# CrossMark

## RESEARCH ARTICLE Invariant death [version 1; referees: 2 approved]

## Steven A. Frank

Department of Ecology and Evolutionary Biology, University of California, Irvine, CA, 92697-2525, USA

V1 First published: 25 Aug 2016, 5:2076 (doi: 10.12688/f1000research.9456.1) Latest published: 25 Aug 2016, 5:2076 (doi: 10.12688/f1000research.9456.1)

#### Abstract

In nematodes, environmental or physiological perturbations alter death's scaling of time. In human cancer, genetic perturbations alter death's curvature of time. Those changes in scale and curvature follow the constraining contours of death's invariant geometry. I show that the constraints arise from a fundamental extension to the theories of randomness, invariance and scale. A generalized Gompertz law follows. The constraints imposed by the invariant Gompertz geometry explain the tendency of perturbations to stretch or bend death's scaling of time. Variability in death rate arises from a combination of constraining universal laws and particular biological processes.

Open Peer Review				
Referee Status: 🗹🗹				
		Invited Referees		
		1	2	
<b>V</b> pt 25	<b>ersion 1</b> <sup>Jblished</sup> 5 Aug 2016	report	report	
1	1 Sean Nee, The Pennsylvania State University USA			
2	<b>Ophélie Ronce</b> , University of Montpellier France			

### **Discuss this article**

Comments (0)

Corresponding author: Steven A. Frank (safrank@uci.edu)

How to cite this article: Frank SA. Invariant death [version 1; referees: 2 approved] *F1000Research* 2016, 5:2076 (doi: 10.12688/f1000research.9456.1)

**Copyright:** © 2016 Frank SA. This is an open access article distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Grant information: National Science Foundation grant DEB-1251035 supports my research.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: No competing interests were disclosed.

First published: 25 Aug 2016, 5:2076 (doi: 10.12688/f1000research.9456.1)

#### Introduction

The coil of a snail's shell expresses the duality of constraint and process. The logarithmic spiral of growth constrains overall form. Particular snails modulate the process of shell deposition, varying the parameters of the logarithmic spiral. To interpret the variety of snail shells, one must recognize the interplay between broad geometric constraint and the special modulating processes of individual types<sup>1</sup>.

The pattern of death in populations follows the same duality of invariant geometric constraint and modulating process. The invariant geometry of death's curve arises from the intrinsic order of large samples<sup>2,3</sup>. A large sample erases underlying randomness, preserving only invariant aggregate values<sup>4</sup>.

I extend the large-sample concept to clarify the invariant geometry of death. I then illustrate the role of particular biological processes in modulating death's curve: the stretch of death's time in nematode response to physiological perturbation<sup>5</sup> and the curvature of cancer's time in response to genetic perturbation<sup>6,7</sup>. The consequences of particular biological perturbations can only be understood within the geometry that constrains change to follow invariant contours.

To restate the puzzle: How can we relate small-scale molecular and physiological process to population consequence? The problem remains unsolved. Finch and Crimmins<sup>8</sup> emphasized: "A key question is how to connect ... [linear] aging processes to the exponential rates of accelerating mortality that set life spans. ... Although we can readily assess molecular aging, such biomarkers of aging are rarely robust as predictors of individual morbidity and mortality risk in populations."

#### **Randomness and invariance**

I begin with the relation between small-scale randomness and large-scale order. The classical theory derives from the principles of statistical mechanics<sup>2</sup>, later developed through aspects of entropy and information<sup>4,9</sup>. Here, I briefly summarize my own extension of classical results based on geometric principles of invariant measurement and scale<sup>10-13</sup>. I then show how the abstract geometry constrains the relation between biological process and the pattern of death's curve.

To understand the probability of dying at a particular age, we begin with the geometry of probability patterns<sup>13</sup>. For an underlying quantity, z, the probability of observing a value near to z is the rectangular area with height  $q_z$ , width  $d\psi_z$ , and area  $qd\psi$ . A probability pattern is a curve with coordinates  $(\psi_z, q_z)$  defined parametrically with respect to z. For the curve of death, the input, z, may be age or time.

Two invariances constrain the geometry of probability curves. First, total probability is invariantly one. Invariant total probability implies that the height of the probability curve has a natural exponential expression<sup>13</sup>

$$q_z = k_a e^{-\lambda(a+T_z)} = k e^{-\lambda T_z},\tag{1}$$

in which  $k = k_a e^{\lambda a}$  remains constant for any *a* to satisfy the requirement that total probability is invariant. The exponential form for the height of the probability curve,  $q_z$ , implies that the probability curve remains invariant to a shift of the fundamental metric,  $T_z$  (see Frank<sup>13</sup>).

In general, we seek metrics for which it does not matter where we set our zero reference point. In geometry, a circle shifted in space retains its invariant form. Similarly, proper geometric scaling for probability patterns is shift invariant. In terms of death, any transformation of time, z, into a fundamental time metric for probability pattern,  $T_z$ , must measure time such that a shift  $a + T_z$  does not alter death's curve. That shift-invariant requirement leads to the exponential expression<sup>13</sup> in equation 1.

The second key invariance is that a uniform stretching or shrinking of the fundamental metric does not alter probability pattern<sup>13</sup>

$$q_z = k e^{-\lambda_b b T_z} = k e^{-\lambda T_z},\tag{2}$$

in which  $\lambda = \lambda_b b$  remains constant for any *b*, causing the probability pattern to be invariant to stretch of  $T_z$ . Stretch invariance is equivalent to invariance of  $\lambda \langle T \rangle_{\psi}$ , the value of  $\lambda$  multiplied by the average value of *T* when probability, *q* d $\psi$ , is measured on the scale,  $\psi$  (see Frank<sup>13</sup>).

To summarize, probability curves remain invariant to shift and stretch of the fundamental metric,  $T_{e}$ , such that

$$T \mapsto a + bT \sim T,\tag{3}$$

in which '~' means invariant with respect to shift and stretch. In geometry, invariance with respect to shift and stretch is affine invariance.

Affine invariance leads to probability pattern described by a sequence of rectangular areas

$$q \,\mathrm{d}\psi = k e^{-\lambda T} \mathrm{d}\psi,$$

in which k and  $\lambda$  are constants that adjust to satisfy, respectively, invariant total probability and invariant average value,  $\lambda \langle T \rangle_{\psi}$ . Many different approaches and interpretations all arrive at this same basic form.

