Statistical Inference in a Two-Compartment Model for Hematopoiesis

Sandra N. Catlin, Janis L. Abkowitz, Peter Guttorp

STAT 518 Jason Xu

April 10, 2013

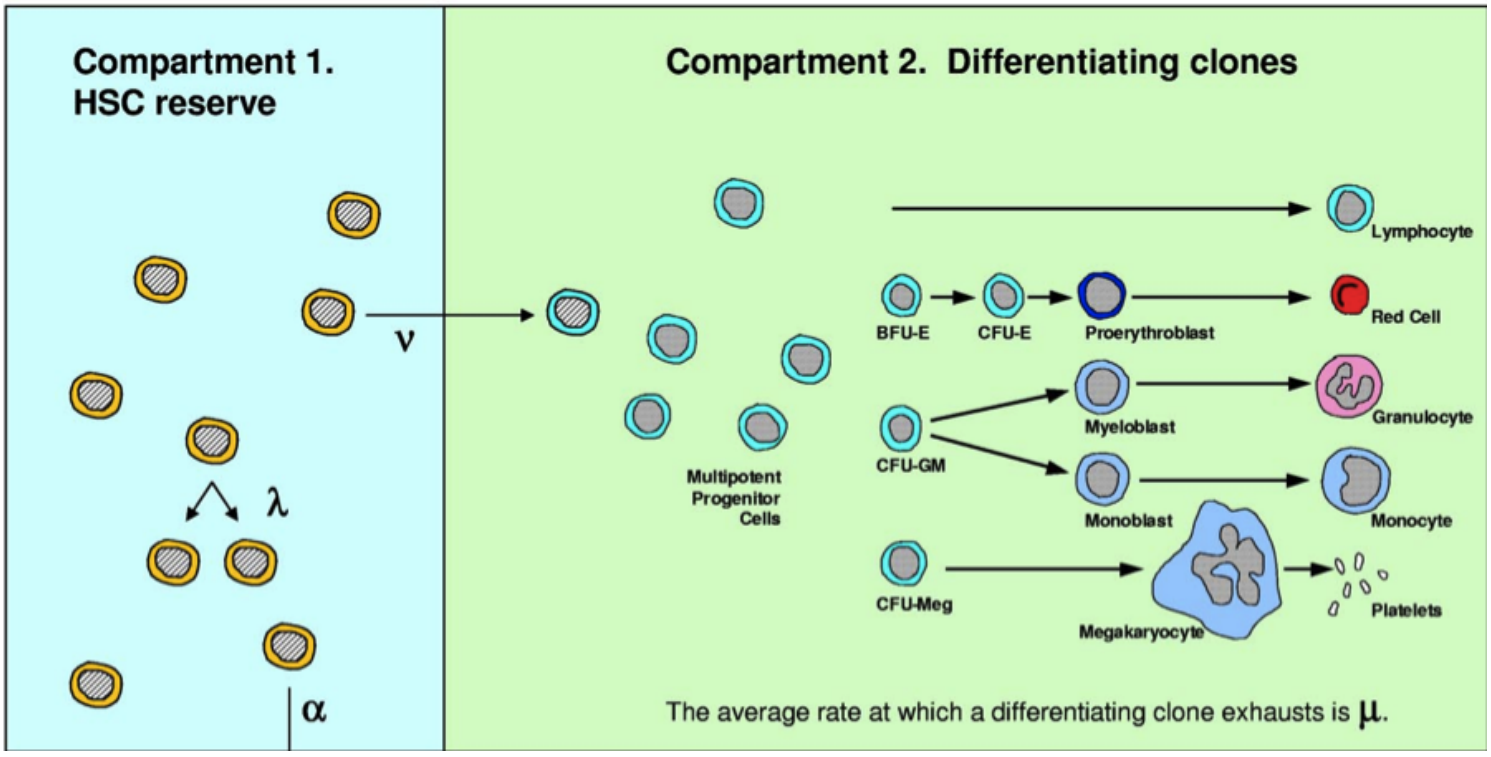
What is Hematopoiesis?

