1 Supplementary Materials 1: Properties of the Original Discrete-time Entropy metric

Constant mortality

We start from the definition of Keyfitz entropy given in equation (5) in the main text,

$$\mathrm{H}_{lx} = -\frac{\sum_{x=0}^{\infty} \log(l_x) l_x}{\sum_{x=0}^{\infty} l_x}. \tag{1}$$

Next we assume mortality is constant, μ, such that survival up to age a is $l_x = \exp(-\mu x)$, and

$$\mathrm{H}_{lx} = \frac{\sum_{x=0}^{\infty} \mu x \exp(-\mu x)}{\sum_{x=0}^{\infty} \exp(-\mu x)}, \tag{2}$$

$$= \frac{\mu \exp(-\mu)}{1 - \exp(-\mu)}, \tag{3}$$

where we used the following two equations for power sums

$$\sum_{k=0}^{\infty} k z^k = \frac{z}{(1-z)^2}, \tag{4}$$

$$\sum_{k=0}^{\infty} z^k = \frac{1}{1-z}, \tag{5}$$

for the numerator and denominator of equation 2, respectively.

In general, H_{lx} is therefore not equal to one for constant mortality. Note however, that as we let μ get very small, a Taylor expansion of equation 3 shows that it becomes approximately one in the limit of small μ. Integrals can be approximated by a series of sums with infinitesimally small step sizes. In the limit of infinitesimally small time steps, survival becomes close to one (or μ close to zero), and therefore the sum approximates the continuous time formula well in this limit.

2 Supplementary Materials 2: Properties of New Discrete Keyfitz Entropy metric

Constant mortality case

We start from the definition of Keyfitz entropy given in equation (11) in the main text,

$$\mathrm{H}_N = \frac{\boldsymbol{\eta}_1^\intercal \mathbf{B}\mathbf{e}_1}{\boldsymbol{\eta}_1^\intercal \mathbf{e}_1}. \tag{6}$$

When mortality is constant throughout life, then life expectancy is the same for each age. That is, $\boldsymbol{\eta}_1^\intercal = c\mathbf{1}^\intercal$ for some constant c. Equation (6) then becomes

$$\begin{aligned} \mathrm{H}_{d2} &= \frac{c\mathbf{1}_1^\intercal \mathbf{B}\mathbf{e}_1}{c\mathbf{1}_1^\intercal \mathbf{e}_1}, & (7)\\ &= \mathbf{1}_1^\intercal \mathbf{B}\mathbf{e}_1, & (8) \end{aligned}$$

since $\mathbf{1}_1^\intercal \mathbf{e}_1 = 1$. The matrix $\mathbf{B}$ contains the distribution of age at death, or put differently, each entry contains the probability of dying at that age. Since death is inevitable, the columns of $\mathbf{B}$ sum to one. That is, for a constant mortality rate

$$\mathrm{H}_{d2} = \mathbf{1}_1^\intercal \mathbf{B}\mathbf{e}_1 = 1. \tag{9}$$

Increasing mortality rate with age

If mortality rate strictly increases with age, then life expectancy decreases with age, so the entries of $\boldsymbol{\eta}_1^\intercal$ get smaller as you go to higher ages. That is, the first entry of $\boldsymbol{\eta}_1^\intercal$, which is the denominator of H_{d2}, is larger than all other entries of $\boldsymbol{\eta}_1^\intercal$. The numerator of H_{d2} is a weighted sum of the entries of $\boldsymbol{\eta}_1^\intercal$, where the weights add to one. Given that the first entry of $\boldsymbol{\eta}_1^\intercal$ is larger than all other entries, this implies that

$$\boldsymbol{\eta}_1^\intercal \mathbf{B}\mathbf{e}_1 < \boldsymbol{\eta}_1^\intercal \mathbf{e}_1, \tag{10}$$

which in turn implies that $H_N < 1$.

Now we will briefly elaborate why the inequality above is true if the entries of $\boldsymbol{\eta}_1^\intercal$ get smaller as you go to higher ages. Lets call the vector with the distribution of age at death for a newborn individual, $\mathbf{B}\mathbf{e}_1 = \mathbf{p}$, such that the entries of $\mathbf{p}$ sum to one, $\sum_i p_i = 1$, because the columns of $\mathbf{B}$ sum to one. We can then write equation (6) as

$$\mathrm{H}_{d2} = \frac{\boldsymbol{\eta}_1^\intercal \mathbf{p}}{\eta_1}, \tag{11}$$

where η_1 is the first entry of the vector $\boldsymbol{\eta}_1^\intercal$. Written out in terms of its components this is equal to

$$\mathrm{H}_{d2} = \frac{\sum_i \eta_i^\intercal p_i}{\eta_1}, \tag{12}$$

For the case of increasing mortality rates, we know that $\eta_1 > \eta_i$ for all $i > 1$. Therefore,

$$\eta_1 p_i > \eta_i^\intercal p_i, \text{ for all } i > 1 \tag{13}$$

which implies that

$$\sum_i \eta_1 p_i > \sum_i \eta_i^\intercal p_i, \tag{14}$$

and therefore that

$$\eta_1 > \sum_i \eta_i^\intercal p_i. \tag{15}$$

Decreasing mortality rate with age

If mortality rate strictly decreases with age, then life expectancy increases with age, so the entries of $\boldsymbol{\eta}_1^\mathsf{T}$ get bigger as you go to higher ages. That is, the first entry of $\boldsymbol{\eta}_1^\mathsf{T}$, which is the denominator of H_{d2}, is smallest of the entries of $\boldsymbol{\eta}_1^\mathsf{T}$. The numerator of H_{d2} is a weighted sum of the entries of $\boldsymbol{\eta}_1^\mathsf{T}$, where the weights add to one. Given that the first entry of $\boldsymbol{\eta}_1^\mathsf{T}$ is smaller than all other entries, this implies that

$$\boldsymbol{\eta}_1^\mathsf{T}\mathbf{B}\mathbf{e}_1 > \boldsymbol{\eta}_1^\mathsf{T}\mathbf{e}_1, \tag{16}$$

which in turn implies that $H_N > 1$.