

생명표 기반 함수로부터 산출된 사망률 지표의 분해 분석 : R을 활용한 단계별 치환 알고리즘과 등치선 치환 분해법 적용

김익한*

초록

인구 집단 간 혹은 같은 집단의 이산적인 두 시점 간 사망률 격차에 대한 요인별 기여도를 산출하기 위하여 분해 분석법이 활용되었다. 사망률 지표 산출 함수가 비선형적 특성을 가지는 경우, 전통적인 가산적 분해 분석법의 적용은 결과에 비뚤림을 불러올 수도 있다. 2000년대에 접어들며 연령표준화 사망률과 기대여명 외에도 생명표 기반 수명 변이 지표들이 이용되기 시작하며, 다양한 지표들을 공통된 방법으로 분해 분석을 수행할 수 있는 일반화 분해 분석법이 개발되었다. 이 연구에서는 일반화 분해 분석 방법 중 단계별 치환 알고리즘에 관하여 소개하고, 통계 소프트웨어 R을 활용한 실제 적용 방법을 고찰하였다. 또한 시점과 집단이 모두 두 개인 경우, 단계별 치환 알고리즘에 기반하여 시작 시점 집단 간 사망률 격차와 집단 간 사망률 변화 추세 차이에 대한 기여도를 합하여 현재 시점의 집단 간 사망률 격차를 산출할 수 있는 등치선 치환 분해법을 이용한 분해법을 함께 소개하였다.

주요 용어: 사망원인, 생명표, 수명 연한, 알고리즘

* 고신대학교 의과대학 (E-mail: ikhan.kim@kosin.ac.kr)

투고일: 2026.1.15./ 수정일: 2026.1.23./ 게재확정일: 2026.2.24.

I. 서론

1. 사망률 측정과 분해 분석

연령이나 사망원인, 교육 수준 등 하위집단에 따른 사망률을 종합하여 한 집단의 사망 수준을 나타내기 위하여 다양한 사망률 지표가 활용되었다. 대표적인 사망률 지표로는 연령표준화 사망률(age-standardized mortality rate), 기대여명(life expectancy), 그리고 수명 변이(lifespan variation)가 있다. 연령표준화 사망률은 개별 집단의 연령 구조가 표준인구의 구조와 같다고 가정했을 때의 사망률로 정의된다. 연령표준화 방법은 산출이 간편하고 집단 간 비교가 쉽지만, 표준인구의 선택에 따라서 다른 결과를 나타낼 가능성이 있다(Bhopal, 2016). 기대여명과 수명 변이 지표는 모두 생명표를 기반으로 한다. 두 지표 모두 사망률 측정에 많은 시간과 비용이 수반되는 코호트 자료를 이용하여 측정되기보다는, 비교적 짧은 기간 동안 측정된 연령군별 사망률을 통하여 구축된 기간 생명표(period life table)에 기반하여 산출되는 경우가 많다. 기간 생명표를 이용한 x 세 기대여명은 x 세 인 사람이 해당 기간의 x 세 이후 연령군별 사망률을 그대로 겪는다고 가정할 때 생존할 것으로 기대되는 연수로 정의된다(Preston et al., 2000). 0세 기대여명은 기대수명(life expectancy at birth)으로도 일컬어진다. 기대여명은 이해가 쉽고, 표준인구가 필요 없다는 장점이 있다(Silcocks et al., 2001). 수명 변이는 집단 내 사망 패턴의 이질성(heterogeneity) 정도를 알려주는 지표들을 통칭한다(van Raalte et al., 2018). 수명 변이 지표의 종류로는 지니 계수, 타일 지수, 표준편차 등이 있다.

동일한 집단의 두 이산적인 시점 간 사망률 차이 혹은 같은 시점에서 서로 다른 두 집단 간 사망률 격차를 설명하기 위하여 분해 분석(decomposition analysis)이 이용되었다(Arriaga, 1984; Andreev et al, 2002; Jdanov et al.,

2017; 2024). 분해 분석은 사망률 차이에 대한 특정 연령군이나 사망원인 등의 기여도를 파악하고자 할 때 이용되는데, 분해 분석을 수행하는 이유는 사망률 변화가 반드시 모든 연령이나 사망원인에서 같은 양상으로 이루어지지 않기 때문이다(Arriaga, 1984). 분해 분석은 1955년 Kitagawa에 의하여 집단 간 조사사망률(crude death rate) 격차를 분석하기 위하여 최초로 수행되었다. 그녀는 인구 집단 간 조사사망률 격차를 실제 사망률 차이와 인구 구성 차이로 나누는 간단한 공식을 제시하였다(Kitagawa, 1955; Jdanov et al., 2017). 1990년대 초, Das Gupta는 Kitagawa의 분해 분석 방법을 확장하여 연령군뿐만 아니라 다른 하위집단을 함께 고려하는 다차원 분해 분석법을 제시하기도 하였다(Das Gupta, 1991; 1994). 기대여명에 대한 분해 분석은 1980년대 일군의 연구자가 독립적으로 집단 간 기대수명 격차에 대한 연령군별 기여도를 산출하는 공식을 도출하여 이루어졌다(Andreev, 1982; Arriaga, 1984; Pollard, 1982; Pressat, 1985). 4명의 연구자가 각기 제시한 방법 중 Andreev, Arriaga, Pressat이 제시한 공식은 결과적으로 같은 결과를 도출한다고 알려져 있다(Jdanov et al., 2017).

2. 사망률 산출 함수의 비선형성이 분해 분석에 미치는 영향

전통적인 분해법에서는 요인별 기여도 총합이 사망률 추정치 격차와 같다고 가정한다. 이 가정은 표준인구의 연령군별 비율과 대상 집단의 연령군별 사망률 간 곱을 모두 더하여 산출하는 연령표준화 사망률에서는 타당하다.¹ 그러나 생명표를 기반으로 하는 기대여명이나 수명 변이 산출 함수는 비선형적(non-linear) 특성을 가지므로, 전통적인 가산적 분해 분석 방법의 적용이 적절하지 않을 수 있다. 기간 생명표를 이용한 x 세 기대여명은 x 세에 진입하는 일정 규모의 코호트를 가정하고 x 세 이

후 각 연령대에 진입하는 생존자 수를 이용하여 산출된다. 따라서 이전 연령에서의 생존 확률은 이후 연령에 누적해서 영향을 미치기 때문에, 특정 연령군의 사망률 변화는 해당 연령군 이후의 생존 확률에도 영향을 미친다. 이러한 특성으로 인하여 영유아기 사망률은 소폭으로 감소하여도 기대수명이 크게 증가할 수 있지만, 노년기 사망률 감소는 영유아기에 비하여 적은 기대수명 증가를 불러온다(Olshansky et al., 1990). 기대여명의 추가적인 변화는 이미 달성된 기대여명 수준에 의하여 결정된다고 할 수 있다(Arriaga, 1984).

