Journal of Statistical
Theory & Methods
1997, Vol. 8, No. 2, pp. 153 ~ 161

한국인의 경험생명표 작성 및 통계적 해석

홍 연 웅[1]· 이 재 만 · 차 영 준[2]

요약

본 연구에서는 우리나라 6대 생명보험회사의 모든 남자의 개인보험 계약자를 대상으로 조사한 1988년부터 1992년까지 25,000,000건의 자료를 가중이동평균모형을 이용한 보간법과 곰페르츠모형을 이용한 보외법을 적용하여 경험생명표를 작성하였다. 적합성 및 평활성 기준에서 가중이동평균모형의 차분수와 항수, 곰페르츠모형의 모수 및 두 모형이 접촉하는 접점의 연령을 산출하였다. 특히 연립방정식을 이용하여 곰페르츠 모형의 모수를 폐쇄형으로 추정하는 방법을 제안하였다.

주제어:경험생명표, 조사망률, 보정된 사망률, 가중이동평균 모형, 곰페르츠 모형

1. 서론

경험생명표(experience life table)는 보험회사가 보험계약자를 대상으로 일정기간동안의 사망현상을 관찰하여 연령에 따른 사망률, 기대여명, 잔존수명 등을 알기 쉽게 나타낸 것으로 생명보험상품의 설계나 보험료 및 책임보험료의 책정 등에 사용된다. 경험생명표는 보험회사의 기본적인 통계자료로 작성결과에 따라 보험시장에서의 경쟁력확보와 재무구조에 결정적 영향을 미치는 인자로 이자율과 더불어 보험사업의 경영에 매우 중요한 역할을 하며 본격화될 보험 시장자율화에 따라 그 중요성은 더욱 커질 전망이다. 우리나라의 보험회사가 경험생명표를 적용한 것은 불과 10년전의 일로 한국보험계리인회가 작성한 제1회경험생명표(1988)와 보험개발원이 계리인회의 업무를 이관 받아 작성한 제2회경험생명표(1991) 및 제3회경험생명표(1996)가 있다. 이러한 생명표는 시간의 경과에 따른 사망률의 감소경향을 반영하기 위하여 수년 간격으로 새롭게 작성되는 완전생명표(complete life table)의 일종으로 6개의 기존 생명보험회사의 2년 이상의 모든 개인보험 가입자를 대상으로 작성하였다. 특히 이들 경험생명표의 작성을 주관한 기관은 상이하지만 산출과정은 동일하고, 우리 나라의 인구곡선에 대한 면밀한 분석이 선행되지 않은 상태에서 일본의 작성

[1]동양대학교 산업공학과
[2]안동대학교 통계학과

방법을 답습한 흔적을 지울 수 없다는 문제를 내포하고 있다. 따라서 우리 나라의 생명보험가입자 사망패턴에 가장 적합한 생명표를 작성할 수 있는 분위기의 조성이 필요하며 이러한 차원에서 본 연구는 계약건수(모집단의 크기)가 25,000,000건에 이르는 제3회 경험생명표의 남자의 기초자료를 이용하여 우리 나라 실정에 맞는 경험생명표의 작성을 위한 통계적인 해석을 하였다. 아울러 곰페르츠 모형을 따른다고 추정되는 연령대의 사망률을 세 개의 부분군으로 나누고 각 부분군에 대한 방정식을 이용하여 곰페르츠 모형의 모수를 폐쇄형으로 추정하는 방법을 제안하였다.

2. 경험생명표의 작성

우리나라의 경험생명표는 안정적인 영업기반을 확보하고 있으며 기초통계자료의 신뢰성이 높은 기존의 6대 생명보험회사의 모든 개인보험 계약자를 대상으로 보험회사별, 성별, 보험상품의 종류별, 계약당시의 진단유무별, 계약자에 대한 배당금지급유무 등으로 분류하여 각 연령(x)별 경과계약(n_x :표본크기와 유사한 개념)과 사망자수 (d_x)를 조사하고 이를 토대로 조사망률 ($u_x = d_x/n_x$: crude death rate)을 산출하는 것에서 출발한다. 이와 같은 조사망률은 대부분 평활(smooth)하지 않으므로 보험료의 계산 등에 불합리한 문제가 발생하며, 이를 해결하기 위하여 평활성과 적합성을 제고시키기 위한 보정(graduation)이 행하여진다.

