Mathematical Hazards Models and Model Life Tables

Formal Demography

Stanford Summer Short Course

James Holland Jones, Instructor

August 12, 2005

Outline

1. Mathematical Hazards Models
   (a) Gompertz-Makeham
   (b) Siler
   (c) Heligman-Pollard
2. Relational Mortality Models
3. Model Life Tables
   (a) Coale-Demeny
   (b) INDEPTH

Basic Quantities in the Analysis of Mortality

Survival Function

$$S(x) = Pr(X > x)$$

for continuous X, strictly decreasing

Complement of the cumulative distribution function of $F(x)$

$$S(x) \equiv l(x) = 1 - F(x), \qquad F(x) = Pr(X \leq x)$$

Also the integral of the probability density function $f(x)$:

$$S(x) = Pr(X > x) = \int_x^\infty f(t)dt$$

Thus, given a survival function, we can calculate the probability density function

$$f(x) \equiv d(x)dt = -\frac{dS(x)}{dx}$$

and $f(x)\Delta x$ is the approximate probability that a death will occur at time x

More Basic Quantities: The Hazard Function

The Hazard is the demographic force of mortality (typically indicated $\mu(x)$)

It is defined as

$$h(x) \equiv \mu(x) = \lim_{\Delta x \to 0} \frac{Pr[x \leq X < x + \Delta x | X \geq x]}{\Delta x}$$

If X is a continuous random variable

$$h(x) = \frac{f(x)}{S(x)} = -d \log[S(x)]$$

The cumulative hazard is

$$H(x) = \int_0^x h(u) du = -\log[S(x)]$$

Thus, for continuous lifetimes

$$S(x) = \exp[-H(x)] = \exp\left[-\int_0^x h(u)du\right]$$

This is exactly as it should be, as it matches our lifetable definitions:

$$l(x) = \exp\left[-\int_0^x \mu(t)dt\right]$$

Gompertz Mortality

Gompertz (1825) suggested that a "law of geometric progression pervades" in mortality after a certain age

Gompertz mortality can be represented as

$$\mu(x) = \alpha e^{\beta x}$$

α is known as the baseline mortality, whereas β is the senescent component

Makeham (1860) extended the Gompertz model by adding a constant γ

Note that since the Gompertz model is for a mortality hazard, we can integrate it to give us the the survival function:

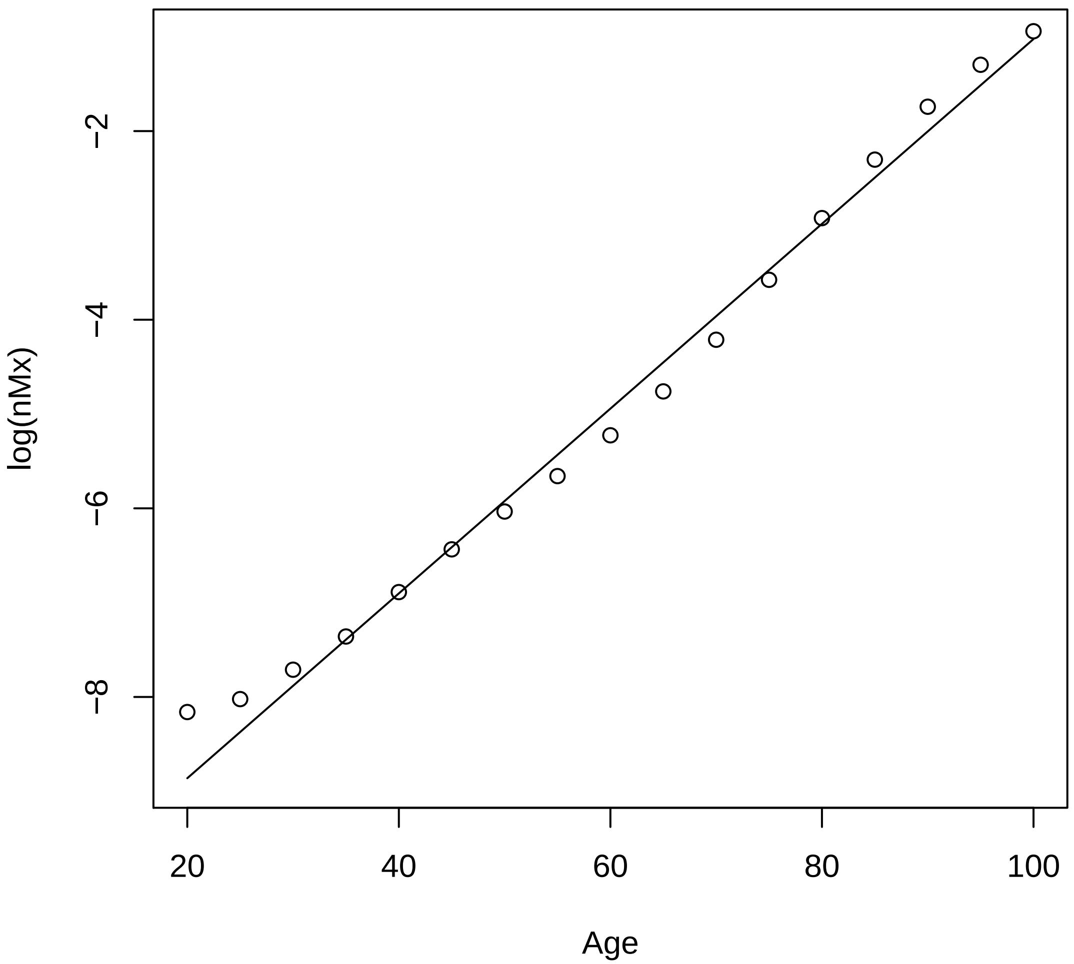
$$h(x) = \alpha e^{\beta x}, \qquad S(x) = \exp\left[\frac{\alpha}{\beta}\left(1 - e^{\beta x}\right)\right]$$

