Cancer Causes Stable Laws

Rick Durrett

...

A report on joint work

with Stephen Moseley (Cornell → consulting in Boston)

with Jasmine Foo, Kevin Leder (Minnesota),

Franziska Michor (Dana Farber Cancer Institute),

and John Mayberry (Cornell postdoc → U. of the Pacific),

and with Kaveh Danesh (Duke Undergraduate)

Laura Havrilesky and Evan Myers

(Obstetrics and Gynecology, Duke Medical Center).

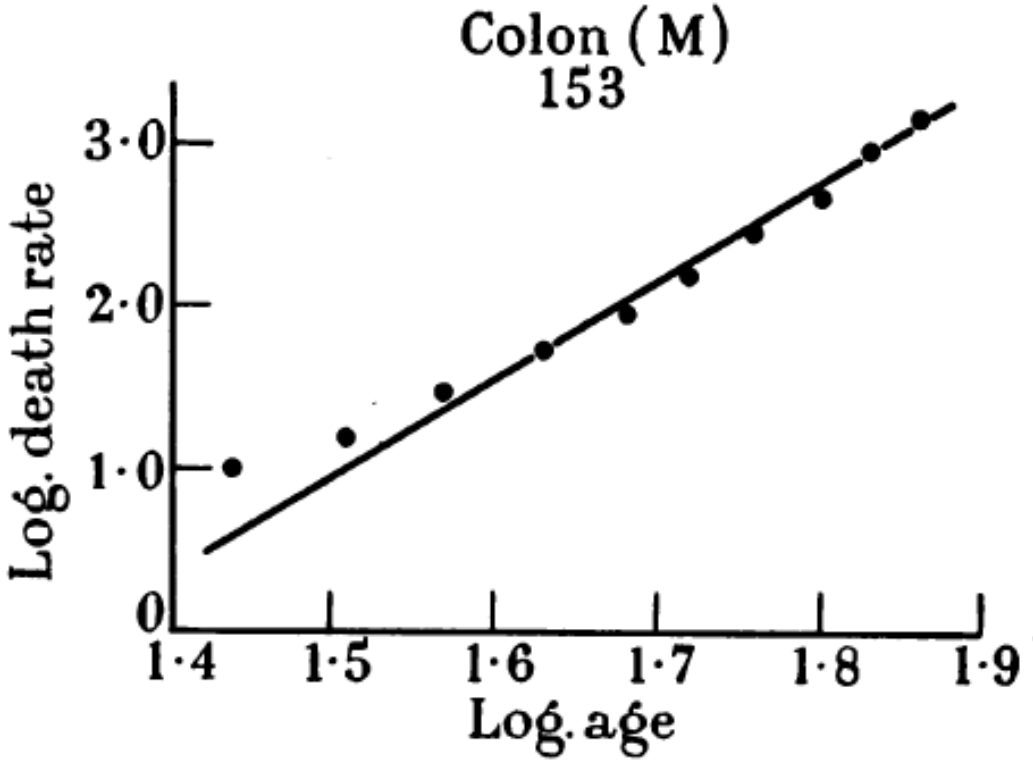
Stan Ulam once said:

"I have sunk so low that my last paper contained numbers with decimal points."

Figure: Feynman, Ulam, von Neumann

Armitage and Doll (1954)

Noticed that log-log plots of cancer incidence data are linear for a large number of cancer types; for example, colorectal cancer incidence has a slope of 5.18 in men and 4.97 in women.



Multi-stage theory of carcinogenesis

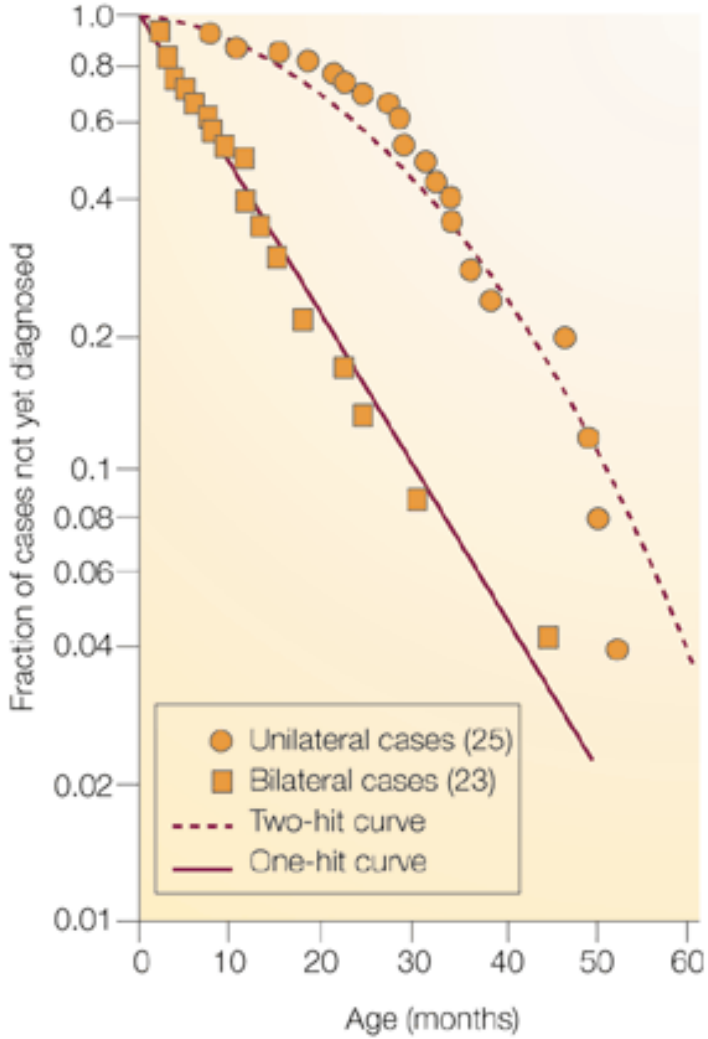
Armitage and Doll (1954) use the observation that the slopes were 5.18 in men and 4.97 in women to argue that colon cancer is a six stage process. The math was very simple