사망률 지표 산출 함수가 비선형성을 가질 때는 분해 분석 결과가 비대칭성(non-symmetrical)과 비이행성(non-transitive)을 가진다(Horiuchi et al., 2008). 비대칭성이란 ‘가’ 집단을 기준으로 ‘나’ 집단을 비교했을 때의 분해 결과와 ‘나’ 집단을 기준으로 ‘가’ 집단을 비교했을 때의 결과가 다르게 나타나는 특성이다. 비이행성이란 ‘가’ 집단과 ‘나’ 집단을 비교한 뒤, 이어서 ‘다’ 집단과 비교한 분해 분석 결과가 ‘가’ 집단과 ‘다’ 집단을 직접 비교한 결과와 일치하지 않음을 의미한다. 비이행성은 시점과 집단이 모두 두 개인 경우, 시작 시점에서 두 집단의 기대여명 격차와 두 집단 간 기대여명 변화 추세 차이에 대한 기여도를 합한 결과가 종료 시점의 두 집단 간 기대여명 차이를 직접 분해 분석한 결과와 다를 수도 있다는 점을 의미한다. 비이행성은 경로 의존성(path-dependence)이라고 표현되기도 하는데, 이 연구에서는 비이행성으로 통일하여 표기하였다(Andreev et al., 2002; Jdanov et al., 2024). 기대여명이나 수명 변이처럼 비선형성을 가지는 함수를 이용했을 때의 사망률 차이를 어떻게 가산적으로 나타낼 것인가는 분해 분석의 중요한 도전 과제라고 할 수 있다(Jdanov et al., 2024).

Arriaga(1984)는 기대여명 산출 함수의 비선형성으로 인하여 연령군별 사망률 변화는 (1) 직접 효과 (2) 간접 효과 (3) 상호작용 효과를 가진다고 주장하였다. 직접 효과는 특정 연령군에서의 사망률 변화가 해당 연령군에서 기대되는 생존 연수(life years)에 미치는 변화를 의미한다. 간접 효과는 특정 연령군에서의 사망률 변화가 해당 연령군 이후로도 각 연령군에 진입하는 생존자 수에 영향을 미침으로 발생한다. 따라서 직접 혹은 간접 효과는 특정 연령군의 사망률이 변화할 때 기대여명에 미치는 영향이라고 할 수 있다. 그러나 연령군별 사망률 변화는 한 연령군에서만 나타나는 것이 아니라 모든 연령군에서 동시에 일어난다. Arriaga는 모든 연령군에서의 동시적 사망률 변화가 일으키는 기대여명 차이를 상호작용 효과로 정의하였다. x 세부터 $x+i$ 세까지 연령군의 사망률 변화에 따라서 변화된 수의 생존자가 $x+i$ 세 이후 연령군에 진입하는데, $x+i$ 세 이후 연령군별 사망률 또한 x 세부터 $x+i$ 세까지의 연령군과 동시에 변화하므로 $x+i$ 세 이후 각 연령군에 진입하는 사람의 수는 $x+i$ 세 이후의 연령군별 사망률이 변화하지 않을 때와 다르다. 연령군별 사망률 변화가 서로 영향을 미치며 기대여명 변화에 기여하는 부분이 상호작용 효과이다.

3. 일반화 분해 분석법의 개발과 발전

2000년대에 접어들며 기대여명뿐만 아니라 수명 변이 지표의 산출과 활용은 보편적으로 사용할 수 있는 분해 분석법의 필요를 불러왔다(Andreev et al., 2002). 일반화 분해 분석법은 비선형적 사망률 산출 함수를 가지는 요약 지표들에 공통적으로 적용 가능한 분해 분석법을

1 연령표준화 사망률(Age-Standardized Mortality Rate, ASMR) 산출 함수는 아래 수식으로 정의할 수 있다.

$$ASMR = \sum_x w_x \cdot r_x$$

이때, w_x 는 표준인구의 x 세 연령군의 비율, r_x 는 사망률이다.

의미한다. 현재 널리 활용되고 있는 두 가지 일반화 분해 분석법 중 먼저 개발된 분해 분석법은 단계별 치환 알고리즘이다. Andreev et al.(2002)은 최저 연령군부터 최고 연령군까지의 사망률을 비교하고자 하는 기준 집단의 연령군별 사망률로 순차적으로 대체할 때의 지표 변화량을 각 연령군의 기여도로 제시하였다. 단계별 치환 알고리즘을 기반으로 하여 Jdanov et al.(2017)은 시점과 집단이 모두 두 개인 경우 적용할 수 있는 등치선 치환 분해법(contour replacement decomposition method)을 제안하였다. 두 번째 방법은 Horiuchi et al.(2008)이 제시한 연속 변화 모형이다. 저자들은 연령이나 사망원인별 사망률이 시점이나 집단에 따라서 연속적으로 변화한다는 가정에 기반하여, 사망률의 변화량을 작은 단위로 나누면 비선형적 특성을 가지는 함수에서도 가산적인 분해 분석이 가능하다고 주장하였다. 실제 분석에서는 단계별 치환 알고리즘과 연속 변화 모형 방법 간 결과 차이는 크지 않다고 알려져 있다(Jdanov et al., 2017; 2024). 또한 연속 변화 모형에 기반해서도 등치선 치환 분해 분석이 가능하다(Jdanov et al., 2024). 이 연구는 단계별 치환 알고리즘과 해당 알고리즘에 기반한 등치선 치환 분해법을 개괄적으로 설명하고, 예시 자료와 통계 소프트웨어 R을 이용하여 실제 두 분해법의 적용 과정을 설명하는 데 목표를 두었다(R Core Team, 2025). 사망률 지표 산출 함수는 기대여명 산출 함수를 이용하였다.

II. 단계별 치환 알고리즘과 등치선 치환 분해법 적용

1. 자료원

이 연구에서는 이전 연구에서 활용했던 자료를 이용하

여 2011년 우리나라 남성의 대도시와 농촌 지역 간 기대수명 격차를 단계별 치환 알고리즘과 2010년 두 지역 간 남성 기대수명 격차 크기를 고려하는 1차원 및 2차원 등치선 분해법을 이용하여 분해 분석하였다(Kim, 2023). 1차원 등치선 분해법은 연령군별 기여도, 2차원 등치선 분해법에서는 연령군과 사망원인별 기여도를 산출하였다. 시군구별로 인구수와 사망원인별 사망자 수는 국가 통계포털(Korea Statistical Information Service)의 주민등록연앙인구와 마이크로데이터 통합서비스(MicroData Integrated Service)의 사망원인통계를 통하여 획득하였다. 인구수와 사망자 수 획득 시 연령군은 5세 단위(0, 1-4, ... 85세 이상)로 구분하였다. 우연에 의한 효과를 상쇄하기 위하여 2010년 시군구 및 연령군별 인구수와 사망원인별 사망자 수는 2009년부터 2011년까지의 자료를 활용하여 가중치를 적용하지 않는 3년 이동 평균 방법을 통하여 평활화하였다. 같은 방법으로 2011년도 자료는 2010-2012년 자료를 이용하여 평활화하였다. 행정구역 분류에 따라서 ‘구’ 지역은 대도시, ‘군’ 지역은 농촌으로 분류한 후, 대도시와 농촌으로 분류된 지역의 연도에 따른 연령군별 인구수와 사망자 수를 모두 합하여 대도시와 농촌 지역의 2010년과 2011년 연령군별 인구수와 연령군 및 사망원인별 사망자 수를 산출하였다. 이 연구에 이용한 자료에 관한 자세한 사항은 다른 문헌에 자세히 설명되어 있다(Kim, 2023). 이 연구에서 사망원인은 분석의 편의를 위하여 암, 심뇌혈관질환, 호흡기계질환, 소화기계질환, 외인사, 기타 등으로 재분류하였다. 모든 분석은 통계 소프트웨어 R을 이용하여 진행하였으며, 분석의 투명성과 재현성을 위하여 Quarto 도구를 이용하여 작성한 코드와 결과를 포함하였다. 분석에 이용한 모든 자료는 웹폴더(https://drive.google.com/drive/folders/10QZPh_jlMwe2alZEnD0vmTVH1VXorWWr?usp=sharing)에서 다운로드 가능하다.

