



*The Abdus Salam
International Centre for Theoretical Physics*



1863-5

**Advanced School and Conference on Statistics and Applied
Probability in Life Sciences**

24 September - 12 October, 2007

Branching Processes and Population Dynamics

Peter Jagers
*Chalmers University of Technology
Mathematical Sciences
Gothenburg
Sweden*



Branching Processes and Population Dynamics

Peter Jagers

Trieste 2007

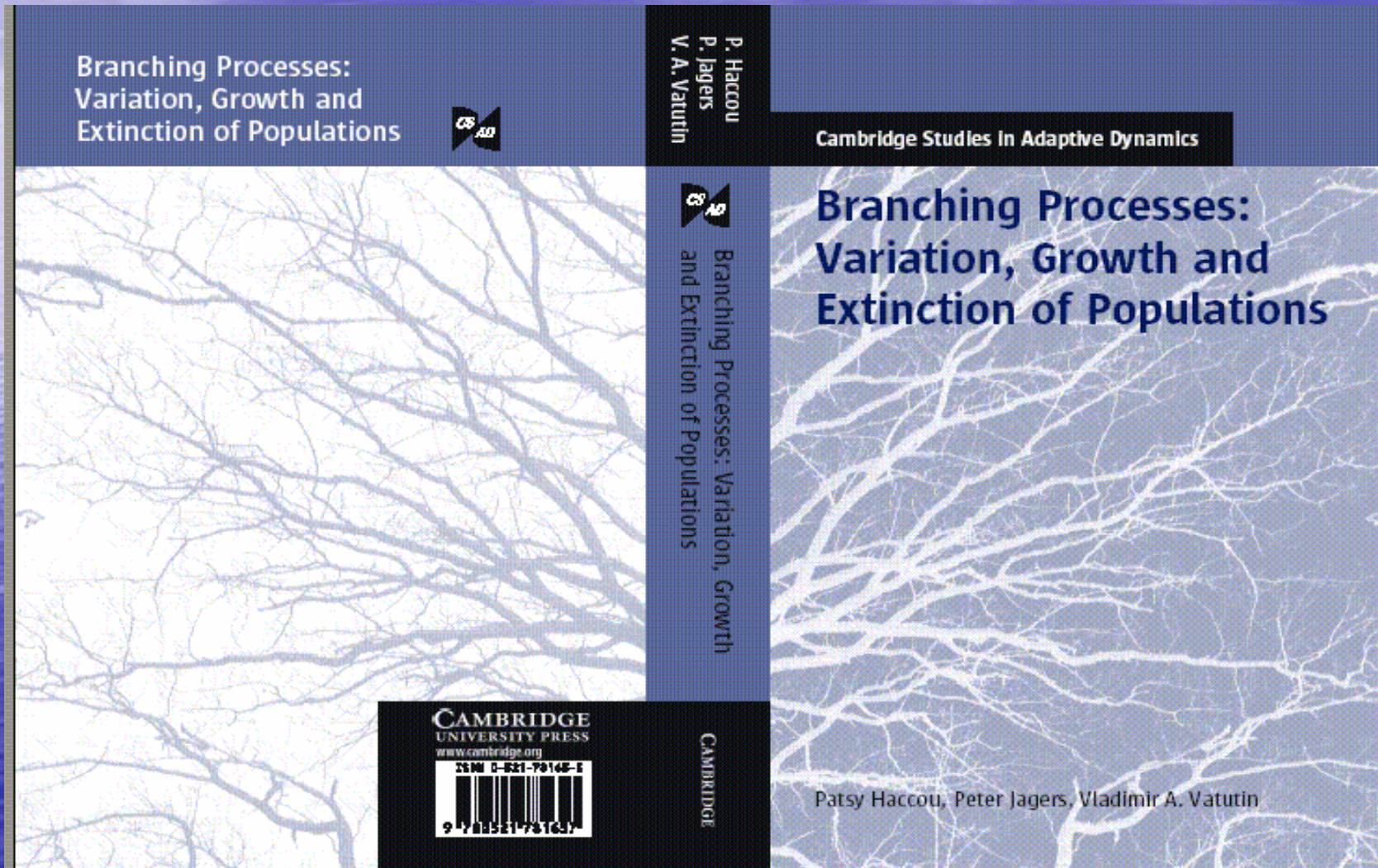
1. Introduction

Generalities about Branching and
Populations

What is a Branching Process?

- Mathematically, a random rooted tree (or forest, usually with nodes branching independently and often even i.i.d.)
- Historically,
 - born in a demographic and biological context, the extinction of family names, Galton, Fisher, Haldane (1850 – 1930);
 - maturing in nuclear physics: the cold war (Harris and Sevastyanov) (1945 – 1965);
 - turning into pure mathematics (Russian school, Dawson, Dynkin, Aldous....)
- But also finding use in computer science – and **population biology** again!

In a book store near you:



More Mathematical Books

- Harris, T. E., The Theory of Branching Processes (1963, recent reprint)
- Sevastyanov, B. A., Vetyashchiesya protsessy (1971 – also in German: Verzweigungsprozesse)
- Mode, C. J., Multitype Branching Processes (1971)
- Athreya, K. B. and Ney, P., Branching Processes (1972)
- Jagers, P., Branching Processes with Biological Applications (1975)
- Asmussen, S. and Hering, H., Branching Processes (1983)
- Guttorp, P., Statistical Inference for Branching Processes (1991)
- Athreya, K. B. and Jagers, P. (eds.), Classical and Modern Branching Processes (1997)
- Lyons, R. and Peres, Y. Probability on Trees and Networks. Under preparation. Manuscript downloadable from <http://mypage.iu.edu/~rdlyons/prbtree/prbtree.html>
- **Biological books:**
 - Taib, Z. , Branching Processes and Neutral Evolution (1992)
 - Kimmel, M. and Axelrod, D. E. , Branching Processes in Biology (2002)

What is a population?

- Originally, a group of humans = people.
- To us a set of **individuals**, who may generate new individuals, and/or disappear.
- Its **dynamics** is how the size and composition changes over time, and possibly stabilises.
- The **size** is usually the number of individuals – but might also be their DNA content or total body mass, or the total number ever alive.
- The **composition** might be over ages or types of individuals, or even family relations. (How probable is it to be first-born?)

The essence is that -

- - changes are initiated by individuals.
- In this sense they act **independently**.
- Literally: individuals live and give birth independently of each other (once born!).
- This is branching processes!
- But "initiated by individuals" does not preclude that individuals influence each other:
 - "Sibling" or "local" interaction
 - Population size dependence
 - Environmental variation

No sex!

- For simplicity: no mating - asexual reproduction. Or just think of the females.
- In small populations and for inheritance of traits mating does matter.
- Extinction risks and times are affected by inbreeding

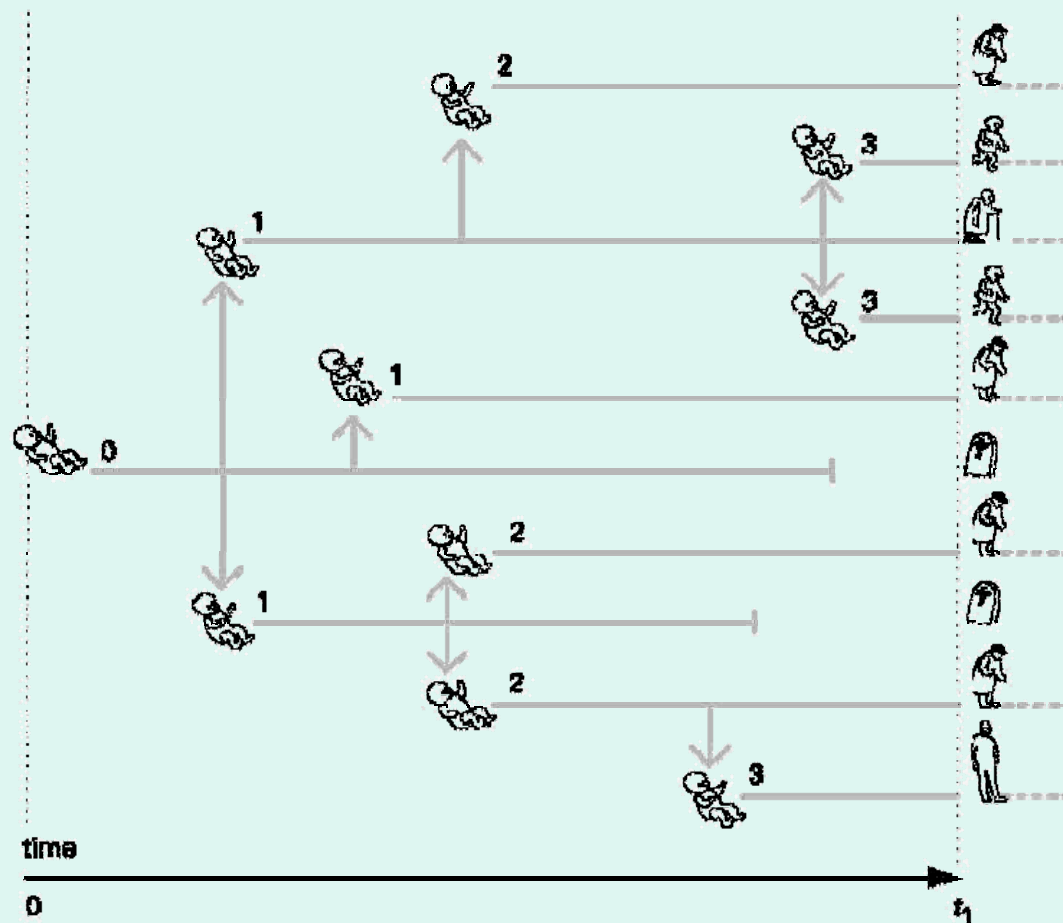


This leads to General Multi-type Branching Processes:

