A Gaussian particle distribution for branching Brownian motion with an inhomogeneous branching rate

by Jason Schweinsberg
University of California San Diego

Joint work with Matt Roberts

Outline of Talk

1. Motivation: a discrete population model
2. A branching Brownian motion process
3. Large deviations heuristics
4. Main tools and proof techniques
5. Main results
6. Future work

Motivation: a discrete population model

Consider a population undergoing selection. Individuals acquire beneficial mutations which increase their fitness. One example of such a model:

- The population has N individuals.
- Each individual independently acquires mutations at rate μ_N.
- Mutations are beneficial. An individual with j mutations (a type j individual) at time t has fitness

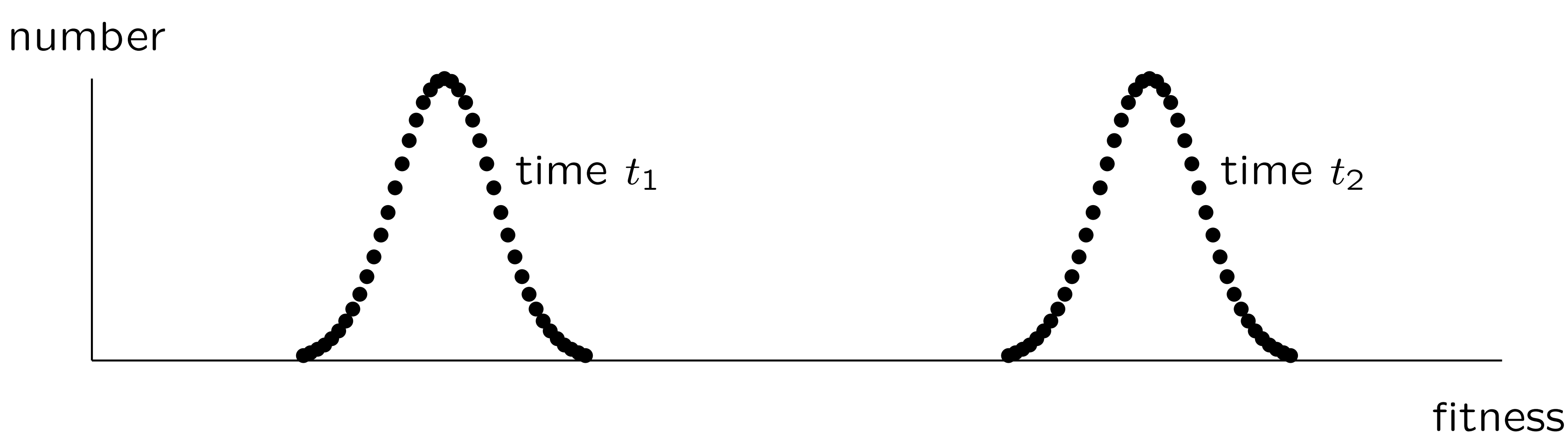
  $$\max\{1 + s_N(j - M(t)),\ 0\},$$

  where $M(t)$ is the average number of mutations of the N individuals at time t. ($s_N =$ fitness benefit from a mutation)
- Each individual independently lives for an exponential(1) time.
- When an individual dies, its replacement is chosen at random from the population, with probability proportional to fitness.

A Gaussian traveling wave

Conjecture: The empirical distribution of the fitness levels of individuals at a fixed time t is Gaussian, leading to a "Gaussian traveling wave":

Tsimring, Levine, and Kessler (1996)
Rouzine, Wakeley, and Coffin (2003)
Desai and Fisher (2007)
Beerenwinkel et. al. (2007)
Brunet, Rouzine, and Wilke (2008)



Rigorous results

1. If $\mu_N \ll 1/(N \log N)$ and $s_N = s > 0$, then selective sweeps happen one at a time.
2. Durrett and Mayberry (2011) considered the case $\mu_N \sim N^{-\beta}$, where $0 < \beta < 1$, and $s_N = s > 0$. They established traveling wave behavior, but only finitely many types present at a time.
3. Schweinsberg (2017) considered faster mutation rates with $\mu_N \sim e^{-(\log N)^c}$ and $e^{-(\log N)^{1-c-\varepsilon}} \ll s_N \leq (\log N)^{-1/2}$, with $1/2 < c < 1$. Fitness distribution has Gaussian-like tails, but most individuals have the same type, as Desai and Fisher (2007) showed by nonrigorous methods.

