Carlo F. Dondena Centre for Research on Social Dynamics

DONDENA WORKING PAPERS

Concentration in survival times and lengevity: The log-scale-location family of failure time models

Chiara Gigliarano, Ugofilippo Basellini, and Marco Bonetti

Working Paper No. 66
September 2014

Carlo F. Dondena Centre for Research on Social Dynamics
Università Bocconi
Via Guglielmo Röntgen 1, 20136 Milan, Italy
http://www.dondena.unibocconi.it

ISSN-2035-2034

Concentration in survival times and longevity: the log-scale-location family of failure time models

Chiara Gigliarano*

DISES, Università Politecnica delle Marche, Ancona

Ugofilippo Basellini†

Just Retirement, London

Marco Bonetti‡

Department of Policy Analysis and Public Management, Bocconi University, Milan

September 17, 2014

Abstract

Evidence suggests that the significantly higher life expectancy levels witnessed over the past centuries are associated with a lower concentration of survival times, both cross-country and over time. The purpose of this work is to study the relationships that exist among models for the evolution of survival distributions, longevity measures, and concentration. We first study relationships between concentration and cohort longevity through empirical comparisons. We then propose a family of survival models that can be used to capture such trends in longevity and concentration across survival distributions.

Keywords: Survival analysis; Longevity; Gini index; Life tables.

1 Introduction

The term *longevity* denotes the long duration of life and is used as a synonym for long life expectancy in demography. It is well-known that a significant increase in longevity has been witnessed during the past

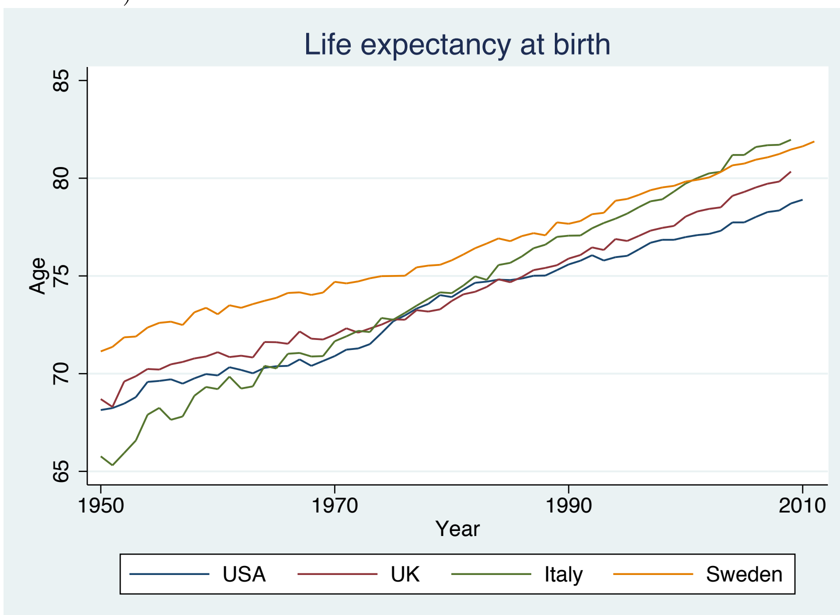
*Email: c.gigliarano@univpm.it

†Email: ugo.basellini@justretirement.com

‡Email: marco.bonetti@unibocconi.it

several centuries. For example, Figure 1 shows how life expectancy has increased since the second half of the 20th century for a selection of OECD countries.

Figure 1: Life expectancy at birth in four OECD countries from 1950 to 2010 (Source: Elaboration on Human Mortality Database)



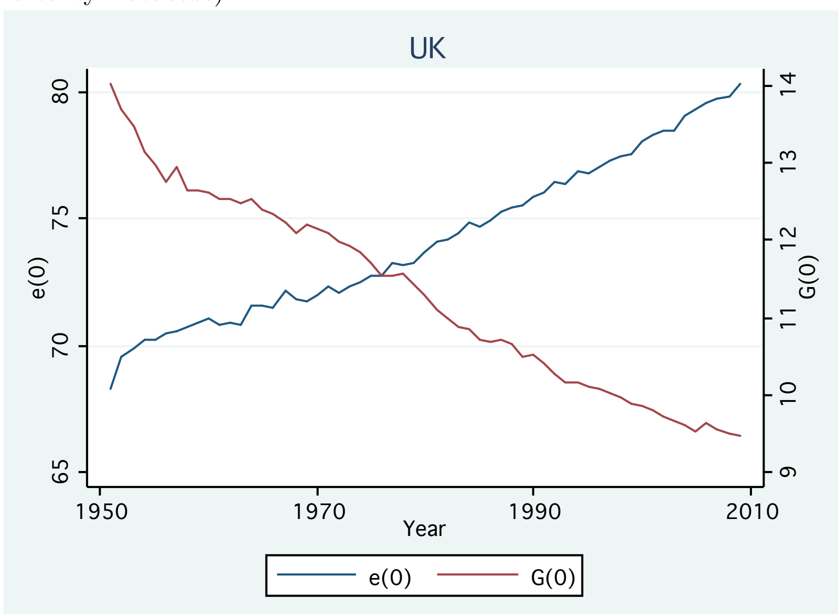
A recent study by Baudisch (2011) emphasizes the importance of distinguishing between two dimensions of aging: *pace* and *shape*. Pace refers to the time aspect of aging, and it is measured by variables like life expectancy and longevity, which summarize the timing of death. Shape refers to the age-pattern of mortality or how mortality changes with age. It captures the time-standardized change in mortality and it reveals whether mortality (on average) increases or decreases over age, and whether these changes are more or less pronounced.