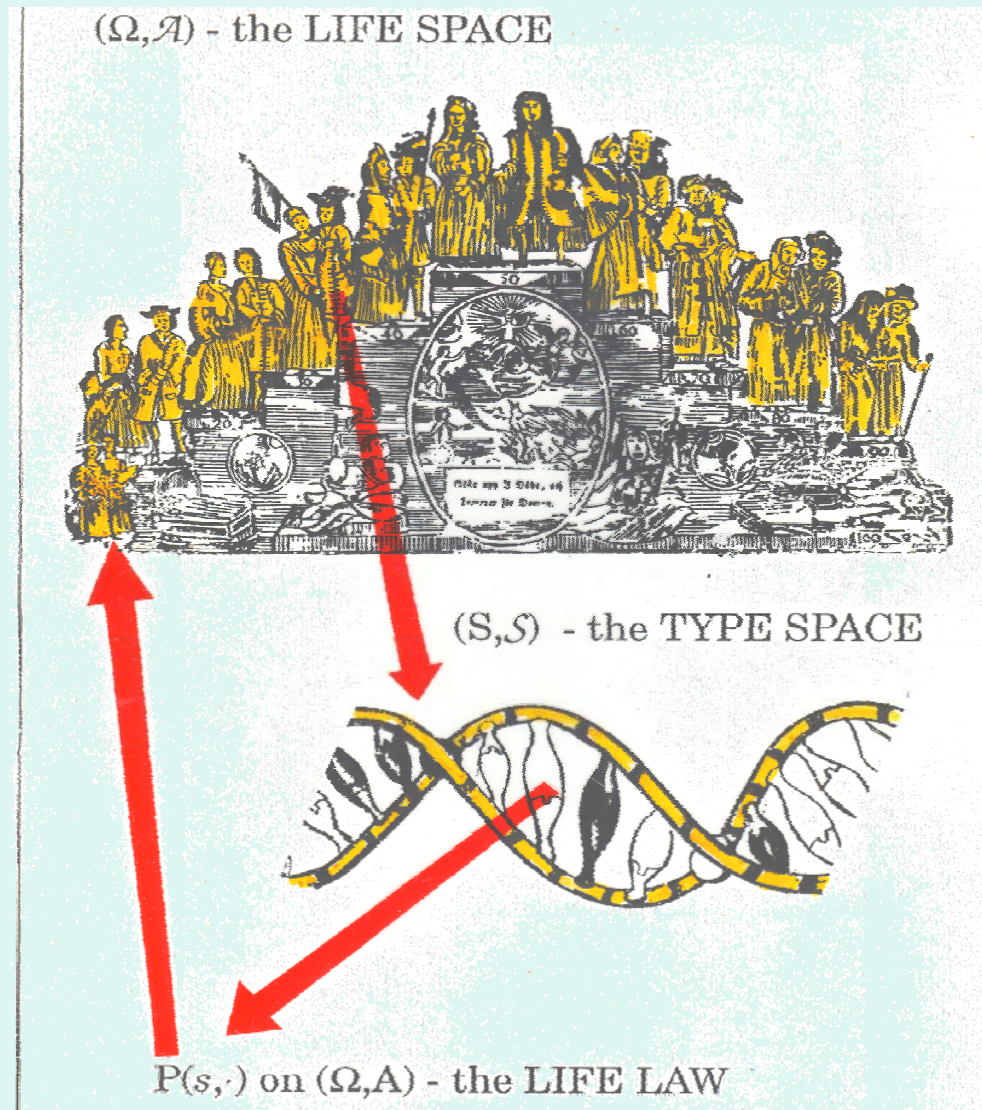
- An individual is characterized by her (geno)type.
- The type defines a probability measure over the set of possible life careers.
- A life career consists of a life span and a reproduction process.
- The life span is a non-negative random variable.
- The reproduction process is a point process, telling at what ages you get children of what types.

General Branching Processes

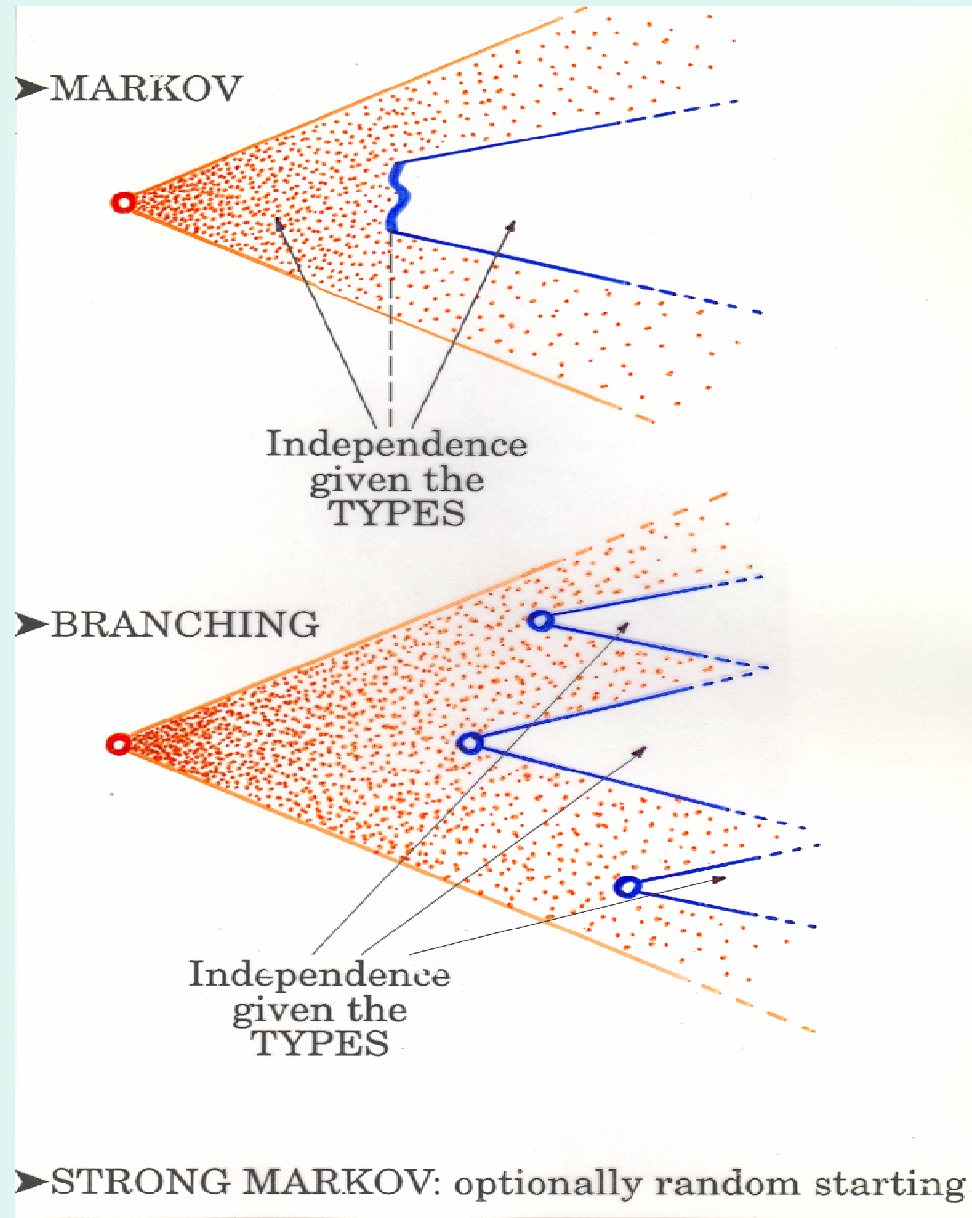
A process, initiated by one individual, might thus start like this, numbers denoting **generations**:



Markov structure, not in time but in the pedigree



Markov Branching Field



Questions

- How is the extinction probability determined from individual reproduction?
- And how the growth rate?
 - Can they be large simultaneously?
- What is the time to extinction (if it occurs)?
- Does population composition stabilise under growth?
 - Age and type distribution?
 - History (like mutational past)?
- Can population size stabilise?

Special Cases:

- Single type populations
- splitting
 - Bellman-Harris age-dependent branching
 - Sevastyanov
 - binary
- Markov branching (in real time)
 - Birth-and-death
 - Splitting Markov
- Discrete time
 - Overlapping generations
 - Non-overlapping: Bienaymé-Galton-Watson (generations or seasons)

2.(Single-type) Galton-Watson processes

- one-season plants or insects, or
- disregarding time structure and counting generations;
- all individuals of the same type.

Defining properties

- Each individual has the same life span (= 1).
- At its end it splits into a random number (in distribution = ξ).
- All individuals reproduce independently.
- Z_n = the number of individuals during season n (in generation n)
- Start from one ancestor, $Z_0 = 1$.

