Variations on the Luria-Delbrück model

Thierry Huillet, LPTM Cergy, France

...

[1] Luria; Delbrück. Mutations of bacteria: from virus sensitivity to virus resistance. Genetics. 1943.
[2] Lea; Coulson. The distribution of the numbers of mutants in bacterial populations. J. of Gen. 1949.
[3] Keller; Antal. Mutant number distribution in an exponentially growing population. JStatMech. 2015.
[4] Dewanji et al. A generalized Luria-Delbrück model. Math Biosci. 2005.
[5] Iwasa et al. Evolution of resistance during clonal expansion. Genetics 172(4), 2006.
[6] Simon. On a class of skew distribution functions. Biometrika. 1955
[7] .... far from exhaustive

IRS Santiago, December 5, 2016

Intro

Luria-Delbrück experiment (Fluctuation Test): genetic mutations of bacteria arise permanently, even in absence of selection, rather than being a response to selection. Mutations do not occur out of necessity (Lamarck), but instead can occur many generations before the selection strikes (Darwin).

Sensitive population ($\#\ x_t$ at t) immune as soon as (i) $N_t > 0$ ($\#\ N_t$ of mutants) or (ii) $N_t > ax_t$.

Lamarck. t instant of viral attack , each of x_t sensitive individuals has proba p to switch instantaneously to a mutant state in response. $\#\ N_t$ of mutants: $N_t \sim \text{bin}(x_t, p)$ mean $\mathbf{E}N_t = x_t p$ and variance $\sigma^2(N_t) = x_t p(1-p)$

-If $x_t \uparrow$, $\mathbf{P}(N_t > 0) = 1 - (1-p)^{x_t} \underset{t\to\infty}{\to} 1$: population will become increasingly immune based on (i).

- $N_t/x_t \overset{\text{a.s.}}{\underset{t\to\infty}{\to}} p$ and $p > a \Rightarrow$ population asymptotically immune, based on (ii).

- $x_t \to \infty$ and $p \to 0$ while $x_t p = \overline{\theta}$ (LPSM $*$−limit) : $N_t \overset{*}{\underset{t\to\infty}{\to}} N_\infty \sim \text{Poi}(\overline{\theta})$, mean=variance.

Darwin-Luria-Delbrück version of this model: more complicated intertwining of $(x_t; N_t)$. In such process, the Yule and Simon distribution pops in.

Simon tail index larger than 1

Naturalist daily records sampled species and occurrences. n campaigns: $N_n(k) = \#$ of species sampled k times, $P_n = \sum_{k=1}^{n} N_n(k)$, # distinct species discovered. Means: $x_n(k)$ and p_n. Step n to $n+1$:
- proba. ρ: sample new species, $N_{n+1}(1) = N_n(1) + 1$, $P_n \sim \text{bin}(n, \rho)$, $\mathbf{E}P_n = p_n = n\rho$.
- With proba. $1-\rho$, outcome of $(n+1)^{\text{th}}$ campaign is species already visited: species k with proba. $kN_n(k)/n$ (reinforcement enhancing species visited often, PA).

$$x_{n+1}(k) = x_n(k) + (1-\rho)(k-1)x_n(k-1)/n - (1-\rho)kx_n(k)/n \text{ if } k \neq 1$$

$\alpha = 1/(1-\rho) > 1$, solutions are $x_n(k) = nx(k)$ with $x(k) = \rho\alpha B(k, \alpha+1)$. Simon:

$$q_k := x_n(k)/p_n = x(k)/\rho = \alpha B(k, \alpha+1), \; k \geq 1 \tag{1}$$

$q_k \underset{k\to\infty}{\sim} \alpha\Gamma(\alpha+1)k^{-(\alpha+1)}$. Obeys $q_{k+1}/q_k = k/(k+\alpha+1)$, $q_1 = \alpha/(\alpha+1)$. Pgf:

$$\sum_{k\geq 1} q_k z^k = \frac{\alpha z}{\alpha+1} F(1,1;2+\alpha;z) =: F_S(z). \tag{2}$$

$$F_S(z) = \alpha \int_0^\infty d\tau \cdot e^{-\alpha\tau} \frac{e^{-\tau}z}{1-(1-e^{-\tau})z} \text{ or } q_k = \alpha \int_0^\infty d\tau \cdot e^{-\alpha\tau} e^{-\tau}(1-e^{-\tau})^{k-1}. \tag{3}$$

$$N_n(k)/P_n \overset{\text{proba}}{\to} q_k. \tag{4}$$

Simon tail index smaller than 1 but rational

Consider a new $\mathbb{N}_0$−valued rv, say $\overline{C}$, $\alpha > 0$, now with pgf

$$\mathbf{E}\left(z^{\overline{C}}\right) = \frac{\alpha}{\alpha+1} F\left(1,1;2+\alpha;z\right).$$

In class of 3-parameters hypergeometric family of pgfs studied in Dacey. When $\alpha < 1$ and α is a rational number, $\overline{C}$ has a Pólya-Eggenberger urn model interpretation: Take an urn with initially b black balls and $w > b$ white balls. Balls are drawn at random one at a time from the urn and each selected ball is returned to the urn along with $r-1$ additional balls of the same color, $r \geq 2$. Repeat sampling procedure. Suppose number of balls returned is $r = w$ and put $\alpha := b/r < 1$.