2. 단계별 치환 알고리즘

단계별 치환 알고리즘은 연령군별 사망률을 한 단계씩 교체할 때, 기대여명이나 수명 변이 같은 사망률 지표가 얼마나 변화하는지로부터 연령군별 기여도를 계산한다(Andreev et al., 2002). 단계별 치환 알고리즘은 연속 변화 모형과 달리 다른 연령군의 사망률은 고정된 상태에서 하나의 연령군별 사망률만 변화한다고 가정한다. \mathcal{X} 세 연령군의 ‘가’ 집단과 ‘나’ 집단 간 기대수명 격차에 대한 기여도는 \mathcal{X} 세 연령군까지 ‘나’ 집단의 연령군별 사망률로 ‘가’ 집단의 사망률을 치환하여 산출한 기대수명으로부터 $\mathcal{X} - n$ 세까지 같은 방식으로 치환하여 산출한 기대수명 산출치를 뺀 값이다. 이때, n 은 연령군의 너비이다. 두 집단 중 어느 집단을 기준으로 두느냐에 따라서 분해 분석 결과가 달라질 수 있기 때문에 ‘가’와 ‘나’ 집단을 기준으로 하는 분석을 모두 수행하고, 두 분석 결과의 평균값을 최종 분해 분석 결과로 간주한다(Andreev et al., 2002). 단계별 치환 알고리즘은 치환 순서에 따라서 값이 달라질 수 있다. 연령의 경우에는 최저 연령군으로부터 최고 연령군까지 오름차순으로 교체한다는 합의에 따라서 분석을 수행한다(Andreev et al., 2002; Jdanov et al., 2017; 2024). 그러나 사망원인이나 교육 수준처럼 명시적인 순서가 없는 경우에는 명확한 순서가 정해져 있지 않다. 이에 관해서는 2차원 등치선 분해법을 다른 부분에서 설명하도록 하겠다.

예시에서는 2011년 우리나라 남성의 대도시와 농촌 지역 간 기대수명 격차에 대한 연령군별 기여도를 단계별 치환 알고리즘을 이용하여 분석하였다. 연령군은 이름이 Age인 벡터에 저장하였다. 추후 연령군과 함께 사망원인에 따른 2차원 분해 분석을 수행하기 위하여 사망원인별로 나눠 사망률 데이터를 구축하였다. 2011년 대도시 지역의 연령군과 사망원인별 사망률을 데이터 A, 농촌 지역의 사망률을 데이터 B로 저장하였다. 연령군이 총 19개, 사망원인이 총 6개이므로 A와 B 데이터 모두

19행 6열로 구성되어 있다. 사망원인 순서는 암, 심뇌혈관질환, 호흡기계질환, 소화기계질환, 외인사, 기타 순이다. A 데이터와 B 데이터의 처음 6개 행의 사망률은 아래와 같다.

```
Age =c(0, 1, seq(5, 85, by = 5))
dim(A)
[1] 19 6
dim(B)
[1] 19 6
head(A)
      [1] [2] [3] [4] [5] [6]
[1.] 3.5409e-05 4.2491e-05 7.0818e-05 1.4164e-05 1.4164e-04 3.0664e-03
[2.] 1.0337e-05 3.6180e-05 3.7903e-05 1.7229e-06 6.0300e-05 7.5806e-05
[3.] 1.4495e-05 2.1083e-05 3.9531e-06 1.3177e-06 4.3484e-05 2.6354e-05
[4.] 6.7409e-06 3.0815e-05 1.9260e-06 1.9260e-06 5.4890e-05 3.9482e-05
[5.] 2.1541e-05 4.2254e-05 2.4855e-06 8.2850e-07 2.3364e-04 3.8111e-05
[6.] 2.2437e-05 5.2353e-05 3.7395e-06 1.8698e-06 3.1973e-04 5.8898e-05

head(B)
      [1] [2] [3] [4] [5] [6]
[1.] 0.0000e+00 6.0259e-05 1.2052e-04 6.0259e-05 4.2181e-04 3.6155e-03
[2.] 2.9547e-05 4.4321e-05 1.4774e-05 1.4774e-05 1.3296e-04 4.4321e-05
[3.] 0.0000e+00 3.4241e-05 1.1414e-05 0.0000e+00 4.5655e-05 4.5655e-05
[4.] 1.7107e-05 2.5661e-05 8.5536e-06 0.0000e+00 8.5536e-05 5.1322e-05
[5.] 7.5692e-06 4.5415e-05 7.5692e-06 0.0000e+00 3.8603e-04 4.5415e-05
[6.] 6.3398e-05 3.1699e-05 0.0000e+00 7.9248e-06 6.6568e-04 7.9248e-05
```

A와 B 데이터를 기대수명 산출 함수인 `ex.per.matrix()`에 투입하여 2011년 우리나라의 대도시와 농촌 지역 기대수명을 산출하고, 두 집단 간 기대수명 격차를 구하였다. `ex.per.matrix()`는 연령군별로 사망원인에 따른 사망률을 모두 합하여 총사망률을 구한 뒤 기대여명을 산출한다. `ex.per.matrix()` 함수는 연령군별 사망원인에 따른 사망률과 연령 벡터를 차례로 투입하여 실행한다.

```
ex.per.matrix(A, Age)
[1] 78.114
ex.per.matrix(B, Age)
[1] 75.624
ex.per.matrix(A, Age) -ex.per.matrix(B, Age)
[1] 2.4903
```

기대수명 산출 결과, 2011년 우리나라 대도시 지역 남성의 기대수명은 78.11세, 농촌 지역은 기대수명이 75.62세로 나타났다. 두 집단 간 기대수명 격차는 2.49

세였다.

다음은 `stepwise_contrib()` 함수를 이용하여 단계별 치환 알고리즘 분해 분석을 수행한다. `stepwise_contrib()` 함수에는 분해 분석하고자 하는 두 집단의 사망률 자료와 연령 벡터를 차례대로 투입한다. A, B 데이터 각각을 기준으로 설정한 분해 분석을 수행한 후, 두 분석 결과의 평균을 구하여 최종 분해 분석 결과를 산출하였다.