Suppose X_i are independent and have an exponential distribution with rates u_i. The sum $X_1 + \cdots + X_k$ has a density function that is asymptotically

$$u_1 \cdots u_k \frac{t^{k-1}}{(k-1)!} \qquad \text{as } t \to 0$$

Incidence of Retinoblastoma

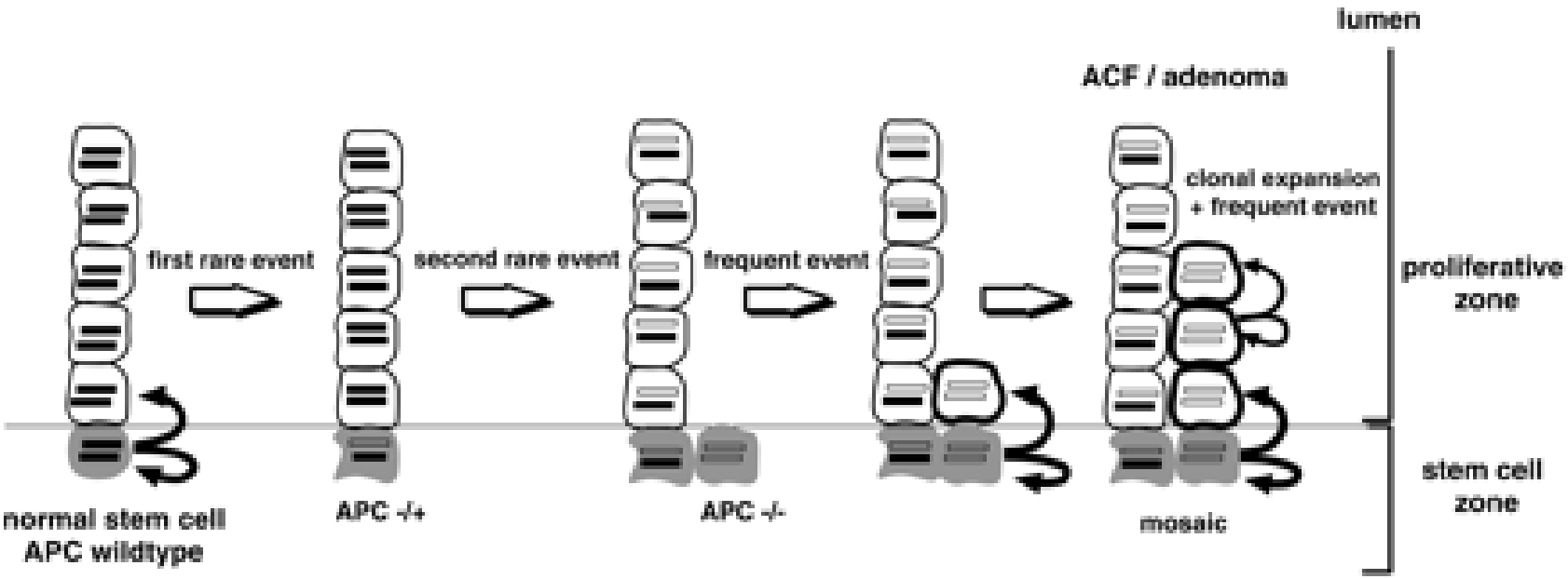
Knudson's two hit hypothesis $\longrightarrow$ tumor-suppressor genes



Nature Reviews | Cancer

Progression to Colon Cancer

Luebeck and Moolgavakar (2002) PNAS fit a four stage model to incidence of colon cancer by age.



What are the stages ?

- In sporadic cases of colon cancer the first two stages are inactivation of the tumor suppressor gene *APC* adenomatous polyposis coli.
- KRAS is an **oncogene** (one mutation turns it on). Once it is turned on it recruits and activates proteins necessary for the propagation of growth factor
- The final stage is thought to involve the inactivation of *TP*53 the gene which makes *p*53. Mutant p53 can no longer bind DNA in an effective way, and as a consequence the p21 protein whose production it stimulates is not made available to act as the 'stop signal' for cell division.

Things are not so simple

- In 20% of colon cancers, *APC* is not mutated but instead the oncogene β-catenin (in the same pathway) is. This and other examples suggest that features of the disease are due to disrupting certain molecular pathways not necessarily specific mutations.
- One of the main aims of large scale sequencing of cancer tumors is to find mutations that are potential drug targets. However many statistically significant mutations are "passengers" that occurred on the same chromosome with a causative mutation.
- Even when mutations are declared to be causative on statistical grounds, the tumor subtypes they define do not correlate well with the behavior of the disease and its response to treatment.

Multitype Markovian binary branching process

$Z_i(t) \equiv$ is the number of type i cells

Type i cell give birth at rate a_i and die at rate b_i.

Yes we have deaths at rate b.

Type i cells produce offspring of type $i+1$ at rate u_{i+1}.

Type 0's are a branching process

Birth at rate a_0, death at rate b_0, $\lambda_0 = a_0 - b_0$.

$$P(Z_0(t) = 0 \text{ for some } t \geq 0) = b_0/a_0$$

As $t \to \infty$, $e^{-\lambda_0 t} Z_0(t) \to W_0$ a.s.

$$W_0 =_d \frac{b_0}{a_0}\delta_0 + \frac{\lambda_0}{a_0} \text{exponential}(\lambda_0/a_0)$$

exponential(r) has density re^{-rt}, mean $1/r$.

If we condition on nonextinction the limit is exponential(λ_0/a_0).

Type 1's: Durrett and Moseley (2009)

$M_t = e^{-\lambda_1 t} Z_1(t) - \int_0^t u_1 e^{-\lambda_1 s} Z_0(s)\, ds$ is a martingale.

Theorem 2. As $t \to \infty$, $e^{-\lambda_1 t} Z_1(t) \to W_1$ a.s. with

$$EW_1 = u_1/(\lambda_1 - \lambda_0)$$

Type 1's: Durrett and Moseley (2009)

$M_t = e^{-\lambda_1 t} Z_1(t) - \int_0^t u_1 e^{-\lambda_1 s} Z_0(s)\, ds$ is a martingale.

Theorem 2. As $t \to \infty$, $e^{-\lambda_1 t} Z_1(t) \to W_1$ a.s. with

$$EW_1 = u_1/(\lambda_1 - \lambda_0)$$

Let $Z_1^*(t)$ be the number of 1's when $Z_0^*(t) = V_0 e^{\lambda_0 t}$, $t \in (-\infty, \infty)$.

Theorem 3. As $t \to \infty$, $e^{-\lambda_1 t} Z_1^*(t) \to V_1$ a.s. with

$$E(\exp(-\theta V_1)|V_0) = \exp(-c_{h,1} u_1 V_0 \theta^{\alpha_1})$$

and $\alpha_1 = \lambda_0/\lambda_1$. $(V_1|V_0)$ is one sided stable law with index $\alpha_1 \in (0,1)$.

Growth rate of type k's

Suppose $Z_0^*(t) = V_0 e^{\lambda_0 t}$ for $t \in (-\infty, \infty)$ where V_0 is exponential(λ_0/a_0).

$$e^{-\lambda_k t} Z_k^*(t) \to V_k \quad \text{a.s.}$$

Let $\mathcal{F}_\infty^{k-1}$ be the σ-field generated by $Z_j^*(t)$, $j \le k-1$, $t \ge 0$.