CLAIM: $\overline{C}_\alpha$ *represents the number of white balls that are drawn till the first black ball is selected in the sampling process.*

With $\overline{q}_k := \mathbf{P}\left(\overline{C} = k\right)$, $k \geq 0$, $\overline{q}_{k+1}/\overline{q}_k = (k+1)/(k+\alpha+2)$, $\overline{q}_0 = \alpha/(\alpha+1)$.

$$\overline{q}_k = \alpha B\left(k+1, \alpha+1\right), \; k \geq 0.$$

The distribution of $\overline{C} = C - 1$ is the distribution of a shifted YS distribution with $\overline{q}_k = q_{k+1}$. Reinforcement entails heavy-tailed with index α.

digression: Sibuya?

Link with Sibuya(α)? $C \geq 1$ integer-valued rv

$$C = \inf \left(l \geq 1 : \mathcal{B}_\alpha (l) = 1 \right),$$

$\left(\mathcal{B}_\alpha (l) \right)_{l \geq 1}$ sequence of independent Bernoulli rvs obeying $\mathbf{P} \left(\mathcal{B}_\alpha (l) = 1 \right) = \alpha / l$, $\alpha \in (0,1)$. First epoch of a success in a Bernoulli trial with probab. of success inversely proportional to the number of the trial.

$$q_k = \mathbf{P} \left(C = k \right) = (-1)^{k-1} \binom{\alpha}{k} = \alpha \left[\overline{\alpha} \right]_{k-1} / k!, \; k \geq 1.$$

Heavy tails: $q_k \underset{k \to \infty}{\sim} \alpha k^{-(\alpha+1)} / \Gamma (1 - \alpha)$ and $q_{k+1} / q_k = (k - \alpha) / (k+1)$, $q_1 = \alpha$.

$$\text{pgf: } \varphi (z) := \mathbf{E} \left(z^C \right) = 1 - (1 - z)^\alpha = \alpha z F (1, 1 - \alpha; 2; z), \; z \leq 1.$$

Scale-free: with $u \circ C$, Bernoulli(u)-thinning of C, C solves (fixed point of)

$$\forall u \in (0,1), \; \left(u \circ C \mid u \circ C \geq 1 \right) \overset{d}{=} C$$

$G(\alpha)$ ~gamma$(\alpha, 1)$, $G(1)$, $G(1-\alpha)$, $G(\alpha)$ mutually $\perp$, Poisson mixture (Devroye)

$$C \overset{d}{=} 1 + \text{Poi} \left(\frac{G(1)\, G(1-\alpha)}{G(\alpha)} \right).$$

Exponential sensitive growth: Luria-Delbrück

WT (sensitive) cells grow at rate $\lambda_t > 0$, $\Lambda_t = \int_0^t ds \cdot \lambda_s < \infty$. Size of the sensitive

$$x_t = x_0 + \Lambda_t, x_0 \geq 0.$$

Each WT cell subject to mutation, rate at which new mutants are being created, one at a time, is $\nu\lambda_t$, ν = mutation proba. of each WT cell.
Mutant population is resistant to a viral attack.
Fix $[0, t]$. Mutations occur at iid times $S_t^{(k)}$ law: $\mathbf{P}(S_t \in ds) = \lambda_s ds / \Lambda_t$.
There are $P(\nu\Lambda_t)$ ~Poi$(\nu\Lambda_t)$ such mutation events.
Once mutant is created, it grows and forms a clone.
$M_t = \#$ mutant sub-population at t given single founder $M_0 = 1$. M_t grows according to BD process.
M_t goes extinct at time τ_e: $\mathbf{P}(M_t > 0) = \mathbf{P}(\tau_e > t)$.
$N_t = \#$ of total mutant pop., summing up all sub-populations contributions.

Global mutant pop. size

$$N_t = \sum_{k=1}^{P(\nu\Lambda_t)} C^{(k)}_{t-S^{(k)}_t}. \tag{5}$$

Pgf:

$$\Phi_t(z) = \mathbf{E}\left(z^{N_t}\right) = \exp\left\{-\nu \int_0^t ds \cdot \lambda_s \left(1 - \phi_{t-s}(z)\right)\right\},\ \phi_t(z) = \mathbf{E}\left(z^{M_t}\right). \tag{6}$$

As well:

$$\text{Compound Poisson: } N_t \stackrel{d}{=} \sum_{p=1}^{P(\nu\Lambda_t)} C_t^{(k)}, \tag{7}$$

$C_t^{(k)}$ iid copies of $C_t \geq 0$, the typical clone size at t with pgf

$$\mathbf{E}\left(z^{C_t}\right) = \frac{1}{\Lambda_t} \int_0^t ds \cdot \lambda_s \phi_{t-s}(z).$$

$C_t \to C_t^+ := C_t \mid C_t > 0$.
Two models for WT population growth: $\lambda_t = \lambda e^{\lambda t}$ and $\lambda_t = \lambda$.

