

An inhomogeneous stochastic rate process for evolution from states in an information geometric neighbourhood of uniform fitness*

C.T.J. Dodson

School of Mathematics, University of Manchester, Manchester M13 9PL, UK
ctdodson@manchester.ac.uk

Abstract

This study elaborates some examples of a simple evolutionary stochastic rate process where the population rate of change depends on the distribution of properties—so different cohorts change at different rates. We investigate the effect on the evolution arising from parametrized perturbations of uniformity for the initial inhomogeneity. The information geometric neighbourhood system yields also solutions for a wide range of other initial inhomogeneity distributions, including approximations to truncated Gaussians of arbitrarily small variance and distributions with pronounced extreme values. It is found that, under quite considerable alterations in the shape and variance of the initial distribution of inhomogeneity in unfitness, the decline of the mean does change markedly with the variation in starting conditions, but the net population evolution seems surprisingly stable.

Keywords: Evolution, inhomogeneous rate process, information geometry, entropy, uniform distribution, log-gamma distribution.

MSC2000: 62P10 92D15

1 Introduction

Consider a population with an inhomogeneous property distribution faced with a sudden environmental change at $t = 0$ and let $N(t)$ represent the declining population of more unfit individuals. In our example the distribution of unfitness a will lie in the range $(0, 1]$ and under selective evolution the fraction $N(t)$ could consist of individuals with unfitness below some threshold value which itself may evolve. However, as $N(t)$ evolves we expect the distribution of unfitness to become skewed increasingly towards smaller values, so improving population fitness. We make use of existing inhomogeneous stochastic rate process theory [9, 10], and information geometry, [1, 2, 5, 6]. In the example we elaborate, the inhomogeneous population N is classified by a smooth family of probability density functions $\{P_t, t \geq 0\}$ with random variable $0 \leq a \leq 1$, having mean $E_t(a)$ and variance $\sigma_t^2(a) = E_t(a^2) - (E_t(a))^2$. Here a represents an unfitness that controls the decline in the frequency of the a -cohort, so we could view fitness as the variable $1 - a$. The model can be reformulated for a vector $(N^i(t))$ representing a composite population with a vector of distributions (P_t^i) and a matrix of fitness variables $[a_{ij}]$.

Let $l_t(a)$ represent the frequency at the a -cohort, then we have

$$N(t) = \int_0^1 l_t(a) da \quad \text{and} \quad P_t(a) = \frac{l_t(a)}{N(t)} \quad (1)$$

$$\frac{dl_t(a)}{dt} = -al_t(a) \quad \text{so} \quad l_t(a) = l_0(a)e^{-at} \quad (2)$$

*Invited paper at 3rd Conference on Information Geometry and its Applications, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig, 2-6 August 2010.

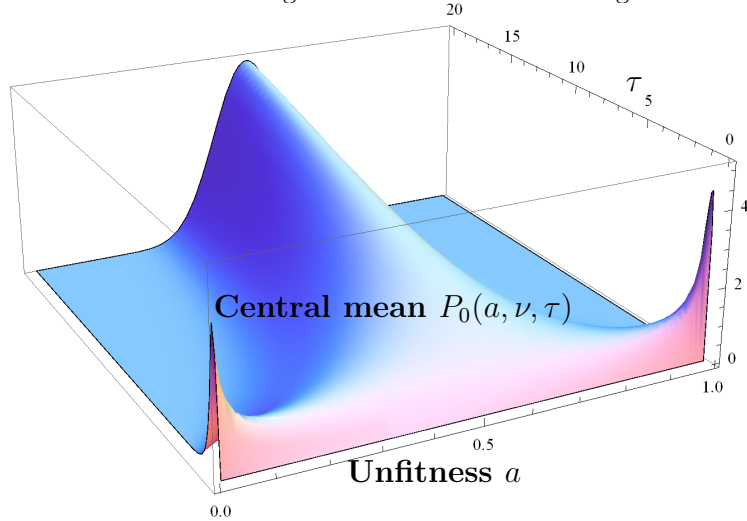


Figure 1: The log-gamma family of probability densities $P_0(a, \nu, \tau)$ from (11) as a surface for the case of central mean $E_0(a) = \frac{1}{2}$. This symmetric surface coincides with the uniform density 1 at $\tau = 1$, and tends to the delta function as $\tau \rightarrow \infty$

