CENSUS SURVIVAL RATIO CONSISTENCY CHECK

Consider two censuses of a closed population taken n-years apart, providing age distributions in n year groups. Define the following notation.

x_i - enumerated number in i-th age group at 1st census

y_i - enumerated number in i-th age group at 2nd census

s_i - observed survivorship ratio, y_{i+1}/x_i

Tx_i - true number in i-th age group at 1st census

Ty_i - true number in i-th age group at 2nd census

Ts_i - true survivorship ratio

a_i - correction factor for 1st census, i-th age group, Tx_i/x_i, implies $a_i x_i = Tx_i$

b_i - correction factor for 2nd census, i-th age group, Ty_i/y_i, implies $b_i y_i = Ty_i$

Now assume that the `age distribution adjustment factors' a_i and b_i at the two censuses are identical, $a_i = b_i$. Then

$$Ts_i = Ty_{i+1}/Tx_i = a_{i+1}y_{i+1}/a_i x_i$$

which implies

$$Ts_i a_i = a_{i+1}s_i$$

whence

$$a_{i+1} = a_i(Ts_i/s_i)$$

i.e.,

$$a_2 = a_1(Ts_1/s_1)$$

$$a_3 = a_2(Ts_2/s_2)$$

$$= a_1(Ts_1/s_1)(Ts_2/s_2)$$

and so on. Given a value for a_1, then, we may compute a_2, a_3, ... by multiplying a_1 by the ratios of the true to the observed survivorship ratios. This determines the shape, but not the level, of the a_i. To set the level we used the observed population totals. In applying these formulas we will use an observed life table, perhaps a model table, considered appropriate for the population in question.

Note that if the second census is under enuemerated relative to the first, and uniformly so by age, the observed survivorship ratios s_i will all be too small, and the a_i for older age groups too small relative to the a_i for younger age groups. Conversely, if the first census is under enumerated relative to the second, and uniformly so by age, the observed survivorship ratios will all be too big, and the a_i for the older age groups will be too large relative to the a_i for the younger age groups.

Similar remarks apply to a life table with too high or too low survivorship ratios.

The first example below illustrates the consistent correction procedure for censuses five years apart, which is the original case considered by Demeny and Shorter. The second example uses the generalized method of synthetic survivorship ratios presented below.

Table 1. The Census Survival Ratio Consistency Check Applied to Japanese Females, 1970-75

Age	AD 1970	AD 1975	s(i)	1970 LLX	Ts(i)	i	Ts(i) /s(i)	IAF	FAF
0-4	4343	4873	1.0025	493,587	.9968	1	.9943	1.0000	1.0261
5-9	4041	4354	1.0010	492,001	.9986	2	.9976	.9943	1.0203
10-14	3911	4045	.9992	491,314	.9984	3	.9992	.9919	1.0178
15-19	4544	3908	.9921	490,526	.9971	4	1.0050	.9911	1.0170
20-24	5383	4508	.9972	489,126	.9962	5	.9990	.9961	1.0221
25-29	4602	5368	1.0041	487,262	.9954	6	.9913	.9951	1.0211
30-34	4226	4621	.9962	485,032	.9939	7	.9977	.9865	1.0123
35-39	4118	4210	.9954	482,084	.9913	8	.9959	.9842	1.0099
40-44	3703	4099	1.0005	477,913	.9872	9	.9867	.9802	1.0058
45-49	3223	3705	.9820	471,815	.9803	10	.9983	.9671	.9924
50-54	2669	3165	.9745	462,538	.9699	11	.9953	.9655	.9907
55-59	2400	2601	.9783	448,594	.9523	12	.9734	.9609	.9860
60-64	1986	2348	.9456	427,208	.9216	13	.9746	.9354	.9598
65-69	1598	1878	.8936	393,717	.8661	14	.9692	.9116	.9354
70-74	1182	1428	.8063	340,995	.7686	15	.9532	.8836	.9067
75-79	744	953	.6734	262,081	.6317	16	.9381	.8423	.8643
80-84	413	501	.5036	165,546	.4615	17	.9164	.7901	.8107
85-89	160	208	.3281	76,401	.3043	18	.9275	.7241	.7430
90-94	44.3	52.5	.1582	23,249	.1818	19	1.1492	.6715	.6890
95-99	4.87	7.01	.0961	4,226	.0956	20	.9948	.7717	.7919
100+	.261	.468	[100-04]	404	.0446	21	-	.7677	.7878
NS	-	15.5	[105-09]	18	.0000	22	-	-	-
Total	53296	56849	-	-	-	-	-	-	-
Check	53295	56848	-	-	-	-	-	-	-

Notes: The computation proceeds from left to right and top to bottom. The following notes indicate the source or computation of the numbers in each column. Columns, exclusive of the label column, are numbered (1) through (9).