On the one hand, populations experienced a prolonged life, which mainly reflects a reduction in the overall level of mortality, or, as Baudisch (2011) argues, a change in the pace of life. On the other hand, populations experienced an exceeding concentration of deaths and a shift of the death hump towards higher ages (Canudas-Romo, 2008; Kannisto, 2000), which implies an increasing steepness of mortality change over the life-course, or, as Baudisch argues, an increasing steepness of the shape of aging.

Well-known measures in demography that account for lifespan disparity are considered good indicators of shape, such as the Gini coefficient or the coefficient of variation. Intuitively, this can be understood by recognizing that a high shape value is analogous to low variability in the age at death. Hence, if everyone died at the same time, variability is zero, and the age-pattern of mortality shows maximum steepness, rising from zero to infinity at the unique age at death. In contrast, constant or falling mortality patterns would lead to high inequality in age at death, many dying early in the life course while a few experience exceptionally long lifespans.

Indeed, evidence suggests that higher life expectancy is associated with a lower concentration of survival times, both cross-country and over time. Figure 2 shows the joint trend of longevity and lifetime concentration for the UK.

Figure 2: Longevity trend ($e(0)$) versus concentration trend ($G(0)$) in the UK, 1950-2010 (Source: Elaboration on Human Mortality Database)



Shkolnikov et al. (2003) perform an empirical analysis for approximately 45 countries for the years 1960-1990, revealing a tight negative association between life expectancy at birth and the Gini coefficient. Specifically, during the first three quarters of the 20th century the inter-individual inequality in length of life has been declining. But, in the last three decades this trend has become weaker, with life expectancy continuing to increase while the decline in the inequality in length of life has slowed down or even stopped in low-mortality countries.

In human demography, this framework could be related to concepts introduced by the shifting mortality and compression of mortality hypotheses. The shifting mortality hypothesis suggests a delay in the mortality schedule, but with a shape that remains the same (Bongaarts and Feeney, 2002; Canudas-Romo, 2008). The compression of mortality hypothesis suggests a change in variability in the age at death (Fries, 1980; Kannisto, 2000). Changes in mortality can then be produced by a change in pace or by a change in shape, and more commonly by changes in both dimensions simultaneously.

Wilmoth and Horiuchi (1999) also explain these observed patterns by a principal historical change in the age pattern of mortality. In the first part of the 20th century, the historical lowering of mortality rates was much more pronounced in the young than in the aged, resulting in more and more concentrated life table deaths at old ages. In other words, this historical reduction of infant mortality caused great equalization of ages at death. Since the Gini coefficient decreases when life table ages at death concentrate around the average age at death, inequality in length of life has been decreasing during this period. However, after a certain point in time (in the 1950s, 1960s or 1970s, depending on the country), at which mortality of young ages had already been reduced to low values, its further reduction was unable to reduce significantly dispersion of ages at death. Moreover, in the 1980s and 1990s, the mortality decline in countries with low mortality was more pronounced at old ages than at middle ages. This process increases inequality in length of life, thus resulting in higher inequality of survival times.

Since the Gini concentration index is scale-invariant, this suggests that the survival times (especially when examined over time for a given country) may not increase uniformly with the same rate across the whole population, but may rather evolve according to different patterns within different subpopulations.

For example, specific advances in health care that cause an increase in survival for a fraction of the population may produce an increase in longevity, which may correspond to the observed decrease in the concentration of survival times. Lastly, the existence of a biological upper limit to human survival could also produce the observed decreases in concentration.

The purpose of this work is to analyze the effect of changes in survival distributions on the concentration of survival times and study the relationship between longevity and concentration. We first review the Gini concentration index for survival data and its link with stochastic dominance (Section 2). We then show that proportional hazard models are not particularly well-suited to the study of the life shape and pace patterns observed in human population (Section 3), and we suggest a family of models that can be used to capture differences in longevity across survival distributions, as well as study how the Gini index evolves under such models (Section 4). Special cases of the family of models proposed are discussed in Section 5. We close with discussion in Section 6.

2 The Gini concentration index and Lorenz ordering

The Gini coefficient is probably the most common statistical index employed in the social sciences for measuring concentration in the distribution of a positive random variable. It is mainly used in Economics as a measure of income or wealth inequality; see, e.g., Gini (1912, 1914), Nygard and Sandröm (1981), Kakwani

(1980). The Gini concentration index of income distributions measures the relative importance of very high incomes in the distribution.

The Gini coefficient has also been used to describe concentration in a distribution of length of life, among different socio-economic groups, and to evaluate inequality in survival times (see, e.g., Hanada 1983, Shkolnikov et al. 2003).