Karev [9] obtained general solutions for these equations giving us

$$N(t) = N(0)L_0(t) \quad \text{where } L_0(t) = \int_0^\infty P_0(a)e^{-at} da \quad (3)$$

$$\frac{dN}{dt} = -E_t(a)N \quad \text{where } E_t(a) = \int_0^\infty a P_t(a) da = -\frac{d \log L_0}{dt} \quad (4)$$

$$\frac{dE_t(a)}{dt} = -\sigma_t^2(a) = (E_t(a))^2 - E_t(a^2) \quad (5)$$

$$P_t(a) = e^{-at} \frac{P_0(a)}{L_0(t)} \quad \text{and } l_t(a) = e^{-at} L_0(t) \quad (6)$$

$$\frac{dP_t(a)}{dt} = P_t(a)(E_t(a) - a). \quad (7)$$

Here $L_0(t)$ is the Laplace transform of the initial probability density function $P_0(a)$ (which is of course zero outside $(0, 1]$) and so conversely $P_0(a)$ is the inverse Laplace transform of the population (monotonic) decay solution $\frac{N(t)}{N(0)}$. See Feller [7] for more discussion of the existence and uniqueness properties of the correspondence between probability densities and their Laplace transforms. We see from (7) that when the unfitness a exceeds its mean $E_t(a)$ then the population density of that cohort declines and conversely the densities of cohorts with $a < E_t(a)$ tend to grow. The Shannon entropy at time t is

$$S_t = -E_t(\log P_t(a)) = -E_t \left(\log \frac{P_0(a)e^{-at}}{L_0(t)} \right) \quad (8)$$

which reduces to

$$S_t = S_0 + \log L_0(t) + E_t(a)t. \quad (9)$$

By using $\frac{dE_t}{dt} = -\sigma_t^2(a)$, the decay rate is then

$$\frac{dS_t}{dt} = -\sigma^2(t)t. \quad (10)$$

This shows how the variance controls the entropy change during quite general inhomogeneous population processes, as we saw in [6]. The system

Karev gave the particular solutions for the cases of initial densities that were Poisson, gamma or uniform. In [6] we studied the case of a bivariate gamma density for an epidemic situation. Here we shall study a family of log-gamma densities that determine a neighbourhood of the uniform distribution [2], so recovering the solution of Karev for the uniform distribution as a special case

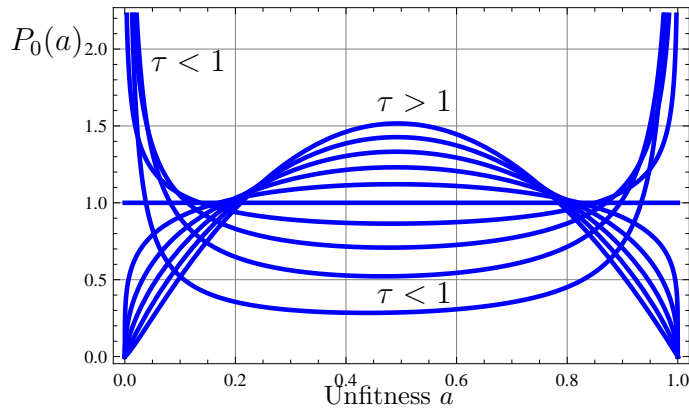


Figure 2: Log-gamma probability density functions $P_0(a)$ from (11) for $a \in (0, 1]$, with central mean $E_0(a) = \frac{1}{2}$, and τ from 0.2 to 2 in steps of 0.2. Note that the parameter τ controls the shape of the graph and for $\tau = 1$ we have $P_0(a) = 1$.

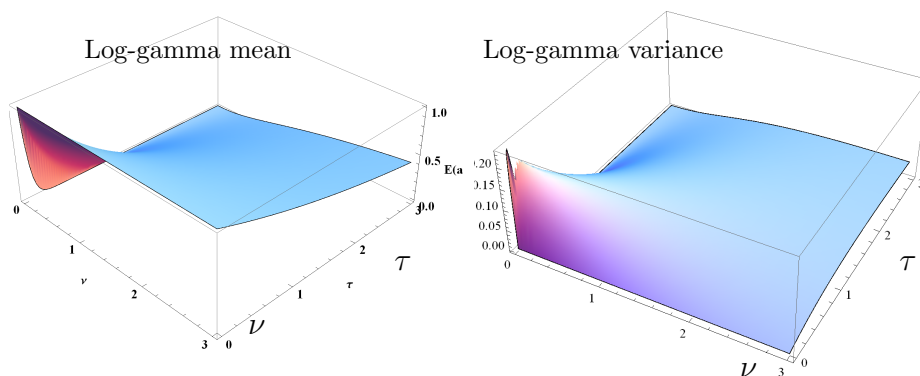


Figure 3: Mean and variance for the log-gamma family.

($\nu = \tau = 1$). The log-gamma family contains also close approximations to Gaussians of arbitrarily small variance truncated to $(0, 1]$ (for $\tau > 1$) as well as densities having outlier cohorts of extreme values that are enhanced ($\tau < 1$), Figure 1.

