) and \( p_2 - p_0 < \frac{p_2}{1-p_0} \),

\[
\tilde{\xi} < \xi < \tilde{\xi} \preceq \xi, 
\]  

(3.21)

\[
\tilde{\xi} < \xi, 
\]  

(3.22)

the last inequality in (3.21) being strict if \( p_0 + p_2 < 1 \). (If \( p_0 + p_2 = 1 \) then \( \tilde{\xi} = \xi = 2 \) with probability 1).

It was shown above that the orderings (1), (2), (3), and (5) hold in general. We now examine (4), (6), and (7) for this quadratic example:

(4) \( \tilde{\mathcal{X}} < \mathcal{X} \) when \( u \leq \frac{1}{2} \), i.e., \( 2p_0 \leq p_2 \), but not when \( u > \frac{1}{2} \), i.e., \( 2p_0 > p_2 \).

From (3.10),

\[
\Pr[\tilde{\xi} \geq 1] < \Pr[\xi \geq 1] 
\]

for all \( u \), while from (3.17) and (3.18),

\[
\Pr[\tilde{\xi} \geq 2] = \Pr[\xi \geq 2] = p_0 \leq p_2 - p_0 = \Pr[\tilde{\xi} = 2] = \Pr[\xi \geq 2] 
\]

16
when \( u \equiv \frac{p_0}{p_2} \leq \frac{1}{2} \). Thus \( \tilde{\xi} \prec \xi \) when \( u \leq \frac{1}{2} \), so (4) holds in this case by Lemma 3.1. When \( u > \frac{1}{2} \), because \( \tilde{X}_n \leq 2^n \) and \( \check{X}_n \leq 2^n \) we see that

\[
\Pr[\check{X}_n \geq 2^n] = \Pr[\check{X}_n = 2^n] = p_0^n > (p_2 - p_0)^n = \Pr[\tilde{X}_n = 2^n] = \Pr[\tilde{X}_n \geq 2^n]
\]

for all \( n \geq 1 \). Therefore \( \check{X}_n \nRightarrow \tilde{X}_n \), hence \( \check{X} \nRightarrow \tilde{X} \) when \( u > \frac{1}{2} \).

(6) \( \check{X} \prec \tilde{X} \):

Both \( \check{X} \) and \( \tilde{X} \) are GW processes, so (6) follows from (3.22) by Lemma 3.1.

(7) \( \check{X} \nRightarrow \tilde{X} \):

Here too \( \check{X}_n \leq 2^n \) and \( \tilde{X}_n \leq 2^n \), so

\[
\Pr[\check{X}_n \geq 2^n] = \Pr[\check{X}_n = 2^n] = \frac{\Pr[X_n = 2^n]}{1 - p_0} = \frac{p_2^n}{1 - p_0}.
\]

Similarly from the mixture representation (2.17),

\[
\Pr[\check{X}_n \geq 2^n] = \Pr[\check{X}_n = 2^n] = \frac{\Pr[X_n = 2^n] - \Pr[\tilde{X}_n = 2^n]}{1 - u} u = \frac{p_2^n - p_0^n (\frac{p_0}{p_2})}{1 - (\frac{p_0}{p_2})}.
\]

Therefore for \( n \geq 1 \),

\[
\Pr[\check{X}_n \geq 2^n] - \Pr[\check{X}_n \geq 2^n] = \frac{p_2^n - p_0^n (\frac{p_0}{p_2})}{1 - (\frac{p_0}{p_2})} - \frac{p_2^n}{1 - p_0} = \frac{p_0 [p_2^n (1 - p_2) - p_0^n (1 - p_0)]}{(p_2 - p_0)(1 - p_0)} = \frac{p_0 [p_2^{n-1} p_2 (1 - p_2) - p_0^{n-1} p_0 (1 - p_0)]}{(p_2 - p_0)(1 - p_0)} > 0
\]

(3.24)

since \( p_2 > p_0 \) and \( p_2 (1 - p_2) \geq p_0 (1 - p_0) \). Thus \( \check{X}_n \nRightarrow \check{X}_n \), so \( \check{X} \nRightarrow \check{X} \).
Summarizing: For a general quadratic pgf $\phi_{p_0,p_2}$,

$$
\tilde{X}' \begin{cases} 
\prec \quad \mathcal{X} & \begin{cases} 
\prec \quad \hat{\mathcal{X}} \\
\prec \quad \hat{\mathcal{X}} 
\end{cases} \\
(2p_0 \leq p_2) 
\prec \quad \hat{\mathcal{X}} & \begin{cases} 
\prec \quad \hat{\mathcal{X}} 
\end{cases} 
\end{cases},
$$

(3.25)

$$
\tilde{X}_1 \prec \hat{X}_1 \text{ if } p_0 + p_2 < 1, \quad \hat{X}_n \nprec \hat{X}_n \text{ for } n \geq 2. \quad \square
$$