Expon. growing WT pop.($\lambda_s = \lambda e^{\lambda s}$): supercrit. mutant ($r > 0$), [3,4,5]

Each mutant duplicates with proba. π_2 or dies with proba. π_0
Global BD rate: $r_e > 0$, $r_b := r_e \pi_2$, $r_s := r_e \pi_1$ and $r_d := r_e \pi_0$, $r_e = r_b + r_d + r_s$.
Mutant BD net rate

$$r = r_b - r_d \text{ and } \rho := \pi_0 / \pi_2 = r_d / r_b,$$

$$\alpha := \lambda / r \text{ and } \mu := \nu \lambda (1 - \rho) / r = \nu \lambda / r_b \text{ scaled mutation proba.}$$

For BD branching proc. $r \neq 0$, pgf $\phi_t(z) := \mathbf{E}\left(z^{M_t}\right)$ is $[\zeta := (z - \rho) / (z - 1)]$

$$1 - \phi_t(z) = e^{rt}(1 - z) / \left(1 + r_b\left(e^{rt} - 1\right)(1 - z)/r\right) = (1 - \rho) / \left(1 - e^{-rt}\zeta\right), \qquad (8)$$

Supercrit. ($r > 0$), extinction occurs with > 0 proba. at time τ_e.

$$1 - \phi_t(0) = \mathbf{P}(\tau_e > t) = e^{rt} / \left(1 + r_b\left(e^{rt} - 1\right)/r\right) \qquad (9)$$

$\rho = \mathbf{P}(\tau_e < \infty)$ proba. extinction of M_t. Given ext. tail of $\tau_e \sim \exp(r)$.

Clone size:

$$\mathbf{E}\left(z^{C_t}\right) = \Lambda_t^{-1} \int_0^t ds \lambda_s \phi_{t-s}(z) \underset{t\to\infty}{\to} \alpha \int_0^\infty d\tau e^{-\alpha\tau} \left(\rho - \zeta e^{-\tau}\right) / \left(\rho - \zeta e^{-\tau}\right) = \mathbf{E}\left(z^{C_\infty}\right)$$

CLAIM: $\mathbf{E}\left(z^{C_\infty}\right)$ *is pgf of an exp(α) mixture (w.r.t. to parameter τ) of a linear-fractional distrib. pgf*

$$\mathbf{E}\left(z^{C}\right) = b_0 + a_0 \frac{az}{1 - bz},$$

($C \overset{d}{=} G(a) \cdot B(a_0)$), G ~geo(a) $\perp$ B ~ber(a_0), with success parameters $\left(a_0 = (1-\rho)/(1-\rho e^{-\tau}),\ a = e^{-\tau}(1-\rho)/(1-\rho e^{-\tau})\right)$.

$$q_k = \mathbf{P}\left(C_\infty^+ = k\right) = \alpha B(k, \alpha+1)(1-\rho) F(k+1, \alpha; k+\alpha+1; \rho) / \int_0^1 \left(1 - \rho z^{1/\alpha}\right) dz,\ k \geq 1.$$

Pgf of the current number of mutants

With $F(\zeta) := F(1,\alpha;1+\alpha;\zeta) = 1+\alpha\sum_{k\geq 1}\frac{\zeta^k}{\alpha+k}$, $\zeta := (z-\rho)/(z-1)$,

$$\Phi_t(z) = \exp -\mu \sum_{k\geq 0} \frac{\zeta^k}{\alpha+k}\left(e^{\lambda t} - e^{-\lambda k t/\alpha}\right) = \exp -\frac{\mu}{\alpha}\left(e^{\lambda t}F(\zeta)\right) - F\left(\zeta e^{-\lambda t/\alpha}\right) \quad (10)$$

$$\mathbf{E}(N_t) = \frac{\mu x_t}{1-\rho}\cdot\begin{cases} \log x_t \text{ if } \alpha = 1 \\ \frac{1}{1-\alpha}\left(x_t^{1/\alpha-1}-1\right) \text{ if } \alpha \neq 1 \end{cases}, \quad (11)$$

$$\sigma^2(N_t) = \frac{\mu x_t}{(1-\rho)^2}\cdot\begin{cases} 2(x_t-1)-(1+\rho)\log x_t \text{ if } \alpha = 1 \\ (1+\rho)\left(x_t^{-1/2}-1\right)+\log x_t \text{ if } \alpha = 2 \\ \frac{2}{2-\alpha}x_t^{2/\alpha-1}+\frac{1+\rho}{\alpha-1}x_t^{1/\alpha-1}+\frac{\rho(2-\alpha)+\alpha}{(2-\alpha)(1-\alpha)} \text{ if } \alpha \neq \{1,2\} \end{cases}. \quad (12)$$

- If $\alpha < 1$, both mean and SD are $O\left(x_t^{1/\alpha}\right)$ (very large fluctuations)
- If $\alpha > 2$, both mean and variance are $O(x_t)$ (a Poissonian regime).
- In all cases $\alpha \leq 2$, the variance exceeds the mean (an overdispersed situation for N_t), with special logarithmic effects when $\alpha \in \{1,2\}$. Limiting (stable) laws of properly scaled versions of N_t [4].

