

# A modified Lee-Carter model based on the projection of the skewness of the mortality

Hangsuck Lee<sup>a,1</sup> · Changryong Baek<sup>b</sup> · Jihyeon Kim<sup>c</sup>

<sup>a</sup>Department of Actuarial Science/Mathematics, Sungkyunkwan University;

<sup>b</sup>Department of Statistics, Sungkyunkwan University;

<sup>c</sup>Long-term Insurance Team, MG Non-Life Insurance

(Received October 27, 2015; Revised December 17, 2015; Accepted December 17, 2015)

---

## Abstract

There have been continuous improvements in human life expectancy. Life expectancy is as a key factor in an aging population and can wreak severe damage on the financial integrity of pension providers. Hence, the projection of the accurate future mortality is a critical point to prevent possible losses to pension providers. However, improvements in future mortality would be overestimated by a typical mortality projection method using the Lee-Carter model since it underestimates the mortality index  $\kappa_t$ . This paper suggests a mortality projection based on the projection of the skewness of the mortality versus the typical mortality projection of the Lee-Carter model based on the projection of the mortality index,  $\kappa_t$ . The paper shows how to indirectly estimate future  $\kappa_t$  trend with the skewness of the mortality and compares the results under each estimation method of the mortality index,  $\kappa_t$ . The analysis of the results shows that mortality projection based on the skewness presents less improved mortality at an elderly ages than the original projection.

Keywords: Lee-Carter model, Longevity risk, the projection of future mortality, the mortality index, skewness

---

## 1. 서론

우리나라는 2000년에 65세 이상 인구 비율이 7% 이상인 고령화 사회에 진입하였으며, 2026년 65세 이상의 인구 비율이 20%를 넘어서는 초고령화 사회에 진입할 것으로 전망된다. 지속적인 사망률 개선으로 인한 수명연장은 낮은 출산율과 함께 우리나라 인구 고령화의 주요인으로 지목되며, 사망률 개선 추세의 지속은 국민연금 등 공적 연금 공급자뿐만 아니라 생명보험 회사와 같은 사적 연금 공급자에게도 리스크 요인으로 부각되고 있다. 사망률의 개선은 연금의 실제 지급기간이 예상 지급기간보다 길어지게 함으로써 연금 공급자의 재정 건전성에 심각한 영향을 미칠 수 있기 때문이다. 또한, 유럽에서 논의 중인 Solvency II는 현행 RBC에서 다루지 않는 장수리스크를 연금보험 판매업자에게 있어 가장 중요한 보험리스크로 보고 별도의 모듈로 지정하고 있다. 따라서 사망률의 개선 추세가 적절히 반영된 사망률

---

This research was supported by the Basic Science Research Program from the National Research Foundation of Korea (NRF), funded by the Ministry of Science, ICT & Future Planning (NRF-2014R1A1A1006025).

<sup>1</sup>Corresponding author: Department of Actuarial Science/Mathematics, Sungkyunkwan University, 25-2, Sungkyunkwan-ro, Jongno-gu, Seoul 03063, Korea. E-mail: [hangsuck@skkku.edu](mailto:hangsuck@skkku.edu)

의 예측은 현 시점에서 선행되어야 할 중요한 과제라고 할 수 있다. 이에 따라 최근 사망률 예측 모형에 대한 연구가 지속적으로 이루어지고 있으며, 이들 모형을 이용한 연금공급자의 리스크 관리에 관한 연구도 활발히 이루어지고 있다.

본 연구는 미래 사망률을 예측하는 대표적인 확률적 사망률 모형인 Lee-Carter 모형(이하 LC 모형)을 사용하여 과거 생명표로 산출한 왜도로 미래 왜도 추세를 예측하여 미래의 사망률 개선 정도를 나타내는 사망률 지수인 모수를 간접적으로 예측하는 방식을 제시한다. 기존의 LC 모형을 이용한 사망률 예측 방식은 사망률 지수  $\kappa_t$ 를 추정하고 미래  $\kappa_t$ 의 수치를 직접 예측함으로써 미래 사망률이 지나치게 개선되는 현상을 보이며, 이를 바탕으로 산출된 연금액과 지급기간 추정 등 연금 공급자의 리스크 관리에 반영되는 장래사망률추계를 과소평가하는 요인으로 작용한다. 본 연구는 기존 예측 방식의 사망률 예측 결과와 제시한 왜도 예측 방식의 사망률 예측 결과를 비교함으로써 기존 사망률 예측 방식의 문제점을 지적하고 개선방안을 논의한다. 미래 사망률 예측 방식에 있어 왜도 예측을 통한 LC 모형의 사망률 예측은 기존 방식보다 사망률 개선효과를 더 적게 반영하며 장수리스크를 덜 왜곡시키는 방안을 제기하는데 의의가 있다고 할 수 있다.

사망률 모형인 LC 모형을 사용하여 미래 사망률을 예측하는 기존의 많은 연구들은 모형 자체의 한계점 수정에 대해 초점을 맞추고 있는 것으로 파악된다. 예를 들어 Kim (2012), Kim (2011), Li와 Lee (2005) 등이 있다. 선행연구들은 미래 사망률의 예측방식에 대해 논의하기보다는 예측에 앞서 선행되어야 할 모수들의 추정방식과 자료의 특성과 관련한 LC 모형의 문제점들을 다루고 있다.

본 연구는 통계청의 1970년부터 2013년까지 총 44개의 남녀 생명표 자료를 사용하여 LC 모형의 추정 방법과 미래 사망률 예측 방법을 구체적으로 살펴보고, 사망률의 추정 및 예측 결과를 바탕으로 추정된 남녀 30세 기준의 년도별  $T_{30}$  왜도와 모수  $\kappa_t$ 의 관계식을 도출하여 사망률 지수  $\kappa_t$ 의 미래 예측수치를 보정하는 방법을 제기함으로써 LC 모형의 사망률 예측의 한계점에 대한 개선방안을 논의한다. 더 자세한 내용은 다음 장에서 설명하도록 하겠다.

## 2. LC 모형과 사망률 예측

### 2.1. 연령별 사망확률의 추정

본 연구의 사망확률 추정을 위한 첫 번째 단계는 사망확률 추정을 위한 기초 데이터의 가공이다. 통계청에서는 1970년부터 1997년까지는 대한민국 국민의 성별 및 5세군단별 간이생명표만을 제공하고 1997년부터 2013년까지는 성별 및 연령별 완전생명표도 간이생명표와 같이 제공하고 있다. 이러한 과거 자료의 부족은 연령별 사망확률 예측을 매우 어렵게 하는 요인이다. 뒤에서 보다 자세히 설명하겠지만 기초 데이터를 가공하지 않고 그대로 사용하여 사망확률 예측 모형의 모수를 추정할 경우 의미 있는 결과를 기대하기 힘들다. 따라서 본 연구에서는 연령별 사망자 수를 나타내는 기초 데이터를 얻기 위해 5세군단별로 정리되어 있는 1970년부터 2013년까지의 간이생명표를 가장 보편적으로 이용하는 3차 Spline 보간법으로 연령별 사망자수를 제공하는 완전생명표로 확장한다 (Joo와 Lee, 2014). 다음은 최종사망연령을 동일하게 고정시킨다. 통계청에서 제공하는 생명표들의 경우 작성시점별로 각기 다른 최종사망연령을 제공한다. 1970년부터 1992년까지는 최종연령이 80세이며, 1993부터 1998년까지는 85세, 1999년과 2000년의 경우는 95세로 이후 2001년부터 현재까지는 100세로 최근에 가까워질수록 최종연령 역시 점차 늘어나고 있기 때문에 보다 편리한 자료의 활용을 위해 각기 다른 최종연령을 동일하게 고정할 필요가 있다. 본 연구에서는 최종연령 수준을 최근 13년간의 최종사망연령과 동일한 100세로 가정한다.

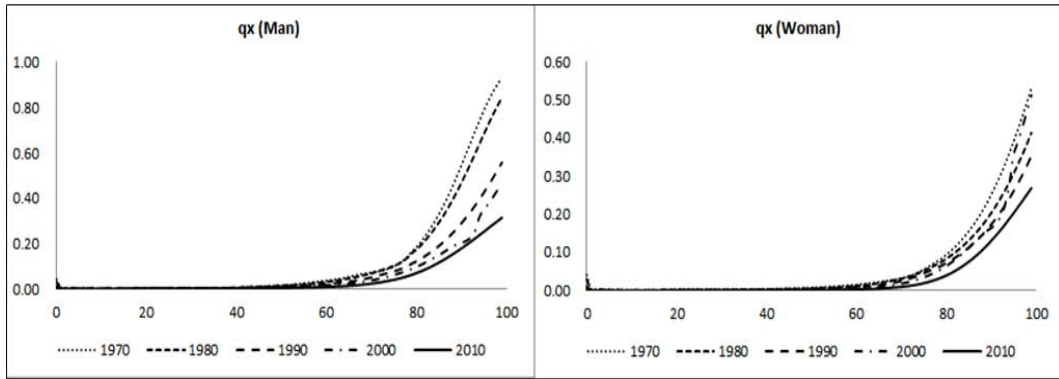


Figure 2.1. Estimated  $q_x$  for given age using Gompertz extrapolation.

## 2.2. Gompertz 보외법을 사용한 고연령 사망확률 추정

3차 Spline 보간법을 이용하여 연령별 사망확률을 추정할 경우 각 생명표 작성 시점에서의 최종연령 내의 연령별 사망확률을 추정할 수 있다. 그러므로 각 생명표의 최종연령 이상 연령의 사망확률은 보외법을 이용하여 추정해야 한다. 본 연구에서는 Gompertz 사망법칙을 가정함으로써 고연령의 사망확률을 추정하였다.

$$m_x \approx \mu_x = BC^x \quad \text{단, } B > 0, C > 1, x = 0, 1, 2, \dots \quad (2.1)$$

위 식의 모수  $B$ 와  $C$ 는 3차 Spline 보간법을 이용해 추정한 연령별 사망확률을 이용, 최소제곱법에 따라 추정할 수 있다.  $m_x$ 는  $x$ 세의 사력  $\mu_x$ 의 추정값으로, 사력  $\mu_x$ 는  $x$ 세의 순간적인 사망확률을 나타낸다. 소수연령 구간에서 사력  $\mu_x$ 가 상수임을 가정하면 사망확률  $q_x$ 는 다음의 식으로 나타낸다.

$$q_x = 1 - \exp^{-BC^x}, \quad x = 0, 1, 2, \dots \quad (2.2)$$

또한, 사망확률  $q_x$ 와 생존확률  $p_x$  간의 관계를 이용하면 다음의 관계식이 유도된다.

$$\ln(-\ln(p_x)) = \ln B + x(\ln C). \quad (2.3)$$

식 (2.3)의 좌변을 치환하게 되면 단순회귀분석을 통해 모수  $B$ 와  $C$ 를 용이하게 추정할 수 있다. 이 단계가 끝나면 다시 식 (2.2)를 이용하여 1970년부터 1992년까지는 80세 이상의 사망확률을, 1993년부터 1998년까지는 85세 이상의 사망확률, 1999년과 2000년은 95세 이상의 고연령에 대한 사망확률을 각 생명표의 75세 이상 사망확률 자료만을 이용하여 각각 계산하게 된다. Figure 2.1은 그 결과에 기반하여 추정한 연령별 사망확률을 나타낸 것이다.