Example 3.2 (A cubic pgf).
Consider the GW process with cubic pgf

$$
\phi(s) = \delta + (1 - 3\delta)s + 2\delta s^3,
$$

(3.27)

where $0 < \delta < \frac{1}{3}$. Then $p_0 = \delta > 0$ and $\mu = \phi''(1) = 1 + 3\delta > 1$ so the process is semiexplosive. The extinction probability $u$ satisfies the cubic equation $\phi(s) = s$, equivalently

$$
2s^3 - 3s + 1 = 0,
$$

(3.28)

which has the unique solution $u \approx 0.36602$ in $(0, 1)$. From (2.13),

$$
\tilde{p}_1 = \frac{1}{1 - u} \sum_{l \geq 1} p_l \binom{l}{1} (1 - u) u^{l-1}
= (1 - 3\delta) + 2\delta \cdot 3u^2,
$$

while from (4.22), $\tilde{p}_1 = \frac{1 - 3\delta}{1 - \delta}$. Therefore

$$
\tilde{p}_1 < \tilde{p}_1 \iff (1 - 3\delta) + 2\delta \cdot 3u^2 < \frac{1 - 3\delta}{1 - \delta}
\iff 6u^2 < \frac{1 - 3\delta}{1 - \delta}
\iff 0.8038... < \frac{1 - 3\delta}{1 - \delta}
\iff \delta < 0.0893....
$$
Because $\tilde{p}_0 = \tilde{p}_0 = 0$, this implies that for all $n \geq 1$,

$$\Pr[\tilde{X}_n \leq 1] = \Pr[\tilde{X}_n = 1] = (\tilde{p}_1)^n < (\tilde{p}_1)^n = \Pr[\check{X}_n = 1] = \Pr[\check{X}_n \leq 1]$$

whenever $\delta < 0.0893\ldots$, hence $\check{X}_n \not\preceq \check{X}_n$, so (6): $\check{X} < \check{X}$ fails in this case. □

4. Geometric and Poisson offspring distributions

In the geometric case, explicit stochastic representations can be obtained for $\check{X}$, $\check{X}$, $\check{X}$, $\check{X}$, and $\check{X}$ that allow direct derivations and/or refinements of some of the stochastic inequalities (1) - (7). This also holds to a limited extent for the Poisson case when $n = 1$.


Suppose that $\xi \overset{d}{=} \text{geometric}(p)$, that is, $p_k = pq^k$, $k = 0, 1, 2, \ldots$, $0 < p < 1$. Here the generating pgf is

$$(4.1) \quad \phi(s) \equiv \phi_p(s) = E(s^\xi) = \sum_{k=0}^{\infty} s^k(pq^k) = \frac{p}{1-qs}$$

and $\mu \equiv E(\xi) = \phi'_p(1) = q/p$, so the process is

$$(4.2) \quad \begin{cases} \text{subexplosive} & \text{if } q \leq p \iff p \geq 1/2; \\ \text{semiexplosive} & \text{if } q > p \iff p < 1/2. \end{cases}$$

For the remainder of this example assume that the process is semiexplosive. The extinction probability is $u = p/q$, the unique solution in $(0, 1)$ to the quadratic equation $\phi(s) = s$. From (2.1), (2.2), (2.19) ($n = 1$), and (2.23), straightforward algebra yields

$$(4.3) \quad \tilde{\phi}_p(s) = \frac{q}{1-ps} \equiv \phi_q(s),$$

$$(4.4) \quad \check{\phi}_p(s) = \frac{us}{1-(1-u)s} \equiv s \cdot \phi_u(s),$$

$$(4.5) \quad \hat{\phi}_{p;[1]}(s) = \frac{pq/s}{(1-qs)(1-ps)} \equiv s \cdot \phi_p(s) \cdot \phi_q(s),$$

$$(4.6) \quad \check{\phi}_p(s) = \frac{ps}{1-qs} \equiv s \cdot \phi_p(s).$$
The first of these is a subcritical geometric pgf because \( q > 1/2 \), while the rest are explosive pgfs. Because the pgf of the sum of independent rvs is the product of their pgfs, (4.3) - (4.6) imply that

\[
\begin{align*}
(4.7) & \quad \bar{\xi} \overset{d}{=} \bar{X}_1 \overset{d}{=} \text{geometric}(q), \\
(4.8) & \quad \bar{\xi} \overset{d}{=} \bar{X}_1 \overset{d}{=} 1 + \text{geometric}(u), \\
(4.9) & \quad \bar{\xi} \overset{d}{=} \bar{X}_1 \overset{d}{=} 1 + \text{geometric}(p) + \text{geometric}(q) \\
(4.10) & \quad \bar{\xi} \overset{d}{=} \bar{X}_1 \overset{d}{=} 1 + \text{geometric}(p),
\end{align*}
\]

where the two geometric rvs on the right side of (4.9) are independent. From (4.7), \( \bar{\mu} \equiv E(\bar{\xi}) = p/q = u \).

The geometric\((p)\) family has strictly decreasing likelihood ratio, hence is strictly stochastically decreasing in \( p \). Thus, because \( u < \frac{1}{1+u} \equiv q \) iff \( u < \frac{\sqrt{5} - 1}{2} \) \((\approx 0.618)\) and \( u > p \), it follows from (4.7) - (4.10) that

\[
\begin{align*}
(1 + \bar{X}_1) & \overset{d}{<} (1 + \tilde{X}_1) \overset{d}{<} (1 + X_1) \overset{d}{=} \bar{X}_1 < \bar{X}_1 \quad \text{if } u < \left(=\right) \frac{\sqrt{5} - 1}{2}, \\
\bar{X}_1 & \overset{\overset{d}{<}}{<} (1 + \tilde{X}_1) \overset{d}{<} (1 + X_1) \overset{d}{=} \bar{X}_1 < \bar{X}_1 \quad \text{if } u > \frac{\sqrt{5} - 1}{2}.
\end{align*}
\]