Hematopoeisis: Process of specialization of stem cells into mature blood cells

- HSCs differentiate (specialize) into progenitor cells: multi-stage process
- Progenitor cells further differentiate to white/red blood cells, platelets, etc. This is well-studied.
- Little is known about early stages: unidentifiability of HSCs

A Stochastic Model

- First birth-death model for hematopoeisis: Till et al, 1963
- Experimentally justified, refined over several studies
- Current paper analyzes hidden two-compartment model



A Stochastic Model

Goal: Develop inferential tools for this problem, and for a class of stochastic population processes

Statistical motivation: Tools for inference in a useful class of models. Hidden compartmental processes include

- SIR models
- Spread of malaria in human host (Gravenor 1998)

Application: Clinical and biological importance

- Cancer therapy: stem cell transplantation
- Gene therapy
- "A remarkable cell renewal process"; close to 1 trillion cells per day supported by HSCs (M. Ogawa)

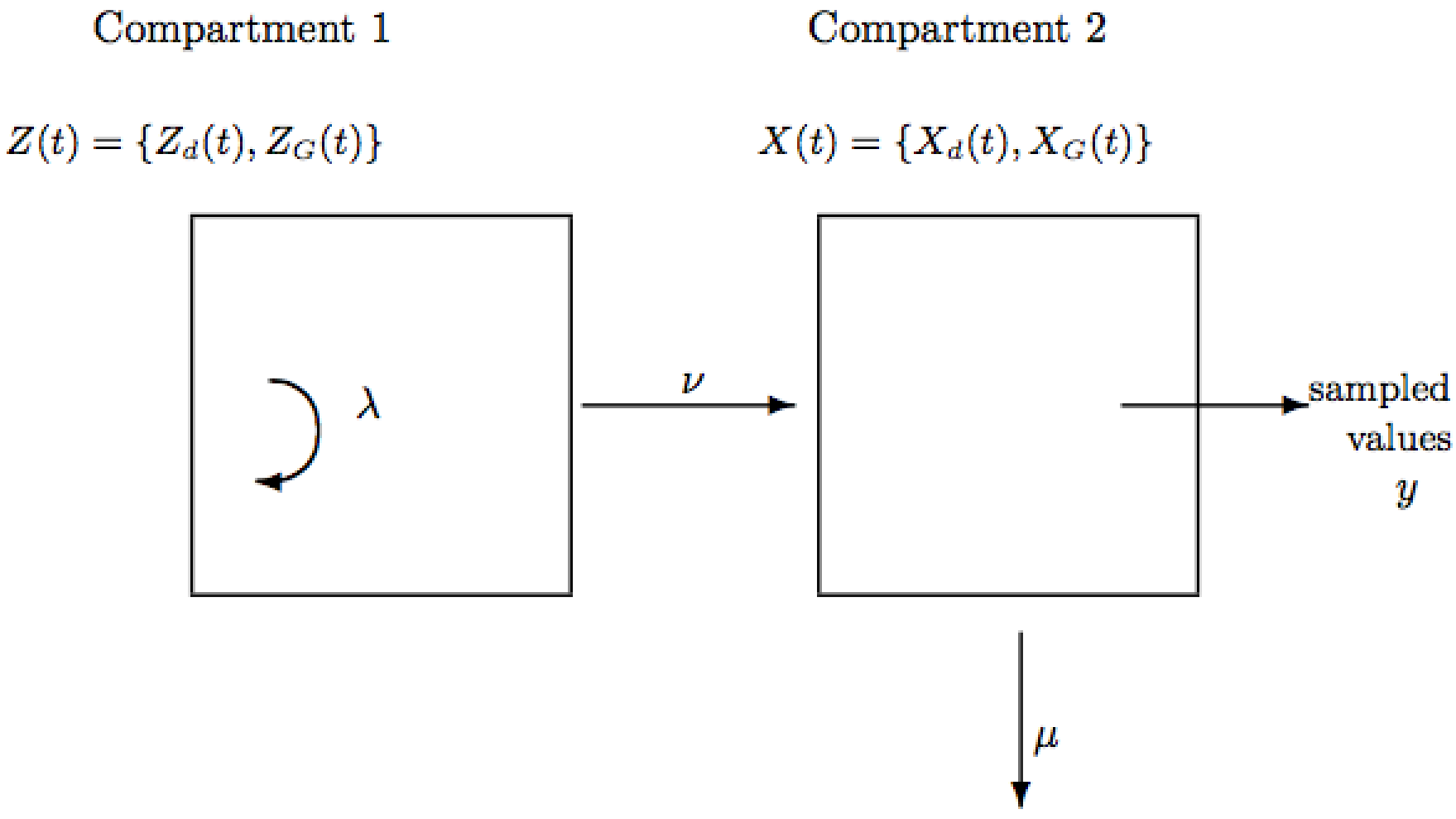