가. 보정의 의의

Miller(1946), Elphinstone(1951), Nesbitt(1965), Kimeldorf & Jones(19-

67), Taylor(1992) 등이 보정을 정의하였는데 이를 종합하면 다음과 같이 간단히 표현할 수 있다. 보정이란 관측된 조사망률 u_x가 주어졌을 때 u_x와 적합(closely approximate)할 뿐만 아니라 자연적 또는 보험적 제약조건을 만족하는 평활한 사망률 v_x를 찾는 과정이다. 여기서 v_x를 보정된 사망률이라 한다. t_x를 x세의 모사망률(true death rate), e_x를 t_x의 추정오차를 나타내는 변수, $\triangle$ 를 전진차분연산자(forward difference operator)라고 하면

$$u_x = t_x + e_x \tag{1}$$

의 관계가 있다. G를 보정연산자(graduation operator)라 하고 (1)식의 양변에 취하면

$$\begin{aligned} G(u_x) &= G(t_x + e_x) \\ &= G(t_x) + G(e_x) \\ &= t_x + e'_x \end{aligned} \tag{2}$$

이다. 한편 $G(u_x) = v_x$이므로

$$v_x = t_x + e'_x \tag{3}$$

이다. (3)은 보정후의 관계이고 (1)은 보정전의 관계식이므로 보정후의 오차 e'_x는 보정전의 오차 e_x보다 작을 것으로 기대된다. 이러한 경우 v_x가 u_x보다 t_x에 더 가까우므로 보정된 사망률은 조사망률보다 좋은 추정치라고 할 수 있다. 그러므로 보정과정은 모집단의 사망률을 구하기 위하여 관측된 조사망률을 개정하는 통계적 추정과정으로 볼 수 있으며 이러한 과정이 일련의 연령에 대하여 체계적으로 수정되어지는데 그 특징이 있다. 여기서 체계적이라는 의미는 다음과 같다: 사망률이 평활하여야하고, 연령이 증가할수록 사망률이 증가하여야하며, 대체로 30세 이상에서는 사망률의 증가율이 증가하여야 한다. 이러한 측면에서 경험생명표의 보정문제는 다음의 두 가지 제약식을 만족하는 $t_i, i = 30, 31, 32, \cdots$의 동시추정문제이다.

$$(a) \quad t_{30} \le t_{31} \le t_{32} \le t_{33} \le \cdots$$

$$(b) \quad \Delta^2 t_x = t_{x+2} - 2t_{x+1} + t_x (\ge 0)$$

이러한 제약식은 30세 이하에 대해서도 성립될 수 있으나, 우리나라를 비롯한 많은 국가의 10세 이하의 저연령이나 사고로 인한 사망의 비율이 높은 20세 부근에서 성립되지 않는 경우가 있으므로 본 연구에서는 30세 이상으로 한정하였다.

나. 보정절차

조사망률을 이용하여 최종사망률을 산출하기까지 4차의 보정을 실시하는데 제1차보정이 Jenkins 5차 접촉보간법을 이용한 보정, 제2차보정이 통계적 위험론에 따른 할증, 제3차보정이 가중이동편균값을 이용한 보정, 제4차보정이 Gompertz 보외(extrapolation)이다. 제1차 보정은 5세 계급별 사망률을 계산하는데 사용되어지고, 제2차 보정은 보험재무적 측면에서 행하여지는 보정이므로 기술적인 의미는 있으나 통계적인 의미가 미약하므로 제3차 및 4차보정에 대하여 중점을 둔다.