2 Evolution of inhomogeneity from an initial log-gamma probability density

We studied in [2] the smooth family of log-gamma distributions and their information geometry. This family has probability density function

$$P_0(a, \nu, \tau) = \frac{a^{\nu-1} \nu^\tau |\log(\frac{1}{a})|^{\tau-1}}{\Gamma(\tau)} \quad (11)$$

for random variable $a \in (0, 1]$ and parameters $\nu, \tau > 0$. The mean and variance, Figure 3, are given by

$$E_0(a) = \left(\frac{\nu}{1+\nu} \right)^\tau \quad (12)$$

$$\sigma_0^2(a) = \left(\frac{\nu}{\nu+2} \right)^\tau - \left(\frac{\nu}{1+\nu} \right)^{2\tau}. \quad (13)$$

The locus in this family of those with central mean $E_0(a) = \frac{1}{2}$ satisfies

$$\nu(2^{\frac{1}{\tau}} - 1) = 1 \quad (14)$$

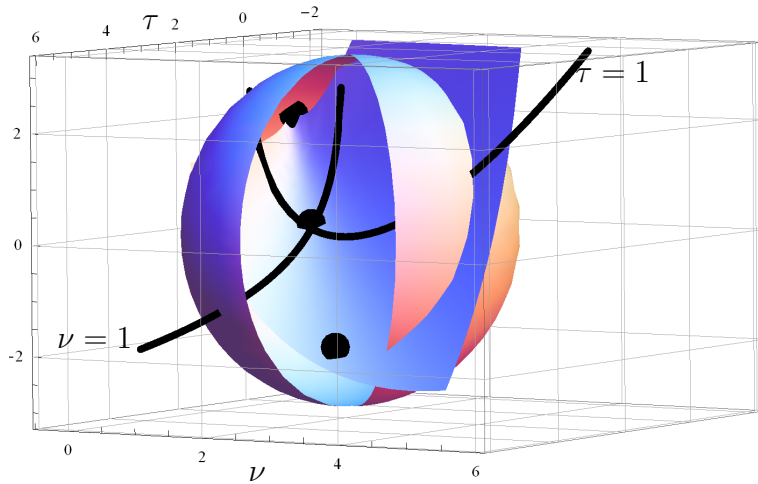


Figure 4: An affine immersion in \mathbb{R}^3 of the 2-manifold of log-gamma probability densities. The black curves in the immersion represent the distributions with $\nu = 1$ and $\tau = 1$ and centred on their intersection is a spherical neighbourhood of the uniform distribution. The other two points are for the two cases $(\nu = 0.1, \tau = 0.289)$ and $(\nu = 2.75, \tau = 2.24)$ cf. Figures 6, 7, 8

and some are shown in Figure 1 The uniform density is the special case with $\tau = \nu = 1$ and this can be seen in Figure 2, with sections through the surface of Figure 1. The log-gamma family yields a smooth Riemannian 2-manifold with coordinates $(\nu, \tau) \in \mathbb{R}^+ \times \mathbb{R}^+$ and metric tensor given by

$$[g_{ij}](\nu, \tau) = \begin{bmatrix} \frac{\tau}{\nu^2} & -\frac{1}{\nu} \\ -\frac{1}{\nu} & \frac{d^2}{d\tau^2} \log(\Gamma) \end{bmatrix}. \quad (15)$$

In fact, this manifold is an isometric diffeomorph of the 2-manifold of gamma densities, with random variable $x = -\log a$ via natural coordinates (ν, τ) (cf. [1]) for which $\tau = 1$ corresponds to the subfamily of exponential densities. Through this smooth diffeomorphism we can therefore represent the manifold of log-gamma densities as a natural affine immersion (cf. [2]) of $\mathbb{R}^+ \times \mathbb{R}^+$ in \mathbb{R}^3

$$(\nu, \tau) \mapsto \{\nu, \tau, \log \Gamma(\tau) - \tau \log \nu\}. \quad (16)$$

This is illustrated in Figure 4 which shows also a spherical neighbourhood in \mathbb{R}^3 centred on the point at the uniform density, $\nu = 1 = \tau$; the other two points are for the two cases $(\nu = 0.1, \tau = 0.289)$ and $(\nu = 2.75, \tau = 2.24)$ which represent densities also having mean value $\frac{1}{2}$ and used in the sequel as initial unfitness distributions, cf. Figures 6, 7, 8.