## 2.3. LC 모형을 이용한 사망률 추정

미래 사망률을 예측하는 대표적인 모형으로는 LC 모형이 있다 (Lee와 Miller, 2001). LC 모형은 1992년 인구통계학자인 Lee와 Carter에 의해 개발된 확률적 사망률 모형이다. LC 모형은 로그변환 사망률을 연령효과(age effect)와 기간효과(period effect)의 선형관계로 설명한다. LC 모형은 모형의 단순성과 추정의 용이함으로 인해 사망률 연구에 가장 기본적인 모형으로 사용되고 있다. LC 사망률 예측을 위한 모형은 다음과 같다 (Lee와 Carter, 1992).

$$\ln \hat{m}_{x,t} = \alpha_x + \beta_x \kappa_t + \epsilon_{x,t} \quad (2.4)$$

여기서  $\hat{m}_{x,t}$ 는  $t$ 해  $x$ 세의 사망률의 관측치이고,  $\epsilon_{x,t}$ 는 평균이 0이고 분산이  $\sigma_\epsilon^2$ 인 정규분포를 따르는 오차항을 나타낸다.  $\alpha_x$ 는 연령별 모수로  $\ln \hat{\mu}_{x,t}$ 의  $t$ 에 대한 평균을 나타낸다.  $\beta_x$ 는 또 다른 연령별 모수로 사망률 지수  $\kappa_t$ 의 변화에 대한 연령별 민감도를 나타내는 지표로 볼 수 있다. 사망률 지수  $\kappa_t$ 는 시간의 흐름에 따른 사망률의 변동 추세를 나타내는 것으로 사망률이 개선되는 것은 사망률 지수  $\kappa_x$ 가 시간의 경과에 따라 감소하는 것으로 확인할 수 있다. LC 모형에서  $\kappa_t$ 역시 모수이기 때문에 일반적인 최소자승법(Ordinary Least Squares; OLS)을 이용하여 추정할 수 없으며, 특이값 분해(Singular Value Decomposition; SVD) 방법으로 추정한다. 또한,  $\beta_x$ 와  $\kappa_t$ 의 유일해를 찾기 위한 조건으로 다음의 제약들을 부과한다.

$$\sum_t \kappa_t = 0, \quad \sum_x \beta_x = 1. \quad (2.5)$$

특이값 분해는 다차원의 변수를 낮은 차원으로 표현하는 주성분 분석(Principal Component Analysis; PCA)의 주성분을 구하는 방법이다. 최소자승법처럼 다음의 식을 최소화하는 모수를 추정하며, 연령별 모수인  $\hat{\alpha}_x$ 는  $\ln \hat{m}_{x,t}$ 의  $t$ 에 대한 평균이다.

$$\sum_{x,t} (\ln \hat{m}_{x,t} - \alpha_x - \beta_x \kappa_t)^2, \quad (2.6)$$

여기서  $\ln(m_{x,t}) - \alpha_x$  행렬에 특이값 분해를 적용하게 되면

$$\text{SVD} [\ln \hat{m}_{x,t} - \hat{\alpha}_x] = \sum_{i=1}^r p_i U_{x,t} V_{t,i}, \quad p_1 \geq p_2 \geq \dots \geq p_r \quad (2.7)$$

로 표현되고, 이 때의  $r$ 은  $\text{rank}[\ln \hat{m}_{x,t} - \hat{\alpha}_x]$ 이고  $p_i$ 는 행렬의 특이값이고  $U_{x,t}$ 와  $V_{x,r}$ 는 각각 좌·우 특이벡터에 해당한다. 즉,  $\ln(m_{x,t}) - \alpha_x$  행렬은 특이값을 기준으로  $r$ 개 행렬의 합으로 분해되기 때문에  $r$ 개 행렬 중 가장 큰 특이값으로 구성되는 첫째 행렬값으로 모수를 추정하게 된다. 다시 말해  $\hat{\beta}_x$ 는 특이값 분해 결과로 나타나는 좌특이벡터의 첫째 열이 되고,  $\hat{\kappa}_t$ 는 가장 큰 특이값  $p_1$ 에 우특이벡터의 첫째 열을 곱한 값이 되는 것이다. 여기에 앞서 제한한 제약식들을 부과하고, 위의 추정결과로 얻은 사망률은 로그사망률이기 때문에 실제 사망 관측치와 기대 사망자수가 일치하도록  $\hat{\kappa}_t$ 를 다음과 같이 보정하여 새로운  $\hat{\kappa}_t$ 를 산출해야 한다.

$$\sum_x D_{x,t} = \sum_x E_{x,t} \exp(\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t). \quad (2.8)$$

위의 수식의 왼쪽 항인  $\sum_x D_{x,t}$ 은  $t$ 년도의 실제 총 사망자 수이며,  $E_{x,t}$ 는  $t$ 년도 나이  $x$ 인 모집단으로 식 (2.3)의 오른쪽 항은 LC 모형에 의해 추정된 총 사망자 수다. 그러므로 두 값들이 서로 일치하도록 하는 것은 전체 기간 동안 실제 사망자 관측치와 추정된 총 사망자 수를 일치하게끔 하고 이를 만족하는  $\hat{\kappa}_t$ 를 구하게 된다.  $\hat{\kappa}_t$ 는 최초 추정된  $\hat{\kappa}_t$ 를 초기값으로 설정하고 위 식이 만족할 때까지 반복적으로 조정하여 산출된다. 관찰시점 이후는 ARIMA 모형을 이용해 추정하여 구하는 방식이 일반적이다.

전통적인 LC 모형에서 특이값 분해를 통한 OLS 추정의 가장 큰 문제점 중 하나는 오차항의 등분산성 가정이다 (Cairns 등, 2008; Park 등, 2013). 그러나 이는 매우 비현실적인 가정으로 실제 관측된 사망률의 로그변환값, 즉  $\ln \hat{m}_{x,t}$ 는 고연령에서 큰 변동을 나타낸다. 이같은 변동은 고연령에서의 실제 사망자 수가 너무 작기 때문에 발생하는 것으로 Alho (2000)는 사망자 수를 포아송 분포로 모델링함으로써 이 문제를 개선하고자 하였다. 즉, 사망자 수는 사건의 수를 나타내는 확률변수(Counting Random Variable)이므로 포아송분포로 적합하는 것이 더 적절할 뿐만 아니라 이를 통해 최대우도추정

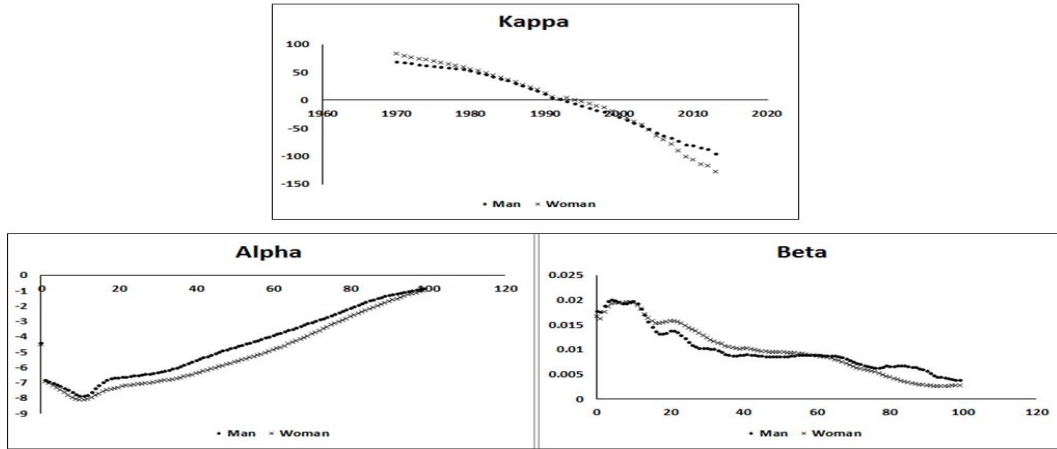


Figure 2.2. Parameter estimates of LC models calculated from LEM (1970-2013).

법(Maximum Likelihood Estimation; MLE)으로 모수를 추정할 수 있다는 것이다. 이에 따르면 사망자 수는

$$D_{x,t} \sim \text{Poisson}(E_{x,t} m_{x,t}) \quad \text{단, } m_{x,t} = \exp(\alpha_x + \beta_x \kappa_t) \quad (2.9)$$

로 표현되고 모수는 로그우도함수인 다음의 식을 최대화하도록 추정된다 (Brouhns 등, 2002).

$$l(\alpha, \beta, \kappa) = \sum_{x,t} D_{x,t}(\alpha_x + \beta_x \kappa_t) - E_{x,t} \exp(\alpha_x + \beta_x \kappa_t) + \text{constant}. \quad (2.10)$$

전통적 LC 모형에서와 마찬가지로 유일해를 찾기 위해 앞서 제약으로 제시한 식 (2.5)가 적용된다. 상기의 방법은 전통적인 LC 모형의 모수를 최대우도법으로 추정하는 것으로 포아송 로그-이중선형 모형(Poisson log-bilinear model)로 불린다. 여기서 모수  $\alpha_x$ ,  $\beta_x$  그리고  $\kappa_t$ 의 해석은 전통적인 LC 모형과 동일하다. 또한 실제 사망자 수가 위의 식에 직접 적용되었기에 전통적인 LC 모형과는 달리 사망률 지수  $\kappa_t$ 의 보정 작업을 필요로 하지 않는다. 뿐만 아니라 ARIMA 모형에 기반한 사망률 지수의 예측을 통해 미래 사망률을 예측하는 절차 또한 동일하다. 따라서 포아송 로그-이중선형 모형은 전통적인 LC 모형의 장점은 그대로 계승하며 LC 모형의 가장 큰 문제점 중의 하나인 등분산성 가정 (Park 등, 2013)의 문제점을 보완한 모델이라 할 수 있겠다. 또한 최대우도추정법은 LC 모형에 코호트 효과를 추가한 확장 모형에서도 모수 추정을 위해 사용되는데 전통적인 LC 모형과 달리 최대로그우도값을 제시함으로써 베이지 정도 기준(Bayesian Information Criterion; BIC)을 이용하여 모형간 적합도를 비교해 볼 수 있다는 장점이 있다.