Also note that log-mortality is a linear function of age

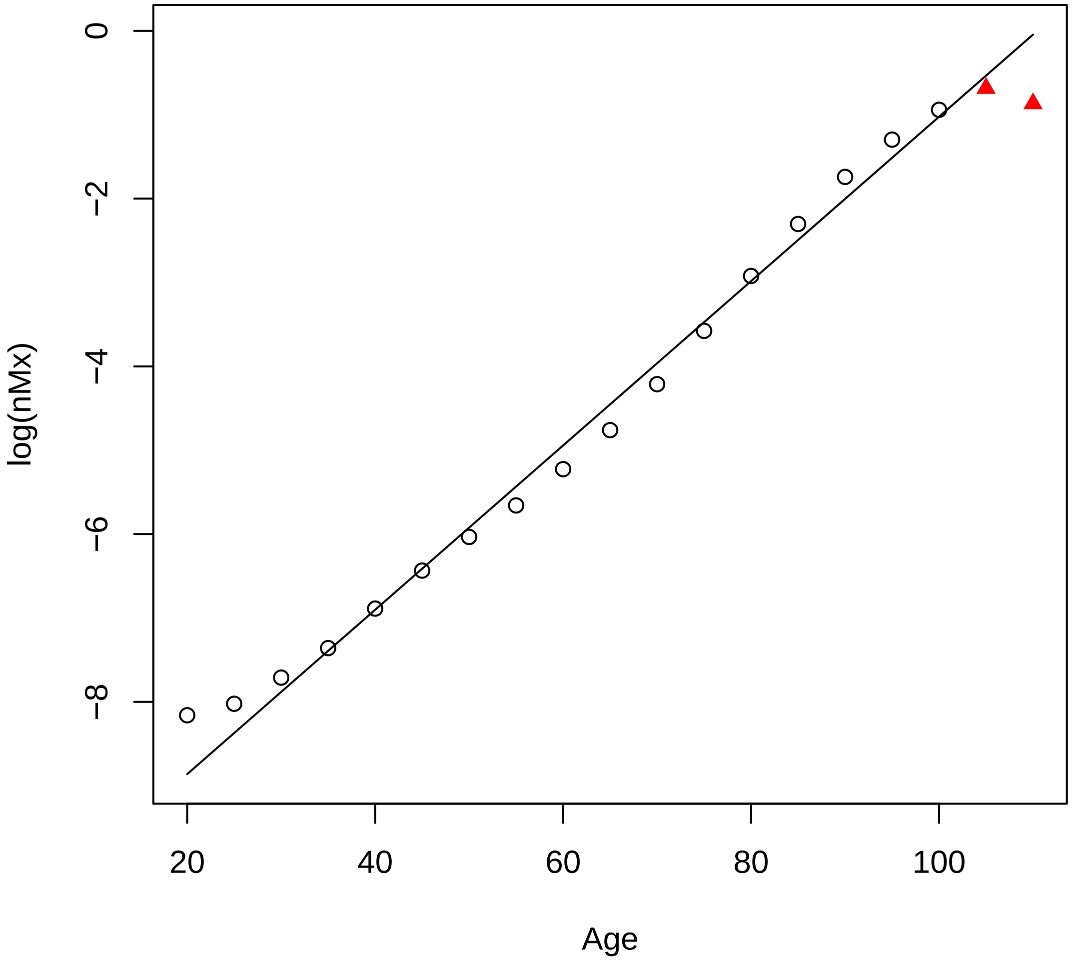
$$\log \mu(x) = \log(\alpha) + \beta x$$

This suggests a regression approach my be useful

Gompertz Fits So Well!



Or Does It?!



Other Mathematical Models of Mortality

Advantages:

- Compact, small number of parameters
- Highly interpretable
- Generalizable
- Good for comparative work (particularly interspecific comparisons)

Disadvantages:

- Hard to fit
- Numerical estimates frequently unstable (correlation between components)
- Almost certainly "wrong"
- What if there is a new source of mortality that is not covered by the model?

Two Models

Despite their shortcomings, two models are of interest:

1. Siler 5-Component Competing Hazard
2. Heligman-Pollard 8-Component Model

Siler 5-Component Competing Hazard Model

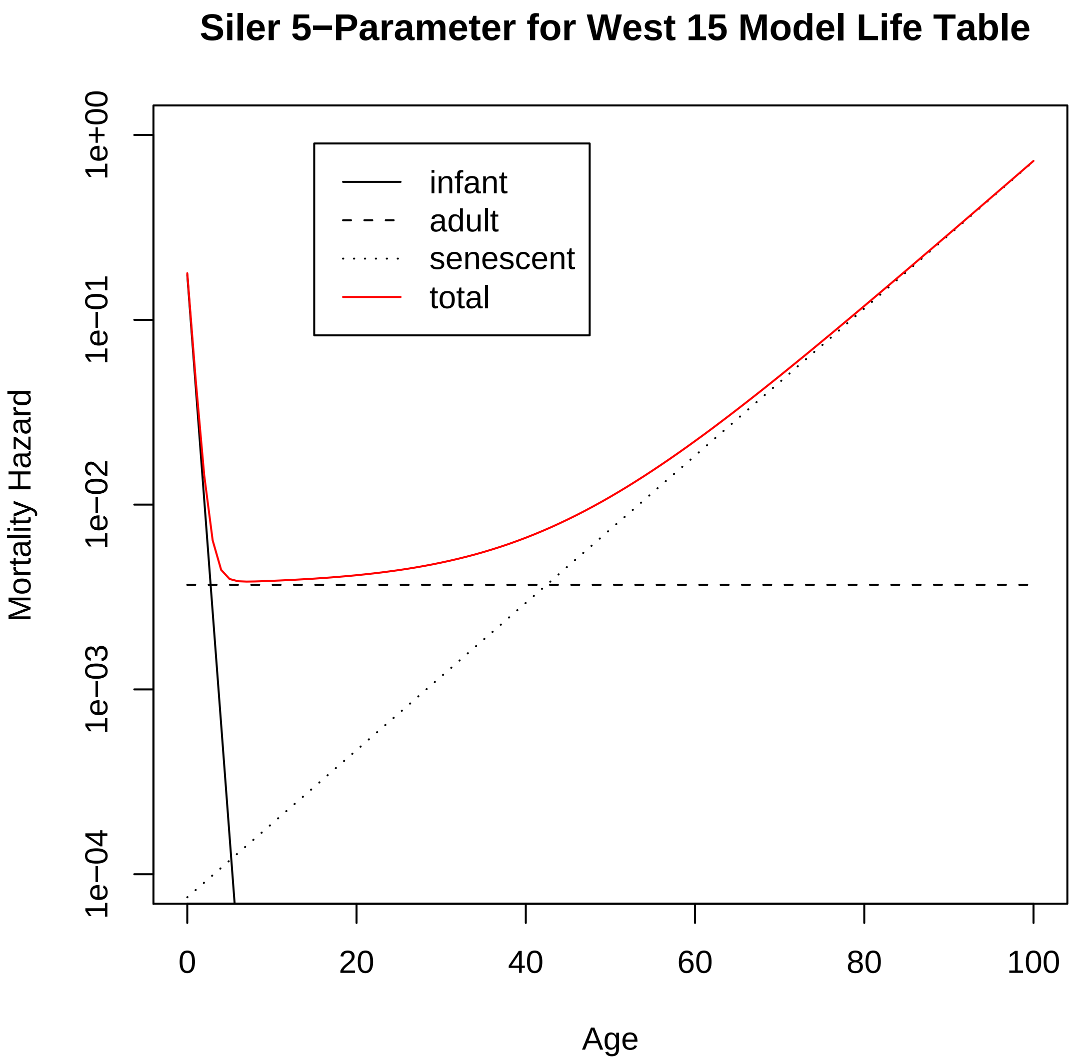
Developed to facilitate interspecific comparisons