The log-gamma entropy (8) is given by

$$\begin{aligned} S_{LG}(\nu, \tau) &= \nu^\tau (\nu + 1)^{-\tau-1} (A(\nu, \tau) + B(\nu, \tau)) \\ \text{with } A(\nu, \tau) &= \tau(\nu + (\nu + 1) \log(\nu + 1) - 1) - (\nu + 1)(\tau - 1)\psi(\tau) \\ \text{and } B(\nu, \tau) &= (\nu + 1) \log \left(\frac{\nu^{-\tau} \Gamma(\tau)}{\nu + 1} \right), \quad \psi(\tau) = \frac{d \log \Gamma(\tau)}{d\tau}. \end{aligned} \quad (17)$$

Figure 5 shows the graph as a surface and as a contour plot with gradient vector field indicated; the main curvature occurs near the origin.

The Laplace transform integral (3) for the log-gamma density with general (ν, τ) seems intractable so we used a series development up to sixth order for the term e^{-at} , bearing in mind that $a \in (0, 1]$. Thus, up to sixth order in $t \geq 0$ we obtain a good approximation to $L_0(t)$, accurate to about 0.1% up to $t = 1$. The corresponding expressions for $L_0(t)$, $P_t(a)$, $E_t(a)$, $\sigma_t^2(a)$ and hence $N(t)$ are known but somewhat cumbersome to present here so we present some graphics for illustration. The decline of the mean $E_t(a)$ as the density $P_t(a)$ develops from three initial

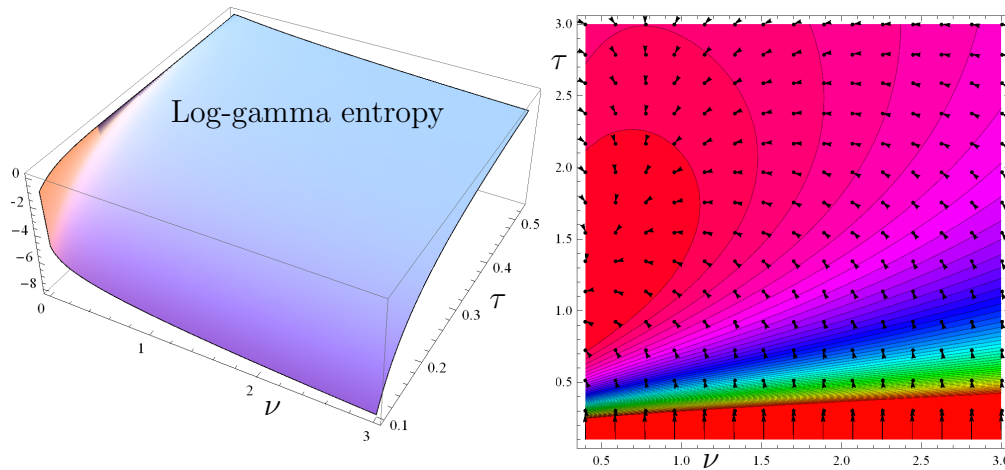


Figure 5: Shannon entropy function for the log-gamma family as a surface (left) and as a contour plot with entropy gradient flow (right).

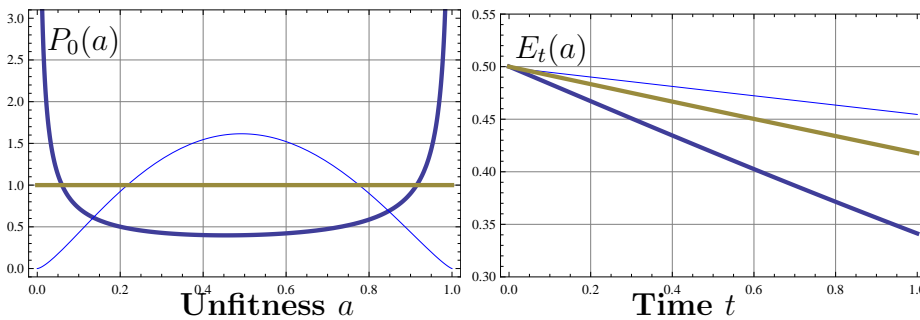


Figure 6: Initial log-gamma densities $P_0(a)$ shown in left panel with central mean $E_0(a) = \frac{1}{2}$ for the cases $\tau = 0.289, \nu = 0.1$ (lower graph), $\tau = 2.24, \nu = 2.75$ (upper graph), and the uniform density $P_0(a) = 1$ for $\tau = \nu = 1$. The right panel shows the decline with time of the mean $E_t(a)$ from (4) for these initial densities.

