ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

D-BSSE
Department of Biosystems
Science and Engineering

Branching processes in biology

Niko Beerenwinkel

Outline

- Galton-Watson process
- Probability generating functions
- Extinction
- Multiple types

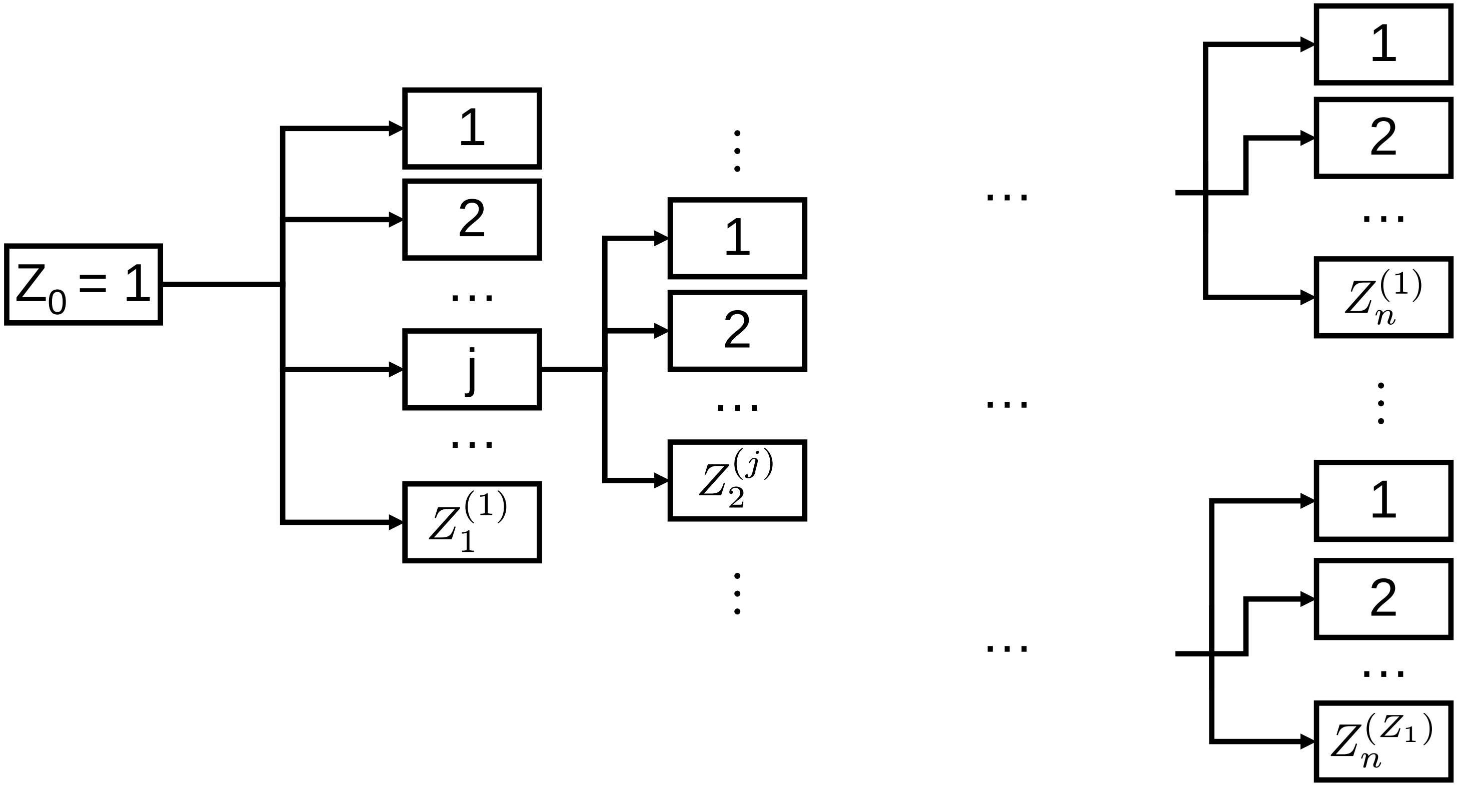
Galton-Watson process

- A single ancestor lives for one unit of time after which it produces a random number of offspring, Z, according to a fixed probability distribution.
- Each offspring behaves independently and identical to the ancestor.
- Let Z_n be the number of individuals in generation n.
  $Z_0 = 1$, $Z_1 = Z$.
- The Galton-Watson process is the Markov chain

$$\{Z_n \mid n = 0, 1, 2, \ldots\}$$

defined on the non-negative integers.