다만 최대우도법에 따른 모수 추정법의 경우 공변수(covariate)없이 이중선형(bilinear)항  $\beta_x \kappa_t$ 를 추정해야 하기 때문에 GLIM 혹은 LEM과 같은 프로그램(soft-ware)을 사용해야 한다. 본 연구에서는 범주형 데이터 분석에 특화된 LEM을 사용하였다 (Vermunt, 1997). Figure 2.2는 남녀별 각 모수들을 위의 방식으로 추정한 결과이다.

LC 모형을 이용한 미래 사망률의 예측은 사망률 추세를 나타내는 사망률 지수  $\kappa_t$ 를 ARIMA 모형으로 식별한 후  $\kappa_t$ 의 예측치를 구함으로써 수행된다. 본 장에서는 우리나라 사망률 자료를 통해 추정된 사망률 지수  $\kappa_t$ 의 ARIMA( $p, d, q$ ) 모형 차수를 식별해보도록 한다. ARIMA 모형의 차수  $p, d, q$ 를 결정하

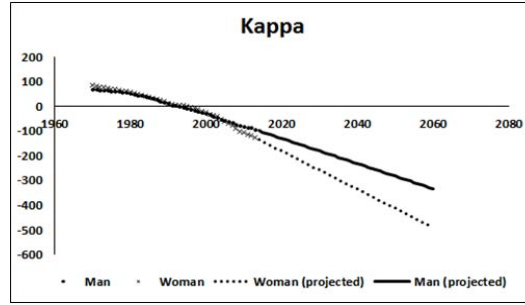


Figure 2.3. Forecasts of  $\kappa_t$  from ARIMA(0, 2, 1) model (1970–2060).

Table 2.1. ARIMA(0, 1, 0) model estimates for  $\kappa_t$

	$\theta$ 추정치	표준오차	통계량	$p$ -value
남성	-5.0301	0.4904	-10.2570	0.0000
여성	-7.9866	0.8016	-9.9636	0.0000

기 위해서는 서로 다른 ARIMA 모형 간 비교 기준이 필요하며 대표적인 비교 기준으로 Akaike 정보기준(Akaike Information Criterion; AIC)과 베이지 정보기준(Bayesian Information Criterion; BIC)이 있다. AIC와 BIC는 아래와 같이 계산된다.

$$\begin{aligned} \text{AIC} &= -2 \ln(l) + 2k, \\ \text{BIC} &= -2 \ln(l) + k(\ln(n)). \end{aligned} \quad (2.11)$$

이 때  $l$ 은 최대로그우도 함수값,  $k$ 는 추정계수의 수,  $n$ 은 추정에 사용된 데이터의 수를 의미하며, 두 정보기준은 모두 값이 작을수록 추정모형의 적합도가 높음을 의미한다. 남성의 경우에는 ARIMA(0, 2, 1) 모형의 AIC 및 BIC 값이 가장 낮게 나타났으며, 여성의 경우 역시 ARIMA(0, 2, 1) 모형의 AIC 및 BIC 값이 가장 낮게 나타남을 알 수 있다. 본 연구에서는 남녀 모두 통계수치상 가장 알맞은 ARIMA(0, 2, 1) 모형을 선택하였다. ARIMA(0, 2, 1) 모형은 다음의 수식으로 표현가능하다.

$$\kappa_t = 2\kappa_{t-1} - \kappa_{t-2} + \theta\epsilon_t, \quad \epsilon_t \sim N(0, \sigma^2). \quad (2.12)$$

남성과 여성의 사망률 지수  $\kappa_t$ 의 ARIMA(0, 2, 1) 모형에서 계수  $\theta$  추정에 대한 검정결과 상  $p$ -value가 유의하지 않아 1970년부터 2013년까지의 남녀 생명표 자료를 ARIMA(0, 2, 1) 모형으로 적합하여 미래 사망률을 예측하는 것은 올바르지 않다고 판단된다.

그러므로 본 연구에서는 2001–2013년 13년간의 최근 사망률 자료만을 선별하여 사망률 지수  $\kappa_t$ 를 ARIMA 모형으로 추정 및 예측을 실행한다. ARIMA 모형 식별은 이용 데이터와 고려하는 연령의 범위에 따라 달라질 수 있는 것으로 위의 방식과 동일하게 최근 데이터에 적용하면 남성과 여성 모두 ARIMA(0, 1, 0) 모형의 AIC 및 BIC 값이 가장 낮게 나타남을 알 수 있다. 본 연구에서는 남녀 모두 통계수치상 가장 알맞은 ARIMA(0, 1, 0) 모형을 선택하였다.

$$\kappa_t = \kappa_{t-1} + \theta + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma^2). \quad (2.13)$$

Table 2.1은 앞서 설명한 ARIMA(0, 1, 0) 모형에서 상수  $\theta$ 의 추정에 대한 검정결과( $z$ -test)를 성별로 나타낸 것이다. 특히  $p$ -value값은 각각 회귀계수가 0인지 아닌지를 검정하는 경우에 계산된다.  $p$ -value값

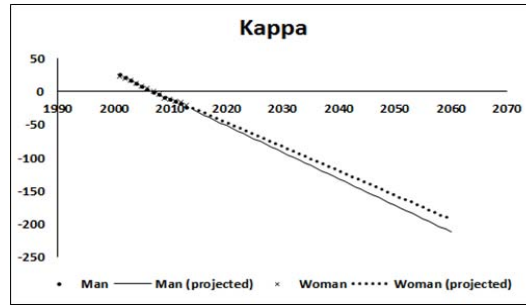


Figure 2.4. Forecasts of  $\kappa_t$  based on ARIMA(0, 1, 0) model in Table 2.1 from 2001 to 2060.

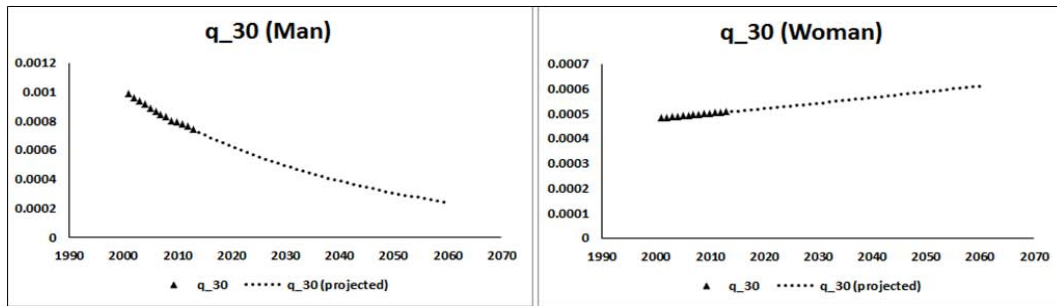


Figure 2.5. Prediction of  $q_{30}$  by gender from 2014 to 2060.

이 아주 작은 경우(일반적유의수준 0.05)에 우리는 그 회귀계수가 유의하다고 말하고 실제로 의미 있는 추정결과라고 간주할 수 있다. 위의 결과에서는 남성과 여성 모두  $p$ -value값이 아주 작으므로 두 경우 모두 사망률 지수  $\kappa_t$ 를 ARIMA(0, 1, 0) 모형으로 추정된 값들이 유의하며 이를 이용하여 장래사망률지수를 추계할 수 있다고 판단된다. Table 2.1의 결과를 이용하여 2014년부터 2060년까지 사망률 지수  $\kappa_t$ 의 미래 예측 수치를 성별로 산출하였으며 Figure 2.4의 그래프는 예측한 값들을 그래프로 나타낸 것이다.

Figure 2.3과 Figure 2.4에서 알 수 있듯이 사망률 지수  $\kappa_t$ 의 예측치는 시간이 지날수록 점차 감소하는 것을 알 수 있으며, 성별과 무관하게 지속적인 선형적 감소가 관찰된다. 최근 13년간의 사망률로 사망률 지수  $\kappa_t$ 를 예측한 Figure 2.4와 앞서 전체자료의 사망률로 예측한 Figure 2.3을 비교해 보면 Figure 2.4의 경우 2060년에 남성의 경우에는 사망률 지수  $\kappa_t$ 가  $-211.94$ 이며 여성의 경우에는  $-193.01$ 이지만, Figure 2.3에서는 2060년 남성과 여성의 사망률 지수  $\kappa_t$ 의 예측치는  $-490.92$ 와  $-334.54$ 로 큰 차이가 발생하는 것을 알 수 있다. 앞서 설명했듯이 이러한 차이는 사망률 지수  $\kappa_t$ 의 정의인 시간의 흐름에 따른 사망률의 개선사항으로 인한 것이다. 사용되는 사망률 데이터의 관측시작점이 오래되었을수록 예측된 장래사망률의 개선속도 역시 더 급속도로 이루어지는 것을 볼 수 있다. 이는 과거 사망률의 개선 속도가 장래사망률 추계에도 반영되기에 나타나는 현상이며 오래된 과거 추세와 최근의 개선 추세와는 상당한 차이가 있으므로 미래 사망률을 예측할 때에는 자료선택에 있어서 좀 더 신중함을 보여야 한다. 본 연구에서는 Figure 2.4처럼 최근 13년간의 사망률 자료를 이용하여 장래사망률추계를 수행하도록 하겠다.

앞서 산출한 Figure 2.2의 모수  $\alpha_x, \beta_x$ 와 Figure 2.4의  $\kappa_t$  예측값을 식 (2.6)에 대입하면 사망률  $m_x$ 의 미래 예측치를 구할 수 있다. 하지만, 생명표의 작성에는 연령별 사망확률  $q_x$ 가 요구되므로 사망률  $m_x$ 로

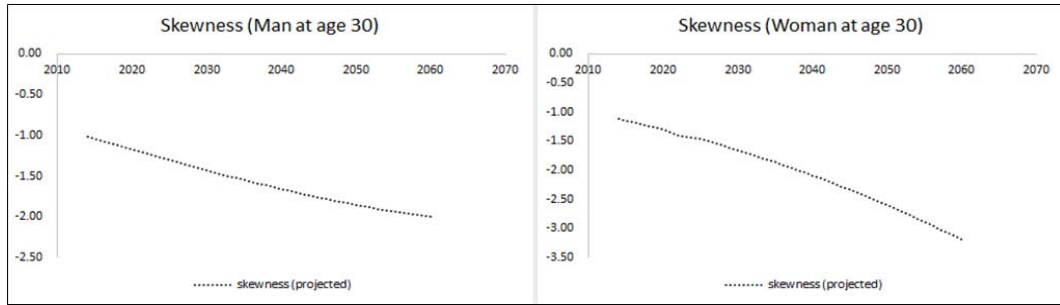


Figure 3.1. Predicted skewness of  $T_{30}$  from 2014 to 2060.

