

# Gompertz사망함수에 의한 고연령 사망률의 추정 절차

## A Process of Old Age Mortality Estimation Using the Gompertz Mortality Function

백 철\* (Cheol Baik)

### <목 차>

- I. 서론
- II. 선행 연구
- III. 추정 절차
- IV. 결론

### <국문초록>

본 연구는 고연령 사망률 추정을 위해 사용되고 있는 사망함수 중 Gompertz사망함수의 모수추정법을 고찰하였다. 실증자료를 통해 관찰값 중에서 모수추정을 위한 적절한 표본 선택과, 선택된 표본에 따른 여러 가지 추정함수 중 최적의 함수를 선정하는 기준을 제시하였다. 또한, 관찰값에 의해 추정된 사망률과 Gompertz사망함수에 의해 추정된 사망률의 접촉연령을 선정하는 기준을 제시하였다.

적정한 모수추정을 위해서는 우선 Gompertz사망함수의 가정사항을 만족하는 조건의 표본 조합을 택해야하며, 이 표본의 사망성향이 Gompertz사망함수로 적절하게 설명되고

---

\* 보험개발원 통계서비스실, e-mail : cbaik@kidi.or.kr

논문투고일 : 2012. 3. 15 논문수정일 : 2012. 6. 10 게재확정일 : 2012. 6. 15

있는지 여부를 검토해야한다. 표본의 선택에 따라 다양한 함수가 도출될 수 있는데 평균  
제곱오차의 변동을 분석하여 이 중 최선의 함수를 선택한다. 선택된 함수는 관찰값에 의  
한 사망률 추정값과 가장 오차가 작은 지점에서 접촉하는 것이 합리적이다.

국문색인어 : 고펜르츠, 사망함수, 생명표, 사망률, 고연령

## I. 서론

생명표 작성시 연령별 사망률을 추정하기 위해서는 해당 위험집단의 생존과 사망에 대한 관찰을 통해 기초사망률을 산출하고, 주변 연령과의 평활성(smoothness) 및 기초사망률과의 적합성(fitness)을 고려한 보정(gradation)과정을 거친다. 또한, 생존과 사망의 관찰값이 충분하지 않거나 부재한 고연령구간의 사망률은 통계가 관찰되는 인접 연령구간의 기초통계를 이용하여 모수적추정을 시행하는 것이 일반적이다.

고연령구간 사망률 추정에 이용되는 사망함수로는 Gompertz, Makeham, Perks, Beard, Kannisto 등이 제안한 사력(the force of mortality)가정이 통상 활용되고 있으며, 우리나라 생명보험가입자의 사망률을 추정한 생명보험 경험생명표에서는 이중 Gompertz 또는 Makeham 사망함수에 의해 고연령 사망률을 추정해왔다.

그런데, 일반적으로 생명표 작성시 사망함수의 모수추정을 통해 고연령의 사망률을 산출하는 과정에서는 활용하는 기초통계에 따라 도출되는 모수추정값이 다양하게 나타나고, 이 중 최선의 모수를 선택하는 기준이 뚜렷하지 않다. 따라서, 최종적인 사망률을 확정하여 생명표를 완성하기 위해서 반복적인 계산을 통해 대량의 함수식을 도출하고 사전에 설정된 특정한 계량적 기준을 통해 최종모수를 선택하는, 산출자의 실무적 경험과 계리적 판단을 요하는 주관적 의사결정 과정을 거쳐왔다.

본 연구에서는 Gompertz사망함수의 모수를 최소제곱법(least square method)에 의해 추정하고 이를 통해 고연령 사망률을 산출하는 절차를 검토하였다. Gompertz 사망함수는 사력은 연령이 증가함에 따라 지수적으로 증가한다는 합리적인 가정이 반영되어 있고, 로그변환시 선형결합의 형태로 나타낼 수 있는 상대적 편의성이 있으며, 동 함수에 의해 사망률을 추정하고 있는 해외사례 등을 고려한 것이다.

한편 그동안의 국내외 생명표 작성사례에서는 고연령 사망률 산출에 적용된 절차에 관한 개괄적인 방법론은 제시되었으나 실증자료에 의한 세부적인 추정절차는 거의 제시되지 않았다. 본 연구에서는 기초통계에 의해 다양하게 산출된 Gompertz 사망함수 중 최선의 함수를 선정하고 이에 따라 실제 고연령 사망률을 추정하는 절차를 실증연구를 통해 제시한다.

