

# 국민연금 가입자의 Markov Chain 보험료 기여 모형

## A Markov Chain Model of the Contribution Behavior of the National Pension Insured Persons

최기홍\*  
Choi Ki-Hong

국민연금으로부터 가입자들이 얻는 수익성과 가입자들 간의 형평성 분석에는 현금흐름을 생성하는 기여-수급 모형이 필요하다. 본 연구는 Hoem(1977), 미국 노동통계국(BLS)의 Markov chain 노동생명표 모형을 응용하여 국민연금 가입자의 보험료 기여모형을 개발하였다. 이러한 모형이 국민연금의 기여-수급 모형에서 필요한 것은 가입자의 사업장과 지역 간의 전이(transition)가 보험료 부과소득과 기여연수에 큰 영향을 미치기 때문이다. 사업장에서 지역으로 전이하면 보험료 부과소득이 하락하며 납부예외, 체납 등으로 기여연수가 감소하는 것이 보편적이다. 본 연구는 이론적 측면에서는 Markov chain 노동생명표 모형에서 체류시간(sojourn time)의 도출에 사용되는 전방향, 후방향 재귀식(recursive formula) 대신 경과 시간별 각 상태에서의 기대 체류시간의 일반식을 도출하여 분석을 간명하게 하였다. 국민연금에 대한 실증분석에서는 사업장, 지역, 그리고 대기자의 가입상태 간 전이확률에 의하여 측정된 가입 이후 경과시간별 사업장과 지역에서의 체류시간으로부터 실제 보험료를 납부한 기여연수를 측정하였다. 또한 사업장과 지역에서의 연령소득곡선(age-earning profile)을 체류시간으로 가중 평균하여 부과소득을 결정하였다. 이러한 Markov chain 모형은 가입자의 전이현상을 반영한 현실성 있는 기여-수급 모형의 구축을 가능하게 한다.

※ 국문 색인어: 국민연금, 기여모형, 기여연수, 마코브체인, 부과소득

\* 국민연금연구원, 연구위원(khchoi@nps4u.or.kr)

## I. 서론

현재 국민연금은 저부담-고급여의 정도가 심하여 보험료율은 높이고 소득대체율은 낮추는 개정이 추진되고 있다. 저부담-고급여 제도에서는 가입자가 국민연금으로부터 얻는 내부수익률이 적립기금의 투자수익률을 초과하여 연금재정은 장기적으로 지속가능하지 않다. 한편, 가입자가 납부하는 연금 보험료에 대한 연금 수급권이 보험료를 초과하므로 보험계리적 관점에서 부채가 발생한다<sup>1)</sup>. 이러한 내부수익률이나 수익비와 같은 수익성 지표의 산정이나 연금부채를 추계하기 위해서는 가입자의 보험료 납부와 연금 수급 현금흐름을 생성하는 기여-수급 모형이 필요하다.

확정급여(defined benefit) 방식의 국민연금에서 기여-수급 모형을 개발할 때 고려해야 하는 사항들은 모두 급여산식으로부터 도출될 수 있다. 국민연금의 급여산식은 기본연금액(BPA: Basic Pension Amount)이라고 하며 다음과 같다<sup>2)</sup>.

$$BPA = 1.8(A+B) [1 + 0.05(n-20)] = 0.09(A+B)n \quad (1)$$

식에서 A는 가입자들의 전체 신고소득(월액)의 평균으로 "A값"이라고 한다<sup>3)</sup>. A값은 산업 전체 임금상승률에 대한 전망으로부터 외생적으로 결정할 수 있다. B는 가입자가 신고한 보험료 부과대상 소득을 A값의 증가율로 현재 가치로 재평가하여 평균한 값이다. 가입자의 보험료 부과소득이  $y_i$ ,  $i \in R$ 이며 R은 국민연금 가입자의 연령범위 (18,60) 이다. 같은 기간 A값이 A라면 시점 t에서 B값은 부과소득  $y_t$ 를 가입자 평균소득 상승률  $A_t/A$ 로 재평가하여 다음과 같이 결정된다.

- 1) 공적연금의 부채는 후세대에 대한 사회적 채무이므로 세계은행의 Holzmann et al.(2001)은 암묵적연금부채(IPD: Implicit Pension Debt)라고 하여 통상적 연금부채와 구별한다.
- 2) 국민연금은 최소 10년은 기여하여야 한다. 10년 이상 20년 미만의 경우는 감액연금 산식이 적용된다. 감액연금과의 차이는 기여연수를 1/2년 감하는 것으로 미소한 편이다.
- 3) 익년 3월경 공시되며 실제로는 지나친 등락을 피하기 위하여 최근 3년 이동평균에 의하여 결정된다.

$$B = \frac{\sum_{i \in R} y_i \frac{A_i}{A_i}}{n} \quad (2)$$

기여연수  $n$ 은 가입기간 중 보험료 부과대상 소득의 이력  $y_i$ 에 의하여 다음과 같이 결정된다. 식에서  $I(y_i)$ 는 지시(indicator) 함수로서 소득  $y_i$ 가 양의 값에서 1, 소득이 없으면 0의 값을 갖는다.

$$n = \sum_{i \in R} I(y_i) \quad (3)$$

그러므로 국민연금 가입자의 기여-수급 모형의 핵심은 가입기간중의 연령별 보험료 부과소득  $y_i$ , 즉 연령소득곡선(age-earning profile)에 대한 가정이다. 문형표(1995), 석재은·김용하(2002), 전병목(2002) 등 기존 연구들은 기여연수  $n$ 을 먼저 가정하고 가입자의 연령범위 [18,60) 내에서 임의의 길이가  $n$ 인 연속구간에 연령소득곡선을 가정하고 있다. 이러한 방법의 문제점은 문형표(1995), 김원식(1996) 등에서 볼 수 있듯이 동일한 기여연수에서도 보험료 현금흐름의 구체적 기여기간 가정에 따라 내부수익률에 적지 않은 차이가 발생할 수 있는 것이다. 최기홍(2005)은 국민연금에 대한 연금 보험료 납부를 일시납 거치연금(single premium deferred annuity)의 구매로 해석하면 보험료 납부로 구매하는 가상적 개별 일시납 거치연금의 내부수익률은 저부담-고급여 상태에서는 연령에 따라 상승하는 특성이 있음을 보였다.

본 연구는 기존 확정적(deterministic) 기여 모형들의 이러한 단점을 개선하고 가입자의 사업장과 지역 간의 이동 현상을 반영할 수 있는 정교한 확률적(probabilistic) 기여 모형을 개발하는 것을 목적으로 하고 있다. 가입자가 사업장에서 지역으로 전이(또는 이동)하면 보험료 부과소득이 하락하며 납부예외, 체납 등으로 기여연수가 감소하는 것이 보편적이어서 현금흐름의 모양에 큰 영향을 미친다. 이러한 목적에는 노동생명표(work-life table)에서 사용되는 Markov chain 모

형이 적합하다.

다음 제Ⅱ장에서는 국민연금 제도 현황을 통하여 보험료 기여 모형에서 가입자의 가입기간 가입상태 전이 현상을 반영하는 것이 필요하다는 것을 보이고 있다. 다음 제Ⅲ장에서는 인구학 및 노동경제학에서 노동기대여명의 추계에 사용되는 Markov chain 모형의 이론적 측면과 실증분석을 위한 기대 체류시간의 개념과 추계를 위한 일반식을 도출한다. 다음 제Ⅳ장에서는 국민연금 가입자의 가입기간 동안의 기여연수와 보험료 부과소득을 Markov chain 모형으로 추계하였다. 마지막 제Ⅴ장에서는 향후 연구과제와 함께 결론을 맺는다.