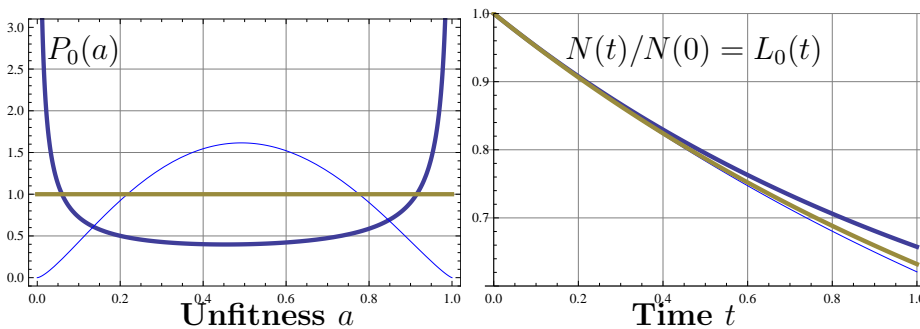


Figure 7: Initial log-gamma densities $P_0(a)$ shown in left panel with central mean $E_0(a) = \frac{1}{2}$ for the cases $\tau = 0.289, \nu = 0.1$ (lower graph), $\tau = 2.24, \nu = 2.75$ (upper graph), and the uniform density $P_0(a) = 1$ for $\tau = \nu = 1$. The right panel shows the fractional decline with time of the population $N(t)/N(0) = L_0(t)$ from (3) for these initial densities.

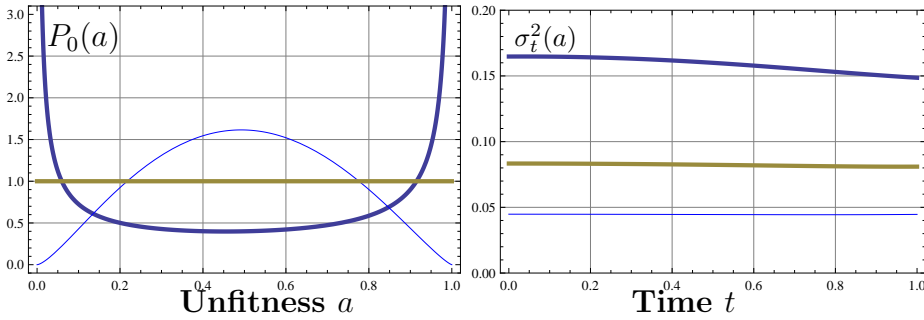


Figure 8: Initial log-gamma densities $P_0(a)$ shown in left panel with central mean $E_0(a) = \frac{1}{2}$ for the cases $\tau = 0.289$, $\nu = 0.1$ (lower graph), $\tau = 2.24$, $\nu = 2.75$ (upper graph), and the uniform density $P_0(a) = 1$ for $\tau = \nu = 1$. The right panel shows the decline with time of the variance $\sigma_t^2(a)$ from (5) for these initial densities.

log-gamma densities $P_0(a)$ with central mean, $E_0(a) = \frac{1}{2}$ is shown in Figure 6. The more rapid decline occurs for the initial density with $\tau = 0.289$ and the slower decline occurs for the case having $\tau = 2.24$ with the uniform case $\tau = \nu = 1$ in between, cf. also Figure 4. The corresponding decline of $N(t)/N(0) = L_0(t)$ is shown in Figure 7 and the variance is shown in Figure 8.

We do have an analytic solution for the special 1-parameter family of log-gamma densities with $\tau = 1$ (the ‘log-exponential’ densities), from (3) and (6):

$$P_t(a, \nu, 1) = \frac{e^{-at} a^{\nu-1} t^\nu}{\Gamma(\nu) - \Gamma(\nu, t)} \quad (18)$$

$$E_t(a, \nu, 1) = \frac{\Gamma(\nu + 1) - \Gamma(\nu + 1, t)}{t(\Gamma(\nu) - \Gamma(\nu, t))} \quad (19)$$

$$\frac{N(t)}{N(0)} = L_0(t) = \nu t^{-\nu} (\Gamma(\nu) - \Gamma(\nu, t)) \quad (20)$$

where $\Gamma(\nu, t)$ is the incomplete gamma function, cf Figure 9. $E_t(a, \nu, 1)$, declines for all choices of parameter ν in the initial density, which includes of course also the uniform density. The evolving family of probability densities (18) has the two positive parameters ν, t and Figure 10 shows the case for $\nu = 10$, $\tau = 1$. The corresponding information metric is known analytically, but the expressions for components $g_{ij}(\nu, t)$ are cumbersome, involving MeijerG functions. Figure 11 shows the evolving approximate information distance in this metric

$$D(\Delta\nu, \Delta t) = \sqrt{g_{11}(\nu, t)\Delta\nu^2 + g_{12}(\nu, t)\Delta\nu\Delta t + g_{22}(\nu, t)\Delta t^2} \quad (21)$$

measured from the initial log-gamma distributions with $\tau = 1, \nu = 0.5$ (left) and $\tau = 1, \nu = 10$ (right), near $t = 2$.