Consider a nonnegative random variable T with cumulative distribution function F, $F(t) = F_T(t) = P(T \leq t)$, survival function S, $S(t) = S_T(t) = 1 - F_T(t)$, probability density function $f(t)$, finite expected value $\mu = \int_{\Re+} S(t)dt$, and variance $Var(T)$. Here we will focus on T as being a survival time.

The Gini coefficient of concentration corresponding to the cumulative distribution function $F(t)$ is defined as

$$G = \frac{\delta}{2\mu} = \frac{\int_{\Re+}\int_{\Re+} |t_1 - t_2|\, dF(t_1)dF(t_2)}{2\mu}$$

(see Gini 1912, 1914, or Kendall and Stuart 1977).

The Gini coefficient varies between 0 (the case of perfect equality) and 1 (perfect inequality), and it is invariant under scale transformations. For length-of-life distributions, it is equal to 0 if all individuals die at the same age, and equal to 1 if all people but one die at age 0 and the one individual dies at a positive age. If a (small) group of individuals lives much longer than the rest of the population, then G will tend to be large.

Several other equivalent ways to define the Gini index exist. In particular, an alternative expression that will be used throughout this paper is given by

$$G = 1 - \frac{\int_0^\infty S^2(t)dt}{\int_0^\infty S(t)dt} \tag{1}$$

(see, e.g., Michetti and Dall'Aglio 1957 and Hanada 1983).

It can be shown that the Gini index is consistent with different orderings of distributions; specifically, such an index is coherent with the ordering induced by the Lorenz curve $L(p)$, defined as

$$L(p) = \frac{1}{\mu}\int_0^p F^{-1}(v)dv, \quad 0 \leq p \leq 1,$$

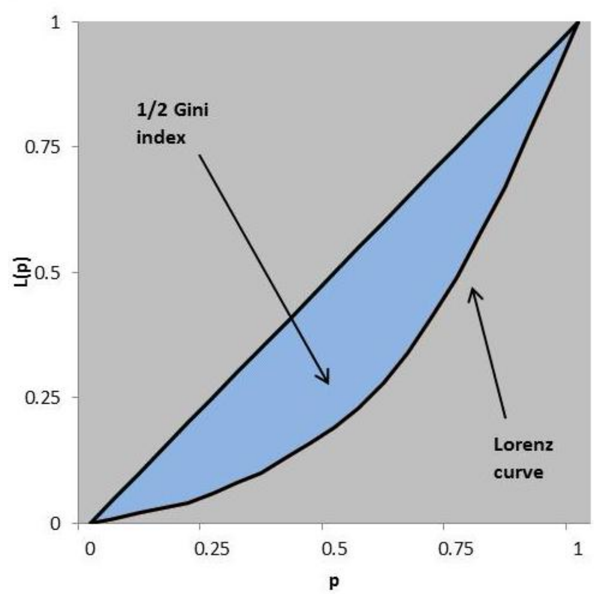
where $F^{-1}(v)$ is the left-continuous version of the inverse of F, defined as $F^{-1}(v) = \inf\{y : F(y) \geq v\}$ (see Lorenz 1905, Pietra 1915, Gastwirth 1971).

The Lorenz curve associates the cumulative proportion of total survival time with the proportion of individuals, arranged in an ascending order of survival time that "receives" such a proportion of cumulated survival time. If the Lorenz curve of the random variable T_1 is never above the Lorenz curve of the random variable T_2, then the Gini index assumes lower values for the distribution of T_2 than for the distribution of T_1.

Indeed, the Gini coefficient can be expressed in terms of the Lorenz curve $L(p)$ as the area between the diagonal (equality) 45-degree segment and the Lorenz curve, divided by the whole area below the diagonal (see Figure 3):

$$G = 1 - 2\int_0^1 L(p)dp.$$

Figure 3: Lorenz curve and Gini index



The Gini index is also consistent with other kinds of stochastic dominance (see for example Atkinson 1970, Muliere and Scarsini 1989, Shaked and Shanthikumar 1994).

3 The proportional hazard regression model with Weibull baseline

We now move to considering families of models in survival analysis that can be used to capture the simultaneous evolution of the concentration in survival times and longevity. Hence, we are interested in studying the behavior of the Gini index as a function of birth year (considered as the covariate x)

$$G(x) = 1 - \frac{\int_0^\infty S^2(t;x)dt}{\int_0^\infty S(t;x)dt}. \tag{2}$$

One of the most common models used in the survival analysis for estimating covariate effects is the Cox proportional hazards model. The Cox proportional hazards model is a semiparametric lifetime regression model proposed by Cox (1972), which assumes the hazard function for the survival time T to be, when considering just one covariate x, of the form

$$r(t;x) = r_0(t)e^{ax},$$

where $r_0(t)$ is an arbitrary baseline hazard function, a is a regression parameter, and x the covariate.

Under this model, the hazard function is assumed to increase exponentially per unit increase of the covariate, so the covariate therefore has a constant multiplicative effect on the hazard rate across all values of t.