## Ⅱ. 국민연금 제도와 가입자의 가입상태 전이

국민연금의 가입자는 1988년 443만 명에서 전 국민 연금화가 달성된 1999년에는 1,626만 명으로 급증하였으며, 2005년 말에는 1,712만 명으로 확대되었다. 국민연금의 가입자는 사업장 가입자와 지역가입자로 구성되어있다. 이중 지역가입자는 사업장 가입자를 제외한 나머지 경제활동 인구로서 자영업자, 비정규직, 실업자 등 이질성이 높은 집단이다. 지역가입자는 1995년 7월 농어촌지역 약 2백만 명을 시작으로 도시지역 나머지 경제활동 인구들로 가입자 범위가 확대된 1999년 4월 급증하여 1천 1백만 명에 달하였으나 2003년 이후 사업장 범위 확대로 지역가입자의 비중이 점차 줄어들고 있다.

〈표 1〉 연도별 가입자 현황

(단위: 명)

구분 연도	총가입자	사업장가입자	지역가입자		
			계	도시	농어촌
1988	4,432,695	4,431,039	-	-	-
1995	7,496,623	5,541,966	1,890,187	-	1,890,187
1999	16,261,889	5,238,149	10,822,302	8,739,152	2,083,150
2003	17,181,778	6,958,794	9,964,234	7,902,223	2,062,011
2005	17,124,449	7,950,493	9,123,675	7,154,658	1,969,017

주: 매년 12월 기준.

국민연금 가입자는 경제활동 기간 중에 사업장과 지역 가입자 간에 이동이 보편적이다. 우리나라에서 흔한 것처럼 사업장에서 퇴직하여 자영업을 시작하는 경우는 사업장에서 지역으로 가입상태가 전이되는 것이다. 이러한 사업장과 지역 간의 가입상태 이동은 국민연금 가입자의 부과소득과 기여연수에 크게 영향을 미친다.

첫째, 지역가입자의 높은 납부예외율의 문제점이다. 2005년 『국민연금 통계연보』에 의하면 910만 명의 지역가입자중 납부예외자가 460만 명에 달하며 이중 실업이 330만 명으로 대다수를 차지한다. 납부예외자가 아니면서 고의로 보험료를 납부하지 않는 경우가 많아 징수율도 낮은 편이다. 가입자중 중 실제 보험료를 납부하는 가입자의 비율을 나타내는 납부율 $[=(1-납부예외율) \times 징수율]$ 은 다음 표와 같다.

〈표 2〉 남성 지역가입자의 보험료 납부율

(단위: %)

연령	납부예외율	징수율	납부율	연령	납부예외율	징수율	납부율
18	0.932	0.518	0.035	39	0.471	0.663	0.351
19	0.966	0.286	0.010	40	0.453	0.675	0.369
20	0.965	0.190	0.007	41	0.440	0.691	0.387
21	0.961	0.174	0.007	42	0.431	0.698	0.397
22	0.943	0.162	0.009	43	0.409	0.710	0.420
23	0.926	0.187	0.014	44	0.393	0.718	0.436
24	0.902	0.229	0.022	45	0.385	0.728	0.448
25	0.861	0.255	0.035	46	0.373	0.739	0.463
26	0.821	0.287	0.051	47	0.366	0.749	0.475
27	0.855	0.310	0.045	48	0.360	0.758	0.486
28	0.847	0.347	0.053	49	0.351	0.769	0.499
29	0.807	0.409	0.079	50	0.340	0.783	0.517
30	0.704	0.492	0.145	51	0.329	0.794	0.533
31	0.629	0.535	0.199	52	0.324	0.809	0.547
32	0.651	0.541	0.189	53	0.309	0.821	0.567
33	0.619	0.571	0.218	54	0.297	0.835	0.587
34	0.589	0.594	0.244	55	0.282	0.848	0.609
35	0.566	0.611	0.265	56	0.266	0.859	0.630
36	0.530	0.623	0.293	57	0.245	0.872	0.659
37	0.510	0.638	0.312	58	0.243	0.884	0.669
38	0.490	0.653	0.333	59	0.237	0.897	0.684

주: 납부율 = (1 - 납부예외율) × 징수율.

자료: 국민연금관리공단 내부자료, 2005년 12월.

20대 후반까지는 병역, 재학 등으로 납부예외율이 80%를 상회한다. 또한 이 연령대에 지역가입자의 징수율도 20대 후반에는 40%까지 증가하지만 전반적으로 매우 낮아서 전체적인 납부율은 50% 부근으로 볼 수 있다. 남성 지역 가입자의 납부율은 50대부터 비로서 50%를 넘는다. 이후 연령의 증가에 따라 납부예외율과 징수율은 점차로 개선되어 연금수급 직전 59세의 납부율은 68%정도에 도달한다. 이러한 낮은 납부율은 기여연수의 부족을 초래하며 노후 연금을 받을 수 없거나 받더라도 무의미한 수준이어서 연금혜택으로부터의 “사각지대” 문제가 제기되고 있다.

둘째, 지역 가입자와 사업장 가입자의 소득에는 큰 격차가 존재한다. 그 이유는 지역가입자는 전체적으로 소득이 적은 계층이 대다수이며 사업장 가입자는 고용주가 보험료의 1/2을 부담해주지만 지역가입자는 전액을 부담하므로 체감하는 보험료 부담이 큰 편이다. 또한 지역가입자의 부과소득은 원칙적으로 본인의 신고에 의하여 결정되므로 제도에 대해 불신하는 경우 소득을 낮게 신고하게 한다.

연도별 추세를 살펴보면, 사업장 대비 지역의 평균소득월액 비율은 전 국민 연금화가 시작된 1999년 4월 이후 60% 이하에서 조금씩이지만 악화되는 추세를 보이고 있다. 가입종별 평균소득월액을 살펴보면, 2005년 12월 기준 전체 평균소득월액은 158.6만원으로 나타났지만, 사업장가입자의 평균소득월액은 188만원, 지역가입자의 경우는 107만원으로 사업장가입자에 비해 지역가입자의 평균소득월액은 57%수준에 불과하다.

이상과 같이 국민연금 제도에는 사업장과 지역의 이질적인 두 종류 가입상태가 존재한다. 그리고 가입자는 전체 근로기간 중 연령에 따른 종사상 지위 변화로 사업장에서 지역 또는 지역에서 사업장으로 이동이 보편적이다. 그런데 지역 가입자의 보험료 부과소득은 낮은 편이며 가입 기간과 실제로 보험료를 납부한 기간이 기여연수 간에 차이가 커서 가입 상태의 변화 패턴을 반영하는 것이 필요하다. 기존의 확정적 보험료 기여 모형에서는 이러한 가입상태의 변화를 반영할 수 없으므로 확률적 보험료 기여 모형이 보다 바람직한 것이다.

〈표 3〉 연도별·가입종별 평균소득월액 현황

(단위: 원)

연도	구분 총가입자	사업장가입자	지역가입자		
			계	도시	농어촌
1995	943,468	1,052,602	577,126	-	577,126
1998	1,095,773	1,439,751	785,159	841,170	628,839
1999	1,290,706	1,607,129	960,552	1,018,510	773,458
2000	1,362,754	1,685,479	1,010,098	1,059,371	844,164
2002	1,525,342	1,816,554	1,058,575	1,095,269	947,853
2004	1,583,176	1,856,353	1,065,592	1,099,420	966,382
2005	1,586,424	1,875,458	1,074,539	1,109,098	972,024

주: 매년 12월 기준.