Model of mortality hazard, $\mu(x)$

Competing hazard model

$$h(x) = ae^{-bx} + c + de^{fx}$$

Siler Fit to Coale-Demeny West 15



Heligman-Pollard 8-Component Model

Heligman-Pollard: models mortality probability q_x

$$\frac{q_x}{1-q_x} = A^{(x+B)^C} + D\exp\left[-E\left\{\log\frac{x}{F}\right\}^2\right] + GH^x$$

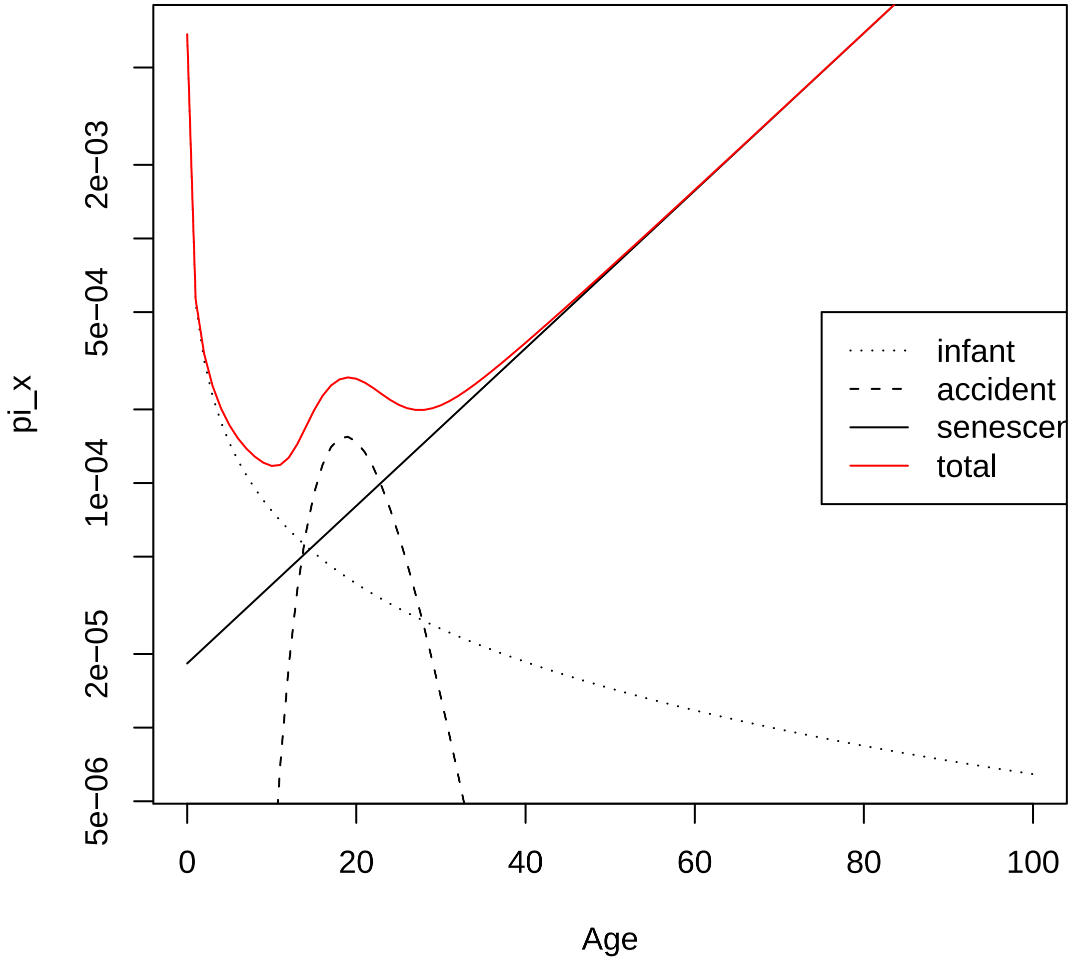
where q_x is the mortality probability at age $x = 0, 1, 2, \ldots, \omega$

The HP model is a beast to fit

Tremendous identifiability problems because the parameters are highly correlated

Dellaportas et al. (2001) suggest a Bayesian strategy to improve model fit and numerical properties

Heligman-Pollard Fit to UK 1995



Relational Life Tables

Sometimes, we actually have some data, but not quite enough to make a full life table

Brass (1971) suggested a regression procedure in which observed mortality was regressed onto a mortality standard, $X^{(S)}$

Mortality probabilities, survivorships, etc. are distributed on $[0, 1]$ and are **not** normally distributed (and the errors associated with them are certainly not)

This makes regression tricky

The usual way to handle this is to use a transformation of the data known as a logit

for some $0 \leq x \leq 1$,

$$\hat{Y} = \mathrm{logit}(x) = \log \left(\frac{x}{1-x} \right)$$

$$x = \frac{e^{\hat{Y}}}{1 + e^{\hat{Y}}}$$

Note that this transforms a variable that ranges over $[0\ 1]$ to one that ranges over $[-\infty\ \infty]$

We now perform the following regression

$$\hat{Y} = \alpha + \beta X^{(S)}$$

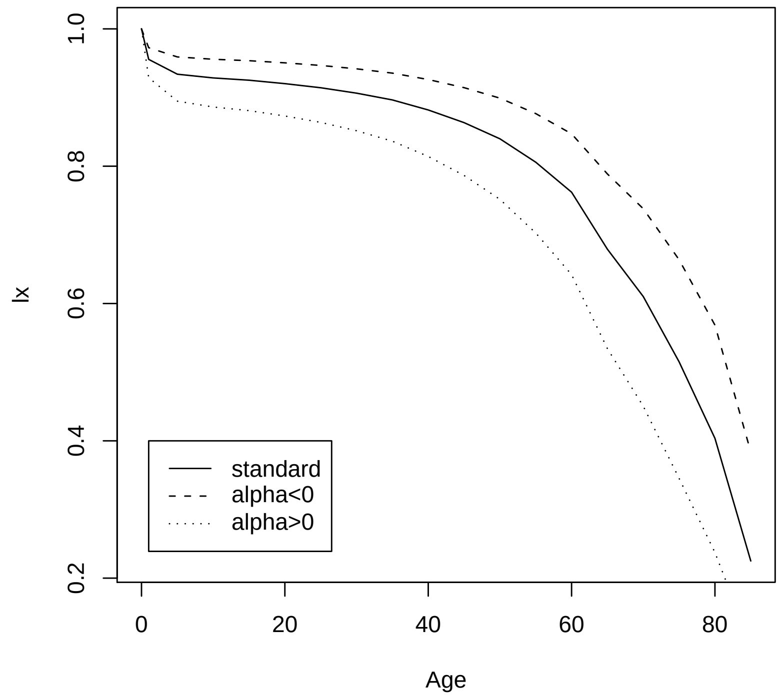
α is interpreted as the "level" of mortality