The proportional hazards regression model assumes the following structure for the conditional (on x) survival function:

$$S(t;x) = S_0(t)^{r(x)}, \tag{3}$$

where $S_0(t)$ is an arbitrary baseline survival function and $r(x)$ is the risk score that is the exponential transformation of the linear predictor e^{ax} (see Lawless, 2003).

As a special case we consider a baseline survival function belonging to the Weibull family, that is $S_0(t) = e^{-(t/\alpha)^\beta}$, with $\alpha, \beta > 0$. We have the following result[1]:

Proposition 3.1. *Assuming a proportional hazards regression model with Weibull baseline with parameters (α, β), that is $S_0(t) = e^{-(t/\alpha)^\beta}$, the Gini index is equal to*

$$G(x) = 1 - (0.5)^{1/\beta},$$

which implies that the Gini index is constant with respect to the covariate x.

Proof. Consider the Gini index as a function of x, defined in (2)

$$G(x) = 1 - \frac{\int_0^\infty S^2(t;x)dt}{\int_0^\infty S(t;x)dt}. \tag{4}$$

Under the Cox model as in (3), the integrand function of the denominator in (4) becomes

$$S(t;x) = e^{-r(x)\cdot(t/\alpha)^\beta} = e^{-(r(x)^{1/\beta}\cdot\frac{t}{\alpha})^\beta} = e^{-(t/\alpha^*)^\beta},$$

with $\alpha^* = \frac{\alpha}{r(x)^{1/\beta}}$. Therefore, $S(t;x)$ is still the survival function of a Weibull distribution with parameters $(\alpha^* = \frac{\alpha}{r(x)^{1/\beta}}; \beta)$.

From Bonetti, Gigliarano and Muliere (2009) we know that if $S(t) = e^{-(t/\alpha)^\beta}$ (Weibull distribution with parameters (α, β)), then the Gini index G as defined in (1) is equal to $G = 1 - (0.5)^{1/\beta}$, hence it does not

[1] Ostasiewicz and Mazurek (2013) provide a numerical estimation of the Gini index under the Weibull distribution.

depend on the scale parameter α. Therefore, $G(x)$ will also not depend on the scale parameter (which is a function of x), but instead only on the shape parameter β, which is indeed constant with respect to x. $\square$

Therefore, the proportional hazards model does not seem to be suitable for capturing the phenomenon of our interest, that is, the decreasing behavior of the Gini index of concentration as a function of the cohort year of birth.

4 The log-scale-location family of failure time models

We now discuss a family of regression models aimed at modeling the joint behavior of life expectancy and concentration in lifetimes over time. We consider the following *log-scale-location model* (see Lawless, 2003):

$$S(t;x) = S_0\left(\frac{\log(t) - u(x)}{b(x)}\right), \tag{5}$$

where $u(x) \in \mathbb{R}$, $b(x) > 0$, and x is any generic covariate. For our purposes, we will use x to indicate the cohort year of birth.

If T has log-scale-location distribution as in (5) and $Y = \log(T)$, then we say that Y has a *scale-location* distribution of the type

$$S(y;x) = S_0\left(\frac{y - u(x)}{b(x)}\right). \tag{6}$$

Common distributions that satisfy model (5) are the following: the Weibull, log-normal, and log-logistic distributions for T correspond, respectively, to extreme value, normal, and logistic distributions for Y.

We now show that the family of models proposed in (5) provides, under a suitable choice of the parameters, a complete ordering of survival distributions in terms of concentration.

Theorem 1. *If $S(t;x) = S_0\left(\frac{log(t)-u(x)}{b(x)}\right)$ with S_0 strictly decreasing and where $b(x)$ is a non-decreasing (non-increasing) function, then the Gini index $G(x) = 1 - \frac{\int_0^\infty S^2(t;x)dt}{\int_0^\infty S(t;x)dt}$ is a non-decreasing (non-increasing) function of x, for any bounded function $u(x)$.*

The proof of Theorem 1 requires some additional preliminary definitions and results from the theory of stochastic orderings (Shaked and Shanthikumar 2002). We first recall the following definitions of *Dispersive order*, *Star order*, and *Lorenz order*:

Definition 4.1. *Let T_1 and T_2 be two random variables with distribution functions F and G, respectively. Let F^{-1} and G^{-1} be the right-continuous inverses of F and G, respectively. Then T_1 is said to be smaller* *than T_2 in the dispersive order (denoted as $T_1 \leq_{disp} T_2$) if*

$$F^{-1}(\beta) - F^{-1}(\alpha) \leq G^{-1}(\beta) - G^{-1}(\alpha),\ \forall\ 0 \leq \alpha \leq \beta \leq 1.$$

The dispersive order compares variables T_1 and T_2 in terms of variability, since it requires the difference between any two quantiles of T_1 to be smaller than the corresponding quantiles of T_2 (see Shaked and Shanthikumar 2002, page 148).