Can we formulate a population model with selection in which fitness distribution converges to a Gaussian?

Consider faster mutation rates. Fitness of an individual changes often due to mutation, like a random walk. We model this by Brownian motion.

Related to infinitesimal model. See Barton, Etheridge, Véber (2017).

Branching Brownian motion with inhomogeneous branching rate

Begin with some configuration of particles on $\mathbb{R}$ at time zero.

Think of the particle as being an individual in a population, and position of the particle is the fitness.

Each particle independently moves according to one-dimensional Brownian motion with drift $-\rho$.

Each particle dies at rate $d(x) = 1$.
A particle at x splits into two at rate $b(x) = 1 + \beta x$.
For $x < -1/\beta$, set $b(x) = 0$, $d(x) = -\beta x$, so $b(x) - d(x) = \beta x$.

As long as $b(x) - d(x) = \beta x$, our results hold if there exists $C > 0$ such that $d(x) \geq C$ for all $x \in \mathbb{R}$ and $b(x) \leq 1/C$ for all $x \leq 1/\beta$.

Main result: Under certain conditions on β and ρ and the initial configuration of particles, after a sufficiently long time, the empirical distribution of particles is approximately Gaussian.

Related work

Model was studied by Neher and Hallatschek (2013).

Brunet, Derrida, Mueller, and Munier (2006) and Berestycki, Berestycki, and Schweinsberg (2013) used branching Brownian motion with killing at 0 to model selection. Particles cluster near 0, does not lead to Gaussian fitness distribution.

Beckman (2019) considered branching Brownian motion with an inhomogeneous branching rate, focused on a shorter time scale.

Harris and Harris (2009) considered branching Brownian motion with no drift, no deaths, and branching at rate $b(x) = \beta|x|^p$. Let R_t be the position of the right-most particle at time t.

1. If $0 < p < 2$, then $t^{-2/(2-p)} R_t \to c$ a.s.
2. If $p = 2$, then $t^{-1} \log R_t \to \sqrt{2\beta}$ a.s.
3. If $p > 2$, the process explodes in finite time.

Berestycki, Brunet, Harris, Harris, Roberts (2015) obtained more detailed results about the number of particles in different regions using large deviations techniques.

Large deviations heuristics

Schilder's Theorem: Let $f : [0,T] \to \mathbb{R}$. The probability that Brownian motion with drift $-\rho$, started from $f(0)$, stays "close" to the function f until time T is approximately

$$\exp\left(-\frac{1}{2}\int_0^T (f'(u)+\rho)^2 \, du \right).$$

Many-to-one Lemma: Recall $b(x) - d(x) = \beta x$. The expected number of particles in branching Brownian motion at time T for which the trajectory of the ancestor stays close to f is roughly

$$\exp\left(\int_0^T \beta f(u) - \frac{1}{2}(f'(u)+\rho)^2 \, du \right).$$

The actual number of particles that stay close to f will be comparable to the expectation, as long as the integral up to t is nonnegative for all $t \in [0,T]$.

If $f(u) = \rho^2/2\beta$ for all $u \in [0,T]$, then the integrand is zero. If we start with one particle near $\rho^2/2\beta$, right-most particle stays near $\rho^2/2\beta$, if the population does not quickly die out.

The Gaussian shape

Start with one particle at $\rho^2/2\beta$. The number of particles near y at time T is approximately

$$\exp\left(\int_0^T \beta f_y(u) - \frac{1}{2}(f_y'(u) + \rho)^2\, du\right),$$

where the function f_y maximizes the integral subject to the conditions $f_y(0) = \rho^2/2\beta$, $f_y(T) = y$, and that the integral up to t is nonnegative for all $t \in [0, T]$.