β is the "shape" of mortality

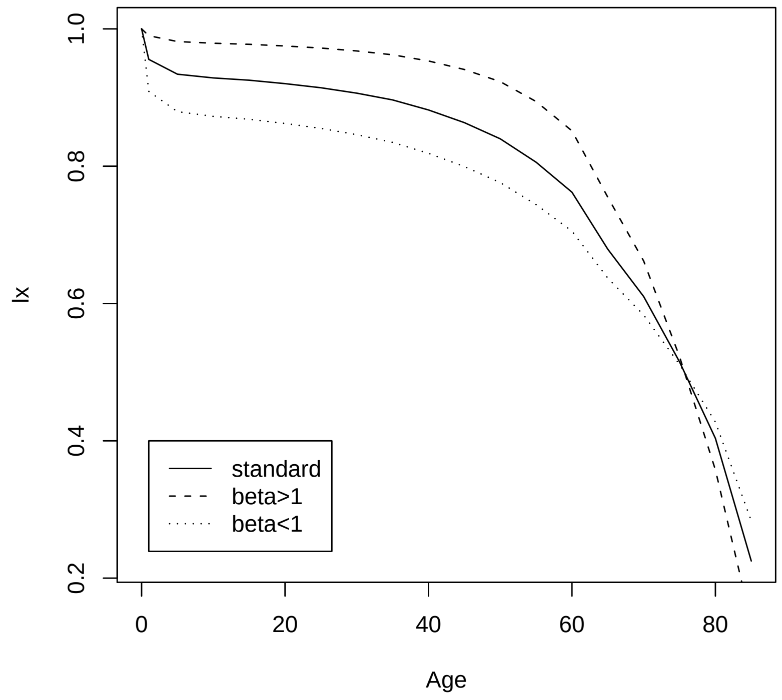
Be Careful! The notational usage in Preston et al. (2001) on this topic is not particularly standard

- q_x is not the standard 5-year mortality probability at age x
- q_x is really $1 - l_x$
- Perhaps a better notation would be ${}_xq_0$

Effect of Changing α



Effect of Changing β



Model Life Tables

There are a variety of model life table systems

1. UN (1958, 1982)
2. Coale & Demeny (1966, 1983)
3. INDEPTH
4. Weiss, 1973 (Obscure anthropological model life tables)

Model Life Table Construction

Model life table construction involves three general steps

- Gather a lot of high quality life tables
- Use Multivariate Analysis Techniques (e.g., PCA) to find clusters
- Combine them to provide a standard

Coale-Demeny Regional Model Life Tables

The Coale & Demeny is broken down into 4 "Regional" Models: North, South, East, and West

North Sweden pre-1920, Norway, Iceland: low infant and post-50 mortality
South Spain, Portugal, Southern Italy: high infant and post-65 mortality
East Austria, Germany, Northern Italy, Hungary, Poland: mortality rates generally high, particularly after 50
West Everything else: Most frequent anthropological application

They are indexed (typically) by life expectancy at birth or age 10

Use of model life tables follows a simple recipe:

1. Use whatever information available to estimate life expectancy
2. Use whatever information available to determine the mortality pattern (or use West model life table)
3. Use the table that fits your assumptions

Model Life Tables Construction

Component mortality models (e.g., Coale-Demeny)

$$\mathbf{m} = \mathbf{C}\mathbf{a}$$

where $\mathbf{m}$ is a vector of $\mathrm{logit}({}_nq_x)$ values for $x \in 0 \ldots k$

$\mathbf{C}$ is a $k \times l + 1$ matrix of loadings of the k ages on the first l components of a PCA (with a leading column of ones)

$\mathbf{a}$ is an $l + 1 \times 1$ vector of coefficients

$\mathbf{a}$ estimated by regressing empirically-derived values of $\mathrm{logit}({}_nq_x)$ on $\mathbf{C}$

AIDS Decremented Model Life Tables: INDEPTH Model 1 & 5

Two model life table families are relevant:

- Model 1: Primarily western Africa (similar to Coale-Demeny North)
- Model 5: Primarily eastern and southern Africa (unique)

AIDS-decremented life tables constructed by adding excess mortality in the characteristic AIDS pattern to the model 1 life table

Regress the difference in $\mathbf{m}$ between model 1 and model 5 life tables with a common $\mathring{e}_0$

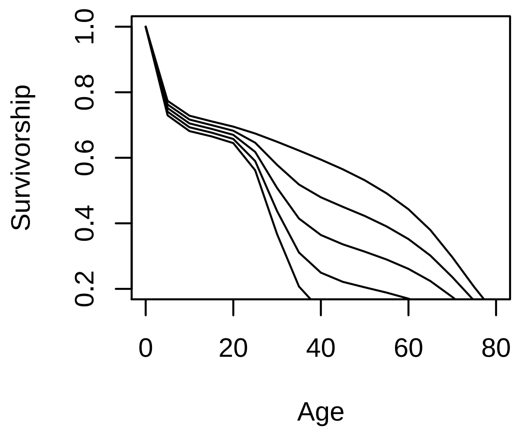
Let $\mathbf{d}$ denote the coefficients of this regression