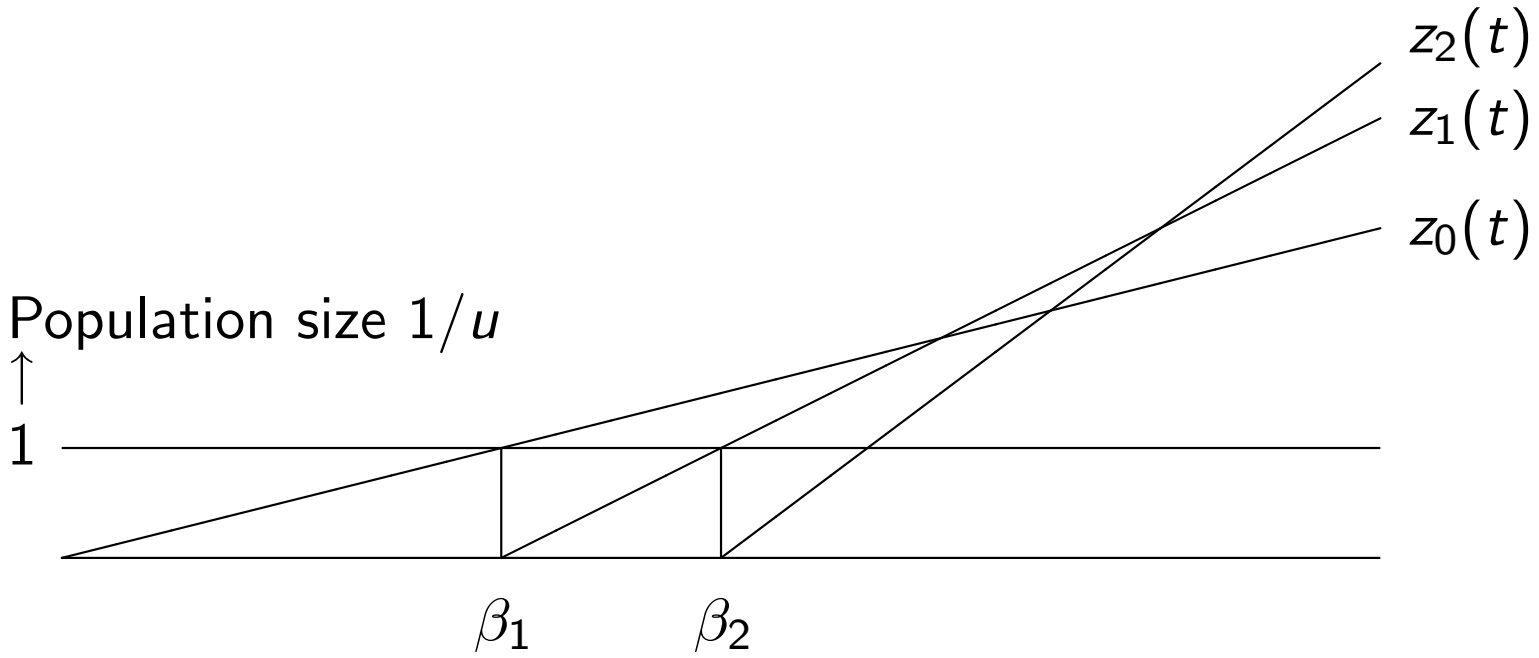
$$E(e^{-\theta V_k} | \mathcal{F}_\infty^{k-1}) = \exp(-c_{h,k} u_k V_{k-1} \theta^{\alpha_k})$$

where $\alpha_k = \lambda_{k-1}/\lambda_k$ and hence $Ee^{-\theta V_k} = \left(1 + c_{\theta,k} \mu_k \theta^{\lambda_0/\lambda_k}\right)^{-1}$.

$$c_{h,k} = \frac{1}{a_k} \left(\frac{a_k}{\lambda_k}\right)^{\alpha_k} \Gamma(\alpha_k)\Gamma(1-\alpha_k)$$

$c_{\theta,k} = c_{\theta,k-1} c_{h,k}^{\lambda_0/\lambda_{k-1}}$, $c_{\theta,0} = a_0/\lambda_0$ and $\mu_k = \prod_{j=1}^k u_j^{\lambda_0/\lambda_{j-1}}$.

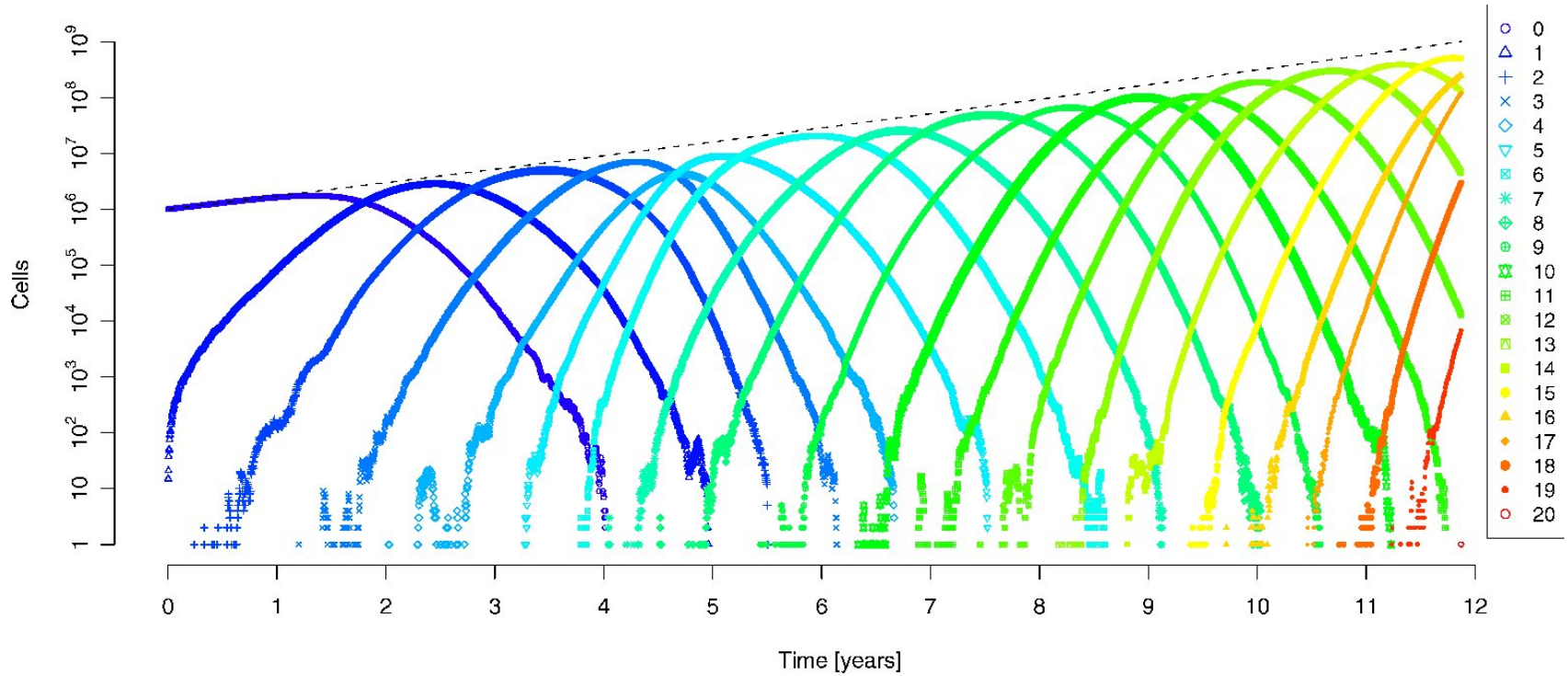
Transitions between waves



$$z_k(t) = \frac{1}{L}\log^+ Z_k(t) \approx \lambda_k(t-\beta_k)^+ \quad L = \log(1/u) \quad \beta_k = \sum_{j=0}^{k-1} 1/\lambda_j$$