(1-2) Age data in first two columns from <u>Summary Results of the 1975 Population Census of Japan</u>, Table 17, pages 270-73. Since the number of age not stated cases in the second census is only 0.00027 of the total population, no proration adjustment is made. Numbers are given in thousands except for older age groups, where digits following the decimal are added to maintain three significant figures. The `total' row gives the total of the figures above it. The `check' row gives the total from the source table rounded to thousands of persons. Small discrepancies due to rounding error.

(3) s(i) = s_i is number in i-th age group in 1970 divided into number in (i+1)st age group in 1975. For the computation of the last survivorship ratio in the column, 0.0961, the 100+ age group is treated as though it were the 100-104 age group, i.e., as though there

were no persons aged 105 or over.

(4) Life table $_5L_x$ values from The 15-th Life Tables (Statistics and Information Department, Minister's Secretariat, Ministry of Health and Welfare) 13th, Life Table, 1970, pages 84-85, by differencing values in the T_x column (this is a complete life table with no five year age group values shown).

(5) Survivorship ratios from life table $_5L_x$ values. Note that these are given for the age groups 100-104 and 105-109.

(6) Index number of age group.

(7) Ratio of true to observed survivorship ratio for age group.

(8) Initial adjustment factors, obtained by setting a_1 = 1 and applying the formula $a_{i+1} = a_i(Ts_i/s_i)$. This calculation may be carried out after all the a_i have been calculated, in which case only four decimal places will be used, or it may be carried out age group by age group, in which all decimal places saved internally in the calculator will be used. The two methods give slightly different results due to cummulating rounding errors.

(9) Final adjustment factors, obtained by multiplying the initial adjustment factors by 1.026130. This constant is calculated as follows. Applying the initial adjustment factors to the first age distribution and summing over all age groups gives 51,998. We impute the adjustment factor for the next to last age group (95-99, 0.9948) to the the last age group (100+, treated here as 100-104). Dividing this into the observed total population at the first census gives 53,296/51,998 = 1.024956. Multiplying the initial adjustment factors by this value will give an adjusted age distribution for the first census for which the total over all ages exactly matches the observed total. The same calculation applied to the second census gives 56,849/55,338 = 1.027305, which factor yields adjustment factors that duplicated the total over all age groups for the second census. The geometric mean (square root of the product) of these two factors is 1.026130, which is applied to the initial adjustment factors to obtain the final adjustment factors.

The Census Survival Ratio Consistency Check Applied to the Total Population of Thailand, Censuses of 1960 and 1970

Age	AD1 1960 (1)	AD2 1970 (2)	ASGR ?1000 (3)	PYL (4)	SSR (5)	LLX (6)	LTSR (7)	LTSR /SSR (8)	IAF (9)	FAF (10)
0-4	4239	5659	29.08	4915	1.0820	4195	.9356	.8647	1.0000	1.0713
5-9	3992	5285	28.24	4608	.9704	3925	.9799	1.0098	.8647	.9264
10-14	3088	4562	39.28	3777	.9903	3846	.9782	.9878	.8732	.9355
15-19	2499	3718	39.99	3068	.9420	3762	.9710	1.0308	.8625	.9240
20-24	2416	2683	10.55	2547	.8861	3653	.9658	1.0899	.8890	.9524
25-29	2071	2241	7.941	2155	.9601	3528	.9609	1.0008	.9690	1.0381
30-34	1754	2124	19.27	1933	.9601	3390	.9563	.9960	.9698	1.0390
35-39	1372	1911	33.36	1627	.9574	3242	.9516	.9939	.9960	1.0349
40-44	1132	1541	31.05	1326	.9290	3085	.9452	1.0174	.9601	1.0286
45-49	976.7	1197	20.47	1083	.8975	2916	.9324	1.0389	.9769	1.0466
50-54	812.0	962.0	17.06	884.9	.8892	2719	.9110	1.0245	1.0149	1.0873
55-59	650.7	790.0	19.53	718.1	.8558	2477	.8761	1.0237	1.0397	1.1138
60-64	473.6	625.0	27.92	545.8	.8148	2170	.8240	1.0113	1.0644	1.1403
65-69	312.6	451.9	37.10	378.0	.7847	1788	.7506	.9565	1.0764	1.1532
70-74	197.9	296.9	40.83	244.1	.7028	1342	.6386	.9087	1.0296	1.1030
75-79	117.1	168.7	36.75	141.3	.5751	857.0	.4996	.8687	.9356	1.0023
80-84	61.98	78.40	23.66	69.87	.5196	428.2	.3447	.6634	.8128	.8708
85-89	25.06	37.59	40.82	30.90	.4569	147.6	.1925	.4213	.5392	.5777
90-94	10.92	13.39	20.53	12.11	.7956	28.41	.0817	-	.2272	.2434
95+	9.049	9.188	1.535	9.118	-	2.32	-	-	-	-
NS	46.09	15.5	-	-	-	-	-	-	-	-
Total	26258	34397	-	-	-	47500	-	-	-	-
Check	26257	34398	-	-	-	47502	-	-	-	-