#### Consequences of affine invariance

Here, by emphasizing the fundamental invariances, we can take the next key step in understanding the geometry of probability patterns and the curves of death. In particular, each successive application of the affine transformation (equation 3) to T leaves the probability pattern unchanged, defining an invariant group of metrics<sup>11</sup>

$$T = \frac{1}{\beta} \left( e^{\beta_{W}} - 1 \right) \sim e^{\beta_{W}}, \tag{4}$$

with  $\beta \to 0$  implying  $T \to w$ . Here, w(z) is a scale for the underlying values, *z*, such that a shift in that scale,  $w \mapsto \alpha + w$ , only changes *T* by a constant multiple, and therefore does not change the probability pattern.

To find the proper metric, T, for a particular probability pattern, we only need to find the proper base scale w for which the probability pattern is shift invariant. If, for example, z is time or age, then we only need to discover the scaling, w(z), for which

$$q \,\mathrm{d}\psi = k e^{-\lambda e^{\beta w}} \mathrm{d}\psi \tag{5}$$

is invariant to a shift in *w*, when allowing adjustment of  $\lambda$ . When  $\beta \rightarrow 0$ , then  $q \rightarrow e^{-\lambda w}$ .

Equation 5 expresses the abstract form of common probability patterns<sup>11</sup>. The abstraction does not specify the two key scaling relations  $\psi(z)$  and w(z) that define the coordinates of the parametric probability curve  $(\psi, q)$  with respect to *z*. However, the invariances that define the geometry do impose strong constraints, leading to a limited set of forms for almost all of the commonly observed probability patterns<sup>11-13</sup>.

We have two scaling relations  $\psi$  and w, but only a single parametric probability curve ( $\psi$ , q) with associated probability  $qd\psi$  in each increment. Thus, many different scales can express the same probability pattern. For each application, there is often a natural scale that has a simple, understandable form for its scaling relations.

#### The invariant ticking of death's clock

A natural scale corresponds to an additional invariance with a simple interpretation. That additional invariance sets the underlying metric for the pair of scaling relations. For death, we can set the probability of dying to be invariant in each increment of the scale,  $d\psi$ , so that  $\psi$  represents the uniform metric of mortality—the invariant ticking of death's clock. This uniform metric extends the theory of extreme values and time to failure<sup>14–17</sup> to a more abstract and general understanding of the invariances that shape all of the common probability patterns<sup>11–13</sup>.

Invariant probability in each increment can be written as  $q d\psi = -dq$ and thus  $\psi = -\log q$ , in which dq is a constant incremental fraction of the total probability. I use a minus sign as a convention to express the total probability as declining with an increase in  $\psi$ .

With regard to dying, we may think of the total probability of being alive as declining by a constant increment of death, -dq, in each increment  $d\psi$ . In classic epidemiology, this definition of q would be expressed as  $q(z) \equiv S(z)$ , in which S(z) is the probability of survival to time z. However, it is important to consider the classic definition as a special case of the deeper abstract geometry, which leads to a more general understanding of the constraints that shape death's curve.

#### **Universal Gompertz geometry**

Given the exponential form for  $q_z$  in equation 2, a constant probability  $q d\psi$  in each increment requires  $d\psi = dT$ . Using the general form of *T* in equation 4, we have  $dT = e^{\beta w} dw = T' dw$ , in which T' > 0 is the derivative of *T* with respect to *w*. With  $\hat{q}dw = q dT$  for the constant probability in each increment, we have

$$\hat{q} \, \mathrm{d}w = k e^{\log T' - \lambda T} \mathrm{d}w = k e^{\beta w - \lambda e^{\beta w}} \mathrm{d}w. \tag{6}$$

This probability pattern is expressed on the scale *w*, in which *w* defines the natural shift-invariant metric. In other words, for some underlying observable value *z*, such as time or age, w(z) transforms *z* to a scale, *w*, that expresses an invariant total probability  $\hat{q} dw$  in each increment, and for which shifts in the scale  $w \mapsto \alpha + w$ , do not change the probability pattern.

The probability pattern in equation 6 has the familiar Gompertz form. I derived that form solely from a few simple geometric invariances. The simple invariances elevate the generalized Gompertzian form to a universal geometric principle for probability patterns<sup>11,13</sup>. By contrast, the Gompertzian pattern is usually derived from descriptive statistics or from particular assumptions about processes of failure or growth.

#### Pattern on the observed scale

We may express the probability pattern on the scale of the underlying observable value, z. For that scale, dw = w'dz, in which w' > 0 is the derivative of w with respect z. The abstract Gompertzian geometry in w becomes the explicit form with respect to the directly measured value z as

$$\tilde{q}\,\mathrm{d}z = ke^{\log w'T' - \lambda T}\mathrm{d}z = kw'e^{\beta w - \lambda e^{\beta w}}\mathrm{d}z.\tag{7}$$

#### The hazard of death

Only living individuals can die. Thus, the hazard of death is the probability of dying in an incremental metric of time divided by the probability of being alive. The incremental metric scale,  $\psi_z$ , transforms the observed value, *z*, which may be time or age, into the abstract incremental scale,  $d\psi$ . The abstract expression for the hazard of dying in an increment  $d\psi$  is

$$h(\psi)d\psi = \frac{q\,d\psi}{1 - \int q\,d\psi}.$$
(8)

In each increment, the probability of dying is  $q d\psi$ . The integral in the bottom is the sum of the probabilities of dying over the period from a starting point until the current period, in which the time metric is described by  $\psi(z)$ .

Three different metrics transform the observable time input, or other measurable input, z, into the scale of analysis: T, w, and z itself. Those three metrics yield three equivalent forms for the hazard, each emphasizing different aspects of the underlying geometric invariances

$$h(T)dT = \lambda dT \propto dT \tag{9}$$

$$\hat{h}(w)dw = \lambda T'dw \propto e^{\beta w}dw \tag{10}$$

$$\tilde{h}(z)dz = \lambda w'T'dz \propto w'e^{\beta w}dz$$
(11)

in which ' $\propto$ ' denotes proportionality. The top form expresses the most general and abstract invariance of death. By transforming time into a general metric,  $z \mapsto T$ , the hazard is invariantly  $\lambda$  in each increment of the metric, d*T*. The metric *T* defines the scale on

which the probability pattern is invariant to the affine transformation  $T \mapsto a + bT$ .