Backward equation



$$Z_{n+1} = \sum_{j=1}^{Z_1} Z_n^{(j)}$$

Transition probabilities

- Set p_k = Prob(Z = k).
- Let P(i, j) = Prob(Z_{n+1} = j | Z_n = i) be the transition probabilities of the time-homogeneous Markov chain.
- Note that P(1, k) = p_k.
- $P(2, j) = p_0\, p_j + p_1\, p_{j-1} + p_2\, p_{j-2} + p_3\, p_{j-3} + \ldots + p_j\, p_0$
- In general, P(0, j) = δ_{0j}, and for i $\geq$ 1,

$$P(i,j) = p_j^{*i} = \sum_{k_1+\cdots+k_i=j} p_{k_1} \cdots p_{k_i}$$

- $\{p_k^{*i}\}_{k\geq 0}$ is the *i-fold convolution* of $\{p_k\}_{k\geq 0}$.

Probability generating function (pgf)

- For the discrete random variable Z ~ $\{p_k\}_{k \geq 0}$, we define the *probability generating function* (pgf)

$$f(s) = \mathrm{E}\left[s^Z\right] = \sum_{k=0}^{\infty} p_k s^k \qquad s \in [0, 1]$$

- The pgf generates the distribution *p*:

$$\frac{d^k f}{ds^k}(0) \;=\; k!\, p_k \qquad k \geq 0$$

Properties of the pgf

- Moments of *Z* :

$$
\begin{aligned}
\mathrm{E}[Z] &= f'(1) \\
\mathrm{Var}[Z] &= f'(1) + f''(1) - f'(1)^2
\end{aligned}
$$

- Powers of *f* :

$$
\begin{aligned}
f(s) &= \sum_j P(1,j)s^j \\
[f(s)]^k &= \sum_j P(k,j)s^j \qquad k \geq 1
\end{aligned}
$$

Iterating and n-step transitions

- Define
$$\begin{aligned} f^{(0)}(s) &= s \\ f^{(1)}(s) &= f(s) \\ f^{(n+1)}(s) &= f(f^{(n)}(s)) \end{aligned}$$

- Let f_n be the pgf of Z_n.
- Denote by $P_n(i, j)$ the n-step transition probabilities.
- The Chapman-Kolmogorov equations assert that

$$P_{n+m}(i,j) = \sum_{k=0}^{\infty} P_n(i,k) P_m(k,j)$$

Proposition: $f_n = f^{(n)}$

$$
\begin{aligned}
f_{n+1}(s) &= \sum_j P_{n+1}(1,j)s^j \\
&= \sum_j \sum_k P_n(1,k)P(k,j)s^j \\
&= \sum_k P_n(1,k) \sum_j P(k,j)s^j \\
&= \sum_k P_n(1,k) f(s)^k \\
&= f_n(f(s)) = \cdots = f^{(n+1)}(s)
\end{aligned}
$$

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

D-BSSE
Department of Biosystems
Science and Engineering

Moments of Z_n

- We assume throughout that $p_0 + p_1 < 1$ and $p_j \neq 1$ for all j.
- If they exist, the moments of Z_n can be expressed in terms of the derivatives of f at $s = 1$.
- Set $m = E[Z] = E[Z_1] = f'(1)$
  and $\sigma^2 = \mathrm{Var}[Z] = f'(1) + f''(1) - f'(1)^2$. Then:

$$\mathrm{E}[Z_n] = m^n$$

$$\mathrm{Var}[Z_n] = \begin{cases} \frac{\sigma^2 m^{n-1}(m^n - 1)}{m-1} & \text{if } m \neq 1 \\ n\sigma^2 & \text{if } m = 1 \end{cases}$$