Berenwinkel et al (2007) PLoS Comp Bio

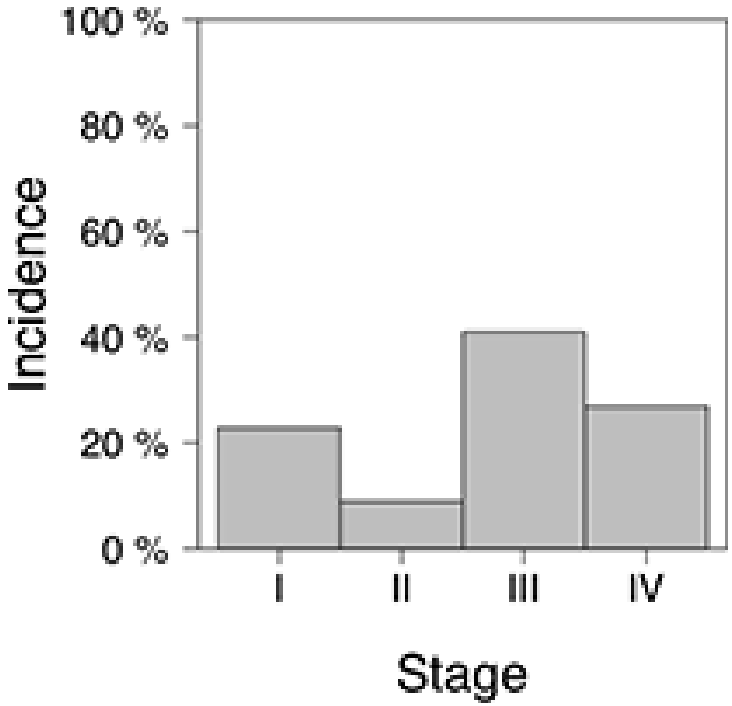
Studied mutational waves in a Moran model with an exponentially growing population. For theorems see D+Mayberry AoAP 21 (2011), 699–744

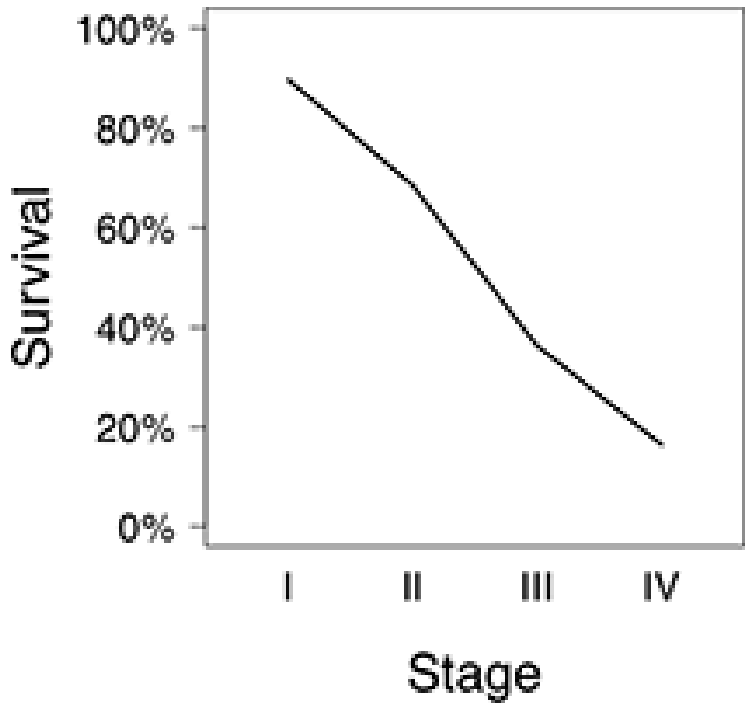


Ovarian Cancer

Ovarian cancer is the fifth leading cause of cancer death among women in the United States.

21,800 new cases and 13,850 deaths in 2010.





Can screening reduce mortality?

In the June 8, 2011 issue of the JAMA, results were published of a study of 78,216 women aged 55–74. Screening used tests for the bio-marker *CA*125 (cancer antigen) and transvaginal ultrasound.

	n	diagnosed	death
annual screening	39,105	212	118
routine care	39,111	176	100

Can screening reduce mortality?

In the June 8, 2011 issue of the JAMA, results were published of a study of 78,216 women aged 55–74. Screening used tests for the bio-marker *CA*125 (cancer antigen) and transvaginal ultrasound.

	n	diagnosed	death
annual screening	39,105	212	118
routine care	39,111	176	100

3,285 women in the screening group had false positive results, 1080 had surgery and 163 had serious complications from surgery.

Multitype model

Lengyel (2010) Am. J. Pathology 177, 1053–1064

An epidermal to mesenchymal transition in the tumor allows its cells to detach. Cells float in the peritoneal fluid as single cells or small groups. Some cells adhere to the omentum, followed by invasion of the extracellular matrix, angiogenesis etc. There are no significant genetic differences between the main tumor and metastases indicating that no additional mutations are needed beyond those that created the initial tumor.

type 0 = primary tumor, type 1 = metastasis. u_1 = rate at which cells leave tumor and attach to the omentum.

