

试论中国人口的死亡水平

冯 方 回

与世界其他国家相比,中国人口的死亡水平如何?第三次人口普查第一次公布了分年龄死亡人口的详细资料,使我们有可能对这个问题进行研究。本文就是利用第三次人口普查的死亡资料,对中国近代的死亡水平(确切地说为1981年的死亡水平)进行分析。模式生命表是研究死亡问题的重要工具。较为准确地搞清中国人口死亡属于何种模式生命表,就易于将中国的死亡水平与其它发展中国家相对比了。利用模式生命表可以研究那些死亡资料缺乏甚至严重缺乏的地区和国家的死亡水平,从而进行人口预测、人口统计等诸方面的估价。虽然中国第三次人口普查公布了较为详细的死亡资料,然而从整体看,死亡资料仍然是不健全的,因而利用模式生命表研究中国的人口同样具有十分重要的意义。

有关中国人口的死亡水平,已有不少中外学者进行了分析研究。美籍南朝鲜学者赵立济(Lee-Jay Cho)认为中国的死亡模式基本上属于科尔-德曼(Coale-Demeny)地区模式生命表(1983)中的北区。中国学者黄荣清和魏进认为,中国城市女性的死亡模式属于科尔-德曼模式生命表(1966)中的西区水平22,农村的女性则属于南区水平20。美国学者班尼斯特(Banister)则认为中国的死亡模式与任何现有的模式生命表都不相符。可见中外学者对此尚无较为一致的见解。本文也试图对这个问题进行探讨,提出一些看法。

一 模式生命表的选择

科尔-德曼的地区模式生命表(1966)为人口的研究提供了十分有力的工具,许多中外学者都曾运用它研究各式各样的人口问题。这是因为,人口死亡的年龄模式成为在这之前三代人的死亡和疾病的历史缩影。它反映了过去的水平、代组的年龄模型及疾病的变化趋势,其结果是恢复或死亡。在分析了发达国家过去150年的死亡经验资料后,安斯利·科尔(Ansley Coale)和保罗·德曼(Paul Demeny)给出了四个地区的死亡模式,相应于欧洲国家的四个地理位置。北区对应于北欧国家,南区对应于南欧国家,东区对应于东欧国家。西区则有些不同,它主要包含了西欧国家,同时还包含了祖先在西欧、而目前在海外居住的人口。

塞缪尔·普雷斯顿后来证明科尔-德曼生命表中的四个死亡模式分别对应着在每个地理区域中的不同的死因类型。他还划分了第五个“非西方”的模式。这个模式几乎包含了所有的拉丁美洲国家,并有着不同的死亡结构原因及由此导出的死亡年龄模型。

在60年代,由于缺少可靠的数据收集系统,阻碍了发展中国家死亡年龄模式的研究。然而70年代以后,发展中国家的人口抽样调查的数量大大增加,并且在生命登记体系中有十分显著的改善。另外,儿童死亡率的估算和按年龄和性别分的成人死亡登记完整性的估价等新的技术也已发展起来。这些技术通常与普查年龄分布和人口增长率联系在一起。这样就有可能详细地去分析发展中国家的死亡率年龄模型,以便区分出不同的年龄模型并建立相应于不同年龄模型的新的模式生命表。

建立模式生命表不仅仅是为了研究死亡率的年龄模型或疾病的过程。它们主要的目的是为了对那些资料有限的国家估计人口学的参数及准备人口预测。在资料缺乏的国家中,估算出生率或死亡率的许多最有用的方法都要依靠已知的或假设的死亡率年龄模型。进行人口预测则要求知道以年为基础的分年龄死亡率,以及死亡率是如何按照年龄和性别变化的假设。模式提供了死亡率变化的典型年龄模型的例子。目前最常使用的模式生命表是科尔-德曼的模式生命表,然而它几乎没有包括欧洲以外的人口。当前的证据表

明, 欧洲以外的死亡年龄模型不同于科尔和德曼的四个组。这种情况促使人口学家考虑需要建立一个新的模式生命表系列, 以更加适应发展中国家的人口学分析。联合国的生命表正是在这种考虑下出版的。虽然科尔和德曼在1983年出版了地区模式生命表第二版, 但与第一版相比没有重大的变化, 只是在老年组部分增加了两个年龄组, 从原来上限为80岁延伸到90岁。

在联合国模式生命表中陈述了发展中国家死亡率的模式年龄模型, 是第一本较为全面概括了全世界发展中国家的死亡模式的模式生命表系列。这些模型是在仔细分析并估价了按年龄、性别分的死亡资料之后建立起来的。联合国运用各种不同的人口学技术估价这些数据, 检查内部和外部的一致性, 剔除了那些可靠性不能得到证明的人口数据, 并且在那些准确的或者可以进行调整的数据中研究所有的年龄模型。当然, 这些可靠数据中构成的生命表不能代表所有发展中国家。然而, 除了撒哈拉非洲外, 几乎包括了所有主要地区和相对广泛的死亡水平资料。根据以上的对比, 笔者认为选择联合国的模式生命表分析中国的死亡水平更为适宜, 而以往的人口学家大都使用科尔-德曼的模式生命表分析中国的死亡水平, 因而得不到满意的结果。