### Ⅲ. Markov chain 노동생명표

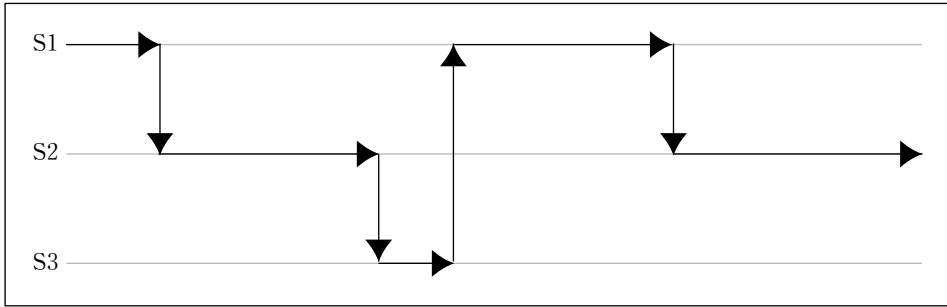
Markov chain은 시간의 경과에 따라 발생하는 상태 간의 전이확률(transition probability)이 전이가 발생하기 직전 상태에만 의존하는 확률과정(stochastic process)이다. Markov chain은 단순하지만 현실에 대한 설명력이 높아서 다양한 분야에 활용되는 대표적 확률과정 모형이다.

인구학자 Hoem(1977)은 근로자의 현재 연령으로 부터 남아있는 근로기간을 의미하는 노동기대여명(work-life expectancy)의 산정에 Markov chain을 이용하는 방법론을 도입하였다. Markov chain 노동생명표에서 근로자의 남아있는 생애는 노동(active)과 비노동(inactive) 두 상태를 전이하는 Markov 과정이다<sup>4)</sup>.

4) 원래는 사망이라는 제3의 상태가 있다. 그러나 Hoem은 사망 상태를 고려하지 않고 이론을 전개하고 나중에 사망확률로 보정하는 접근법을 취하고 있다.



〈그림 1〉 상태 간 이동과 체류시간



노동생명표(work-life table) 모형에서 노동(active) 상태에 머무는 시간이 노동기대여명이다. 미국 노동통계국 BLS(Bureau of Labor Statistics)는 Smith(1982, 1986)에 의하여 기존 Wolfbein and Wool(1950)의 방법 대신 남녀 모두에게 적용이 가능한 Hoem(1977), Schoen and Woodrow(1980)의 Markov chain 노동생명표 방법론을 채택하였다.

Markov 과정에서 상태의 수는 모두  $m$ 종류이며  $i, j=1, 2, \dots, m$ 로 나타낸다.  $x$ 세에  $i$ 상태에 있는 개인이  $t$ 시간 경과 후 상태  $j$ 에 있을 확률을  $t$ -단계 전이확률이라고 하며  $p_x^{ij}$ 로 나타낸다.

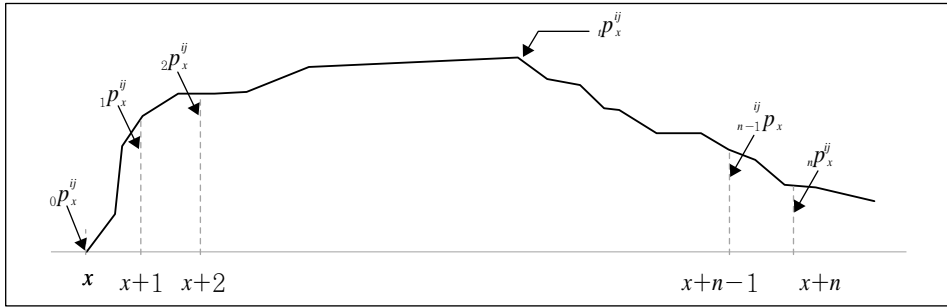
$x$ 세에 상태  $i$ 에서 출발하여  $x+n$ 세에 이르기 까지 상태  $j$ 에 머무는 시간으로 정의되는  $e_{x:n}^{ij}$ 는 다음과 같이  $t$ -단계 전이확률의 적분으로 나타낼 수 있으며 체류시간(sojourn time)이라고 한다<sup>5)</sup>.

$$e_{x:n}^{ij} = \int_0^n p_x^{ij} dt \quad (4)$$

체류시간은  $t$ -단계 전이확률의 적분으로 정의되었으므로 다음 그림과 같이 선분  $[x, x+n]$ 과  $p_x^{ij}$  사이의 면적을 의미한다.

5) 본 원고에서 수식의 기호체계는 Hoem(1977)을 대부분 따르고 있다.

〈그림 2〉 체류시간  $e_{x:n}^{ij}$ , ( $i \neq j$ )의 수치적분



Hoem(1970)에서 기술되고 있듯이 Markov chain에서 전이확률  $p_x^{ij}$ 는 연령  $x$ 에 대해 Kolmogorov 상미분방정식으로 정의되므로 시간  $t$ 는 구간  $(0,1)$ 에서만 의미를 갖는다<sup>6)</sup>. 그러므로 Hoem(1977)은  $t$ -단계 전이확률  $p_x^{ij}$ 에서  $t=1$ 로 하여 연령별 1-단계 전이확률  $p_x^{ij}$ ,  $p_{x+1}^{ij}$ ,  $p_{x+2}^{ij}$ , ...,  $p_{x+n-1}^{ij}$ 를 해석적으로 구하고  $x$ 세로부터  $t$ -단계 전이확률  $p_x^{ij}$ 는 연령별 1-단계 전이확률과 다음과 같은 Chapman-Kolmogorov 관계식에 의하여 구하였다.

$$p_x^{ij} = \sum_{s=1}^m {}_{t-1}p_x^{is} p_{x+t-1}^{sj}, \quad t=1, 2, 3, \dots, n \quad (5)$$

식 (5)와 같이  $t$ -단계 전이확률이 주어지면 체류시간은 다음 식 (6)과 같이 사다리꼴 공식으로 수치적 근사식으로 구할 수 있다. 다음 식에서 〈그림 2〉의  $n$ 개의 개별 사다리꼴의 면적들을  $d_{x+k-1|x}^{ij}$ 로 나타내었다.

$$e_{x:n}^{ij} = \sum_{k=1}^n d_{x+k-1|x}^{ij}, \quad d_{x+k-1|x}^{ij} \equiv \frac{{}_{k-1}p_x^{ij} + {}_k p_x^{ij}}{2} \quad (6)$$

6) Kolmogorov forward differential equation with constant coefficients in each age group.

식 (6)은 전체 체류시간  $e_{x:n}^j$ 이 개별 단위기간의 체류시간  $d_{x+k-1,x}^j$ 의 합으로 결정되는 구조를 갖고 있다. 식에서 기호 ' $x$ '는  $x$ 세에서 기대 값이다. 본 연구는 모든 분석에 기존 문헌들에 비해 간명한 식 (6)을 사용하였다. 다음은 식 (6)과 기존 문헌, 특히 미국 노동통계국 BLS의 방법론과의 관계를 보고자 한다.