Further analytic solutions arise for other initial log-gamma densities with positive integer τ (the ‘log-Pearson Type III’ distributions) cf. Figure 2, as generalized hypergeometric functions; for example,

$$\tau = 2 \Rightarrow \frac{N(t)}{N(0)} = L_0(t) = {}_2F_2(\nu, \nu; \nu + 1, \nu + 1; -t) \quad (22)$$

$$\tau = 3 \Rightarrow \frac{N(t)}{N(0)} = L_0(t) = {}_3F_3(\nu, \nu, \nu; \nu + 1, \nu + 1, \nu + 1; -t). \quad (23)$$

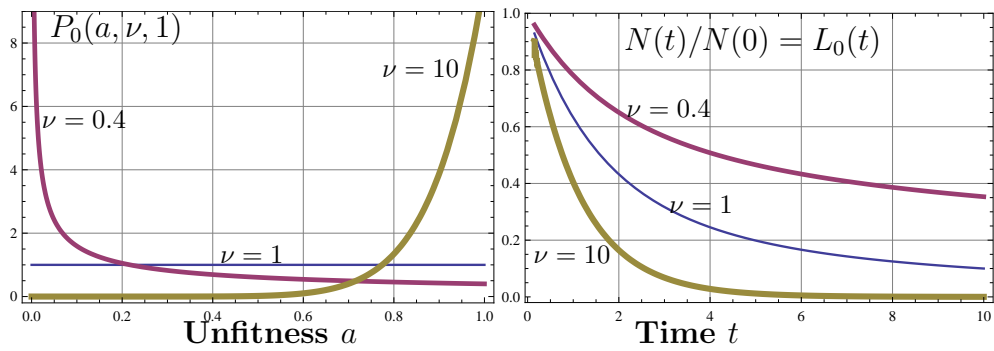


Figure 9: Initial log-gamma densities $P_0(a, \nu, 1)$ from (18) are shown in the left panel for the uniform density $\tau = \nu = 1$, and also for $\tau = 1$ with $\nu = 0.4$ and 10; Figure 10 shows the evolution of the case with $\nu = 10$. The right panel shows the corresponding fractional decline with time of the population $N(t)/N(0)$ from (20) for these initial densities. Figure 10 shows the evolution from $P_0(a, 10, 1)$.

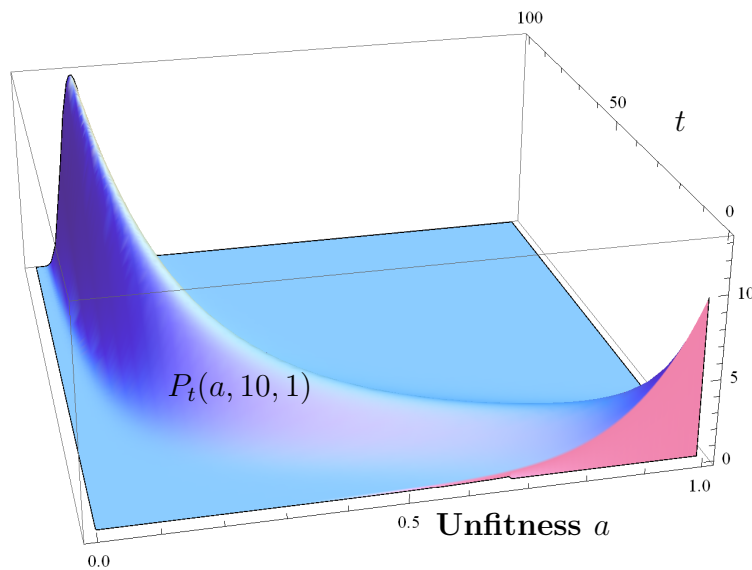


Figure 10: The evolution of probability density $P_t(a, 10, 1)$ from (18).