We know the scale of death's curve when we can transform our underlying observation, z, such as age, to the affine-invariant scale, T. Often, adding a constant to age or multiplying age by a constant,  $z \mapsto a + bz$ , changes the pattern of death's curve, so using age itself as the metric is usually not sufficient. We must find some transformation of age.

The middle expression in equation 10 describes the generalized Gompertzian geometry in the most direct way. In this case, when we transform  $z \mapsto w$ , changing an observation such as age, z, to the metric, w(z), we only require that death's curve be invariant to a shift,  $w \mapsto a + w$ .

#### The force of death and the curvature of time

We are partitioning the scaling of death's curve into two steps,  $z \mapsto w \mapsto T$ . Once we have the shift-invariant scaling of time, w, then  $T = e^{\beta w}$  changes w into the ultimate affine-invariant scaling, T. To make that last change, we need to know  $\beta$ , which is

$$\beta = \frac{T''}{T'} = \frac{d\log \hat{h}}{dw},\tag{12}$$

in which each prime denotes the derivative with respect to w. This value of  $\beta$  defines the curvature for the geometry of death and time. Once we have the shift-invariant scaling for time,  $z \mapsto w$ , we can consider death's invariant curvature in the transformation  $w \mapsto T$ .

The expression T'' is the acceleration, or absolute curvature, of T. The expression T' is the rate or velocity at which T is changing. Thus, T''/T' can be thought of as the acceleration relative to the velocity.

Acceleration, curvature and force are ultimately equivalent. In terms of death, for a given velocity or rate at a particular age, T', the value of  $\beta$  is the relative force that bends death's curve. The bending of death's curve may also be described as

$$\mathscr{A} = \frac{\mathrm{d}\log\hat{h}}{\mathrm{d}w} = \frac{\hat{h}'}{\hat{h}},\tag{13}$$

which is the change in the hazard of death relative to the current hazard. The hazard,  $\hat{h}$ , is the rate, or velocity, of death on the scale *w*. Thus, the change in relative velocity,  $\mathscr{A}$ , describes the acceleration of mortality in terms of the relative bending of death's rate.

The invariant geometry of death's curve in equation 12 may be expressed as a balance,  $\beta - \mathcal{A} = 0$ , between force and acceleration. That balance is roughly analogous to Newton's second law of motion,  $F = m\mathcal{A}$ , relating force to acceleration.

#### Inference

The invariant geometry does not tell us the form for the shiftinvariant scaling of death's time, w, or the value of the invariant force,  $\beta$ , that bends death's curve. However, the invariances strongly constrain the likely form of death's curve and the meaningful metrics of death's time. Importantly, these expressions allow us to transform data about rates or motions into expressions that emphasize force and causal interpretations<sup>18,19</sup>. In biology, we rarely can predict trajectories. Instead, we focus on interpreting the changes in observed trajectories with respect to hypothesized forces<sup>7,20</sup>.

The abstract geometry is correct unto itself. In application, the geometry provides a tool that we may use for particular problems. A tool is neither right nor wrong. Instead, a tool is helpful or not according to its aid in providing insight. Below I discuss some examples. A few comments prepare for the discussion.

If we knew the correct scaling for age, w(z), then within that frame of reference, the force,  $\beta$ , and acceleration,  $\mathscr{A}$ , of mortality would be constant with respect to w. Thus, the frame of reference, w, provides valuable insight. However, w may turn out to be a weirdly nonlinear scaling of measured time, z, in which the form of w is difficult to determine directly. In practice, we can derive w from equation 11 by relating the hazard,  $\tilde{h}(z)$ , to w by

$$\beta w = \log \int h dz, \tag{14}$$

or  $T_z \sim \int \tilde{h} dz$ , the affine similarity of  $T_z$  to the accumulated hazard on the z scale.

I now discuss the time scaling of mortality in nematodes and cancer. I consider these applications only to illustrate general aspects of mortality's temporal geometry. See Stroustrup *et al.*<sup>5</sup> for details about nematodes and Frank<sup>7</sup> for details about cancer.

#### Nematode mortality and the stretch of time

Stroustrup *et al.*<sup>5</sup> conclude from their study of nematode mortality:

[W]e observe that interventions as diverse as changes in diet, temperature, exposure to oxidative stress, and disruption of [various] genes ... all alter lifespan distributions by an apparent stretching or shrinking of time. To produce such temporal scaling, each intervention must alter to the same extent throughout adult life all physiological determinants of the risk of death.

I begin with the apparent *stretching or shrinking of time*. I will arrive at the same description of the nematode mortality pattern as given by Stroustrup *et al.*<sup>5</sup>, but framed within my more general understanding of mortality's invariant geometry. From that broader perspective, the observed stretching or shrinking of time in the nematode study can be seen as a special case of the various temporal deformations that arise with respect to mortality's invariant scale.

The perspective of my general framing calls into question the second conclusion that *each intervention must alter to the same extent throughout adult life all physiological determinants of the risk of death.* I present a simple counterexample consistent with the observed patterns. My counterexample may not be the correct description of process in nematode mortality. The counterexample does, however, emphasize key aspects of the logic by which we must evaluate the relations between pattern and process in mortality.

My framework analyzes the sequence of transformations  $z \mapsto w \mapsto T$ . The initial input, *z*, typically represents what we measure, such as a standard description of time or age. We then seek a transformation, w(z), such that the parametric curve,  $(w, \hat{q})$ , for observed or assumed probability pattern is shift-invariant with respect to *w* (equation 6). In other words, the shift  $w \mapsto \alpha + w$  does not alter the probability curve. When we find the shift-invariant scale for *w*, we have an expression for the probability pattern in terms of the Gompertzian geometry of equation 6.