Extinction

- $Z_n = 0$ is an absorbing state.

$$\begin{aligned}
\rho &= \text{Prob}(Z_i = 0 \text{ for some } i \geq 0) \\
&= \lim_{n\to\infty} \text{Prob}(Z_i = 0 \text{ for some } 1 \leq i \leq n) \\
&= \lim_{n\to\infty} \text{Prob}(Z_n = 0) \\
&= \lim_{n\to\infty} f_n(0)
\end{aligned}$$

- Thus, we have to study the limit behavior of the pgf.

Properties of the pgf

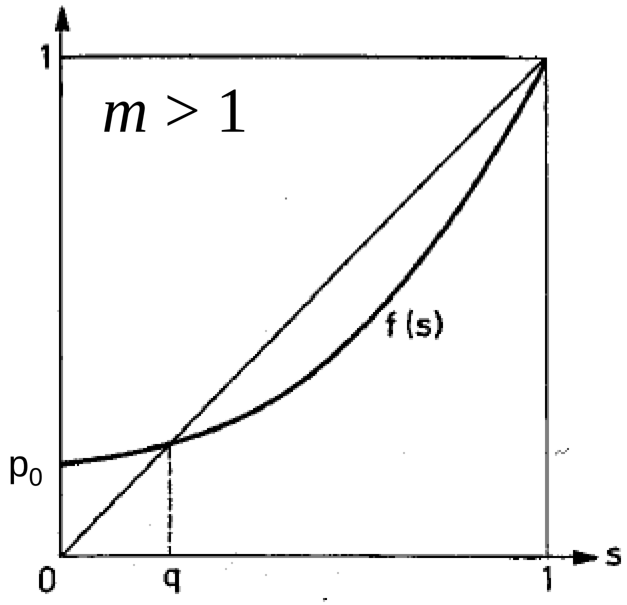
- The pgf, f, is a power series with non-negative coefficients $\{p_k\}_{k \geq 0}$ adding up to 1 (and $p_0 + p_1 < 1$). Hence:

i. f is strictly convex and increasing in [0,1]
ii. $f(0) = p_0$ and $f(1) = 1$
iii. If $m = f'(1) \leq 1$, then $f(s) > s$ for $s \in [0, 1)$
iv. If $m > 1$, then $f(s) = s$ has a unique root in $[0, 1)$

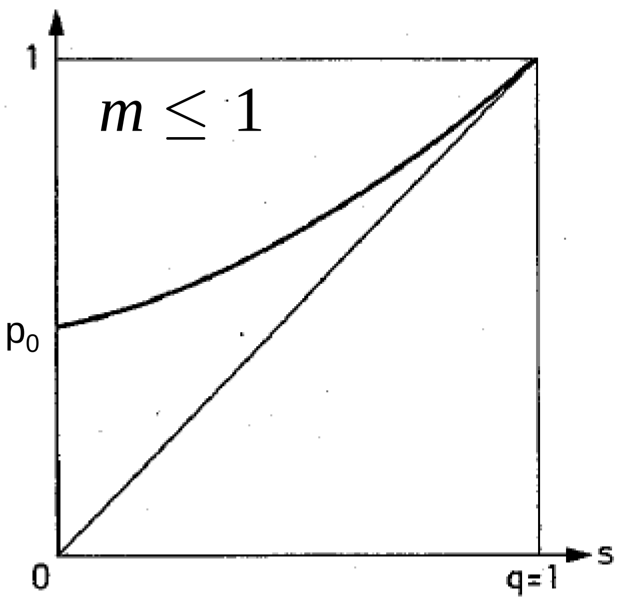
- Let q be the smallest root of $f(s) = s$ for $s \in [0, 1]$.

