

A discrete Weibull model of population turnover

Bas Cornelissen

August 9, 2017

Population turnover is commonly modelled by replacing one random agent in every round. Such a constant mortality rate is not very realistic, and this appendix proposes an alternative, discrete Weibull model. It is reparametrised such that the mean life expectancy is the only parameter.

How can we realistically model the mortality-rate in a population? This is a central question in *survival analysis* (e.g. Rog  r  uez 2007). In the context of a model of language evolution, if one agent dies in every iteration, the probability that any given agent dies at time t is thus γ — *given that it was still alive at time $t - 1$* . If T is the random variable that measures the time of death of an agent, this means $p(T = t \mid T \geq t) = \gamma$. This conditional probability is known as the *hazard rate* $h(t)$, as it measures the rate of death (hazard) occurring at t .¹ In our example the hazard rate was constant. But, as explained in chapter ??, models with constant hazard rates are poor models of life-expectancy in human populations and demographers usually adopt either the *Weibull* or *Gompertz* distribution (Juckett and Rosenberg 1993). We here consider the simpler Weibull distribution, mainly because several discrete analogues have been proposed (Nakagawa and Osaki 1975; Stein and Dattero 1984; Almalki and Nadarajah 2014).

DISCRETE WEIBULL DISTRIBUTION The Weibull distribution (Weibull 1951) is a continuous distribution parametrized by a scale parameter $\kappa > 0$ and a shape parameter $\lambda > 0$. If $\kappa > 1$ the distribution is unimodal, meaning that most agents die around the same age, which is in turn determined by λ (see figure 1). For completeness, the density of a Weibull-distributed random variable T is

$$p(t \mid \lambda, \kappa) = \kappa/\lambda \cdot (t/\lambda)^{\kappa-1} \cdot \exp\left(- (t/\lambda)^\kappa\right). \quad (1)$$

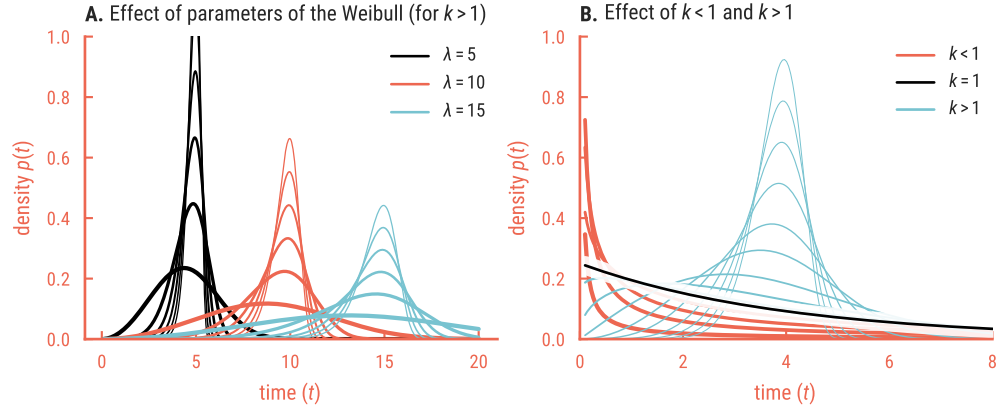
Since language games are discrete time models, we use a discrete approximation known as the *Discrete Weibull* distribution² (Nakagawa and Osaki 1975), which preserves the so called *survival* function of the continuous distribution. The *survivor* function $S(t) =$

² I reparametrized the distribution by taking $\beta := \kappa$ and $q := \exp(-\lambda^{-\kappa})$, which is both computationally and conceptually more convenient.

¹ It is usually defined for continuous T with density f as $h(t) = \frac{f(t)}{1-F(t)}$, in which case it is not a conditional probability but the *rate* of instantaneous hazard.

FIGURE 1 The Weibull distribution can model the probability that an agent dies at time t . **A.** Varying the parameters of a Weibull distribution illustrates that λ is a scale parameter and κ a shape parameter. **B.** If $\kappa > 1$ the Weibull is a unimodal distribution, whose variance decreases with higher κ (thinner lines), but for $\kappa < 1$ the distribution has no mode. When $\kappa = 1$ the Weibull reduces to an exponential distribution.

FIG04



$p(T \geq t)$ measures the probability of surviving to at least time t . The Weibull distribution, this function takes the form

$$S(t) = \exp\left(-\left(t/\lambda\right)^\kappa\right), \quad (2)$$

and the Discrete Weibull is defined as the discrete distribution with the same survival function. This can be done, since the probability mass function is fully determined by the survival function:

$$p(T = t) = S(t) - S(t + 1) = \exp\left(-\left(t/\lambda\right)^\kappa\right) - \exp\left(-\left((t+1)/\lambda\right)^\kappa\right). \quad (3)$$

It should be stressed that the resulting distribution *approximates* the Weibull distribution, which for our purposes is sufficient.

