Analysis of the branching process relies heavily on the use of p.g.f.'s.

Formulation
Suppose that each individual in a population independently produces $n$ offspring in the succeeding generation with probability $p(n)$ for $n = 0, 1, \ldots$. Consider the discrete time, discrete, space stochastic process $\{X_n : n = 0, 1, \ldots\}$, where $X_n$ represents the number of individuals in the $n$-th generation. Then $\{X_n\}$ is a time homogeneous Markov Chain, but the t.p.m. is difficult to manipulate directly.

Let
$$G(z) = \sum_{n=0}^{\infty} p(n) z^n$$

be the p.g.f. of the number of progeny produced by an individual and let $G_n(z)$ denote the p.g.f. of $X_n$.

We assume initially that $x_0 = 1$, then
$$G_0(z) = z \quad \text{and} \quad G_1(z) = G(z).$$
Proposition: If the initial population size \( X_0 = 1 \), then
\[
G_n(2) = G_{n-1}(G(2)) = G(G(\cdots G(2) \cdots))
\]

Proof: \(E(2^{X_n} | X_{n-1}) = E\{G(2)^{X_{n-1}}\} = E[G_{n}(2)^{X_{n-1}}] = G_{n-1}(G(2))\).

Hence, \(G_n(2) = G_{n-1}(G(2)) = G_{n-2}(G(G(2))) = \cdots = G_0(G(G(\cdots G(2) \cdots))) = G(G(\cdots G(2) \cdots)) [\because G_0(2) = 2] \)

Mean and Variance in the \(n\)-th generation

Let \(\mu\) and \(\kappa\) denote the mean and variance of the number of offspring produced by a single individual and define
\[
\mu_n = E(X_n) \quad \text{and} \quad \kappa_n = \text{var}(X_n)
\]

Proposition: If the initial population size \( X_0 = 1 \), then, for \( n = 1, 2 \),
\[
\mu_n = \mu^n, \quad \kappa_n = \begin{cases} \kappa \mu^{n-1} (1-\mu^n) (1-\mu), & \mu \neq 1 \\ \kappa, & \mu = 1 \end{cases}
\]
proof: \[ E(X_n) = E \left\{ E(X_n | X_{n-1}) \right\} \]
\[ = \mu E(X_{n-1}) \]

Similarly, \[ \text{var}(X_n) = \text{var} \left\{ E(X_n | X_{n-1}) \right\} + E \left\{ \text{var}(X_n | X_{n-1}) \right\} \]
\[ K_n = \mu K_{n-1} + \mu^{n-1} \lambda, \quad n = 1, 2, \ldots \]

and these recurrence equations can be solved to obtain \( K_n \).

However, an instead step analysis provides an alternative set of equations when \( \mu 
eq 1 \).

\[ \text{var}(X_n) = \text{var} \left\{ E(X_n | X_1) \right\} + E \left\{ \text{var}(X_n | X_1) \right\} \]
\[ K_n = \mu^{2n-2} K + \mu^{n-1} \lambda, \quad n = 1, 2, \ldots \]

and we can eliminate \( K_{n-1} \) from the two eqns. for \( K_n \) to get \( K_n \). When \( \mu = 1 \), both sets of eqns. give \( K_n = K_{n-1} + \lambda \) and hence \( K_n = n \lambda \).

**Extinction Probabilities**

Let \( x_n = \Pr(X_n = 0) \) denote the probability of extinction by the \( n \)-th generation. Note that the probability of extinction at the \( n \)-th generation is not \( x_n \) but \( x_n - x_{n-1} \).
Proposition: If \( x_0 = 1 \), then \( x_0 = 0 \) and the recursion 
\[ x_n = G(x_{n-1}) \]
determines \( x_n \) for \( n = 1, 2, \ldots \).

Proof: Clearly \( x_0 = 0 \) and 
\[ x_n = G(x_n - 1) = G(G(x_{n-1})) = G(G(x_{n-2})) \]
for \( n = 1, 2, \ldots \).

Proposition: Let \( \alpha \) denote the probability of ultimate extinction. Then if the initial population size \( x_0 = 1 \), \( \alpha \) is the smallest root of \( G(\alpha) = 0 \) in \([0, 1]\).

Proof: First, note that \( \alpha = \lim_{n \to \infty} x_n \)
because, \( 0 = x_0 \leq x_1 \leq x_2 \leq \cdots \leq 1 \).

Therefore, we can take the limit as \( n \to \infty \) in the recursion for \( x_n \) and deduce that \( \alpha \) satisfies \( \alpha = G(\alpha) \). It remains to show that \( \alpha \) is the smallest root of \( G(\alpha) = 0 \) in \([0, 1]\).

