



Max-Planck-Institut für demografische Forschung
Max Planck Institute for Demographic Research
Konrad-Zuse-Strasse 1 · D-18057 Rostock · GERMANY
Tel +49 (0) 3 81 20 81 - 0; Fax +49 (0) 3 81 20 81 - 202;
<http://www.demogr.mpg.de>

MPIDR TECHNICAL REPORT 2010-001
MAY 2010 (REVISED MARCH 2023)

**Spreadsheet for calculation
of life-table dispersion measures**

Vladimir M. Shkolnikov (shkolnikov@demogr.mpg.de)
Evgeny M. Andreev

For additional material see www.demogr.mpg.de/tr/

© Copyright is held by the authors.

Technical reports of the Max Planck Institute for Demographic Research receive only limited review. Views or opinions expressed in technical reports are attributable to the authors and do not necessarily reflect those of the Institute.

Spreadsheet for calculation of life-table dispersion measures

by Vladimir M. Shkolnikov and Evgeny M. Andreev

Abstract

Measures for the estimation of disparity in respect to length of life within populations are increasingly used. They are calculated from the conventional life table and reflect greater or smaller amounts of inter-individual differences across the life-table cohort. We provide an Excel spreadsheet for building the Lorenz curve and the calculation of life-table dispersion measures. A complete life table is used as input data. The output includes seven measures: the Gini coefficient (concentration index), the average inter-individual difference in age at death (absolute concentration index), the inter-quartile range, the standard deviation, the coefficient of variation, the e-dagger, and Keyfitz's entropy.

Keywords: life table dispersion, inequality, age at death, Gini concentration index, inter-quartile range, standard deviation, lifetime losses, e-dagger, mortality compression

Background

In recent decades, demographers have become increasingly interested in analyses of measures expressing inter-individual disparities with respect to length of life (or age at death) within populations. These measures can be computed from the conventional life table. They show to what extent individuals from the life-table cohort differ from each other in their prospects for longevity. The more dispersed the life-table deaths are along the age axis, the higher the values of the measures.

Different studies have focused on different substantive aspects of these length-of-life disparities. Wilmoth and Horiuchi (1999) considered the historical decline of the life-table dispersion as a part of the longevity revolution, which graphically expresses itself as a rectangularization of the survival curve.

Anand et al. (2001); Shkolnikov, Andreev, and Begun (2003); and Shkolnikov et al. (2009) considered life-table dispersion as a reflection of inter-individual inequality in the face of death. The dispersion measurement captures the total amount of inter-individual differences independently of specific socioeconomic or demographic groupings. This approach differs considerably from numerous other studies on socioeconomic inequality in mortality that are focusing on inter-group differences.

Vaupel and Canudas Romo (2003) showed that the magnitude of longevity gains produced by age-specific mortality decreases depends not only on the average rate of these improvements, but also on the age-at-death disparity. More recently, Zhang and Vaupel (2009) considered the properties of total lifetime losses, which include components connected with mortality compression and mortality extension.

Finally, Edwards and Tuljapurkar (2005) noticed that a greater or smaller degree of lifespan dispersion means a greater or smaller degree of uncertainty in individuals' expectations about

their length of life. That in turn can influence people's behaviors with respect to savings, pensions, medical insurance, and other important aspects of financial planning.

So far, the calculation of life-table dispersion measures has not become a conventional procedure that is widely known among potentially interested students and researchers. Therefore, we decided to provide a spreadsheet intended to serve as a simple tool that can be used directly for computations, and also for practical instruction on how to construct the most useful measures of life-table dispersion.

Dispersion-Life-Table-complete.xls

The spreadsheet consists of five worksheets. Each worksheet is related to a specific type of measurement, and allows for the calculation of relative and/or absolute life-table dispersion measures from a standard complete life table. As an example, we use life tables extracted from the Human Mortality Database (HMD) that are available at www.mortality.org or www.humanmortality.de.

The first worksheet, "Lorenz curve-Gini-AID," builds the Lorenz curve and calculates from it two measures: the life-table Gini coefficient (also called the concentration index or the Gini index) as a relative measure, and the life-table average inter-individual difference in age at death, AID (also called the absolute concentration index) as an absolute measure. Both measures are computed here for the entire range of ages. The second worksheet, "Gini-AID-Hanada," allows us to calculate the same two measures at every age x . G_x and AID_x measure the relative and the absolute inter-individual inequality in age at death over age x and older ages. In this worksheet, calculations are based on formula by Hanada (1983).

The third worksheet, "IQR," allows us to compute the inter-quartile range IQR for the distribution of ages at death over age x and older ages. For computation of IQR for the range of ages beginning from age x , the user should specify the lower age limit x in cell L1. For example, $L1=0$ corresponds to IQR_0 calculated over the entire range of ages. $L1=10$ corresponds to calculation of IQR_{10} over ages 10+.

The fourth worksheet, "STD-CoefVar," allows us to compute the standard deviation and the coefficient of variation of ages at death over age x , and older ages as absolute and relative measures, respectively.

The fifth worksheet, "e-dagger-H," allows us to compute the lifetime losses e-dagger and the life table entropy H_x over age x and older ages.