A probability pattern that remains the same except for a constant stretching or shrinking of time corresponds to  $w(z) = \log z$ , because a constant stretch or shrink of time by  $a = e^{\alpha}$  yields  $w(az) = \alpha + w(z)$ . If we express the associated parametric probability curve as the relation between time and probability,  $(z, \tilde{q})$ , as in equation 7 with  $w = \log z$ , we obtain a curve that is invariant to a constant stretching or shrinking of the temporal scale, *z*, as

$$\tilde{q} \propto z^{\beta - 1} e^{-\lambda z^{\beta}},$$

in which the parameter  $\lambda$  and the constant of proportionality both adjust to cancel any stretch or shrink of time by a > 0 (see Frank<sup>13</sup>). This curve is the Fréchet probability distribution, corresponding to the power law hazard in equation 11 as

$$\tilde{h} \propto z^{\beta-1}$$

Stroustrup *et al.*<sup>5</sup> concluded that the Fréchet distribution is the best overall match to their nematode studies. However, they invoked the Gompertz-Fréchet family of distributions by appeal to traditional epidemiology and by appeal to the general form of extreme value distributions for failure times. By contrast, I derived those distributions simply as the inevitable consequence of basic assumptions about the invariant geometry of meaningful scales<sup>13</sup>.

#### The deformation of death's time

Stroustrup *et al.*<sup>5</sup> discussed the stretching or shrinking of death's time by a single constant value. My framework generalizes the deformation of time in relation to death. We begin with *T*, the universal frame of reference for the scaling of death's time. On the temporal scale, *T*, the hazard of death, h(T), remains constant at all times (equation 9). Thus, *T* represents the invariant ticking of mortality's clock.

Given that universal frame of reference for time, we may then consider other temporal scales in terms of the way in which they deform the invariant frame of reference. In this case, we work inversely, by starting with *T* in equation 4, and then inferring the deformations with respect to the underlying scale of description, *z*. We can then think of the shape of the curve  $(T_z, z)$  as describing how measured time, *z*, is deformed in relation to the universal invariant scale of mortality's time,  $T_z$ .

Ideally, we first infer the shift-invariant scale, w(z), and then use w in equation 4 to determine the relation between T and z. In the nematode case,  $w(z) = \log z$  achieved shift invariance. Thus  $T_z \sim e^{\beta w} = z^{\beta}$ . The power law curve  $(z^{\beta}, z)$ , with curvature determined by  $\beta$ , describes the deformation of time. The different experimental treatments did not significantly alter the curvature associated with  $\beta$ .

We can relate increments of the measured input, dz, to increments of mortality's universal measure, dT, by starting with equation 11 as  $\tilde{h} \propto dT/dz$ , and then writing

$$\mathrm{d}z \propto \mathrm{d}T/h. \tag{15}$$

For a case such as the nematodes in which  $T_z \sim z^{\beta}$ , the measured temporal increments, dz, scale in relation to the universal temporal frame as  $dz \propto dT/z^{\beta-1}$ . For  $\beta > 1$ , measured temporal increments, dz, shrink as time passes relative to the constant ticking of mortality's clock at dT. When we think of dT as mortality's constant temporal frame of reference, then the deformation of measured time is

$$dz \propto \frac{1}{z^{\beta-1}}.$$

The shrinking of measured time corresponds to the increase in the rate of measured mortality, in other words, the same amount of mortality, dT, is squeezed into smaller temporal increments, dz, increasing the density of mortality per measured unit.

In other cases, the relation of measured inputs, z, to mortality's universal scale,  $T_z$ , will have different functional forms. Those different functional forms may correspond to non-uniform stretching and shrinking of the observed temporal scale at different magnitudes of z relative to the universal frame of reference for mortality on the scale  $T_z$ . If possible, we first infer the shift-invariant scale, w(z), for example by equation 14, and then use w to determine the relation between T and z, as in the nematode example. However, in practice, it may be easier to go directly from the invariant clock, T, to the deformed time scale, z, by using the relation  $dz \propto dT/\tilde{h}$ . The following critique of the conclusions by Stroustrup *et al.*<sup>5</sup> about nematode mortality provides an example.

#### Invariant pattern and underlying process

Stroustrup *et al.*<sup>5</sup> claimed that all physiological determinants of the risk of death change in the same way with each experimental intervention. I present simple counterexamples. Although my counterexamples may not describe the true underlying process, they do highlight two important points. First, commonly observed patterns often express invariances that are consistent with many alternative underlying processes<sup>21,22</sup>. Second, consideration of the alternative processes with the same observable invariances leads to testable predictions about the underlying causal processes.

In these examples, suppose that death follows a multistage process, as is often discussed in cancer progression<sup>23</sup>. Following Frank<sup>7</sup>, p. 98, we may write the dynamics of progression toward mortality as a sequence of transitions

$$\begin{aligned} \dot{x}_0(z) &= -u_0 x_0(z) \\ \dot{x}_i(z) &= u_{i-1} x_{i-1}(z) - u_i x_i(z) \qquad i = 2, \dots, n-1 \\ \dot{x}_n(z) &= u_{n-1} x_{n-1}(z), \end{aligned}$$

where  $x_i(z)$  is the fraction of the initial population born at time z = 0 that is in stage *i* at measured time, *z*. Assume that when the cohort is born, at z = 0, all individuals are in stage 0, that is,  $x_0(0) = 1$ , and the fraction of individuals in other stages is zero.

As time passes, some individuals move into later stages of progression toward death. The rate of transition from stage *i* to stage *i* + 1 is  $u_i$ . The  $\dot{x}$ 's are the derivatives of *x* with respect to *z*. Death occurs when individuals transition into stage *n*. A fraction  $x_n(z)$  of individuals has died at time *z*, and the rate of death at time *z* is  $\dot{x}_n(z) \equiv \tilde{q}$ , in which  $\tilde{q}$  has the probability interpretation of equation 7.

If the transition rates are constant and equal,  $u_i = u$  for all *i*, then we can obtain an explicit solution for the multistage model<sup>24</sup>. This solution provides a special case that helps to interpret more complex assumptions. The solution is  $x_i(z) = e^{-uz}(uz)^{i/i!}$  for i = 0, ..., n - 1, with the initial condition that  $x_0(0) = 1$  and  $x_i(0) = 0$  for i > 0. Note that the  $x_i(z)$  follow the Poisson distribution for the probability of observing *i* events when the expected number of events is uz.

In the multistage model above, the derivative of  $x_n(z)$  is given by  $\dot{x_n}(z) = ux_{n-1}(z)$ . From the solution for  $x_{n-1}(z)$ , we have

$$\tilde{q} = \dot{x}_n(z) = ue^{-uz} (uz)^{n-1} / n - 1!,$$

which is a gamma probability distribution. One can think of the gamma distribution as the waiting time for the *n*th event, in which each event occurs at constant rate, *u*. However, many other processes lead to the same gamma distribution.