For sufficiently large T, the number of particles near y is approximately $\exp(g(y))$, where

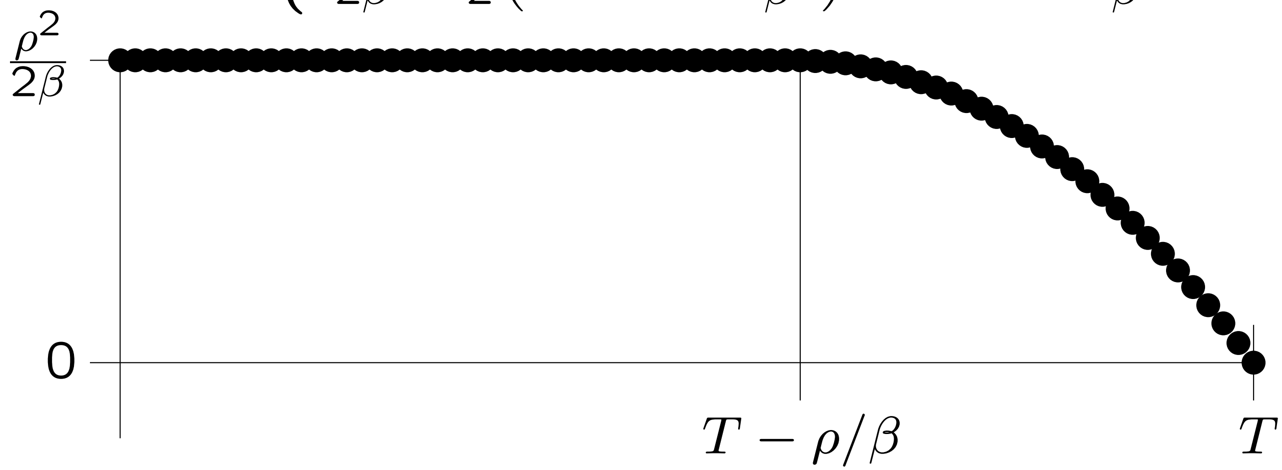
$$g(y) = \frac{\rho^3}{2\beta} - \rho y - \frac{2\sqrt{2\beta}}{3}\left(\frac{\rho^2}{2\beta} - y\right)^{3/2}.$$

Because $g'(0) = 0$, $g''(0) = -\beta/\rho$, the distribution of particles at time T is approximately Gaussian with mean 0 and variance ρ/β.

The trajectory of particles

Particles near 0 at time t typically follow the trajectory f_0:

$$f_0(t) = \begin{cases} \frac{\rho^2}{2\beta} & \text{if } 0 \leq t \leq T - \frac{\rho}{\beta} \\ \frac{\rho^2}{2\beta} - \frac{\beta}{2}\Big(u - (T - \frac{\rho}{\beta})\Big)^2 & \text{if } T - \frac{\rho}{\beta} \leq u \leq T \end{cases}$$



Particles with highest fitness should be close to $\rho^2/2\beta$.
It takes approximately time ρ/β for Gaussian shape to emerge.

For these heuristics to be accurate, we need:

1. The standard deviation of the Gaussian $\sqrt{\rho/\beta}$ must be much smaller than $\rho^2/2\beta$. That is, $\rho^3 \gg \beta$.
2. βx should be small when $x = \rho^2/2\beta$. That is, $\rho \ll 1$.

More precise methods are needed to provide rigorous proofs.

Density calculations

Define $p_t(x,y)$ so that if there is one particle at x at time zero, the expected number of particles in A at time t is

$$\int_A p_t(x,y)\,dy.$$

Define $q_t(x,y)$ to be the density for the process with no drift.

Let $(B_t)_{t\geq 0}$ be Brownian motion. By the Many-to-one Lemma,

$$\begin{aligned} q_t(x,y) &= E_x\left[\exp\left(\int_0^t \beta B_s\,ds\right); B_t \in dy\right] \\ &= \frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{(y-x)^2}{2t} + \frac{\beta(y+x)t}{2} + \frac{\beta^2 t^3}{24}\right). \end{aligned}$$

See, for example, Borodin and Salminen (1996).

A Girsanov transformation gives

$$p_t(x,y) = \exp\left(\rho x - \rho y - \frac{\rho^2 t}{2}\right) q_t(x,y).$$

Moment calculations

Let $N(t)$ be the number of particles at time t, and denote by $X_1(t) \geq X_2(t) \geq \cdots \geq X_{N(t)}(t)$ the positions of particles.