二 联合国生命表简介

联合国的生命表按照性别和年龄分为5个群, 它们分别为拉丁美洲、智利、南亚、远东和一般5个大群。联合国从不同的发展中国家收集了72张生命表, 其中男性、女性各占一半。然后采用了两种统计方法和一种绘图方法对这72张生命表逐个进行分析和归类。两种统计方法分别是线性最优化剖面结构法 (linearly optimal profile construction) 和动态群组分析法 (dynamic clustering analysis)。绘图法则先计算比率 $R(x) = nq_x / nq_x^0$, 式中 nq_x 是在给定的生命表中年龄 x 的死亡概率, nq_x^0 是科尔和德曼西区模式生命表中, 在年龄10岁与给定生命表有相同预期寿命的相应水平在年龄 x 的死亡概率。然后将每张生命表以 $R(x)$ 为纵坐标, 年龄 x 为横坐标绘出图形, 并将有类似形状的生命表放在一起, 归为一类。所有的

表1 中国男性简略生命表 (1981年)

年龄组	死亡概率	尚存人数	死亡人数	平均生存人数	x 岁以上人 年数合计	x 岁的平均 预期寿命
Age	q_x	l_x	d_x	L_x	T_x	e_x^0
0	0.03321	100 000	3 321	97 596	6 612 528	66.13
1	0.01603	96 679	1 550	382 831	6 514 932	67.39
5	0.00649	95 130	618	473 901	6 132 101	64.46
10	0.00392	94 512	371	471 621	5 658 200	59.87
15	0.00540	94 141	509	369 545	5 186 579	55.09
20	0.00707	93 633	662	466 518	4 717 034	50.38
25	0.00729	92 971	678	463 163	4 250 517	45.72
30	0.00881	92 293	813	459 504	3 787 354	41.04
35	0.01215	91 480	1 112	454 758	3 327 849	36.38
40	0.01746	90 369	1 578	448 113	2 873 091	31.79
45	0.02664	88 790	2 365	438 422	2 424 978	27.31
50	0.04343	86 425	3 754	423 345	1 986 556	22.99
55	0.07087	82 672	5 859	399 707	1 563 211	18.91
60	0.11920	76 813	9 156	362 371	1 163 504	15.15
65	0.18383	67 656	12 437	308 670	801 133	11.84
70	0.29325	55 219	16 193	236 399	492 463	8.92
75	0.42335	39 026	16 522	153 370	256 064	6.56
80	0.62070	22 504	13 968	74 887	102 694	4.56
85+	1.00000	8 536	8 536	27 807	27 807	3.26

资料来源: 由作者根据1981年第三次全国人口普查数据计算并编制。

方法都得出了相同的群，即上面提到的5大类。出生预期寿命的范围从小于45岁到大于70岁。因此联合国生命表无论从地区性还是从预期寿命都有着广泛的代表性。

联合国模式生命表的一个特点是具有较大的灵活性，不仅可以利用印出的表格进行对照，而且还可以通过计算求出更为满意的结果。这是其它模式生命表所不具备的。

在联合国生命表的书中印出的所有生命表都是1-分量的生命表，出生预期寿命从35岁到75岁，间隔为1岁，因此每一个群按性别分各有41张表。每张生命表都由9个变量组成。