Hoem(1977)과 BLS의 노동생명표 문헌에는 식 (6)과 같은 체류시간의 일반해는 제시되지 않는다. Hoem(1977, p.13)은  $n$ 개의 1기간 체류시간  $e_{x:1}^j, e_{x+1:1}^j, \dots, e_{x+n-1:1}^j$ 들과 식 (7)과 같은 후방 재귀식(backward recursion formula)으로  $x$ 세에서  $x+n$ 세까지 체류시간  $e_{x:n}^j$ 을 추계하였다.

$$e_{x:n}^j = e_{x:1}^j + \sum_{k=1}^m p_x^{ik} e_{x+1:n-1}^{kj} \quad (7)$$

노동통계국 BLS는 다음 식 (8)과 같은 전방 재귀식(forward recursion formular)을 사용하고 있다. BLS의 전방 재귀식은 Hoem의 기호 체계로 나타내면 다음 식 (8)과 같다. 식 (7)과 식 (8)의 재귀식을 비교하면 식 (7)에는 1-단계 전이 확률이 사용되지만 식 (8)에는 다단계 전이확률이 사용된다. 그러므로 재귀식 (8)에는 병행하여 다단계 전이확률을 위한 Chapman-Kolmogorov 재귀식 (5)가 사용된다.

$$e_{x:n}^j = e_{x:n-1}^j + \sum_{k=1}^m p_x^{ik} e_{x+n-1:1}^{kj} \quad (8)$$

Hoem의 식 (7)과 BLS의 식 (8)에는 공통적으로 단위기간 중 체류시간  $e_{x:1}^j$ 이 사용되고 있지만 차이가 있다. Hoem(1977)은 식 (4)에 의하여  $e_{x:1}^j$ 을 해석적으로 구하였지만 Wessels(2006, p.6), Millimet et al(2003, p.88)에서는 <표 4>와 같은  $e_{x:1}^j$ 의 근사식을 사용하고 있다. <표 4>의 근사식은 상태 간 전환이 있는 경우

즉,  $i \neq j$ 에는 각 상태에 0.5씩 머무는 것으로 가정하는 것이다<sup>7)</sup>.

〈표 4〉 BLS의 단위기간 체류시간  $e_{x:\overline{1}}^{ij}$ 의 근사식( $m=2$ )

$i \setminus j$	active(a)	inactive(i)
active(a)	$1 \times p_x^{aa} + 0.5 \times p_x^{ai}$	$0.5 \times p_x^{ai}$
	<i>ia</i>	<i>ii</i> <i>ia</i>

그런데 〈표 4〉는  $i = j$ 이면  ${}_0p_x^{ij} = 1$ ,  $i \neq j$ 이면  ${}_0p_x^{ij} = 0$ 이라는 사실을 이용하면 다음과 같이 요약된 형태로 나타낼 수 있다. 식 (9)는 앞서 제시된 수치적분에 의한 체류시간의 일반식 (6)에서  $n = 1$ , 즉  $d_{x|1}^{ij}$ 에 해당한다.

$$e_{x:\overline{1}}^{ij} = \frac{{}_0p_x^{ij} + {}_1p_x^{ij}}{2} \tag{9}$$

식 (6)은 BLS 전망 재귀식의 일반적 해임을 다음과 같이 보일 수 있다. 여기서 3번째 식은 Chapman-Kolmogorov 관계식과 식 (9)의 정의에 의해서 유도된다.

$$\begin{aligned}
 e_{x:\overline{n}}^{ij} - e_{x:\overline{n-1}}^{ij} &= \sum_{s=1}^m {}_{n-1}p_x^{is} e_{x+n-1:\overline{1}}^{sj} \tag{10} \\
 &= 0.5 \times \sum_{k=1}^m {}_{n-1}p_x^{is} ({}_0p_{x+n-1}^{sj} + {}_1p_{x+n-1}^{sj}) \\
 &\qquad \qquad \qquad ij \qquad \qquad ij \qquad \qquad ij
 \end{aligned}$$

7) “the standard ID(Increment/Decrement) assumption that a person who is observed to be active at the beginning of the year and inactive at the end of the year is active half the year”, Wessels(2006, p.6).

$$= 0.5 \times ({}_{n-1}p_x + {}_n p_x) \equiv d_{x+n-1|x}$$

다음 식 (11)은 역시 앞서 제시된 식 (6)이 Hoem의 후방 재귀식 (8)의 해임을 개략적으로 보이고 있다. 식 (7)에 식 (6)과 식 (9)를 대입하면 다음과 같이 항등식으로 되는 것을 확인할 수 있다.

$$\begin{aligned} \sum_{k=1}^n \frac{{}_{k-1}p_x + {}_k p_x}{2} &= \frac{{}_0 p_x + {}_1 p_x}{2} + \sum_{s=1}^m p_x \sum_{k=1}^{n-1} \frac{{}_{k-1}p_{x+1} + {}_k p_{x+1}}{2} \quad (11) \\ &= \frac{{}_0 p_x + {}_1 p_x}{2} + \sum_{k=1}^{n-1} \sum_{s=1}^m p_x \frac{{}_{k-1}p_{x+1} + {}_k p_{x+1}}{2} \\ &= \frac{{}_0 p_x + {}_1 p_x}{2} + \sum_{k=1}^{n-1} \frac{{}_k p_x + {}_{k+1} p_x}{2} \end{aligned}$$

결론적으로 체류시간의 일반식 (6)은 미 BLS의 전방 재귀식을 정확히 만족시킨다. 또한 단위기간의 체류시간  $e_{x:1}$ 의 근사식을 식 (9)와 같이 하면 Hoem의 후방 재귀식을 만족시킨다. 이와 같은 사실들은 식 (6)과 같은 간명한 산식이 기존 Markov chain 노동생명표 모형과 일관성을 갖음을 보인다. 또한 이론적으로 식 (6)의 형태는 Markov chain 노동생명표 모형이 원래 생명표의 확장임이 자연스럽게 드러내는 장점이 있다.

#### IV. 국민연금에 대한 적용: 기여연수와 부과소득

국민연금제도에서 가입자는 상용 임금근로자인 경우 가입상태는 사업장으로 분류되고 자영자, 임시직 근로자, 실업자 등은 지역으로 분류된다. 그러므로 국민연금 가입자는 가입기간 중 종사상 지위 등 근로 상태의 변동에 따라 “사업장”과 “지역”

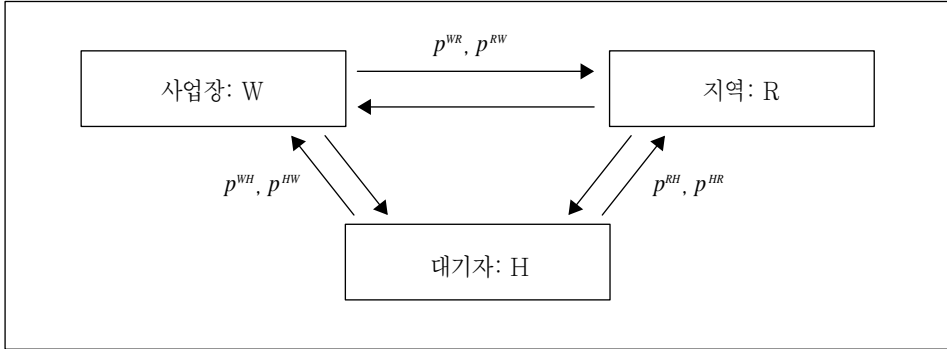
을 전이하는 것이 보편적이다. 또한 가입자는 국민연금에서 탈퇴하여 “대기자”로 이동할 수 있다<sup>8)</sup>. 국민연금에서 대기자는 현재는 가입자의 자격을 상실하였지만 국민연금에 보험료 납부 실적은 남아있는 자를 의미한다<sup>9)</sup>.