If $f : (0, \infty) \to \mathbb{R}$, then

$$E_x\left[\sum_{i=1}^{N(t)} f(X_i(t))\right] = \int_0^\infty f(y) p_t(x, y)\, dy.$$

Take $f = \mathbf{1}_A$ to get expected number of particles in a set A.

One can also calculate second moments to estimate fluctuations (Ikeda, Nagasawa, Watanabe (1969)):

$$E_x\left[\left(\sum_{i=1}^{N(t)} f(X_i(t))\right)^2\right] = \int_0^\infty f(y)^2 p_t(x, y)\, dy +$$

$$2 \int_0^t \int_0^\infty \int_0^\infty \int_0^\infty f(y_1) f(y_2) p_s(x, z) p_{t-s}(z, y_1) p_{t-s}(z, y_2)\, dy_1\, dy_2\, dz\, ds.$$

Moment calculations can be dominated by rare events in which a particle moves far to the right and has many offspring. We need to apply truncation: kill particles that reach L.

Density calculations with truncation

Let $q_t^L(x,y)$ be the density with no drift, killing at L. Then

$$\frac{\partial}{\partial t} q_t^L(x,y) = \frac{1}{2}\frac{\partial^2}{\partial y^2} q_t^L(x,y) + \beta y q_t^L(x,y).$$

Looking for solutions of the form $e^{\lambda t} h(y)$, we have

$$\lambda h(y) = \frac{1}{2} h''(y) + \beta y h(y), \quad h(L) = 0, \quad \lim_{y \to -\infty} h(y) = 0.$$

Let $\cdots < \gamma_2 < \gamma_1 < 0$ denote the zeros of the Airy function. Let $\varphi_k(x) = Ai((2\beta)^{1/3}(L - x) + \gamma_k)$. Then (Salminen, 1988):

$$p_t^L(x,y) = (2\beta)^{1/3} \sum_{k=1}^{\infty} \frac{e^{(\beta(L+(2\beta)^{-1/3}\gamma_k) - \rho^2/2)t}}{(Ai'(\gamma_k))^2} e^{\rho x} \varphi_k(x) e^{-\rho y} \varphi_k(y).$$

Choose

$$L = \frac{\rho^2}{2\beta} - (2\beta)^{-1/3}\gamma_1.$$

If t is large enough that the $k = 1$ term dominates, then

$$p_t^L(x,y) \approx \frac{(2\beta)^{1/3}}{(Ai'(\gamma_1))^2} e^{\rho x} \varphi_1(x) e^{-\rho y} \varphi_1(y).$$

Observations related to density formula

$$p_t^L(x,y) \approx \frac{(2\beta)^{1/3}}{(Ai'(\gamma_1))^2}\, e^{\rho x}\varphi_1(x)\, e^{-\rho y}\varphi_1(y).$$

1) Let

$$Z(t) = \sum_{i=1}^{N(t)} e^{\rho X_i(t)} Ai((2\beta)^{1/3}(L - X_i(t)) + \gamma_1) \mathbb{1}_{\{X_i(t)<L\}}.$$

When particles at L are killed, $(Z(t), t \geq 0)$ is a martingale. We can use $Z(t)$ to measure the "size" of the process at time t.

2) The density of particles near y is roughly proportional to

$$e^{-\rho y} Ai((2\beta)^{1/3}(L - y) + \gamma_1).$$

Because

$$Ai(z) \sim z^{-1/4} e^{-(2/3)z^{-3/2}},$$

this is roughly $\exp(g(y))$,

$$g(y) = \frac{\rho^3}{2\beta} - \rho y - \frac{2\sqrt{2\beta}}{3}\left(\frac{\rho^2}{2\beta} - y\right)^{3/2}.$$

This is the same g obtained from the large deviations heuristic.

Strategy for studying configuration of particles

To study the particles within $O(\beta^{-1/3})$ of L, consider the process in which particles are killed at L:

- Show that particles do not reach L frequently.
- Use first moment calculations to show that the "density" of particles near y is proportional to $e^{-\rho y} Ai((2\beta)^{1/3}(L-y)+\gamma_1)$.
- Use second moment bounds to control fluctuations, show configuration of particles near L is relatively stable over time.