1) 가중이동평균을 이용한 보정

가중이동평균치를 이용한 보정은 Greville보정이라고도 불리며 다음 식과 같이 표현된다.

$$v_x = \sum_{r=-n}^{n} a_r u_{x+r} \tag{4}$$

여기서 a_r은 t_x에 대한 v_x의 불편성과 오차의 가중합 $e'_x = \sum a_r e_{x+r}$을 최소로 하는 조건하에서 결정된다. 이를 구체화하면 첫째, $a_0 + 2\sum_{r=1}^{n} a_r = 1$, 둘째, $\sum_{r=1}^{n} r^2 a_r = 0$을 만족하고 셋째, $R_z^2 = \frac{z!z!}{2z!}\sum_{r=-n-z}^{n}(\Delta^z a_r)^2$를 최소화하는 것이다. 여기서 z 는 차분의 차수를 나타내며, z가 2, 3 또는 4일 때 평활성이나 적합성 측면에서 만족스런 결과가 얻어진다고 알려져있다. 앞서의 세 가지 조건을 동시에 만족시키는 가중치 a_r은, $z = 2$이면

$$a_r = \frac{3(3n^2 + 3n - 1) - 15r^2}{(2n-1)(2n+1)(2n+3)} \tag{5}$$

이고, $z=3$이면

$$a_r = \frac{315[(n+1)^2-r^2][(n+2)^2-r^2][(n+3)^2-r^2][2n^2+12n-4-11r^2]}{8(n+2)[(n+2)^2-1][4(n+2)^2-1][4(n+2)^2-9][4(n+2)^2-25]} \tag{6}$$

이다(London, 1985, pp. 42-45). 표1은 $z=3$ 일 때 a_r값을 나타낸 것이다. $r=0$ 일 때 a_r값이 가장 크며 r이 클수록 a_r 이 감소함을 알 수 있다. 가중이동평균값으로 x세의 사망률을 보정할 때 u_x에 가장 큰 비중을 주며, x세 이상이나 이하로 멀어질수록 가중치가 감소한다. 특히 말단값(end value)에는 음의 가중치를 주어 평활성이 개선되게 하는 역할을 한다. 식(4)에서 a_r이 (5) 또는 (6)의 계수를 가질 때 (4)를 z차분 $(2n+1)$항 가중이동평균식 이라 고한다.

< 표 1 > 3차분 $(2n+1)$항 가중이동편균식의 a_r값

	$2n+1=7$	9	11	13	15	17
r=0	.412588	.331140	.277944	.240058	.211542	.189232
1	.293706	.266557	.238693	.214337	.193742	.176390
2	.058741	.118470	.141268	.147360	.145904	.141112
3	-.058741	-.009873	.035723	.065492	.082918	.092293
4		-.040724	-.026792	.000000	.024028	.042093
5			-.019350	-.027864	-.014134	.002467
6				-.019350	-.024499	.018640
7					-.013730	-.020370
8						-.009961

이제 가중이동편균값을 이용한 보정의 적정항수를 결정하기 위하여 < 표 2 >의 조사망률에 대하여 식(4)로 보정한 결과를 저연령과 고연령으로 구분하여 나타내면 그림1 및 2와 같다.

< 표 2 > 3차분 $(2n+1)$항 가중이동평균모형의 적합성과 평활성

항수	7	9	11	13	15
적합도	165.70935	144.22476	188.60379	101.53020	104.11510
평활성	0.00000552	0.00000250	0.00000203	0.00000170	0.00000117

2) 모수적 외삽모형

모수적 외삽모형에는 곰페르츠모형, 와이불모형, 로지스틱모형 등이 있지만 본 연구에서는 우리나라에서 많이 사용하는 곰페르츠모형을 적용한다. 곰페르츠모형은 로지스틱모

형과 더불어 초기발전단계, 중간성숙단계,최종포화단계로 성장형태를 보이는 현상을 설명하는데 많이 사용되는 모형으로 이들 방법의 핵심은 최종포화단계에서의 크기를 주고 관찰치를 연령(시간대)에 따라 가장 적절히 설명하는 모수를 찾는 것이다. 곰페르츠모형의 생존함수 $S(x)$는 다음과 같다.