第1个变量是年龄。从5岁开始，年龄组的间隔是5年。最后一个年龄组是85岁及以上。从0岁到1岁的间隔是1年，从1岁到5岁的间隔是4年。

第2个变量是年龄别中心死亡率。一般的记号是 nmx ，其中 x 是在一个年龄段的起始年龄， n 是这个年龄段的长度，下同。

第3个变量是死亡概率，一般的记号是 nqx 。

第4个变量是存活人数（当基数是100 000时），一般的记号是 lx 。

第5个变量是死亡人数，一般的记号是 ndx 。

第6个变量是存活人年数，一般的记号是 Lx 。

第7个变量是累积存活人数，一般的记号是 Tx 。

第8个变量是预期寿命，一般的记号是 e_x^0 。

第9个变量是在该年龄段内死亡的人在该年龄段内的平均存活年数，一般的记号是 ax 。

该生命表的输入变量只有一个，就是年龄别中心死亡率。也就是说只要有了 nmx 值就可以算出其它变量值。

表2 中国女性简略生命表（1981年）

年龄组	死亡率	尚存人数	死亡人数	平均生存人数	x岁以上人年数合计	x岁的平均预期寿命
Age	q_x	l_x	d_x	L_x	T_x	e_x^0
0	0.03159	100 000	3 159	97 713	6 904 473	69.04
1	0.01781	96 841	1 724	383 025	6 806 760	70.29
5	0.00537	95 117	511	474 077	6 423 735	67.54
10	0.00327	94 606	310	472 268	5 949 658	62.89
15	0.00466	94 296	440	470 481	5 477 390	58.09
20	0.00667	93 856	626	467 748	5 006 909	53.35
25	0.00739	93 230	689	464 425	4 539 161	48.69
30	0.00842	92 541	779	460 804	4 074 736	44.03
35	0.01081	91 762	992	456 398	3 613 932	39.38
40	0.01448	90 770	1 314	450 722	3 157 534	34.79
45	0.02102	89 456	1 880	442 877	2 706 812	30.26
50	0.03282	87 576	2 874	431 087	2 263 935	25.85
55	0.05049	84 702	4 277	413 513	1 832 848	21.64
60	0.08417	80 425	6 769	386 122	1 419 335	17.65
65	0.13086	73 656	9 639	345 619	1 033 213	14.03
70	0.21645	64 017	13 856	286 655	687 594	10.74
75	0.32358	50 161	16 231	211 055	400 939	7.99
80	0.51359	33 930	17 426	124 708	189 884	5.60
85+	1.00000	16 504	16 504	65 176	65 176	3.95

资料来源：由作者根据1981年第三次全国人口普查数据计算并编制。

三 中国生命表的归类

联合国生命表的特点使得它十分适合估计中国的死亡水平。利用联合国生命表模式对中国的生命表加以分析将会得到较为满意的结果。根据1982年第三次人口普查的数据,首先编制出1981年的中国分男、女的简略生命表(见表1和表2),其年龄范围从0岁到85岁及以上。这两个表的出生平均预期寿命与中国人民大学编制的简略生命表的数值十分接近,仅在小数点之后有些差别。

由于中国的男、女生命表与联合国已列出的所有1-分量的生命表有较大的差距,所以作者采取了计算的方法。按照拉丁美洲、智利、南亚、远东和一般5个大群(分男女性别)分别同中国的男、女死亡模式加以对比,产生了有关生命表。在每张表中又分别计算了1-分量,2-分量和3-分量的生命表。还分别计算了平均平方误差值(mean square error),其值越小说明拟合程度越好。

这种计算采用了主分量法或最小二乘法。当生命表包含了从0~1岁到80~84岁组的全部数据时,主分量拟合(principal-component fit)和最小二乘拟合(least-squares fit)是一致的。

$$\text{主分量模型的定义如下: } {}_nY_x = {}_n\bar{Y}_x^s + \sum_{m=1}^k a_m U_{mx}$$

其中 ${}_nY_x$ 是从某个经验表格(即实际计算出来的)中得到的 $\logit[nq_x]$, ${}_n\bar{Y}_x^s$ 是联合国建立的五个模

表3A 计算中国1981年男性模拟一般模式的载荷因子值(a_m 值)

年龄 x	nq_x 值 中国男性	${}_nY_x$	${}_n\bar{Y}_x^s$ (一般模式)	${}_nY_x - {}_n\bar{Y}_x^s$	U_{1x}	U_{2x}	U_{3x}	$({}_nY_x - {}_n\bar{Y}_x^s)U_{1x}$		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
0	.03321	-1.68563	-1.27638	-.40925	.23686	-.46007	.09331	-.09694	.18828	-.03819
1	.01603	-2.05865	-1.78957	-.26908	.36077	-.68813	-.29269	-.09708	.18516	.07876
5	.00649	-2.51517	-2.35607	-.15910	.33445	.06414	-.47139	-.05321	-.01020	.07500
10	.00392	-2.76862	-2.55527	-.21335	.30540	.12479	-.17403	-.06516	-.02662	.03713
15	.00540	-2.60764	-2.34263	-.26501	.28931	.24384	.10715	-.07667	-.06462	-.02840
20	.00707	-2.47223	-2.16193	-.31030	.28678	.10713	.28842	-.08899	-.03324	-.08950
25	.00729	-2.45711	-2.09109	-.36602	.27950	.06507	.33620	-.10230	-.02382	-.12306
30	.00881	-2.36173	-2.00215	-.35958	.28023	.03339	.33692	-.10077	-.01201	-.12115
35	.01215	-2.19902	-1.86781	-.33121	.26073	.02833	.21354	-.08636	-.00938	-.07073
40	.01746	-2.01499	-1.70806	-.30693	.23626	.06473	.15269	-.07251	-.01987	-.04686
45	.02664	-1.79926	-1.52834	-.27092	.20794	.08705	.06569	-.05634	-.02358	-.01780
50	.04343	-1.54606	-1.33100	-.21506	.17804	.10620	.00045	-.03829	-.02284	-.00010
55	.07087	-1.28669	-1.12934	-.15735	.15136	.11305	-.03731	-.02382	-.01779	.00587
60	.11920	-.99999	-.91064	-.08935	.13217	.09467	-.10636	-.01181	-.00846	.00950
65	.18383	-.74532	-.68454	-.06078	.12243	.10809	-.11214	-.00744	-.00657	.00682
70	.29325	-.43983	-.45685	.01702	.11457	.14738	-.22258	.00195	.00251	-.00379
75	.42335	-.15451	-.23002	.07551	.10445	.21037	-.19631	.00789	.01588	-.01482
80	.62070	.24625	.00844	.23781	.08878	.30918	-.38123	.02111	.07353	-.09066