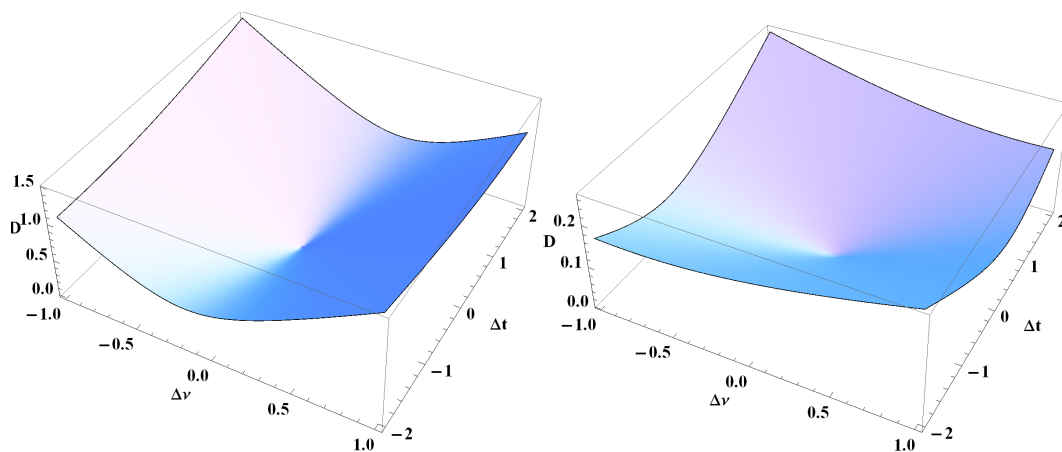


Figure 11: 11. The evolving approximate information distance D measured from the initial log-gamma distributions with $\tau = 1, \nu = 0.5$ (left) and $\tau = 1, \nu = 10$ (right), near $t = 2$.

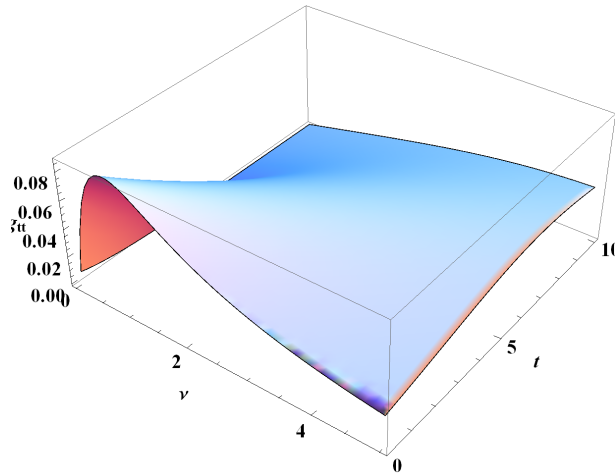


Figure 12: The time component $g_{tt}(\nu, t)$ of the evolving information metric for log-gamma with $\tau = 1$.

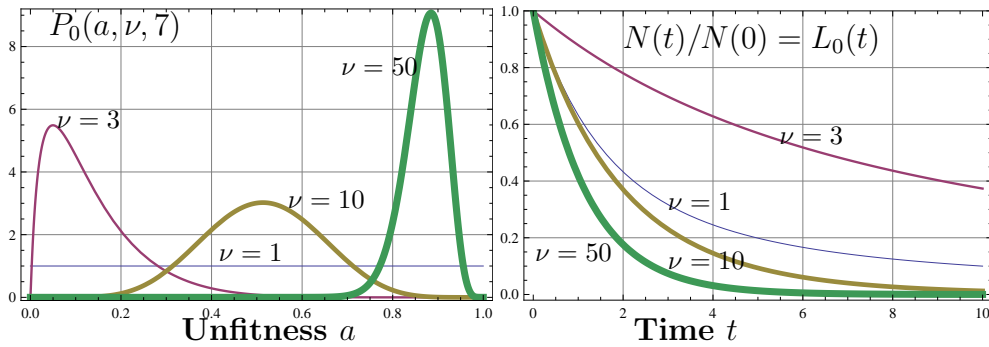


Figure 13: Initial log-gamma densities $P_0(a, \nu, 7)$ are shown in the left panel for the uniform density $\tau = \nu = 1$, and also for $\tau = 7$ with $\nu = 3, 10, 50$ in graphs with increasing thickness. The right panel shows the corresponding fractional decline with time of the population $N(t)/N(0)$ from (25) for these initial densities.

3 Discussion

The initial states that we have investigated are

Symmetric uniform ($\tau = 1$ with variance $\sigma_0^2(a) = \frac{1}{12} \approx 0.083$)

Symmetric, higher variance ($\tau = 0.289$ with $\sigma_0^2(a) = 0.164$)

Symmetric, lower variance ($\tau = 2.24$ with $\sigma_0^2(a) = 0.045$)

Asymmetric log-exponential ($\tau = 1, \nu > 0$)

Asymmetric unimodal ($1 < \tau \in \mathbb{N}, \nu > 0$).