For \( n \geq 2 \) we can obtain explicit stochastic representations for \( \bar{X}_n, \tilde{X}_n, \) and \( \bar{X}_n \), as well as indirect representations for \( X_n \) and \( \tilde{X}_n \), that allow more direct derivations and/or refinements of the stochastic orderings (2) - (7). Because \( p = 1/(\mu + 1) \), we can rewrite the geometric\((p)\) pgf \( \phi_p \) as

\[
(4.11) \quad \phi_p(s) = \frac{p}{1 - qs} = \frac{1}{1 + \mu(1 - s)}. 
\]

The second iterate \( \phi_{p;2} \) of \( \phi_p \) is easily seen to be

\[
\phi_{p;2}(s) = \frac{1 + \mu(1 - s)}{1 + (\mu^2 + \mu^2)(1 - s)},
\]

and by induction the \( n \)th iterate \( \phi_{p;n} \) is given by

\[
\phi_{p;n}(s) = \frac{1 + (\mu + \cdots + \mu^{n-1})(1 - s)}{1 + (\mu + \cdots + \mu^n)(1 - s)}.
\]
\[(4.12) \quad \frac{1 + \mu_{n-1}(1 - s)}{1 + \mu_n(1 - s)} \]
\[(4.13) \quad \equiv \frac{\phi_{\pi_n}(s)}{\phi_{\pi_{n-1}}(s)},\]
where
\[(4.14) \quad \mu_n = \mu + \cdots + \mu^n,\]
\[(4.15) \quad \pi_n = \frac{\mu - 1}{\mu^{n+1} - 1}.\]
Therefore \(\phi_{p; n}(s) \cdot \phi_{\pi_{n-1}}(s) = \phi_{\pi_n}(s)\), so we obtain the stochastic representation
\[(4.16) \quad X_n + \text{geometric}(\pi_{n-1}) \overset{d}{=} \text{geometric}(\pi_n),\]
where the two rvs on the left are independent.

Next, from (4.7), simply interchange \(p\) and \(q\) in (4.16) to obtain
\[(4.17) \quad \tilde{X}_n + \text{geometric}(\tilde{\pi}_{n-1}) \overset{d}{=} \text{geometric}(\tilde{\pi}_n),\]
where the two rvs on the left are again independent and where
\[(4.18) \quad \tilde{\pi}_n = \frac{1}{1 + \tilde{\mu}_n} \equiv \frac{1}{1 + \tilde{\mu} + \cdots + \tilde{\mu}^n} = \frac{1 - \tilde{\mu}}{1 - \tilde{\mu}^{n+1}},\]
\[(4.19) \quad = \frac{1 - u}{1 - u^{n+1}} = \frac{1}{1 + u + \cdots + u^n} \equiv \frac{1}{1 + u_n}.\]
(Recall that \(\tilde{\mu} = u = p/q < 1\).) Since \(\tilde{\pi}_n = \mu^n \pi_n\), (4.17) can be rewritten as
\[(4.20) \quad \tilde{X}_n + \text{geometric}(\mu^{n-1} \pi_{n-1}) \overset{d}{=} \text{geometric}(\mu^n \pi_n),\]
(compare to (4.16)).

Next, from (4.4) the second iterate \(\tilde{\phi}_{p; 2}\) of \(\tilde{\phi}_p(s) = s \phi_u(s)\) is
\[\tilde{\phi}_{p; 2}(s) \equiv \tilde{\phi}_p(\tilde{\phi}_p(s)) = s \phi_u(s) \phi_u(s \phi_u(s))\]
\[= \frac{su}{1 - (1 - u)s} \cdot \frac{u}{1 - (1 - u)(\frac{su}{1 - (1 - u)s})}\]
\[= \frac{su^2}{1 - (1 - u^2)s} \equiv s \cdot \phi_{u^2}(s).\]
Similarly by induction, the $n$th iterate $\tilde{\varphi}_{p,n}$ of $\tilde{\varphi}_p(s) \equiv s\varphi_u(s)$ is
\begin{equation}
\tilde{\varphi}_{p,n}(s) = \frac{s u^n}{1 - (1 - u^n)s} \equiv s \cdot \varphi_u^n(s),
\end{equation}
hence
\begin{equation}
\tilde{X}_n \overset{d}{=} 1 + \text{geometric}(u^n).
\end{equation}

Now rewrite (4.12) as
\begin{equation}
\varphi_{p,n}(s) = 1 - \frac{\mu^n(1 - s)}{1 + \mu_n(1 - s)},
\end{equation}
so that by (2.19), the pgf of $\tilde{X}_n$ is
\begin{equation}
\hat{\varphi}_{p,[n]}(s) = \frac{1}{1 - u} \left[ \frac{\mu^n(1 - us)}{1 + \mu_n(1 - us)} - \frac{\mu^n(1 - s)}{1 + \mu_n(1 - s)} \right]
= s \cdot \frac{1}{1 + \mu_n(1 - s)} \cdot \frac{\mu^n}{1 + \mu_n(1 - s)}
= s \cdot \frac{1}{1 + \mu_n(1 - s)} \cdot \frac{\mu^n}{1 + \mu_n}
\equiv s \cdot \varphi_{\pi_n}(s) \cdot \varphi_{\mu^n\pi_n}(s) \quad \text{[apply (4.10)]}
\end{equation}
\begin{equation}
\equiv s \cdot \varphi_{\mu^n\pi_n}(s) \cdot \varphi_{\pi_n}(s).
\end{equation}

Therefore
\begin{equation}
\tilde{X}_n \overset{d}{=} 1 + \text{geometric}(\pi_n) + \text{geometric}(\mu^n\pi_n)
\end{equation}
\begin{equation}
\equiv 1 + \text{geometric}(\tilde{\pi}_n) + \text{geometric}(\tilde{\pi}_n),
\end{equation}
where the two geometric rvs on the right of each equation are independent.