$$\alpha_1 = -.94672 \quad \alpha_2 = .18636 \quad \alpha_3 = -.43197$$

注: 根据方程(1a), 可得到1—2—和3—分量的拟合:

$$\alpha_1 = \alpha_1 = -.94672$$

$$\alpha_2 = \alpha_2 = .18636$$

$$\alpha_3 = \alpha_3 = -.43197$$

式中的一个的平均 $\text{logit}[nq_x]$ 值,或者是从另一个人口得到的 $\text{logit}[nq_x]$ 值,而该人口被假定与要测算的人口有着相似的死亡模式, a_m 是未知的分国家的因子载荷,它需要加以估算。 a_m 可以用最小二乘法加以估算,即使得经验的 nY_x 值与估计的值之间的平方偏差之和为最小。 U_{mx} 则是按年龄组分的主分量值,一般只用前3个主分量值就可以了。

假设有一个3-分量的拟合,要使以下函数的值为最小:
$$\sum_x [nY_x - n\bar{Y}_x^2 - \sum_{m=1}^3 a_m U_{mx}]^2$$

则 a_m 的最小二乘估计量可按以下步骤得到:

$$\left. \begin{aligned} \alpha_1 &= \frac{\alpha_1(\gamma_2\gamma_3 - \beta_3^2) + \alpha_2(\beta_2\beta_3 - \beta_1\gamma_3) + \alpha_3(\beta_1\beta_3 - \beta_1\gamma_2)}{D} \\ \alpha_2 &= \frac{\alpha_1(\beta_2\beta_3 - \beta_1\gamma_3) + \alpha_2(\gamma_1\gamma_3 - \beta_3^2) + \alpha_3(\beta_1\beta_2 - \beta_3\gamma_1)}{D} \\ \alpha_3 &= \frac{\alpha_1(\beta_1\beta_3 - \beta_2\gamma_2) + \alpha_2(\beta_1\beta_2 - \beta_3\gamma_1) + \alpha_3(\gamma_1\gamma_2 - \beta_1^2)}{D} \end{aligned} \right\}$$

表3B

用一般模式拟合1981年中国男性数据1-, 2-, 3-, 分量

年龄x	中国男性 观察值 nq_x	拟合 nq_x 值			平均平方误差		
		1分量	2分量	3分量	1分量	2分量	3分量
0	.03321	.04737	.04021	.03721	.12621	.03659	.01295
1	.01603	.01389	.01079	.01385	.02038	.15690	.02141
5	.00646	.00475	.00486	.00729	.09816	.08382	.01329
10	.00392	.00337	.00353	.00410	.02280	.01095	.00204
15	.00540	.00531	.00581	.00530	.04048	.05073	.00396
20	.00707	.00764	.00795	.00621	.00593	.01361	.01710
25	.00729	.00891	.00913	.00684	.04048	.05073	.00396
30	.00881	.01061	.01075	.00805	.03487	.03962	.00798
35	.01215	.01435	.01450	.01209	.02772	.03130	.00003
40	.01746	.02056	.02106	.01850	.02668	.03496	.00333
45	.02664	.03076	.03174	.03004	.02070	.03073	.01448
50	.04343	.04747	.04929	.04927	.00788	.01599	.01590
55	.07087	.07274	.07564	.07792	.00068	.00424	.00900
60	.11920	.11189	.11545	.12517	.00401	.00102	.00238
65	.18383	.16786	.17356	.18790	.00826	.00330	.00048
70	.29325	.24404	.25432	.29247	.03373	.02028	.00001
75	.42335	.34123	.35907	.39896	.04650	.02712	.00352
80	.62070	.46227	.49101	.57282	.08685	.05494	.00644
					.61218	.62140	.13466
					.03401	.03452	.00748

注: 1. 拟合 nq_x 值将根据以下的方程计算:

1分量拟合—— $\text{logit}(nq_x) = n\bar{Y}_x^s + a_1 U_{1x}$

2分量拟合—— $\text{logit}(nq_x) = n\bar{Y}_x^s + a_1 U_{1x} + a_2 U_{2x}$

3分量拟合—— $\text{logit}(nq_x) = n\bar{Y}_x^s + a_1 U_{1x} + a_2 U_{2x} + a_3 U_{3x}$

以上方程各项都可以从表A中得到。

2. 平均平方误差的计算公式为:
$$\frac{\sum (\ln(nq_x) - \ln(nq_x))^2}{18}$$
 其中 nq_x 为观察值, nq_x 为拟合值, 18是年龄组数

(共18个年龄组)。

$$\text{其中: } \gamma_i = \sum_{\text{U}_x} U_{ix}^2 \quad i=1,2,3; \quad \alpha_i = \sum_{\text{U}_x} ({}_nY_x - {}_n\bar{Y}_x^i) U_{ix} \quad i=1,2,3;$$

$$\beta_1 = \sum_{\text{U}_x} U_{1x} U_{2x}; \quad \beta_2 = \sum_{\text{U}_x} U_{1x} U_{3x}; \quad \beta_3 = \sum_{\text{U}_x} U_{2x} U_{3x};$$

$$\text{同时, } D = \gamma_1 \gamma_2 \gamma_3 - \gamma_3 \beta_1^2 - \gamma_2 \beta_2^2 - \gamma_1 \beta_3^2 + 2\beta_1 \beta_2 \beta_3$$