Experimental design

Female safari cat study

- Distinct G6PD phenotype expressed as d or G
- Retained after replication/differentiation; neutral
- Provides binary marker of each cell and its clones

Observing proportion of, say d, allows us to "track" HSC behavior

The model


Compartment 1
Compartment 2
$Z(t) = \{Z_d(t), Z_G(t)\}$
$X(t) = \{X_d(t), X_G(t)\}$
λ
ν
sampled values y
μ

The model

Simple continuous time, discrete state process:

- Compartment 1 is a linear birth-death (BD) process
- Compartment 2 is a non-homogeneous immigration-death process
- Inference: rates λ, ν, μ

Likelihood: $L(\lambda, \nu, \mu) \propto \lambda^{B_T} \nu^{E_T} \mu^{D_T} \exp(-(\lambda + \nu) S_T^z - \mu S_T^x)$

- $B_T =$ births, $E_T =$ emigrations, $D_T =$ deaths, $S_T^i =$ total time in i
- MLEs available: $\hat{\lambda} = B_T / S_T^z$, $\hat{\nu} = E_T / S_T^Z$, $\hat{\mu} = D_T / S_T^x$; nice asymptotic properties

Difficulty: Partial Observations

We only have sampled values from the second compartment: $Y(t)$, the total cells marked d, is a hidden Markov process