- If $\alpha = \lambda/r = 1$ (NEUTRALITY)
$F(\zeta) = 1 - F_s(\eta) = -\frac{1-\eta}{\eta}\log(1-\eta)$, $\eta = \zeta/(\zeta-1) = (z-\rho)/(1-\rho)$ and

$$\Phi_t(z) = \left(1 - \left(1 - e^{-\lambda t}\right)(z-\rho)/(1-\rho)\right)^{\mu e^{\lambda t}\frac{1-z}{z-\rho}}.$$

$$\mathbf{E}N_t \sim \frac{\mu}{1-\rho} x_t \log x_t \text{ and } \sigma^2(N_t) \sim \frac{2\mu}{(1-\rho)^2} x_t^2.$$

If in addition (pure birth): $\rho = 0$ and $\mu = \nu\alpha = \nu$: (Luria-Delbrück relations)

$$\begin{aligned} \mathbf{E}N_t &\sim \nu x_t \log x_t \text{ and } \sigma^2(N_t) \sim 2\nu x_t^2. \\ \sigma^2(N_t)/\mathbf{E}N_t &\sim 2x_t/\log x_t \gg 1 \text{ and } \sigma(N_t)/\mathbf{E}N_t \sim 1/\left(\sqrt{\nu/2}\log x_t\right) \end{aligned}$$

contrasting with $N_t \sim$bin(x_t, p)!.

The large population, small mutation $*-$limit

When $t \to \infty$, $\nu \to 0$ while $\mu e^{\lambda t} \sim \mu x_t = \theta > 0$ (the $*-$limit),

$$\text{Compound Poisson:} \quad \begin{aligned} \Phi_t(z) \underset{*}{\to} \Phi_\infty(z) &:= \mathbf{E}\left(z^{N_\infty}\right) = \exp\left\{-\frac{\theta}{\alpha} F(\zeta)\right\} \\ &= \exp\left\{-\frac{\theta}{\alpha} F(\rho)\left(1-\left(1-F(\zeta)/F(\rho)\right)\right)\right\}, \end{aligned} \tag{13}$$

$$\varphi(z) = \mathbf{E}\left(z^{C_\infty^+}\right) = 1 - F(\zeta)/F(\rho) = \frac{F_S(\eta) - F_S(\rho/(\rho-1))}{1 - F_S(\rho/(\rho-1))}.$$

CLAIM: (i) *The joint prob.* $P_{n,p} := \mathbf{P}(N_\infty = n, P = p)$ *obeys the five-term recurrence:*

$$\begin{cases} F(\rho)\,\rho\,(n+1)\,P_{n+1,p} = \left(p\alpha F(\rho)\,\overline{\rho} + n(\rho+1)\,F(\rho)\right) P_{n,p} \\ +\alpha\overline{\theta}\left[1 - F(\rho)\,\overline{\rho}\right] P_{n,p-1} - (n-1)\,F(\rho)\,P_{n-1,p} - \alpha\overline{\theta} P_{n-1,p-1}. \end{cases}$$

(ii) q_k *obeys the three-term recurrence* ($q_0 = 0$)*:*

$$\rho\,(k+1)\,q_{k+1} = \left(\alpha\overline{\rho} + k(\rho+1)\right) q_k - (k-1)\,q_{k-1}, \;\; k \geq 1.$$

$$\mathbf{E}\left(N_\infty\right)=\begin{cases}\infty \text{ if } 0<\alpha\leq 1\\ \frac{\theta}{(1-\rho)(\alpha-1)} \text{ if } \alpha>1\end{cases} \text{ and } \sigma^2\left(N_\infty\right)=\begin{cases}\infty \text{ if } 0<\alpha\leq 2\\ \frac{\theta}{(1-\rho)^2}\frac{\rho(2-\alpha)+\alpha}{(\alpha-2)(\alpha-1)} \text{ if } \alpha>2\end{cases}. \quad (14)$$

- $\alpha>1$, $\sigma^2\left(N_\infty\right)=\mathbf{E}\left(N_\infty\right)\frac{\rho(2-\alpha)+\alpha}{(1-\rho)(\alpha-2)}>\mathbf{E}\left(N_\infty\right)$. - $0<\alpha\leq 1$, both $=\infty$.

CLAIM: *N_∞ is discrete-self-dec. (SD) and thus unimodal. With $\theta_{\max}:=\alpha\left(1-\rho\right)/F_S'\left(-\rho/\left(1-\rho\right)\right)$, it has its mode at the origin if $\theta<\theta_{\max}$ and two modes at $n=\{0,1\}$ if $\theta=\theta_{\max}$.*

$\theta\leq\theta_{\max}\Rightarrow$ SD and unimodal near the origin. $\theta>\theta_{\max}$, N_∞ still SD thus unimodal but with mode away from origin.