The hazard rate of the Discrete Weibull distribution can be computed as

$$h(t) = p(T = t \mid T \geq t) = \frac{S(t) - S(t + 1)}{S(t)} = 1 - \exp\left(\left(t/\lambda\right)^\kappa - \left((t+1)/\lambda\right)^\kappa\right). \quad (4)$$

Recall that the hazard rate is the probability that an agent dies at time t , given that it hasn't died yet. Therefore, if we want to model a population where the probability of dying at time T approximately follows a Weibull distribution, we should in every round flip a coin with weight $h(t)$ to decide whether the agent dies.

MODELLING POPULATION TURNOVER The Discrete Weibull gives us a more realistic model of population turnover, but its behaviour is regulated by two parameters: λ and κ . Ideally a *single* parameter would interpolate between immediate death (iterated learning) and immortality (naming game). This can be done by defining λ and κ in terms of a $\gamma \geq 1$:

$$\kappa(\gamma) := \log(\gamma) + \kappa_0 \quad \text{and} \quad \lambda(\gamma) := \frac{\gamma}{\Gamma(1 + 1/\kappa(\gamma))}, \quad \gamma \geq 1 \quad (5)$$

where Γ is the gamma function. Figure 2 illustrates the effect of γ . Three reasons underly this reparametrization. First, the term $\Gamma(1 + 1/\kappa(\gamma))$ makes γ interpretable: γ is the

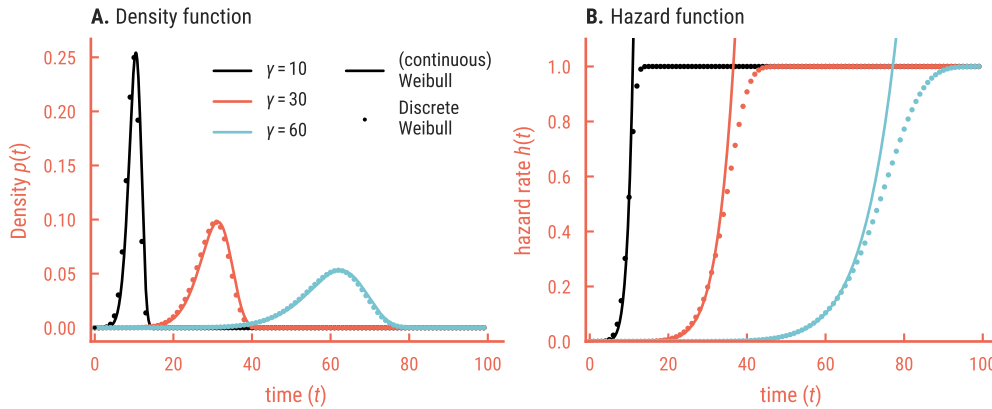


FIGURE 2 The single-parameter version of the continuous and discrete Weibull distribution. A. The distributions closely line up and γ is easily interpretable. B. The hazard rate increases with time, thus capturing ageing effects. Note that a continuous hazard rate $h(t)$ is not a distribution and exceeds 1.

FIG04

mean of the continuous distribution $\text{Weibull}(\lambda(\gamma), \kappa(\gamma))$. Second, scaling $\kappa(\gamma)$ logarithmically with γ results in a realistic mortality distribution for all $\gamma \geq 1$. Third, the constant κ_0 guarantees that for $\gamma = 1$ the hazard rate is approximately 1, corresponding to instant death in iterated learning. I opt for³ $\kappa_0 = 5$.

In sum, we have defined a discrete Weibull model, approximating the continuous Weibull, but parametrised by a single parameter γ , the average life expectancy. When used in combination with a random walk through a population of fixed size, γ thus interpolates between iterated learning ($\gamma = 1$) and a naming game ($\gamma = \infty$).

Bibliography

- Almalki, Saad J. and Saralees Nadarajah (2014). “A New Discrete Modified Weibull Distribution”. In: *IEEE Transactions on Reliability* 63.1, pp. 68–80. DOI: 10.1109/TR.2014.2299691. arXiv: arXiv:1307.3925v1.
- Juckett, David A. and Barnett Rosenberg (1993). “Comparison of the Gompertz and Weibull Functions as Descriptors for Human Mortality Distributions and their Intersections”. In: *Mechanisms of Ageing and Development* 69.1-2, pp. 1–31. DOI: [http://doi.org/10.1016/0047-6374\(93\)90068-3](http://doi.org/10.1016/0047-6374(93)90068-3).
- Nakagawa, Toshio and Shunji Osaki (1975). “The Discrete Weibull Distribution”. In: *IEEE Transactions on Reliability* R-24.5, pp. 300–301. DOI: 10.1109/TR.1975.5214915.
- Rogríguez, Germán (2007). “Survival Models”. In: *Lecture Notes on Generalized Linear Models*. Chap. 7.
- Stein, William E. and Ronald Dattero (1984). “A New Discrete Weibull Distribution”. In: *IEEE Transactions on Reliability* R-33.2, pp. 196–197. DOI: 10.1109/TR.1984.5221777.
- Weibull, W. (1951). “A statistical distribution function of wide applicability”. In: *Journal of applied mechanics* 18, pp. 293–297.

³ This results in the hazard rate $h(1 | \lambda = 1) > 1 - 10^{-8}$, which seems sufficiently close to 1.