# Tutorial: Temporal improvements in elephant longevity challenge the notion of compromised survival in zoos

Fiona Sach, Hannah Jenkins, Dalia A. Conde, Johanna Staerk,
Morgane Tidière, Malcolm Fitzpatrick, Fernando Colchero

## Brief introduction

Here we present code to visualize the life tables presented on the manuscript, to calculate life expectancy and lifespan equality from Siler mortality parameters, and to calculate Kullback-Leibler discrepancies from un- and bi-variate distributions.

The .zip file includes three directories, namely 01code/, 02data, and 03tutorial. The 01code directory includes two R script files, ElephantLongevityFunctions.R and ElephantLongevityCode.R. This last is the code you should work with, unless you choose to change any of the analysis functions.

## Preparing code

First you will need to point the R code in ElephantLongevityCode.R to the path where you have saved the directories:

```r
# Set working directory:
setwd("path to directory...")

# Source functions:
source("01code/ElephantLongevityFunctions.R")
```

The second line calls the analysis functions while also loading the relevant R libraries. The following lines load the different data tables used for this tutorial:

```r
# Load life tables:
lifeTabs <- read.csv(file = "02data/ElephantLifeTables.csv",
                     header = TRUE, stringsAsFactors = FALSE)

# Load Siler mortality parameters:
thetaMat <- read.csv(file = "02data/ElephantSilerParams.csv",
                     header = TRUE, stringsAsFactors = FALSE)

# Load life expectancy-lifespan equality:
leMat <- read.csv(file = "02data/ElephantLifeExpLifespEq.csv",
                  header = TRUE, stringsAsFactors = FALSE)
```

## Plotting life tables

All the life tables used in the study are stored in the lifeTabs object. As an example, we show how to plot the life table survival for female African elephants (*Loxodonta africana*):

```r
# ==================================== #
# ==== PLOT LIFE TABLE SURVIVAL: ==== #
# ==================================== #
# Select species:
```

```
sps <- "Loxodonta africana"

# Select sex:
sx <- "Female"
```
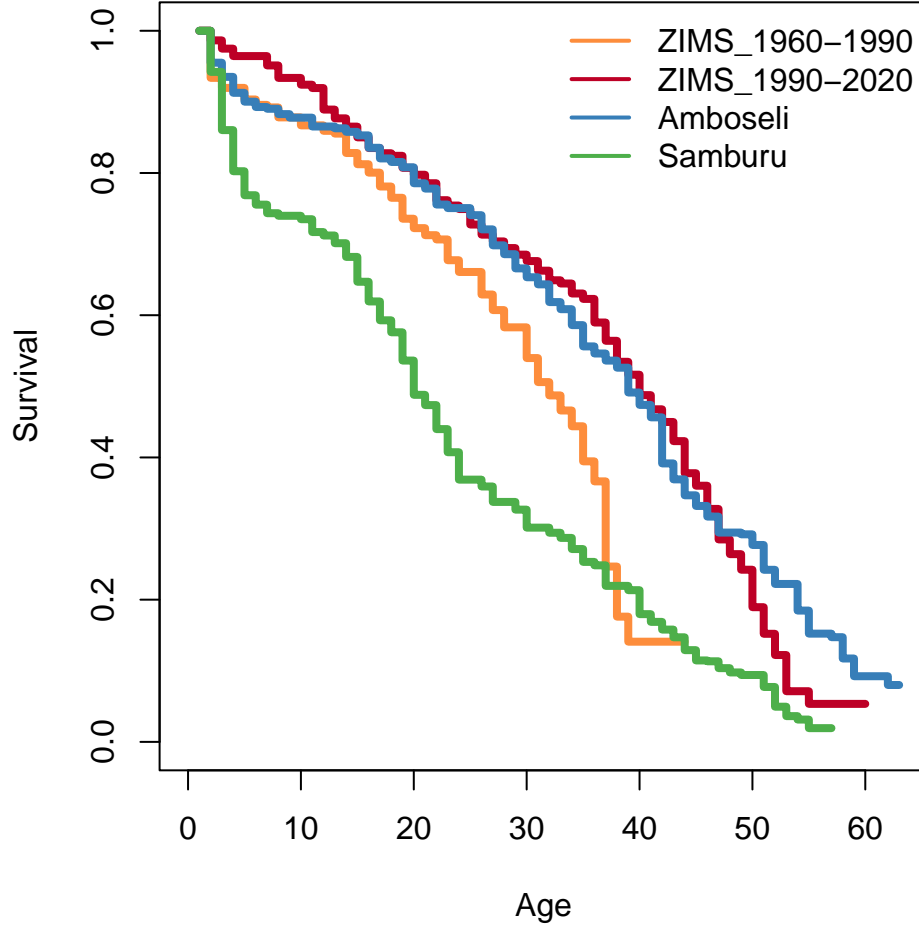
The code that follows produces the plot below:



Figure 1: Example of life table survival plot for female African elephants starting at age 1.

