

Mathematical Appendix

Preliminaries

Let x be age, $\mu(x)$ mortality hazard, and

$$\Lambda_\mu(x_0, x) = \Lambda(x) - \Lambda(x_0) = \int_{x_0}^x \lambda(a) da \quad (1)$$

the cumulative mortality hazard contingent on surviving to age x_0 , where for Λ_μ and other functions we adopt the convention $\Lambda_\mu(0, x) = \Lambda_\mu(x)$. The survivorship function contingent on surviving to age x_0 is

$$l(x_0, x) = \frac{l(x)}{l(x_0)} = e^{-\Lambda(x_0, x)} \quad (2)$$

and the stable age distribution [Caswell, 2010] contingent on surviving to age x_0 is

$$c(x_0, x) = \frac{e^{-r[x-x_0]}l(x_0, x)}{\int_{x_0}^{\infty} e^{-r[a-x_0]}l(x_0, a) da} = \frac{e^{-rx}l(x)}{\int_{x_0}^{\infty} e^{-ra}l(a) da}, \quad (3)$$

where r is the population growth rate. There are two pertinent probability distributions to consider with respect to age-at-death: (1) the individual age-at-death distribution and (2) the assemblage age-at-death distribution. The former gives the probability density of ages-at-death for an individual who has survived to age x_0 and depends on only the mortality hazard, whereas the latter gives the probability density of ages-at-death for a sample in an archaeological assemblage and depends on both the mortality hazard and population growth rate. The distributions are identical if the population growth rate is zero. The individual age-at-death density contingent on surviving to age x_0 is proportional to the survival rate from age x_0 to x times the mortality rate at age x ,

$$f(x_0, x) = \frac{l(x_0, x)\mu(x)}{\int_{x_0}^{\infty} l(x_0, a)\mu(a) da} = \frac{l(x)\mu(x)}{\int_{x_0}^{\infty} l(a)\lambda(a) da}. \quad (4)$$

The integral in the denominator ensures that $f(x_0, x)$ is a proper probability density that integrates to 1. Similarly, the archaeological age-at-death distribution is proportional to the stable age distribution times the mortality,

$$g(x_0, x) = \frac{c(x_0, x)\mu(x)}{\int_{x_0}^{\infty} c(x_0, a)\mu(a) da} = \frac{e^{-rx}l(x)\mu(x)}{\int_{x_0}^{\infty} e^{-ra}l(a)\mu(a) da}. \quad (5)$$

Gompertz-Makeham Likelihood

Michael Holton Price

December 2019

1 Hazard model

Let x be age, $\lambda(x)$ mortality hazard, and

$$\Lambda(x_0, x) = \Lambda(x) - \Lambda(x_0) = \int_{x_0}^x \lambda(a) da \quad (1)$$

the cumulative hazard contingent on surviving to age x_0 , where for Λ and other functions we adopt the convention $\Lambda(0, x) = \Lambda(x)$. The survivorship function contingent on surviving to age x_0 is

$$l(x_0, x) = \frac{l(x)}{l(x_0)} = e^{-\Lambda(x_0, x)} \quad (2)$$

and the stable age distribution [Caswell, 2010] contingent on surviving to age x_0 is

$$c(x_0, x) = \frac{e^{-r[x-x_0]}l(x_0, x)}{\int_{x_0}^{\infty} e^{-r[a-x_0]}l(x_0, a) da} = \frac{e^{-rx}l(x)}{\int_{x_0}^{\infty} e^{-ra}l(a) da}. \quad (3)$$

The age-at-death probability density, again contingent on surviving to age x_0 , is

$$g(x_0, x) = \frac{c(x_0, x)\lambda(x)}{\int_{x_0}^{\infty} c(x_0, a)\lambda(a) da} = \frac{e^{-rx}l(x)\lambda(x)}{\int_{x_0}^{\infty} e^{-ra}l(a)\lambda(a) da}. \quad (4)$$

We assume a Gompertz-Makeham mortality hazard,

$$\lambda(x) = b_1 + b_2 e^{-b_3 x}, \quad (5)$$

for which the cumulative hazard and survivorship function are, respectively,

$$\Lambda(x) = b_1 x + \frac{b_2}{b_3} [e^{b_3 x} - 1] \quad (6)$$

and

$$l(x) = \exp \left[\frac{b_2}{b_3} \right] * \exp [-b_1 x] * \exp \left[-\frac{b_2}{b_3} e^{b_3 x} \right]. \quad (7)$$

Define $z(x)$ to be the numerator of Equation ?? given the Gompertz-Makeham parameterization and excluding the constant term,

$$z(x) = \exp \left[-(r + b_1)x - \frac{b_2}{b_3} e^{b_3 x} \right] (b_1 + b_2 e^{b_3 x}). \quad (8)$$

The integral of $z(x)$,

$$Z(x) = \int_0^x z(a) da, \quad (9)$$

satisfies $Z(x_0, x) = Z(x) - Z(x_0)$, where again we adopt the convention $Z(0, x) = Z(x)$. To evaluate $Z(x)$, apply the change of variable $u = b_2/b_3 \cdot e^{b_3 x}$,

$$Z(x) = \frac{1}{b_3} \left[\frac{b_2}{b_3} \right]^{\frac{r+b_1}{b_3}} \int_{\frac{b_2}{b_3}}^{\frac{b_2}{b_3} e^{b_3 x}} u^{-1} u^{-\frac{r+b_1}{b_3}} e^{-u} [b_1 + b_3 u] du. \quad (10)$$