부터 사망확률  $q_x$ 를 계산해야 한다. 본 연구에서는 소수연령 구간에서의 사망자 수는 균등하게 분포한다는 가정(Uniform Distribution of Deaths; UDD) 하에 다음의 식으로 2060년까지의 장래사망률을 추계한다.

$$q_x = \frac{m_x}{1 + 0.5m_x}. \quad (2.14)$$

다음 Figure 2.5는 남녀별 30세의 사망확률을 2014년부터 2060년까지 예측해 본 결과이다.

### 3. 사망률 지수 $\kappa_t$ 의 미래 예측값 산출방식 고안

#### 3.1. 왜도예측방식의 소개와 기존예측방식과의 비교

앞서 예측한 사망률 지수  $\kappa_t$ 를 사용하여  $q_{30}$  외의  $T_{30}$  기대여명, 표준편차, 왜도, 첨도 등의 다양한 통계량들을 예측할 수 있다.  $T_{30}$ 은 30세가 되는 시점부터 사망시점에 이르기까지의 기간이다. 따라서 30세의 사망확률  $q_{30}$ 은 다음과 같이 정의할 수 있다.

$${}_tq_{30} = \Pr(T_{30} \leq t).$$

Figure 3.1은  $T_{30}$ 의 왜도를 나타낸 것이다.  $T_{30}$  왜도는  $q_{30}$ 의 확률분포의 대칭성을 측정한다. 왜도가 0에 가까울수록  $q_{30}$ 의 확률분포는 정규분포와 비슷하며, 위의 결과처럼 음수로 나타난다면  $q_{30}$ 의 확률분포는 오른쪽 분포에 비해 왼쪽 분포가 더 길게 뻗어있는 형태로 중앙값을 포함한 값들이 오른쪽으로 편중되어 분포해 있는 것을 알 수 있다. 또한,  $T_{30}$  왜도의 절대값이 점차 증가하는 추세이므로 이러한  $q_{30}$ 의 편중이 시간이 지날수록 점점 심해지는 것을 알 수 있다. 따라서 미래에는 고연령에서 사망이 더 집중적으로 발생할 것으로 판단된다.

이러한 문제는  $\kappa_t$  장래추계의 계속된 선형적 감소 예측에 따른 것으로 지나친 사망률 개선효과의 반영으로 인해 미래 사망확률을 과소평가하는 결과를 가져온다. 새로운 사망률 예측방식의 필요성은 기존의 연구 (Jeong과 Kim, 2011)에서도 언급되었으며 기대여명을 이용하여 사망률 예측 모형을 선택하는 방안에 대해 진행되었다. 그러나 본 연구에서는 사망률 예측 모형의 선택보다는  $\kappa_t$ 의 예측방식의 수정에 대한 관점에서 이 문제를 살펴보고자 한다. 자세한 과정은 Figure 3.2를 살펴보기 바란다.

본 연구는 왜도와 사망률 지수  $\kappa_t$ 의 관계에서 새로운  $\kappa_t$ 를 산출하는 방안을 제안한다. 앞선 과정에서  $T_{30}$  왜도와  $\kappa_t$  모두 전체 예측기간 동안 지속적으로 감소하는 것을 살펴보았다. 또한, 두 값 모두 사망확률의 분포와의 강한 연관성이 있는 점은 이 두 값 사이의 관계로 새로운  $\kappa_t$ 를 산출하는 방안의 제안에



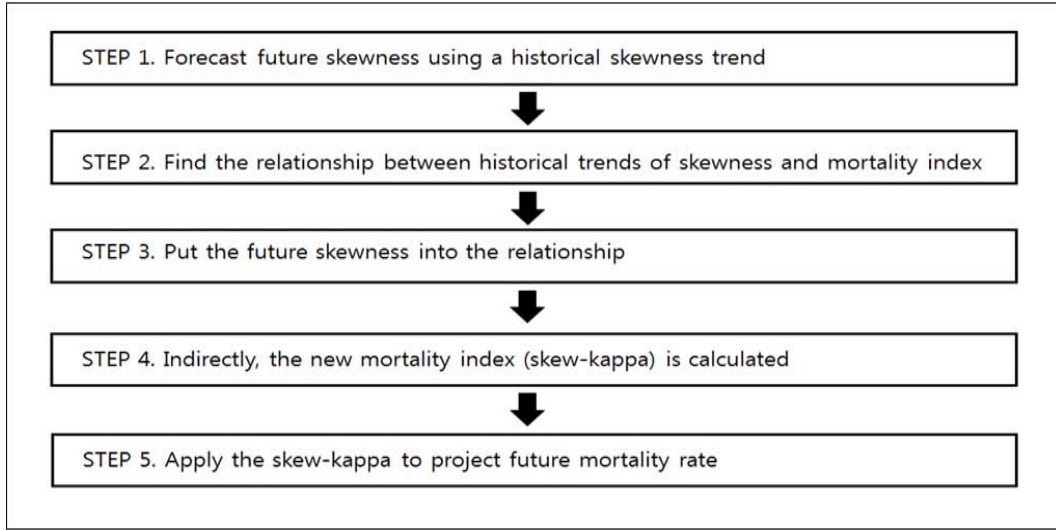


Figure 3.2. Diagram of proposed prediction method based on skew- $\kappa_t$ .

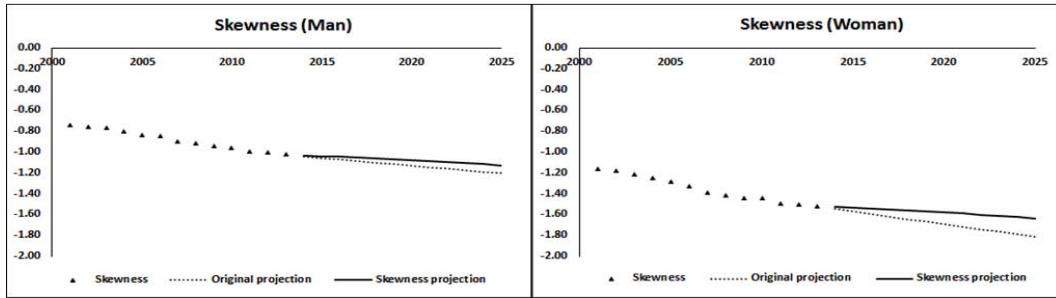


Figure 3.3. The comparison of predicted skewness of  $T_{30}$  (2014–2024).

강한 근거로작용한다. 새로운  $\kappa_t$ 는 과거 자료의 왜도 추세로 예측한 미래 왜도(이하 왜도예측)와 과거  $\kappa_t$ 의 추정치를 이용하여 예측한 미래  $\kappa_t$ 의 관계식에서 간접적으로 산출한다. 본 연구에서는 이 제안한 간접적인 방식으로 예측한  $\kappa_t$ 를 Skew- $\kappa_t$ 라 하겠다. Figure 3.2는 이러한 산출 과정을 단계별로 나타낸 것이다.

또한, Figure 3.3은 왜도예측방식과 기존예측방식을 이용하여 계산한 2014–2024년 사이의 미래  $T_{30}$  왜도예측을 비교한다. Figure 3.3의 왜도 예측을 이용한 미래  $T_{30}$  왜도의 경우 남녀  $T_{30}$  왜도는 각각 ARIMA(0, 1, 1)과 (0, 2, 1)에 적합하여 예측했다. Figure 3.3을 살펴보면, 남녀  $T_{30}$  왜도의 예측 결과가 왜도 예측방식이 기존 예측방식보다 적게 감소하는 것을 알 수 있다. 그러므로 왜도예측방식으로 장래사망률을 추계하는 경우 기존예측방식보다 덜 편향된(less left-skewed)  $q_{30}$ 의 사망확률분포를 예측할 것으로 판단된다.

Figure 3.4의 그래프들은 기존예측방식에 따른  $T_{30}$  왜도수준과  $\kappa_t$  예측수준 간의 관계식을 남녀별로 나타낸 것이다. Figure 3.4에 따르면 왜도와  $\kappa_t$ 의 예측치 사이에는 정비례관계가 성립하는 것을 알 수 있다. 앞서 Figure 3.2에서 언급한 것처럼, 위의 관계식에 왜도예측방식으로 산출한  $T_{30}$  왜도예측수치를

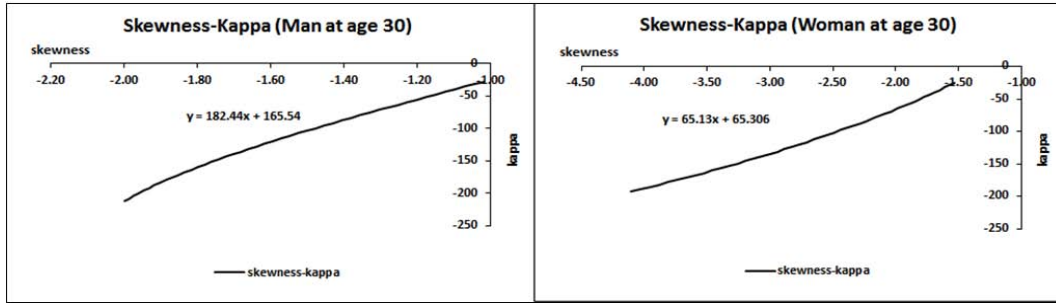


Figure 3.4. Relationship between  $T_{30}$  skewness and  $\kappa_t$  (2014–2060).

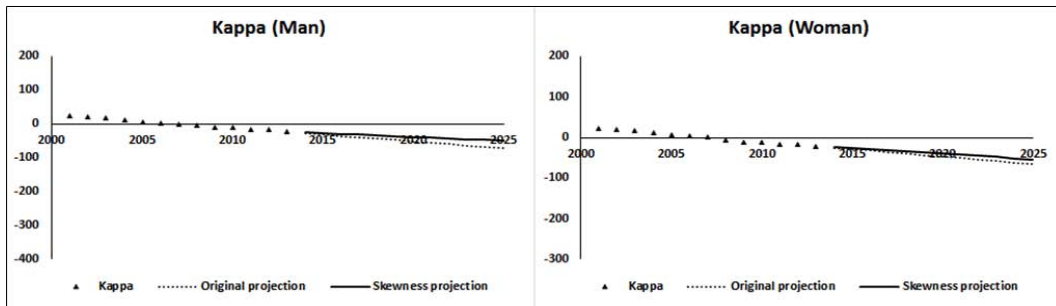


Figure 3.5. The comparison of predicted  $\kappa_t$  (2014–2024).