## Calculate life expectancy and lifespan equality

### Theoretical background

Here we use parametric mortality functions to calculate life expectancy and lifespan equality. These two measures are summary statistics of the distribution of ages at death. Thus, let $X$ be a random variable for ages at death, with realizations $x \geq 0$. The age-specific hazard rate or mortality function is

$$\mu(x) = \lim_{\Delta x \to 0} \frac{\Pr[x < X < x + \Delta x | X > x]}{\Delta x}. \tag{1}$$

From the hazard rate or mortality function in Eq.~1, we calculate the cumulative hazard, given by

$$U(x) = \int_0^x \mu(t)dt, \tag{2}$$

from which the survival or cumulative survival function is

$$S(x) = e^{-U(x)}, \tag{3}$$

with complement given by the cumulative distribution function $F(x) = 1 - S(x)$ The probability density function of ages at death is given by

$$f(x) = \mu(x)S(x). \tag{4}$$

Life expectancy is, as the name shows, the expected age at death and, as an expectaction, which, from birth is given by

$$
\begin{aligned}
e_0 &= \int_0^\infty x f(x) dx \\
&= \int_0^\infty S(x) dx. \tag{5}
\end{aligned}
$$

Lifespan equality requires first calculating lifespan inequality as proposed by Demetrius (1978), given by

$$
\begin{aligned}
H &= -\frac{\int_0^\infty S(x) \ln[S(x)] dx}{e_0} \\
&= \frac{\int_0^\infty S(x) U(x) dx}{e_0}. \tag{6}
\end{aligned}
$$

In short, lifespan inequality provides a weighted average of the accumulation in the risk of death with age, weighted by the survival function. As deaths accumulate earlier, the higher the lifespan inequality. Lifespan equality is

$$\varepsilon_0 = -\ln(H) \tag{7}$$

and thus is inversely related to lifespan inequality, which implies that $\varepsilon_0$ measures the concentration of deaths at older ages (Colchero *et al.* 2021).

**Summary variables from Siler model**

Here, we use the Siler mortality function (Siler 1979), given by

$$\mu(x|\boldsymbol{\theta}) = \exp(a_0 - a_1 x) + c + \exp(b_0 + b_1 x), \tag{8}$$

where $\boldsymbol{\theta}^\top = [a_0, a_1, c, b_0, b_1]$ is a vector of Siler mortality parameters, where $a_0, b_0 \in \mathbb{R}$ and $a_1, c, b_1 \geq 0$.

The cumulative hazard of the Siler model is

$$U(x|\boldsymbol{\theta}) = \frac{e^{a_0}}{a_1}(1 - e^{-a_1 x}) + cx + \frac{e^{b_0}}{b_1}(e^{b_1 x} - 1). \tag{9}$$

Here we use a left-hand quadrature to find the integrals in Eqs.~5 and 6. Thus, in the R script section on Calculating life expectancy and lifespan equality, we show how to estimate both variables for female African elephants in ZIMS during the 1990-2020 period, but the code can be adapted to specify a different period or wild population, males or Asian elephants:

```
# ================================================ #
# ==== CALCULATE LIFE EXPECT.-LIFESP. EQUAL. ==== 
# ================================================ #
# Select species:
sps <- "Loxodonta africana"

# Select sex:
sx <- "Female"

# Select population:
pop <- "ZIMS_1990-2020"
```

The section of the code returns the following output to the console:

```
>
> =============================
> Species: Loxodonta africana
> Sex: Female
> - Life exp.  = 31.05
> - Lifesp eq. = 0.674
> =============================
```

## Calculating uni- and bi-variate Kullback-Leibler discrepancies

Kullback-Leibler (Kullback and Leibler 1951) discrepancies measure the information lost when approximating a given distribution $f_i$ by means of another distribution $f_j$. They are calculated as

$$K_{ij} = \int_{-\infty}^{\infty} f_j(x) \ln \left[ \frac{f_j(x)}{f_i(x)} \right] dx, \tag{10}$$

and are therefore a measure of entropy, whereby if $f_j(x) = f_i(x)$ for all $x \in \mathbb{R}$ then $K_{ij} = 0$, this is, both distributions are equal. The bivariate Kullback-Leibler is simply obtained as

$$K_{ij} = \int_{\Omega} \int_{\Lambda} f_j(x,y) \ln \left[ \frac{f_j(x,y)}{f_i(x,y)} \right] dxdy, \tag{11}$$

where $f_j(x,y)$ and $f_i(x,y)$ are two bivariate distributions, and where $X$ and $Y$ are random variables with sample spaces $\Lambda$ and $\Omega$, respectively.

Here, we use Kullback-Leibler discrepancies to estimate the differences in the posterior distributions of life expectancy and lifespan equality between populations of elephants of the same species and sex. We use a transformation proposed by McCulloch (1989) which simplifies interpretability by providing values between 0 and 1, where 0 implies perfect matching between distributions, and 1 that both distributions are entirely different.

For instance, to estimate the Kullback-Leibler discrepancies between the posterior densities of life expectancy of African elephants in both periods in ZIMS you can run the following code:

which produces the following output

```
>         KL         qKL
> 13.0628390  0.9999989
```

where the first term is the $K_{ij}$ and the second term is the transformation.

To calculate the bivariate value simply run:

```
# Calculate bivariate Kullback-Leibler discrepancy:
KLbiv <- CalcBivKL(m1 = le1$Mean, m2 = le2$Mean,
                cv1 = as.matrix(le1[, c("Cov1", "Cov2")]),
                cv2 = as.matrix(le2[, c("Cov1", "Cov2")]))
KLbiv
```

```
>       KL       qKL
> 32.77289   1.00000
```

## References

Colchero F. *et al.*, The long lives of primates and the 'invariant rate of ageing' hypothesis. Nature Communications. **12**, 3666 (2021).

Demetrius, L. Entropy and life table. Naturwissenschaften **65**, 435-436 (1978)

McCulloch, R. Local model influence. Journal Of The American Statistical Association **84**, 473-478 (1989)

Kullback S., R. A. Leibler, On Information and Sufficiency. The Annals of Mathematical Statistics. **22**, 79-86 (1951).

Siler, W. A competing-risk model for animal mortality. Ecology. **60**, 750-757 (1979).