Age-specific incidence is the hazard<sup>7</sup>

$$\tilde{h}(z) = \frac{\dot{x}_n(z)}{1 - x_n(z)} = \frac{u(uz)^{n-1}/n - 1!}{\sum_{i=0}^{n-1} (uz)^i / i!}.$$
(16)

We can express the scaling of measured time, dz, relative to the constant ticking of mortality's time, dT, from equation 15, by taking dT as constant and thus

$$dz \propto \frac{1}{(uz)^{n-1}} \left( 1 + uz + \frac{(uz)^2}{2} + \dots + \frac{(uz)^{n-1}}{n-1!} \right).$$
(17)

#### Simultaneity and temporal deformation

When measured time, z, is small, during the initial period of the process, the deformation of time is approximately the same as the Fréchet pattern,  $dz \propto 1/z^{n-1}$ . This deformation in the gamma process describes the force of simultaneity. Early in the process, all components that protect against mortality remain in the initial working state. Thus, mortality requires the nearly simultaneous failure of *n* independent events, which creates a force that deforms the constant ticking of mortality's clock by rescaling the measured increments, dz.

As measured time increases, the increments dz shrink, compressing the same amount of mortality, dT, into smaller measured temporal increments. As z becomes larger, the increments dz in equation 17 shrink less, because of the reduced force of simultaneity that deforms mortality's constant clock. With larger z, the higher-power terms of the sum increasingly dominate, until the largest power term dominates and dz then ticks at a constant rate, with  $dz \propto dT$ .

The changing deformation of dz and the associated force of mortality can be thought of roughly as follows<sup>7</sup>. Early in the gamma process, mortality requires the nearly simultaneous failure of nindependent events, creating a force of simultaneity such that  $dz \propto 1/z^{n-1}$ . As time passes, many individuals suffer failure of some of the n processes, leaving in aggregate the equivalent of n-1 remaining protective components, and a force of simultaneity such that, approximately,  $dz \propto 1/z^{n-2}$ . As more time passes, additional components fail, and the remaining force of simultaneity diminishes, until eventually only one protective component remains for those still alive, at which point dz then ticks at a constant rate, so that  $dz \propto dT$ .

We may also express the scaling of time on the shift-invariant Gompertzian scale, w, in which  $\beta$  is a relative measure of the acceleration of mortality (equation 12), by using the general expression in equation 14 and the specific form of  $\tilde{h}$  in equation 16 to yield

$$\beta w = \log \int \tilde{h} dz = \log \log \Gamma(n, uz)^{-1},$$

in which  $\Gamma$  is the incomplete gamma function.

#### Alternative models of nematode mortality

With this understanding of the gamma process, we can consider alternative interpretations of the nematode data<sup>5</sup>. I present these alternatives to illustrate the logic of mortality's temporal scaling and the potential relation to underlying process. The data do not provide information about whether or not these alternative interpretations are the correct description of nematode mortality. The point here is that these alternatives, or some other structurally similar alternative, might be correct, and therefore the strong conclusions of the original article may be false.

To repeat the key conclusion from Stroustrup *et al.*<sup>5</sup>, *each intervention must alter to the same extent throughout adult life all physiological determinants of the risk of death.* 

That conclusion is true for the simple gamma process, as summarized by equation 17. In that equation, the value of *u* represents the rate at which each of the *n* processes fails and contributes to overall mortality. If we substitute  $uz \mapsto \xi$ , then the scaling of measured time, expressed as  $dz \propto d\xi$ , changes only by a constant of proportionality as the rate, *u*, changes.

I now consider two variations on the underlying gamma process for mortality. Each of these variations leads to a constant rescaling of time, dz. However, that constant rescaling arises from underlying processes of mortality that change in different ways in response to perturbations. These examples show that the constant rescaling of time does not imply that an intervention alters to the same extent throughout adult life all physiological determinants of the risk of death. The first example considers two distinct sets of underlying processes that influence mortality, each set composed of *n* processes. Mortality occurs only after the failure of all 2*n* processes. Before experimental perturbation, one set of *n* processes has a relatively slow failure rate per process of *u*. The other set of *n* processes has a relatively fast failure rate of  $u' \gg u$ .

In this case, the fast processes will all tend to fail early in life, almost always before all of the slow processes fail. Thus, the fast processes have little influence on mortality. The mortality rate will closely follow the gamma process with *n* steps, each step at rate *u*, as analyzed above<sup>7</sup>.

Now suppose that an intervention influences all of the slow steps but none of the fast steps. The intervention changes the previously slow rate processes into fast processes,  $u \mapsto u'' \gg u'$ . After intervention, the mortality rate will closely follow the gamma process with *n* steps, with each step at rate u'. The mortality pattern remains unchanged, except for a constant rescaling of time.

However, the underlying physiological processes that determine mortality have changed completely. Previously unimportant rate processes with respect to mortality now completely dominate, and previously important rate processes no longer influence mortality.

The second example considers a set of *n* underlying processes that influence mortality. Each process has a different failure rate of  $u_i$  for i = 1, ..., n, with  $u_i < u_{i+1}$ . As before, mortality occurs only after the failure of all *n* processes. Frank<sup>7</sup> presented numerical studies for this heterogeneous rate process model. Typically, the faster rate processes fail early in life and have relatively little influence. The slower processes dominate the overall temporal pattern.

With n equal rate processes, the curvature declines with time, as in equation 17. With a heterogeneous set of rate processes, the curvature tends to decline more quickly as the fastest processes fail earlier, typically leaving a progressively smaller set of remaining protective mechanisms as time passes, reducing the force of simultaneity.

Now suppose that the heterogeneous set of *n* processes has a simple hierarchy of rates, such that  $u_{i+1} = \gamma u_i$ , in which  $\gamma > 1$  is the factor by which each rate increases relative to its slower neighbor. If the only effect of an experimental intervention is abrogation of the slowest process,  $u_1$ , then the hierarchy is effectively altered only by multiplying each rate in the set by  $\gamma$ , because the fastest processes typically have almost no influence on pattern.