## II. 선행 연구

Gompertz는 연령  $x$ 에서의 사력(the force of mortality)을 다음과 같이 나타내었다.

$$\mu_x = BC^x \quad B > 0, \quad C > 1 \quad (2-1)$$

이에 대해, 생존확률  $p_x$ 는 사력  $\mu_x$ 의 정의에 따라 다음과 같이 유도된다.

$$\begin{aligned} p_x &= e^{-\int_x^{x+1} \mu_t dt} = e^{-\int_x^{x+1} BC^t dt} \\ &= e^{\frac{-BC^x(C-1)}{\ln C}} \end{aligned} \quad (2-2)$$

이를 사망률  $q_x$ 로 표현하면,

$$\begin{aligned} q_x &= 1 - p_x = 1 - e^{\frac{-BC^x(C-1)}{\ln C}} \\ &= 1 - g^{C^x(C-1)} \end{aligned} \quad (2-3)$$

이고, 여기서  $g = e^{-\frac{B}{\ln C}}$ 이다.

### 1. King-Hardy방법에 의한 Gompertz사망함수의 모수추정

모수  $B$ 와  $C$ 를 찾는 방법으로 보험개발원(2005) 등에서는 King-Hardy 방법을 이용하여 왔으며 이는 다음과 같은 절차에 의해 유도되었다.

(2-4)로부터 (2-1)을 적용하면, (2-5)를 얻을 수 있다.

$$l_x = l_0 \cdot \exp\left(-\int_0^x \mu_y dy\right) \quad (2-4)$$

$$\begin{aligned}
l_x &= l_0 \cdot \exp\left(-\int_0^x B \cdot C^y dy\right) = l_0 \cdot \exp\left(-B \cdot \frac{C^y}{\ln C}\right) \Bigg|_0^x \\
&= l_0 \cdot \exp\left\{-\frac{B}{\ln C}(C^x - 1)\right\}
\end{aligned} \tag{2-5}$$

여기서  $-\frac{B}{\ln C} = \ln g$  라 하고 (2-5)에 적용하면, 다음과 같다.

$$l_x = l_0 \cdot g^{C^x - 1} = k \cdot g^{C^x}, \text{ 단, } k = \frac{l_0}{g} \tag{2-6}$$

(2-6)의 양변에 로그를 취하면, 다음을 얻는다.

$$\ln l_x = \ln k + C^x \cdot \ln g \tag{2-7}$$

연령  $x$ 로부터  $n$ 개 연령씩 3개군  $\{x, x+1, \dots, x+n-1\}$ ,  $\{x+n, x+n+1, \dots, x+2n-1\}$ ,  $\{x+2n, x+2n+1, \dots, x+3n-1\}$ 로 나누고 각 군에 대하여 합하면 다음과 같은 상수를 구할 수 있다.

$$\alpha = \sum_{i=0}^{n-1} \ln l_{x+i} = n \cdot \ln k + \frac{C^x(C^n - 1)}{C - 1} \cdot \ln g \tag{2-8}$$

$$\beta = \sum_{i=0}^{n-1} \ln l_{x+n+i} = n \cdot \ln k + \frac{C^{x+n}(C^n - 1)}{C - 1} \cdot \ln g \tag{2-9}$$

$$\gamma = \sum_{i=0}^{n-1} \ln l_{x+2n+i} = n \cdot \ln k + \frac{C^{x+2n}(C^n - 1)}{C - 1} \cdot \ln g \tag{2-10}$$

(2-8), (2-9) 및 (2-10)을 이용하여  $C$ 와  $g$ 를 구하면 각각 다음과 같다.

$$C = \left(\frac{\gamma - \beta}{\beta - \alpha}\right)^{1/n} \tag{2-11}$$

$$g = \exp\left[\frac{C - 1}{C^x(C^n - 1)^2} \cdot (\beta - \alpha)\right] \tag{2-12}$$

King-Hardy방법에 의해 함수식을 추정할 때는  $x$ 로부터  $n$ 개씩 3개군의 관찰값을 사용하여 여러 가지의 함수식을 도출하고, 그 중 적합성과 평활성, 최종연령 등을 고려하여 실제 생명표의 고연령 사망률 추정에 적용할 함수식을 채택한다.

모수를 추정하는 과정 중 (2-8), (2-9), (2-10)에서와 같이, 추정에 사용되기 시작하는 관찰값의 연령  $x$ , 하나의 모수를 추정하기 위해 사용하는 관찰값의 개수  $n$ 의 결정에 따라 매우 다양한 함수가 도출될 수 있다.

## 2. 최소제곱법에 의한 Gompertz 사망함수의 모수 추정

Gompertz사망함수의 모수를 추정하는 또다른 방법으로 최소제곱법을 이용할 수 있다.

(2-2)에 로그를 취하여 다음의 식을 유도할 수 있다.

$$\ln p_x = \frac{-BC^x(C-1)}{\ln C} \quad (2-13)$$

(2-13)으로부터 양변에 로그를 취한  $\ln(-\ln p_x)$ 는 다음과 같이 상수와 연령  $x$ 에 의한 선형결합함수로 표현된다.

$$\begin{aligned} \ln(-\ln p_x) &= \ln B + \ln(C-1) - \ln(\ln C) + x \ln C \\ &= \alpha + \beta x \end{aligned} \quad (2-14)$$

즉,  $\alpha = \ln B + \ln(C-1) - \ln(\ln C)$ ,  $\beta = \ln C$  인 단순선형회귀모형으로 볼 수 있다.