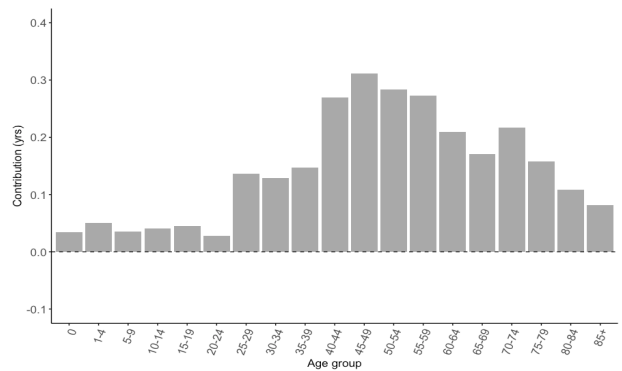
```
diff.AB = stepwise_contrib(A, B, Age)
diff.BA = stepwise_contrib(B, A, Age)
contrib_mean <- (diff.AB + (-diff.BA)) / 2
print(contrib_mean)

[1] 0.0694621 0.0172722 0.0091185 0.0168795 0.0455267 0.1059792
0.0708362 0.1397710 0.1481597 0.2028468 0.2593483 0.2923513 0.2715947
0.2095837 0.2006616 0.1870613 0.1591493 0.0496839 0.0349654

sum(contrib_mean)

[1] 2.4903
```

기대수명 격차에 대한 연령군별 기여도는 `contrib_mean` 벡터에 저장되어 있다. 분석 결과, 기여도 합계(약 2.49세)는 두 집단 간 기대수명 격차를 산출했을 때와 동일하였다. 연령군별 기여도가 양수이면 대도시 지역, 음수이면 농촌 지역에 유리한 결과이다. 연령군별 기대수명 격차에 대한 기여도를 막대 그래프로 나타낸 결과는 <그림 1>과 같다. 2011년 우리나라 남성의 대도시 지역과 농촌 지역 간 기대수명 격차에 크게 기여한 연령군은



주: 기여도가 양수이면 대도시, 음수이면 농촌 지역에 유리한 결과이다.

[그림 1] 단계별 치환 알고리즘을 이용한 2011년 우리나라 남성의 대도시와 농촌 지역 간 기대수명 격차에 대한 연령군별 기여도

45-49, 50-54, 55-59세 순으로 나타났다. 0-24세 연령군에서는 비교적 적은 차이를 나타냈고, 80세 이상 연령군 또한 두 집단 간 기대수명 격차에 기여하였다.

3. 등치선 치환 분해법

1) 1차원 등치선 치환 분해법

등치선 치환 분해법은 두 집단의 현재 기대여명 격차를 시작 시점에서의 차이와 변화 추세 차이에 대한 기여도로 분해 분석하는 방법이다(Jdanov et al., 2017). 1차원 등치선 치환 분해법에서는 기여 요인으로 연령군만을 고려하였다. 등치선 치환 분해법은 현재 시점에서의 두 집단 간 기대여명 격차가 과거 시점에서의 기대여명 격차와 시간에 따른 변화 추세 차이의 산물이라는 생각으로부터 시작되었다(Jdanov et al., 2017). 등치선 치환 분해법을 이용하는 이유는 기대여명 산출 함수의 비선형성으로 인하여 현재 시점의 기대수명 격차를 직접 분해 분석한 결과와 과거 시점 연령군별 기여도와 집단 간 변화 추세 차이에 대한 기여도를 합한 결과가 다를 가능성이 있기 때문이다. Jdanov 등은 2017년 발표한 연구에서 1980년과 2010년 잉글랜드 및 웨일스와 미국 간 기대수명 격차 분해 분석 결과를 예로 들었다(Jdanov et al., 2017). 저자들이 전통적인 기대수명 분해 분석법을 이용하여 산출한 1980년의 두 국가 간 기대수명 격차에 대한 40-59세 연령군의 기여도는 0.50세였다. 2010년 자료를 이용했을 때는 해당 연령군의 기여도가 0.92세로 나타나, 1980년 대비 2010년 40-59세 연령군의 기여도가 0.42세 더 컸다. 그러나 국가별로 1980년과 2010년 기대수명 차이에 대하여 40-59세 연령군의 기여도를 산출했을 때는 잉글랜드 및 웨일스는 1.59세, 미국은 1.29세로 나타나, 두 국가의 40-59세 기여도는 약 0.30세의 차이를 보였다. 이 결과는 연구 기간 범위가 넓을수록 시작 시점의 기대수명 격차에 대한 요인별 기여도와 종료 시

점의 기여도를 단순히 비교할 수 없음을 보여주며, 시작 시점에서의 기여도를 종료 시점의 사망률 구조에 맞춰 보정할 필요가 있음을 알려준다(Jdanov et al., 2017). 이를 위하여 1차원 등치선 치환 분해법은 단계별로 등치선을 그리듯이, 비교하고자 하는 집단의 시작과 종료 시점, 시작 시점의 비교 집단과 기준 집단 간, 마지막으로 기준 집단의 시작과 종료 시점의 연령군별 사망률을 차례대로 치환한다. 등치선 치환 분해법에 대한 더 자세한 설명은 Jdanov et al.(2017)의 논문을 참조할 수 있다.

등치선 분해법 분석을 통하여 앞서 분석하였던 2011년 우리나라 남성의 대도시와 농촌 지역 간 기대수명 격차를 2010년 두 집단 간 기대수명 격차와 2010-2011년 동안 두 집단의 기대수명 변화 추세 차이에 대한 연령군별 기여도를 합하여 나타냈다. 2011년 대도시(A)와 농촌(B) 지역 데이터에 더하여 2010년 대도시와 농촌 지역의 연령군과 사망원인별 사망률을 데이터 a와 b로 저장하고, 지역별로 2010년과 2011년 자료를 합하여 각각 데이터 c1과 c2로 저장하였다. c1과 c2 데이터의 첫 번째 열을 연령군별 시작 나이로 설정하였다.

```
dim(a)
[1] 19 6
dim(b)
[1] 19 6
head(a)
      [1]      [2]      [3]      [4]      [5]      [6]
[1.] 3.6328e-05 4.3594e-05 7.2656e-05 7.2656e-06 1.9617e-04 3.5020e-03
[2.] 1.7316e-06 1.3853e-05 2.2511e-05 5.1949e-06 7.2729e-05 8.4850e-05
[3.] 8.6021e-06 2.9493e-05 3.6866e-06 2.4577e-06 4.6697e-05 3.0722e-05
[4.] 1.2022e-05 2.7742e-05 9.2474e-07 0.0000e+00 5.0860e-05 3.0516e-05
[5.] 1.3244e-05 4.6352e-05 5.7940e-06 8.2772e-07 2.2845e-04 4.8008e-05
[6.] 3.0294e-05 3.7867e-05 3.7867e-06 3.7867e-06 3.8625e-04 5.3014e-05
head(b)
      [1]      [2]      [3]      [4]      [5]      [6]
[1.] 1.2325e-04 0.0000e+00 6.1626e-05 0.0000e+00 4.9301e-04 3.6359e-03
[2.] 0.0000e+00 2.9694e-05 0.0000e+00 1.4847e-05 1.9301e-04 1.3362e-04
[3.] 0.0000e+00 6.3929e-05 0.0000e+00 0.0000e+00 9.5893e-05 6.3929e-05
[4.] 1.6661e-05 2.4992e-05 8.3306e-06 1.6661e-05 1.1663e-04 6.6645e-05
[5.] 3.8239e-05 3.0592e-05 0.0000e+00 0.0000e+00 3.1356e-04 1.1472e-04
[6.] 3.1236e-05 4.6854e-05 0.0000e+00 7.8090e-06 4.9978e-04 3.1236e-05
c1 = cbind(Age, cbind(a, A))
c2 = cbind(Age, cbind(b, B))
```