Once again, the overall mortality pattern will change only by a constant rescaling of time, even though the underlying physiological processes have changed significantly with respect to their influence on mortality. In this case, the most important process that limited mortality before intervention was in effect knocked out after intervention, whereas all other processes did not change.

#### Different processes lead to same invariances

The actual biology of nematodes will, of course, not follow exactly either of these two example cases. The examples do show, however, that a constant rescaling of measured time for mortality can arise by heterogeneous changes in the underlying physiological determinants of the risk of death.

How should we interpret the match between variants of the multistage gamma process model and the observed scaling of nematode mortality? The correct view is that the invariances expressed by the gamma model are approximately the same as the invariances that arise by the true physiological processes. Those invariances dominate the shape of the observed patterns. The examples of the gamma models are helpful, because they show the sorts of underlying processes that generate the required invariances.

Ultimately, the theoretical challenge is to understand the full set of underlying processes that lead to the same invariances and thus the same observed pattern<sup>22</sup>. The empirical challenge is, of course, to figure out which particular processes occur in each particular case. Success in the empirical challenge will likely depend on further progress on the theoretical challenge, because the theoretical frame strongly influences how one goes about solving the empirical problem.

#### Cancer incidence and the curvature of time

I now turn to genetic knockouts in cancer that change the curvature of time. Cancer incidence often follows a pattern roughly consistent with a multistage gamma process<sup>7,23</sup>. Again, that match does not mean that the underlying physiological processes truly follow the assumptions of the gamma model. Instead, the correct view is that the invariances expressed by the gamma model are approximately the same as the invariances that arise by the true physiological processes.

Consider the simplest gamma process, in which cancer arises only after n protective mechanisms fail. Each mechanism fails at the same rate, u. I gave the explicit solution for that process earlier. In interpreting that solution for cancer, it is important to note an essential distinction between mortality and cancer incidence.

Everyone dies but only a small fraction of individuals develop a particular form of cancer. Thus, we must analyze mortality by running the measured time, or age, z, out to a large enough value so that the cumulative probability of dying approaches one. In the gamma models above, that means letting uz increase significantly above one. By contrast, if only a small fraction of the population develops cancer before dying of other causes, then we must run z only up to a time at which the cumulative probability of cancer remains small. That limit on total incidence typically means capping uz below one.

With a small maximum value of uz, the age-specific hazard simplifies approximately to  $\tilde{h} \propto z^{n-1}$ , and the scaling of measured

time simplifies to  $dz \propto 1/z^{n-1}$ . Those scalings match the Fréchet model when we equate the curvature of time,  $\beta$ , with the number of steps, *n*, and we interpret force and curvature with respect to the shift-invariant scaling of time,  $w(z) = \log z$ .

We can think of  $n = \beta$  as the force imposed on the logarithmic time scale, log *z*, caused by the requirement for the nearly simultaneous failure of *n* protective processes. The greater *n*, the greater the protective force, and the greater the bending of observed time relative to cancer's invariant clock, ticking in increments of d*T*.

All of that may seem to be a very abstract theory in relation to the actual physiological processes of cancer. However, certain empirical studies suggest that the simple geometric theory of cancer's time does in fact capture key aspects of cancer's real physiology and genetics. In particular, certain inherited genetic mutations correspond almost exactly to the predicted theoretical change in the force of simultaneity and the temporal curvature of incidence.

If a mutational knockout reduces the number of protective mechanisms by one, such that  $n \mapsto n - 1$ , then the approximate pattern of incidence changes from  $\tilde{h} \propto z^{n-1}$  to  $\tilde{h} \propto z^{n-2}$ . In other words, the force and associated curvature,  $\beta$ , is reduced by one.

Two classic studies of cancer incidence made exactly that comparison. Ashley<sup>25</sup> compared colorectal cancer incidence between groups with and without an inherited mutation that predisposes to the disease. Similarly, Knudson<sup>26</sup> compared retinoblastoma incidence between groups with and without a predisposing inherited mutation.

I analyzed those same cancers with additional data that became available after the original studies<sup>6</sup>. My analysis showed that, in each case, the groups carrying the inherited predisposing mutation had a pattern of incidence that changed relative to the control groups by reducing the estimated value of  $\beta$  by approximately one. Thus, the genetic knockouts reduced time's curvature by almost exactly the amount predicted by the reduced force of diminished simultaneity in the protective mechanisms.

#### Conclusion

A few simple invariances shape the patterns of death. That geometry does not tell us exactly how biological mechanisms influence mortality. But the geometry does set the constraints within which we must analyze the relation between pattern and process.

I started with the temporal frame of reference, dT, on which mortality has a constant rate, or velocity. That temporal frame, with unchanging rate, expresses the ticking of mortality's clock in the absence of any apparent force that would change velocity.

Given that frame without apparent force, we can then evaluate other temporal scales in terms of the forces that must be applied to change mortality's rate relative to the force-free scale. That approach focuses attention on the forces of mortality, rather than the incidence or "motion" alone, because the pattern of motion is inherently confounded with the particular temporal frame of reference<sup>18,19</sup>.

Mortality's temporal frame leads to a natural expression of invariant death with respect to a universal Gompertzian geometry. That geometric expression separates the uniform application of force from the additional distortions of time with respect to observed pattern.

The examples of nematodes and cancer illustrated how to parse observable deformations of mortality's clock with respect to invariant aspects of pattern and potential underlying explanations about process.

Until biologists can see the constraints of Gompertzian geometry on the curves of death as clearly as they can see the constraints of the logarithmic spiral on the growth curves of snail shells and goats' horns, we will not be able to read properly the relations between the molecular causes of failure and the observable patterns of death.

Put another way, geometry does not tell one how to build a bridge. But one would not want to build a bridge without understanding the constraints of geometry. Properly interpreting the duality of constraint and process with respect to pattern is among the most difficult and most important aspects of science.

#### Author contributions

SAF did all the research and wrote the article.

#### Competing interests

No competing interests were disclosed.