Last, as in (4.21), from (4.6) the $n$th iterate $\tilde{\varphi}_{p,n}$ of $\tilde{\varphi}_p(s) \equiv s\varphi_p(s)$ is
\begin{equation}
\frac{s p^n}{1 - (1 - p^n)s} \equiv s \cdot \varphi_{p^n}(s),
\end{equation}

22
hence

\[(4.31) \quad \tilde{X}_n \overset{d}{=} 1 + \text{geometric}(p^n).\]

The stochastic representations \((4.16), (4.17), (4.20), (4.22), (4.28), (4.29),\) and \((4.31)\) are now applied to strengthen and/or refine the stochastic orderings \((2) - (7)\) in this geometric example, at least for fixed \(n \geq 1.\)

\((2^*) \quad X_n \prec \tilde{X}_n - 1:\)
From \((4.16)\) and \((4.28),\) \(1 + X_n \prec 1 + \text{geometric}(\pi_n) \prec \tilde{X}_n.\)

\((3^*) \quad X_n \prec \tilde{X}_n - 1:\)
Note that

\[
\frac{1}{p^n} = \frac{(p + q)^n}{p^n} = (1 + \mu)^n > 1 + \mu + \cdots + \mu^n = \frac{1}{\pi_n},
\]
hence \(\pi_n > p^n.\) Then from \((4.16)\) and \((4.31),\)

\(1 + X_n \prec 1 + \text{geometric}(\pi_n) \prec 1 + \text{geometric}(p^n) \overset{d}{=} \tilde{X}_n.\)

\((4^*) \quad \tilde{X}_n \prec \tilde{X}_n - 1 \text{ if } n \geq n(u) = \frac{\log 1 - u}{\log u} - 1; \text{ in particular this holds for all } n \geq 1 \text{ if } u \leq \frac{\sqrt{5} - 1}{2} \approx 0.618.\)
From \((4.17),\)

\(1 + \tilde{X}_n \prec 1 + \text{geometric}(\bar{\pi}_n) \preceq 1 + \text{geometric}(u^n) \overset{d}{=} \tilde{X}_n\)
whenever \(\bar{\pi}_n \geq u^n.\) Because \(\bar{\pi}_n = \frac{1 - u}{1 - u^n + 1}\) by \((4.18),\) this inequality is equivalent to \(1 - u \geq u^{n+1},\) equivalently \(n \geq \frac{\log 1 - u}{\log u} - 1.\) The second statement holds because \(\frac{\log 1 - u}{\log u} - 1 \leq 1 \text{ iff } 1 - u \geq u^2,\) equivalently iff \(u \leq \frac{\sqrt{5} - 1}{2}.\)

\((5) \quad \tilde{X}_n \prec \tilde{X}_n:\)
Because \(u^n = \frac{1}{\mu^n} > \frac{1}{1 + \mu + \cdots + \mu^n} = \pi_n,\) \((4.22)\) and \((4.28)\) imply that

\(\tilde{X}_n \overset{d}{=} 1 + \text{geometric}(u^n) \prec 1 + \text{geometric}(\pi_n) \prec \tilde{X}_n.\)
(6) \( \tilde{X}_n \prec \tilde{X}_n \):
Because \( u^n > p^n \), this follows from (4.22) and (4.31).

(7) \( \tilde{X}_n \prec \tilde{X}_n \) for \( n \geq \max(n_1(u), n_2(u)) \) (defined below):
By (4.29) and (4.31), it suffices to show that for sufficiently large \( n \),

\[
(4.32) \quad V + W \prec Z,
\]

where, since \( \tilde{\mu} = u \) and \( p = \frac{u}{1 + u} \) in this example,

\[
V \overset{d}{=} \text{geometric}(\tilde{\mu}^n \tilde{\pi}_n) = \text{geometric}(\frac{u^n}{1 + u_n}),
\]

\[
W \overset{d}{=} \text{geometric}(\tilde{\pi}_n) = \text{geometric}(\frac{1}{1 + u_n}),
\]

\[
Z \overset{d}{=} \text{geometric}(p^n) = \text{geometric}(\left(\frac{u}{1 + u}\right)^n),
\]

with \( V \) and \( W \) independent. Because

\[
(4.33) \quad \Pr[\text{geometric}(\nu) \geq r] = (1 - \nu)^r \quad \text{for } r \geq 1,
\]

the upper tail probabilities of \( Z \) and \( V + W \) can be expressed as follows:

\[
\Pr[Z \geq r] = \left(1 - \left(\frac{u}{1 + u}\right)^n\right)^r;
\]