2010년 우리나라 남성의 대도시와 농촌 지역 기대수명을 ex.per.matrix() 함수를 이용하여 산출한 결과, 대도시 지역은 77.75세, 농촌 지역은 75.02세로 나타나 두 집단 간 기대수명 격차는 2.73세였다.

```
ex.per.matrix(a, Age)
[1] 77.745
ex.per.matrix(b, Age)
[1] 75.017
ex.per.matrix(a, Age) - ex.per.matrix(b, Age)
[1] 2.7285
```

등치선 분해법은 decomp.contour() 함수를 이용하였다. decomp.contour() 함수는 등치선 분해법을 제안한 Jdanov 등이 만들어 공용 배포한 코드를 이 연구의 데이터 분석에 맞춰 일부 변형하였다(Jdanov & Shkolnikov, 2014). decomp.contour() 함수에 차례대로 c1과 c2 데이터, 그리고 기대수명 산출 함수명인 ex.per.matrix를 입력한다. 분석 결과는 데이터 프레임 형태로 산출되는데, 시작 시점(2010년)의 기대수명 격차에 대한 연령군별 기여도, 두 집단의 기대수명 변화 추세 차이에 대한 연령군별 기여도, 그리고 종료 시점(2011년)에서의 두 집단 간 기대수명 격차에 대한 기여도가 포함된다. 등치선 분해법은 시작 시점의 두 집단 간 기대수명 격차에 대한 기여도와 변화 추세 차이에 대한 기여도를 합하여 현재 시점에서의 두 집단 간 기대수명 격차에 대한 기여도를 산출하므로 분해 분석 결과 마지막 열의 종료 시점 기여도는 비교 및 확인용으로 쓰인다.

```
contour_res = decomp.contour(c1, c2, FUN = ex.per.matrix)
print(contour_res)
      initial      trend      conventional
0  0.034870  0.03459247  0.0694621
1  0.050319 -0.03304723  0.0172722
5  0.035418 -0.02629908  0.0091185
10 0.041181 -0.02430165  0.0168795
15 0.045912 -0.00038490  0.0455267
20 0.027783  0.07819649  0.1059792
25 0.137163 -0.06632659  0.0708362
30 0.130645  0.00912550  0.1397710
35 0.148938 -0.00077837  0.1481597
```

```

40 0.272302 -0.06945561 0.2028468
45 0.316006 -0.05665802 0.2593483
50 0.288761 0.00359073 0.2923513
55 0.278148 -0.00655328 0.2715947
60 0.214152 -0.00456829 0.2095837
65 0.175291 0.02537069 0.2006616
70 0.222630 -0.03556847 0.1870613
75 0.163439 -0.00428958 0.1591493
80 0.114246 -0.06456253 0.0496839
85 0.084699 -0.04973399 0.0349654

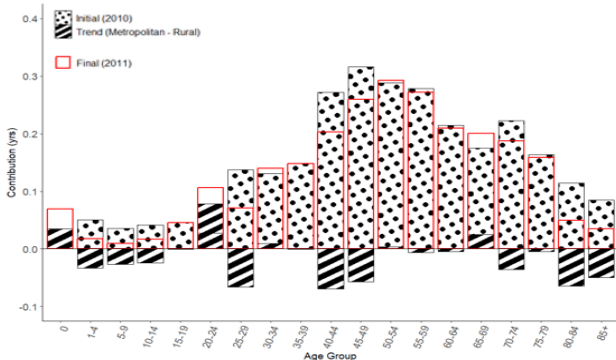
```

```

sum(contour_res[,1])
[1] 2.7819
sum(contour_res[,2])
[1] -0.29165
sum(contour_res[,3])
[1] 2.4903

```

시작 시점인 2010년 두 집단 간 기대수명 격차에 대한 연령군별 기여도를 모두 합한 값은 2.78세, 두 집단의 2010-2011년 기대수명 변화 추세 차이에 대한 기여도를 모두 합한 값은 -0.29세로 나타났다. 등치선 치환 분해 분석 결과를 그림으로 나타낸 결과는 <그림 2>와 같다. 시작 시점인 2010년과 종료 시점인 2011년 사이 대도시와 농촌 지역 간 남성의 기대수명 격차 감소에는 25-29세와 40-49세 연령군과 함께 80세 이상 연령군의 기여도가 비교적 컸다.



주: Initial은 시작 시점인 2010년, Final은 종료 시점인 2011년이다. 기여도가 양수이면 대도시, 음수이면 농촌 지역에 유리한 결과이다.

[그림 2] 1차원 등치선 치환 분해법을 이용한 2011년 우리나라 남성의 대도시와 농촌 지역 간 기대수명 격차에 대한 연령군별 기여도

2) 2차원 등치선 치환 분해법

연령군과 함께 사망원인에 따른 기여도를 함께 산출하고자 할 때는 2차원 등치선 치환 분해법을 이용할 수 있다(Jdanov et al., 2024). 2차원 등치선 치환 분해법의 기본적인 원리는 1차원 등치선 치환 분해법과 같다. 다만, 사망원인과 같은 요인이 추가되기 때문에 단계별 치환 횟수가 많아진다. 연령의 경우, 1차원 등치선 치환 분해법과 마찬가지로, 낮은 연령부터 높은 연령까지 순서대로 치환한다. 그러나 사망원인처럼 명확한 순서가 없는 경우에는 가능한 모든 순서에 따라서 치환 분석을 진행한 후 결과의 평균을 구한다. 2차원 등치선 치환 분해법 또한 1차원 분해법과 마찬가지로, 시작 시점과 추세 차이에 대한 기여도를 합하여 종료 시점의 기여도를 산출한다. 2차원 등치선 치환 분해법에 대한 자세한 이론적인 설명은 Jdanov et al. (2024)의 선행 연구에 나타나 있다.