Definition 4.2. *Consider non-negative random variables T_1 and T_2 as defined in Definition 2.1. We say that T_1 is smaller than T_2 in the star order (denoted by $T_1 \leq_* T_2$) if*

$$G^{-1}F(t_1)/t_1 \text{ is increasing in } t_1 \geq 0.$$

Definition 4.3. *Consider non-negative random variables T_1 and T_2 as defined in Definition 2.1. We say that T_1 is smaller than T_2 in the Lorenz order (denoted $T_1 \leq_{Lorenz} T_2$) if*

$$\frac{1}{E\left(T_1\right)} \int_0^{F^{-1}(u)} t_1 dF(t_1) \geq \frac{1}{E\left(T_2\right)} \int_0^{G^{-1}(u)} t_1 dG(t_1),\ \forall u \in (0,1].$$

Note that the Lorenz order $T_1 \leq_{Lorenz} T_2$ means that the Lorenz curve of T_1 always lies above the Lorenz curve of T_2, and therefore the Gini index of T_1 is always smaller than the Gini index of T_2.

The following result from Shaked and Shanthikumar (2002) shows the relationship existing among the previous orders:

Theorem 2. *[Shaked and Shanthikumar, 2002] Let T_1 and T_2 be two non-negative random variables. Then*

a) $T_1 \leq_* T_2 \Leftrightarrow log(T_1) \leq_{disp} log(T_2)$ (p.214)

b) $T_1 \leq_* T_2 \Rightarrow T_1 \leq_{Lorenz} T_2$ (p.223)

Therefore, Theorem 2 reveals that star order is equivalent to the dispersive order between the log transformation of the random variables, and that Lorenz order (and hence Gini order) is implied from the former two orderings.

We can now prove Theorem 1 above.

Proof of Theorem 1. Without loss of generality in the proof we take $b(x)$ non-decreasing.
Let $b_1 = b(x_1)$, $b_2 = b(x_2)$, $u_1 = u(x_1)$, $u_2 = u(x_2)$, with $x_1 < x_2$. Therefore, $b_1 \leq b_2$. Let $Y_1 = log(T_1)$ and $Y_2 = log(T_2)$ and $\beta = F_{Y_1}(y_1) = 1 - S_0\left(\frac{y_1 - u_1}{b_1}\right)$ and $\beta = F_{Y_2}(y_2) = 1 - S_0\left(\frac{y_2 - u_2}{b_2}\right)$.
The inverse functions are $F_{Y_1}^{-1}(\beta) = u_1 + b_1 S_0^{-1}(1-\beta)$ and $F_{Y_2}^{-1}(\beta) = u_2 + b_2 S_0^{-1}(1-\beta)$.
Therefore,

$$F_{Y_1}^{-1}(\beta) - F_{Y_1}^{-1}(\alpha) = u_1 + b_1 S_0^{-1}(1-\beta) - u_1 - b_1 S_0^{-1}(1-\alpha) = b_1(S_0^{-1}(1-\beta) - S_0^{-1}(1-\alpha)),$$

for $0 \leq \alpha \leq \beta \leq 1$. Assuming that S_0 is a strictly decreasing function and since $1 - \beta < 1 - \alpha$, then $S_0^{-1}(1-\beta) - S_0^{-1}(1-\alpha) > 0$.

Condition $F_{Y_1}^{-1}(\beta) - F_{Y_1}^{-1}(\alpha) \leq F_{Y_2}^{-1}(\beta) - F_{Y_2}^{-1}(\alpha)$ is equivalent to $b_1(S_0^{-1}(1-\beta) - S_0^{-1}(1-\alpha)) \leq b_2(S_0^{-1}(1-\beta) - S_0^{-1}(1-\alpha))$, which reduces to $b_1 \leq b_2$. Therefore, dispersion ordering as in Definition 2.1.($Y_1 \leq_{disp} Y_2$) is satisfied, as is, from Theorem 2, Lorenz ordering as in Definition 2.3. ($T_1 \leq_{Lorenz} T_2$); see Section 2. This implies that $G(x_1) \leq G(x_2)$, given $b(x)$ non-decreasing as a function of x. □

Theorem 1 shows that under the class of models as in (5), the Lorenz ordering, and hence the Gini ordering, are always satisfied. Specifically, the Gini index always decreases as the cohort year of birth increases, if the shape parameter $b(x)$ increases; it decreases, otherwise.

We now examine the behavior of the survival distributions of two random variables satisfying the Gini ordering.

Corollary 1. *Under the model in Theorem 1 with $u(x) = u$, for any $u \in \mathbb{R}$ constant, and with $b(x)$ non-decreasing (non-increasing), the Gini index $G(x)$ is a non-decreasing (non-increasing) function of x, and:*

(i) for $t > e^u$, $S(t;x)$ is a non-decreasing (non-increasing) function of x;

(ii) for $t < e^u$, $S(t;x)$ is a non-increasing (non-decreasing) function of x; and

(iii) for $t = e^u$, $S(t;x)$ is constant with respect to x.

Proof. Without loss of generality in the proof we take $b(x)$ non-decreasing.