#### Grant information

National Science Foundation grant DEB-1251035 supports my research.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

#### References

- Thompson DW: On Growth and Form. 2nd ed (Canto reprint of 1961 revised edition). Cambridge University Press, Cambridge, 1992.
   Publisher Full Text
- Feynman RP: Statistical Mechanics: A Set Of Lectures. Westview Press, New York, 2nd edition, 1998.
   Reference Source
- Gnedenko BV, Kolmogorov AN: Limit Distributions for Sums of Independent Random Variables. Addison-Wesley, Reading, MA, 1968. Reference Source
- Jaynes ET: Probability Theory: The Logic of Science. Cambridge University Press, New York, 2003.
   Reference Source
- Stroustrup N, Anthony WE, Nash ZM, et al.: The temporal scaling of Caenorhabditis elegans ageing. Nature. 2016; 530(7588): 103–107. PubMed Abstract | Publisher Full Text | Free Full Text
- Frank SA: Age-specific incidence of inherited versus sporadic cancers: a test of the multistage theory of carcinogenesis. Proc Natl Acad Sci U S A. 2005; 102(4): 1071–1075.
   PubMed Abstract | Publisher Full Text | Free Full Text
- Frank SA: Dynamics of Cancer: Incidence, Inheritance, and Evolution. Princeton University Press, Princeton, NJ, 2007.
   PubMed Abstract
- Finch CE, Crimmins EM: Constant molecular aging rates vs. the exponential acceleration of mortality. Proc Natl Acad Sci U S A. 2016; 113(5): 1121–1123. PubMed Abstract | Publisher Full Text | Free Full Text
- Pressé S, Ghosh K, Lee J, et al.: Principles of maximum entropy and maximum caliber in statistical physics. Rev Mod Phys. 2013; 85(3): 1115–1141.
   Publisher Full Text
- Frank SA, Smith DE: Measurement invariance, entropy, and probability. *Entropy.* 2010; 12(3): 289–303.
   Publisher Full Text
- Frank SA, Smith E: A simple derivation and classification of common probability distributions based on information symmetry and measurement scale. J Evol Biol. 2011; 24(3): 469–484.
   PubMed Abstract | Publisher Full Text
- Frank SA: How to read probability distributions as statements about process. Entropy. 2014; 16(11): 6059–6098.
   Publisher Full Text
- 13. Frank SA: Common probability patterns arise from simple invariances. Entropy.

2016; 18(5): 192. Publisher Full Text

- Embrechts P, Kluppelberg C, Mikosch T: Modeling Extremal Events: For Insurance and Finance. Springer Verlag, Heidelberg, 1997; 33.
   Publisher Full Text
- Kotz S, Nadarajah S: Extreme Value Distributions: Theory and Applications. World Scientific, Singapore, 2000; 196.
   Publisher Full Text
- Coles S: An Introduction to Statistical Modeling of Extreme Values. Springer, New York, 2001.
   Publisher Full Text
- Gumbel EJ: Statistics of Extremes. Dover Publications, New York, 2004. Reference Source
- Lanczos C: The Variational Principles of Mechanics. Dover Publications, New York, 4th edition, 1986.
   Reference Source
- Frank SA: d'Alembert's direct and inertial forces acting on populations: The Price equation and the fundamental theorem of natural selection. *Entropy*. 2015; 17(10): 7087–7100.
   Publisher Full Text
- 20. Frank SA: The inductive theory of natural selection. arXiv: 1412.1285, 2014. Reference Source
- 21. Frank SA: The common patterns of nature. *J Evol Biol.* 2009; 22(8): 1563–1585. PubMed Abstract | Publisher Full Text | Free Full Text
- 22. Frank SA: Generative models versus underlying symmetries to explain biological pattern. J Evol Biol. 2014; 27(6): 1172–1178. PubMed Abstract | Publisher Full Text
- Armitage P, Doll R: The age distribution of cancer and a multi-stage theory of carcinogenesis. Br J Cancer. 1954; 8(1): 1–12.
   PubMed Abstract | Publisher Full Text | Free Full Text
- 24. Frank SA: A multistage theory of age-specific acceleration in human mortality. BMC Biol. 2004; 2: 16.
- PubMed Abstract | Publisher Full Text | Free Full Text
  25. Ashley DJ: Colonic cancer arising in polyposis coli. J Med Genet. 1969; 6(4): 376–378.
- PubMed Abstract | Publisher Full Text | Free Full Text
   Knudson AG Jr: Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci U S A. 1971; 68(4): 820–823.
   PubMed Abstract | Publisher Full Text | Free Full Text

## **Open Peer Review**

## **Current Referee Status:**

 $\checkmark$ 

Version 1

Referee Report 07 October 2016

doi:10.5256/f1000research.10185.r16246

Ophélie Ronce

French National Center for Scientific Research, University of Montpellier, Montpellier, France

This paper by S. A. Frank builds on his recent work on general common patterns of statistical distributions to comment on some invariance rules when examining the distribution of mortality data. In particular, the present paper shows that there always exists an adequate transformation of time (or age), such as, when measured on this scale, the shape of the survival curve is unaffected by any shift on this scale. Interestingly, Frank here also shows that on this scaling, the survival curve has necessarily a Gompertz form. On this scaling, the acceleration of mortality with (transformed) time is constant.

This is indeed interesting to realize that any mortality distribution would have these properties. So not much could be inferred about the mechanisms of mortality and aging just because one transformation of time would lead to the distribution of mortality satisfying those properties. The remaining question is whether we can infer something from the fact that one transformation of time rather than another produces this invariance.

As an illustration, the present paper by S. A Frank then comments on a recent study by Stroustrup et al. (2016, hereafter S2016), which showed that various genetic or environmental interventions affected mortality patterns in nematodes populations by modifying the time scale of mortality, leaving the shape of survival curves unchanged. More precisely, a log transformation of time then satisfies the invariance properties mentioned above. Interestingly, a few interventions in their data set do not exhibit invariance after a log-transformation of time. In the supplementary material of the S2016, the authors have examined a diversity of mechanistic models of mortality and checked under which perturbations of the parameters of these models one would observe the specific invariance pattern seen in their data. In some of these models, the specific invariance pattern emerges only when all sub-processes affecting mortality are affected to the same extent by the intervention (e.g. competing risk models), or when all parameters in the models are affected to a similar extent by the intervention (diffusion models). In some other models (e.g. network models), the specific invariance can emerge when only part of the processes are affected by the intervention. These conclusions appear therefore more subtle than the claim made in the main text of S2016 that "each intervention must alter to the same extent throughout adult life all physiological determinants of the risk of death". It is furthermore not very clear to me what is meant exactly by "all physiological determinants of the risk of death". Can we say that the drift and variance terms in a diffusion model for vitality correspond to different physiological determinants of the risk of death? Again we could imagine different mechanistic sub-models that would create different functional relationships between the drift and variance parameters.