2차원 등치선 분해법에 이용할 데이터는 1차원 등치선 분해법에 이용한 c1과 c2 데이터를 그대로 이용하였다. 2차원 등치선 분해법은 `cod_decomposition()` 함수를 이용하였다. `cod_decomposition()` 함수 또한 Jdanov(2024)가 선행 연구를 통하여 제시한 코드를 변형하여 이용하였다. `cod_decomposition()` 함수 또한 분석하고자 하는 데이터와 기대여명 산출 함수를 차례대로 입력한다.

```

two_dim_contour_res =cod_decomposition(c1, c2, FUN =ex.per.matrix)
Loading required package: gtools
two_dim_contour_res
$InitialEffect

```

	c1	c2	c3	c4	c5	c6
0	0.00665029	-0.00333523	-0.0008439	-0.00055587	0.022710	0.0102443
1	-0.00051166	0.00468048	-0.0066516	0.00285190	0.035540	0.0144106
5	-0.00298423	0.01194638	-0.0012790	-0.00085264	0.017067	0.0115201
10	0.00149441	-0.00088587	0.0023854	0.00536654	0.021184	0.0116369
15	0.00743088	-0.00468540	-0.0017225	-0.00024607	0.025303	0.0198320
20	0.00025681	0.00244983	-0.0010323	0.00109651	0.030949	-0.0059370
25	0.00171745	0.00613359	-0.0015421	0.00759965	0.106828	0.0164261
30	0.01010856	0.00607247	0.0058629	0.00744416	0.076009	0.0251485
35	0.01030287	0.02123716	0.0032713	0.01406609	0.071087	0.0289734
40	0.01804651	0.03436572	0.0089505	0.05410152	0.106209	0.0506296
45	0.03264776	0.04049024	0.0071424	0.06455568	0.121014	0.0501565

```

50 0.02697102 0.04992155 0.0116549 0.05150619 0.100927 0.0477799
55 0.02465514 0.07176937 0.0139820 0.03784359 0.089781 0.0401173
60 0.03874242 0.05803867 0.0169033 0.01183381 0.078536 0.0100981
65 -0.00384941 0.07202615 0.0098146 0.01825244 0.067833 0.0112145
70 0.00692710 0.10243254 0.0327102 0.01711286 0.067160 -0.0037134
75 0.00549278 0.04782280 0.0289514 0.00626305 0.069357 0.0055517
80 0.01188463 0.02923509 0.0269665 0.00887881 0.046616 -0.0093341
85 -0.00453784 -0.01308955 0.0343693 0.00783269 0.025367 0.0347577

```

\$TrendEffect

	c1	c2	c3	c4	c5	c6
0	-0.0093555	0.00469287	4.6447e-03	4.0802e-03	-0.00126842	0.03179866
1	0.0061856	-0.00227783	-1.8471e-04	1.0046e-03	-0.01406603	-0.02370879
5	-0.0020448	-0.00738042	3.8670e-03	3.9561e-04	-0.01631288	-0.00482355
10	0.0018450	-0.00077476	-2.5077e-04	-5.9862e-03	-0.01131143	-0.00782352
15	-0.0115821	0.00562407	3.2333e-03	-2.3362e-07	0.01999183	-0.01765180
20	0.0109032	-0.00807846	1.2876e-05	5.5434e-04	0.06333128	0.01147326
25	0.0017859	0.00250624	2.6680e-03	-2.9948e-03	-0.05720867	-0.01308322
30	0.0029089	0.01414091	-6.0471e-03	-2.7524e-03	0.00784157	-0.00696635
35	-0.0043307	0.00072968	3.1455e-03	-3.4478e-03	0.01753083	-0.01440586
40	-0.0012375	-0.00390037	8.0146e-04	-2.6697e-02	-0.03357271	-0.00484919
45	-0.0031023	-0.00290063	2.0754e-03	-2.2614e-02	-0.02220645	-0.00791000
50	0.0063667	-0.00160269	1.1529e-03	1.2020e-02	-0.01475587	0.00040939
55	0.0040728	0.00640784	3.4834e-03	-6.4856e-03	-0.02322723	0.00919553
60	-0.0106057	-0.01429984	-1.8345e-03	2.0660e-02	-0.00053143	0.00204319
65	0.0179589	-0.02661274	2.2424e-02	8.1570e-04	0.01188422	-0.00109893
70	-0.0033749	-0.03204386	7.1123e-03	-1.5994e-03	0.01475798	-0.02042054
75	0.0031442	0.01732596	1.5606e-02	-4.2817e-03	-0.02568828	-0.01039564
80	-0.0189989	-0.02638072	1.2760e-02	3.1383e-03	-0.01150932	-0.02357195
85	-0.0043845	-0.02289259	-1.1812e-02	-2.2065e-03	0.00659350	-0.01503187

\$CDecomp

```

[1] 0.0694621 0.0172722 0.0091185 0.0168795 0.0455267 0.1059792
0.0708362 0.1397710 0.1481597 0.2028468 0.2593483 0.2923513
0.2715947 0.2095837 0.2006616 0.1870613 0.1591493 0.0496839 0.0349654

```

sum(two_dim_contour_res[[1]])

[1] 2.7819

sum(two_dim_contour_res[[2]])

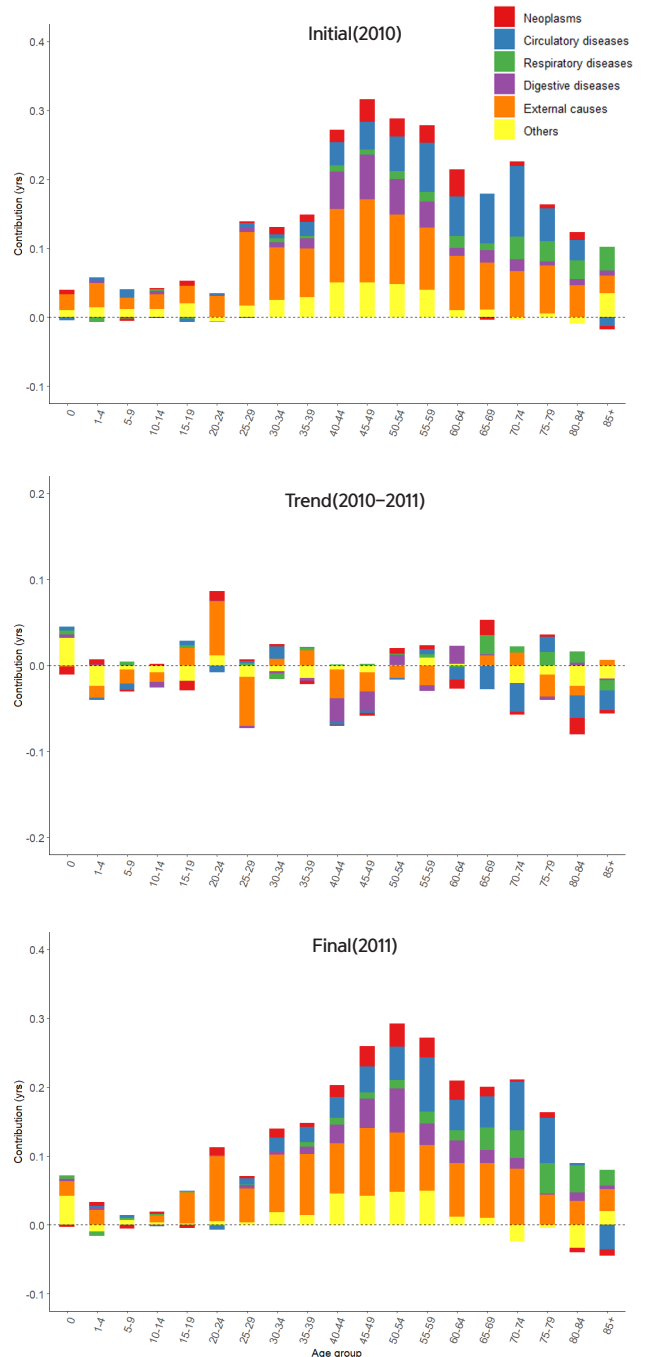
[1] -0.29165

sum(two_dim_contour_res[[3]])