(i) If $t > e^u$, hence $\log(t) > u$ and assuming $x_1 < x_2$ then $S(t;x_1) = S_0\left(\frac{\log(t)-u}{b(x_1)}\right) \leq S_0\left(\frac{\log(t)-u}{b(x_2)}\right) = S(t;x_2)$;

(ii) If $t < e^u$, hence $\log(t) < u$ and assuming $x_1 < x_2$ then $S(t;x_1) = S_0\left(\frac{\log(t)-u}{b(x_1)}\right) \geq S_0\left(\frac{\log(t)-u}{b(x_2)}\right) = S(t;x_2)$;

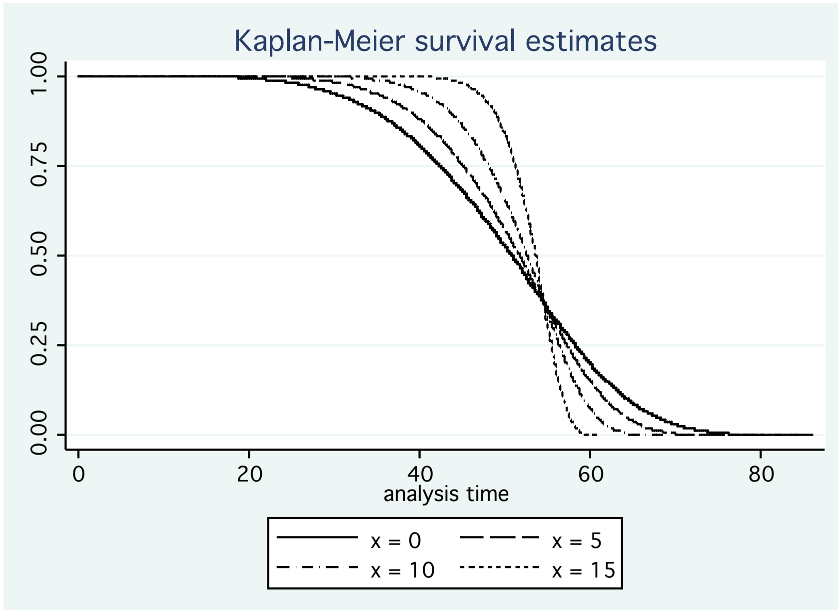
(iii) If $t = e^u$, hence $\log(t) = u$, $S(t;x) = S_0\left(\frac{0}{b(x)}\right)$ is constant with respect to x.

□

Corollary 1 reveals that survival functions will cross when Gini (or Lorenz) ordering is satisfied. A graphical illustration is provided in Figure 4.

We now move on studying how life expectancy behaves under the log-scale location model in (5). The following result provides a clear expression of life expectancy as a function of the year of birth x.

Figure 4: Illustration of Corollary 1



Source: Simulated data from a Weibull distribution (sample of size $n = 10,000$), with $b(x) = 0.2 - x/100$ and $u = -4$.

Corollary 2. *Consider a random variable T following a log-scale-location distribution $S(t; x)$ as in (5). Then the expected value of T, as a function of the covariate x, satisfies the following condition:*

$$\mu(x) = E(T; x) = e^{u(x)} E\left(e^{R \cdot b(x)}\right), \tag{7}$$

where R is a standardized random variable following the distribution $S_0(r)$ as in (5).

Proof.

$$\begin{aligned}
\mu(x) &= E(T; x) = \int_{\mathbb{R}_+} S(t; x)dt = \int_{\mathbb{R}_+} S_0\left(\frac{log(t) - u(x)}{b(x)}\right) dt \\
&= b(x) \int_{\mathbb{R}} S_0(r) e^{u(x) + r \cdot b(x)} dr \quad \left(\text{change of variable: } r = \frac{log(t) - u(x)}{b(x)}\right) \\
&= e^{u(x)} b(x) \int_{\mathbb{R}} S_0(r) e^{r \cdot b(x)} dr \\
&= e^{u(x)} b(x) \left[\frac{e^{r \cdot b(x)}}{b(x)} S_0(r) \Big|_{-\infty}^{\infty} + \int_{\mathbb{R}} \frac{e^{r \cdot b(x)}}{b(x)} f_0(r) dr \right] \ (\text{integrating by parts}).
\end{aligned}$$

We now examine the convergence of the first addend:

$$\lim_{r \to \infty} e^{rb(x)} S_0(r) = \lim_{r \to \infty} e^{rb(x)} e^{-e^r} = \lim_{r \to \infty} e^{rb(x) - e^r} = 0.$$

Therefore,

$$\mu(x) = e^{u(x)} b(x) \int_{\mathbb{R}} \frac{e^{rb(x)}}{b(x)} f_0(r) dr = e^{u(x)} E\left(e^{R \cdot b(x)}\right).$$

$\square$

Note that Corollary 2 implies that the life expectancy $\mu(x)$ corresponds to a linear transformation of the moment generating function of the standardized random variable $R = (log(T) - u(x))/b(x)$ with the baseline distribution $S_0(r)$ of the log-scale-location family (5), evaluated at $b(x)$.

5 Some special cases

We now focus on some special cases of the log-scale-location family introduced in (5).