\[
\Pr[V + W \geq r] = \mathbb{E}\{\Pr[V \geq r - W | W]\}
\]
\[
= \left(\frac{1}{1 + u_n}\right)^{r-1} \sum_{w=0}^{r-1} \left(1 - \frac{u^n}{1 + u_n}\right)^{-w} \left(\frac{u_n}{1 + u_n}\right)^w + \sum_{w=r}^{\infty} 1 \cdot \left(\frac{1}{1 + u_n}\right) \left(\frac{u_n}{1 + u_n}\right)^w
\]
\[
= \sum_{w=0}^{r-1} \left(1 - \frac{u^n}{1 + u_n}\right)^{-w} \left(\frac{1}{1 + u_n}\right) \left(\frac{u_n}{1 + u_n}\right)^w + \left(1 - \frac{1}{1 + u_n}\right)^r
\]
\[
= \left(1 - \frac{u^n}{1 + u_n}\right)^r \left(\frac{1}{1 + u_n}\right) \sum_{w=0}^{r-1} \left(\frac{u_n}{1 + u_{n-1}}\right)^w + \left(1 - \frac{1}{1 + u_n}\right)^r
\]
\[
= \left(\frac{1 - u^n}{1 - u^{n+1}}\right)^r \sum_{w=0}^{r-1} u^w + u^r \left(\frac{1 - u^n}{1 - u^{n+1}}\right)^r
\]

24
\[
\begin{align*}
&= \left(\frac{1-u^n}{1-u^{n+1}}\right)^r \left[\left(\frac{1-u}{1-u^{n+1}}\right)^{r-1} \sum_{w=0}^{r-1} u^w + u^r\right] \\
&= \left(\frac{1-u^n}{1-u^{n+1}}\right)^r \left[\left(\frac{1-u}{1-u^{n+1}}\right) \left(\frac{1-u^r}{1-u}\right) + u^r\right] \\
&= \left(\frac{1-u^n}{1-u^{n+1}}\right)^r \left(\frac{1-u^{r+n+1}}{1-u^{n+1}}\right).
\end{align*}
\]

Thus (4.32) holds iff

\[
(4.34) \quad \left(\frac{1-u^{r+n+1}}{1-u^{n+1}}\right)^\frac{1}{r} \leq \left(\frac{1-u^{n+1}}{1-u^n}\right) \left(1 - \left(\frac{u}{1+u}\right)^n\right)
\]

for all \(r \geq 1\), with strict inequality for at least one \(r\). Rewrite (4.34) in the equivalent form

\[
(4.35) \quad \left(1 + \frac{u^{n+1}(1-u^r)}{1-u^{n+1}}\right)^\frac{1}{r} \leq \left(1 + \frac{u^n(1-u)}{1-u^n}\right) \left(1 - \left(\frac{u}{1+u}\right)^n\right),
\]

then take logarithms to obtain the equivalent inequality

\[
(4.36) \quad \frac{1}{r} \log \left(1 + \frac{u^{n+1}(1-u^r)}{1-u^{n+1}}\right) \leq \log \left(1 + \frac{u^n(1-u)}{1-u^n}\right) + \log \left(1 - \left(\frac{u}{1+u}\right)^n\right).
\]

Now apply the inequalities

\[
\begin{align*}
x - \frac{x^2}{2} &< \log(1+x) < x, \quad \text{if } 0 < x \leq 1, \\
-x - \frac{x^2}{2(1-x)} &< \log(1-x), \quad \text{if } 0 < x < 1.
\end{align*}
\]

These are applicable to the three logarithms in (4.36) provided that

\[
\frac{u^{n+1}(1-u^r)}{1-u^{n+1}} \leq 1 \text{ for all } r \geq 1 \quad \text{and} \quad \frac{u^n(1-u)}{1-u^n} \leq 1,
\]
equivalently provided that

\[
(4.37) \quad n \geq \max \left(\frac{2}{\log \frac{1}{u}} - 1, \frac{2-u}{\log \frac{1}{u}}\right) = n_1(u).
\]

25
For such \( n \) the left side (LS) and right side (RS) of (4.36) respectively satisfy

\[
\text{LS} < \frac{1}{r} \frac{u^{n+1}(1-u^n)}{1-u^{n+1}} \leq \frac{u^{n+1}(1-u)}{1-u^{n+1}},
\]
\[
\text{RS} > \frac{u^n(1-u)}{1-u^n} - \frac{1}{2} \left( \frac{u^n(1-u)}{1-u^n} \right)^2 - \left( \frac{u}{1+u} \right)^n - \frac{(\frac{u}{1+u})^{2n}}{2(1-(\frac{u}{1+u})^n)},
\]

so (4.36) holds if

\[
\frac{u(1-u)}{1-u^{n+1}} \leq \frac{u^n(1-u)^2}{2(1-u^n)^2} - \left( \frac{1}{1+u} \right)^n - \frac{(\frac{u}{1+u})^{n}}{2((1+u)^n-u^n)},
\]
equivalently, if

\begin{align*}
(4.38) & \quad \left( \frac{1}{1+u} \right)^n + \frac{u^n(1-u)^2}{2(1-u^n)^2} + \frac{(\frac{u}{1+u})^{n}}{2((1+u)^n-u^n)} \\
(4.39) & \quad \leq \frac{u(1-u)}{1-u^n} - \frac{u^n(1-u)}{1-u^{n+1}} \equiv \frac{(1-u)^2}{(1-u^n)(1-u^{n+1})}.
\end{align*}

As \( n \to \infty \), (4.38) approaches 0 while the right side of (4.39) approaches \((1-u)^2 > 0\), hence the inequality in (4.39) is satisfied for sufficiently large \( n \), say \( n \geq n_2(u) \). Therefore (4.36), and thus (4.32), holds for all \( n \geq \max(n_1(u), n_2(u)) \).