(2-14)로부터, 최소제곱추정법에 의하여,  $\hat{\alpha}$ 와  $\hat{\beta}$ 에 대해 다음을 최소화하는 값을 추정할 수 있다.

$$\min_{\hat{\alpha}, \hat{\beta}} \sum_x \{ \ln(-\ln p_x) - \hat{\alpha} - \hat{\beta}x \}^2 \quad (2-15)$$

(2-15)를 다음과 같이  $\hat{\alpha}$ 과  $\hat{\beta}$ 로 각각 편미분하면 다음과 같다.

$$\frac{\partial \sum \{\ln(-\ln p_x) - \hat{\alpha} - \hat{\beta}x\}^2}{\partial \hat{\alpha}} = -2 \sum \{\ln(-\ln p_x) - \hat{\alpha} - \hat{\beta}x\} \quad (2-16)$$

$$\frac{\partial \sum \{\ln(-\ln p_x) - \hat{\alpha} - \hat{\beta}x\}^2}{\partial \hat{\beta}} = -2 \sum x \cdot \{\ln(-\ln p_x) - \hat{\alpha} - \hat{\beta}x\} \quad (2-17)$$

(2-16)과 (2-17)을 0으로하는 추정값  $\hat{\alpha}$ 과  $\hat{\beta}$ 를 구하면, 이로부터 선형회귀모형의 정규방정식(normal equation)이 유도되어 다음과 같이 모수를 추정할 수 있다.

$$\hat{\beta} = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2}, \quad y = \ln(-\ln p_x) \quad (2-18)$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \quad (2-19)$$

(2-18)과 (2-19)의 결과값을 (2-14)에 각각 대입하면 최소제곱법에 의해 Gompertz 함수의 모수 추정값을 산출할 수 있다.

최소제곱법에 의한 모수추정과정은 King-Hardy방법과 마찬가지로 어떤 관찰값들  $(x, p_x)$ 을 이용하여 모수를 추정하느냐에 따라 그 함수가 다양하게 도출되므로, 사전에 이들 중 최선의 함수를 선정하는 기준을 설정하는 것이 중요하다.

### 3. Gompertz사망함수 적용에 대한 적합성 검토

표본에 의한 고연령 사망률 추정을 위해서는 우선 표본의 관찰값 등이 Gompertz 사망함수를 따르는가에 대한 적합성을 검토해야한다. 즉, 일종의 분포에 대한 적합성 검정을 시행함으로써 해당 집단의 사망성향이 Gompertz가 제시한 가정과 부합하는 경우는 일련의 추정을 시행하지만, 표본의 사망현상이 Gompertz사망함수로 설명하기 어렵다면 서론에서 언급한 다른 사망함수를 이용하여 추정해야할 것이다.

Gompertz사망함수의 적합성을 검정하는 절차로 Yue(2002)는 다음과 같은 방법을 제시하였다.

$x$ 세의 생존율  $p_x$ 는 각 연령의 기초통계에 의한 생존율의 불편추정량이며 정규분포근사의 가정하에서 일정한 신뢰구간을 추정을 할 수 있다.

(2-13)으로부터 다음을 얻을 수 있으며,

$$\frac{\ln p_{x+1}}{\ln p_x} = C \quad (2-20)$$

95% 신뢰수준에서  $p_x$  는 다음과 같이 하한값( $p_x^L$ )과 상한값( $p_x^U$ )을 갖는다.

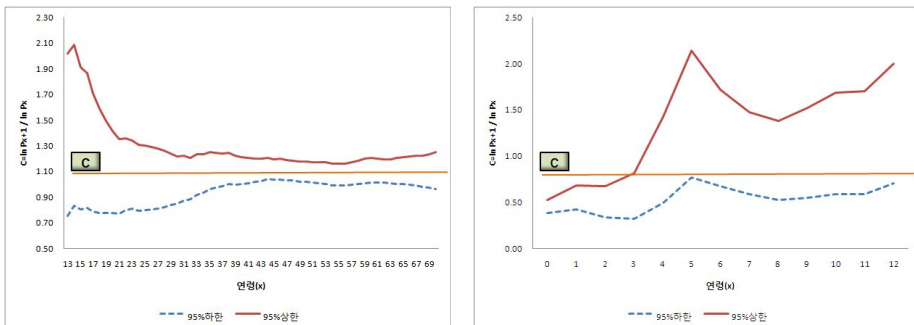
$$p_x^L = p_x - 1.96 \sqrt{\frac{p_x(1-p_x)}{N_x}}, \quad p_x^U = p_x + 1.96 \sqrt{\frac{p_x(1-p_x)}{N_x}} \quad (2-21)$$

이를 (2-20)에 적용하면 각 연령  $x$ 에서의 상수값  $C$ 의 신뢰구간의 상한과 하한을 산출할 수 있다.