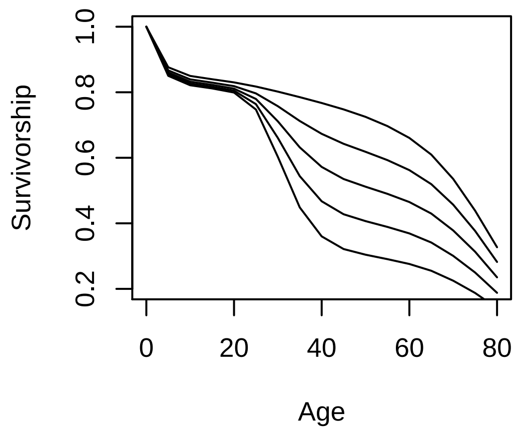
The AIDS decremented model life table is generated by:

$$\mathbf{m} = \mathbf{C}(\mathbf{a} + \alpha\mathbf{d})$$

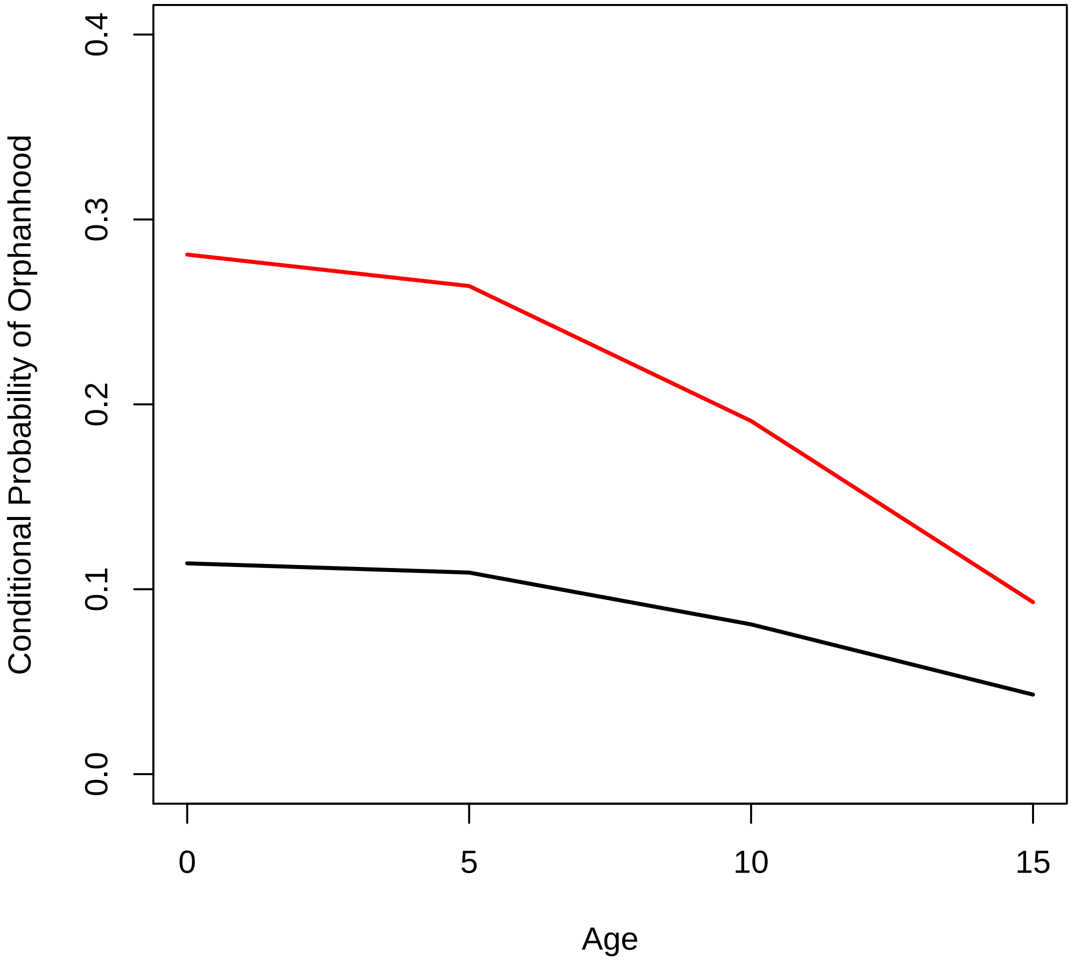
where α is a scale parameter determining the extent of mortality excess