When a life table is entered in cells A3:J113 on worksheet "Lorenz curve-Gini-AID," the results of the calculations will appear instantly in this and the other four worksheets.

Mathematical expressions

Lorenz curve, Gini coefficient and average inter-individual difference

The Lorenz curve (Figure 1) expresses the cumulative share Φ_x in years of life as a function of the share in the life-table population F_x .

$$F_x = \frac{\sum_{t=0}^x d_t}{\sum_{t=0}^{\omega} d_t} = 1 - \frac{l_{x+1}}{l_0} \quad (1.1)$$

$$\Phi_x = \frac{\sum_{t=0}^x d_t \cdot \bar{t}}{\sum_{t=0}^{\omega} d_t \cdot \bar{t}} \quad (1.2)$$

Both shares vary from 0 (at age 0) to 1 at the last (maximum) age ω .

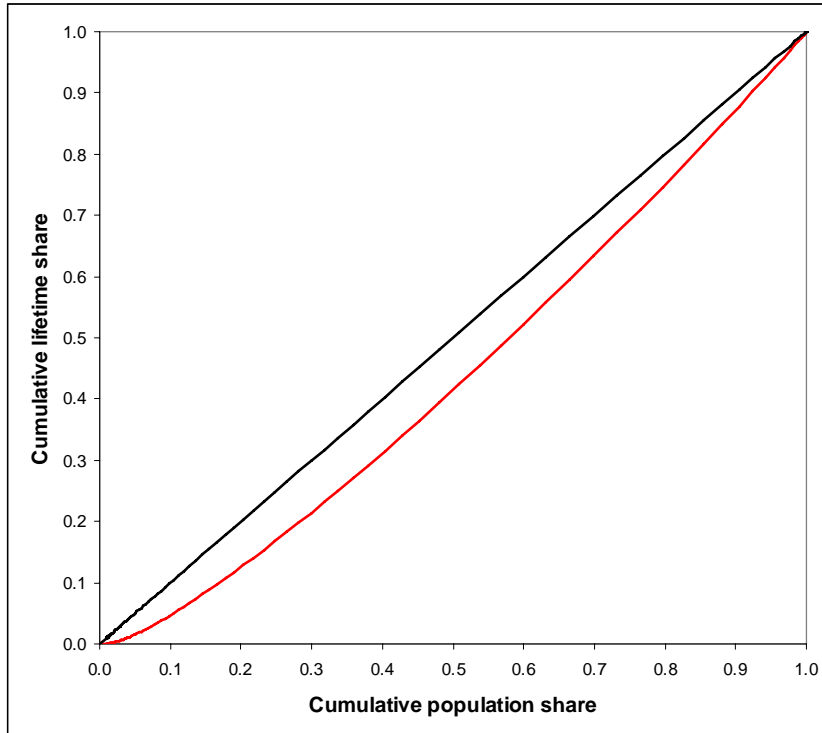


Figure 1. Lorenz curve based on data from the HMD life table for the male population of France in 1990.

The greater the deviation of the Lorenz curve from the diagonal, the greater the degree of inequality in age at death. In the situation of perfect equality (when everyone has the same length of life), the Lorenz curve is exactly the diagonal. Geometrically, the Gini coefficient is defined as a half of the area between the Lorenz curve and the diagonal. This definition leads to the following formula for the Gini coefficient:

$$G_x = 1 - \sum_{x=0}^{\omega-1} (F_{x+1} - F_x)(\Phi_{x+1} + \Phi_x) \quad (1.3)$$

Correspondingly, the average inter-individual difference *AID* is expressed as

$$AID_x = G_x \cdot e_x \quad (1.4)$$

While *AID* is measured in years of length of life, *G* is expressed as a share of the mean length of life.

Formulae (1.1)-(1.4) are used in the worksheet “Lorenz curve-Gini-AID.”

A compact expression (2.1) for the life table Gini by Hanada (1983) in its discrete form (Shkolnikov, Andreev, Begun, 2003) is especially useful for practical calculations.

$$G_x = 1 - \frac{1}{e_x(l_x)^2} \int_x^\infty [l(t)]^2 dt \cong 1 - \frac{1}{e_x(l_x)^2} \sum_{t=x}^{\omega-1} [(l_{t+1})^2 + a_x((l_t)^2 - (l_{t+1})^2)] \quad (2.1)$$

This formula is used in the worksheet “Gini-AID-Hanada.”

Further insights on *G*, *AID*, and related measures and their mathematical properties, can be found in papers by Le Grand and Rabin (1986), Anand et al. (2001), and Shkolnikov, Andreev, and Begun (2003).

Inter-quartile range

IQR is a simple and intuitively transparent measure of absolute disparity in length of life. It was promoted by Wilmoth and Horiuchi (1999). Although this measure has certain weaknesses (Shkolnikov, Andreev, Begun, 2003), it was also shown that variations in *IQR* across time are highly correlated with corresponding variations in more sophisticated measures (Wilmoth and Horiuchi, 1999).

$$IQR = Q_{25} - Q_{75}, \text{ where } Q_{25} = x, \text{ such that } l_x = 0.25 \\ \text{and } Q_{75} = x, \text{ such that } l_x = 0.75 \quad (3.1)$$

Calculation of the *IQR*_{*x*} is realized in the worksheet “*IQR*.”