이러한 가입자의 가입상태 전이 현상이 국민연금 보험료 기여 모형에서 중요한 것은 첫째, 지역가입자의 높은 납부예외율과 낮은 징수율 때문에 가입자 자격을 유지한 가입연수와 실제 보험료를 납부한 기여연수 간에는 큰 괴리가 발생한다. 둘째, 실제 저소득 계층이 많고 보험료 전액을 부담하므로 사업주가 1/2을 보조하는 사업장 가입자에 비해 보험료에 대한 체감하는 부담이 커서 소득을 실제보다 낮게 신고하는 경향 때문에 지역 가입자와 사업장 가입자 간에는 현저한 소득 격차가 존재하는 현실의 반영이 필요하기 때문이다.

가입자 또는 대기자의 사업장(W), 지역(R), 대기자(H) 세 가지 상태 간 이동을 <그림 3>과 같은 Markov chain으로 가정하였으며  $x$ 세  $i$ 상태에서  $x+1$ 세에  $j$ 상태로 전이확률은  $p_x$ ,  $i, j = W, R, H$ 로 나타낸다. 전이확률은 국민연금 재정추계 모형을 위하여 추정된 박성민(2006)의 추정치를 사용하였으며 <부록 1>에 수록하였다. 국민연금 재정추계 모형은 전이확률에 의하여 가입자를 기여연수별로 분류하여 가입자로부터 신규 연금수급자를 발생시키는 시뮬레이션 접근법을 근간으로 한다<sup>10)</sup>.

- 8) 배우자가 국민연금 가입자인 경우 국민연금의 당연 가입대상에서 제외되므로 취업 경험 이 있는 대부분의 전업주부들이 대기자에 해당한다. 이 경우 남성들도 지역 가입 대신 대기자를 선택하는 경우가 많다.
- 9) 대기자는 60세에 도달하면 10년 이상 기여실적이 있는 경우 연금을 받고 그렇지 못한 경우에는 납부한 보험료를 정기 금리에 의한 원리합계를 일시금으로 받는다.
- 10) 엄밀하게는 가정의 타당성을 검증하기 위하여 가입자 데이터에 대한 통계적 적합성 검증의 필요성이 있을 것이다.

〈그림 3〉 국민연금 가입 상태 간 이동



Markov chain 모형이 처음 도입된 노동생명표에서는 상태가 노동(active)과 비노동(inactive)의 두 종류이지만 국민연금 기여 모형에서는 사업장, 지역, 대기자의 세 종류이다. 그렇지만 단위기간 체류시간의 근사식에 대한 정의는 다음 〈표 5〉와 같이 쉽게 확장될 수 있다. 각 항은  $i$  상태에서 시작하여 단위기간 안에  $j$  상태에 머무는 체류시간을 나타낸다.

〈표 5〉 단위기간 체류시간  $e_{x:1}^{ij}$ 의 근사식

$i \setminus j$	W	R	H
W	$p_x^{WW} + 0.5 \times (p_x^{WR} + p_x^{WH})$	$0.5 \times p_x^{WR}$	$0.5 \times p_x^{WH}$
R	$0.5 \times p_x^{RW}$	$p_x^{RR} + 0.5 \times (p_x^{RW} + p_x^{RH})$	$0.5 \times p_x^{RH}$
H	$0.5 \times p_x^{HW}$	$0.5 \times p_x^{HR}$	$p_x^{HH} + 0.5 \times (p_x^{HW} + p_x^{HR})$

앞서 밝혀진 바와 같이 위와 같은 단위기간 체류시간의 근사식을 사용하는 경우 전방, 후방 재귀식에 의한  $n$ 시간 중 체류시간  $e_{x:n}^{ij}$ 은 모두 동일하며 식 (6)과 같다. 이 때  $x$ 세에  $i$ 상태에 있는 개인이  $x+n$ 세, 즉  $(x+n, x+n+1)$  기간에 기대되는  $j$ 상태에서의 체류시간  $d_{x+n|x}^{ij}$ 은 다음과 같다.

$$d_{x+n|x}^{ij} = \frac{{}_nD_x^{ij} + {}_{n+1}P_x^{ij}}{2}, \quad i=W, R, \quad j=W, R, H \quad (12)$$

국민연금은 강제가입이 원칙이지만 지역가입자는 실업, 질병 등으로 소득이 없는 경우 납부예외를 인정받을 수 있다. 또한 납부예외 신청 등 정당한 사유없이 보험료를 체납하는 경우가 많다. 이러한 경우 국민연금 가입자의 자격을 유지하는 가입기간과 실제 급여산식에 적용되는 기여연수 간에 큰 차이가 발생한다. 이러한 차이점은 지역에서의 체류시간  $d_{x+n|x}^{ik}$ 에 보험료 납부율  $\epsilon_{x+n}$ 을 곱하여 쉽게 반영될 수 있다.

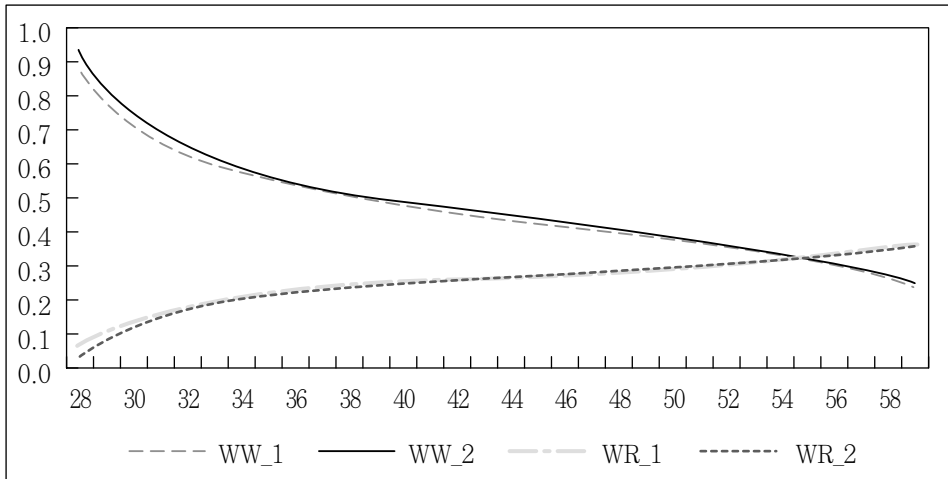
다음 <그림 4>는 사업장에서 28세에 시작한 가입자의 59세까지 경과 시점별 단위 기간의 체류시간을 나타낸다. 계열 1은 전이확률  ${}_nD_x^{ij}$ 만으로 산정한 것이며 계열 2는 전이확률의 사다리꼴 공식  $0.5 \times ({}_nD_x^{ij} + {}_{n+1}P_x^{ij})$ 으로 산정한 체류시간이다. 두 가지 방식의 실제 차이는 작은 것으로 나타나고 있다. 본 연구에서 앞으로 모든 결과는 보다 정교한 사다리꼴 공식에 의한 “계열 2”를 기본으로 사용하기로 한다.

그림에 의하면 최초 사업장으로 가입한 가입자와 지역으로 가입한 가입자 모두 사업장에서의 체류시간은 점감하며 지역에서의 체류시간은 증가하는 것으로 나타난다. 그것은 연령이 증가함에 따라 사업장에서 지역으로 전이하는 확률이 커지는 것이 반영된 것이다. 이러한 현상은 우리나라 근로자들이 40대 후반부터 임금근로자에서 자영업으로 근로형태를 바꾸는 것과 관련된다.

