Health Transition Review, Supplement to Volume 6, 273-282

Appendices to Kunitz on mortality calculations

K. Ruben Gabriel

Department of Statistics, University of Rochester, Rochester, NY

Observed survival

Survey data provided information on survival of the survey period for each case. The proportions surviving, for each sex and age, could be estimated from these data, but rates for individual ages, or even for five-year age groups, were very unreliable because of the small number of cases. A more reliable estimate of the age trend in the proportion of each sex surviving was therefore obtained by fitting logistic curves, i.e., for each sex, the probability of surviving from age **A** was modelled as

$$\text{Prob(surviving from age A)} = 1/[1+\exp(-a-bA)] \quad \text{for some a and b.}$$

For each sex, logistic regression was used to estimate the parameters α and β, yielding estimated models

$$\mathbf{Prob}(\text{surviving from age } \mathbf{A}) = 1/[1+\exp(-3.931+.04831\mathbf{A})] \qquad \text{for males,}$$

and

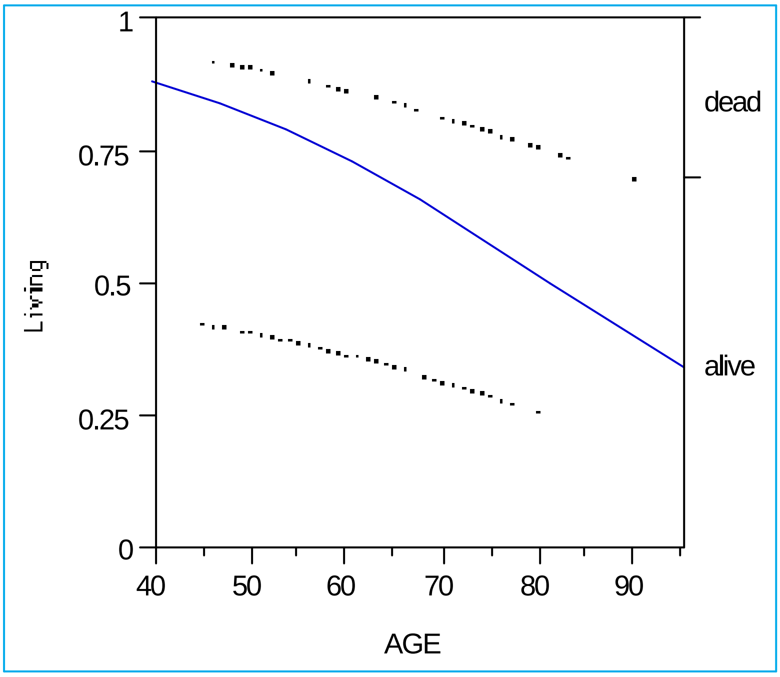
$$\mathbf{Prob}(\text{surviving from age } \mathbf{A}) = 1/[1+\exp(-9.243+.12310\mathbf{A})] \qquad \text{for females.}$$

$$= 1/[1+\exp(1.38-5.3114\mathbf{S}-.02647\mathbf{A}+.07479\mathbf{SA})]\ .$$

Figures 1 and 2 show these logistic regression fits: the solid curve always is the modelled proportion of survival; the dots above and below it give an indication of where the individual observations were.

In summary, female proportions surviving decrease rapidly with age, as one would expect. Male proportions decrease much more slowly: Figure 3 shows that at ages up to 70 they are below those of females, as one would expect, but above age 70 male proportions surviving decrease very slowly and indeed exceed those of females, a most surprising result. It is difficult to know how to account for this result which does not seem to be due to a too small number of men above 70: about a third of those surveyed were in this age group. In any case it is obviously spurious as it suggests a probability of more than 40 per cent for males to survive to age 90. Clearly, therefore, the data for males at advanced ages must be discounted. Whether this casts doubt also on the data for earlier ages, and for females, is unclear.