Reproduction law and generating function

- Write $p_k = P(\xi = k)$, $k=0, 1, 2, \dots$, the reproduction law ;
- $f(s) = \sum_{k=0}^{\infty} p_k s^k$, $0 \leq s \leq 1$; the reproduction generating function ;
- $m = E[\xi] = \sum k p_k = f'(1)$, the reproduction mean, $\sigma^2 = \text{Var}[\xi]$, the reproduction variance.
- The process is supercritical if $m > 1$, critical if $m = 1$, and subcritical if $m < 1$.

Benchmarks

- **Binary splitting:** $\xi = 0$ or 2 . The generating function becomes $f(s) = 1-p + ps^2$, $m = 2p$, $\sigma^2 = 4p(1-p)$.
- **Geometric:** $p_k = (1-p)p^k$, $k=0, 1, 2, \dots$;
 $f(s) = (1-p)/(1-ps)$, $m = p/(1-p)$.
- **Poisson:** $p_k = e^{-m} m^k/k!$, $f(s) = e^{m(s-1)}$.

To what extent are these biologically relevant?

Population development

- Let $\mathcal{B}_n = \sigma\{Z_0, Z_1, \dots, Z_n\}$ = the past.
- The GW process is **Markov**.
- Hence $E[Z_n | \mathcal{B}_{n-1}] = E[Z_n | Z_{n-1}] = mZ_{n-1}$,
and $E[Z_n] = m E[Z_{n-1}] = \dots = m^n$.
 - $\rightarrow \infty$ if $m > 1$,
 - $= 1$ if $m = 1$,
 - $\rightarrow 0$ if $m < 1$.

But how much does the expectation tell us?

$$\begin{aligned}\text{Var}[Z_n] &= \text{Var}[E[Z_n | \mathcal{B}_{n-1}]] + E[\text{Var}[Z_n | \mathcal{B}_{n-1}]] = \\ &= \text{Var}[E[Z_n | Z_{n-1}]] + E[\text{Var}[Z_n | Z_{n-1}]] = \\ &= \text{Var}[mZ_{n-1}] + E[\sigma^2 Z_{n-1}] = \\ &= m^2 \text{Var}[Z_{n-1}] + \sigma^2 m^{n-1} = \dots = \\ &= \sigma^2 m^{n-1} (m^n - 1) / (m - 1), \text{ if } m \neq 1, \text{ and} \\ &= \sigma^2 n, \text{ if } m = 1.\end{aligned}$$

Thus, the coefficient of variation (the ratio between the standard deviation and the expectation) explodes in the critical case, stabilises otherwise.

Tentative Conclusions

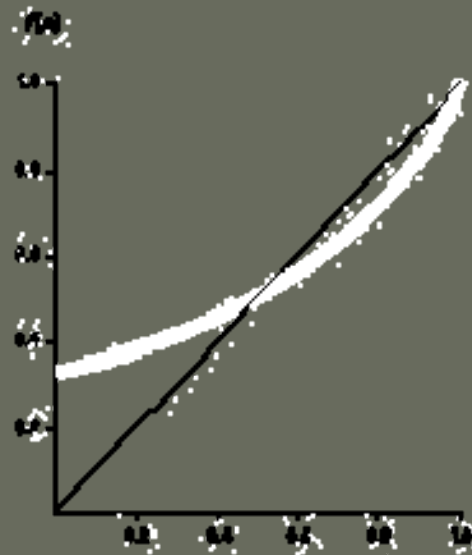
- If $m > 1$, exponential growth, like m^n .
(Malthus's Law, 1798; Euler 1748)
- If $m = 1$, ???
- If $m < 1$, **extinction**.
- But is this correct? And what happens in the critical case? And is the time structure relevant? Yes, for one season plants, no for generation counting.

Extinction

- A general branching process dies out if and only if its generation process does!
- Malthus 1798: in the city of Berne 379 out of 487 bourgeois families died out 1583-1783
- Bienaymé 1845, Galton 1873: (Problem 4001, Educational Times) "A large nation of whom we will only concern ourselves with the adult males, N in number, and who each bear separate surnames, colonise a district. Find (1) what proportion of surnames will have become extinct....."

The Extinction Probability

Write $q_n = P(Z_n=0)$. Then $q_{n+1} = \sum_k p_k q_n^k = f(q_n)$,
 $q_n \uparrow q = P(\text{extinction}); q = f(q)$:



The Extinction Theorem

The equation $f(q) = q$ has one solution in $[0, 1)$ if $m \leq 1$ and $p_0 < 1$, none if $m > 1$.

The smallest non-negative solution is the extinction probability. It is always true that $f(1) = 1$.

Watson's mistake: $f(1) = 1$ implies that ALL (branching) populations die out. Strange?

In the real world extinction is frequent!

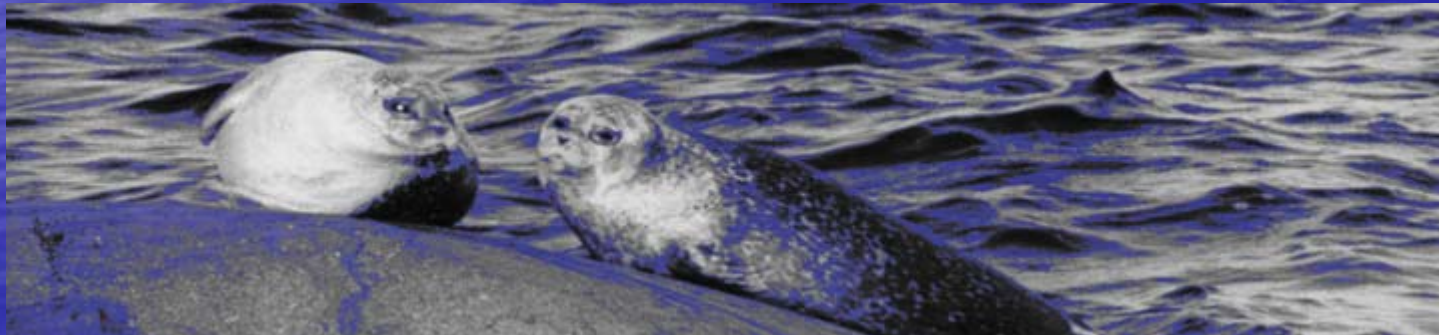
- Bienaymé and Galton considered extinction of (noble) family names.
- The paleontologist Raup (1991) claims that 99.99% (!!) of all species are extinct.
- For GW branching processes it is easy to obtain examples where m is large but the extinction probability q still close to 1.

Supercritical processes dying out are -- subcritical!

- With $Q := \{Z_n \rightarrow 0\}$ and p_{jk} the transition probability, $P(Z_{n+1} = j | Z_n = i, Q) = P(Z_{n+1} = k, Z_n = j, Q) / P(Z_n = j, Q) = P(Z_n = j) p_{jk} q^k / P(Z_n = j) q^i = p_{jk} q^{k-i}$.
- The Markov property can be checked similarly, and $\{Z_n | Q\}$ has the reproduction law $\{p_k q^{k-1}\}$.
- Strangely, this property extends to general processes! (Lagerås, 2007)

A Seal Example

- Survival probabilities for seals look like
0.6 0.8 0.95 0.95 0.95,
possibly slightly lower after 30 years or so.
- The first three years no children are born. The fourth year, the probability of a daughter is 0.2, and then it is 0.45 per year.
- The extinction probability for the family line from one female is $q=0.65$. The imbedded GW $m=3$, (The Malthusian parameter is $\alpha = 0.11$. Doubling time 6 years!)



But what if a population does not die out?

- The merciless dichotomy of population dynamics:
- Consider non-negative random variables X_1, X_2, \dots such that $X_n = 0 \Rightarrow X_{n+1} = 0$. Suppose that there is a history-independent risk of extinction: For any x there is a $\delta > 0$ such that $P(\exists n; X_n = 0 \mid X_1, \dots, X_k) \geq \delta$, if only $X_k \leq x$.
- Then, with probability one either there is an n such that all $X_k = 0$ for $k \geq n$ or $X_k \rightarrow \infty$ as $k \rightarrow \infty$.

Proof

- Lévy: $P(A|X_1, X_2, \dots, X_n) \rightarrow P(A|X_1, X_2, \dots)$, as $n \rightarrow \infty$.
- Hence, $A \in \sigma(X_1, X_2, \dots) \Rightarrow P(A|X_1, \dots, X_n) \rightarrow 1_A$.
- But $D = \{\exists n; X_n = 0\} \in \sigma(X_1, X_2, \dots)$.
- By assumption, $X_n \leq x$ infinitely often $\Rightarrow 0 < \delta \leq P(D|X_1, \dots, X_n) \rightarrow 1_D \Rightarrow 1_D = 1$
- Thus, $1 = P(X_n \rightarrow \infty) + P(\bigcup_{x \leq 1} \{X_n \leq x \text{ i.o.}\}) \leq P(X_n \rightarrow \infty) + P(D)$.

How quick is growth?

- $W_n := Z_n/m^n$. Then, $E[W_n | \mathcal{B}_{n-1}] = E[W_n | Z_{n-1}] = W_{n-1}$ and
- $\text{Var}[W_n] = \text{Var}[Z_n] / m^{2n} =$
 $= \sigma^2 (1 - m^{-n}) / (m - 1)m$, if $m \neq 1$,
- and $= \sigma^2 n$, if $m = 1$.
- Thus: If $m \leq 1$ (and $p_1 < 1$), $W_n \rightarrow 0$.
If $m > 1$ (and σ^2 is finite), $W_n \rightarrow$ some $W \geq 0$, a.s. and in mean square. $P(W=0) = q$,
 $E[W] = 1$, $\text{Var}[W] = \sigma^2 / (m - 1)m$.