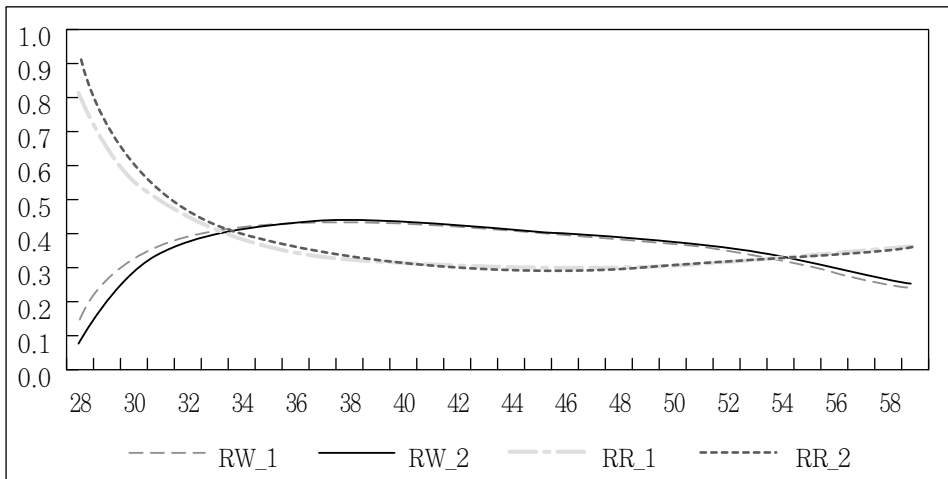


〈그림 4〉 연령별 상태별 체류시간:  $d_{x+t|x}^j$

- 사업장으로 가입한 가입자



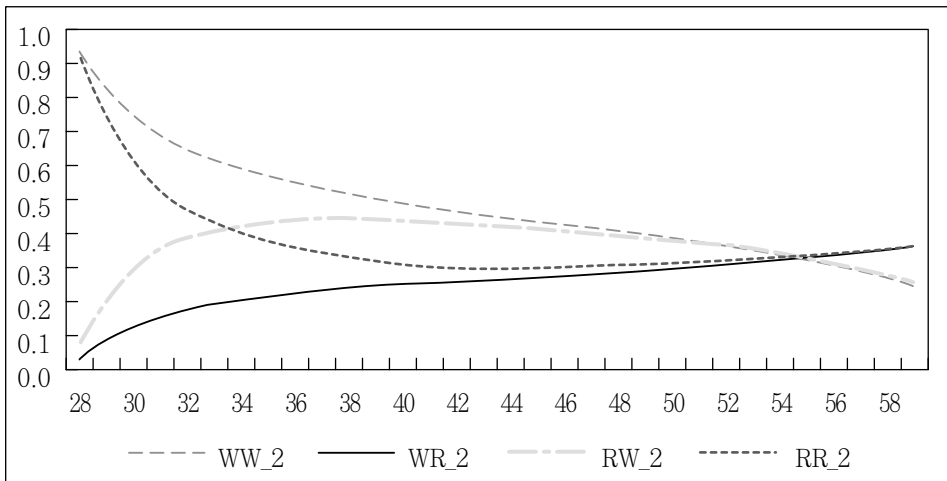
- 지역으로 가입한 가입자



주: W\_k, R\_k, k=1, 2는 각각 계산 방법의 차이를 나타내며, 2는 사다리꼴 공식임.

다음 <그림 5>는 앞서 최초 가입 상태별 두 그림을 합성한 것이다. 그림에서 중요한 특징은 가입 이후 12년이 경과한 40세 부터는 상태별 체류시간이 최초 28세 가입상태와 무관해 지는 것이다. 국민연금의 전이확률이 연령에 따라 변하는 “in-homogeneous” 한 경우이므로 수평선은 아니다. 55세 무렵부터는 지역에서 체류하는 시간이 사업장에서의 체류시간을 상회하며 이는 통념과 일치하는 결과이다.

<그림 5> 경과시점별 단위기간 중 상태별 체류시간:  $d_t^j$



다음 <표 6>은 <부록 1>의 국민연금 재정추계 모형에서 사용되는 전이확률과 앞서 식 (12)의 산식에 의하여 계산된 기대 체류시간  $d_{x+n|x}^j$ 이다. 표에서 지역\_1은 단위기간 지역에서의 체류시간이며 지역\_2는 ‘지역\_1×납부율’로 계산된 실제 보험료를 납부한 기간이다. 다음 <표 6>은 앞으로 가입자의 경과 시점별 누적 기여연수와 그에 대응되는 가중평균 보험료 부과소득을 결정하는 기초자료가 된다.

〈표 6〉 경과 시점별 상태별 체류시간:  $d_{x+t|x}^j$

(단위: 연)

$i :$	$j :$	사업장에서 시작					지역에서 시작				
		사업	지역_1	지역_2	대기	전체	사업	지역_1	지역_2	대기	전체
2005	28	0.934	0.033	0.002	0.033	0.969	0.078	0.899	0.048	0.023	0.149
2006	29	0.823	0.088	0.007	0.089	0.919	0.209	0.725	0.057	0.066	0.332
2007	30	0.745	0.126	0.018	0.129	0.893	0.295	0.605	0.088	0.100	0.483
2008	31	0.690	0.153	0.030	0.158	0.878	0.348	0.524	0.104	0.128	0.580
2009	32	0.648	0.174	0.033	0.178	0.859	0.382	0.469	0.089	0.149	0.620
2010	33	0.616	0.191	0.042	0.193	0.851	0.405	0.428	0.093	0.167	0.665
2011	34	0.590	0.204	0.050	0.207	0.846	0.420	0.397	0.097	0.183	0.700
2012	35	0.567	0.215	0.057	0.219	0.842	0.430	0.374	0.099	0.197	0.726
2013	36	0.547	0.224	0.066	0.229	0.841	0.435	0.356	0.104	0.210	0.749
2014	37	0.529	0.232	0.072	0.239	0.840	0.438	0.341	0.106	0.221	0.765
2015	38	0.514	0.239	0.080	0.248	0.841	0.438	0.330	0.110	0.232	0.780
2016	39	0.499	0.244	0.086	0.257	0.841	0.436	0.321	0.113	0.243	0.792
2017	40	0.485	0.250	0.092	0.265	0.842	0.433	0.314	0.116	0.253	0.802
2018	41	0.472	0.254	0.098	0.274	0.844	0.429	0.308	0.119	0.263	0.811
2019	42	0.461	0.258	0.102	0.281	0.845	0.425	0.304	0.121	0.272	0.817
2020	43	0.450	0.262	0.110	0.288	0.848	0.420	0.301	0.126	0.279	0.826
2021	44	0.440	0.266	0.116	0.294	0.850	0.415	0.299	0.130	0.286	0.831
2022	45	0.430	0.271	0.121	0.299	0.851	0.409	0.299	0.134	0.292	0.835
2023	46	0.421	0.275	0.127	0.304	0.853	0.403	0.299	0.138	0.298	0.840
2024	47	0.412	0.279	0.132	0.309	0.854	0.397	0.300	0.142	0.304	0.843
2025	48	0.403	0.284	0.138	0.314	0.854	0.390	0.301	0.146	0.309	0.845