Detailed mathematical background and empirical analyses of *IQR* can be found in Wilmoth and Horiuchi (1999).

Standard deviation

This is a conventional dispersion measure in mathematical statistics. When applied to the life table, *STD* reflects mean of the weighted deviations of ages at death from the mean age at death:

$$STD_x = \sqrt{\frac{\sum_{y=x}^{\omega} d_y (\bar{y}_y)^2}{l_x} - (x + e_x)^2} \quad (4.1)$$

Note that, due to the insufficient precision of values of life-table functions, formula (4.1) sometimes does not work at high ages *x* (90, 95, or 100) because quantity under the square root becomes negative. This is not very important because in most cases *STD* is being calculated for a wide ranges of ages, beginning from ages 0, 10, or 30.

The coefficient of variation represents the absolute variant of the same measurement

$$CoefVar_x = \frac{STD_x}{x + e_x} \quad (4.2)$$

STD has some useful properties. For example, for a population consisting of several sub-groups, STD can be decomposed into the between-group and the within-group components.

A study by Edwards and Tuljapurkar (2005) provides a detailed explanation of STD as a measure, and also shows its application to analyses of the life-table dispersion in the US.

The calculation of the STD and the CoefVar from a complete life table is realized in the worksheet “STD.”

Lifetime losses

This type of measurement has a long tradition in demography. It was first obtained by N.Keyfitz (1977) in his derivation of a formal relationship between a small change in age-specific mortality rates and its effect on the life expectancy. More recently, e^\dagger (e-dagger) came up in a non-conventional decomposition of temporal change in life expectancy by Canudas, Romo, and Vaupel (2003). Apart from being a measure of the life-table dispersion, e-dagger has important implications for public health. Its values quantify the average life expectancy losses due to death. It generally follows Keyfitz’s idea that “everybody dies prematurely,” since every death “deprives the person involved of the reminder of his expectation of life” (Keyfitz 1977:61-68).

In the worksheet “e-dagger,” we use the following discrete formula for e^\dagger :

$$e_x^\dagger = \frac{1}{l_x} \cdot \sum_{y=x}^{\omega-1} [d_y (e_{y+1} + 1 - a_y)] + \frac{l_\omega}{2l_x} e_\omega \quad (5.1)$$

The relative variant of this measure is the classic life-table entropy H by Nathan Keyfitz (1977)

$$H_x = - \frac{\int_x^\infty l(\zeta) \cdot \ln(l(\zeta)) d\zeta}{e(x)} \cong \frac{e_x^\dagger}{e_x} \quad (5.2)$$

Calculations corresponding to formulae (5.1) and (5.2) are realized using the worksheet “e-dagger-H.”

Detailed information about the mathematics of H and e^\dagger and their applications can be found in Keyfitz (1977); Goldman and Lord (1986); Canudas Romo and Vaupel (2003); Zhen Zhang and Vaupel (2009); and Shkolnikov et al. (2010).

References

Edwards, R.D. and Tuljapurkar, S. (2005). Inequality in life spans and a new perspective on mortality convergence across industrialized countries. *Population and Development Review* 31(4): 645–674.

Goldman, N. and Lord, G. (1986). A new look at entropy and the life table. *Demography* 23(2): 275–282.

Keyfitz, N. (1977). *Applied Mathematical Demography*. New York: John Wiley, 1st ed.

Shkolnikov, V., Andreev, 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(11): 305–358.

- Vaupel, J.W. (1986). How change in age-specific mortality affects life expectancy. *Population Studies* 40(1): 147–157.
- Vaupel, J.W. and Canudas-Romo, V. (2003). Decomposing changes in life expectancy: A bouquet of formulas in honor of Nathan Keyfitz's 90th birthday. *Demography* 40(2): 201–216.
- Edwards, R.D. and Tuljapurkar, S. (2005). Inequality in life spans and a new perspective on mortality convergence across industrialized countries. *Population and Development Review* 31(4): 645–674.
- Wilmoth, J.R. and Horiuchi, S. (1999). Rectangularization revisited: Variability of age at death with human populations. *Demography* 36(4): 475–495.
- Shkolnikov, V. M.; Andreev, E. M.; Zhang, Z.; Oeppen, J.; Vaupel, J. W. (2009). Losses of expected lifetime in the US and other developed countries: methods and empirical analyses. Rostock, MPIDR Working Paper WP-2009-042.
- Anand, S., F. Diderichsen, T. Evans, V. Shkolnikov, and M. Wirth. (2001). "Measuring Disparities in Health: Methods and Indicators." Pp. 49-67 in *Challenging Inequities in Health: From Ethics to Action*, edited by T. Evans, M. Whitehead, F. Diderichsen, A. Bhuiya. New York: Oxford University Press.
- Zhang, Z.; Vaupel, J. W. (2009). The age separating early deaths from late deaths. *Demographic Research*, 20:29, 721-730.