모수추정을 위해 사용하는 표본연령구간내에서  $C$ 의 신뢰구간 상한의 최소값이 신뢰구간 하한의 최대값보다 크다면, 즉, 모든 연령에 걸쳐 신뢰구간 범위내에서 공통으로 적용할 수 있는  $C$ 가 존재한다면, 이는 동 표본이 Gompertz사망함수에 의해 설명이 가능하다 볼 수 있다. 한편, 이러한  $C$ 가 존재할 수 없다면 이는 해당 사망현상이 Gompertz사망함수로 설명할 수 없음을 의미하므로 서론에서 제시한 다른 사망함수를 이용하거나 비모수적방법에 의해 고연령 사망률을 추정하는 것이 바람직하다.

<그림 2-1>은  $C$ 가 존재하는 연령구간과 그렇지 않은 연령구간의 예를 나타낸 것이다. 왼쪽의 그림은 해당 연령구간의 사망현상을 설명할 수 있는 Gompertz사망함수의 모수  $C$ 가 존재함을 보이지만, 오른쪽 그림은 모수  $C$ 가 해당연령 구간의 사망성향을 설명할 수 없다는 것을 나타낸다.

<그림 2-1> 예 : 연령구간별  $C$ 의 상하한





### Ⅲ. 추정 절차

본 연구에서는 고연령 사망률 추정절차를 실증적으로 제시하기 위해 2006~2008 관찰년도의 생명보험 기초통계자료 중 남자의 사망통계를 이용하였다. 그러나, 동 추정절차는 사망현상을 관찰한 다른 다양한 기초통계자료에도 동일하게 적용할 수 있을 것이다.

관찰값이 충분한 연령구간은 기초통계에 의해 기초사망률을 산출하고 보정작업(smoother)을 시행한다. 우리나라 통계청의 생명표 및 경험생명표에서는 Greville의 3차  $n$ 항식에 의한 평활작업을 시행하는데, Greville보정법은 일종의 가중이동평균법으로서 인접연령간의 기초사망률을 이용하여 특정연령에서 나타날 수 있는 기초사망률의 이상치를 제거 또는 완화하는 과정이다.

본 연구에서는 관찰값의 기초사망률에 Greville 3차 13항식을 시행한 결과값을 기초값으로 설정하였다.

또한, Gompertz사망함수의 모수추정을 위한 절차는 기존까지의 경험생명표 등에서 적용한 King-Hardy방법 대신 최소제곱법에 의한 추정을 시행하였는데, 이는 추정절차의 합리성, 표본선정의 용이성, 반복적인 계산의 간소성 등을 고려한 선택이다.

#### 1. 표본의 선정

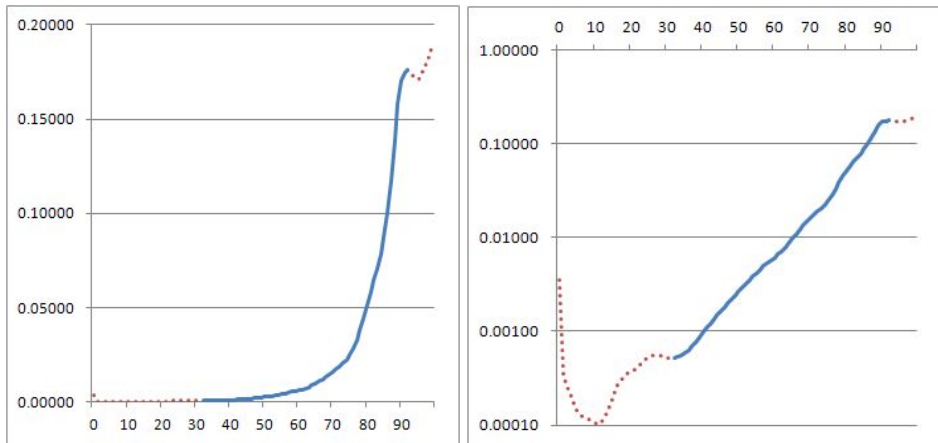
Gompertz사망함수의 기본가정은 연령이 증가함에 따라 사력이 증가한다는 것인데 영아사망이 영향을 미치는 저연령구간이나 재해사망이 많은 20대 등의 연령 등에서는 연령의 증가따라 사망률이 증가하지 않을 수 있다. 이는 동 연령구간의 사망성향이 Gompertz사망함수에 의해 설명되지 않는 구간이라 할 수 있다.