Brown and Palmer (2009) PLoS Medecine Vol. 6, Issue 7, e1000114

$\lambda_0 = (\ln 2)/4$, $\lambda_1 = (\ln 2)/2.5$ per month, $u_1 = ?$

Brown and Palmers' parameter estimation

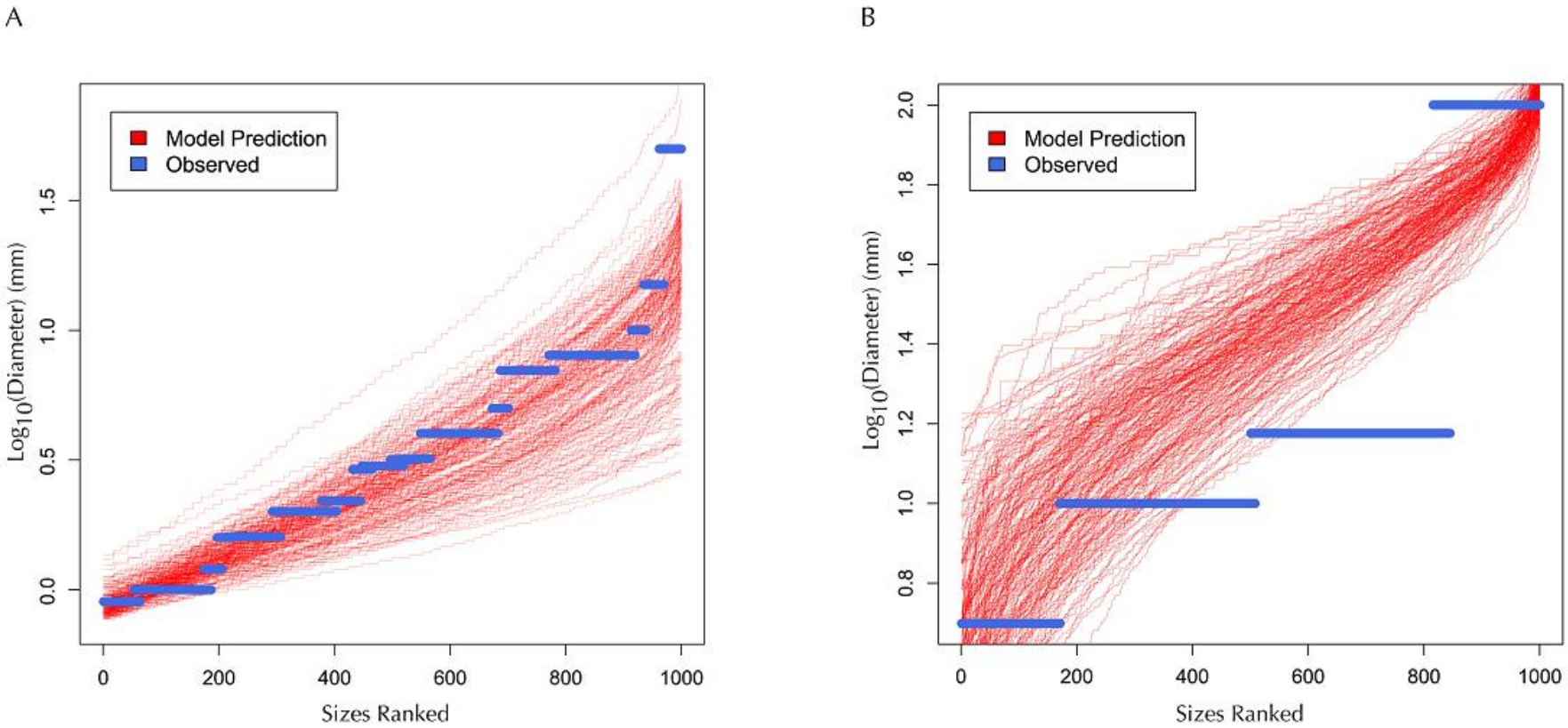


Figure: A. Growth in stages I and II, B. Stages III and IV.

What is the detection window?

People tell us that $1\text{cm}^3 = 10^9$ cells.

Why is this true? $1\text{cm}^3 = 1$ gram. Cells are mostly water. 10 trillion cells (10^{13}) in the human body (and 100 million microbes live on or in us according to Susan Holmes) 100 Kilograms $= 10^5$, which gives $1\text{cm}^3 = 10^8$ cells.

Let T_0 be the time the primary tumor (type 0) has a diameter of 0.5 cm, corresponding to $Z_0(t) = 6.5 \times 10^7$ cells. $V = (\pi/6)d^3$.

At this point it would be hard to detect by transvaginal ultrasound.

Let T_1 be the time that there are $Z_1(t) = 10^9$ metastatic cells (one gram).

Size of the detection window

$Z_0(t) = e^{\lambda_0 t}$ (absorb V_0 by shifting the time origin). $\lambda_0 = 0.1733$.

$$T_0 = \frac{1}{\lambda_0} \ln(6.7 \times 10^7) = 103.8 \text{ months } = 8.65 \text{ years}$$

Let $\gamma_0 = (2/3)\lambda_0$ (leave from surface). First type 1 at $\approx s_1$ where

$$1 = \int_0^{s_1} u_1 e^{\gamma_0 t}\, dt \approx \frac{u_1}{\gamma_0} e^{\gamma_0 s_1}$$

If $u_1 = e^{-a}$ solving gives $s_1 = (1/\gamma_0)(a + \ln(\gamma_0)) = 8.65(a - 2.158)$

$$T_1 - s_1 = \frac{1}{\lambda_1} \ln(10^9) = 74.74 \quad \text{where } \lambda_1 = 0.2773$$

Picking u_1

$T_0 = 103.8,\ T_1 - s_1 = 74.74,\ u_1 = e^{-a},\ s_1 = 8.65(a - 2.158)$

Size of primary at time s_1 is $0.04e^{3a/2}$ cells. During $[s_1, T_1]$ the size of the primary increases by a factor of

$$\exp(\lambda_0(9 \cdot \ln 10)/\lambda_1) \approx 423,000$$

Brown and Palmer: mean size entering stage III is 3 cm. 1.4×10^{10} cells.

When $u_1 = 10^{-4}$ the window $T_1 - T_0 = 2.66$ years.

Size of primary at s_1 is 4×10^4. (Too small?) Diameter = 0.84 mm.

Size at time T_1 is 1.68×10^{10} in agreement with Brown and Palmer.

Primary tumor follows logistic growth?

Let $R = (3V/4\pi)^{1/3}$ be the tumor radius.

$$\frac{dV}{dt} = \lambda \int_0^R 4\pi r^2 f(R-r)\, dr$$

where $f(x)$ describes nutrient availability x cm from the surface and $f(x) \downarrow 0$ exponentially fast as $x \uparrow \infty$.

$V(t)$ grows exponentially fast early but for large t,

$$V'(t) \sim BR^2 \quad \text{i.e.,} \quad V(t) \sim Ct^3.$$

Size of the Primary at $T_1 = \inf\{t : Z_1(t) > 10^9\}$

In the previous calculations we ignored the fact that $Z_1(t) \sim V_1 e^{-\lambda_1 t}$.