然而, 如果所有的18个年龄组 (5q₈₀) 都能用上, 则以上的方程还可以简化, 因为此时对所有的 i, β_i 都等于零, γ_i 都等于1。简化的方程为:

$$\left. \begin{aligned} \alpha_1 &= \alpha_1 = \sum_{\text{U}_x} ({}_nY_x - {}_n\bar{Y}_x^1) U_{1x} \\ \alpha_2 &= \alpha_2 = \sum_{\text{U}_x} ({}_nY_x - {}_n\bar{Y}_x^2) U_{2x} \\ \alpha_3 &= \alpha_3 = \sum_{\text{U}_x} ({}_nY_x - {}_n\bar{Y}_x^3) U_{3x} \end{aligned} \right\} \quad (1a)$$

表4A 计算中国1981年男性模拟拉美模式的载荷因子值 (a_m值)

年龄 x	nq _x 值 中国男性	${}_nY_x$	${}_n\bar{Y}_x^i$ (拉美模式)	${}_nY_x - {}_n\bar{Y}_x^i$	U_{1x}	U_{2x}	U_{3x}	$({}_nY_x - {}_n\bar{Y}_x^i) U_{1x}$		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
0	.03321	-1.68563	-1.12977	-.55586	.23686	-.46007	.09331	-.13166	.25574	-.05187
1	.01603	-2.05865	-1.49127	-.56738	.36077	-.68813	-.29269	-.20469	.39043	.16607
5	.00649	-2.51517	-2.13005	-.38512	.33445	-.06411	-.47139	-.12880	-.02470	.18154
10	.00392	-2.76862	-2.40748	-.36114	.30540	.12479	-.17403	-.11029	-.04507	.06285
15	.00540	-2.60764	-2.21892	-.38872	.28931	.24384	.10715	-.11246	-.09479	.04165
20	.00707	-2.47223	-2.01157	-.4606	.28678	.10713	.28842	-.13211	-.04935	-.13286
25	.00729	-2.45711	-1.93591	-.52120	.27950	.06507	.33620	-.14568	-.03391	-.17523
30	.00881	-2.36173	-1.86961	-.49212	.28023	.03339	.33692	-.13791	-.01643	-.16581
35	.01215	-2.19902	-1.76133	-.43769	.26073	.02833	.21354	-.11412	-.01240	-.09346
40	.01746	-2.01499	-1.64220	-.37279	.23626	.06473	.15269	-.08807	-.02413	-.05692
45	.02664	-1.79926	-1.49651	-.30275	.20794	.08705	.06569	-.06295	-.02635	-.01989
50	.04343	-1.54506	-1.34160	-.20446	.17804	.10620	.00045	-.03640	-.02171	-.00009
55	.07087	-1.28669	-1.15720	-.12949	.15136	.11305	-.03731	-.01960	-.01464	.00483
60	.11920	-.99999	-.96945	-.03054	.13217	.09467	-.10636	-.00404	-.00289	.00325
65	.18383	-.74532	-.74708	.00176	.12243	.10809	-.11214	.00022	.00019	-.00020
70	.29325	-.43983	-.52259	.08276	.11457	.14738	-.22258	.00948	.01220	-.01842
75	.42335	-.15451	-.29449	.13998	.10445	.21037	-.19631	.01462	.02945	-.02748
80	.62070	.24625	-.04031	.28656	.08878	.30918	-.38123	.02544	.08860	-.10925
								$\alpha_1 = -1.37903 \quad \alpha_2 = .41022 \quad \alpha_3 = -.47459$		

注: 根据方程(1a), 可得到1—2—和3—分量的拟合:

$$\alpha_1 = \alpha_1 = -1.37903$$

$$\alpha_2 = \alpha_2 = .41022$$

$$\alpha_3 = \alpha_3 = -.47459$$

由于中国的男女生命表(1981年)都有完整的年龄组(从0岁到80岁),所以在计算上可以用简化的方程(1a)。通过计算,得到表3至表1,其中表3、表4是中国男性生命表与联合国生命表的拟合,表5、表6则是中国女性生命表及其拟合。使用的算法和具体的方程都已在上面介绍。

通过对表3、表7的分析可以得出,中国男性的死亡模式与联合国一般模式最为接近,其中3分量的拟合最为理想。一般模式1分量、2分量和3分量的拟合 nq_x 值的平均平方误差分别为0.03401、0.03452和0.00748,可见3分量的平均平方误差最小(见表3B)。3分量的每个 nq_x 值与中国男性相比较,也都十分接近,最大的平均平方误差也只有0.0241(在1-4岁组),特别在老年组的拟合非常理想,这是其它模式所不具备的。