<그림 3-1>은 관찰값의 연령별 기초사망률이다. 왼쪽의 그림에서와 같이 연령별 사망률은 연령이 증가에 따라 대체로 우상향하는 형태가 나타나는데, 저연령과 고연령의 사망률 편차가 크기 때문에 사망률의 눈금간격을 조정하여 오른쪽과 같이 다시 나타내었다. 그림에서 볼 수 있듯이 실제 자료에서 30세 이전의 연령구간에서는 사망률이 단조증가하지 않는 현상이 보인다. 이는 앞서 언급한 원인이 실제 자료에서도 나타난 것으로 볼 수 있다. 고연령에서도 일부 단조증가하지 않는 구간이 나타나는데, 이는 고연령자의 생존계약 및 사망에 대한 관찰건수가 충분하지 않아

기초율이 불규칙하게 나타나는 현상으로 볼 수 있다.

따라서, Gompertz사망함수를 적용하기 위해서는 함수의 기본가정이 만족하는 연령구간, 즉 사망률이 연령에 따라 단조증가하는 연령구간의 기초통계를 표본으로 선택하는 것이 합리적이라 하겠다. 그러므로, 0~99세의 관찰값 중 저연령과 고연령구간(점선)을 제외한 32세~92세의 자료(실선)를 일차적으로 표본으로 선택할 수 있다.

<그림 3-1> 연령별 기초사망률

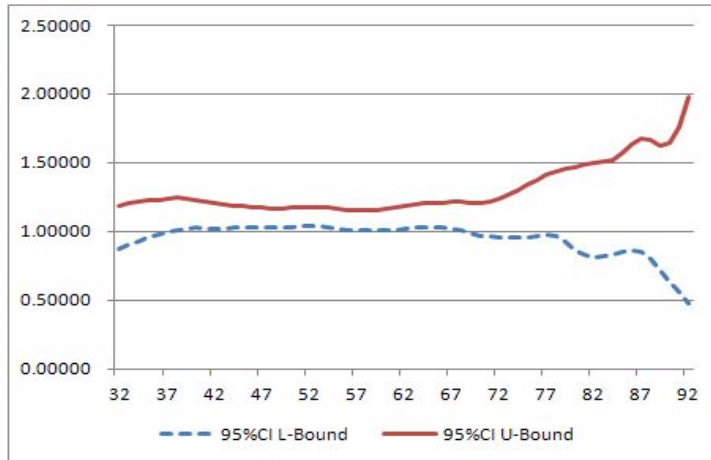


## 2. Gompertz사망함수의 유의성 검토

일차적으로 표본으로 선정된 관찰값이 Gompertz사망함수로 설명가능한지 여부를 검토하고, 만일 적합하지 않은 것으로 판단될 경우에는 다른 사망함수를 이용하거나 다른 추정방법을 적용해야할 것이다.

(2.21)을 관찰값에 적용하면 <그림 3-2>와 같은 결과를 얻을 수 있다. 주요 연령에 대한 값은 <표 3-1>에 표기하였다.

<그림 3-2> 상수값  $C$ 의 신뢰구간 상하한(32~92세)



<표 3-1> 주요 연령의  $C$ 의 신뢰구간 상하한(32~92세)

주요연령	35	45	52	55	58	65	75	85
하한	0.951	1.029	<u>1.037</u>	1.018	1.003	1.032	0.955	0.848
상한	1.222	1.180	1.174	1.162	<u>1.149</u>	1.208	1.339	1.572

이는 Gompertz사망함수의 모수인  $C$ 의 추정값은 연령에 따라 상한 및 하한의 범위가 각기 달라지지만, 상한과 하한의 값이 서로 겹치지 않는 것을 볼 때, 기초통계의 연령 전체에 걸쳐 적용할 수 있는 어떠한  $C$ 가 존재한다는 것을 의미한다. 즉, 관찰값의 사망현상을 Gompertz사망함수 가정으로 설명할 수 있다고 볼 수 있다.

즉, 표에서와 같이  $C \in [1.037, 1.149]$ 로 추정할 수 있다.

### 3. 각 함수식의 모수추정값 산출

Gompertz사망함수로 설명가능한 연령구간의 관찰값의 조합에 의해, (2-18) 및 (2-19)와 같이 최소제곱법에 의한 모수를 추정할 수 있다.

그런데, 모수추정값은 어떠한 관찰값을 사용하느냐에 따라 각기 다르게 산출될 수

있다. 예를 들면, 32세~92세 통계를 이용하여 추정한 모수와 33세~92세 통계를 이용하여 추정한 모수는 다르며, 33세~92세 통계를 이용하여 추정한 모수와 34~92세 통계를 이용하여 추정한 모수가 각각 다르다. 고연령의 사망률을 추정하는 절차이므로 관찰값 중에서 고연령구간의 자료를 최대한 모수추정에 사용하는 것이 합리적일 것이다.

추정에 사용될 관찰값을 확정하기 위하여 우선 각 관찰값을 다양하게 사용한 함수식을 각기 산출하였다. 실증자료에서는 32~92세의 대상 관찰값이 있으므로, 고연령을 최대한 포함하는 제약조건하에서는 60종류(=92-32)의 Gompertz사망함수식이 도출된다.