$$S(x) = \frac{K}{\exp(1+\exp(\alpha+\beta x))} \tag{7}$$

이와 같은 곰페르츠모형의 변곡점은 K/e 이고 상한과 하한은 각각 K와 0이다. 곰페르츠모형이 로지스틱 모형과 다른 점은 변곡점을 중심으로 비대칭적인 형태를 갖는다는 점이다. 그림(1)의 연령에 따른 사망률 추이를 보면 모형의 형태가 대칭인 로지스틱모형보다는 비대칭인 곰페르츠 모형이 고연령사망률의 예측에 유용하다고 판단되므로 곰페르츠모형으로 사망률예측을 하도록 한다. 식(7)에서K값의 결정에 개입될 주관성을 일부라도 배제하기 위하여 $\beta = \ln c, K = ek, \alpha = \ln(\ln 1/g)$라고 치환하면 (7)식은

$$S(x) = kg^{c^x} \tag{8}$$

로 다시쓸 수 있다. 식(8)의 모수를 추정하기 위하여 양변에 대수를 취하고 곰페르츠 모형을 따르는 연령구간인 $y+1$세부터 $y+3n$세까지를 $(y+1, y+n), (y+n, y+2n), (y+2n, y+3n)$의 세 그룹으로 나누고 각 그룹에 대하여 합하면 다음과 같다.

$$\begin{aligned} s_1 &= n\ln k + c^{y+1}(c^n-1)/(c-1)\ln g \\ s_2 &= n\ln k + c^{n+y+1}(c^n-1)/(c-1)\ln g \\ s_3 &= n\ln k + c^{2n+y+1}(c^n-1)/(c-1)\ln g \end{aligned} \tag{9}$$

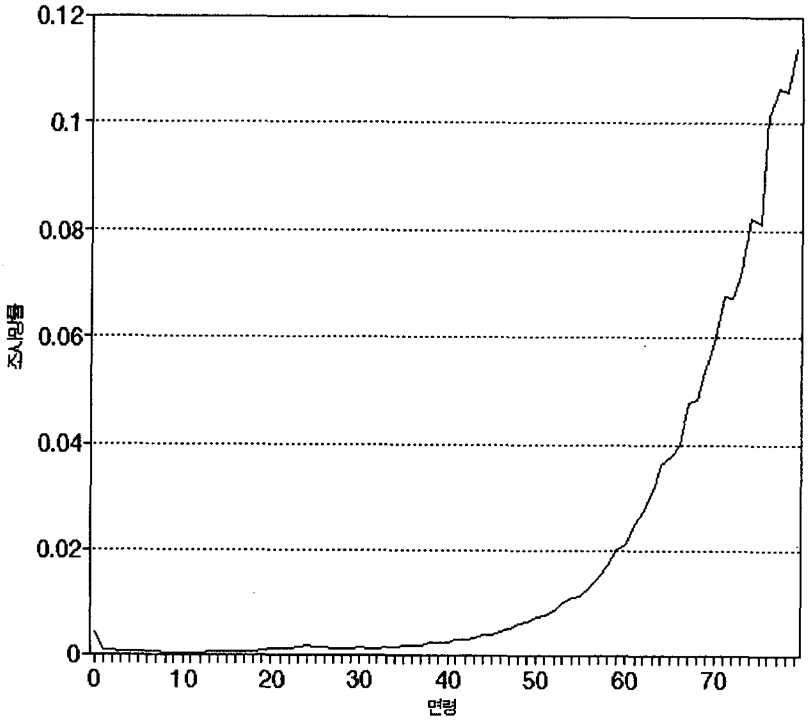
$s_1 = \sum_{i=1}^n \ln(S(y+i))$, $s_2 = \sum_{i=1}^n \ln(S(y+n+i))$, $s_3 = \sum_{i=1}^n \ln(S(y+2n+i))$이다. 위의 방정식을 연립하여 c와 g를 구하면

$$c = \left(\frac{s_3-s_2}{s_2-s_1}\right)^{\frac{1}{n}}, \quad g = \exp\left\{\frac{(c-1)(s_3-s_2)}{c^{n+y+1}(c^n-1)^2}\right\} \tag{10}$$

이고, k는 (9)의 세가지 식을 합한 것에 대입하여 정리하면 다음과 같다.

$$k = \exp\left\{[s_1+s_2+s_3-c^{y+1}(c^{3n}-1)/(c-1)](3n)^{-1}\ln g\right\} \tag{11}$$