Roots of the pgf

- Let q be the smallest root of f(s) = s. Then:

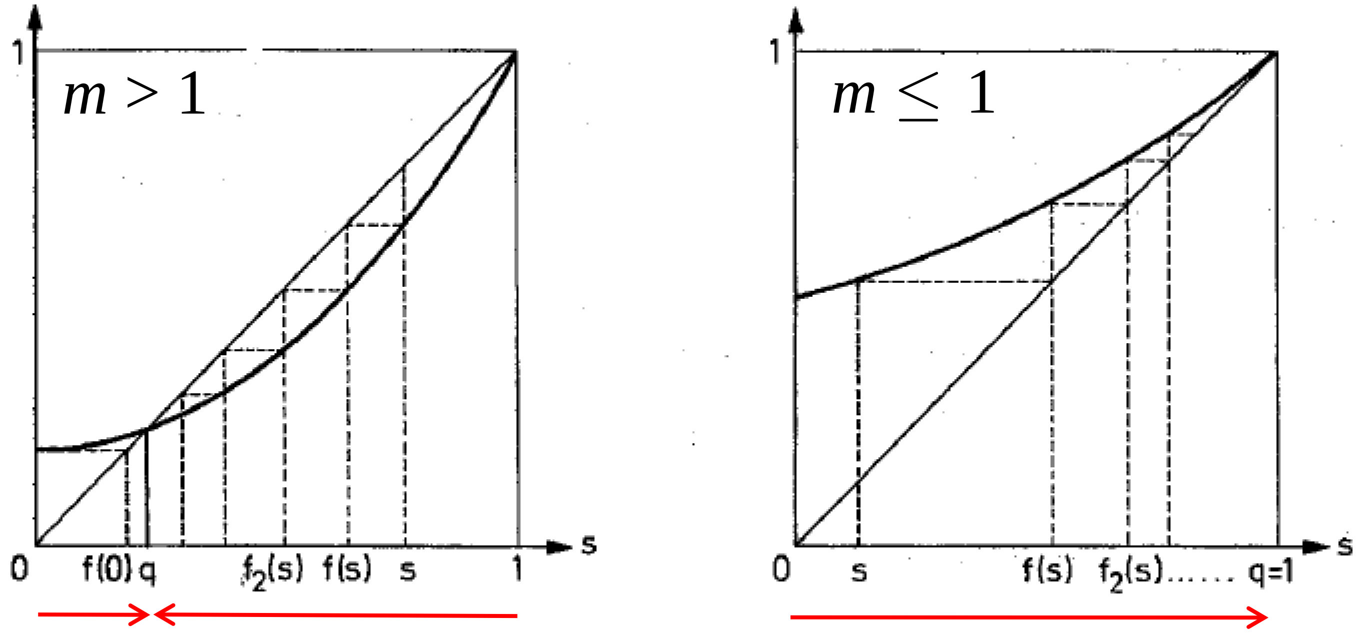


$\Rightarrow \quad q < 1$



$\Rightarrow \quad q = 1$

Limit behavior



- If $s \in [0, q)$ then $f_n(s) \nearrow q$ as $n \to \infty$.
- If $s \in (q,1)$ then $f_n(s) \searrow q$ as $n \to \infty$.
- If $s = q$ or $s = 1$ then $f_n(s) = s$ for all n.

Extinction probability

- **Theorem**:
  The extinction probability of the Galton-Watson process $\{Z_n\}$ is the smallest non-negative root q of the equation f(s) = s.
  If $m \leq 1$ then q = 1. If m > 1 then q < 1.

- Criticality:

supercritical	$m > 1$	$\mathrm{E}[Z_n] \nearrow \infty$	$q < 1$
critical	$m = 1$	$\mathrm{E}[Z_n] = 1$	$q = 1$
subcritical	$m < 1$	$\mathrm{E}[Z_n] \searrow 0$	$q = 1$

All positive states are transient

$$\mathrm{Prob}(Z_{n+i} \neq k \mid Z_n = k,\, i \geq 1) \geq$$

$$\geq \left\{ \begin{array}{ll} P(k,0) & \text{if } p_0 > 0 \\ 1 - P(k,k) & \text{if } p_0 = 0 \end{array} \right\} > 0$$