INDEPTH AIDS-Decremented Model Life Tables

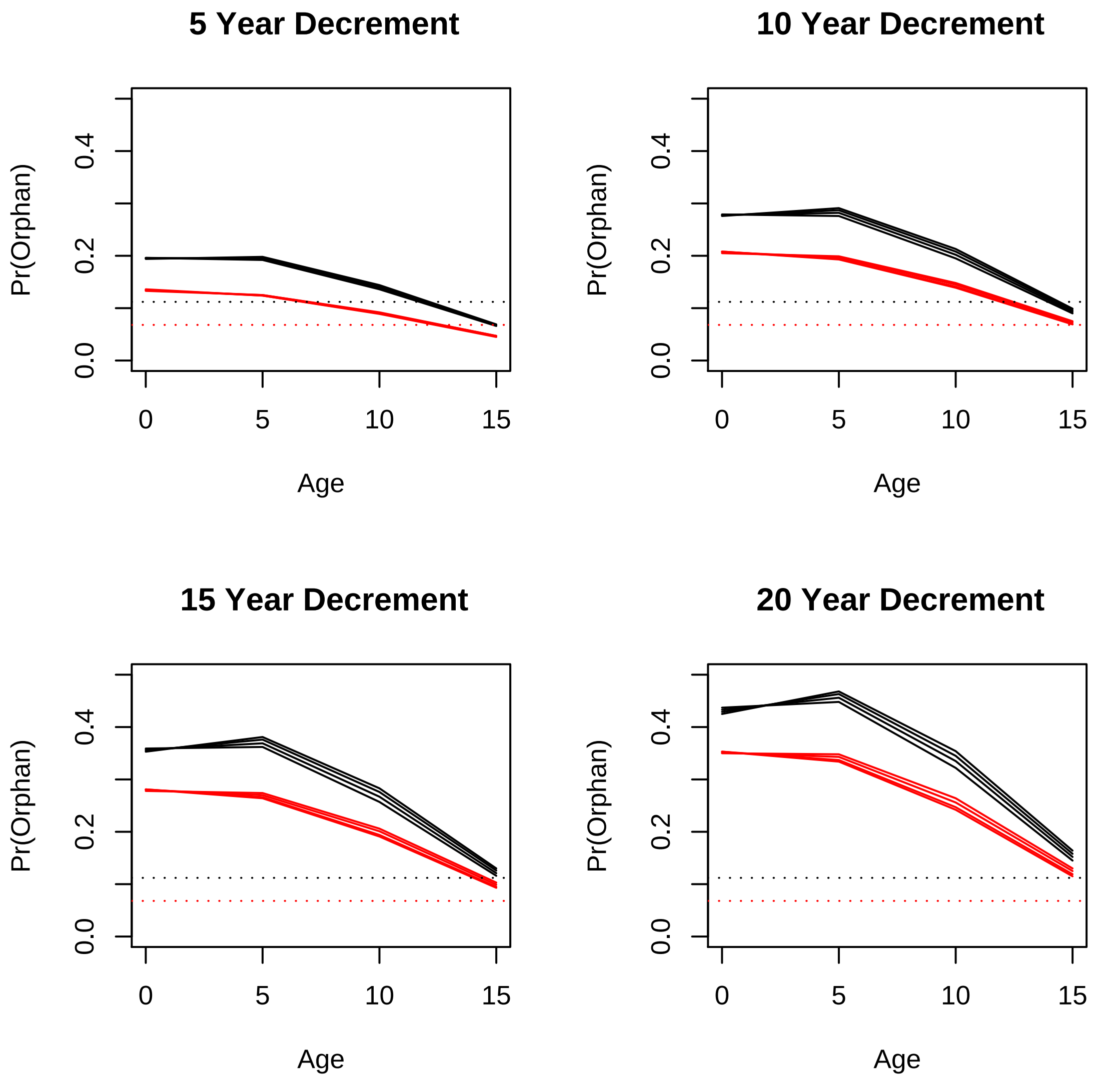




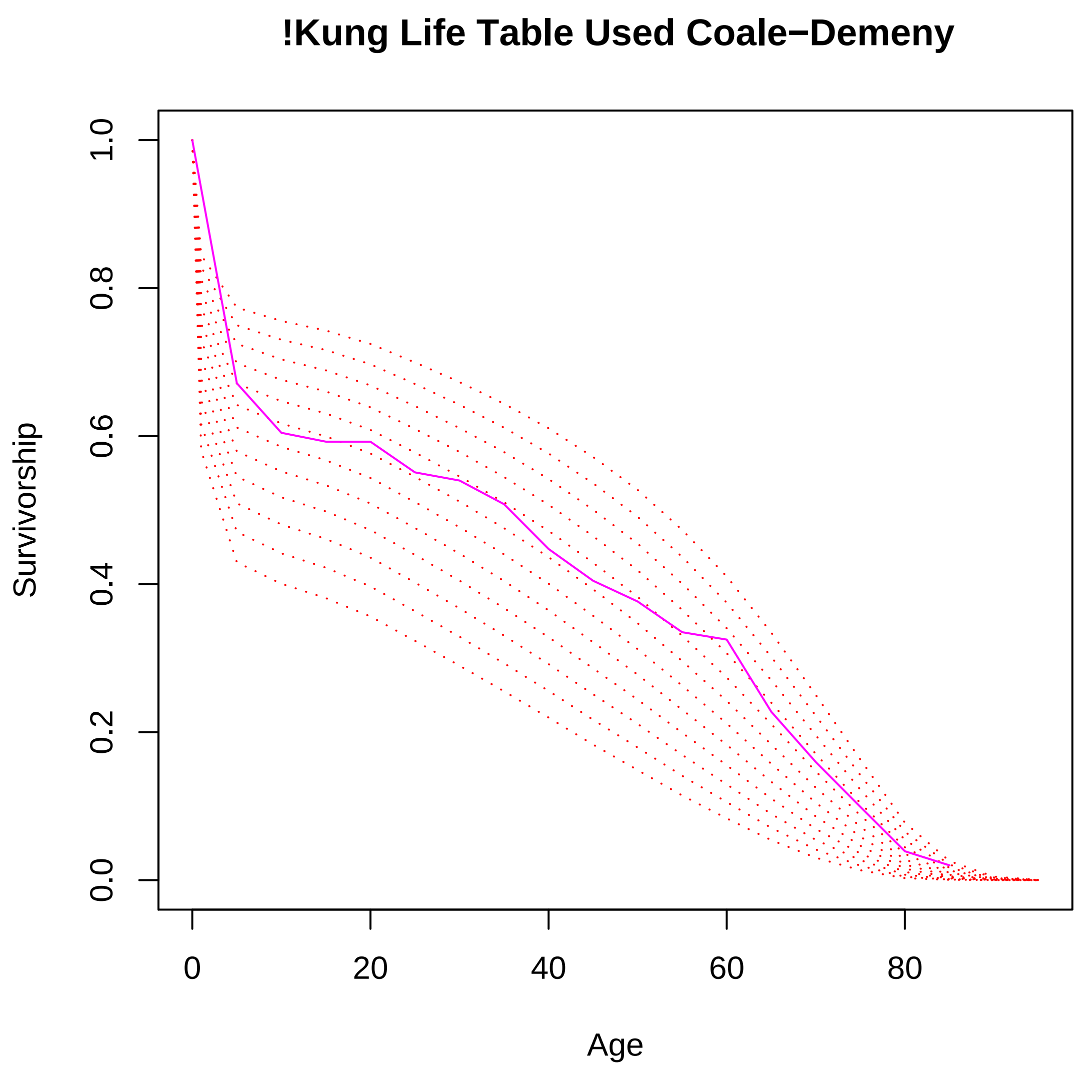