Notes: Input required for the calculation consists of the two census age distributions shown in columns (1) and (2) and the life table ${}_5L_x$ shown in column 6. The numbers of persons with age not stated, if non-neglible, should be prorated (this has not been done here). The computation proceeds from left to right and top to bottom. The following notes indicate the source or computation of the numbers in each column. The age groups 0-4, 5-9, ..., are indexed by i = 1, 2, ...

(1-2) The census dates are April 25, 1960 (1960.315) and April 1, 1970 (1970.249). To compute the intercensal interval, express the two census dates in decimal form and subtract. To compute April 25 as a fraction of the year elapsed, add 25 to the number of days in each preceding month and divide by 365, (31+28+31+25)/365 = 0.315. Similarly, April 1 coresponds to 0.249. The intercensal interval is sthus 1970.249 - 1960.315 = 9.934 years.

(3) The average intercensal growth rate for the age group, denoted r_i and computed as $(1/t)\ln(y_i/x_i)$, where t denotes the length of the

intercensal interval. They are expressed here per thousand.

(4) Average person years lived during the intercensal period by the persons in the given age group, computed as $(y_i - x_i)/tr_i$.

(5) The synthetic survivorship ratio SSR_i, computed as

$$\frac{PYL_i \exp\{2.5?r_i\}}{PYL_{i+1} \exp\{-2.5?r_{i+1}}$$

In calculating the synthetic survivorship ratio for survival from 90-94 to 95-99, the open ended group 95+ is treated as though it were the group 95-99, i.e., we effectively assume no persons over 100.

(6) Model life table $_5L_x$ values from the Coale-Demeny West Female table with expectation of life at birth equal to 47.5 years. Page 47 of Ansley J. Coale and Paul Demeny, with Barbara Vaughan, Regional Model Life Tables and Stable Populations, 2nd Edition, Academic Press, New York, 1983.

(7) The life table survivorship ratio (LTSR) is computed from the $_5L_x$ values in the usual way.

(8) The ratio of the life table survivorship ratio (LTSR, column 7) to the synthetic survivorship ratio (SSR, column 5).

(9) The initial adjustment factor IAF_i is set equal to 1 for the 0-4 age group and calculated for subsequent age groups as the cummulative produce of the values in column (8). Calculating these products as the ratios in column (8) are obtained minimizes rounding errors, and this has been done in the calculation above.

(10) The final adjustment factors are computed by multiplying the initial adjustment factors by 1.0713, this value computed as follows. Apply the initial adjustment factors (IAF, column 9) to the first census age distribution (AD1, column 1) and sum over all age groups to obtain an implied total population of 24,534 (impute the adjustment factor for the 90-94 group to the 95+ group). The ratio of the total population for the first age distribution to this implied total, 1.0703, when applied to the initial adjustment factors, gives an adjusted age distribution summing to the enumerated value at the first census. The same calculation for the second census gives a ratio of 1.0724. The geometric mean (square root of the product) of these two factors equals 1.0713.

Note added 1997-08-24

No date on original, maybe 1992 when last offered my course.