<i>i</i> :	사업장에서 시작					지역에서 시작					
	<i>j</i> :	사업	지역_1	지역_2	대기	전체	사업	지역_1	지역_2	대기	전체
2026	49	0.393	0.289	0.144	0.318	0.855	0.382	0.304	0.152	0.314	0.848
2027	50	0.384	0.294	0.152	0.322	0.858	0.374	0.307	0.159	0.318	0.852
2028	51	0.374	0.300	0.160	0.326	0.860	0.366	0.311	0.166	0.323	0.855
2029	52	0.365	0.305	0.167	0.330	0.862	0.358	0.315	0.172	0.327	0.857
2030	53	0.355	0.310	0.176	0.334	0.866	0.350	0.319	0.181	0.331	0.862
2031	54	0.340	0.316	0.186	0.343	0.869	0.336	0.324	0.190	0.340	0.866
2032	55	0.323	0.323	0.197	0.354	0.874	0.319	0.329	0.201	0.351	0.871
2033	56	0.305	0.332	0.209	0.363	0.877	0.302	0.338	0.213	0.360	0.875
2034	57	0.284	0.344	0.227	0.372	0.883	0.282	0.349	0.230	0.369	0.881
2035	58	0.264	0.354	0.237	0.382	0.883	0.262	0.358	0.240	0.380	0.882
2036	59	0.249	0.360	0.246	0.391	0.886	0.247	0.363	0.248	0.389	0.885

## 1. 기여연수

다음은  $x=28$ 세 남자 가입자의 다음과 같이 정의되는 연령별 누적 가입 연수와 누적 기여연수  $n_{x+t|x}^j$ 를 나타내고 있다.

$$n_{x+t|x}^j = \sum_{k=0}^{t-1} \sum_{j=W,R} \frac{{}_k p_x^j + {}_{k+1} p_x^j}{2}, \quad i=W, R \quad (13)$$

다음의 <표 7>은 28세에 사업장으로 가입한 경우 19.00년, 지역으로 가입한 경우 16.05년으로 나타나서 약 3년의 기여연수의 차이를 보인다. 누적 가입연수와 기여

연수와의 차이는 가입자 자격은 유지하지만 보험료는 납부예외 또는 체납으로 실제는 보험료를 납부하지 않은 기간을 의미한다.

〈표 7〉 경과 시점별 누적 가입연수와 기여연수

(단위: 연)

연령	사업장에서 시작					지역에서 시작				
	사업장	지역_1	지역_2	가입연수	기여연수	사업장	지역_1	지역_2	가입연수	기여연수
28	0.934	0.033	0.002	0.967	0.936	0.078	0.899	0.048	0.977	0.126
29	1.757	0.121	0.009	1.878	1.765	0.287	1.624	0.105	1.911	0.392
30	2.502	0.247	0.027	2.748	2.529	0.582	2.229	0.193	2.811	0.775
31	3.192	0.399	0.057	3.591	3.249	0.930	2.754	0.297	3.683	1.227
32	3.840	0.573	0.090	4.413	3.930	1.311	3.222	0.386	4.534	1.697
33	4.456	0.764	0.132	5.220	4.588	1.716	3.651	0.479	5.367	2.195
34	5.046	0.967	0.181	6.013	5.227	2.136	4.048	0.576	6.184	2.712
35	5.612	1.182	0.238	6.794	5.851	2.566	4.421	0.675	6.987	3.241
36	6.159	1.406	0.304	7.565	6.463	3.001	4.777	0.779	7.778	3.780
37	6.688	1.638	0.376	8.327	7.065	3.439	5.118	0.886	8.557	4.324
38	7.202	1.877	0.456	9.079	7.658	3.877	5.448	0.995	9.325	4.873
39	7.701	2.121	0.542	9.822	8.243	4.314	5.768	1.108	10.082	5.422
40	8.186	2.371	0.634	10.557	8.820	4.747	6.082	1.224	10.829	5.971
41	8.658	2.625	0.732	11.284	9.391	5.176	6.390	1.343	11.566	6.519
42	9.119	2.883	0.835	12.002	9.954	5.601	6.693	1.464	12.294	7.064
43	9.569	3.145	0.945	12.714	10.514	6.021	6.994	1.590	13.015	7.610
44	10.010	3.411	1.061	13.421	11.070	6.435	7.293	1.720	13.729	8.155
45	10.440	3.682	1.182	14.122	11.622	6.844	7.592	1.854	14.436	8.698

연령	사업장에서 시작					지역에서 시작				
	사업장	지역_1	지역_2	가입연수	기여연수	사업장	지역_1	지역_2	가입연수	기여연수
46	10.861	3.956	1.309	14.818	12.170	7.247	7.891	1.992	15.138	9.240
47	11.273	4.235	1.441	15.509	12.715	7.644	8.191	2.135	15.834	9.779
48	11.676	4.519	1.579	16.195	13.255	8.033	8.492	2.281	16.525	10.315
49	12.069	4.807	1.723	16.877	13.792	8.416	8.796	2.433	17.212	10.849
50	12.453	5.101	1.875	17.554	14.328	8.790	9.104	2.592	17.893	11.382
51	12.827	5.401	2.035	18.228	14.862	9.156	9.415	2.758	18.571	11.914
52	13.192	5.706	2.202	18.898	15.394	9.515	9.730	2.930	19.244	12.445
53	13.547	6.016	2.378	19.564	15.925	9.864	10.049	3.111	19.913	12.975
54	13.888	6.333	2.563	20.220	16.451	10.200	10.373	3.301	20.572	13.501
55	14.210	6.656	2.760	20.866	16.970	10.519	10.702	3.502	21.221	14.021
56	14.516	6.988	2.969	21.504	17.485	10.821	11.039	3.714	21.861	14.536
57	14.800	7.332	3.196	22.132	17.996	11.103	11.388	3.944	22.491	15.047
58	15.064	7.686	3.433	22.750	18.497	11.366	11.746	4.184	23.112	15.549
59	15.313	8.046	3.679	23.358	18.992	11.613	12.109	4.432	23.722	16.045

주: 지역\_1은 가입연수, 지역\_2는 실제 보험료를 납부한 기여연수.

원칙적으로 국민연금 가입자의 가입자 자격을 유지한 가입연수와 가입기간 중에 실제 보험료를 납부한 기여연수는 국민연금 재정추계모형에 의해서 측정될 수 있다. 다음 <표 8>은 국민연금 재정추계 모형으로부터 추계된 평균 기여연수이다. 표에서 기여연수는 노령연금 수급자를 기준으로 한 것이다. 연도에 따라 조금씩 기여연수가 증가하는 것은 1988년 제도 도입에 의한 초기 조건의 영향, 그리고 경제활동참가율의 증가, 납부예외율, 징수률의 개선이 혼합된 효과로 볼 수 있다. 보통 국민연금의 기여연수로서 인용되는 21.7년은 이러한 변동하는 값들이 안정되는 2070년 수급자를 기준으로 한 것이다.

〈표 8〉 재정추계 모형에 의한 평균 기여연수

연도	2010	2020	2030	2050	2070
평균가입 연수	15.0	15.8	17.6	20.7	21.7

자료: 국민연금발전위원회, 2003. 6, p.105.

국민연금 재정추계 모형은 가입자의 상태에 따라 납부예외율, 징수율을 적용하여 매 시점별 가입자 cohort별<sup>11)</sup> 보험료 기여금액을 추정하므로 실질적으로 Markov chain 모형과 동일한 정보를 사용하고 있다. 그러나 국민연금 재정추계 모형은 재정수지의 전망에 초점을 두고 개발되었으며 제도의 세밀한 부분들을 묘사하는 복잡한 시뮬레이션 모형이므로 별도의 분석에 의한 집계が必要하며 추계 결과의 해석이 어렵다<sup>12)</sup>.