Figure 1
Survival by age - males



Converged by Gradient

Whole–Model Test

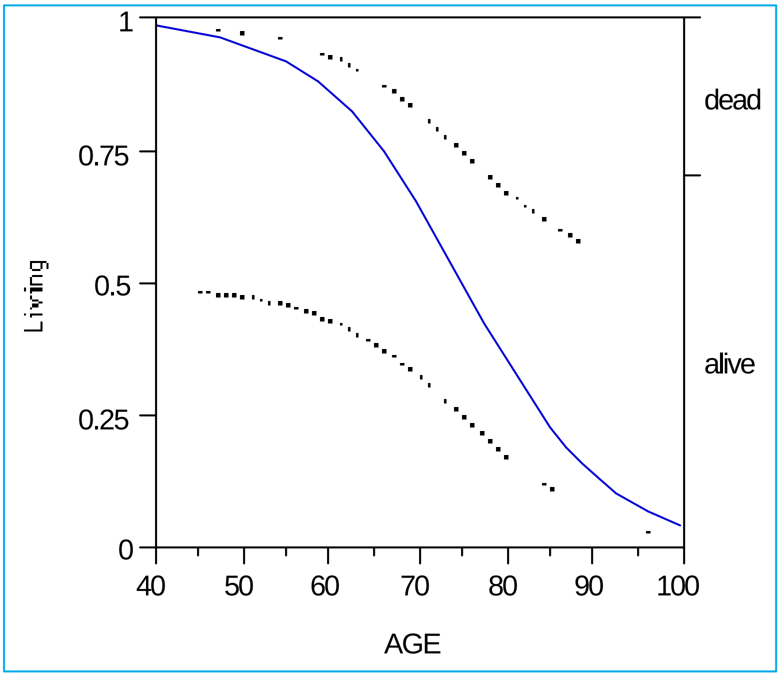
Source	DF	–LogLikelihood	ChiSquare	Prob>ChiSq
Model	1	3.036037	6.072075	0.013734
Error	137	82.126750		
C Total	138	85.162788		

RSquare (U)	0.0356
Observations (or Sum Wgts)	139

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	3.93132203	1.3163508	8.92	0.0028
AGE	-0.0483117	0.0200902	5.78	0.0162

Figure 2
Survival by age - females



Converged by Gradient

Whole–Model Test

Source	DF	–LogLikelihood	ChiSquare	Prob>ChiSq
Model		23.35248	46.70497	0.000000
Error	170	81.20280		
C Total	171	104.55529		

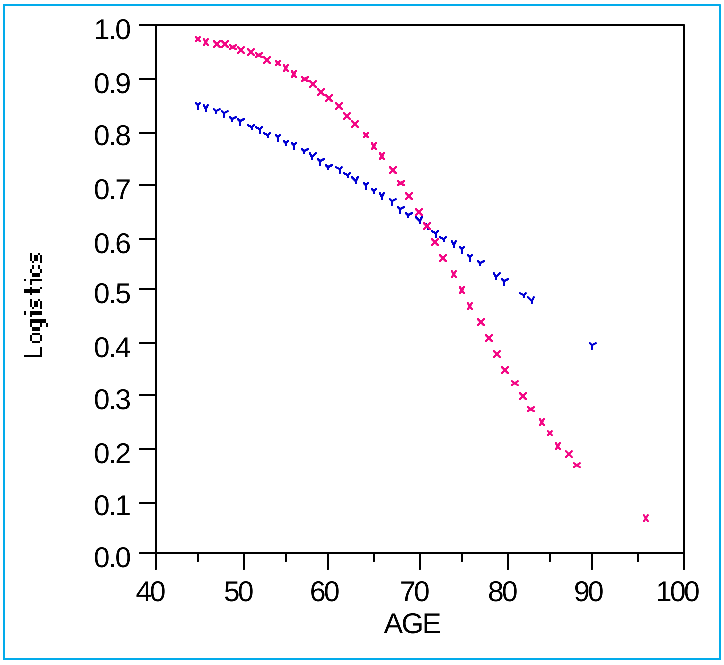
RSquare (U)	0.2234
Observations (or Sum Wgts)	172

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	9.24278023	1.5285463	36.56	0.0000
AGE	-0.1230984	0.0216793	32.24	0.0000

Figure 3
Survival by Age and Sex

1/(1+exp[1.38-5.3114*SEX-0.02647*AGE+.07479*AGE*SEX])



Footnote   Male - Y, SEX=1,   Female - X, SEX=2

Survival probabilities: life table computations and interpolation

The 1979 mortality experience of Yugoslavia was used as a baseline for comparing the actual mortality experience of the surveyed group. The proportions surviving during the survey were compared to corresponding estimated probabilities of survival according to the 1979 experience.