The new results show how evolution of an inhomogeneous rate process reacts to perturbations of the uniform density using the smooth two parameter (ν, τ) family of log-gamma probability density functions [2] to represent unfitness measured by parameter $a \in (0, 1]$. We have illustrated the evolution from three initial states having central mean $E_0(a) = \frac{1}{2}$ but different variances (cf. Figures 2 and 4). These first three cases have starting log-gamma densities in the left hand panes of Figures 6,7,8 show widely differing shapes; the higher variance arises from peaks at the extremities of unfitness and the lower variance case has a symmetric bell-like curve about the mean. The evolution of features from the initial uniform density from Karev [9] is—cf. equation (6) above,

$$P_0(a) = 1 \Rightarrow P_t(a) = \frac{te^{-at}}{1 - e^{-t}}, \quad E_t(a) = \frac{1}{t} + \frac{1}{1 - e^{-t}}, \quad \frac{N(t)}{N(0)} = \frac{1 - e^{-t}}{t}. \quad (24)$$

To this we add our analytic results (18), (19) for the initial log-gamma cases with $\tau = 1$, which agrees with (24) at $\nu = 1$ and the generalized hypergeometric solutions (22), (23) along with corresponding versions for integer $\tau > 3$. Our series approximation agrees with the analytic results for the early development. Figure 6 shows that in each of our cases the evolution of mean unfitness $E_t(a)$ does indeed follow the expected equation (5), which expresses Fisher's law of natural selection [8]—see also Cziko [4] for discussion of selection theory in a wide range of biological processes from synapses to cells and organisms. Also, in Figure 8 the very slow early evolution of its variance $\sigma_t^2(a)$ reflects the corresponding lack of curvature in the mean. In fact, though the decline of the mean fitness does change with the variation in starting conditions, the net population evolution $N(t)$ in Figure 7 seems surprisingly stable under quite considerable alterations in the shape and variance of the initial density of inhomogeneity in fitness. The rate process evidently has a strong smoothing effect when we begin from a central mean, even though the variances differ widely.

We can however obtain considerable changes in the evolution if we depart from an initial central mean. We illustrate this using our analytic solutions to allow evolution over longer periods. For example, the initial log-gamma density for $\tau = 1$, equations (18),(19),(20), gives the examples shown in Figure 9. Here we see the effect of ν which acts as a location parameter for the initial density. Figure 10 shows the evolution of the initial case with $\nu = 10$ into the probability densities $P_t(a, 10, 1)$ from (18). Figure 11 shows the evolving approximate information distance measured from the initial log-gamma distributions with $\tau = 1, \nu = 0.5$ (left) and $\tau = 1, \nu = 10$ (right), near $t = 2$.

Also, the initial unimodal log-gamma probability densities for $\tau = 7, P_t(a, \nu, 7)$, give this generalized hypergeometric solution

$$\frac{N(t)}{N(0)} = L_0(t) = {}_7F_7(\nu, \nu, \nu, \nu, \nu, \nu, \nu; \nu + 1, \nu + 1, \nu + 1, \nu + 1, \nu + 1, \nu + 1, \nu + 1; -t) \quad (25)$$

and again ν serves as a location parameter. Figure 13 illustrates cases for $\nu = 3, 10, 50$ as well as that for an initial uniform density.

Our approach has considered the determinate rate process for the evolution of individual types with an inhomogeneous fitness distribution; Baake and Georgii [3] consider the multi-dimensional systems of many types. Their mutation and differential reproduction processes yield n-dimensional systems of differential equations which they analyse using variational methods.

References

- [1] S-I. Amari and H. Nagaoka. **Methods of Information Geometry**, American Mathematical Society, Oxford University Press, Oxford, 2000.
- [2] Khadiga Arwini and C.T.J. Dodson. **Information Geometry Near Randomness and Near Independence**. Lecture Notes in Mathematics, Springer-Verlag, New York, Berlin 2008.
- [3] E. Baake and H.-O. Georgii. Mutation, selection, and ancestry in branching models: a variational approach. Preprint 2006, <http://arxiv.org/abs/q-bio/0611018>
- [4] G. Cziko. **Without Miracles: Universal Selection Theory and the Second Darwinian Revolution**, MIT Press, Cambridge, Massachusetts, 1995.
- [5] C.T.J. Dodson. On the entropy flows to disorder. In Proc CHAOS 2009, C. H. Skiadas and I. Dimotikalis, Editors, **Chaotic Systems: Theory and Applications** World Scientific, Singapore 2010. Cf. also <http://arxiv.org/abs/0811.4318>
- [6] C.T.J. Dodson. Information geometry and entropy in a stochastic epidemic rate process. Preprint, 2009. <http://arxiv.org/abs/0903.2997>
- [7] W. Feller. **An Introduction to Probability Theory and its Applications**, Volume II 2nd Edition, Wiley, New York, 1971.
- [8] R.A. Fisher. **The Genetical Theory of Natural Selection**, 2nd Edition, Dover, New York, 1958. Cf. also S.A. Frank, *J. Evol. Biol.* 22 (2009) 231-244.
- [9] G.P. Karev. Inhomogeneous models of tree stand self-thinning. *Ecological Modelling* 160 (2003) 23-37.
- [10] G.P. Karev. Replicator equations and the principle of minimal production of information. Preprint, 2009. <http://arxiv.org/abs/0901.2378>