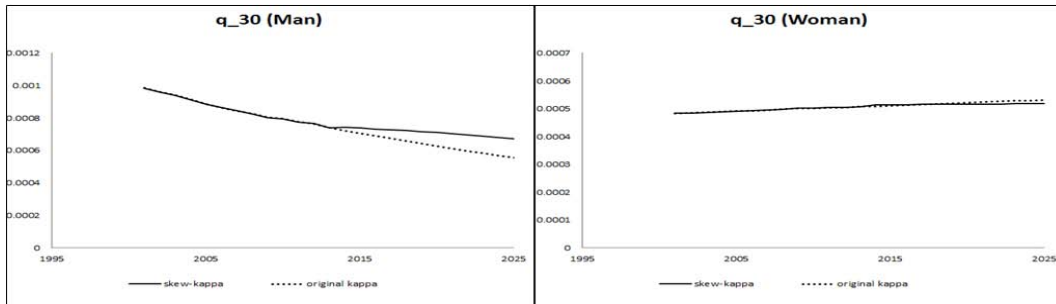


Figure 3.6. The comparison of predicted  $q_{30}$  (2014–2024).

대입하면 그에 대응하는  $Skew-\kappa_t$ 를 간접적으로 구할 수 있다. 다음의 그림은  $Skew-\kappa_t$ 와 기존- $\kappa_t$ 의 예측수치를 비교한 것이다.

Figure 3.5에서 역시 Figure 3.3에서 유추할 수 있었던 결과처럼  $Skew-\kappa_t$ 가 기존- $\kappa_t$ 에 비해 덜 왜곡되고 있는 것을 볼 수 있다.

Figure 3.7에서  $T_{30}$ 의 장래기대여명추계를 보면 성별에 무관하게 왜도예측방식이 기존예측방식에 비해 사망률 개선이 덜 반영된 것을 알 수 있다. 예를 들어 2024년에 왜도예측방식의 남녀 기대여명은 각각 50.88과 57.46을 기록하여 각각 53.46과 59.14로 나타난 기존예측방식에 비해 5%와 3%씩 감소한 수치를 나타낸다. 그러므로 본 연구에서 제안한 왜도예측방식이 기존방식에 비해 장래사망률을 덜 왜곡하는 것으로 판단된다.

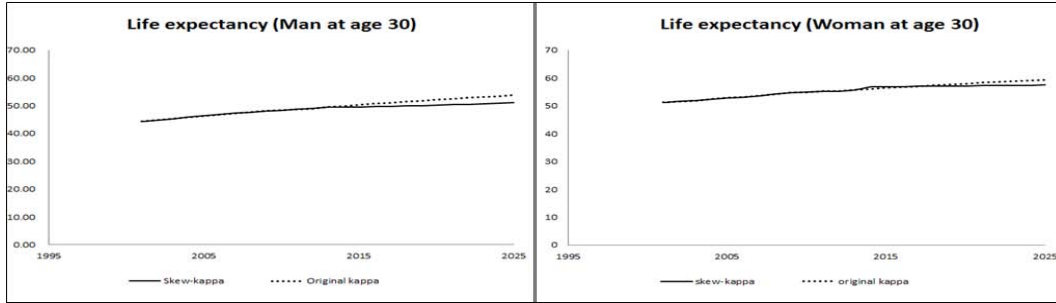


Figure 3.7. The comparison of the predicted life expectancy of  $T_{30}$  (2014–2024).

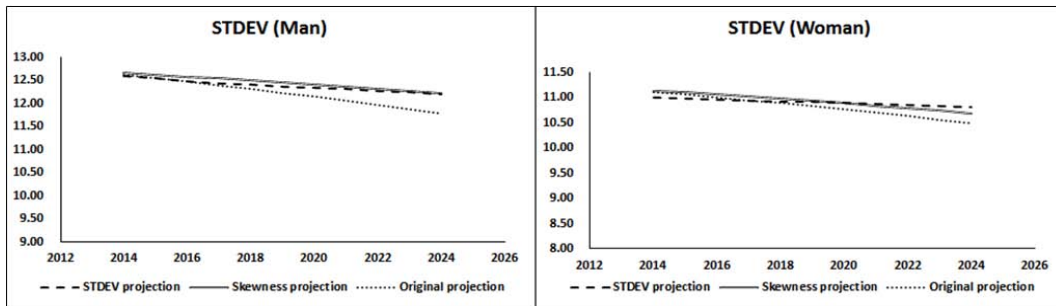


Figure 3.8. The comparison of predicted standard deviation of  $T_{30}$  by gender.

### 3.2. 왜도 예측과 표준편차 예측 및 첨도 예측의 비교

본 연구에서는 왜도가 보유한 사망확률분포와의 연관성과  $\kappa_t$ 와의 관계에서 출발한 왜도예측방식을 제안하였다. 그리고 제안한 방식으로 간접적으로 예측한  $\kappa_t$ 와 기존예측방식의  $\kappa_t$ 를 비교하였으며, 그 결과 의도했던 것처럼 덜 왜곡된 장래사망률추계를 얻을 수 있었다. 다음은 과거자료의 표준편차 추세를 이용하여 미래 표준편차를 예측하는 표준편차예측방식과 이와 비슷하게 미래 첨도를 첨도의 추세를 이용하여 예측하는 첨도예측방식을 살펴보겠다. 이 방식들로 예측한 표준편차와 첨도를 왜도예측방식과 기존예측방식으로 산출한 값과의 비교함으로써 왜도예측방식의 적합성에 대해 알아보도록 하겠다.

Figure 3.8은 표준편차예측, 왜도예측 및 기존예측방식을 이용해 예측한 미래  $T_{30}$  표준편차추계의 비교 결과를 나타낸 것이다. 그 결과 표준편차예측과 왜도예측을 이용한 표준편차의 장래추계는 2024년 남성의 경우 0.2% 이하의 근소한 차이를 보이지만 기존예측방식의 표준편차와 앞선 두 방식은 약 4%의 차이가 벌어지는 것을 알 수 있다. 여성의 경우도 이보다는 근소하지만 비슷하게 나타난다 (Figure 3.9).

또한, 이와 유사하게 왜도예측과 기존예측방식을 이용하여 산출한  $T_{30}$  첨도의 장래추계와 2001–2013년 과거  $T_{30}$  첨도 추세로 미래 첨도를 예측한 첨도예측 역시 앞선 표준편차 장래추계의 비교결과처럼 첨도 예측과 왜도예측을 이용한 장래첨도의 추계는 근사하였으나 기존 방식으로 예측한 첨도수치와는 큰 차이가 생기는 것으로 나타난다. 예를 들어 2024년의 남성은 왜도예측과 첨도예측으로 산출한 첨도 예측값에서 약 2.2%의 차이가 발생하였으나 기존예측과 위 두 예측값들과는 12%의 큰 차이가 나타났다. 여성의 경우 각각 3%와 9%의 차이를 보였다. 따라서 장래사망률추계에 대한 왜도예측방식의 사용은 높은 타당성을 갖는 것으로 볼 수 있다.

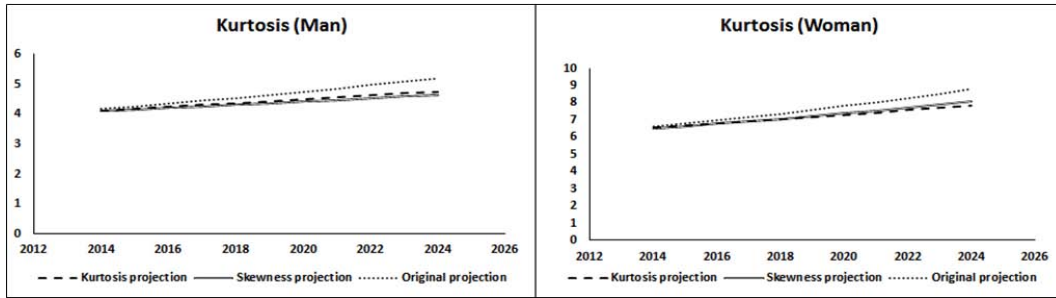


Figure 3.9. The comparison of predicted kurtosis of  $T_{30}$  by gender.

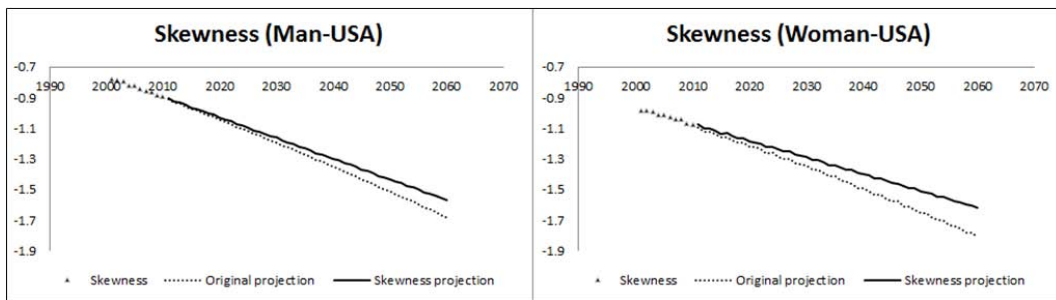


Figure 3.10. Projected skewness of  $T_{30}$  in US by gender (2011-2060).

### 3.3. 왜도예측방식을 적용한 미국-캐나다의 장래사망률추계

**3.3.1. 왜도 예측을 적용한 미국의 장래사망률추계** 미국의 1933-2010년의 78개의 남녀 생명표 자료를 이용하여 앞서 수행한 LC 모형의 모수추정 방식을 동일하게 적용하여 왜도를 예측하였다. 위의 그림은 본 연구에서 제안한 왜도예측방식과 기존예측방식을 사용하여 2011-2060년까지 장래 사망률 지수 추계를 각각 예측해 본 결과이다. 기존 예측방식과 왜도 예측방식은 AIC와 BIC가 모두 최소가 되는 ARIMA(2, 0, 0)으로 남녀 모두 적합하였다. 앞서 우리나라 자료를 이용한 결과와 유사하게 미국 자료에서 역시 기존 예측의 왜도가 왜도 예측의 왜도에 비해 과소평가되는 것을 볼 수 있었으며 이는  $T_{30}$ 의 사망확률분포가 시간이 지남에 따라 고연령에 더 많이 편중될 것을 의미하는 것으로 판단된다. 미국의 사망률 지수  $\kappa_t$ 를 왜도예측을 이용하여 간접적으로 추정한  $Skew-\kappa_t$ 와 기존예측방식으로 예측한 기존- $\kappa_t$ 를 비교해 본다면 그 결과는 Figure 3.11 같다.