"xlog x"

- $E[\xi \log_+ \xi] < \infty$ is the famous "xlog x" condition.
- In the supercritical case, "xlog x" iff $W_n \rightarrow W$ a.s. and in the mean. If "xlog x" does not hold $W_n \rightarrow 0$.
- For subcritical processes, "xlog x" iff $P(Z_n > 0) \sim m^n$. Further, $Z_n | Z_n > 0$ has a limit in distribution, as $n \rightarrow \infty$. The latter has a finite mean precisely under "xlog x".

Critical populations

- If $m = 1$ and $\sigma^2 < \infty$, then as $n \rightarrow \infty$
 - $P(Z_n > 0) \sim 2/(n\sigma^2)$, and
 - $Z_n/n | Z_n > 0 \rightarrow \text{Exp}(2/\sigma^2)$ in distribution.
- But note: this holds only if $m=1$ **exactly**.
The form of the other results are robust against small deviations.

A comment on multi-type GW

- A branching process is **multi-type** if individuals can be of finitely many types $i = 1, 2, \dots, d$. In the GW case, all still have life span one but the off-spring random variable ξ is replaced by a vector, whose distribution is determined by the mother's type.
- The rôle of m is taken by $M = (m_{ik}) = (E[\text{no of } k\text{-children to } i\text{-individual}])$.
- If **types communicate** (some $M^n \gg 0$), not much is changed.

A decomposable process

- Some populations where types do not communicate are of great interest, like:
- Mutation to survival (Iwasa-Haccou-Serra). Consider a single-type subcritical GW process, where children are supercritical mutants with a little probability, or become so after a number of mutations (cancer).
- What is the probability of non-extinction? The time to an out-break?
- In the conference S. Sagitov will tell more about such matters

3. General processes

- still (mostly) single-type
- but time enters: individuals can have different life spans and give birth at different ages
- dividing cells, animals with yearly litters.

Basics

- Each life-span is distributed like λ and reproduction like a point process ξ on \mathbb{R}_+ .
- Then $\xi(a)$ is the number of children born up to age a , $\xi(\infty)$ is the total off-spring, $m = E[\xi(\infty)]$, $\mu(a) := E[\xi(a)]$ is called the **reproduction measure**. (Usually $\xi(\lambda) = \xi(\infty)$, of course.)
- Assumptions: $E[\xi(0)] < 1$, $m < \infty$, and $\sigma^2 = \text{Var}[\xi(\infty)] < \infty$, mostly.

The Malthusian parameter

- Is there an intrinsic time-scale, doubling/halving time?
- A process is called **Malthusian** if there is a number α (the Malthusian parameter) with $\int_0^\infty e^{-\alpha t} \mu(dt) = E[\int_0^\infty e^{-\alpha t} \xi(dt)] = 1$.
- Supercritical and critical processes are always Malthusian. Only mathematicians can think of subcritical non-Malthusian populations!

3.1 Growth and Stabilisation

of general, single-type,
supercritical populations

Renewal theory and expected size

- The key renewal theorem: If $f(t) = h(t) + \int_0^t f(t-u)g(du)$,
 - h is "Riemann integrable", and
 - g is a "non-lattice" distribution function on \mathbb{R}_+ ,then $f(t) \rightarrow \int_0^\infty h(u)du / \int_0^\infty ug(du)$, as $t \rightarrow \infty$.
- But: $E[Z_t] = P(\lambda > t) + \int_0^t E[Z_{t-u}] \mu(du)$;
- $e^{-\alpha t} E[Z_t] = e^{-\alpha t} P(\lambda > t) + \int_0^t e^{-\alpha(t-u)} E[Z_{t-u}] e^{-\alpha u} \mu(du)$.
- Write $\beta := \int_0^\infty ue^{-\alpha u} \mu(du)$, "the mean age at child-bearing". Then, as $t \rightarrow \infty$,
- $E[Z_t] \sim e^{\alpha t} \int_0^\infty e^{-\alpha u} P(\lambda > u) du / \beta = e^{\alpha t} (1 - E[e^{-\alpha \lambda}]) / \alpha \beta$.

Once again: How close is the process to its expectation?

- Let \mathcal{B}_n be the σ -algebra generated by all lives of the n first generations (as before).
- $\text{Var}[Z_t] = \text{Var}[E[Z_t|\mathcal{B}_0]] + E[\text{Var}[Z_t|\mathcal{B}_0]] = \dots$
complicated renewal formulas, leading to $\text{Var}[Z_t] \sim \text{constant} \times e^{2\alpha t}$.
- So, like in the GW case, the coefficient of variation stabilises for supercritical populations
- Indeed, $e^{-\alpha t} Z_t \rightarrow \text{some } W, \text{ a.s. (and in mean square, if } \sigma^2 < \infty)$, as time passes. $P(W=0) = q$ under an "xlog x" condition.

Example: Splitting

- An individual can give birth only at death.
- Bellman-Harris: life-span (λ) and number of children (ξ) are independent.
- Benchmarks:
 - Binary splitting (cell proliferation):
 - Markov branching (\Leftrightarrow life is exponentially distributed \Leftrightarrow no aging);
 - Birth-and-death (can be viewed as splitting or not).

Bellman-Harris

- The reproduction law remains $\{p_k\}$, write L for the life span distribution. Then, $\mu = mL$, insertion into the renewal equation for the expected population size m_t is easy,
- $m_t = 1 - L(t) + m \int_0^t m_{t-u} L(du)$, and
- $v_t = \text{Var}[Z_t] = \text{Var}[E[Z_t | \mathcal{B}_0]] + E[\text{Var}[Z_t | \mathcal{B}_0]] = \text{Var}[1_{[0, \lambda)}(t) + mm_{t-\lambda}] + E[mv_{t-\lambda}] = L(t)(1 - L(t)) - 2m(1 - L(t)) \int_0^t m_{t-u} L(du) + m^2(\int_0^t m_{t-u}^2 L(du) - (\int_0^t m_{t-u} L(du))^2) + m \int_0^t v_{t-u} L(du)$.

Markov

- If $L(t) = 1 - e^{-at}$, then
- $m_t = e^{-at} + m \int_0^t m_{t-u} a e^{-au} du = e^{-at} + m e^{-at} \int_0^t m_u a e^{au} du.$
- $m_t' = (m-1)am_t$ and $m_t = e^{(m-1)at}$ exactly; $\alpha = a(m-1).$
- Differential equations also for second moments, and even the generating function of $Z_t.$

Measuring populations

- We have seen that the number of individuals alive grows exponentially. The expected proportionality constant was $(1 - E[e^{-\alpha \lambda}]) / \alpha \beta = E[\int_0^\infty e^{-\alpha u} 1_{[0, \lambda]}(u) du] / \beta = \int_0^\infty e^{-\alpha u} (1 - L(u)) du / \beta$.
- In this any individual aged u at the time of counting is given the weight $1_{[0, \lambda]}(u)$, λ being her life span.
- More generally, if $\chi(u)$ is the "size" at age u , and this is determined by the individual and her progeny, the population size thus measured will behave like
- $e^{\alpha t} E[\int_0^\infty e^{-\alpha u} \chi(u) du] / \beta = e^{\alpha t} \int_0^\infty e^{-\alpha u} E[\chi(u)] du / \beta$

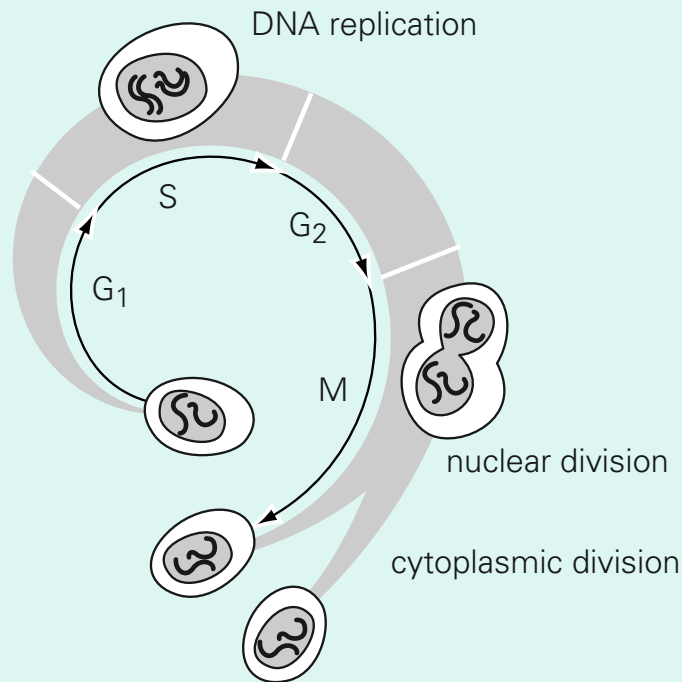
Asymptotic composition

- If Y_t is the χ -counted population, then $Y_t/Z_t \rightarrow \alpha E[\int_0^\infty e^{-\alpha u} \chi(u) du] / (1 - E[e^{-\alpha \lambda}])$, on the set of non-extinction.
- The stable age distribution: The choice $\chi(u) = 1_{[0, a \wedge \lambda]}(u)$ yields
- $\int_0^a e^{-\alpha u} (1 - L(u)) du / \int_0^\infty e^{-\alpha u} (1 - L(u)) du$.
- This age distribution is "stable". (Euler, 1760)

The probability of being first-born

- Let τ be the age at giving birth to the first-born (infinite for a childless individual), and Λ the child's life span.
- $1_{[0,\Lambda]}(u-\tau)$ counts the number of first-borns.
- Since Λ and τ are independent,
$$E\left[\int_{\tau}^{\tau+\Lambda} e^{-\alpha u} du\right] = E[e^{-\alpha\tau}](1-E[e^{-\alpha\Lambda}])/\alpha.$$
- Since life spans have the same distribution, the proportion of first-borns tends to $E[e^{-\alpha\tau}]$.
- The doubling time $\gg \tau \Rightarrow \alpha \tau \ll \ln 2 \Rightarrow E[e^{-\alpha\tau}] \gg 0.5$. "Most people are firstborn."