한편 위의 방법은 y와 n값의 선택방법이나, (9)식을 연립하는 방법 등에 따라서 k, g 및 c값이 다를 수 있으나 대체로 그 차이는 지극히 작은 편이며, 간혹 유의한 차이를 보이는 경우 적합성과 평활성을 고려하여 결정한다. 본 연구에서는 (y = 51세, n = 5세) 및 (y = 46세, n = 10세)의 두 가지 경우에 모형(8)의 모수를 추정한 결과 ($c = 1.132197, g = 0.999989, k = 0.999150$)와 ($c = 1.120855, g = 0.999120, k = 0.999220$) 으로 나타났으나 ($y$ = 51세, n = 5세)인 경우의 추정치를 이용한 모형의 적합성과 평활성이 우수한 것으로 나타났다. 곰페르츠모형을 이용한 보정된 사망률은 표3에서 70세 이후의 사망률이다.

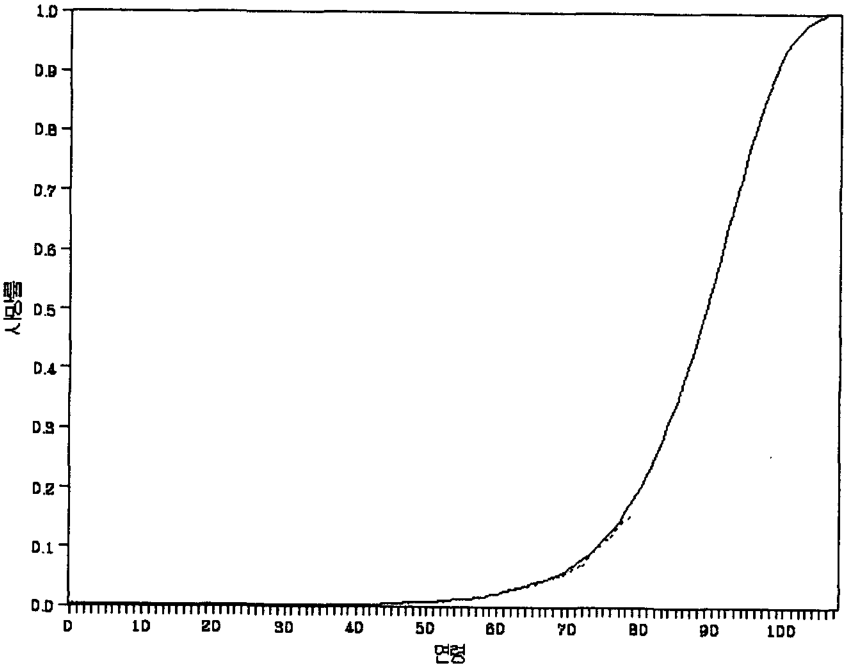


<그림 1> 조사망률(1988-1992, 남자)

3) 사망률의 접속 가중이동평균을 이용한 보정결과와 모수적 외삽모형을 이용한 보정(예측) 결과가 중복이 되는 연령구간이 발생하게 된다. 즉 특정 연령을 경계로 두 가지 모형을 연결할 접점을 결정해야하는데 일반적으로 x세를 경계로 상하 d세까지의 두 가지 모형에 의한 보정된 사망률의 절대차의 합이 최소가 되는 기준을 사용한다. 본 연구에서는 $d = 5$세로 했을 때 $x = 70$이 됨을 알 수 있었다. 즉, 두 모형의 접속연령이 70세 이다. 이상의 과정을 통하여 산출된 경험생명표의 사망률은 표 3 및 그림 2와 같다.