$$T_1 = \frac{1}{\lambda_1} \ln(10^9/V_1)$$

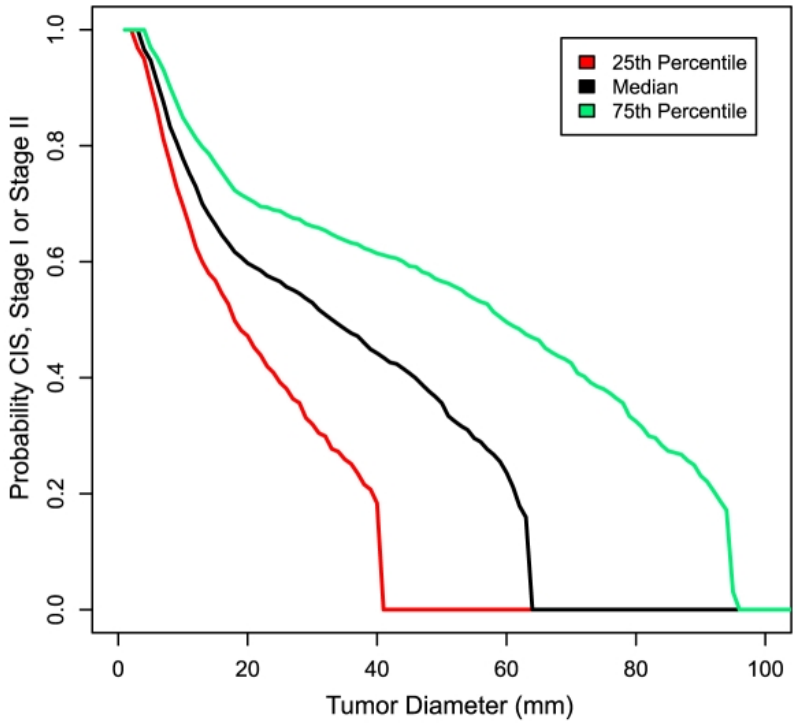
At this time

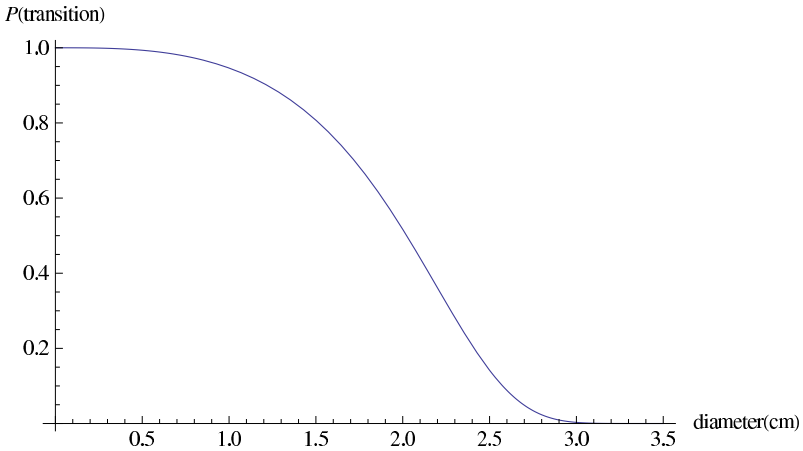
$$Z_0(T_1) = \exp(\lambda_0 T_1) = (10^9 V_1)^\alpha \quad \alpha = \lambda_0/\lambda_1$$

Brockwell and Brown (1978) ZfW (aka PTRF), $V_1^{-\alpha}$ has density

$$\sum_{k=0}^{\infty} \frac{(-x)^k}{\Gamma(k+1)\Gamma(1-\alpha-\alpha k)}$$

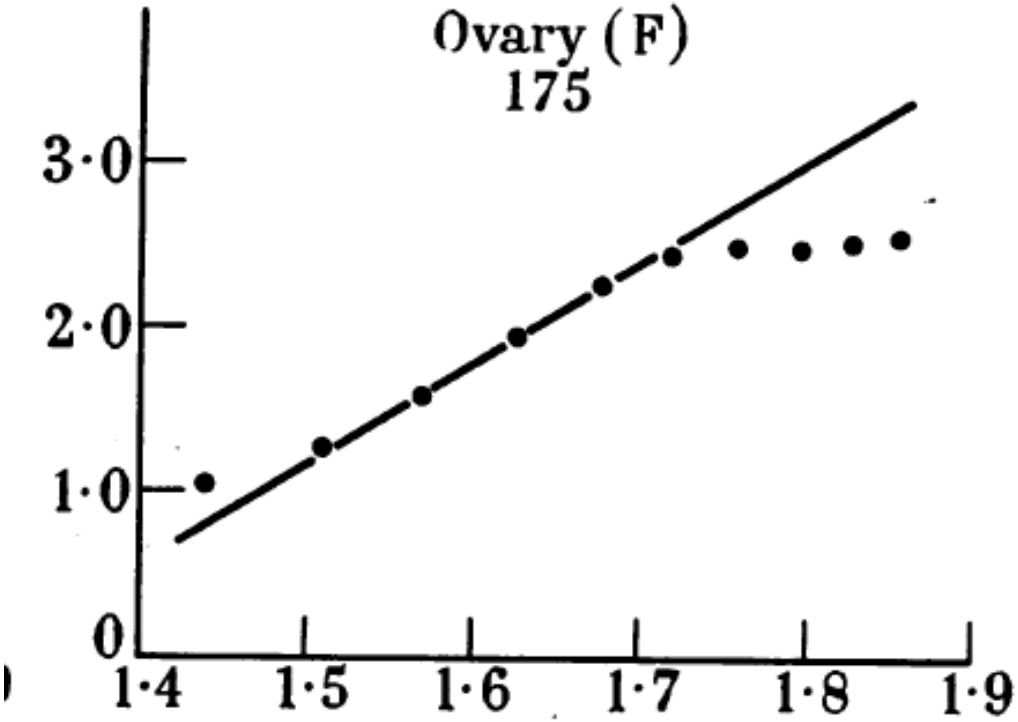
Probability of Progression vs. Size





Work in progress

Use branching process model and a model of detection to predict incidence curve. Picture from Armitage and Doll. We will use SEER data.

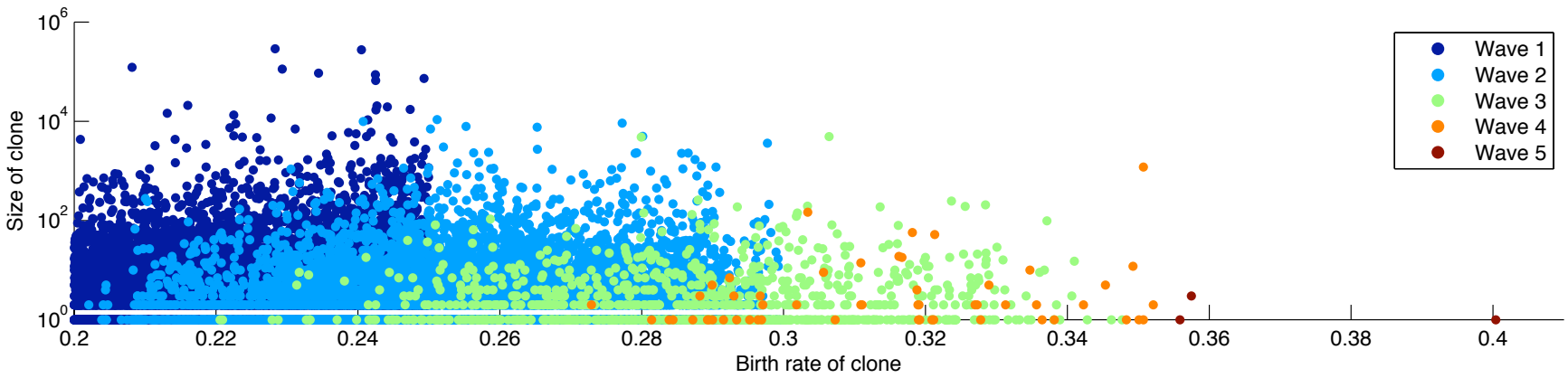


Within Tumor Heterogeneity

Problems in cancer treatment caused by intra-tumor diversity:

- Different subpopulations within a tumor may have varying types of response to any given treatment, making total tumor reduction and prevention of resistance difficult.
- Heterogeneity levels are associated with aggressiveness of disease (e.g., in Barrett's esophagus and prostate cancer).
- Studies have also shown mutational heterogeneity between primary tumors and metastases in breast cancer, explaining lack of response to EFGR antibody therapy in patients that appeared to have no mutation in KRAS.