To study the configuration of particles near 0 at a large time T:

- Observe that particles near 0 at time T are descended from ancestors within $O(\beta^{-1/3})$ of L at time $T - \rho/\beta$. These ancestors are in a stable configuration as noted above.
- Use additional first and second moment estimates to show the empirical distribution of particles at time T is approximately Gaussian with mean 0 and variance ρ/β.

Main Theorem (Roberts and Schweinsberg, 2020)

Consider a sequence of branching Brownian motion processes indexed by n. Suppose

$$\lim_{n\to\infty} \frac{\rho_n^3}{\beta_n} = \infty, \qquad \lim_{n\to\infty} \rho_n = 0.$$

Assume the initial configuration is such that

$$Z_n(0) \asymp_p \frac{\beta_n^{1/3}}{\rho_n^3} e^{\rho_n L_n},$$

where $\asymp_p$ means the ratio of the two sides is tight, and

$$Y_n(t) = \sum_{i=1}^{N_n(t)} e^{\rho_n X_{i,n}(t)}, \qquad \rho_n^2 e^{-\rho_n L_n} Y_n(0) \to_p 0.$$

For $t \geq 0$, define the random probability measure

$$\zeta_n(t) = \frac{1}{N_n(t)} \sum_{i=1}^{N_n(t)} \delta_{X_{i,n}(t)\sqrt{\beta_n/\rho_n}}.$$

Let $1 < a < \infty$, and let $t_n = a\rho_n/\beta_n$. Let μ be the standard normal distribution. Then $\zeta_n(t_n) \Rightarrow \mu$ as $n \to \infty$.

Interpretation of drift

Long-run empirical distribution of particle locations is approximately normal with mean 0 and variance ρ/β.

Our model is equivalent to one with no drift and, at time t,

$$b(x) - d(x) = \beta(x - \rho t).$$

Then the empirical distribution of particles after a long time t will be approximately normal with mean ρt. That is, ρ is the speed at which the Gaussian traveling wave advances.

Fisher's Fundamental Theorem of Natural Selection (1930): rate at which fitness increases = variance of fitness distribution.

1. Rate at which fitness increases: $\beta\rho$.
2. Variance of fitness distribution: $\beta^2(\rho/\beta) = \beta\rho$.

Note: Number of particles in long-run is

$$N \asymp \frac{\beta^{1/3}}{\rho^3} \exp\left(\frac{\rho^3}{6\beta} - \rho(2\beta)^{-1/3}\gamma_1\right),$$

so we can solve for ρ given N and β.

Particle configurations near right edge

(Roberts and Schweinsberg, 2020)

Most particles within $O\left(\sqrt{\rho/\beta}\right)$ of the origin.

Next result considers particles within $O(\beta^{-1/3})$ of L, which are the particles that will have descendants alive far into the future.

Make the same assumptions on the parameters and initial conditions as before. For $t \geq 0$, define the random probability measure

$$\xi_n(t) = \frac{1}{Y_n(t)} \sum_{i=1}^{N_n(t)} e^{\rho_n X_{i,n}(t)} \delta_{(2\beta_n)^{1/3}(L_n - X_{i,n}(t))}.$$

Let $0 < a < \infty$, and let $t_n = a\rho_n/\beta_n$. Then $\xi_n(t_n) \Rightarrow \nu$, where ν is the probability measure on $(0, \infty)$ with probability density function

$$h(y) = \frac{Ai(y + \gamma_1)}{\int_0^\infty Ai(z + \gamma_1)\, dz}.$$

Future Work

Genealogy of the Population

- We conjecture that the genealogy of the population is given by the Bolthausen-Sznitman coalescent. This was predicted by Neher and Hallatschek (2013).
- This should follow from arguments similar to those used by Berestycki, Berestycki, and Schweinsberg (2013) in the case of branching Brownian motion with absorption at zero.

Discrete Population Models

- Results for branching Brownian motion should hold for the discrete population model when μ_N is large and s_N is small.
- Correspondence should hold even if s_N is negative or random.