Figure 3.11의 그래프들은 기존예측방식에 따른 왜도 예측치와  $\kappa_t$  예측치의 관계식을 남녀별로 나타낸 것이다. Figure 3.11에 따르면 왜도와  $\kappa_t$ 의 예측치 사이에는 정비례관계가 성립하는 것을 알 수 있다. 앞서 Figure 3.2에서 언급한 것처럼, 위의 관계식에 왜도예측으로 산출한 미래 왜도값을 대입하면 그에 대응하는  $Skew-\kappa_t$ 를 간접적으로 구할 수 있다. 그 결과 각각의 예측방식을 통해 왜도를 예측해 본 결과에서 알 수 있었던  $Skew-\kappa_t$ 가 기존- $\kappa_t$ 보다 남녀 장래사망률의 개선을 좀 더 보수적으로 반영하는 경향을 보이는 것을 알 수 있다 (Figure 3.12).

좀 더 자세히 살펴보기 위해 위의  $\kappa_t$  예측을 이용하여  $T_{30}$ 의 사망확률 및 기대여명을 산출해 보았다. 앞선 사망률 지수  $\kappa_t$ 의 예측 그래프로부터 알 수 있었던 미국 남녀의  $q_{30}$ 은 왜도 예측이 기존 예측에 비해 더 높게 나타나는 것을 알 수 있었다. 2024년의 남성  $q_{30}$ 은 왜도예측방식이 기존방식에 비해 4% 높게 예측되었으며 여성은 12%나 높게 예측되었다 (Figure 3.13). 이에 따른 결과로 남녀  $T_{30}$ 의 기대여

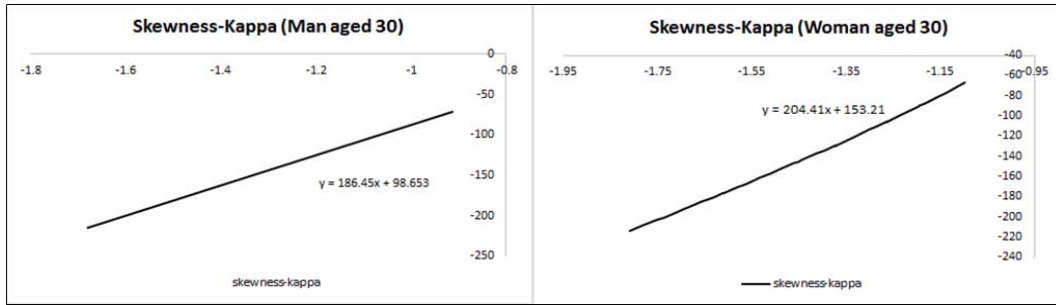


Figure 3.11. Relationship between  $\kappa_t$  and skewness of  $T_{30}$  (2011–2060).

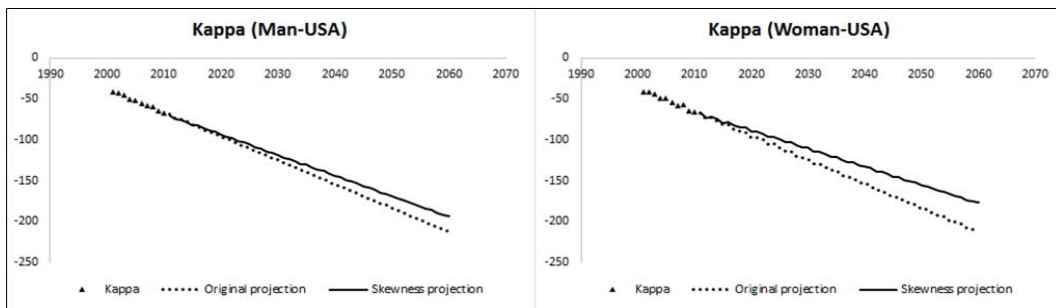


Figure 3.12. Projected mortality rate  $\kappa_t$  in US by gender (2011–2060).

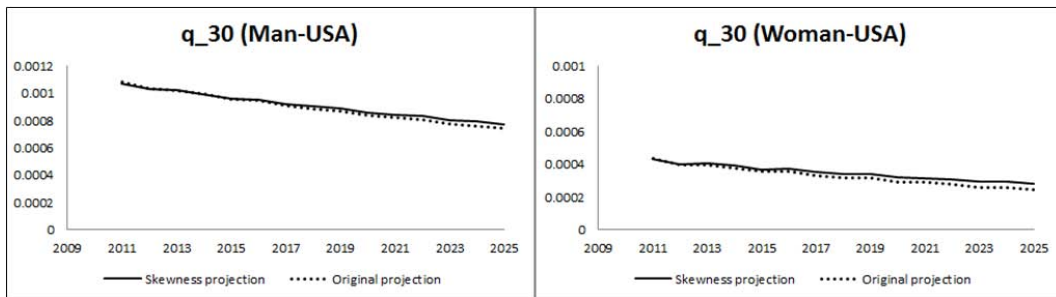


Figure 3.13. Projected  $q_{30}$  by gender in US by gender (2011–2025).

명이 더 낮게 예측되어 과도하게 측정된 사망률 개선 효과를 보정하는 효과를 보이는 것을 알 수 있다. 2024년에 예측된 남녀의 기대여명은 각각 약 0.61%와 1%씩 기존에 비해 감소하는 것을 볼 수 있었다 (Figure 3.14).

**3.3.2. 왜도 예측을 적용한 캐나다의 장래사망률추계** 캐나다의 1921–2011년의 91개의 남녀 생명표 자료를 이용하여 앞서 수행한 LC 모형의 모수추정 방식을 동일하게 적용하여 위의 Figure 3.15와 같이 왜도를 예측하였다. 위의 그림은 본 연구에서 제안한 왜도 예측방식과 기존 예측방식을 사용하여 2012–2060년까지 장래 사망률 지수 추계를 각각 예측해 본 결과이다. 기존 예측방식은 AIC가 최소화 되는 ARIMA(0, 1, 0)과 (0, 0, 1) 모형으로 남녀에 각각 적합하였고 왜도 예측방식은 ARIMA(0, 0, 2)와 (0, 0, 1)로 적합하였다. 앞서 우리나라 자료를 이용한 결과와 유사하게 캐나다 자료에서 역시 기존 예측

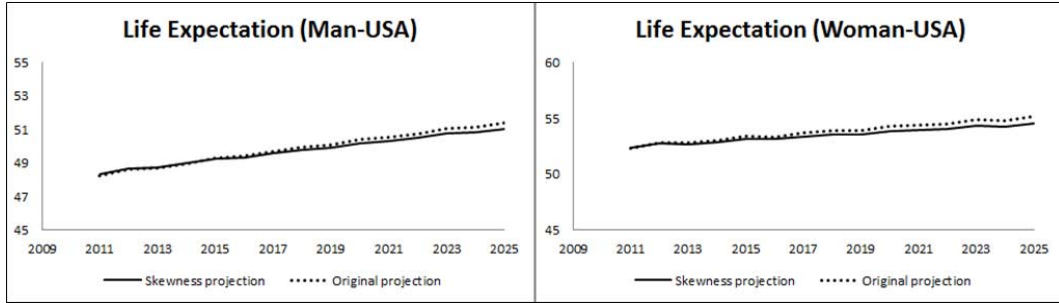


Figure 3.14. Projected life expectancy of  $T_{30}$  in US by gender (2011–2025).

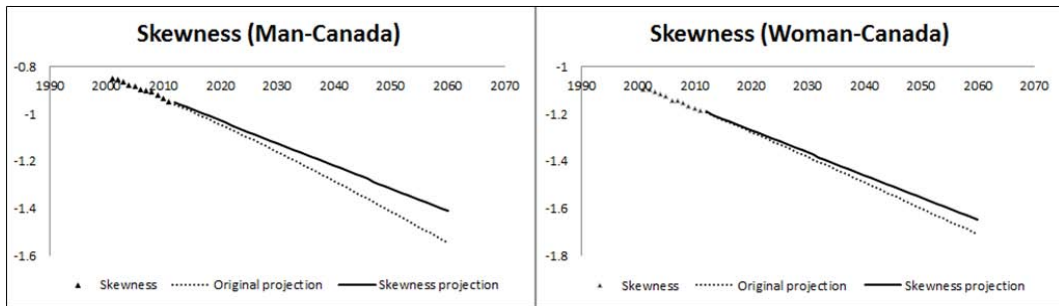


Figure 3.15. Projected skewness of  $T_{30}$  in Canada by gender (2001–2060).

의 왜도가 왜도 예측의 왜도에 비해 과소평가되는 것을 볼 수 있었으며 이는  $T_{30}$ 의 사망확률분포가 시간이 지남에 따라 고연령에 더 많이 편중될 것을 의미하는 것으로 판단된다. 캐나다의 사망률 지수  $\kappa_t$ 를 왜도예측을 이용하여 간접적으로 추정된  $Skew-\kappa_t$ 와 기존 예측방식을 이용한 기존- $\kappa_t$ 를 통해 비교해 본다면 그 결과는 다음의 그림과 같다.

Figure 3.16의 그래프들은 기존예측방식에 따른 왜도 예측치와  $\kappa_t$  예측치의 관계식을 남녀별로 나타낸 것이다. Figure 3.16에 따르면 왜도와  $\kappa_t$ 의 예측치 사이에는 정비례관계가 성립하는 것을 알 수 있다. 앞서 Figure 3.2에서 언급한 것처럼, 위의 관계식에 왜도예측으로 산출한 미래 왜도값을 대입하면 그에 대응하는  $Skew-\kappa_t$ 를 간접적으로 구할 수 있다. 그 결과 각각의 예측방식을 통해 왜도를 예측해 본 결과에서 알 수 있었듯  $Skew-\kappa_t$ 가 기존- $\kappa_t$ 보다 남녀 장래사망률의 개선을 좀 더 보수적으로 반영하는 경향을 보이는 것을 알 수 있다.

좀 더 자세히 살펴보기 위해 위의  $\kappa_t$  예측을 이용하여  $T_{30}$ 의 사망확률 및 기대여명을 산출해 보았다. 앞선 사망률 지수  $\kappa_t$ 의 예측 그래프로부터 알 수 있었듯 캐나다 남녀의  $q_{30}$ 은 왜도 예측이 기존 예측에 비해 더 높게 나타나는 것을 알 수 있었다. 2024년의 남성  $q_{30}$ 은 왜도예측방식이 기존방식에 비해 12% 높게 예측되었으며 여성은 6% 높게 예측되었다. 이에 따른 결과로 남녀  $T_{30}$ 의 기대여명이 더 낮게 예측되어 과도하게 측정된 사망률 개선 효과를 보정하는 효과를 보이는 것을 알 수 있다. 2024년에 예측된 남녀의 기대여명은 각각 약 0.82%와 0.3%씩 기존에 비해 감소하는 것을 볼 수 있었다.