从对表5、表6的分析中可以得出,中国女性的死亡模式与联合国生命表的拉美模式最为接近,其中也是以3分量的拟合最为理想。因为在该模式的3个分量拟合中,平均平方误差分别为0.03131、0.00817和0.00759,也是以3分量的平均平方误差为最小(见表6B)。平均平方误差最大的年龄组也是1~4岁组,其值为0.02974。该模式的特点也是在老年组拟合的相当理想。

无论是男性数据还是女性数据,相对来说低年龄组的误差比其他年龄组的误差要大。

无论是男性数据还是女性数据,相对来说低年龄组的误差比其它年龄组的误差要大。造成这种情况的原

表4B 用拉美模式拟合1981年中国男性数据, 1-, 2-, 3-分量

年龄x	中国男性 观察值 nq_x	拟合 nq_x 值			平均平方误差		
		1分量	2分量	3分量	1分量	2分量	3分量
0	.03321	.05152	.03591	.03266	.19300	.00611	.00005
1	.01603	.01839	.01054	.01387	.01886	.17579	.02097
5	.00649	.00558	.00588	.00917	.02289	.00979	.11913
10	.00392	.00348	.00385	.00454	.01430	.00031	.02158
15	.00540	.00529	.00646	.00584	.00042	.03187	.00600
20	.00707	.00805	.00878	.00669	.01674	.04689	.00304
25	.00729	.00954	.01006	.00733	.07250	.10376	.00003
30	.00881	.01086	.01115	.00813	.04380	.05587	.00646
35	.01215	.01448	.01451	.01188	.02378	.03137	.00052
40	.01746	.01915	.02018	.01750	.00851	.02082	.00000
45	.02664	.02748	.02945	.02772	.00097	.01010	.00160
50	.04343	.04015	.04364	.04362	.00619	.00002	.00002
55	.07087	.06112	.06666	.06890	.02191	.00375	.00079
60	.11920	.09084	.09746	.10671	.07381	.04056	.01226
65	.18383	.13802	.14891	.16291	.08214	.04436	.01459
70	.29325	.20405	.22439	.26328	.13151	.07162	.01162
75	.24335	.29379	.33082	.37329	.13348	.06082	.01584
80	.62070	.41934	.48205	.75201	.15379	.06390	.00667
					1.01864	.77773	.24119
					.05659	.04321	.01340

注: 1. 拟合 nq_x 值将根据以下的方程计算:

$$1\text{分量拟合} \text{---} \logit(nq_x) = n\bar{Y}_x + a_1 U_{1x}$$

$$2\text{分量拟合} \text{---} \logit(nq_x) = n\bar{Y}_x + a_1 U_{1x} + a_2 U_{2x}$$

$$3\text{分量拟合} \text{---} \logit(nq_x) = n\bar{Y}_x + a_1 U_{1x} + a_2 U_{2x} + a_3 U_{3x}$$

以上方程各项都可以从表A中得到。

2. 平均平方误差的计算公式为: $\frac{\sum (\ln(nq_x) - \ln(nq_x))^2}{18}$ 其中 nq_x 为观察值, nq_x 为拟合值, 18是年龄组数(共18个年龄组)。

因很可能是由于低年龄组不易获得准确的资料所致。出乎意料的是,男女在老年组的拟合均相当理想。这说明普查时老年组的年龄相当准确,没有许多发展中国家存在的严重的年龄堆积现象。

总之,中国男性的死亡模式与联合国制定的模式生命表中的一般模式最为接近,而中国女性的死亡模式与拉美模式最为接近。两者都以3分量的拟合最佳。而其它模式生命表,如科尔-德曼地区模式生命表在分析中国的数据时都不如联合国模式生命表拟合得好。以上的结论证明,本文开始提出的“联合国模式生命表在分析发展中国家的数据时比其它模式生命表具有更好的代表性”的假设是正确的。然而十分遗憾的是,联合国在编制模式生命表时尚缺乏中国的资料。如果吸收了中国的资料,那么拟合可能会更理想,而且较低分量的拟合可能就会得到满意的结果,从而省去大量复杂的计算。

表5A 计算中国1981年女性模拟一般模式的载荷因子值(a_m 值)