Cell proliferation



- The fraction of cells in mitosis (m.i.) is an important parameter, fig 3.4
- $2pE[e^{-\alpha T}] = 1$
- The characteristic counting the number of cells in mitosis is $1_{[T-M, T]}$, $M =$ duration of mitosis, $T =$ cycle time.
- If T and M are independent, $E[\int_0^\infty e^{-\alpha u} \chi(u) du] = E[e^{-\alpha T}] (E[e^{\alpha M}] - 1) \approx \alpha E[M]/2p$
- Hence, **m.i. $\approx \alpha E[M]/(2p-1)$.**

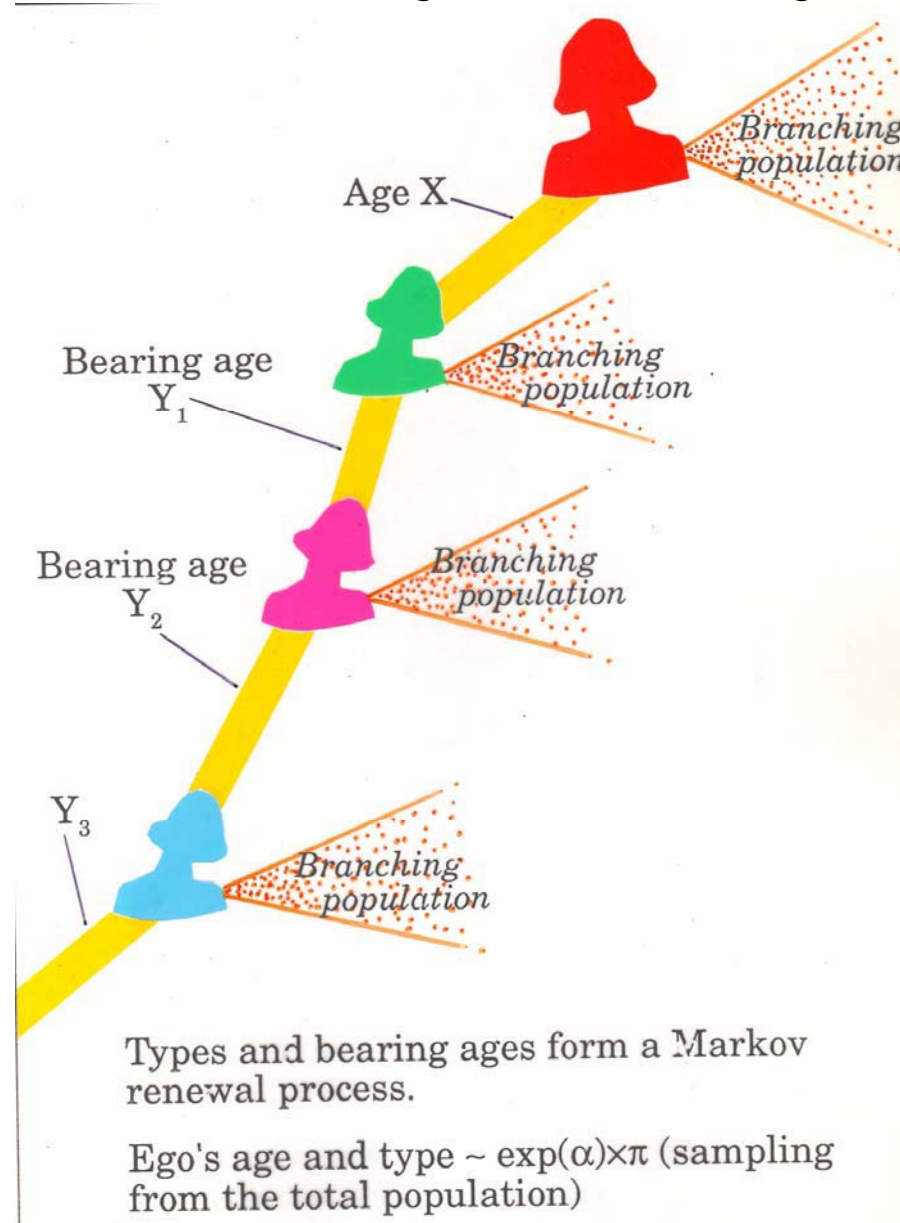
Doubling time

- What is the doubling time d in a cell tissue in balanced exponential growth?
- $e^{\alpha d} = 2$ tells us that $d = (\log 2)/\alpha$.
- Jensen: $-\alpha E[T] = \log e^{-\alpha E[T]} < \log E[e^{-\alpha T}] = -\log 2p$. Or $E[T] > d + (\log p)/\alpha$.
- "The cycle time $E[T]$ is larger than the doubling time, if there is no cell death."

Another interpretation: sample an individual ("Ego") at random.

- from the **accumulated** population (live or dead). Since the population grows $\sim e^{\alpha t}$, her age A is exponentially distributed with parameter α .
- from those alive, and you must condition upon being alive to get the stable age distribution:
- $P(A \leq a | \lambda > A) = \int_0^a \alpha e^{-\alpha u} (1 - L(u)) du / \int_0^\infty \alpha e^{-\alpha u} (1 - L(u)) du$.
- In the multi-type case, there is also a **stable type distribution**, π .

Family history



The renewal structure

- This is rather intuitive,
- but proving it in full generality is very technical.
- So we proceed to an application:

Mutational history

- In the one-type case the ancestral process is a renewal process, with intensity $1/\beta$, β = the generation time = the average age at child-bearing.
- In the neutral mutation, infinite alleles case the mutational history process remains of renewal type, the generation time replaced by the time between mutations: β/p , if p is the mutation probability.
- Thus, the evolution rate is independent of population size!

Molecular Clock of Evolution

- Independence of population size. But not of the time scale β (in general).
- However, if mutation risk is age-dependent, $p = p(a) = c \times a \sim 1 - e^{-ca}$, a = mother's age at bearing, this will also cancel, and:
- **Evolution rate = mutation rate = c** , a molecular clock of evolution, independently of both population size and reproductive mechanism.

Evolution viewed as branching or classically

- + General reproduction mechanism.
 - + No fixed population size.
 - Exponential growth for billions of years.
 - No sex.
- Simplistic time scale and reproduction
 - Fixed population size, since dawn of time.
 - Random mating

3.2 Extinction

How long does it take until a (large)
subcritical process dies out?

And what is the path to extinction?

Preliminaries

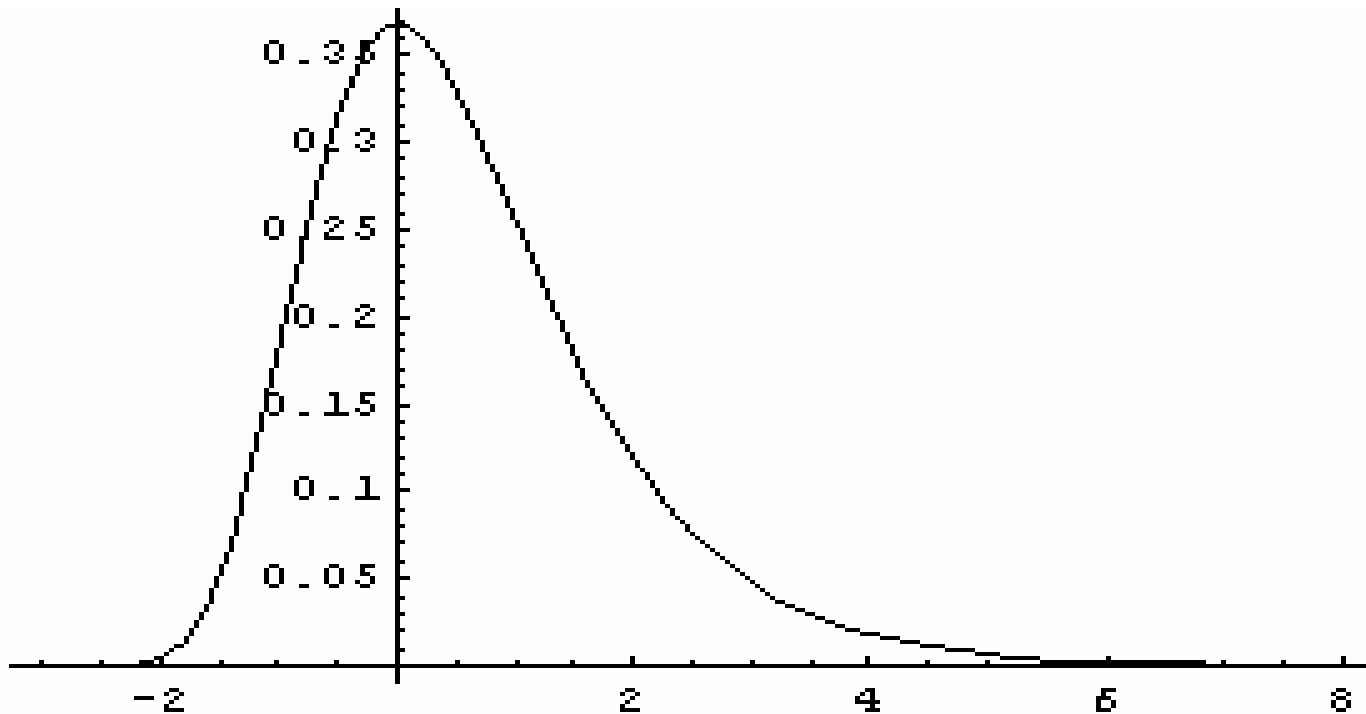
- General, non-lattice, single-type branching process Z_t^x starting from x , $Z_t = Z_t^1$,
- subcritical with the Malthusian parameter $-r$
- and having a finite $x \log x$ or second moment of reproduction (and technical assumptions).
- Then Yaglom's (generalised) theorem holds:
- $P(Z_t > 0) \sim ce^{-rt}$, $1 > c > 0$
- $c = C/b$, $C = \lim e^{rt} E[Z_t]$, $b = \lim E[Z_t | Z_t > 0]$
- Recall: A supercritical process bound to die out is a subcritical process!