<표 3> 연령별 조사망률 및 보정된 사망률(남자, 1988-1992)

연령	조사망률	보정된 사망률	연령	조사망률	보정된 사망률	연령	보정된 사망률
0	0.00431	0.00151	40	0.00247	0.00258	80	0.20360
1	0.00081	0.00156	41	0.00287	0.00285	81	0.22712
2	0.00084	0.00131	42	0.00298	0.00315	82	0.25291
3	0.00075	0.00092	43	0.00351	0.00347	83	0.28108
4	0.00067	0.00062	44	0.00384	0.00383	84	0.31169
5	0.00062	0.00054	45	0.00399	0.00425	85	0.34478
6	0.00052	0.00053	46	0.00467	0.00474	86	0.38033
7	0.00041	0.00046	47	0.00515	0.00530	87	0.41825
8	0.00039	0.00039	48	0.00594	0.00592	88	0.45840
9	0.00037	0.00032	49	0.00642	0.00658	89	0.50051
10	0.00027	0.00029	50	0.00725	0.00730	90	0.54426
11	0.00021	0.00030	51	0.00782	0.00810	91	0.58918
12	0.00029	0.00037	52	0.00856	0.00896	92	0.63472
13	0.00052	0.00049	53	0.01011	0.00985	93	0.68022
14	0.00063	0.00060	54	0.01087	0.01079	94	0.72493
15	0.00076	0.00066	55	0.01123	0.01194	95	0.76805
16	0.00073	0.00068	56	0.01297	0.01342	96	0.80877
17	0.00060	0.00069	57	0.01475	0.01526	97	0.84632
18	0.00059	0.00073	58	0.01716	0.01744	98	0.88001
19	0.00085	0.00083	59	0.02038	0.01983	99	0.90933
20	0.00103	0.00098	60	0.02109	0.02239	100	0.93398
21	0.00107	0.00115	61	0.02510	0.02529	101	0.95390
22	0.00127	0.00132	62	0.02742	0.28044	102	0.96930
23	0.00148	0.00145	63	0.03104	0.03176	103	0.98063
24	0.00158	0.00152	64	0.03611	0.03527	104	0.98850
25	0.00151	0.00154	65	0.03773	0.03879	105	0.99363
26	0.00149	0.00149	66	0.04123	0.04239	106	0.99673
27	0.00136	0.00143	67	0.04776	0.04636	107	0.99847
28	0.00134	0.00138	68	0.04838	0.05096		
29	0.00128	0.00135	69	0.05457	0.05704		
30	0.00138	0.00135	70	0.05982	0.06422		
31	0.00136	0.00136	71	0.06749	0.07229		
32	0.00134	0.00140	72	0.07546	0.08135		
33	0.00141	0.00147	73	0.07243	0.09149		
34	0.00151	0.00156	74	0.08245	0.10284		
35	0.00173	0.00169	75	0.08119	0.11552		
36	0.00181	0.00184	76	0.10135	0.12966		
37	0.00192	0.00216	77	0.10664	0.14540		
38	0.00217	0.00216	78	0.10594	0.16287		
39	0.00231	0.00235	79	0.11420	0.18222		



<그림 2 > 조사망률 및 보정된 사망률(1988-1992, 남자)

참고문헌

1. 보험개발원 (1992). 제2회 경험생명표

2. 보험개발원 (1996). 제3회 경험생명표작성결과 보고서

3. Elphinstone, M. D. W. (1951). Summation and some other methods of graduation: The foundation of theory, TSA, XX.

4. Kimeldorf, G. S. and Jonse, D. A. (1967). Bayesian graduation, TSA, XIX

5. Miller, M. D. (1946). *Elements of Graduation,* ASA.

6. Nesbitt, D. (1965). *A least squares methods for determining the Makeham constants,* M. S. Thesis, Northeastern Univ.

7. Taylor, G. (1992). A Bayesian interpretation of Whittaker-Henderson Graduation, *Mathematics and Economics,* 11, 7-16

Construction of Korean Experiance Life Table

Yeon Woong Hong [3] · Jae Mann Lee · Young Jun Cha [4]

Abstract

A Korean experience life table(male) is constructed by using a mixture of weighted moving average(WMA) model and Gompertz' parametric survival model based on 25,000,000 insured of major 6 life insurance companies from 1988 to 1992. The graduated values are taken as those which minimize the composite measure of fittness and smoothness. Moreover, we propose closed form estimators for three parameters of Gompertz' model.

[3]Department of Induatrial Engineering, Dongyang University
[4]Department of Statistics, Andong University