추정에 사용한 기초통계에 따른 Gompertz사망함수의 모수추정결과는 <표3-2>와 같다. 편의상 각 함수에는 추정사용연령의 시작점을 id로 부여하였다. 즉 32세~92세의 관찰치로 추정한 함수식 id는 <32>로 표기하였다.

#### 4. 최적 함수의 선택

고연령 사망률 추정을 위해서는 <표3-2>의 여러 가지 함수식 중 최적이라고 판단되는 함수를 선택해야 할 것이다. 다시 말하면, 고연령 사망률을 추정에 사용할 표본을 확정해야하는 것이다.

많은 관찰값을 사용(낮은 연령에서부터 추정에 사용)하는 경우는 모수추정을 위한 정보를 많이 얻을 수 있어 안정적인 함수식을 얻을 수 있으나, 고연령 사망성향과 연관성이 낮은 저연령의 관찰값이 포함될 수 있고, 실제값과 추정에 의한 값의 적합도가 떨어지는 단점이 있다.

반면, 너무 적은 관찰값을 사용(높은 연령에서부터 추정에 사용)하는 경우는 실제값과 추정값의 적합도가 높고 고연령 사망성향이 추정에 많이 반영되지만, 관찰된 정보가 충분히 활용되지 못하며 특정 연령의 이상치(outlier)에 영향을 받을 수 있다. 따라서, 이 중 최선의 함수를 선택하여 생명표를 작성하는 것이 중요하다.

추정한 각 함수의 오차의 정도를 평가하기 위하여 평균제곱오차(mean square error)를 산출하였다. 평균제곱오차는 실제 기초사망률과 추정값의 편차의 제곱합의 크기를 기준으로 하였는데, 이를 함수추정에 사용된 연령갯수로 나눔으로서, 1개 연령의 평균제곱오차가 기준이 되게 하였다.

각 함수별 평균제곱오차를 계산하고(<표3-3>) 이를 그림으로 나타내면(<그림 3-3>) 다음과 같다.

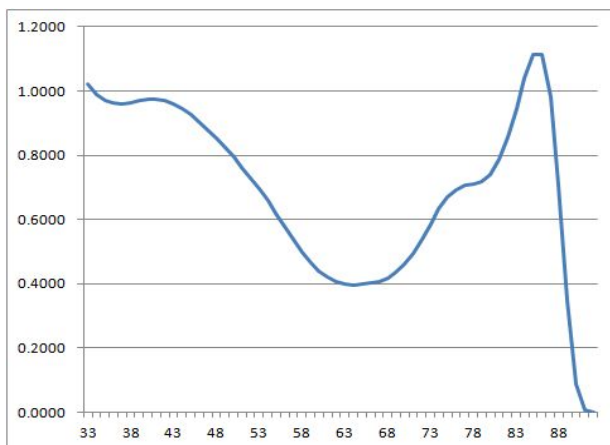
<표 3-2> 사용된 관찰값에 따른 Gompertz함수 모수추정값

id	$\hat{C}$	$\hat{B}$ $\times 10E-05$	id	$\hat{C}$	$\hat{B}$ $\times 10E-05$	id	$\hat{C}$	$\hat{B}$ $\times 10E-05$
32	1.1055	1.5957	52	1.1130	0.9499	72	1.1344	0.1949
33	1.1059	1.5515	53	1.1138	0.8971	73	1.1352	0.1842
34	1.1062	1.5195	54	1.1147	0.8418	74	1.1354	0.1817
35	1.1065	1.4971	55	1.1156	0.7845	75	1.1348	0.1893
36	1.1066	1.4812	56	1.1167	0.7269	76	1.1336	0.2075
37	1.1067	1.4693	57	1.1178	0.6710	77	1.1319	0.2377
38	1.1068	1.4587	58	1.1189	0.6186	78	1.1299	0.2777
39	1.1070	1.4465	59	1.1200	0.5711	79	1.1282	0.3155
40	1.1071	1.4301	60	1.1211	0.5292	80	1.1276	0.3325
41	1.1074	1.4082	61	1.1220	0.4931	81	1.1279	0.3230
42	1.1076	1.3806	62	1.1229	0.4619	82	1.1287	0.3052
43	1.1080	1.3484	63	1.1237	0.4346	83	1.1287	0.3045
44	1.1084	1.3125	64	1.1246	0.4091	84	1.1267	0.3581
45	1.1088	1.2741	65	1.1254	0.3836	85	1.1209	0.5668
46	1.1093	1.2333	66	1.1264	0.3568	86	1.1092	1.4679
47	1.1098	1.1902	67	1.1275	0.3281	87	1.0905	6.8993
48	1.1103	1.1449	68	1.1288	0.2980	88	1.0652	58.64
49	1.1109	1.0980	69	1.1303	0.2672	89	1.0382	611.71
50	1.1116	1.0500	70	1.1318	0.2382	90	1.0171	4028.90
51	1.1122	1.0007	71	1.1332	0.2133	91	1.0060	11124.19