#### 3.4. 미국 자료를 이용한 Back-testing 검증 결과

앞서 제안한 왜도예측방식과 기존예측방식으로 산출한 남녀 30세의 왜도( $T_{30}$  왜도)를 실제 자료상의

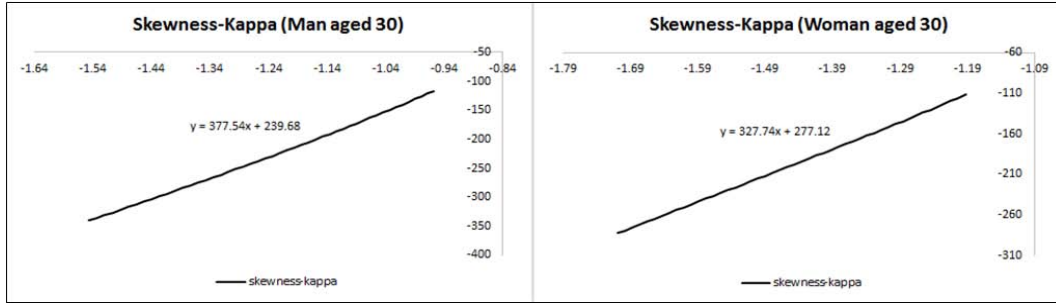


Figure 3.16. Relationship between  $\kappa_t$  and skewness of  $T_{30}$  in Canada by gender (2012–2060).

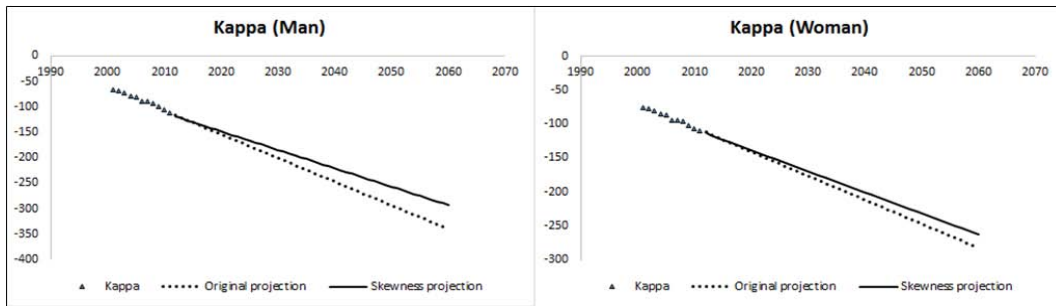


Figure 3.17. Projected mortality rate  $\kappa_t$  in Canada by gender (2011–2060) (2012–2060).

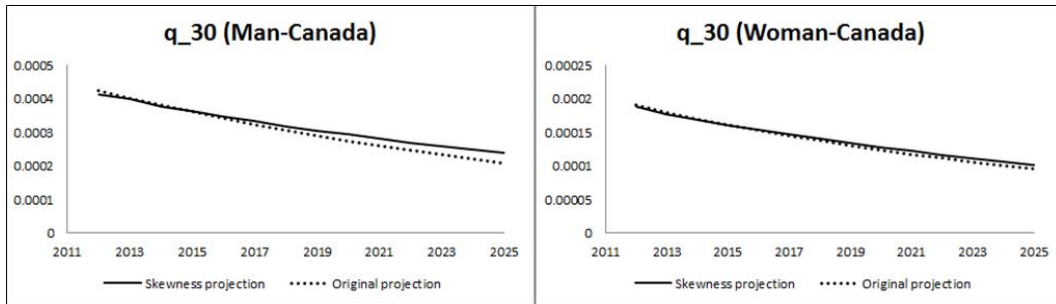


Figure 3.18. Projected  $q_{30}$  by gender in US by gender (2012–2024).

$T_{30}$  왜도와 비교하여 예측방식간의 정확도를 살펴보기 위해 Back-testing을 이용하였다. 본 연구에서는 미국 남녀  $T_{30}$  왜도를 사용하였으며 전체 생명표 자료 중 최근 2001년부터 2010년까지의 10년간의 자료만을 이용하였다. Back-testing방식은 다음과 같다. 2001–2007년까지의 7년간의 자료만으로 2008, 2009, 2010년 3년 동안의  $T_{30}$  왜도를 왜도 추세로 직접 예측하는 왜도예측방식과 기존의  $\kappa_t$ 의 추세를 이용하여 미래  $\kappa_t$ 를 예측하고 이를 LC 모형에 대입하여 왜도를 예측하는 기존예측방식으로 각각 예측한다. 그 후, 2001–2010년 10년간의 자료로 구한 실제  $T_{30}$ 의 왜도를 대조군으로 삼아 각 예측방식으로 산출한  $T_{30}$ 의 왜도와 비교한다. 다음 Figure 3.20은 그 결과를 그래프로 나타낸 것이다.

Figure 3.20에서도 알 수 있듯이, 2008–2010년의 실제 자료상 미국 남녀  $T_{30}$ 의 왜도와 왜도예측방식의  $T_{30}$ 의 왜도가 거의 일치하는 것을 볼 수 있다. 하지만, 기존예측방식의  $T_{30}$ 의 왜도는 실제 자료 및 왜도

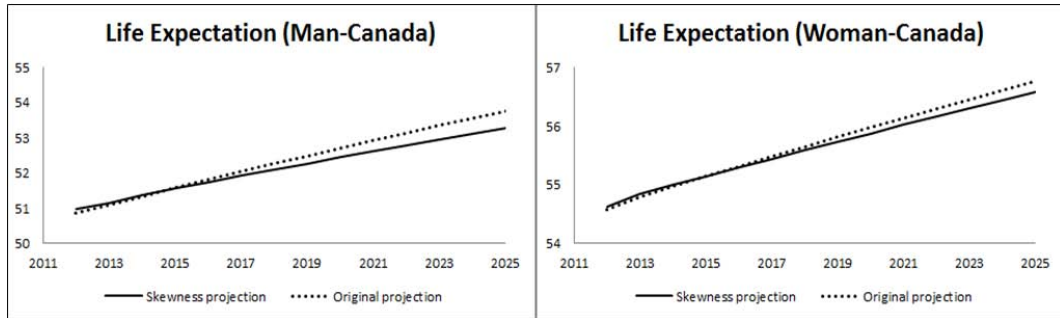


Figure 3.19. Projected life expectancy of  $T_{30}$  in Canada by gender (2012–2024).

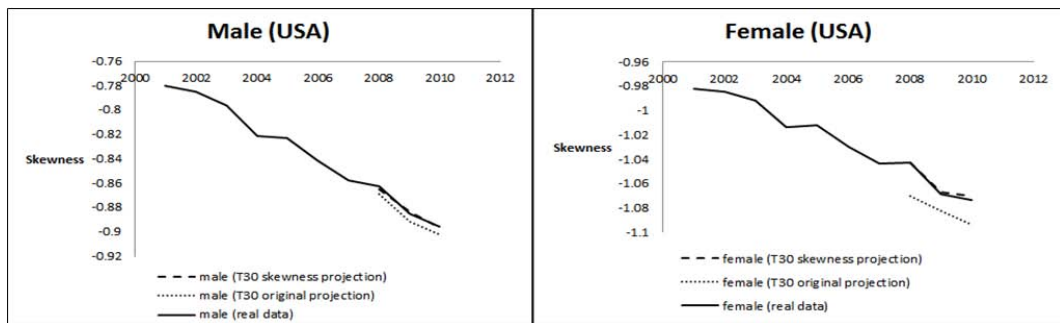


Figure 3.20. Back-testing of  $T_{30}$  skewness in US by gender (2001–2010).

예측방식의 결과와는 큰 차이가 벌어지는 것을 알 수 있다. 특히, 여성의 경우 남성에 비해 훨씬 큰 격차가 벌어지는 것을 볼 수 있으며 급격한 왜도의 감소추세로 이어지는 것을 볼 수 있다. 이러한 Back-testing 검증 결과는 본 연구에서 제안한 왜도예측방식이 기존예측방식에 비해 훨씬 실제 자료와 비슷한 추세로 미래 사망확률을 예측한다는 것을 입증할 수 있다.

### 3.5. 왜도 예측방식과 기존 예측방식 간 가중치 부여

본 연구에서는 예측치상의 과도하게 측정되는 사망률 개선 효과를 적절한 수준으로 보정하고자 근래의 사망확률분포의 편중을 반영한 왜도예측방식을 제안하였다. 그 결과, 기존예측방식을 적용할 경우 나타나는 사망률 개선 효과의 과대평가로 인한 사망률예측의 왜곡을 보정할 수 있었다. 하지만, 추정된 사망률지수의 차이가 미래 사망률의 예측치에 미치는 영향이 크기 때문에 본 연구에서 제안한 방식이 고려하는 장점들을 포함하면서 기존방식의 단점들을 줄이는 두 방식간의 중간점을 찾는 방안에 대해 고려해야한다. 본 장에서는 왜도예측방식과 기존예측방식으로 예측한  $k_t$ 의 장래추계 사이에 생기는 차이를 줄이기 위해 임의의 상수 가중치  $\alpha$ 를 부여한다. 가중치  $\alpha$ 를 적용한 다양한 통계량들의 예측치를 위의 그래프에 나타내 보고 그 차이를 살펴보았다 (단,  $\alpha = 0, 0.2, 0.4, \dots, 1$ ). Figure 3.21, Figure 3.22와 Figure 3.23은 앞서 언급했듯 두 방식간의 중간점을 찾기 위해 도입한 상수 가중치  $\alpha$ 에 따른 남녀 30세 기준의 사망확률과 기대여명, 왜도 등을 표현한 것이다.

하지만 임의적으로 부여한 가중치에 따른 사망확률 및 통계량이 다양하게 추정되므로 향후에는 적정 가중치 수준을 판단하기 위한 연구 및 적정 수준에 대한 적합성 검토과정이 필요할 것으로 보인다.



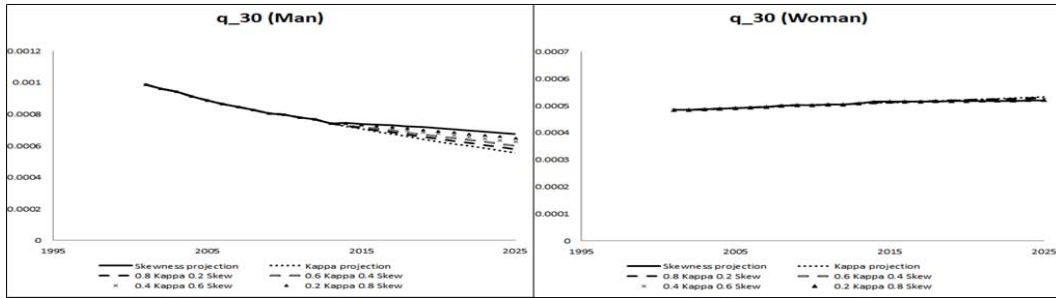


Figure 3.21. Weighted projection of  $q_{30}$  by gender (2001–2024).