Expected time to extinction

- $T_x = T = \inf \{t \geq 0; Z_t^x = 0\}$
- $P(T_x > t) = 1 - P(T_1 \leq t)^x = 1 - P(Z_t = 0)^x = 1 - (1 - c_t e^{-r t})^x$,
 $c_t \rightarrow c$.
- $E[T_x] = \int_0^\infty P(T_x > t) dt = (\ln x + \ln c + \gamma_x)/r$, $\gamma_x \rightarrow$ Euler's γ .
- $\gamma \approx 0.577$, $c = C/b$ often quite small ($C = 1$ in the Markov case, and if the population has not died out, then it is probably large). In the benchmark case of geometric reproduction in GW, it can be calculated explicitly to be $1 - m$.
- So if $x = 10\,000$ and $c = 0.2$, $\ln x = 9.2$, $\ln c = -1.5$, and γ may also matter! “Log of a large number is 10.”

The Actual Time

- $T_x = (\ln x + \ln c + \eta_x)/r$,
- Since $P(\eta_x < y)$ has an exponential tail,
 $P(\eta_x < y) \rightarrow \exp(-e^{-y})$, $y \in \mathbb{R}$, $x \rightarrow \infty$. Gumbel!

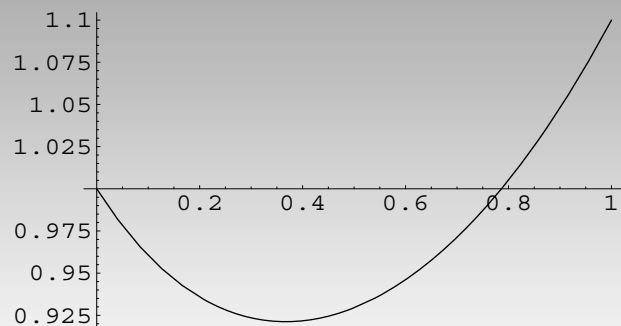


This can be used to calculate survival intervals:

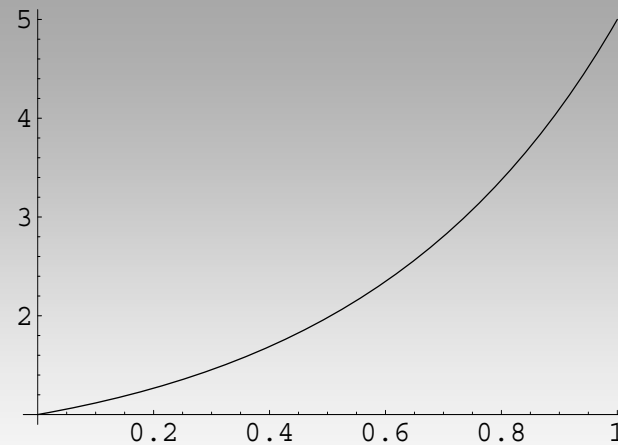
- Go back to the seal example, but with slightly (0.1) lower yearly survival probabilities,
0.5 0.85 0.85 0.85...
- The first three years no children are born. The fourth year, the probability of a daughter is 0.2, and then it is 0.45 per year, all as before.
- Then the population is subcritical, $m=0.784$, and $r = -\alpha = 0.018$.
- If $x = 10\ 000$ and $c = 1-m = 0.216$, $\ln x + \ln c = 7.7$.
- The Gumbel approximation yields 381 years as a 90% survival interval. (But, this is the lattice case.)

Between Dawn and Demise

- $x^{u-1}Z_{uT}^x \rightarrow C^{1-u}b^ue^{-u\eta}$, $0 < u < 1$, as $x \rightarrow \infty$, in distribution.
- Typically, the normed shape is thus exponentially decreasing, provided $\eta > \ln(b/C)$.
- The expected path is $C^{1-u}b^u\Gamma(u+1)$:



$b=1.1, C=1$

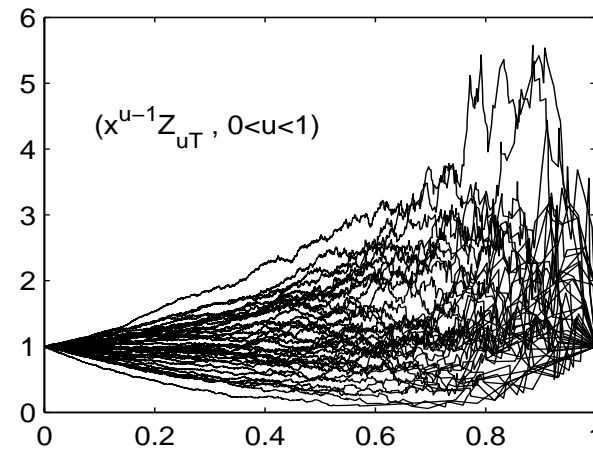
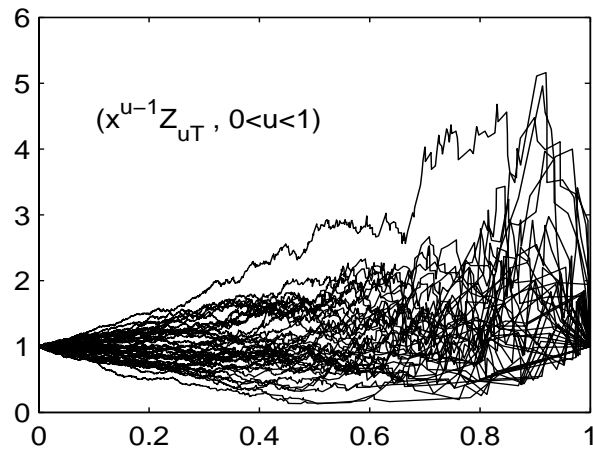


$b=5, C=1$

Steps of Proof

- We know that $T_x = (\ln x + \ln c + \eta_x)/r$.
- Therefore, check $x^{u-1} E[Z_{u(\ln x + t)/r}^x]$
 $\sim x^{u-1} x C x^{-u} e^{-ut} = C e^{-ut}$
- and $x^{2(u-1)} \text{Var}[Z_{u(\ln x + t)/r}^x] \rightarrow 0$.
- Under a boundedness condition on the intensity of births, the process $\{x^{u-1} Z_{u(\ln x + t)/r}^x ; t \in \mathbb{R}\}$ is tight.
- Hence, the random variables $\eta_x \xrightarrow{d} \eta (+ \ln c)$ can be inserted, and
- $x^{u-1} Z_{uT} \xrightarrow{d} C c^{-u} e^{-u\eta}$, in for fixed $0 < u < 1$.

And realisations:



Markov branching with $p_0 = 0.75$ and $p_2 = 0.25$, life expectancy = 1, $x = 1\ 000$ and $10\ 000$, respectively.

4. Dependence in Galton-Watson type processes

Random environments, sibling interaction, size dependence

Sibling dependence

- Can be analysed by a simple trick:
- Turn sibships into "macro individuals".
- These have independent reproduction and constitute a multi-type process with type = sibship size.
- Growth properties remain but extinction probability can be severely affected.

Random environments

- Each season provides its own "weather", determining the off-spring law. Given that, contemporary individuals reproduce i.i.d. Thus, consider $\{Z_n\}$ adapted to a filtration of environments $\{\mathcal{E}_n\}$.
- Given \mathcal{E}_n , Z_{n+1} is the sum of Z_n i.i.d. non-negative integer valued random variables each with conditional
 - mean M_n
 - and variance σ_n^2 .
- Then $Z_n/(M_1 \dots M_n)$, $n = 0, 1, \dots$ forms a martingale.