AIDS Creates More Orphans



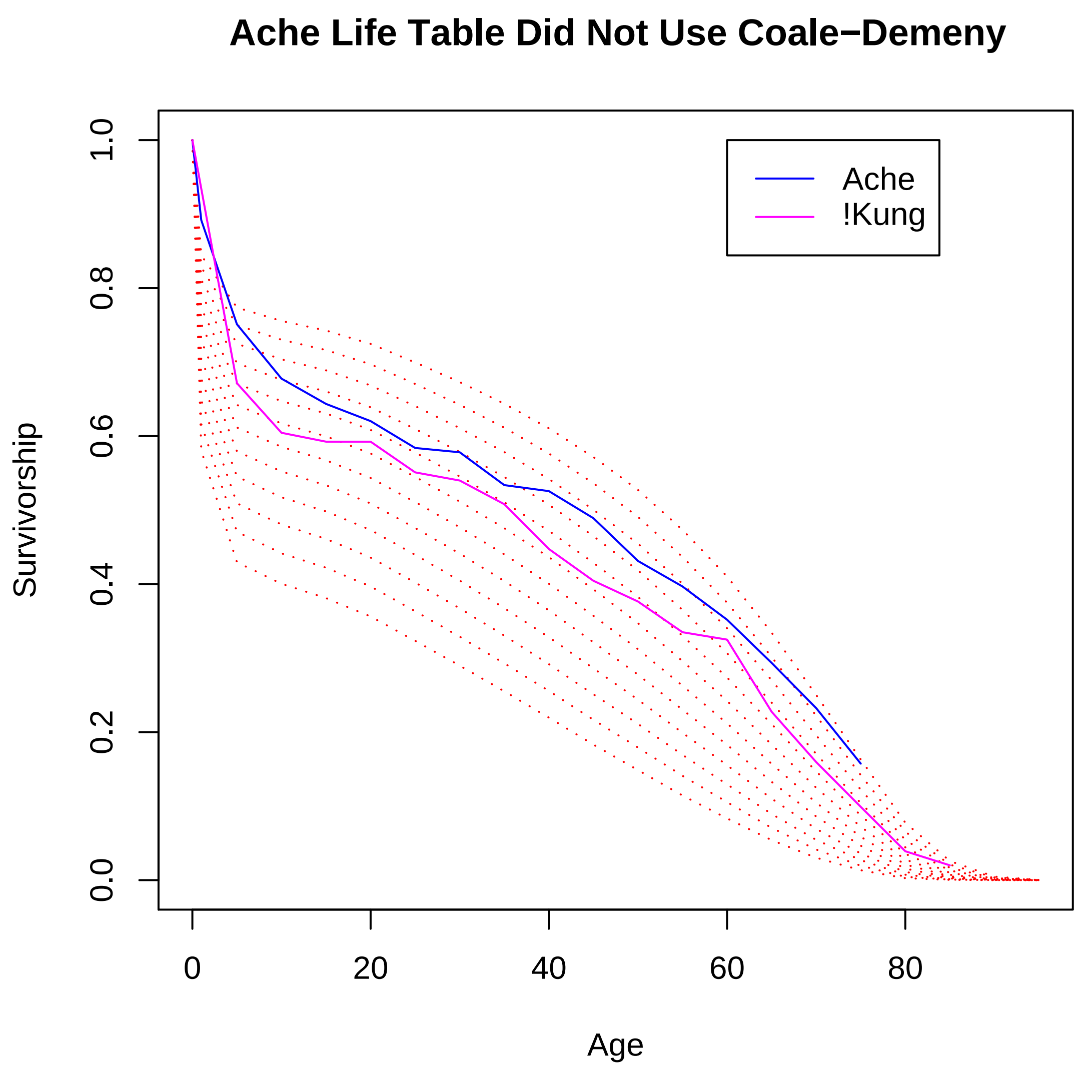
It's the Pattern of Mortality that Matters