Inspecting (11) and (12) closer, in the $*-$limit

$$\mathbf{E}\left(N_t\right)\underset{*}{\sim}\frac{\theta}{1-\rho}\cdot\begin{cases}\frac{1}{1-\alpha}x_t^{1/\alpha-1} \text{ if } 0<\alpha<1\\ \log x_t \text{ if } \alpha=1\\ \frac{1}{\alpha-1} \text{ if } \alpha>1\end{cases}, \quad \sigma^2\left(N_t\right)\underset{*}{\sim}\frac{\theta}{(1-\rho)^2}\cdot\begin{cases}\frac{2}{2-\alpha}x_t^{2/\alpha-1} \text{ if } 0<\alpha<2\\ \log x_t \text{ if } \alpha=2\\ \frac{\rho(2-\alpha)+\alpha}{(2-\alpha)(1-\alpha)} \text{ if } \alpha>2\end{cases}$$

Time spent in the mutant-free state: local exctinctions

$I_t = \int_0^t \mathbf{1}(N_s = 0)\, ds$ fraction of time interval $[0, t]$ free of mutants (length of the set $\mathcal{I}_t$ uncovered by the mutants), $I_t^c = \int_0^t \mathbf{1}(N_s > 0)\, ds = t - I_t$, length of the covered set $\mathcal{I}_t^c$, with

$$\mathcal{I}_t = [0, t] \cap \mathcal{I}_t^c \ ; \ \mathcal{I}_t^c = \bigcup_{k=1}^{P(\nu \Lambda_t)} \left] S_t^{(k)}, S_t^{(k)} + \tau_e^{(k)} \right[\cap [0, t], \qquad (15)$$

$\tau_e^{(k)}$ are iid copies of τ_e. $\mathbf{E}(I_t) = \int_0^t \Phi_s(0)\, ds$ and putting $z = 0$, $\zeta = \rho$ in (10)

$$\Phi_t(0) = \exp\left\{ -\mu \sum_{k \geq 0} \frac{\rho^k}{\alpha + k} \left(e^{\lambda t} - e^{-rkt} \right) \right\} = \exp\left\{ -\frac{\mu}{\alpha} \left(e^{\lambda t} F(\rho) - F\left(\rho e^{-\lambda t/\alpha} \right) \right) \right\}.$$

And, $\mathbf{E}(I_t) = \int_0^t \Phi_s(0)\, ds \to \mathbf{E}(I_\infty) < \frac{e^{\mu/\alpha F(\rho)}}{\lambda} E_1\left(\frac{\mu}{\alpha}\right) < \infty$.

The pure birth Yule case ($\pi_0 = 0,\ \rho = 0$)

$$\mathbf{E}\left(z^{C_t}\right) \underset{t\to\infty}{\to} \alpha\int_0^\infty d\tau \cdot e^{-\alpha\tau}\frac{ze^{-\tau}}{1-z+ze^{-\tau}} \overset{YS}{=} \frac{\alpha z}{\alpha+1}F\left(1,1;2+\alpha;z\right),$$

$$\begin{aligned}\Phi_t\left(z\right) &= \exp\left\{-\tfrac{\mu}{\alpha}\left(e^{\lambda t}F\left(z/\left(z-1\right)\right)\right) - F\left(z/\left(z-1\right)e^{-\lambda t/\alpha}\right)\right\} \\ &= \exp\left\{-\tfrac{\mu}{\alpha}\left(e^{\lambda t}\left(1-F_S\left(z\right)\right) - \left(1-F_S\left(\tfrac{e^{-\lambda t/\alpha}z}{1-z\left(1-e^{-\lambda t/\alpha}\right)}\right)\right)\right)\right\}\end{aligned} .$$

$$\text{Compound-Poisson: } \Phi_t\left(z\right) \underset{*}{\to} \Phi_\infty\left(z\right) = e^{-\frac{\theta}{\alpha}\left(1-F_S\left(z\right)\right)},$$

CLAIM: (i) $\mathbf{P}\left(N_\infty = n, P = p\right)$ *obeys the three-term recurrence:*

$$\left(p+\frac{n}{\alpha}\right)\mathbf{P}\left(N_\infty = n, P = p\right) = \frac{\theta}{\alpha}\mathbf{P}\left(N = n-1, P = p-1\right) + \frac{n-1}{\alpha}\mathbf{P}\left(N = n-1, P = p\right)$$

(ii) $q_k = \mathbf{P}\left(C_\infty^+ = k\right)$ *obeys 2-term recurrence:* $\left(k+\alpha+1\right)q_{k+1} = kq_k,\ k \geq 1.$

(iii) $P_n = \left(P \mid N_\infty = n\right) \underset{n\to\infty}{\overset{d}{\to}} 1+\mathrm{Poi}\left(\overline{\theta}\right)$ at rate $n^{-\alpha\wedge 1}$.