[1] 2.4903

2차원 등치선 분해 분석 결과(two_dim_contour_res) 또한 시작 시점, 변화 추세, 그리고 종료 시점의 격차에 대한 기여도로 구성되어 있다. 각 요소에는 연령군과 사망원인별 기여도가 행렬의 형태로 저장되어 있는데 행의 수는 연령군의 개수와 같고, 열의 수는 사망원인 수와 같다. 1차원 등치선 분해법과 마찬가지로 시작 시점의 기여도와 추세 차이에 대한 기여도 결과의 타당성을 검증하기 위하여 종료 시점의 기여도가 함께 산출된다. 종료

시점 기여도는 연령군별 기여도만 산출된다. 연령군과 사망원인에 따른 기여도를 그림으로 나타낸 결과는 <그림 3>과 같다. 분석 결과, 2010년 우리나라 남성의 대도



주: Initial은 시작 시점인 2010년, Final은 종료 시점인 2011년, Trend는 두 집단의 기대수명 변화 추세 차이에 대한 기여도이다. 기여도가 양수이면 대도시 지역, 음수이면 농촌 지역에 유리한 결과이다.

[그림 3] 2차원 등치선 분해 분석법을 이용한 우리나라 남성의 2011년 대도시와 농촌 지역 간 기대수명 격차에 대한 연령군과 사망원인별 기여도

시와 농촌 지역 간 기대수명 격차에는 외인사, 심뇌혈관 질환, 소화기계질환, 암 등의 기여도가 컸다. 2010-2011년 기간 동안 대도시와 농촌 지역 남성의 기대수명 변화 추세 차이에 대한 기여도를 분석한 결과, 40-49세 연령군의 외인사와 소화기계질환, 그리고 65세 이상 연령군에서의 심뇌혈관질환이나 암 등이 두 집단 간 기대수명 격차 감소에 비교적 크게 기여하였다.

III. 고찰

이 연구는 생명표에 기반하여 비선형성을 가지는 기대여명이나 수명 변이의 분해법 가운데, 단계별 치환 알고리즘에 대하여 고찰하고 통계 소프트웨어 R을 이용하여 해당 분해법을 실제 적용하는 방법을 예시 자료와 함께 설명하였다. 또한 단계별 치환 알고리즘에 기반하여 두 집단의 시간에 따른 사망률 변화 양상의 차이를 분해 분석에 고려할 수 있는 등치선 치환 분해법의 적용 방법 또한 함께 제시하였다. 단계별 치환 알고리즘과 등치선 치환 분해법은 비선형성을 특징으로 하는 사망률 산출 함수에 대한 분해 분석 시 나타날 수 있는 비대칭성과 비이행성의 영향을 최소화한 결과를 산출할 수 있도록 한다.

단계별 치환 알고리즘과 등치선 치환 분해법은 요인의 치환 순서에 따라서 결과가 영향을 받을 수 있다. 연령군의 경우 낮은 연령군부터 높은 연령군 순으로 차례대로 치환하는 방법이 받아들여지고 있다(Jdanov et al., 2024). 그러나 사망원인이나 교육 수준처럼 명확한 순서가 없는 요인의 기여도를 파악하고자 할 때는 가능한 모든 순서에 따른 치환을 수행하고, 결과의 평균을 최종 결과로 삼는다(Jdanov et al., 2024). 하지만 요인이 일정한 개수 이상의 범주를 가질 때는 계산 부하량이 커져서 분석에 상당한 시간이 필요하다. 예를 들면, 사망원인이 여섯 가지

범주로 분류될 때는 연령군별로 720번의 치환을 수행한다. 따라서 연령군의 개수가 19개인 간이 생명표 기반 기대수명의 분해 분석을 위해서는 13,680회의 치환이 수행되어야 한다. 만약, 사망원인의 범주를 더 세분화하여 15개로 분류한다면 모든 가능한 치환 순서는 약 1.3조에 달한다. 이때 모든 경우의 분석을 수행하는 것은 물리적 계산 자원의 한계로 인하여 대체로 불가능하므로, 그중 10,000개의 순서를 무작위로 뽑아 분석을 수행하는 방안이 해결책으로 제시되었다(Jdanov et al., 2024).

등치선 치환 분해법은 시작 시점과 종료 시점을 언제로 설정하느냐에 따라서 결과에 큰 차이를 보일 수 있다. 등치선 치환 분해법은 시작 시점의 사망률 차이에 따른 기여도를 현재 시점의 사망률 구조로 보정하기 때문에, 시작과 종료 시점 간 시간차가 현저히 크면 실제 격차 크기와는 크게 차이가 나는 값이 산출될 수 있다. 또한 등치선 치환 분해법은 수행 원리에 따라서 시점 간 선형적인 변화 관계를 가정하는데, 만약 실제로는 비선형적 변화가 있는 경우 결과에 비뚤림을 불러올 수 있다. 선행 연구들에서는 이 문제에 대한 해결책으로 가능한 한 짧은 세부 기간으로 전체 기간을 나눈 뒤, 각 세부 기간별로 분해 분석을 수행하고 이를 합하여 전체 기간의 분해 분석 결과로 제시할 것을 권유하였다(Andreev et al., 2002; Bergeron-Boucher et al., 2015). 본 연구에서도 같은 자료를 활용한 선행 연구와 달리 등치선 분해 분석의 시작과 종료 시점 간 간격을 1년으로 설정하였다(Kim, 2023).

예시 분석 결과에 따르면 2011년 우리나라 남성의 대도시와 농촌 지역 간 기대수명 격차에는 외인사, 심뇌혈관질환, 소화기계질환 및 암 등의 기여도가 비교적 컸다. 2010-2011년 기간 동안 40-49세 연령군의 외인사 및 소화기계질환과 65세 이상 연령군의 심뇌혈관질환 및 암으로 인한 사망이 대도시와 농촌 지역의 기대수명 격차 감소에 비교적 크게 기여하였다. 이러한 결과는 국가 수준의 교통사고 대책과 자살 예방 전략 도입, 간염 백

신 접종, 흡연을 감소 및 의료 접근성 확대와 기술 발전 등에 따라서 대도시 지역과 농촌 지역 간 절대적인 사망률 격차가 감소하였기 때문에 여겨진다(Yang & Kim, 2003; Jang, 2009; Cha et al., 2015; Park et al., 2016; Kim, 2021). 그러나 비교적 긴 기간 동안 변화 양상을 분석한 앞선 연구와 달리 이 연구 결과는 짧은 기간의 변화를 분석한 결과이므로 해석에 주의가 필요하다(Kim, 2023).