Growth in random environments

- Growth during the n seasons is determined by $M_1 \dots M_n$, and is equivalent to time homogeneous growth with conditional mean reproduction $(M_1 \dots M_n)^{1/n}$.
- $(M_1 \dots M_n)^{1/n} = \exp\{(\log M_1 + \dots + \log M_n)/n\}$.
- **If the M_i are i.i.d the law of large numbers holds:**
 $(M_1 \dots M_n)^{1/n} \rightarrow \exp\{E[\log M]\}$ and
- $Z_n \sim W \times M_1 \dots M_n \sim W \exp\{nE[\log M]\}$.
- **Explosion or extinction according as $E[\log M] >$ or < 0 .**
- This remains true for, say, stationary ergodic environments. (But there are also results for deteriorating ones etc.)
- Note: $E[M] = E[\exp(\log M)] > \exp(E[\log M])$ (Jensen).
- More about branching in random environments in V. Vatutin's lecture!

Population size dependence

- Now identify environment with one particular (but highly non-stationary) aspect: population size, so that $Z_{n+1} | \epsilon_n = Z_{n+1} | Z_n$.
- Consider **near-critical** processes:
- $M_n = 1 + C_n/Z_n + R_n$, C_n and R_n in ϵ_n .
- Then, $E[Z_{n+1} | \epsilon_n] = Z_n + C_n + Z_n R_n$.
- Assume:
 - C_n independent of Z_n and
 - $R_n = o(1/Z_n)$,
- and write $c_n = E[C_n]$.

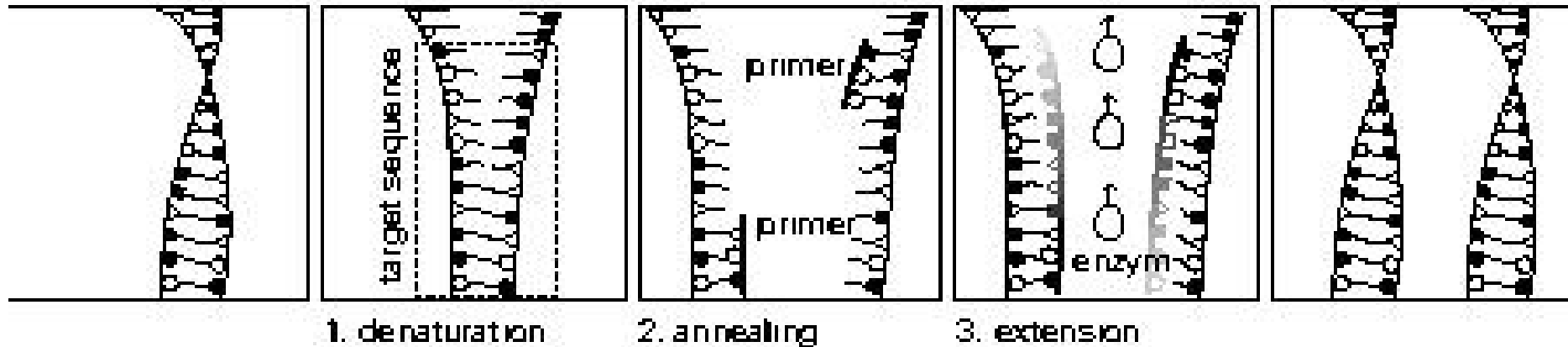
Must growth be exponential?

- Recall that the true dichotomy is between extinction and unbounded growth.
- Summing the expectation of $E[Z_k | \epsilon_{k-1}] = Z_{k-1} + C_{k-1} + Z_{k-1} R_{k-1}$ yields
- $E[Z_n] = E[Z_0] + \sum_{k=0}^n c_k P(Z_k > 0) + o(\sum_{k=0}^n c_k P(Z_k > 0))$.
- If $q = \lim P(Z_n = 0)$, $E[Z_n] - E[Z_0] \approx (1-q) \sum_{k=0}^n c_k := (1-q)L_n = (1-q)qcn$, if all $c_k = c$.
- $Z_n \sim (1-q)L_n$? This depends on the reproduction variances, σ_n^2 !

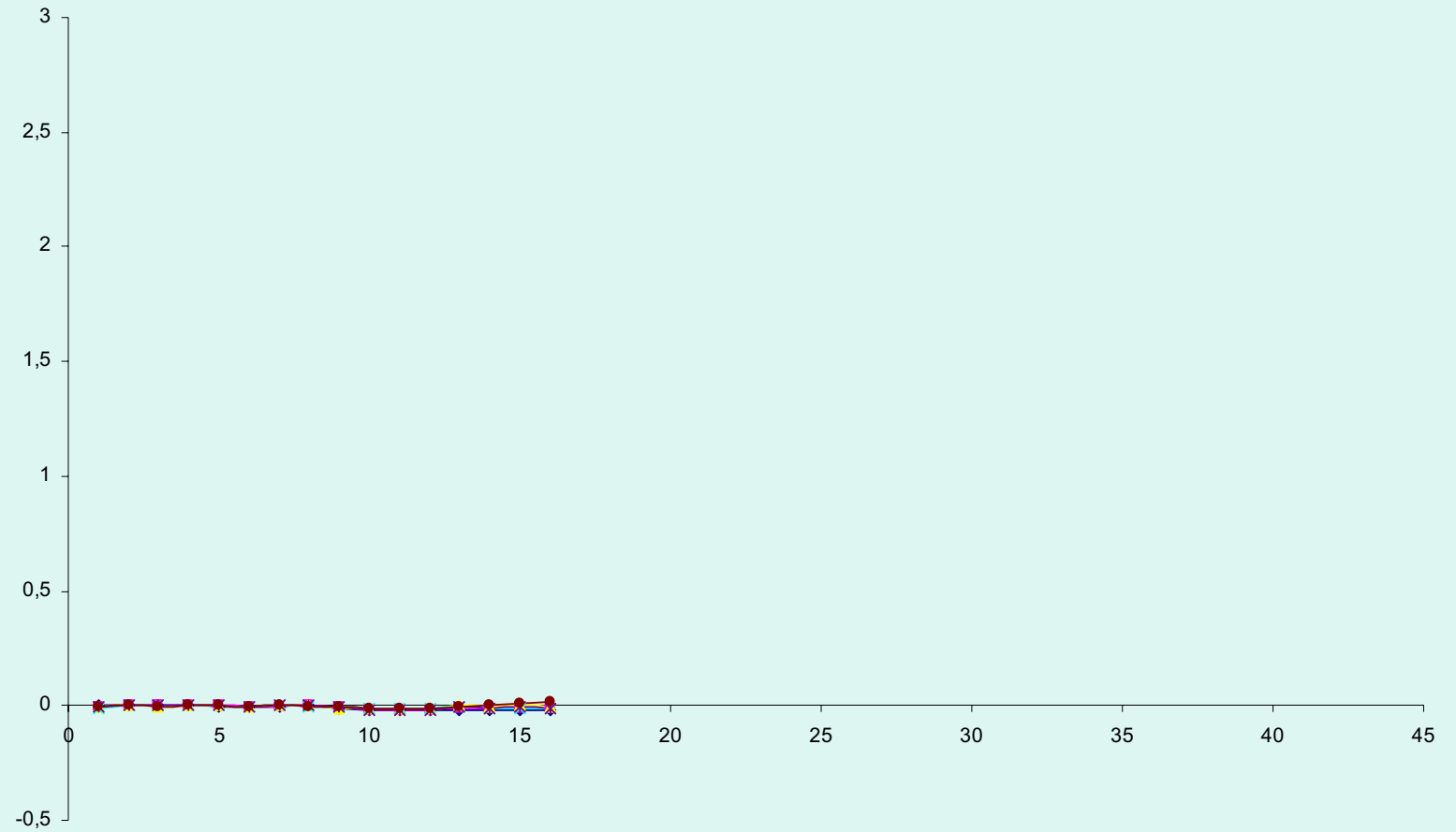
Process convergence

- Assume that $L_n \rightarrow \text{infinity}$ and $L_n / L_{n-1} \rightarrow 1$,
- that all conditional offspring moments are bounded,
- that conditional reproduction variances stabilize in large populations: $\sigma_n = V_n + o(1)$, as Z_n tends to infinity,
- and with $v_n = E[V_n]$, $v_n \sim ac_n$, $a \geq 0$, as $n \rightarrow \text{infinity}$.
- Then, it can be shown that $Z_n / L_n \rightarrow^d \Gamma(2/a, 2/a)$, if $a > 0$, and $\rightarrow 1$, if $a = 0$, provided the population does not die out.