Intra-tumor diversity generated by model



In this simulation, $b_i = 0.1$, $a_0 = 0.2$, $a_i - a_{i-1} \sim U([0, 0.05])$, $u = 0.001$.

Point process representation of V_1

$Z_0(t) = V_0 e^{\lambda_0 t}$, V_0 nonrandom.

Define a two dimensional point process $\mathcal{X}_t$ with a point at (s, w) if there was a mutation to type 1 at time s and the resulting type 1 branching process $\tilde{Z}_1(t)$ has $e^{-\lambda_1(t-s)}\tilde{Z}_1(t) \to w$.

A point at (s, w) contributes $e^{-\lambda_1 s} w$ to $V_1 = \lim_{t\to\infty} e^{-\lambda_1 t} Z_1(t)$.

$V_1 = \sum_{(s,w)\in\mathcal{X}_t} e^{-\lambda_1 s} w$, the sum of points in a Poisson point process with mean measure $\mu(z, \infty) = A_1 u_1 V_0 z^{-\alpha}$ where $\alpha = \lambda_0/\lambda_1$.

True for $(V_k | \mathcal{F}_\infty^{k-1})$ with $\alpha = \lambda_{k-1}/\lambda_k$.

Flash back to stable laws

Let $Y_1, Y_2, \ldots$ be independent and identically distributed nonnegative random variables with $P(Y_i > x) \sim cx^{-\alpha}$ with $0 < \alpha < 1$. Let $S_n = Y_1 + \cdots + Y_n$. Then

$$S_n/n^{1/\alpha} \to V$$

where V is the sum of points in a Poisson process with mean measure $\mu(z, \infty) = cx^{-\alpha}$

Simpson's index

We define Simpson's index to be the probability two randomly chosen individuals in wave k are descended from the same mutation.

$$R = \sum_{i=1}^{\infty} \frac{X_i^2}{V_k^2}$$

where $X_1 > X_2 > \ldots$ are points in the Poisson process and V_k is the sum.

Theorem. $ER = 1 - \alpha$ where $\alpha = \lambda_{k-1}/\lambda_k$ for wave k.

Proof. Apply results of Fuchs, Joffe and Teugels (2001) about convergence to stable laws. ER does not depend on V_{k-1}.