For this purpose, the 1979 age and sex specific mortality rates were used to compute a life table for each sex, resulting in function $\mathbf{l_A}$ which shows the probability of surviving from birth to exact age **A**, computed at five-yearly intervals - see accompanying life tables. In order to be able to interpolate the function $\mathbf{l_A}$ for other ages, and be able to compute the probability of surviving the years of the survey, the logits of $\mathbf{l_A}$ for each sex and for ages 40 and above (ages of people in the survey) were fitted by a cubic in age - see Figures 4 and 6. The fitted cubics were then back transformed (inverse of logit) to yield fitted $\mathbf{l_A}$'s

$$\mathbf{l_A\ fit} = 1/\{1+\exp[-(5.532-21.654\,\mathbf{a}+41.593\,\mathbf{a}^2-29.943\,\mathbf{a}^3)]\}\ \text{ for males,}$$

and

$$\mathbf{l_A\ fit} = 1/\{1+\exp[-(6.189-25.274\mathbf{a}+49.388\mathbf{a}^2-34.059\mathbf{a}^3)]\}\ \text{ for females,}$$

where $\mathbf{a}=\mathbf{A}/100$.

These back-transformed cubics were found to yield the function very closely for ages 40 and above - see **l** and **l Fit** functions in life tables and Figures 5 and 7. It was therefore felt that these functions could be used with confidence to gauge probabilities of survival after age 40.

Table 1
Male Life Table 1979, with interpolation

Initial Age	m	p	l	logit l	logit l Fit	l Fit
0	0.0092	0.99084213	1	•	5.532	0.99605754
1	0.035	0.86915888	0.99084213	4.68394118	5.31958936	0.99512908
5	0.0012	0.99401795	0.86119923	1.82528629	4.54953962	0.98953853
10	0.0005	0.99750312	0.85604749	1.78284241	3.752587	0.97708064
15	0.0004	0.998002	0.85391004	1.76560329	3.11868488	0.95765693
20	0.0009	0.9955101	0.85220393	1.75199244	2.625376	0.93247699
25	0.0014	0.99302441	0.84837762	1.72193277	2.25020313	0.90466805
30	0.0014	0.99302441	0.84245969	1.67664446	1.970709	0.87768725
35	0.0018	0.99104032	0.83658304	1.63302083	1.76443638	0.85376441
40	0.0027	0.98659051	0.82908752	1.57917414	1.608928	0.8332625
45	0.004	0.98019802	0.81796989	1.50265339	1.48172662	0.81483324
50	0.0066	0.96753566	0.80177246	1.39740931	1.360375	0.79582064
55	0.0101	0.95074372	0.77574345	1.24103115	1.22241588	0.77248842
60	0.0151	0.92724645	0.73753321	1.03318658	1.045392	0.73988906
65	0.0226	0.89304307	0.68387505	0.77163769	0.80684613	0.69143703
70	0.0358	0.83570445	0.61072987	0.45038124	0.484321	0.6187677
75	0.0571	0.75016408	0.51038967	0.04156468	0.05535938	0.51383631
80	0.0873	0.64169916	0.382876	-0.4773588	-0.502496	0.37695428
85	0.1622	0.42298115	0.24569121	-1.1217263	-1.2117024	0.22939997
90	0.2191	0.29219835	0.10392275	-2.1543788	-2.094717	0.10961136