This evaluates to

$$Z(x) = \frac{1}{b_3} \left[\frac{b_2}{b_3} \right]^{\frac{r+b_1}{b_3}} (b_1 [\gamma(s, \frac{b_2}{b_3} e^{b_3 x}) - \gamma(s, \frac{b_2}{b_3})] + b_3 [\gamma(s+1, \frac{b_2}{b_3} e^{b_3 x}) - \gamma(s+1, \frac{b_2}{b_3})]) \quad (11)$$

where $\gamma(\cdot, \cdot)$ is the lower incomplete gamma function and $s = -\frac{r+b_1}{b_3}$. Using the identity $\gamma(s+1, u) = s\gamma(s, u) - u^s e^{-u}$, Equation 11 simplifies to

$$Z(x) = -\frac{r}{b_3} \left[\frac{b_2}{b_3} \right]^{\frac{r+b_1}{b_3}} [\gamma(-\frac{r+b_1}{b_3}, \frac{b_2}{b_3} e^{b_3 x}) - \gamma(-\frac{r+b_1}{b_3}, \frac{b_2}{b_3})] - e^{-(r+b_1)x} e^{-\frac{b_2}{b_3} e^{b_3 x}} + e^{-\frac{b_2}{b_3}} \quad (12)$$

The age-at-death probability density is for Gompertz-Makeham mortality is

$$g(x_0, x) = \frac{z(x)}{Z(x_0, \infty)}. \quad (13)$$

The cumulative distribution function for age-at-death is

$$G(x_0, x) = \frac{Z(x_0, x)}{Z(x_0, \infty)}. \quad (14)$$

References

[Caswell, 2010] Caswell, H. (2010). Reproductive value, the stable stage distribution, and the sensitivity of the population growth rate to changes in vital rates. *Demographic Research*, S8(19):531–548.

Gompertz-Makeham parameterization

We assume a Gompertz-Makeham mortality hazard,

$$\mu(x) = b_1 + b_2 e^{-b_3 x}, \quad (6)$$

for which the cumulative mortality hazard and survivorship function are, respectively,

$$\Lambda_\mu(x) = b_1 x + \frac{b_2}{b_3} [e^{b_3 x} - 1] \quad (7)$$

and

$$l(x) = \exp \left[\frac{b_2}{b_3} \right] * \exp [-b_1 x] * \exp \left[-\frac{b_2}{b_3} e^{b_3 x} \right]. \quad (8)$$

Define $z_f(x)$ and $z_g(x)$ to be the numerators of Equations 4 and 5 given the Gompertz-Makeham parameterization and excluding constant terms,

$$z_f(x) = \exp \left[-b_1 x - \frac{b_2}{b_3} e^{b_3 x} \right] (b_1 + b_2 e^{b_3 x}) \quad (9)$$

and

$$z_g(x) = \exp \left[-(r + b_1)x - \frac{b_2}{b_3} e^{b_3 x} \right] (b_1 + b_2 e^{b_3 x}). \quad (10)$$

The integral of $z_f(x)$,

$$Z_f(x) = \int_0^x z_f(a) da, \quad (11)$$

satisfies $Z_f(x_0, x) = Z_f(x) - Z_f(x_0)$ (and similarly for Z_g), where again we adopt the convention $Z(0, x) = Z(x)$. Although both Z_f and Z_g have analytic formulas that depend on the upper incomplete gamma function, we have found that calculation of the upper incomplete gamma function (at least, using the R package `pracma`) is sometimes numerically inaccurate when it relies on the recursion formula $\gamma(s+1, u) = s \gamma(s, u) - u^s e^{-u}$. For this reason, we utilize direct numerical integration to calculate $Z_f(x_0, x)$ and $Z_g(x_0, x)$. The individual age-at-death probability density for Gompertz-Makeham mortality is

$$f(x_0, x) = \frac{z_f(x)}{Z_f(x_0, \infty)}, \quad (12)$$

whereas that for the assemblage age-at-death probability density is

$$g(x_0, x) = \frac{z_g(x)}{Z_g(x_0, \infty)}. \quad (13)$$

For a single observation x , the contribution to the likelihood (which we maximize) is $g(x_0, x)$.

Accounting for multiple observations

While Equation 13 gives the contribution to the likelihood of a single observation, for some individuals we have estimates of the age-at-death coming from multiple teeth. Hence, the likelihood must be modified to account for the possibility of multiple estimates of the age-at-death. To do so, we assume that the probability of a measurement $x^{(s)}$ for sample s is normally distributed with mean x (an unobserved true age) and standard deviation σ . Given this, the contribution of an individual to the likelihood is

$$L = \int_{x_0}^{\infty} g(x_0, a) \prod_s \frac{e^{-\frac{[x^{(s)} - a]^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} da = \frac{\int_{x_0}^{\infty} z_g(a) \prod_s \frac{e^{-\frac{[x^{(s)} - a]^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} da}{Z_g(x_0, \infty)}. \quad (14)$$

References

- [Caswell, 2010] Caswell, H. (2010). Reproductive value, the stable stage distribution, and the sensitivity of the population growth rate to changes in vital rates. *Demographic Research*, S8(19):531–548.