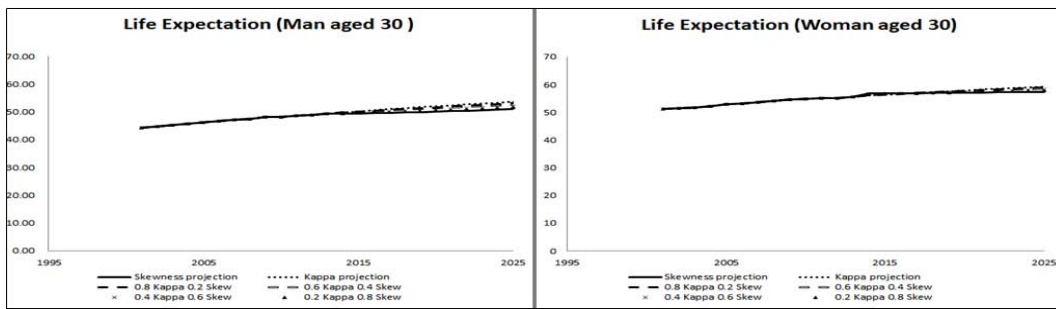


Figure 3.22. Weighted projection of life expectation of  $T_{30}$  by gender (2001–2024).

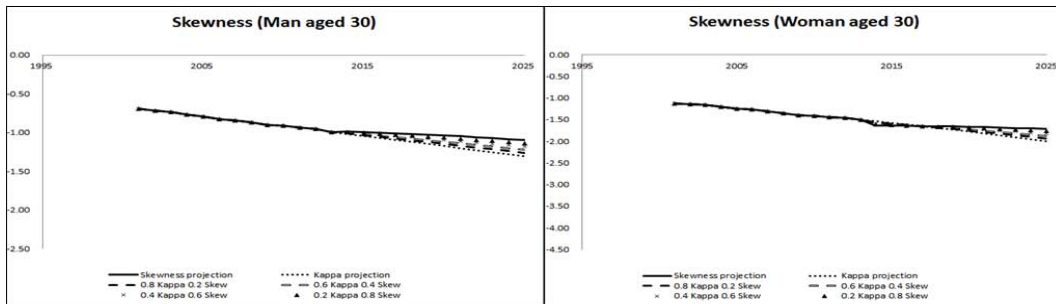


Figure 3.23. Weighted projection of  $T_{30}$  skewness by gender (2001–2024).

#### 4. 결론

본 연구는 미래 사망률을 예측하는 대표적인 확률적 사망률 모형인 LC 모형을 사용하여 과거자료의 왜도로 미래 왜도 추세를 예측하고, 기존방식으로 예측한 미래 왜도와 사망률 지수  $\kappa_t$ 의 비례관계를 이용하여 새로운 사망률 지수  $\kappa_t$ 를 간접적으로 예측하는 왜도예측방식을 제안한다. 기존의 LC 모형을 이용한 사망률 예측 방식은 미래 사망률이 지나치게 개선되는 현상을 보이며, 이를 바탕으로 산출된 연금액과 지급기간 추정 등 연금 공급자의 리스크 관리에 반영되는 장수리스크를 왜곡하게 하는 요인으로 작용한다. 본 연구에서는 과대평가되고 있는 사망률 개선 효과를 적절한 수준으로 보정하고자 사확률분포의 편증을 나타내는 왜도를 직접 사용하여 장래의 사망확률 및 기대여명을 추정하는 왜도 예측방식을 제안하였다. 왜도예측방식을 적용하여 산출한 각각표준편차와 첨도의 예측치는 과거 표준편차와 첨도 자료

등의 통계량으로 직접 예측한 각각의 장래추계와 매우 근사한 값을 보이지만, 기존의 LC 모형의 예측방식으로 산출한 통계량들과는 상당한 차이점이 나타난다. 이에 본 연구가 제안한 왜도예측방식을 이용한 장래사망률추계의 타당성 및 적합성을 뒷받침할 수 있다. 또한, 기존예측방식과 왜도예측방식을 미국과 캐나다의 생명표 자료에도 적용하였다. 그 결과 미국과 캐나다의 사망률 예측 역시 왜도예측방식이 기존에 비해 사망률 개선효과의 과도한 반영을 감소시킬 수 있었고, 장수리스크의 측정을 덜 왜곡시키는 것으로 판단되었다.

하지만, 왜도예측방식과 기존예측방식으로 도출된 사망률 지수들 간 차이가 현존하는 장래사망률추계와의 다소 큰 차이를 발생하게 한다. 본 연구에서는 추정된 사망률지수의 차이가 미래 사망률의 예측치에 미치는 영향이 크기 때문에 이를 감안하여 본 연구에서 제안한 방식이 고려하는 장점들을 포함하면서 기존방식의 단점들을 줄이는 두 방식간의 타협점을 찾는 방안에 대해 고려하였다. 위의 두 예측방식 간 임의의 상수 가중치  $\alpha$ 를 부여하여 산출한  $T_{30}$ 의 사망확률 및 다양한 통계량들의 예측치를 그래프를 통해 그 차이를 살펴보았다. 하지만 임의적으로 부여한 가중치에 따른 사망확률 및 통계량이 다르게 추정되므로 향후에는 적정 가중치 수치를 적용하기 위한 연구 및 검토과정이 필요할 것으로 보인다. 또한, 단기간의 집적 자료를 이용하여 자료 집적기간보다 4배 혹은 5배 이상의 기간을 예측하는 것은 너무 비약적이므로 주의가 필요하다. 본 연구에서는 2001년부터 2013년까지의 최근 13년간 생명표 자료를 사용하여 2014년부터 2025년까지의 통계량들을 예측해보았으나 위와 동일한 자료로 더 긴 기간을 예측할 경우에는 자료의 짧은 집적기간으로 인해 일시적인 원인에 의한 특정 년도의 사망률 감소현상이 큰 영향을 주어 전반적인 사망률 추세 예측을 왜곡시킬 수 있기 때문이다. 따라서 국민 생명표의 사망률 예측수치와의 비교를 통해 사망률의 차이 원인이 선택효과인지 표본오차인지 검토가 필요하다.

## References

- Alho, J. (2000). Discussion of "The Lee-Carter method for forecasting mortality, with various extension and applications" by R. Lee., *North American Actuarial Journal*, **4**, 91–93.
- Brouhns, N., Denuit, M., and Vermunt, J. (2002). A Poisson log-bilinear regression approach to the construction of projected lifetables, *Insurance: Mathematics and Economics*, **31**, 373–393.
- Cairns, A., Blake, D., and Dowd, K. (2008). Modeling and management of mortality risk: A review, *Scandinavian Actuarial Journal*, **2**, 79–113.
- Jeong, S. H. and Kim, K. W. (2011). A comparison study for mortality forecasting models by average life expectancy, *Korean Journal of Applied Statistics*, **24**, 115–125.
- Joo, H. C. and Lee, H. S. (2014). The effect of mortality improvement on benefit reserves, *The Journal of Risk Management*, **25**, 3–33.
- Kim, S. J. (2012). A study on the prediction of mortality rate using Lee-Carter model, *The Journal of Actuarial Science*, **4**, 48–66.
- Kim, S. Y. (2011). Mortality forecasting for the republic of Korea: the Coherent Lee-Carter method, *Korea Journal of Population Studies*, **34**, 157–177.
- Lee, R. D. and Carter, L. (1992). Modeling and forecasting US mortality, *Journal of the American Statistical Association*, **87**, 659–671.
- Lee, R. and Miller, T. (2001). Evaluating the performance of the Lee-Carter method for forecasting mortality, *Demography*, **38**, 537–549.
- Li, N. and Lee, R. (2005). Coherent mortality forecasts for a group of populations: an extension of the Lee-Carter method, *Demography*, **42**, 575–594.
- Park, Y. S., Jang, S. W., and Kim, S. Y. (2013). VECM-LC model for forecasting mortality in Korea, *The Korean Association for Survey Research*, **14**, 19–47.
- Vermunt, J. K. (1997). *LEM: A General Program for the Analysis of Categorical Data*, Department of Methodology and Statistics, Tilburg University.

# 왜도 예측을 이용한 Lee-Carter 모형의 사망률 예측

이항석<sup>a,1</sup> · 백창룡<sup>b</sup> · 김지현<sup>c</sup>

<sup>a</sup>성균관대학교 보험계리학과, <sup>b</sup>성균관대학교 통계학과, <sup>c</sup>MG 손해보험

(2015년 10월 27일 접수, 2015년 12월 17일 수정, 2015년 12월 17일 채택)

---

## 요약

지속적인 사망률 개선으로 인한 평균 수명연장은 인구 고령화의 주요인이며 연금 공급자의 재정건전성에 심각한 영향을 미치는 원인으로 지목되기에 정확한 미래 사망률의 예측은 현 시점에서 선행되어야 할 중요한 과제다. 본 연구는 미래 사망률을 예측하는 대표적인 확률적 사망률 모형인 Lee-Carter 모형을 사용하여 과거 생명표로 산출한 왜도를 기반으로 미래 사망률 지수를 간접적으로 예측하는 왜도예측방식을 제시한다. 기존의 Lee-Carter 모형을 이용한 사망률 예측방식은 사망률 지수를 추정하고 미래값을 직접 예측함으로써 미래 사망률이 지나치게 개선되는 현상을 보이며, 이를 바탕으로 산출된 연금액과 지급기간 추정 등 연금 공급자의 리스크 관리에 영향을 미친다. 본 연구는 기존 예측 방식의 사망률 예측 결과와 제시한 왜도 예측 방식의 사망률 예측 결과를 비교함으로써 기존 사망률 예측 방식의 문제점을 지적한다. 분석결과 왜도 예측을 통한 Lee-Carter 모형의 사망률 예측은 기존 방식보다 사망률 개선효과를 더 적게 반영하며 장수리스크를 덜 왜곡한다는 데 의의가 있다고 할 수 있다. 하지만 기존 방식 간 차이를 감안하여 적절한 미래 사망률 수준을 찾기 위해 임의로 부여한 가중치에 대해 향후 검토가 필요할 것으로 보인다.

주요용어: Lee-Carter 모형, 장수리스크, 사망률 예측, 사망률 지수, 왜도

---

이 논문은 2014년도 정부 (미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구 사업임 (NRF-2014R1A1A1006025).

<sup>1</sup>교신저자: (03063) 서울특별시 종로구 성균관로 25-2, 성균관대학교 보험계리학과/수학과.

E-mail: hangsuck@skku.edu.