<표 3-3> 함수별 평균제곱오차

id	MSE	id	MSE	id	MSE
32	1.0218	52	0.6946	72	0.5863
33	0.9889	53	0.6578	73	0.6321
34	0.9698	54	0.6186	74	0.6697
35	0.9614	55	0.5779	75	0.6936
36	0.9606	56	0.5374	76	0.7053
37	0.9646	57	0.4995	77	0.7097
38	0.9705	58	0.4665	78	0.7167
39	0.9751	59	0.4398	79	0.7392
40	0.9757	60	0.4200	80	0.7856
41	0.9708	61	0.4069	81	0.8567
42	0.9602	62	0.3996	82	0.9465
43	0.9449	63	0.3972	83	1.0417
44	0.9259	64	0.3981	84	1.1128
45	0.9042	65	0.4018	85	1.1149
46	0.8801	66	0.4085	86	0.9829
47	0.8537	67	0.4191	87	0.7065
48	0.8249	68	0.4355	88	0.3480
49	0.7943	69	0.4602	89	0.0874
50	0.7625	70	0.4947	90	0.0067
51	0.7294	71	0.5382	91	0.0000

<그림 3-3> 함수별 평균제곱오차



저연령의 표본값까지 사용하여 모수추정에 사용하는 기초통계가 많으면(함수id <32>에 가까운 함수들) 활용하는 정보가 많으므로, 추정한 사망률의 신뢰도가 높아지고 기초통계의 이상치에 대해 추정결과가 로버스트(robust)하다. 그러나, <그림 3-3>에서와 같이 평균제곱오차는 비교적 크게 나타난다. 또한, 고연령대의 통계만을 사용(함수id <91>에 가까운 함수들)하면 평균제곱오차가 작지만 특정 연령의 기초통계의 영향을 크게 받는다. 또한, 고연령대의 통계만을 이용하는 것이 항상 작은 평균제곱오차를 갖는 것은 아님을 실증자료는 나타내고 있다.

그러므로, 평균제곱오차가 적으면서 충분한 관찰값을 이용하여 산출한 함수를 선택하는 것이 합리적이라 할 것이다.

<그림 3-3>에서 평균제곱오차는 함수id <63> 및 <88> 이후에서 작게 나타난다. 따라서, 위의 실증자료에서는 작은 평균제곱오차를 가지면서 많은 표본을 이용하여 산출한 함수id <63>식을 채택한다.

즉, 고연령구간의 사력함수는 다음과 같다.

$$\mu_x = BC^x = 0.4346 \times 10^{-5} \cdot 1.1237^x \quad (3-1)$$

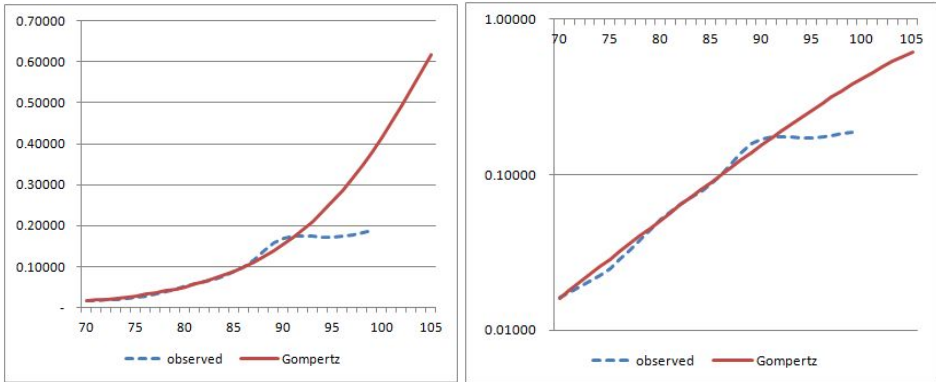
## 5. 접촉 연령 설정

관찰값에 의해 추정된 사망률과 Gompertz사망함수에 의해 추정된 사망률의 비교는 <그림 3-4>와 같다. 동일한 그래프를 사망률 눈금의 스케일을 달리하여 두 가지로 나타내었다.

접촉연령은 관찰값에 의해 추정된 사망률 구간에서 Gompertz사망함수에 의해 추정된 사망률 구간으로 추정값이 변경되는 연령을 말한다. 위의 그림에서 관찰값과 사망함수로부터 추정된 값의 차이가 가장 작은 지점을 접촉연령으로 택한다. 이는 두 사망률 구간이 접촉되는 연령에서 불연속성을 최소화하기 위함이다. 즉, 다음을 가장 작게하는 연령  $x$ 를 접촉연령으로 택하는 것이 합리적이다.