- Thus, with probability 1,

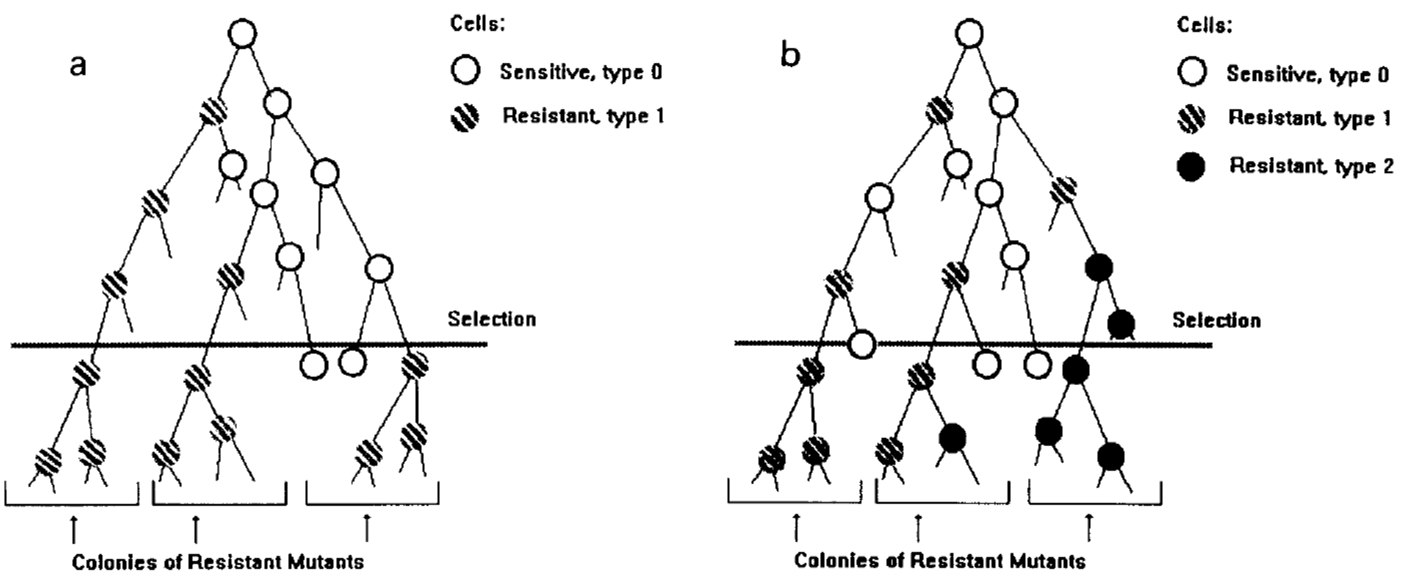
$$Z_n \to 0 \quad \text{or} \quad Z_n \to \infty.$$

Instability

- **Theorem:**

$$
\begin{aligned}
\lim_{n\to\infty} \mathrm{Prob}(Z_n = k) &= 0 \qquad\qquad k \geq 1 \\
\mathrm{Prob}\left(\lim_{n\to\infty} Z_n = 0\right) &= q \\
\mathrm{Prob}\left(\lim_{n\to\infty} Z_n = \infty\right) &= 1 - q
\end{aligned}
$$

Multiple types



The multi-type Galton Watson process

- We consider two types: type 0 (wild type) and type 1 (mutant) with counts $Z_0(t)$ and $Z_1(t)$, respectively, in generation $t \in \{0, 1, 2, \ldots\}$. [Note the change in notation!]
- Each cell at the moment of division gives birth to two daughter cells. A type 0 cell has type 1 offspring with probability α, the mutation rate.
- The mutation is irreversible: type 1 cells can not produce type 0 offspring.