年龄 x	nq_x 值 中国女性	nY_x	$n\bar{Y}_x^s$ (一般模式)	$nY_x - n\bar{Y}_x^s$	U_{1x}	U_{2x}	U_{3x}	$(nY_x - n\bar{Y}_x^s)U_{1x}$		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
0	.03159	-1.71140	-1.35963	-.35177	.18289	-.51009	.23944	-.06434	.17943	-.08423
1	.01781	-2.00515	-1.77385	-.23130	.31406	-.52241	-.11117	-.07264	.12083	.02571
5	.00537	-2.61050	-2.39574	-.21476	.31716	.08947	.07566	-.06811	-.01921	-.01625
10	.00327	-2.85917	-2.64549	-.21368	.30941	.03525	.06268	-.06611	-.00753	-.01339
15	.00466	-2.68192	-2.44766	-.23426	.32317	.03132	-.26708	-.07570	-.00734	-.06257
20	.00667	-2.50146	-2.28991	-.21155	.32626	.07843	-.39053	-.06902	-.01659	.08262
25	.00739	-2.45024	-2.18850	-.26174	.30801	.06762	-.28237	-.08062	-.01770	.07391
30	.00842	-2.38425	-2.08535	-.29890	.29047	.00482	-.14277	-.08682	-.00144	.04267
35	.01081	-2.25836	-1.97231	-.28605	.25933	-.01409	-.05923	-.07418	.00403	.01694
40	.01448	-2.11032	-1.84731	-.26301	.22187	-.02178	.18909	-.05835	.00573	-.04973
45	.02102	-1.92056	-1.69291	-.22765	.19241	.01870	.24773	-.04380	-.00426	-.05640
50	.03282	-1.69168	-1.50842	-.18326	.17244	.04427	.33679	-.03160	-.00811	-.06172
55	.05049	-1.46707	-1.30344	-.16363	.15729	.08201	.34121	-.02574	-.01342	-.05583
60	.08417	-1.19365	-1.08323	-.11028	.14282	.08061	.38290	-.01575	-.00889	-.04223
65	.13086	-.94669	-.84402	-.10267	.12711	.15756	.26731	-.01305	-.01618	-.02745
70	.21645	-.64324	-.59485	-.04839	.11815	.24236	.14442	-.00572	-.01173	-.00699
75	.32358	-.36868	-.34158	-.02710	.11591	.30038	.09697	-.00314	-.00817	-.00263
80	.51359	.02718	-.06493	.09211	.09772	.50530	-.13377	.00900	.04654	-.01232
								$\alpha_1 = -.84571$	$\alpha_2 = .21600$	$\alpha_3 = -.12475$

注: 根据方程(1a), 可得到1—2—和3—分量的拟合:

$$\alpha_1 = \alpha_1 = -.84571$$

$$\alpha_2 = \alpha_2 = .21600$$

$$\alpha_3 = \alpha_3 = -.12475$$

编辑说明: 本文原有12张表, 因版面有限, 我们删去其中6张表。它们是: 1A. “中国1981年男性模拟南亚模式的载荷因子值(a_m 值)”, 1B. “用南亚模式拟合1981年中国男性数据1-, 2-, 3-分量”; 2A. “中国1981年男性模拟远东模式的载荷因子值(a_m 值)”, 2B. “用远东模式拟合中国男性数据, 1-, 2-, 3-分量”; 3A. “中国1981年男性模拟智利模式的载荷因子值(a_m 值)”, 3B. “用智利模式拟合1981年中国男性数据; 1-, 2-, 3-分量”; 4A. “中国1981年女性模拟南亚模式的载荷因子值(a_m 值)”, 4B. “用南亚模式拟合1981年中国女性数据, 1-, 2-, 3-分量”; 5A. “中国1981年女性模拟远东模式的载荷因子值(a_m 值)”, 5B. “用远东模式拟合1981年中国女性数据, 1-, 2-, 3-分量”; 6A. “中国1981年女性模拟智利模式的载荷因子值(a_m 值)”, 6B. “用智利模式拟合1981年中国女性数据, 1-, 2-, 3-分量”。如有读者对此文感兴趣, 并想了解整个数据表, 可与作者联系。

参考书目:

- Banister, Judith: "An Analysis of Recent Data on the Population of China," Population and Development Review, Vol.10, No.2, June 1984: 241-71, 392-4 pp. New York, N.Y.
- Cho Lee-Jay; Han Jing Qing and Li-Bo Hua: Application of Own-children Method to the Household Data of the 1982 National Fertility Survey of the People's Republic of China: Paper Prepared for Presentation at the International Symposium on China's One-per-thousand Population Sample Fertility Survey, Beijing, China, October 14~18, 1985.
- Coale, A.J. and P. Demeny: Regional Model Life Tables and Stable Populations (Princeton, N.J., Princeton University Press, 1966).
- 冯方回: 《1981年女性死亡记录完整性估价》, 《人口动态》1989年第1期, pp.19~27, p.31.
- 黄荣清和魏进: 《一九八一年全国女性初婚表的制作与分析》, 《人口研究》1985年第5期, pp.54~57.
- preston, S.H., Mortality Patterns in National Populations (New York, Academic Press, 1976).
- United Nations, Model Life Tables for Developing Countries, New York, 1982.
- 中国人口情报研究中心: 《中国人口资料手册》(1987) pp.462~463. 1987年12月. 北京.