$\mathbf{a}_i$=age at beginning of interval/100,

$\mathbf{m}_i$=death rate in interval,

$\mathbf{p}_i=(1-int*\mathbf{m}_i/2)/(1+int*\mathbf{m}_i/2)$ =survival in interval

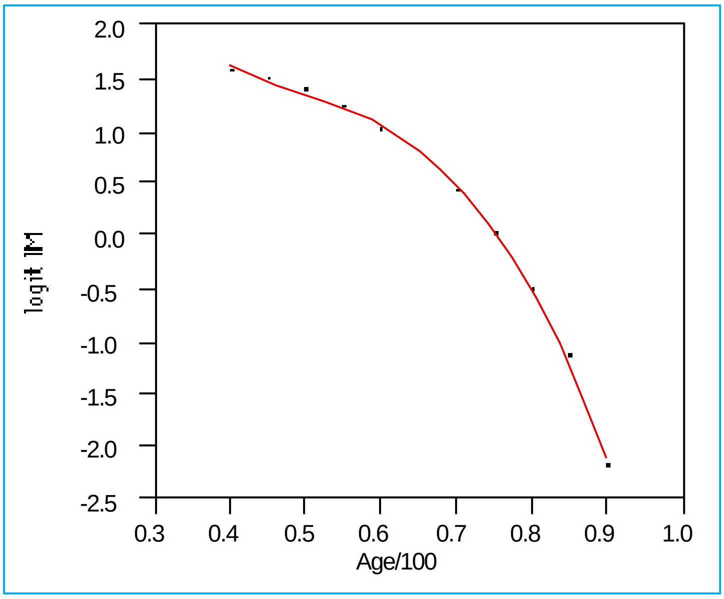
$\mathbf{l}_i=\Pi_{e=1}^{i-1}\,\mathbf{p}_e$ = survival from birth to beginning of age interval

$\textbf{logit } \mathbf{l}_i=\log[\mathbf{l}_i/(1-\mathbf{l}_i)]$

$\textbf{logit } \mathbf{l}_i \textbf{ fit} = 5.532-21.654\,\mathbf{a}_i+41.593\,\mathbf{a}_i^2-29.943\,\mathbf{a}_i^3$,

$l_i \text{ Fit} = 1/[\,1+\exp(-\textbf{logit } \mathbf{l}_i \textbf{ fit})\,]$,

Figure 4
Interpolation of survival probabilities by a logistic polynomial fit for l_M from age 40

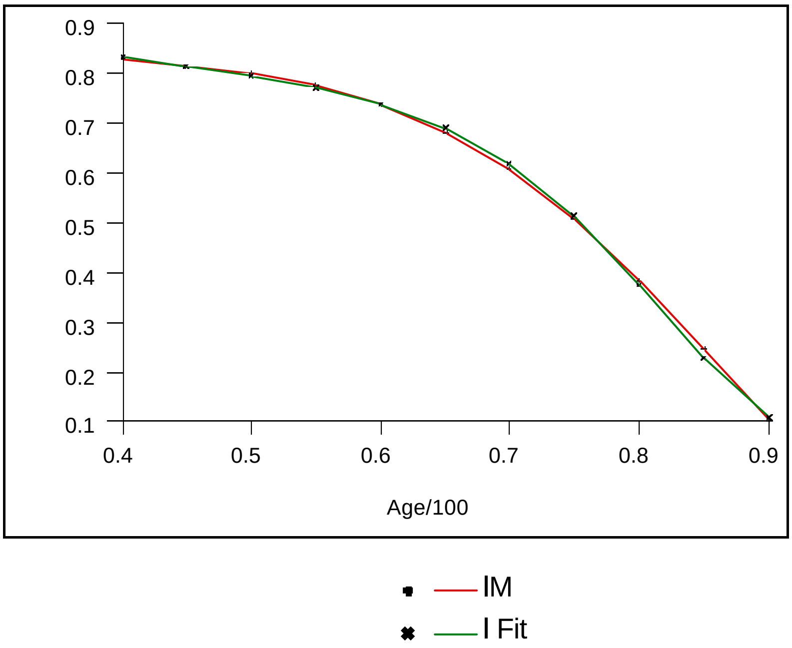


Polynomial Fit, degree=3 **RSquare 0.998762**

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	5.5322484	1.30126	4.25	0.0038
Age/100	-21.65365	6.38662	-3.39	0.0116
Age/100^2	41.593488	10.1038	4.12	0.0045
Age/100^3	-29.94314	5.16919	-5.79	0.0007