Probability generating function

- The components of the pgf $F = (F_0, F_1)$ are

$$F_0(s_0, s_1; t) = \mathrm{E}\left[s_0^{Z_0(t)} s_1^{Z_1(t)} \mid Z_0(0) = 1,\ Z_1(0) = 0\right]$$

$$F_1(s_0, s_1; t) = \mathrm{E}\left[s_0^{Z_0(t)} s_1^{Z_1(t)} \mid Z_0(0) = 0,\ Z_1(0) = 1\right]$$

- We also write $F_i(t) = F_i(s; t)$, where $s = (s_0, s_1)$.

Recurrence equations

$$\begin{aligned} F_0(s;t) &= [(1-\alpha)F_0(s;t-1) \\ &\qquad + \alpha F_1(s;t-1)]^2 \end{aligned}$$

$$F_1(s;t) = [F_1(s;t-1)]^2$$

Differentiation of the recurrence equations

- Differentiation w.r.t. s_0 yields at s = (1,1), for F_1 and F_0 resp.,

$$\mathrm{E}\left[Z_0(t) \mid Z_i(0) = \delta_{1i}\right] = 2\mathrm{E}[Z_0(t-1) \mid Z_i(0) = \delta_{1i}] = 0$$

$$\mathrm{E}\left[Z_0(t) \mid Z_i(0) = \delta_{0i}\right] = 2(1-\alpha)\mathrm{E}[Z_0(t-1) \mid Z_i(0) = \delta_{0i}]$$

$$\Rightarrow \quad \mathrm{E}[Z_0(t) \mid Z_i(0) = \delta_{0i}] = [2(1-\alpha)]^t$$

the expected total number of wild type cells at time t.

The number of cells at time t

- The expected total number of cells is

$$\begin{aligned} N(t) &= \mathrm{E}[Z_0(t) + Z_1(t) \mid Z_i(0) = \delta_{0i}] \\ &= 2^t \end{aligned}$$

- Thus, the expected number of mutant cells is

$$\begin{aligned} r(t) &= \mathrm{E}[Z_1(t) \mid Z_i(0) = \delta_{0i}] \\ &= 2^t - [2(1-\alpha)]^t \\ &= 2^t[1-(1-\alpha)^t] \end{aligned}$$

The probability of a mutant-free population

- The probability of mutant cells being absent from the population at time t is

$$\begin{aligned} P_0(t) &= F_0(1, 0; t) \\ &= \mathrm{E}\left[1^{Z_0(t)}\, 0^{Z_1(t)} \mid Z_i(0) = \delta_{0i}\right] \end{aligned}$$

where

$$0^{Z_1(t)} = \begin{cases} 1 & \text{if } Z_1(t) = 0 \\ 0 & \text{else} \end{cases}$$

Recurrence equations at s = (1, 0)

- Set $P_1(t) = F_1(1, 0; t)$.
- The recurrence equations at s = (1, 0) yield

$$\begin{aligned} P_0(t) &= [(1-\alpha)P_0(t-1) + \alpha P_1(t-1)]^2 \\ P_1(t) &= [P_1(t-1)]^2 \end{aligned}$$

with initial conditions $P_0(0) = 1$ and $P_1(0) = 0$.

- We find $P_1(t) = 0$ for all t = 0, 1, 2, …
- Then,

$$\begin{aligned} P_0(1) &= (1-\alpha)^2 \\ P_0(2) &= [(1-\alpha)(1-\alpha)^2]^2 = (1-\alpha)^2(1-\alpha)^4 \\ P_0(3) &= (1-\alpha)^2(1-\alpha)^4(1-\alpha)^8 \\ &\ldots \end{aligned}$$