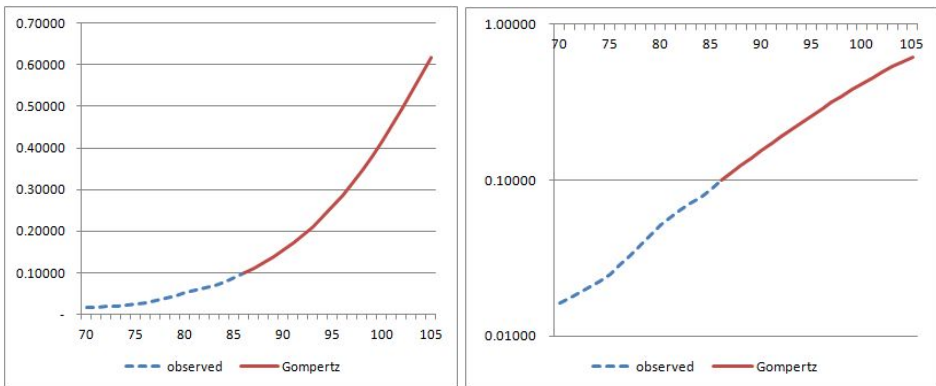
$$\frac{|q_x^{obs} - q_x^{Gompertz}|}{q_x^{obs}} \quad (3-2)$$

<그림 3-4> 관찰값과 보외값의 비교, 70세이후



위 실증자료에서는 86세가 접촉연령이 되며, 접촉연령 전후의 사망률 추정값을 이어서 그림으로 나타내면 <그림 3-5>와 같다. 관찰값과 사망함수에 의한 추정값이 연속적으로 연결되고 있음을 알 수 있다.

<그림 3-5> 접촉연령 전후의 사망률 추정값, 70세이후





## IV. 결론

본 연구에서는 Gompertz사망함수에 의해 고연령 사망률을 추정하는 절차를 실증적으로 제시하였다.

함수의 모수추정을 위해서는 우선 연령이 증가할수록 사망률이 단조증가한다는 Gompertz사망함수의 기본가정을 만족하는 연령의 관찰값들을 선택해야할 것이다. 또한 동 관찰값들이 Gompertz사망함수에 의해 설명될 수 있는지 적합성을 확인하고 난 후 모수추정작업을 시행함이 타당하다.

선택된 관찰값을 이용하여 Gompertz사망함수의 모수를 추정하는 방법은 여러 가지가 있을 수 있으나, 본 연구에서는 최소제곱법에 의한 산출방법과 산출결과를 제시하였다. 표본선택의 조합에 따라 추정모수와 추정함수식은 다양하게 도출될 수 있다. 이 중 최선의 함수를 선정하기위하여 평균제곱오차를 측정하였는데, 평균제곱오차가 작으면서도 가능한 많은 정보를 이용한 함수를 선택하는 것이 합리적이다. 선택된 함수에 의한 사망률과 관찰값에 의해 산출된 사망률은 두 사망률의 차이가 가장 적은 연령을 접촉연령으로 선정하는 것이 합리적이다.

## 참고문헌

- 백철, “경험생명표의 최종연령 확장 검토”, 생명보험 상품조사 제4호, 보험개발원, 2011.
- 오창수, 「최신보험수리학」, 박영사, 1995.
- 보험개발원, 「제5회 경험생명표」, 2005.
- 통계청, “생명표 작성방법”, [www.kostat.go.kr](http://www.kostat.go.kr)
- Thatcher, A. R., “The Force of Mortality at ages 80 to 120”, *Odense Monograph on Population Aging*, 1998.
- Yue, J. C., “Oldest-Old Mortality Rates and the Gompertz Law : A Theoretical and Empirical Study Based on Four Countries”, *Living to 100 and Beyond Symposium*, 2002.
- Doray, L. G., “Inference for Logistic-type Models for the Force of Mortality”, *Living to 100 and Beyond Symposium*, 2008.
- Society of Actuaries, “Report of the Individual Life Insurance Valuation Mortality Task Force”, 2001.

### <Abstract>

This study reviews the parameter estimation method of the Gompertz mortality function. The Gompertz mortality function is one of various models that can estimate old age mortality.

Through the use of empirical data, this paper suggests sample selection criterion for appropriate parameter estimation as well as function selection criterion from selected samples.

Additionally, this study puts forward the criterion for setting the age where the estimated mortality value based on observed data and the Gompertz mortality function come into contact.

For appropriate parameter estimation, first, a sample mix satisfying the assumptions of the Gompertz function must be selected. Then, it must be investigated whether the mortality characteristics of the selected sample can be explained reasonably by the Gompertz mortality function.

A wide range of functions can be derived depending on the combination of sampling observations. The best function among those results is selected upon analyzing the mean square error.

The mortality estimate from the selected function is a reasonable fit when it meets the estimate based on observed data where the difference is the smallest.

*Key words* : Gompertz, Mortality Function, Mortality Table