表5B 用一般模式拟合1981年中国女性数据, 1-, 2-, 3-分量

年龄x	中国女性 观察值 nq_x	拟合 nq_x 值			平均平方误差		
		1分量	2分量	3分量	1分量	2分量	3分量
0	.03159	.04615	.03736	.03527	.14368	.02818	.01216
1	.01781	.01664	.01333	.01370	.00455	.08395	.06884
5	.00537	.00483	.00502	.00493	.01132	.00461	.00752
10	.00327	.00298	.00302	.00297	.00915	.00648	.00923
15	.00466	.00431	.00437	.00467	.00605	.00413	.00000
20	.00667	.00587	.00607	.00669	.01637	.00889	.00001
25	.00739	.00741	.00762	.00818	.00001	.00099	.01028
30	.00842	.00936	.00938	.00972	.01114	.01158	.02042
35	.01081	.01233	.01226	.01244	.01740	.01585	.01974
40	.01448	.01679	.01664	.01588	.02202	.01936	.00860
45	.02102	.02386	.02405	.02264	.01612	.01818	.00555
50	.03282	.03528	.03594	.03314	.00523	.00823	.00009
55	.05049	.05351	.05533	.05105	.00337	.00838	.00012
60	.08417	.08256	.08524	.07808	.00037	.00016	.00564
65	.13086	.12977	.13765	.12992	.00007	.00256	.00005
70	.21645	.19948	.21673	.21067	.00666	.00000	.00073
75	.32358	.29334	.32104	.31579	.00693	.00006	.00059
80	.51359	.42674	.48080	.48913	.03431	.00435	.00238
					.31744	.22596	.17196
					.01764	.01255	.00955

注: 1. 拟合 nq_x 值将根据以下的方程计算:

$$1\text{分量拟合} \text{---} \logit(nq_x) = n\bar{Y}_x^3 + a_1 U_{1x}$$

$$2\text{分量拟合} \text{---} \logit(nq_x) = n\bar{Y}_x^3 + a_1 U_{1x} + a_2 U_{2x}$$

$$3\text{分量拟合} \text{---} \logit(nq_x) = n\bar{Y}_x^3 + a_1 U_{1x} + a_2 U_{2x} + a_3 U_{3x}$$

以上方程各项都可以从表A中得到。

2. 平均平方误差的计算公式为: $\frac{\sum (\ln(nq_x) - \ln(nq_x))^2}{18}$ 其中 nq_x 为观察值, nq_x 为拟合值, 18是年龄组数 (共18个年龄组)。

表6A

计算中国1981年女性模拟拉美模式的载荷因子值(a_m 值)

年龄 x	nq_x 值 中国女性	nY_x	$n\bar{Y}_x^*$ (拉美模式)	$nY_x - n\bar{Y}_x^*$	U_{1x}	U_{2x}	U_{3x}	$(nY_x - n\bar{Y}_x^*)U_{1x}$		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
0	.03159	-1.71140	-1.22452	-.48688	.18289	-.51009	.23944	-.08905	.24835	-.11658
1	.01781	-2.00515	-1.45667	-.54848	.31406	-.52241	-.11117	-.17226	.28653	.06097
5	.00537	-2.61050	-2.13881	-.47169	.31716	.08947	.07566	-.14960	-.04220	-.03569
10	.00327	-2.85917	-2.46676	-.39241	.30941	.03525	.05268	-.12142	-.01383	-.02460
15	.00466	-2.68192	-2.31810	-.36382	.32317	.03132	-.26708	-.11757	-.01139	.09717
20	.00667	-2.50146	-2.14505	-.35641	.32626	.07843	-.39053	-.11628	-.02795	.13919
25	.00739	-2.45024	-2.03883	-.41141	.30801	.06762	-.28237	-.12672	-.02782	.11617
30	.00842	-2.38425	-1.93924	-.44501	.29047	.00482	-.14277	-.12926	-.00214	.06353
35	.01081	-2.25836	-1.83147	-.42689	.25933	-.01409	-.05923	-.11071	.00601	.02528
40	.01448	-2.11032	-1.74288	-.36744	.22187	-.02178	.18909	-.08152	.00800	-.06948
45	.02102	-1.92056	-1.62385	-.29671	.19241	.01870	.24773	-.05709	-.00555	-.07351
50	.03282	-1.69168	-1.47924	-.21244	.17244	.04427	.33679	-.03663	-.00940	-.07155
55	.05049	-1.46707	-1.28721	-.17986	.15729	.08201	.34121	-.02829	-.01475	-.06137
60	.08417	-1.19351	-1.07443	-.11908	.14282	.08061	.38290	-.01701	-.00960	-.04560
65	.13086	-.94669	-.83152	-.11517	.12711	.15756	.26731	-.01464	-.01815	-.03079
70	.21645	-.64324	-.59239	-.05085	.11815	.24236	.14442	-.00601	-.01232	-.00734
75	.32358	-.36868	-.35970	-.00898	.11591	.30138	.09697	-.00104	-.00271	-.00087
80	.51359	-.02718	-.08623	-.11341	.09772	.50530	-.13377	.01108	-.05731	-.01517
									$a_1 = -1.36401$	$a_2 = -.40838$
									$a_3 = -.05022$	

注: 根据方程(1a), 可得到1—2—和3—分量的拟合

$$\alpha_1 = \alpha_1 = -1.36401$$

$$\alpha_2 = \alpha_2 = -.40838$$

$$\alpha_3 = \alpha_3 = -.05022$$

(本文责任编辑: 王跃生)

(作者工作单位: 中国人口情报中心)