Solution

- Because

$$1 + 2 + 4 + 8 + \cdots + 2^t = 2^{t+1} - 1$$

we have for all t = 0, 1, 2, …,

$$P_0(t) = (1-\alpha)^{2^{t+1}-2}$$

$$P_1(t) = 0$$

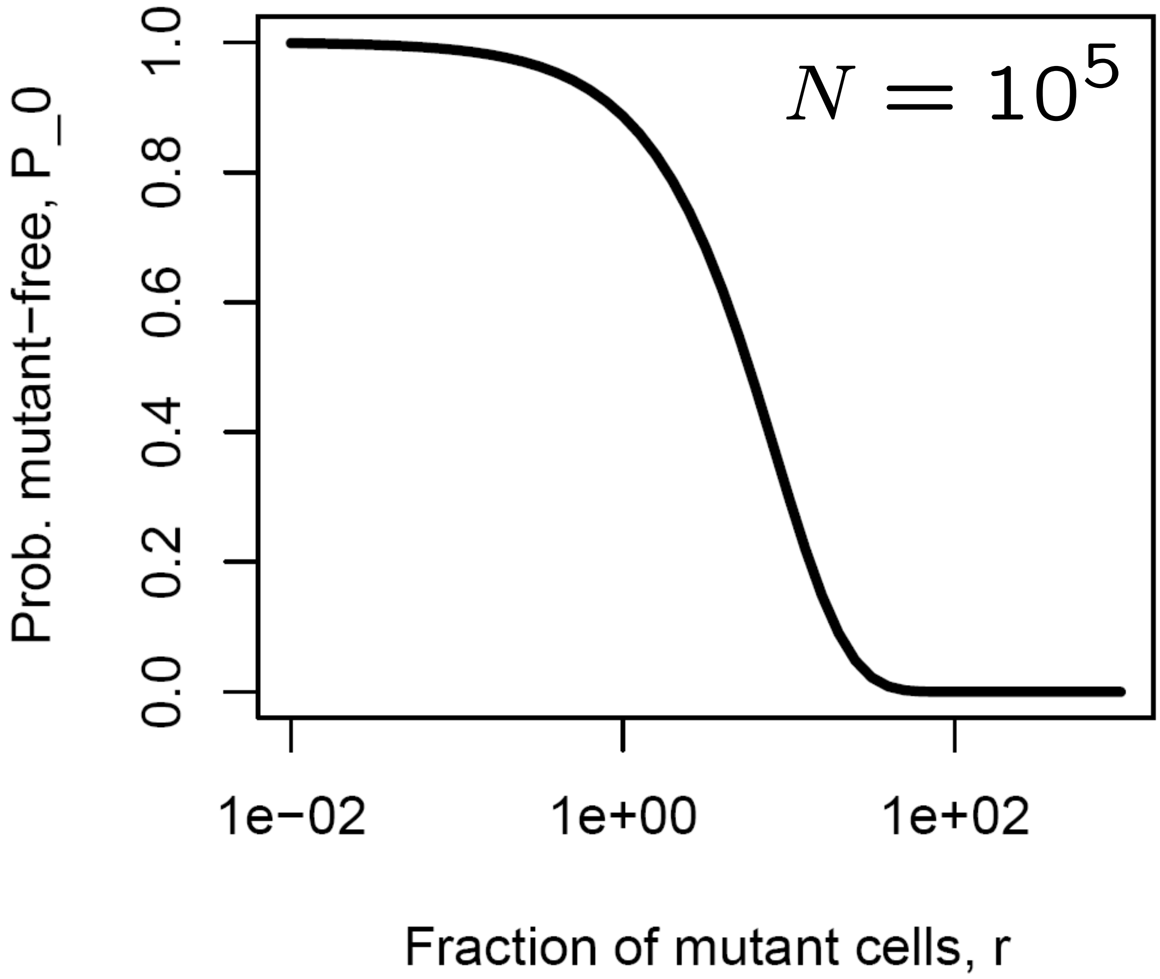
Summary of the irreversible 2-type GW process

$$
\begin{aligned}
N(t) &= 2^t \\
r(t) &= 2^t \left[1 - (1-\alpha)^t\right] \\
P_0(t) &= (1-\alpha)^{2(2^t-1)}
\end{aligned}
$$

- For each fixed N, we can solve for P_0 to obtain

$$
P_0(r) = \left(1 - \frac{r}{N}\right)^{\frac{2(N-1)}{\log_2 N}}
$$

r–P_0 plot


$N = 10^5$
Prob. mutant−free, P_0
0.0
0.2
0.4
0.6
0.8
1.0
1e−02
1e+00
1e+02
Fraction of mutant cells, r