Figure 5
Male survival from birth

1979 life table (l_M) and logistic interpolation from age 40



Interpolation formula l_M Fit = 1/[1+exp(-logit fit)] , where

logit fit = 5.532-21.654a+41.593a^2-29.943a^3, with a = age/100.

Table 2
Female Life Table 1979, with interpolation

Initial Age	m	p	l	logit l	Fitted logit	l Fit
0	0.008	0.99203187	1	•	6.189	0.99795232
1	0.0318	0.88040617	0.99203187	4.82430572	5.94116474	0.99737793
5	0.0011	0.99451508	0.87339098	1.93127955	5.04451263	0.99359667
10	0.0004	0.998002	0.8686005	1.88864102	4.121421	0.98403749
15	0.0003	0.99850112	0.86686504	1.87351993	3.39418087	0.96752218
20	0.0005	0.99750312	0.86556571	1.8623078	2.837248	0.94465576
25	0.0005	0.99750312	0.8634045	1.8438593	2.42507812	0.91871975
30	0.0006	0.99700449	0.86124868	1.82570006	2.132127	0.89398676
35	0.0008	0.99600798	0.85866881	1.80427728	1.93285037	0.87356458
40	0.0013	0.99352106	0.85524099	1.77631292	1.801704	0.85835624
45	0.002	0.99004975	0.84969993	1.73224951	1.71314362	0.84724358
50	0.0031	0.9846192	0.84124521	1.66752234	1.641625	0.83775593
55	0.005	0.97530864	0.82830618	1.57367011	1.56160387	0.82658338
60	0.0077	0.96222713	0.80785418	1.43612699	1.447536	0.80961893
65	0.013	0.937046	0.7773392	1.2502273	1.27387713	0.78140573
70	0.0218	0.89663348	0.7284026	0.98653308	1.015083	0.73401372
75	0.0388	0.82315406	0.65311015	0.63273851	0.64560937	0.65602036
80	0.0638	0.72488141	0.53761027	0.15072579	0.139912	0.53492105
85	0.1355	0.49393091	0.38970369	-0.4485579	-0.5275534	0.37108771
90	0.2067	0.31860887	0.1924867	-1.4339325	-1.382331	0.20063489

$\mathbf{a}_i$=age at beginning of interval/100,

$\mathbf{m}_i$=death rate in interval,

$\mathbf{p}_i=(1-int*\mathbf{m}_i/2)/(1+int*\mathbf{m}_i/2)$ = survival in interval

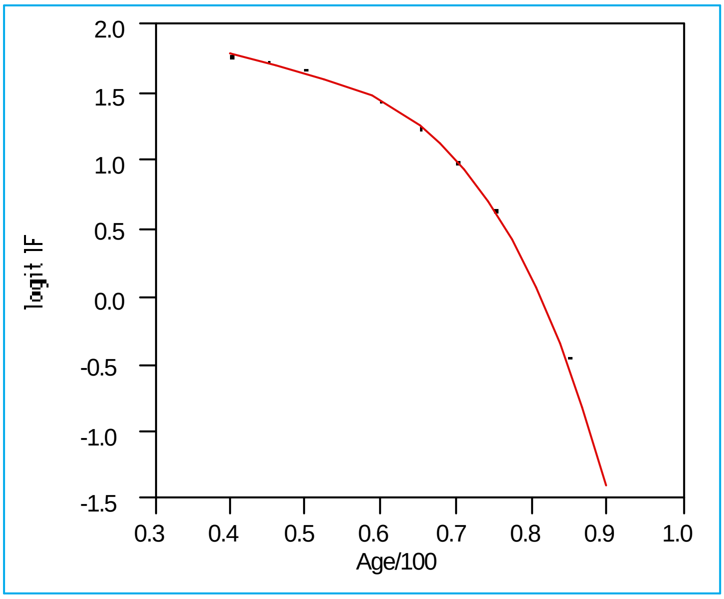
$\mathbf{l}_i=\Pi_{e=1}^{i-1}\mathbf{p}_e$ = survival from birth to beginning of age interval