CLAIM: *N_∞ is discrete-SD and thus unimodal. It has its mode at the origin if $\theta < 1 + \alpha$ and 2 modes at $n = 0, 1$ if $\theta = 1 + \alpha$.*

If $r_e = r = \lambda \Rightarrow \alpha = 1$, using $\frac{z}{2} F\left(1, 1; 3; z\right) = 1 + \frac{1-z}{z} \log\left(1 - z\right)$

$$q_k = \mathbf{P}\left(C_\infty^+ = k\right) = 1/\left(k\left(k+1\right)\right).$$

And

$$\Phi_t\left(z\right) = \left(1 - \left(1 - e^{-\lambda t}\right) z\right)^{\mu e^{\lambda t}(1-z)/z}.$$

Compound-Poisson pgf of # mutants in $*$–limit:

$$\Phi_\infty\left(z\right) = \left(1 - z\right)^{\theta(1-z)/z}.$$

And

$$\mathbf{E}\left(N_t\right) \overset{*}{\sim} \theta \log x_t \text{ and } \sigma^2\left(N_t\right) \overset{*}{\sim} 2\theta x_t > \mathbf{E}\left(N_t\right) \text{ contrasting with Poi}(\overline{\theta}).$$

The subcritical case ($r < 0$)

$\mathbf{P}(\tau_e > t) \sim e^{-r_d t}$, exp. tails. A.s. extinction. $\kappa = -\frac{r_b}{r} > 0$, $\alpha = -\lambda / r > 0$.
CLAIM: *With B ~Bernoulli(a_0) distributed with success prob.*
$a_0 = e^{-\tau} / (1 + \kappa (1 - e^{-\tau}))$, $\perp$ of $G \sim$ geometric(a) distributed with success prob.
$a = 1/(1 + \kappa (1 - e^{-\tau}))$, C_∞ is an exp(α) mixture (with respect to τ) of $C \overset{d}{=} G \cdot B$.

$$\xi = (z-1)/(z-\rho) = 1/\zeta \text{ and } \Phi_t(z) = \exp\left\{-\frac{\mu}{\alpha}\left(e^{\lambda t} F(\xi)\right) - F\left(\xi e^{tr}\right)\right\}.$$

The $*$–limit: $\nu \to 0$ and $\mu e^{\lambda t} \sim \mu x_t = \theta$, $\rho_* = 1/\rho$, pgf of a CP rv with intensity $\overline{\theta} := \theta F(\rho_*)/\alpha$ and clone size with pgf

$$\varphi(z) = \mathbf{E}\left(z^{C_\infty^+}\right) = 1 - F(\xi)/F(\rho_*), \text{ with } \varphi(0) = 0.$$

$\eta = \xi/(\xi - 1) = \rho_*(z-1)/(1-\rho_*)$, $\mathbf{E}\left(z^{C_\infty^+}\right)$ is pgf with all falling fact. mom. $< \infty$

$$\mathbf{E}\left[\left(C_{\infty}^{+}\right)_{k}\right] = \frac{k!}{F(\rho_*)}\left[(z-1)^k\right]\frac{\alpha\eta}{\alpha+1}F(1,1;2+\alpha;\eta) = \frac{k!}{F(\rho_*)}\left(\frac{\rho_*}{1-\rho_*}\right)^k \alpha B(k,\alpha+1).$$

CLAIM: *In the subcrit. regime $r<0$, # of mutants in $*-$limit is a compound-Poisson($\theta F(\rho_*)/\alpha$) rv, with clone size C_{∞}^{+} having all its moments.*

$$\textbf{Pure death}:\ \pi_2 = 0 :\ \Phi_t(z) = \exp\left\{-\frac{\nu\lambda(1-z)}{\lambda+r_d}\left(e^{\lambda t}-e^{-r_d t}\right)\right\}.$$

$*-$limit: $\nu \to 0,\ t\to\infty$ while $\nu e^{\lambda t} = \theta$

$$\Phi_{\infty}(z) = \exp\left\{-\frac{\theta\lambda(1-z)}{\lambda+r_d}\right\},$$

pgf of Poisson rv with intensity $\overline{\theta} := \theta\lambda/(\lambda+r_d)$.
With $\alpha = \lambda/r = \lambda/(-r_d) < 0$,

$$\mathbf{E}(I_t) \underset{t\to\infty}{\to} \frac{1}{\lambda}\int_1^{\infty}\frac{du}{u}e^{-\frac{\nu\lambda}{\lambda+r_d}\left(u-u^{1/\alpha}\right)} < -\frac{1}{r_d}e^{-\frac{\nu\lambda}{\lambda+r_d}}E_1\left(-\frac{\nu\lambda}{\lambda+r_d}\right) < \infty.$$

Linearly growing sensitive ($\lambda_t = \lambda$): supercritical BPI

With $\mu := \nu\lambda / r_b$, $p_t / (1 - p_t) = r_b (e^{rt} - 1) / r$, Neg. Bin.

$$\Phi_t(z) = \exp\left\{-\nu\lambda(1-z)\int_0^t ds \cdot \frac{e^{rs}}{1 + \frac{r_b}{r}(e^{rs}-1)(1-z)}\right\} = \left(\frac{1-p_t}{1-p_t z}\right)^{\mu}$$

$$\mathbf{E}(N_t) = \mu p_t / (1-p_t) \sim \nu\lambda e^{rt}/r,\ \sigma^2(N_t) = \mu p_t/(1-p_t)^2 \sim \nu\lambda r_b e^{2rt}/r^2.$$

$$\mathbf{E}(I_t) \underset{\text{if } r_b \neq \nu\lambda}{\sim} \frac{1}{r}\left(\frac{r_d}{r}\right)^{-\mu} \frac{1}{-\mu}\left[\left(\frac{r_b}{r_d}e^{rt}\right)^{-\mu} - \left(\frac{r_b}{r_d}\right)^{-\mu}\right]$$

CLAIM: • *if $r_b \neq \nu\lambda$ ($\mu \neq 1$):* $\mathbf{E}(I_t) \underset{t\to\infty}{\sim} \frac{1}{\mu r}\left(\frac{r_b}{r}\right)^{-\mu}$*: a constant portion of* $\mathbb{R}_+$.