5.1 Weibull baseline hazard function

We first assume that lifetime T follows a Weibull distribution, with p.d.f. written in the form

$$f(t; \alpha, \beta) = \frac{\beta}{\alpha} \left(\frac{t}{\alpha}\right)^{\beta - 1} \cdot \exp\left(-(t/\alpha)^{\beta}\right),\ t \geq 0$$

where $\alpha > 0$ and $\beta > 0$ are the scale and the shape parameter, respectively. The expected value of a Weibull distribution is $E(T) = \alpha \cdot \Gamma(1 + \frac{1}{\beta})$, with $\Gamma(\cdot)$ being the Gamma function. See Lawless (2003, page 218).

It can be shown that the random variable $Y = \log(T)$ follows the extreme value distribution and belongs to the location-scale family of distributions with p.d.f.

$$f(y; u, b) = \frac{1}{b} e^{(y-u)/b} \exp\left(-e^{(y-u)/b}\right),\ -\infty < y < \infty,$$

where $u = \log(\alpha)$ and $b = \beta^{-1}$. See Lawless (2003, page 218). The moment generating function of Y is $M_Y(\theta) = \Gamma(1 + \theta)$; see Lawless (2003, page 21).

Here we allow the parameters u and b to depend on a covariate x; therefore, we consider the following p.d.f. for Y

$$f(y; u(x), b(x)) = \frac{1}{b(x)} e^{(y-u(x))/b(x)} \exp\left(-e^{(y-u(x))/b(x)}\right),\ -\infty < y < \infty, \tag{8}$$

If $Y = log(T)$ has an extreme value distribution as in (8), then the standardized random variable $R = \frac{log(T) - u(x)}{b(x)}$ has survival distribution $S_0(r) = e^{-e^r}$ and p.d.f. $f_0(r) = e^{-e^r} \cdot e^r$ for $r \in \mathbb{R}$.

Assuming $u(x) \geq 0$ w.l.o.g., then, from Corollary 2 we have

$$E(T; x) = \mu(x) = e^{u(x)} \cdot \Gamma(1 + b(x)), \tag{9}$$

where $\Gamma()$ is the Gamma function.

Therefore, under the Weibull/extreme value model $G(x)$ is always decreasing in x, but this is not necessarily true for life expectancy $\mu(x)$, whose behavior depends on the assumptions around the functional form of $b(x)$ and of $u(x)$ as in (9).

In particular, if $u(x) = u$ is constant, then: *(i)* for any decreasing function $b(x)$ such that $0 < b(x) < 0.4616$, $G(x)$ is a decreasing function and $\mu(x)$ is a increasing function of x, while *(ii)* for any decreasing function $b(x)$ such that $b(x) \geq 0.4616$, both $G(x)$ and $\mu(x)$ are decreasing functions of x.

5.2 Log-normal baseline hazard function

If T follows a log-normal distribution with parameters m_0 and σ^2, then $Y = log(T)$ has a normal distribution with parameters m_0 and σ^2. Under this setting, for any non-increasing function $b(x)$ and any non-increasing function $u(x)$, both $G(x)$ and $\mu(x)$ are non-increasing functions of x.

In particular, from Corollary 2 we have that $\mu(x) = e^{u(x)} E(e^{R \cdot b(x)})$, where R has a standard normal distribution. Therefore,

$$\mu(x) = e^{u(x) + \frac{1}{2} b(x)^2},$$

which is an increasing function of $b(x)$ and $u(x)$, and therefore a non-increasing function of x.

5.3 Log-logistic baseline hazard function

Consider the lifetime T having a log-logistic distribution with p.d.f. written in the form

$$f(t; \alpha, \beta) = \frac{(\beta/\alpha)(t/\alpha)^{\beta-1}}{(1 + (t/\alpha)^{\beta})^2}, \ t \geq 0$$

with parameters $\alpha, \beta > 0$. See Lawless (2003, page 23).

It can be shown that if T follows a log-logistic distribution, then the transformation $Y = \log(T)$ has logistic distribution with p.d.f.

$$f(y; u, b) = \frac{b^{-1} \exp(-(y - u)/b)}{(1 + \exp(-(y - u)/b))^2}, \ -\infty < y < \infty,$$

where $u = \log(\alpha)$ and $b = \beta^{-1}$. See Lawless (2003, page 23).

The standardized random variable $R = (Y - u)/b$ has p.d.f. $f_0(r) = \frac{e^{-r}}{(1+e^{-r})^2}$ for $r \in \mathbb{R}$ and moment generating function $M_R(\theta) = Be(1-\theta; 1+\theta)$ for $\theta \in (-1, 1)$, where $Be(\cdot;\cdot)$ is the Beta function; see Johnson-Kotz-Balakrishnan (1995, page 113).

If $Y = log(T)$ has a logistic distribution with parameters $b = b(x)$ and $u = u(x)$, and assuming $u(x) \geq 0$ w.l.o.g., then $E(T|x) = \mu(x) = e^{u(x)} \cdot Be(1 - b(x); 1 + b(x))$.

Since $Be(1 - b(x); 1 + b(x))$ is a decreasing function of $b(x)$ for negative values of $b(x)$ that are out of the range of $b(x)$, and it is an increasing function of $b(x)$ for positive values of $b(x)$, we conclude that for any decreasing and positive function $b(x)$ and any non-increasing and positive function $u(x)$, both $G(x)$ and $\mu(x)$ are decreasing functions of x.