$\textbf{logit } \mathbf{l}_i = \log[\mathbf{l}_i/(1-\mathbf{l}_i)]$

$\textbf{logit } \mathbf{l}_i \textbf{ fit} = 6.189-25.274\mathbf{a}_i+49.388\mathbf{a}_i^2-34.059\mathbf{a}_i^3$,

$l_i \text{ Fit} = 1/[1+\exp(-\textbf{logit } \mathbf{l}_i \textbf{ fit})]$,

Figure 6
Interpolation of survival probabilities by a logistic polynomial fit for l_F from age 40

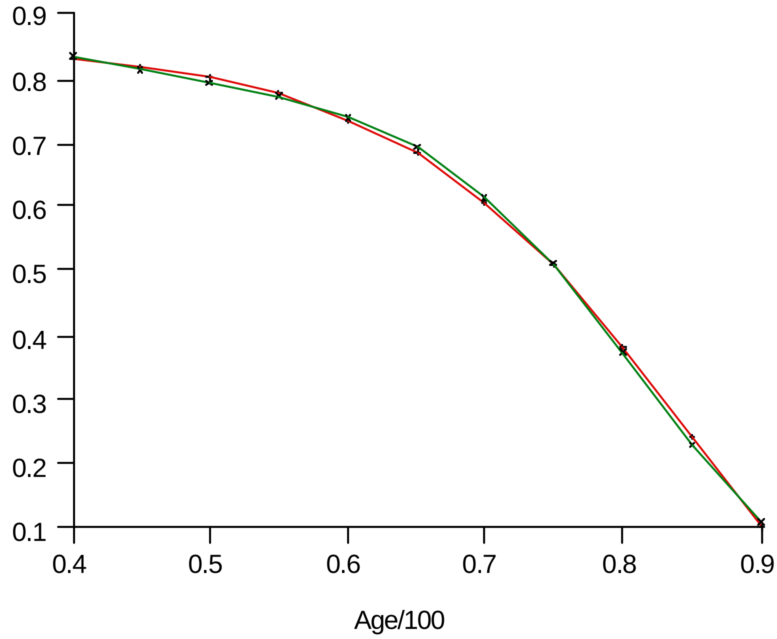


Polynomial Fit, degree=3 **RSquare** **0.99884**

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	6.1885767	1.08327	5.71	0.0007
Age/100	-25.27399	5.31673	-4.75	0.0021
Age/100^2	49.387916	8.41117	5.87	0.0006
Age/100^3	-34.05944	4.30325	-7.91	0.0001

Figure 7
Female Survival from Birth,

1979 life table (l_F) and logistic interpolation from age 40



Interpolation formula l_F Fit = 1/[1+exp(-logit fit)] , where

logit fit = 6.189-25.274$\mathbf{a}$+49.388$\mathbf{a}^2$-34.059$\mathbf{a}^3$, with $\mathbf{a}$ = age/100.

Comparison of observed and expected survival.

Survival of the period (9 1/8 years) of the survey was compared to expected survival based on the 1979 death rates by age and sex. The plots in the text show, separately for each sex and as a function of age, the actual survival proportions (**+**) of the survey and the corresponding probabilities (**x**) of surviving 9.125 years according to the 1979 death rates. Both the proportions and the probabilities are interpolated in the available observations. The former were fitted by logistic regression on age, as discussed in 'Observed Survival'. The latter were calculated as $(l_{A+9.125}\ \mathbf{fit})/(l_A\ \mathbf{fit})$ by interpolating in the survival function l_A, as discussed in 'Survival Probabilities'.