• *if $r_b = \nu\lambda$ ($\mu = 1$):* $\mathbf{E}(I_t) = \frac{1}{r_d}\left[\log\frac{u-1}{u}\right]_{\frac{r_b}{r_d}}^{\frac{r_b}{r_d}e^{rt}} \underset{t\to\infty}{\sim} \frac{1}{r_d}\log\frac{r}{r_b}$*: a constant portion of* $\mathbb{R}_+$.

The *-limit ($\nu \to 0$, $x_t \sim \lambda t \to \infty$, $\nu x_t = \theta$). $r_b, r_d \to 0$, ($r \to 0^+$), $t \to \infty$ in such a way that $r_b t = \kappa > 0$ and $r_d t = \kappa(1 - o(1)) \to \kappa$, so that $rt \to 0$. Suppose in addition $\nu \to 0$ while $\nu / r_b = \theta / (\lambda\kappa)$. Then

$$\Phi_t(z) = \left(1 + \frac{r_b}{r}(e^{rt}-1)(1-z)\right)^{-\nu\lambda/r_b} \to \Phi_\infty(z) = (1 + \kappa(1-z))^{-\theta/\kappa},$$

neg. bin parameters κ and $\overline{\theta} := \theta/\kappa$.

Subcritical case $(r<0)$

$$\Phi_t(z)=\exp\left\{-\nu\lambda(1-z)\int_0^t ds\cdot\frac{e^{rs}}{1+\frac{r_b}{r}\left(e^{rs}-1\right)(1-z)}\right\}\underset{t\to\infty}{\to}\left(1-\frac{r_b}{r}(1-z)\right)^{-\mu},$$

$$\text{Neg.Bin. } \mathbf{E}(N_\infty)=\mu=-\nu\lambda/r,\ \sigma^2(N_\infty)=-\nu\lambda/r\left(1-r_b/r\right).$$

CLAIM: *N_∞ discrete-SD, so unimodal. ($\kappa:=-\frac{r_b}{r}>0$) mode at origin if $\mu<(1+\kappa)/\kappa$, 2 modes at $n=0,1$ if $\mu=(1+\kappa)/\kappa$.*

CLAIM: $\mathbf{P}(P_n=p):=\mathbf{P}(P=p\mid N_\infty=n)=\frac{\mu^p|s_{n,p}|}{[\mu]_n}$, *where $|s_{n,p}|$ are the absolute first-kind Stirling numbers.*

Proba. that p species are being visited when taking an uniform n−sample from the PD(μ) partition of [0,1] representing species abundances with $\infty-$ many species.

Matches with ESF in pop. gen. P_n is # of mutations explaining $N_\infty=n$.

CLAIM (ESF): (i) $\mathbf{P}(P_{n+1}=p+1\mid P_n=p)=\frac{\mu}{\mu+n}$ and $\mathbf{P}(P_{n+1}=p\mid P_n=p)=\frac{n}{\mu+n}$,*gives proba. that a new mutation occurred (the transition $p\to p+1$) or not (the transition $p\to p$) when observing one more terminal mutant (the transition $n\to n+1$), and (ii)* $\frac{P_n}{\log n}\underset{n\to\infty}{\overset{\text{a.s.}}{\to}}\mu$.

CLAIM: If $r<0$, $\mathbf{E}(I_t)=\int_0^t\Phi_s(0)\,ds\underset{t\to\infty}{\sim}\left(1-\frac{r_b}{r}\right)^{-\mu}t$. Constant **fraction** of $\mathbb{R}_+$.

Critical case ($r = 0$)

$$\Phi_t(z) = \exp\left\{-\nu\lambda(1-z)\int_0^t \frac{1}{1+r_b s(1-z)}ds\right\} = (1+r_b t(1-z))^{-\mu}$$

$$SD - \text{Neg. Bin.: } \mathbf{E}(N_t) = \nu\lambda t,\ \sigma^2(N_t) = \nu\lambda t(1+r_b t) \sim \nu\lambda r_b t^2.$$

$$\mathbf{E}\left(e^{-\omega N_t/(\nu\lambda t)}\right) = \left(1+r_b t\left(1-e^{-\omega/(\nu\lambda t)}\right)\right)^{-\mu} \underset{t\to\infty}{\sim} \left(1+\omega\frac{r_b}{\nu\lambda}\right)^{-\mu}, \text{ gamma}(\mu,\mu)$$

- $r_b \to 0$, $\Phi_t(z) \to e^{-\nu\lambda t(1-z)}$, Poisson($\nu\lambda t$).
- $*$–limit: $r_b \to 0$, $x_t = \lambda t \to \infty$, $r_b t \sim r_b x_t/\lambda = \kappa > 0 \Rightarrow \Phi_t(z) \overset{\text{Neg. Bin}}{\to} (1+\kappa(1-z))^{-\mu}$.