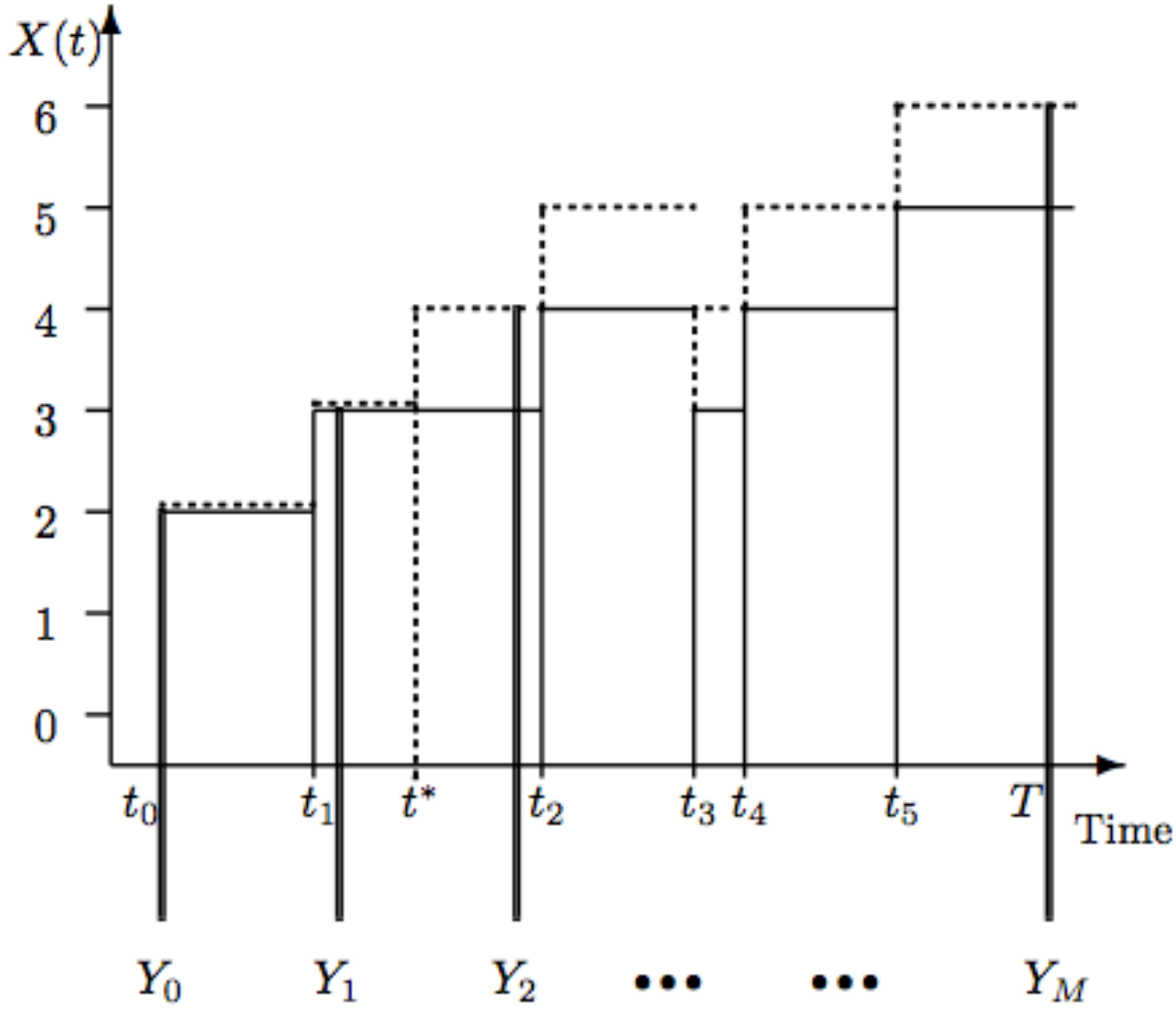
$$[Y(t)|(x(t), z(t))] \sim Binom(N_t, \frac{x_d(t)}{x_d(t) + x_G(t)})$$

- Distribution of this binomial proportion mathematically difficult
- Exact likelihood methods infeasible
- No successful attempts in obtaining transition probabilities

Other Approaches

- Abkowitz (1996): vary parameters and simulate realizations: compare simulations to true data
- Catlin (1997): normal approximation of transition probabilities
- Golinelli (2006): Bayesian inference via RJMCMC
  - Integrate over paths between discrete observations
  - Most precise estimates and effective use of data at computational cost

Other Approaches



Current Method

Estimating equation approach

- Calculate moments of process by solving Kolmogorov forward equation
- Create estimating function relating these expressions and data
  - Method of moments cannot be used directly: differing population sizes over realizations at given time
- Solve using nonlinear least squares

Discussion

- Simulations starting with estimated rates close to observed data
- Parameter estimates very similar to results from other studies
- Minor discrepancies: theoretical and simulated errors
- Advantages: not restricted to large population sizes
  - Accurate parameter estimates without much computational cost
  - Provides standard error estimates
- Drawbacks: does not utilize all data efficiently
  - Dependent on number of realizations
  - Biological shortcomings

Discussion

Closing remarks: While not able to make as efficient use of data as stochastic integration methods, provides a more "elegant" solution that is accurate and applicable when MCMC methods become infeasible.

Studying hematopoiesis via two-compartment stochastic model has provided much insight to understanding the complex behavior of HSCs.