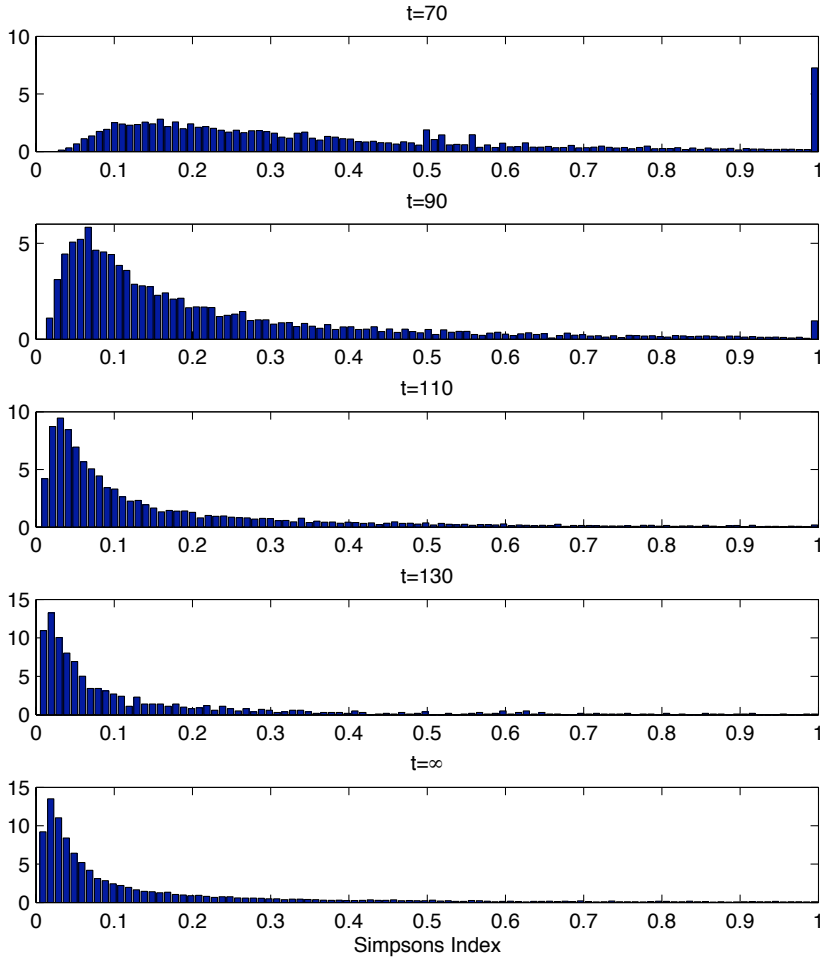


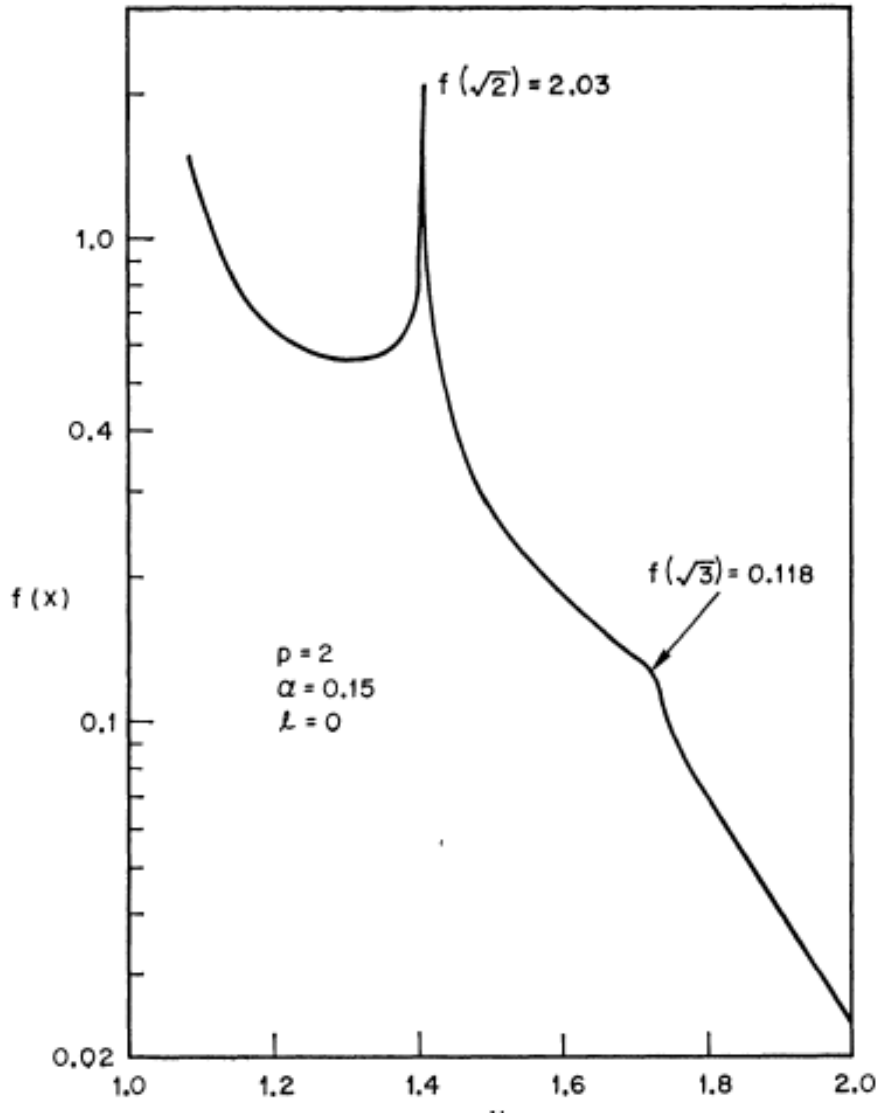
Figure: Empirical distribution of Simpson's Index for wave 1 at times $t = 70, 90, 110, 130, \infty$. Parameters: $b_i = 0.1$, $a_0 = 0.2$, $a_i - a_{i-1} \sim U([0, 0.01])$, mean is $1 - \alpha = 1/11$.

Logan, Mallows, Rice and Shepp (1973)

Consider the "self-normalized sums"

$$S_n(p) = \frac{\sum_{i=1}^n X_i}{(\sum_{j=1}^n X_j^p)^{1/p}} \qquad S_n(2) = R_n^{-1/2}$$

They proved convergence in distribution and identified the Fourier transform of the limit.


f (√2) = 2.03
f (√3) = 0.118
p = 2
α = 0.15
ℓ = 0
f (x)
1.0
0.4
0.1
0.02
1.0
1.2
1.4
1.6
1.8
2.0

Largest clone

Let $U_n = \max_{1 \le i \le n} Y_i / S_n$ be the contribution of the largest term to the sum.

Darling (1952). Theorem 5.1. As $n \to \infty$, $U_n^{-1} \to T$ where T has characteristic function $e^{it}/f_\alpha(t)$ where

$$f_\alpha(t) = 1 + \alpha \int_0^1 (1 - e^{itu}) u^{-(\alpha+1)} \, du$$

$ET = 1/(1-\alpha)$ and $\text{var}\,(T) = 2/(1-\alpha)^2(2-\alpha)$.

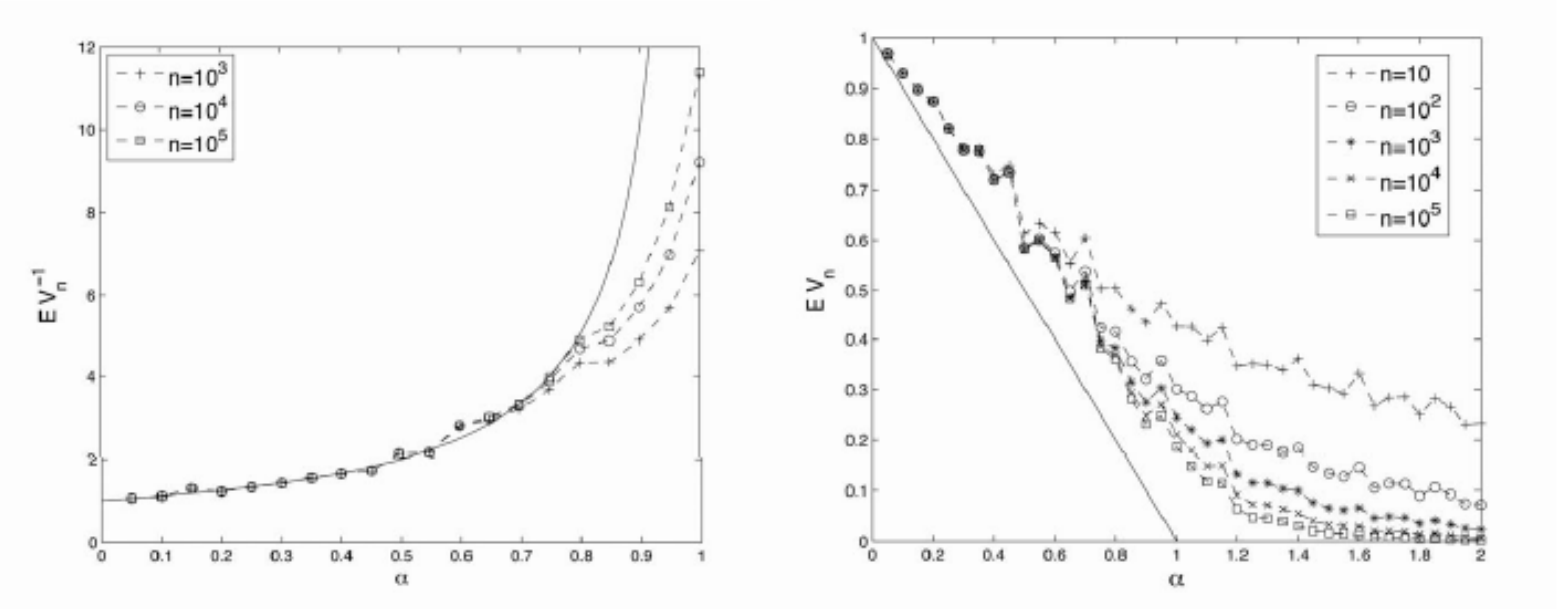


Figure: Monte Carlo estimates for $E(1/U_n)$ and EU_n plotted versus $1/(1-\alpha)$ and $1-\alpha$. $T = \lim 1/U_n$ has $ET = 1/(1-\alpha)$ and $E(1/T) > 1-\alpha$.

Summary

Growth, progression, and metastasis of cancer can be modeled with multi-type branching processes, and these models can be used to evaluate screening strategies and treatment regimens.

$(e^{-\lambda_k t} Z_k(t) | \mathcal{F}_{k-1}^{\infty}) \to V_k$ where V_k is one-sided stable with index $\alpha_k = \lambda_{k-1}/\lambda_k$.

Results about stable laws can be used to obtain results about tumor heterogeneity and other quantities of interest.

In contrast to simulation, our analytical results are exact, and reveal the dependence on underlying parameters.