Frank uses S2016 in two ways: first, as an illustration of a particular type of invariance exhibited by real

data, second, as a warning against the risk of over-interpretation of such invariances. In particular, Frank reacts to the claim cited above, which is indeed confusing. As S2016 did in their supplementary material, Frank exhibits a mechanistic model of mortality (but yet a different one from those considered S2016), which shows the same properties of invariance when only part of the processes of mortality are affected by the intervention. Both the theoretical exercise of Frank and those of S2016 show that there are many different mechanistic models that can exhibit the same specific invariance and thus raise together strong doubts about what we can infer from invariances in mortality patterns. The set of models examined by S2016 is certainly not exhaustive despite its diversity, and claims based upon this set (already more complex than the simple cited argument suggests) cannot really be generalized to all models of mortality.

I therefore agree with the general message of caution of the present paper by S.A. Frank. It however remains frustrating that the general framework that he proposes does not really help to get a better general grasp at what features of a model would produce one type of invariance rather than another, and for instance help generating hypotheses that would explain why some interventions did fail to produce the same type of invariance in S2016. This what S2016 attempted in their supplementary material.

In particular, I would be interested in a clearer illustration of the claim made by Frank that "consideration of the alternative processes with the same observable invariances leads to testable predictions about the underlying causal processes".

Interestingly, the model put forward by Frank is shift-invariant on a scale, which is not a simple log-transformation of time (see equation 17 and that following on the right column of page 6). While early in the process, the deformation of time to achieve invariance would resemble that expected under a log-scaling, this is not true later. I therefore failed to understand exactly why this model would actually fit the data presented by S2016. More explanation would be necessary here.

The whole argument page 7 about interventions affecting part of the processes when slow and fast processes determine mortality makes intuitive sense but would be more convincing if illustrated (as was done for instance with examples in the supplementary material of S2016). I found these arguments about slow and fast processes to be quite disconnected actually from the general arguments about invariance presented before.

To conclude, I found the contribution of Frank novel and interesting, but still a bit frustrating about how this perspective could help us extract more information from patterns of mortality. The message of caution is an important one. The criticism of S2016 is justified by the over-simplistic claim included in their abstract and main text, but I am a bit concerned that it may misrepresent what these authors have achieved. This claim was actually motivated, not directly by the examination of invariance in the data, but by the comparison of several models to data in a spirit similar to what Frank proposes here. That several models could exhibit the same invariance rule for different reasons was clearly shown already in S2016 (supplementary material). The common properties of models exhibiting that specific invariance rule however appear even less clear than suggested by the initial exploration of S2016 and it could be good that the present paper communicates more precisely about this point.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.

Referee Report 14 September 2016

#### doi:10.5256/f1000research.10185.r15879



### Sean Nee

The Pennsylvania State University, Pennsylvania, PA, USA

Survival analysis is a field with a long history, mainly associated with the name David Cox, and is of interest far outside of biology, e.g. the reliability analysis of machines. One element of it concerns making inferences about underlying mechanisms from observed processes. The historical development of the subject seems to have ended in a cul-de-sac which we see in Stroustrup's worm paper, discussed here, although clearly explored in an earlier Ricklefs paper, discussed by no-one. Frank offers an entirely new approach with a different starting point. It seems sensible to discuss the Frank paper at the junction where the new collides with the old, which is the Stroustrup worm paper, published in *Nature* magazine this year.

The heart of the matter, from my perspective, is this. The old development of the subject ended up at this point -

- 1. Survival data is typically well described by *both* Weibull and Gompertz distributions, which are described in terms of the 'hazard' function derived from them, describing your future chance of surviving beyond a particular time, given you have survived to that time so far.
- 2. If these hazards change for some reason (e.g. temperature for Stroustrup), we can model the effect of these changes in one of two ways, dreamed up by Cox because they 'work' for statistical analysis. One of these ways (AFT) is a time scaling of the hazard, in which everything happens in 'dog years' for example. The other (PH), as noted by Cox, is very hard to give a biological interpretation to in terms of underlying mechanisms, although it is very useful for prediction, evaluating drug efficacy etc.

Now, the Weibull distribution can be proved analytically to allow interpretation under both these ways of changing circumstance, one of which is, itself, intrinsically hard to interpret. So if the data are Weibull, which fits the Stroustrup data, you cannot really say anything about the underlying mechanism. Unless you get very inspired and allow "imagination" (Ricklefs) and "preferences" (Stroustrup Supp Info) to assist you. This is not necessarily a bad thing, of course, but it sounds like an admission that you have gone as far as you can by the traditional route. I note that the Supp. Info Figure 1.1, which is central to the narrative flow, uses a Weibull distribution, which has a log-log hazard and the Weibull plays a large role in their statistics/simulations etc. The Gompertz has a measly log hazard, which rather screws up the simple scaling story.

Frank now approaches the whole subject in a way which is new, thought-provoking, challenging and very welcome, starting from a rethinking of the basic meanings of statistical distributions. It is far to soon to have an opinion on the likely success of this approach in leading us out of the cul-de-sac. But I certainly have the opinion it should be published.

The Ricklefs paper deserves some acknowledgement and will also flag this topic for a larger audience of evolutionary ecologists, as it concerns the evolution of longevity, species variation in longevity etc. etc.

Ricklefs RE, Scheuerlein A. 2002. Biological implications of the Weibull and Gompertz models of aging. J Gerontol A Biol Sci Med Sci. 2002 Feb;57(2):B69-76. doi: 10.1093/gerona/57.2.B69

#### References

1. Ricklefs R, Scheuerlein A: Biological Implications of the Weibull and Gompertz Models of Aging. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*. 2002; **57** (2): B69-B76 Publisher Full Text

2. Stroustrup N, Anthony WE, Nash ZM, Gowda V, Gomez A, López-Moyado IF, Apfeld J, Fontana W: The temporal scaling of Caenorhabditis elegans ageing.*Nature*. 2016; **530** (7588): 103-7 PubMed Abstract | Publisher Full Text

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

*Competing Interests:* No competing interests were disclosed.