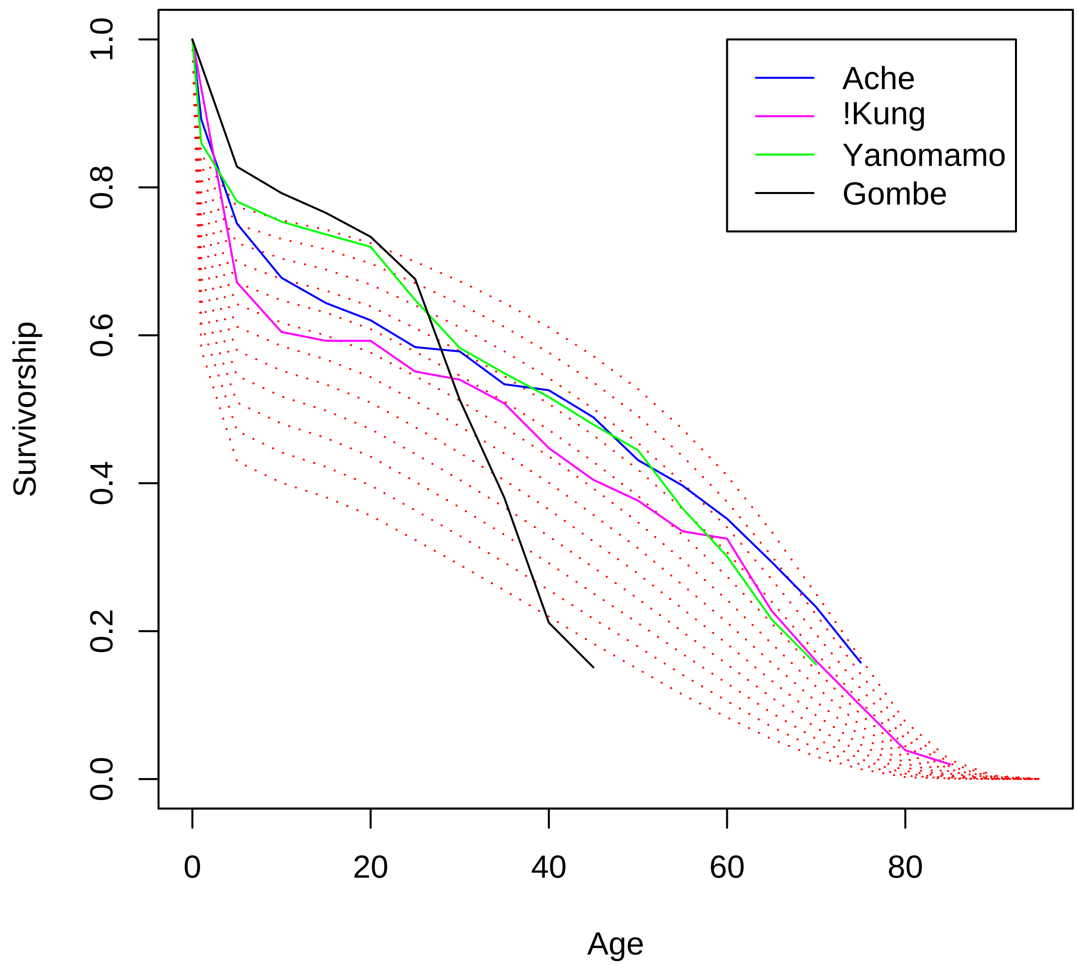


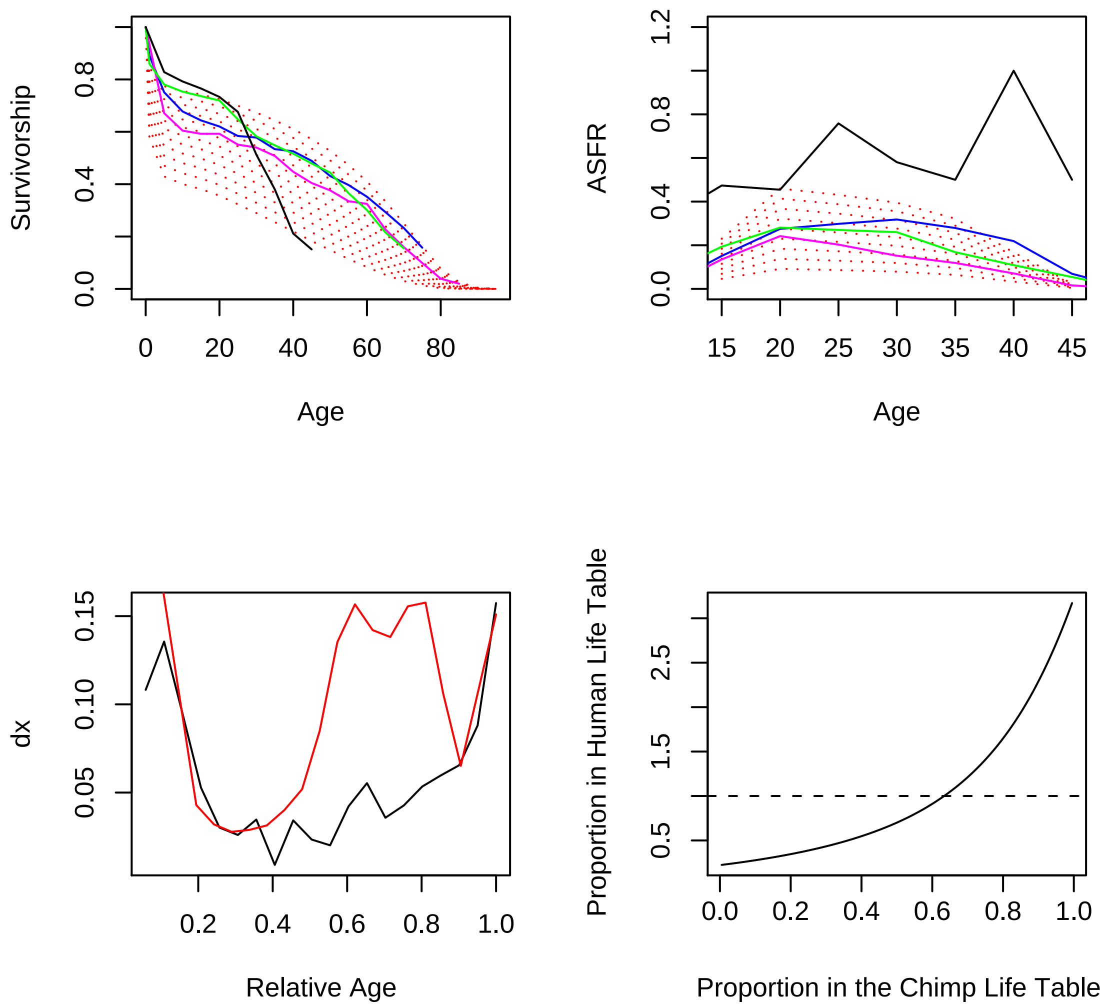
Howell Used Coale-Demeny Model Life Tables



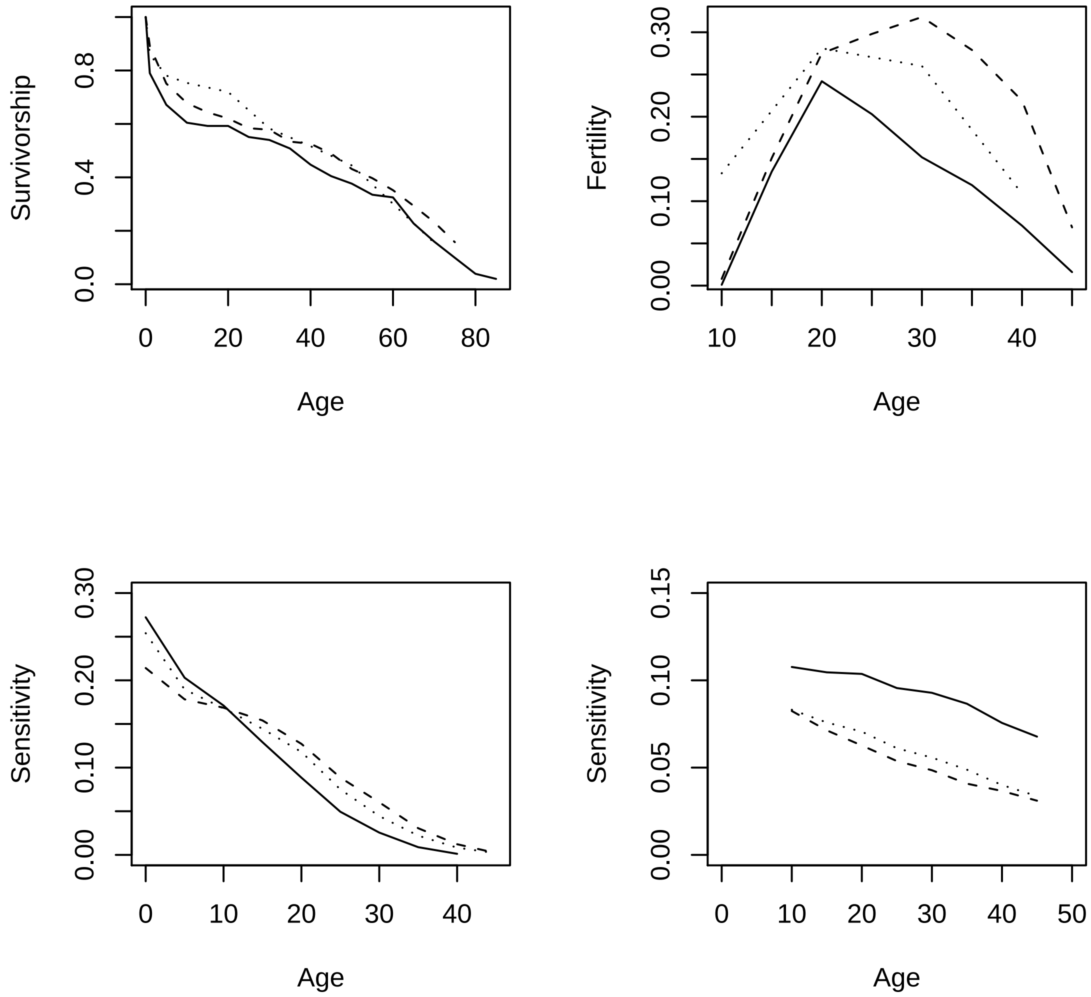
Hill & Hurtado Did Not




Ache
!Kung
Yanomamo
Gombe
Survivorship
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
Age



Survivorship
0.0
0.4
0.8
0
20
40
60
80
Age
ASFR
0.0
0.4
0.8
1.2
15
20
25
30
35
40
45
Age
dx
0.05
0.10
0.15
0.2
0.4
0.6
0.8
1.0
Relative Age
Proportion in Human Life Table
0.5
1.5
2.5
0.0
0.2
0.4
0.6
0.8
1.0
Proportion in the Chimp Life Table


Survivorship
0.0
0.4
0.8
0
20
40
60
80
Age
Fertility
0.00
0.10
0.20
0.30
10
20
30
40
Age
Sensitivity
0.00
0.10
0.20
0.30
0
10
20
30
40
Age
Sensitivity
0.00
0.05
0.10
0.15
0
10
20
30
40
50
Age