Critical case ($r = 0$), c'tnd

$$\begin{aligned} \mathbf{E}\left(I_t\right) &= \int_0^t \Phi_s\left(0\right) ds = \int_0^t \left(1 + r_b s\right)^{-\mu} ds \\ &= \begin{cases} \frac{1}{r_b(1-\mu)} \left(\left(1 + r_b t\right)^{1-\mu} - 1\right) & \text{if } \mu \neq 1 \\ \frac{1}{r_b} \log\left(1 + r_b t\right) & \text{if } \mu = 1. \end{cases} \end{aligned}$$

CLAIM: *All cases:* $\mathbf{E}\left(I_t\right)/t \underset{t\to\infty}{\to} 0$. *Safe!*

- *if* $r_b > \nu\lambda$ ($\mu < 1$): $\mathbf{E}\left(I_t\right) \underset{t\to\infty}{\sim} \frac{1}{r_b^{\mu}(1-\mu)} t^{1-\mu}$: *sub-linear power-law growth.*
- *if* $r_b = \nu\lambda$ ($\mu = 1$): $\mathbf{E}\left(I_t\right) = \frac{1}{r_b} \log\left(1 + r_b t\right) \underset{t\to\infty}{\sim} \frac{1}{r_b} \log t$: *logarithmic growth.*
- *if* $r_b < \nu\lambda$ ($\mu > 1$): $\mathbf{E}\left(I_t\right) \underset{t\to\infty}{\sim} \frac{1}{r_b(\mu-1)}$: *constant portion of* $\mathbb{R}_+$.

Critical case ($r=0$), c'tnd

Variance of I_t: $t_2 > t_1$, $\phi_{t_1,t_2}(z_1,z_2) = \mathbf{E}\left(z_1^{M_{t_1}} z_2^{M_{t_2}}\right) = \phi_{t_1}\left(z_1 \phi_{t_2-t_1}(z_2)\right)$
Joint pgf of (N_{t_1}, N_{t_2}) is (Parzen, Theorem $5A$, page 146):

$$\mathbf{E}\left(z_1^{N_{t_1}} z_2^{N_{t_2}}\right) = \exp -\nu\lambda \left\{ \int_0^{t_1} ds \left(1 - \phi_{t_1-s}\left(z_1 \phi_{t_2-t_1}(z_2)\right)\right) + \int_{t_1}^{t_2} ds \left(1 - \phi_{t_2-s}(z_2)\right) \right\}.$$

$$\mathbf{P}\left(N_{t_1} = 0, N_{t_2} = 0\right) = \Phi_{t_1,t_2}(0,0) = \left(1 + r_b t_1\right)^{-\mu} \left(1 + r_b\left(t_2 - t_1\right)\right)^{-\mu}.$$

CLAIM: $B := \frac{\Gamma(1-\mu)\Gamma(2-\mu)}{\Gamma(3-2\mu)} = B(1-\mu, 2-\mu)$:

- *if* $r_b > \nu\lambda$ $(\mu < 1)$: $\sigma^2(I_t) = \mathbf{E}\left(I_t^2\right) - \mathbf{E}(I_t)^2 \underset{t\to\infty}{\sim} \frac{1}{r_b^{2\mu}(1-\mu)^2}\left(2(1-\mu)B - 1\right) t^{2(1-\mu)}$.

Standard deviation same order as $\mathbf{E}(I_t) \sim \frac{1}{r_b^{\mu}(1-\mu)} t^{1-\mu}$.

- *if* $r_b < \nu\lambda$ $(\mu > 1$, *non-int.*$)$: $\sigma^2(I_t) \underset{t\to\infty}{\sim} \frac{1}{r_b^2(\mu-1)^2}$: *here,* $I_t \underset{t\to\infty}{\overset{d}{\to}}$ *finite non-degen. rv.*

CLAIM: *If* $\mu = 1$,

$$\sigma^2(I_t) \underset{t\to\infty}{\sim} \left(\frac{\log t}{r_b}\right)^2 \underset{t\to\infty}{\sim} \mathbf{E}(I_t)^2.$$

critical, c'tnd

Covariances of the vacancy process $\{B_t\} = \{\mathbf{1}(N_t = 0)\}$**:** With $B_{t_1} := \mathbf{1}(N_{t_1} = 0)$ and $B_{t_2} := \mathbf{1}(N_{t_2} = 0)$, $t_2 > t_1 > 0$. With $\tau = t_2 - t_1 > 0$, for each fixed t_1, we have (long-range power-law covariances)

CLAIM:

$$0 < \mathrm{Cov}\left(B_{t_1}, B_{t_1+\tau}\right) \underset{\text{large } \tau}{\sim} C\left(t_1\right)\tau^{-(1+\mu)}.$$

THANK YOU