Summarizing: In the geometric case,

\[
(4.40) \quad \tilde{X}_n \begin{cases} \prec X_n \begin{cases} \prec \tilde{X}_n - 1 \\ \prec \tilde{X}_n - 1 \end{cases} \\ (n \geq u(u)) \end{cases} \begin{cases} \prec \tilde{X}_n - 1 < \tilde{X}_n \\ \prec \tilde{X}_n \end{cases},
\]
\[
(4.41) \quad \tilde{X}_1 \prec \tilde{X}_1, \quad \tilde{X}_n \prec \tilde{X}_n \text{ for } n \geq \max(n_1(u), n_2(u)). \quad \Box
\]

**Example 4.3a** (Poisson offspring distribution).

Suppose that \( \xi \sim \text{Poisson}(\lambda) \), that is, \( p_k = p_{\lambda,k} = e^{-\lambda} \lambda^k / k! \) for \( k = 0, 1, \ldots \).

Here \( \phi(s) = e^{\lambda(s-1)} \) and \( \mu \equiv E(\xi) = \lambda \), so extinction occurs with probability 1 iff \( \lambda \leq 1 \) (subexplosive). Here we assume that \( \lambda > 1 \).
(semiexplosive), so extinction occurs with probability \( u \), where by (1.5), \( u \) is the unique value in \((0,1)\) that satisfies

\[
e^\lambda(u-1) = u. \tag{4.42}
\]

An explicit solution is not available so this must be solved numerically. (Upper and lower bounds for \( u \) are given in Remark 4.1.) Also, by (4.42),

\[
\phi'_\lambda(u) = \lambda e^{\lambda(u-1)} = u\lambda. \tag{4.43}
\]

From (2.1) and (4.42),

\[
\tilde{\phi}_\lambda(s) = \frac{\phi_\lambda(us)}{u} = \frac{e^{\lambda(us-1)}}{e^{\lambda(u-1)}} = e^{u\lambda(s-1)} = \phi_{u\lambda}(s). \tag{4.44}
\]

By (4.44) and (4.43),

\[
\tilde{\mu} = \tilde{\phi}_\lambda'(1) = \phi'_{u\lambda}(1) = u\lambda = \phi'_\lambda(u) < 1 \tag{4.45}
\]

(see Figure 2), so \( \tilde{\phi}_\lambda \equiv \phi_{u\lambda} \) is a subcritical pgf.

To describe the pgfs \( \tilde{\phi} \), \( \hat{\phi} \), and \( \tilde{\phi} \), let \( K_\lambda \) denote a Poisson(\( \lambda \)) rv and

\[
K^*_\lambda \overset{d}{=} (K_\lambda - 1) \mid \{K_\lambda \geq 1\} \tag{4.46}
\]

a truncated conditional Poisson(\( \lambda \)) rv with range \( \{0, 1, 2, \ldots\} \) and pgf

\[
\phi^*_\lambda(s) = \frac{e^{\lambda s} - 1}{s(e^\lambda - 1)}. \]

Denote the distribution of \( K^*_\lambda \) by Poisson*(\( \lambda \)). Then from (2.2) and (4.42),

\[
\begin{align*}
\tilde{\phi}_\lambda(s) &= \frac{\phi_\lambda((1-u)s + u) - u}{1-u} = \frac{e^{\lambda((1-u)s+u-1)} - u}{1-u} \\
&= \frac{e^{\lambda(u-1)(1-s)} - u}{1-u} = \frac{u^{(1-s)} - u}{1-u} \\
&= \frac{u^{-s} - 1}{u^{-1} - 1} = \frac{e^{(1-u)\lambda s} - 1}{e^{(1-u)\lambda} - 1} = s \cdot \phi^*_{(1-u)\lambda}. \tag{4.47}
\end{align*}
\]
Next, from (2.19) \((n = 1)\) and (4.42),

\[
\dot{\phi}_{\lambda;[1]}(s) = \left[ \frac{e^{\lambda(s-1)} - e^{\lambda(us-1)}}{1 - e^{\lambda(u-1)}} \right] = \left[ \frac{e^{(1-u)\lambda s} - 1}{e^{(1-u)\lambda} - 1} \right] e^{u\lambda(s-1)} = s \cdot \phi^{*}_{(1-u)\lambda}(s) \cdot \phi_{u\lambda}(s).
\]

(4.48)

Last, from (4.23),

\[
\ddot{\phi}_{\lambda}(s) = \frac{e^{\lambda(s-1)} - e^{-\lambda}}{1 - e^{-\lambda}} = \frac{e^{\lambda s} - 1}{e^{\lambda} - 1} \equiv s \cdot \phi^{*}_{\lambda}(s).
\]

(4.49)

Thus (4.44), (4.47), (4.48), and (4.49) yield the representations

\[
\begin{align*}
\dot{\xi} & \overset{d}{=} X_1 \overset{d}{=} K_\lambda, \\
\ddot{\xi} & \overset{d}{=} \dot{X}_1 \overset{d}{=} K_{u\lambda} \\
\dddot{\xi} & \overset{d}{=} \ddot{X}_1 \overset{d}{=} 1 + K^{*}_{(1-u)\lambda}, \\
\dddot{\xi} & \overset{d}{=} \dddot{X}_1 \overset{d}{=} 1 + K^{*}_{(1-u)\lambda} + K_{u\lambda}, \\
\dddot{\xi} & \overset{d}{=} \dddot{X}_1 \overset{d}{=} 1 + K^{*}_{\lambda},
\end{align*}
\]

where the two rvs on the right side of (4.53) are independent.