비선형성을 특징으로 하는 생명표 기반 기대여명이나 수명 변이와 같은 사망률 지표를 분해 분석할 때는 일반화 분해 방법을 이용한 분해 분석이 필요하다. 단계별 치환 알고리즘과 등치선 치환 분해법의 적용은 집단 간 사망률 격차나 동일 집단 내 시간에 따른 사망률 변화 요인을 보다 완전히 이해하는 데 도움을 줄 수 있을 것이다. 또한 분해 분석 결과는 집단 간 사망률 격차에 인과적으로 영향을 미치는 원인에 대한 가설을 설정하고, 이러한 원인을 규명하거나 사망률 격차 감소를 위한 중재 방법을 모색하는 데 근거 자료로도 활용될 수 있을 것으로 기대된다.

연구윤리

이 연구는 고신대학교 복음병원 기관생명윤리위원회의 승인을 받았다(KUGH 2023-07-031). 본 연구는 2차 자료를 이용하였으므로 연구 대상자의 사전 동의는 면제되었다.

이해상충선언

본 논문에는 이해관계 충돌의 여지가 없음.

사사표기

이 연구는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(RS-2023-00212260).

이 연구는 2025년 8월 국립창원대학교에서 개최된 세미나 발표 내용을 바탕으로 발전시켰으며, 참석자들의 소중한 의견에 감사 드린다.

참고문헌

- Andreev, E. M. (1982). Metod komponent v analize prodoljitelnosti zizni [The method of components in the analysis of length of life]. *Vestnik Statistiki*, 9, 42-47.
- Andreev, E. M., Shkolnikov, V. M., & Begun, A. Z. (2002). Algorithm for decomposition of differences between aggregate demographic measures and its application to life expectancies, healthy life expectancies, parity-progression ratios and total fertility rates. *Demographic Research*, 7(article 14), 499-522. <https://doi.org/10.4054/DemRes.2002.7.14>
- Arriaga, E. E. (1984). Measuring and explaining the change in life expectancies. *Demography*, 21(1), 83-96.
- Bergeron-Boucher, M., Ebeling, M., & Canudas-Romo, V. (2015). Decomposing changes in life expectancy: Compression versus shifting mortality. *Demographic Research*, 33(14), 389-424.
- Bhopal, R. S. (2016). *Concepts of Epidemiology*. Third ed. Oxford University Press.
- Cha, E. S., Chang, S. S., Gunnell, D., Eddleston, M., Khang, Y. H., & Lee, W. J. (2015). Impact of paraquat regulation on suicide in South Korea. *Int J Epidemiol*, 45(2), 470-479.
- Das Gupta, P. (1991). Decomposition of the difference between two rates and its consistency when more than two populations are involved. *Mathematical Population Studies*, 3, 105-125.
- Das Gupta, P. (1994). Standardization and decomposition of rates from cross-classified data. *Genus*, 50(3/4), 171-196.
- Horiuchi, S., Wilmoth, J. R., & Pletcher, S. D. (2008). A decomposition method based on a model of continuous change. *Demography*, 45(4), 785-801. <https://doi.org/10.1353/dem.0.0033>
- Jang J. W. (2009). Current status of liver diseases in Korea: liver cirrhosis. *Korean J Hepatol*, 15(6), 40-49.
- Jdanov, D. A. (2024). Two-dimensional contour decomposition. <https://github.com/djdanov/cdecomp>.
- Jdanov, D. A., & Shkolnikov, V. M. (2014). An R-script for the assessment of the cross-sectional and the longitudinal components of a difference between two values of an aggregate demographic measure by contour replacement (MPIDR Technical Report TR-2014-003). Max Planck Institute for Demographic Research.
- Jdanov, D. A., Jasilionis, D., & Shkolnikov, V. M. (2024). Two-dimensional contour decomposition: Decomposing mortality differences into initial difference and trend components by age and cause of death. *Demographic Research*, 50(41), 1223-1246. <https://doi.org/10.4054/DemRes.2024.50.41>
- Jdanov, D. A., Shkolnikov, V. M., van Raalte, A. A., & Andreev, E. M. (2017). Decomposing Current Mortality Differences Into Initial Differences and Differences in Trends: The Contour Decomposition Method. *Demography*, 54(4), 1579-1602.

<https://doi.org/10.1007/s13524-017-0599-6>

- Kitagawa, E. M. (1955). Components of a difference between two rates. *Journal of the American Statistical Association* 50, 1168-1194.
- Kim, H. C. (2021). Epidemiology of cardiovascular disease and its risk factors in Korea. *Glob Health Med*, 3(3), 134-141.
- Kim, I. (2023). Contributions of the life expectancy gap reduction between urban and rural areas to the increase in overall life expectancy in South Korea from 2000 to 2019. *Int J Equity Health*, 22(1), 141. <https://doi.org/10.1186/s12939-023-01960-0>
- Olshansky, S. J., Carnes, B. A., & Cassel, C. (1990). In search of Methuselah: estimating the upper limits to human longevity. *Science*, 250(4981), 634-640.
- Park, C., Jee, Y. H., & Jung, K. J. (2016). Age-period-cohort analysis of the suicide rate in Korea. *J Affect Disor*, 194, 16-20.
- Pollard, J. (1982). The expectation of life and its relationship to mortality. *Journal of the Institute of Actuaries*, 109, 225-240.
- Pressat, R. (1985). Contribution des écarts de mortalité par âge à la différence des vies moyennes [Contribution of age-specific mortality differentials to mean lifetimes]. *Population (French ed.)*, 40, 766-770.
- Preston, S., Heuveline, P., & Guillot, M. (2000). *Demography: Measuring and Modelling Population Processes*. Wiley-Blackwell.
- R Core Team. (2025). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Silcocks, P. B., Jenner, D. A., & Reza, R. (2001). Life expectancy as a summary of mortality in a population: statistical considerations and suitability for use by health authorities. *Journal of epidemiology and community health*, 55(1), 38-43. <https://doi.org/10.1136/jech.55.1.38>
- van Raalte, A. A., Sasson, I., & Martikainen, P. (2018). The case for monitoring life-span inequality. *Science*. 362(6418), 1002-1004.
- Yang, B. M. & Kim, J. (2003). Road traffic accidents and policy interventions in Korea. *Inj Control Saf Promot*, 10(1-2), 89-94.

Decomposition of Mortality Measures Derived from Life Table-Based Functions

: Application of Stepwise Replacement and Contour Decomposition Using R

Ikhan Kim*

Abstract

Decomposition methods have been widely used to quantify the contributions of specific factors, such as age groups or causes of death, to mortality differences between populations or to temporal changes in mortality within the same population. However, when summary measures of mortality are derived from nonlinear functions, the application of traditional additive decomposition methods becomes problematic. Since the early 2000s, life table-based indicators of lifespan variation, in addition to age-standardized mortality rates and life expectancy, have been increasingly utilized, creating a demand for generalized decomposition methods applicable to a wide range of mortality indicators. Among the two major approaches to generalized decomposition, this study introduces the stepwise replacement algorithm and examines its practical application using illustrative data. Furthermore, for settings involving two populations observed at two time points, we present the contour replacement decomposition method, which extends the stepwise replacement algorithm to quantify contributions to both cross-sectional mortality differences between populations at the initial time point and differences in mortality change trajectories between populations over time.

Keywords: Cause of Death, Life Tables, Longevity, Algorithms

* College of Medicine, Kosin University (ikhan.kim@kosin.ac.kr)