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

D-BSSE
Department of Biosystems
Science and Engineering

Drug resistance data

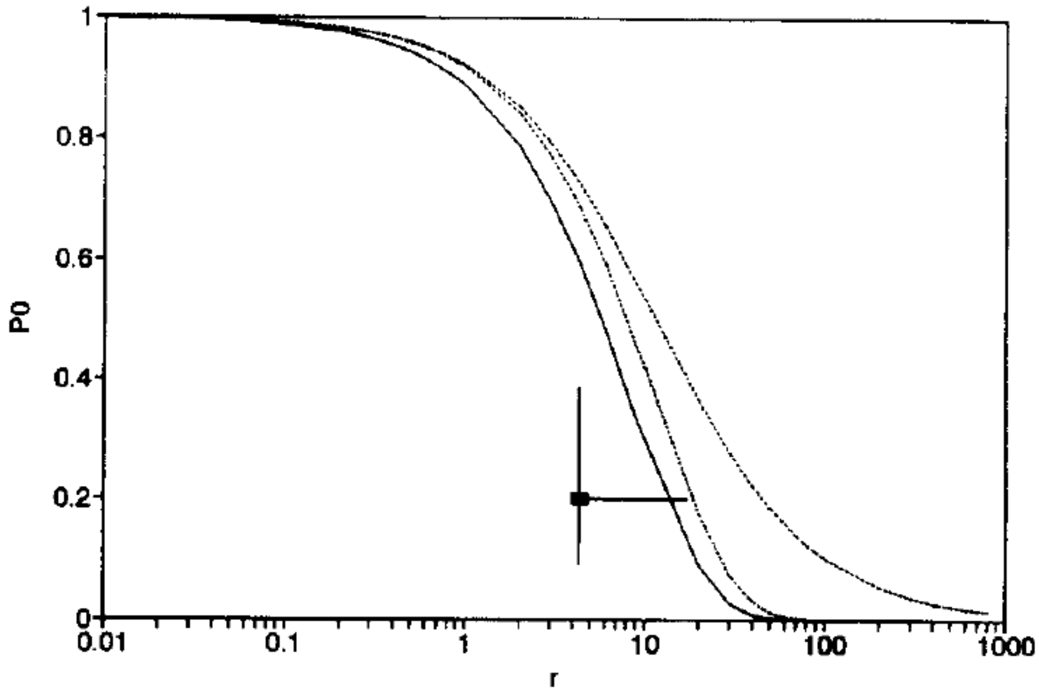


Fig. 5. The drug resistance data for $M - Mc$ mouse cells from experiment 1 of Morrow (1970) and the $r - P_0$ plots ($N = 10^5$) of the Galton–Watson, Luria–Delbrück and Markov branching process models (from left to right).

Classification of branching processes

- Lifetime
  - $\tau = 1$ (Galton-Watson process)
  - $\tau \sim \mathrm{Exp}(\lambda)$
  - τ any distribution (Bellman-Harris process)
- Type space
  - single type
  - multi-type
  - denumerable {1, 2, 3, …}
  - continuous
  - abstract
- Offspring distribution

Summary

- A branching process models an evolving population of finite, fluctuating size with i.i.d. offspring distribution.
- The Galton-Watson process defines a Markov chain on the non-negative integers.
- The probability generating function is the main mathematical tool to study branching processes.
- Branching processes are inherently instable: They predict extinction or indefinite growth of the population.
- The fate of different types (mutants) can be studied using multi-type branching processes.

Further reading

- Athreya KB, Ney PE. Branching processes. Dover, 1972.
- Kimmel M, Axelrod DE. Branching Processes in Biology. Springer, 2000.
- Kimmel M, Axelrod DE. Fluctuation test for two-stage mutations: application to gene amplification. Mutat Res 306:45-60, 1994.
- Haccou P, Jagers P, Vatutin VA (Eds.). Branching processes: Variation, growth, and extinction of populations Cambridge University Press, 2005.