We now establish several stochastic inequalities among the Poisson(\(\lambda\)) and Poisson*\((\lambda)\) families of distributions.

(i) It follows from (4.46) that

\[
K^{*}_{\lambda} + 1 \overset{d}{=} (K_\lambda \mid K_\lambda \geq 1) \succ K_\lambda.
\]

(ii) The Poisson(\(\lambda\)) family \(\{K_\lambda \mid \lambda > 0\}\) has strictly increasing likelihood ratio: for \(0 < \lambda_1 < \lambda_2\) and \(k \geq 0\),

\[
\frac{p_{\lambda_2,k}}{p_{\lambda_1,k}} = \frac{e^{-\lambda_2} \lambda_2^k/k!}{e^{-\lambda_1} \lambda_1^k/k!} = e^{\lambda_1 - \lambda_2} \left( \frac{\lambda_2}{\lambda_1} \right)^k,
\]

28
which is strictly increasing in \( k \). Therefore the Poisson(\( \lambda \)) family is strictly stochastically increasing in \( \lambda \).

(iii) Let \( p_{\lambda,k}^* = \Pr[K_\lambda^* = k] \ (k \geq 0) \) denote the pmf of a Poisson*\( (\lambda) \) rv. Then

\[
p_{\lambda,k}^* = \Pr[K_\lambda - 1 = k \mid K_\lambda \geq 1] = \frac{e^{-\lambda} \lambda^{k+1}}{(k+1)! (1 - e^{-\lambda})},
\]

so for \( 0 < \lambda_1 < \lambda_2 \) and \( k \geq 0 \),

\[
\frac{p_{\lambda_2,k}^*}{p_{\lambda_1,k}^*} = \frac{e^{\lambda_1} - 1}{e^{\lambda_2} - 1} \left( \frac{\lambda_2}{\lambda_1} \right)^{k+1},
\]

which is also strictly increasing in \( k \). Therefore the Poisson*\( (\lambda) \) family \( \{K_\lambda^* \mid \lambda > 0\} \) also is strictly stochastically increasing in \( \lambda \).

(iv) For fixed \( \lambda > 0 \), \( K_{(1-\alpha)\lambda}^* + K_{\alpha\lambda}^* \) is strictly stochastically increasing in \( \alpha \), \( 0 \leq \alpha \leq 1 \). Here \( K_{(1-\alpha)\lambda}^* \) and \( K_{\alpha\lambda}^* \) are assumed independent. As a consequence, for any \( \mu, \nu > 0 \), if \( K_\mu^* \) and \( K_\nu^* \) are independent then

\[
K_{\mu+\nu}^* < K_\mu^* + K_\nu^* < K_{\mu+\nu}.
\]

To establish the asserted result, note that for \( k \geq 0 \),

\[
\Pr[K_{(1-\alpha)\lambda}^* + K_{\alpha\lambda}^* = k]
\]

\[
= \Pr[K_{(1-\alpha)\lambda}^* - 1 + K_{\alpha\lambda}^* = k \mid K_{(1-\alpha)\lambda}^* \geq 1]
\]

\[
= \frac{\Pr[K_{(1-\alpha)\lambda}^* + K_{\alpha\lambda}^* = k + 1, K_{(1-\alpha)\lambda}^* \geq 1]}{\Pr[K_{(1-\alpha)\lambda}^* \geq 1]}
\]

\[
= \frac{\Pr[K_{(1-\alpha)\lambda}^* + K_{\alpha\lambda}^* = k + 1] - \Pr[K_{\alpha\lambda}^* = k + 1, K_{(1-\alpha)\lambda}^* = 0]}{\Pr[K_{(1-\alpha)\lambda}^* \geq 1]}
\]

\[
= \frac{e^{-\lambda} \lambda^{k+1}}{(k+1)!} - \frac{e^{-\alpha (\lambda) (k+1)} e^{-(1-\alpha) \lambda}}{(k+1)!}
\]

\[
= \frac{1 - e^{-(1-\alpha) \lambda}}{1 - e^{-(1-\alpha) \lambda}}
\]

\[
\frac{e^{-\lambda} \lambda^{k+1}}{(k+1)!} \left[ 1 - \alpha^{k+1} \right]
\]

\[
(4.56)
\]

where \( K_{(1-\alpha)\lambda}^* \) and \( K_{\alpha\lambda}^* \) are independent, so \( K_{(1-\alpha)\lambda}^* + K_{\alpha\lambda}^* \overset{d}{=} K_{(1-\alpha)\lambda} \). It follows from (4.56) that for \( 0 \leq \alpha < \beta \leq 1 \), the likelihood ratio is

\[
\frac{\Pr[K_{(1-\beta)\lambda}^* + K_{\beta\lambda}^* = k]}{\Pr[K_{(1-\alpha)\lambda}^* + K_{\alpha\lambda}^* = k]} = c \cdot \frac{1 - \beta^{k+1}}{1 - \alpha^{k+1}} = c' \cdot \frac{1 + \beta + \cdots + \beta^k}{1 + \alpha + \cdots + \alpha^k},
\]

\[
29
\]
which is strictly increasing in \( k \), which implies the asserted result.