Summing up, we have seen that the log-scale-location family is a flexible model that allows for modeling different scenarios. On the one hand, some of the distributions belonging to that family allow for modeling a situation where, as cohort year of birth x increases, lifetimes become less concentrated (i.e., Gini index decreases) while life expectancy $\mu(x)$ increases. This is the case of Weibull baseline distribution under suitable choices of the range of function $b(x)$. On the other hand, other members of the log-scale-location family allow for modeling situations where, as cohort year of birth x increases, lifetimes become more concentrated (i.e., Gini index increases) and life expectancy $\mu(x)$ increases. This happens under the models with log-normal or the log-logistic baseline distributions. According to empirical evidence, one may observe that over time and/or across countries, the log-scale-location family provides a semi-parametric method for estimating the phenomenon under observation.

6 Conclusions

The class of models that we have focused on allows the study of some of the relationships that exist among survival distributions and concentration. Longevity measures may also be studied within such a promising framework. Indeed, from the second half of the 20th century, we have seen that the general increase in longevity around the globe has been accompanied by a broad decline in concentration of survival times.

Further work will focus on the extension of the results to other families of distributions and on the estimation of model parameters. With respect to the second direction, we will consider two different types of data: (i) population datasets, where we will work with people in different birth cohorts, and (ii) subject-specific datasets derived from pension funds, where we will work in the presence of right-censoring.

References

Atkinson, T. (1970). On the Measurement of Inequality. *Journal of Economic Theory*, 2, 244-263.

Baudisch, A. (2011). The pace and shape of ageing. *Methods in Ecology and Evolution*, 2(4), 375-382.

Bonetti, M., Gigliarano, C. and Muliere, P. (2009). The Gini concentration test for survival data. *Lifetime Data Analysis*, 15, 493-518.

Bongaarts, J. and Feeney, G. (2002). How long do we live? *Population and Development Review*, 28(1), 13-29.

Canudas-Romo, V. (2008). The modal age at death and the shifting mortality hypothesis. *Demographic Research*, 19(30), 1179-1204.

Cox, D. R. (1972). Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society, Series B*, 34, 187 - 220.

Fries, J. F. (1980). Aging, natural death, and the compression of morbidity. *New England Journal of Medicine*, 303(3), 130-135.

Gastwirth, J.L. (1971). A General Definition of the Lorenz Curve. *Econometrica*, 31, 1037-1039.

Gini, C. (1912). Variabilità e Mutabilità. Contributo Allo Studio Delle Distribuzioni e Relazioni Statistiche. *Studi Economico-Giuridici dell'Università di Cagliari*. III.

—— (1914), Sulla Misura Della Concentrazione e Della Variabilità dei Caratteri. *Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti*, LXXIII(part 2), 1203-1248.

Hanada, K. (1983). A Formula of Gini's Concentration Ratio and its Applications to Life Tables. *Journal of Japanese Statistical Society*, 19, 293-325.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1995), *Continuous univariate distributions*, Volume 2, Wiley & Sons ed.

Kakwani, N. C. (1980). *Income Inequality and Poverty: Methods of Estimation and Policy Applications*. Oxford: Oxford University Press.

Kannisto, V. (2000). Measuring the compression of mortality. *Demographic Research*, 3(6), 1-24.

Kendall, M. and Stuart, A. (1977). *The Advanced Theory of Statistics*, volume I. New York: Mac Millan Publishing.

Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*. New York: Wiley.

Lorenz, M. O. (1905). Methods of Measuring the Concentration of Wealth. *Publications of the American Statistical Association*. 9(70), 209-219.

Michetti, B. and Dall´Aglio, G. (1957). La Differenza Semplice Media. *Statistica*, 7(2), 159-255.

Muliere, P. and Scarsini, M. (1989). A Note on Stochastic Dominance and Inequality Measures. *Journal of Economic Theory*, 49, 314-323.

Nygard, F. and Sandröm, A. (1981). *Measuring Income Inequality*. Stockholm: Almqvist and Wilsell International.

Ostasiewicz K. and Mazurek E. (2013). Comparison of the Gini and Zenga indexes using some theoretical income distributions abstract. *Operations Research and Decisions*, 1, 37-62.

Pietra, G. (1915). Delle Relazioni tra gli Indici di Variabilità, I, II. *Atti del Reale Istituto Veneto di Scienze, Letter ed Arti*, LXXIV(II), 775–804.

Shaked, M. and Shanthikumar, J.G. (1994). *Stochastic Orders and Their Applications*. Boston: Academic Press, Inc..

Shkolnikov, V. M., Andreev, E. E., and Begun, A. Z. (2003). Gini Coefficient as a Life Table Function: Computation From Discrete Data, Decomposition of Differences and Empirical Examples. *Demographic Research*, 8, 305-358.

Wilmoth, J.R. and Horiuchi, S. (1999). Rectangularization revisited: variability in age at death within human populations, *Demography*, 36(4), 475-95.