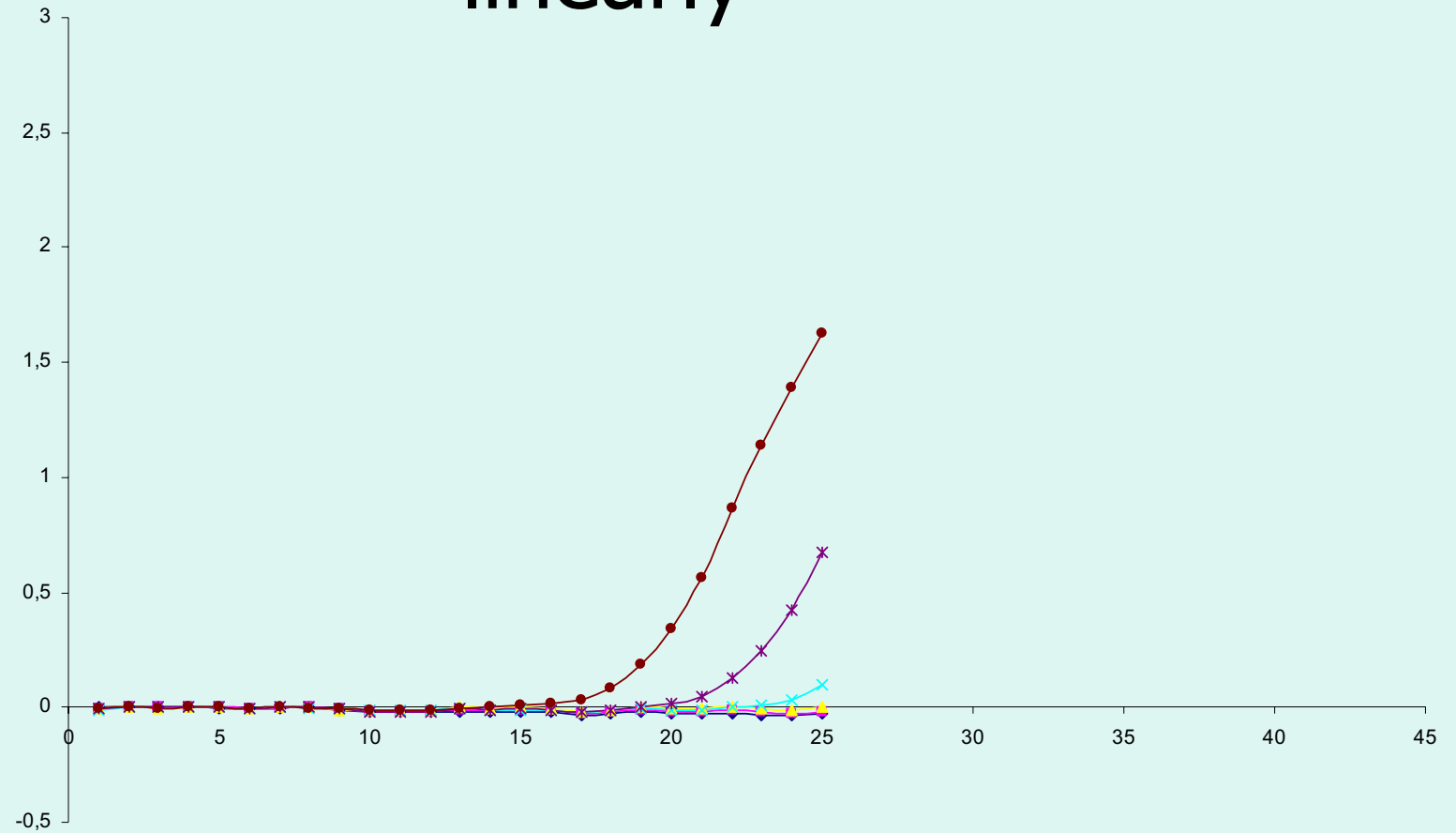
The Polymerase Chain Reaction (PCR)



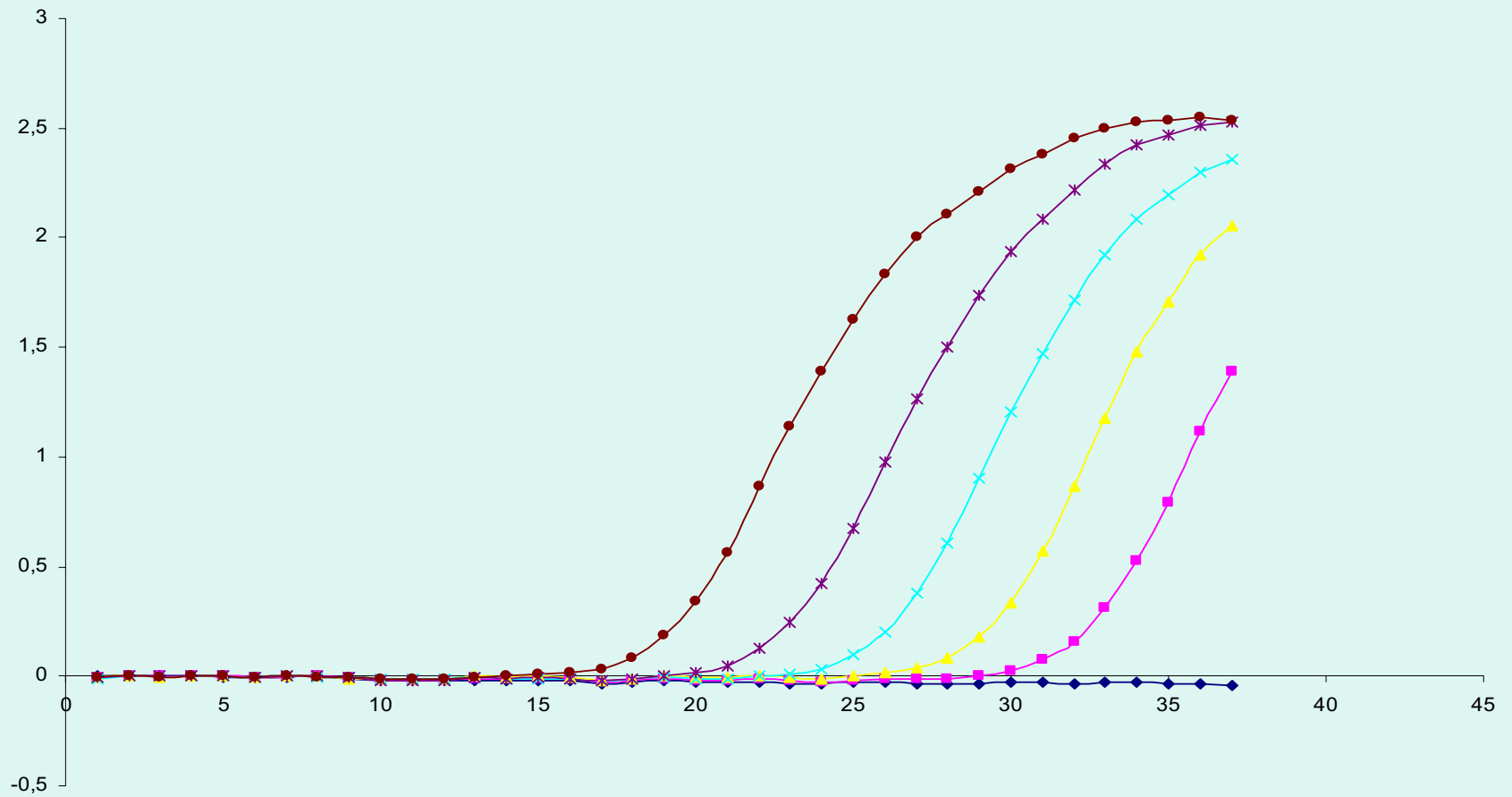
In the beginning you can't see much....



But soon the number starts to grow ---
linearly



But everything has an end...



Why?

- The probability of successful replication = $P(\text{two "children"}) = p(\text{DNA, primers, polymerase, deoxynucleic triphosphate, MgCl}_2, \dots) = K/(K+Z)$, $Z=\text{DNA}$, $K=\text{random or constant (usually)}$.
"Michaelis-Menten kinetics".
- If replication does not succeed, the molecule remains = one "child".
- Mean reproduction: $M_n = 1 + K_n/(K_n+Z_n) = 1 + K_n/Z_n + o(1/Z_n)$, K_n i.i.d., Z_n no. molecules in cycle n .
- Mean variance: $\sigma_n^2 = K_n Z_n / (K_n + Z_n)^2$.

Then:

- With $c_n = E[K_n] = c$, $L_n = cn$, and $\sigma_n \rightarrow 0$, as Z_n grows.
- Hence, $Z_n/(cn) \rightarrow 1$. Linear growth, in probability and actually a.s.!
- Directly: $E[Z_n|Z_{n-1}] = Z_{n-1} + KZ_{n-1}/(K+Z_{n-1}) \approx 2Z_{n-1}$, as long as $Z_{n-1} \ll K$.
- But $Z_n \rightarrow \infty$, and rather $E[Z_n|Z_{n-1}] \approx Z_{n-1} + K$.
- By dominated convergence indeed
- $E[Z_n] = E[Z_{n-1}] + E[KZ_{n-1}/(K+Z_{n-1})] = \dots$
 $E[Z_0] + \sum_{k=0}^{n-1} E[KZ_k/(K+Z_k)] \sim Kn, n \rightarrow \infty$.

And the variances:

- With z molecules, the replication variance is $\sigma^2(z) = 4p(z) + 1 - p(z) - (1 + p(z))^2 = p(z)(1 - p(z)) = Kz / (K + z)^2$, $p(z) = K / (K + z)$.
- $\text{Var}[Z_n] = E[\text{Var}[Z_n | Z_{n-1}]] + \text{Var}[E[Z_n | Z_{n-1}]] = E[KZ_{n-1}^2 / (K + Z_{n-1})^2] + \text{Var}[Z_{n-1} + KZ_{n-1} / (K + Z_{n-1})] \sim K + \text{Var}[Z_{n-1} + K] = K + \text{Var}[Z_{n-1}] = Kn + \text{Var}[Z_0]$.
- $E[(Z_n/n - K)^2] = \text{Var}[Z_n/n] + (E[Z_n/n] - K)^2 \rightarrow 0$.
- Actually, $Z_n/n \rightarrow K$ also a.s.

But many areas remain....

- How does inbreeding affect the time to extinction?
- Can we model inbreeding at all? Other mating effects?
- Can branching help in understand biological evolution?
- E.g., is sympatric speciation possible?
- And many more questions.

Thank you for your
attention

and Good Luck!