Let \( \beta \) denote any non-negative root of 
\[ G(\beta) = \beta^p(0) + \beta^p(1)2 + \beta^p(2)2^2 + \cdots \]
is non-decreasing for \( \beta > 0 \). We prove by induction
that \( x_n \leq \beta \) for \( n = 0, 1, \ldots \). Thus, for \( n = 0, 1, \)
\[ x_0 = G(0) \leq G(\beta) = \beta \]
because \( \beta > 0 \). Now suppose that \( \alpha_n \leq \beta \) then
\[
\alpha_{n+1} = g(\alpha_n) \leq g(\beta) = \beta
\]
again using the fact that \( g(2) \) is non-decreasing
for \( z \geq 0 \). This completes the proof that \( \alpha_n \leq \beta \) for all \( n \), which implies that \( \alpha \leq \beta \) and \( \alpha \) is
the smallest non-negative root of \( g(2) = 2 \).
A root in \([0, 1]\) is ensured, because \( g(1) = 1 \).

**Corollary:** \( \alpha < 1 \) if \( \mu > 1 \) and \( \alpha = 1 \) if \( \mu \leq 1 \).

**Proof:** \( \mu = g'(1) \) and \( g(2) \) is non-negative, non-decreasing and convex in \([0, 1]\), the last because \( g''(z) \) is non-negative for \( z \geq 0 \). Thus if \( g'(1) < 1 \), \( g(0) > 0 \) and the only root of \( g(2) = 2 \)
in \([0, 1]\) is \( z = 1 \) and so \( \mu < 1 \implies \alpha = 1 \).
If \( g'(1) > 1 \), there must be a (unique) root of \( g(2) = 2 \)
in \([0, 1]\) and hence \( \mu > 1 \) \( \implies \alpha < 1 \). Now, if \( g'(1) = 1 \), there is again
a single root \( z = 1 \), unless \( g(2) = 2 \), in which case each individual has a single child with force 1 and of course \( \alpha \leq 0 \).
Explicit distributions: the simplest case

Assume \( p(z) = \left(1 - \frac{z}{y}\right) \mu^y \frac{1}{(1 + \mu)^{y+1}} \), \( y = 0, 1, \ldots \)

where \( \mu = \frac{p}{1-p} \).

Hence, \( g(z) = \frac{z}{1 + \mu - \mu g(z)} \)

and, assuming \( x_0 = 1 \) as usual, \( g_n(z) = g(g_{n-1}(z)) \) implies that \( g_n(z) = \frac{z}{1 + \mu - \mu g_{n-1}(z)} \), \( n = 1, 2, \ldots \)

with \( g_0(z) = z \). Following approach provides a way of solving the above equations:

Consider, \( w_n = \frac{aw_{n-1} + b}{cw_{n-1} + d} \), \( n = 1, 2, \ldots \)

or, instead, the linear recurrence relations:

\[
\begin{align*}
  u_n &= au_{n-1} + b u_{n-1} & (n = 1, 2, \ldots) \\
  v_n &= c u_{n-1} + d u_{n-1}
\end{align*}
\]

If we can find \( u_n \) and \( v_n \), then \( w_n = u_n / v_n \), with \( w_0 = w_0 \) and \( u_0 = 1 \), must solve the previous set of equations.
We can rewrite the equations for \( u_n \) and \( w_n \) in matrix form by defining the matrix

\[
Q = \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} 0 & -\mu \\ 1 & 1+\mu \end{pmatrix}
\]

in our particular case. Then

\[
(u_n, w_n) = (u_{n-1}, w_{n-1})Q = (u_{n-2}, w_{n-2})Q^2 = \ldots = (u_0, w_0)Q^n,
\]

so that we merely require the \( n \)th power of \( Q \).

The eigenvalues of \( Q \) are 1 and \( \mu \), so that, for \( \mu \neq 1 \),

\[
Q^n = A + B\mu^n, \quad n = 0, 1, \ldots
\]

with the matrices \( A \) and \( B \) determined by

\[
Q^0 \text{ and } Q^1. \quad \text{Thus,}
\]

\[
Q^n = \frac{1}{(1-\mu)} \left\{ \begin{pmatrix} -\mu & -1 \\ 1 & 1 \end{pmatrix} + \mu^n \begin{pmatrix} 1 & \mu \\ -1 & -\mu \end{pmatrix} \right\},
\]

with \((u_0, w_0) = (2, 1)\). Hence

\[
\alpha_n(2) = \frac{1 - \mu 2 - \mu^n (1-2)}{1 - \mu 2 - \mu^{n+1}(1-2)} > \mu \neq 1.
\]
and $X_n$ has a zero-modified distribution with

$$P(X_n = 0) = \alpha_n = \text{Geom}(0) = \frac{1 - \mu^n}{1 - \mu^{n+1}},$$

$$P(X_n = x) = \frac{(1 - \mu)^{x+1} \mu^{n+1} x^{-1} (1 - \mu^n)^x - 1}{(1 - \mu^{n+1})^{x+1}}, \quad x = 1, 2, \ldots,$$

provided $\mu \neq 1$. 