(v) The likelihood ratio

\[
\frac{\Pr[K^*_\nu = k]}{\Pr[K_\lambda = k]} = \frac{p_{\nu,k}^*}{p_{\lambda,k}} = \frac{\nu e^\lambda}{e^\nu - 1} \left(\frac{\nu}{\lambda}\right)^k \frac{1}{k+1}
\]

is strictly increasing in \( k \) if \( \nu \geq 2\lambda \), in which case \( K^*_\nu \succ K_\lambda \). \( \Box \)

These results can be combined with (4.50) - (4.54) to obtain the following stochastic comparisons for \( n = 1 \):

\[
\left\{
\begin{array}{c}
\tilde{X}_1 \prec X_1 \\
\prec 1 + \tilde{X}_1 \\
\tilde{X}_1 \overset{(u \leq \frac{1}{2})}{\prec} \tilde{X}_1
\end{array}\right. \prec \tilde{X}_1 \prec X_1 \prec (1 + X_1);
\]

(4.57)

\[
1 + \tilde{X}_1 \overset{(u \leq \frac{1}{2})}{\prec} \tilde{X}_1, \quad \tilde{X}_1 \overset{(u \geq \frac{1}{2})}{\prec} 1 + \tilde{X}_1.
\]

(4.58)

For \( n \geq 2 \) in the Poisson case, unlike the geometric case it is not possible to obtain explicit stochastic representations for \( X_n, \tilde{X}_n \), etc. The only general result in (3.13) - (3.14) that can be strengthened using (4.57) - (4.58) is (6): \( \mathcal{X} \prec \mathcal{X} \), because both \( \mathcal{X} \) and \( \mathcal{X} \) are GW processes hence Lemma 3.1 can be applied. \( \Box \)

**Remark 4.1.** (Poisson offspring distribution, continued).

The following exponential bounds for the extinction probability \( u \equiv u_\lambda \) hold in the semiexplosive case \( (\lambda > 1) \):

\[
\frac{1}{e^\lambda - 1} < u_\lambda < \frac{1}{e^{\lambda-1}}.
\]

(4.59)

To derive the first inequality, apply (3.15) to see that

\[ u > \frac{\lambda e^{-\lambda}}{\lambda - \lambda u(1 - e^{-\lambda})} = \frac{e^{-\lambda}}{1 - u(1 - e^{-\lambda})}, \]

equivalently,

\[ (e^\lambda - 1)u^2 - e^\lambda u + 1 < 0, \]

30
which implies that \( u > \frac{1}{e^{\lambda - 1}} \). To derive the second inequality in (4.59), apply (4.42) and (4.45) to see that \( u = e^{\lambda u - \lambda} < e^{1 - \lambda} \).

\[ \square \]

5. Predicting extinction vs. explosion

Suppose that, based on observations \( X_1, \ldots, X_n \) up to time \( n \) from a semiexplosive GW process, one wishes to predict whether ultimate extinction or ultimate explosion will occur. Guttrop and Perlman [4] formulate this as a simple hypothesis testing problem: test

\[
H_{0, p}^n : X_n \overset{d}{=} \tilde{X}_n \text{ (extinction)} \quad \text{vs.} \quad H_{1, p}^n : X_n \overset{d}{=} \hat{X}_n \text{ (explosion)}.
\]

They show that for each fixed generating distribution \( p \equiv (p_0, p_1, p_2, \ldots) \), the most powerful size \( \alpha \) prediction procedure has the simple form

\[
(5.1) \quad \begin{cases} \text{predict extinction} & \text{if } X_n \leq c_{n, \alpha}(p), \\ \text{predict explosion} & \text{if } X_n > c_{n, \alpha}(p), \end{cases}
\]

where \( c_{n, \alpha}(p) \equiv c \) is a positive integer. It follows from the general result (3.13) that \( X_n \prec \hat{X}_n \), hence this prediction procedure is strictly unbiased:

\[
\Pr[\text{predict extinction} \mid \text{extinction}] = \Pr[X_n \leq c \mid \text{extinction}] = \Pr[\tilde{X}_n \leq c] > \Pr[\hat{X}_n \leq c] = \Pr[X_n \leq c \mid \text{explosion}] = \Pr[\text{predict explosion} \mid \text{explosion}] .
\]

Perhaps surprisingly, they also show that the simple choice \( c_{n, \alpha}(p) = 1/\alpha \equiv c(\alpha) \) provides a (conservative) level \( \alpha \) test that is universal, i.e., depends on neither \( n \) nor \( p \). Because \( \mathcal{X} \) and \( \hat{X} \) are subcritical and explosive GW processes respectively and \( \mathcal{X} \prec \hat{X} \) (cf. (3.13)), both error probabilities are small:

\[
\Pr[\hat{X}_n > c(\alpha)] \rightarrow 1 \quad \text{and} \quad \Pr[\hat{X}_n \leq c(\alpha)] \rightarrow 0 \quad \text{as } n \rightarrow \infty,
\]

hence this prediction procedure is consistent.
References