## 2. 보험료 부과대상 소득

국민연금의 보험료는 정률 체계로서 부과대상 소득의 9%를 보험료로 납부하게 되어있다. 가입자의 생애 중 보험료 부과소득 또는 연령소득곡선(age-earning profile)의 모형에서 보편적으로 사용되는 방법은 초기 소득을 전반적인 임금상승률과 생애 연령별 임금곡선을 복합적으로 반영하는 것이다. 국민연금 가입자의 연령소득곡선으로 문형표(1995), 전병목(2002)은 국민연금 연령별 평균 부과소득(표준소득월액) 통계를 사용하였으며 석재은·김용하(2002)는 임금통계를 사용하였다. 손해보험 상실소득 추계와 관련하여 마승렬(2002)은 임금통계를 사용하였다.

다음 식 (14)는  $x$ 세 가입 이후  $t$ 시간 경과 시점의 가입 상태별( $j=W, R$ )로 보험료

11) 구체적으로 재정추계 모형의 가입자 cohort는 성별, 연령별, 가입종별로 세분되어 관리된다.

12) 체계적으로 재정추계 결과로부터 기여연수를 추계하는 과정을 정리하면 Markov 추계 모형과 유사한 산식으로 나타날 것으로 예상된다.

부과소득  $z_{x+t}^j$ 를 나타낸 것이다. 여기서는 전체적인 임금상승률과 국민연금 가입자의 연령별 평균 부과소득에 의하여 다음과 같이 나타냈다.

$$z_{x+t}^j = s_{x+t}^j \frac{A_{x+t}}{A_x}, \quad j = W, R \quad (14)$$

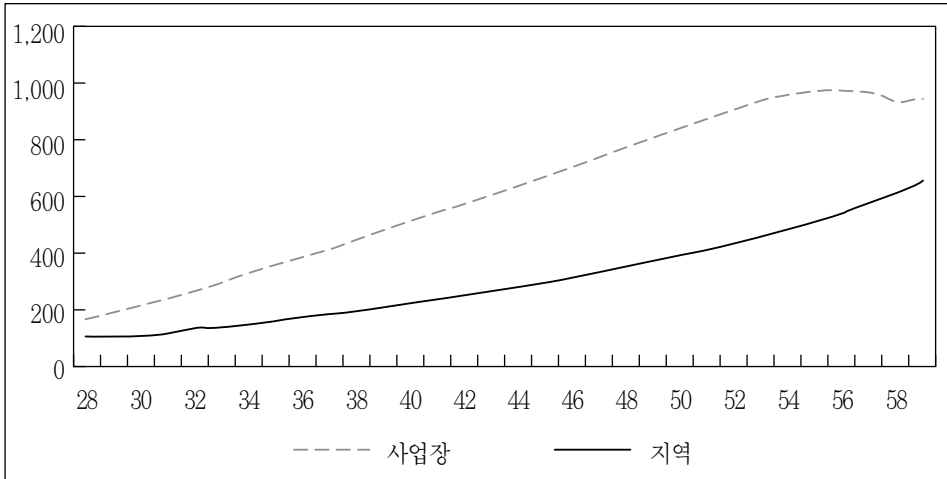
$s_{x+t}$ 는 2005년도  $x$ 세 국민연금 가입자가  $t$ 년 경과 후의 가입 상태별( $j = W, R$ ), 연령별 평균 표준소득월액이다. 국민연금 가입자의 전체 평균 표준소득월액  $A$ 값의 상승률  $A_{x+t}/A_x$ 는 경제의 전반적인 임금상승률에 해당한다. 국민연금 가입자 통계에 의하면 2005년 28세 부과소득은 사업장 가입자는 170.7만원, 지역가입자는 102.8만원이다. 다음 <그림 6>은 식 (14)에 의해서 산정한 가입 상태별 보험료 부과소득이다. 그림에서 횡축은 연령과 연도를 동시에 나타내는 것이다. 평균적 개인을 상징하고 있으므로 28세는 동시에 2005년을, 59세는 동시에 2036년을 나타낸다.  $A$ 값의 상승률은 2003년 재정계산제도에서 가정된 거시경제변수의 해당 연도 임금상승률을 적용하였다<sup>13)</sup>.

13) 국민연금발전위원회, 2003. 6. p.105.



〈그림 6〉 가입종별 연령별 표준소득월액

(단위: 만원)



가입상태  $i$ 에서 시작한 가입자의  $t$ 년 경과한  $x+t$ 세, 즉  $(x+t, x+t+1)$  기간 중에 사업장과 지역에서의 보험료 부과소득은 식 (12)에서 정의된 사업장과 지역 체류 시간  $d_{x+t|x}^{iW}, d_{x+t|x}^{iR}$ 에 각각에 적합한 식 (14)의 보험료 부과소득  $z_{x+t}^i$ 를 곱하면 다음과 같이 정해진다. 다음 식 (15)는 각 상태별 보험료 부과소득을 체류시간으로 가중 평균한 값이다.

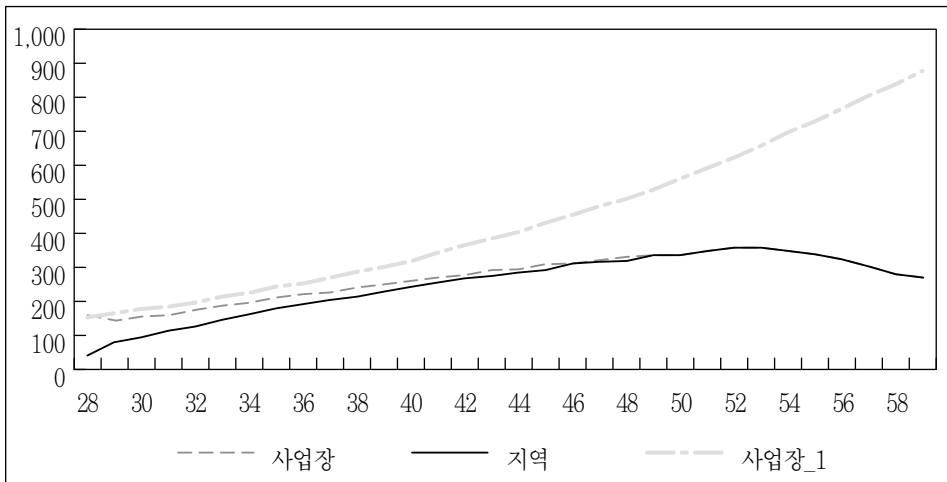
$$y_{x+t}^i = d_{x+t|x}^{iW} \times z_{x+t}^W + d_{x+t|x}^{iR} \times z_{x+t}^R, \quad i = W, R \quad (15)$$

다음 〈그림 7〉은 식 (15)에 의해서 산정된 사업장과 지역에서 시작한 가입자의 보험료 부과소득  $y_{x+t}^i$ 을 나타내고 있다. 가입 이후 약 15년 경과할 때까지 사업장에서 시작한 가입자의 가중평균 부과소득( $y_{x+t}^W$ )이 지역에서 시작한 가입자의 가중평균 부과소득( $y_{x+t}^R$ )에 비해 다소 높지만 이후는 거의 같아지는 것으로 나타나고 있다. 그림에서 사업장\_1은 비교를 위하여 나타난 사업장에서 시작한 가입자의 부과소득  $z_{x+t}^W$ 이다. 사업장과 지역의 부과소득  $y_{x+t}^i$  ( $i=W, R$ )는 부과소득  $z_{x+t}^W$ 을 많이 미

달하는 것으로 나타나고 있다. 그것은 여기서 제시되는 부과소득은 대기자에서의 체류시간, 납부예외 또는 미납 기간이 제외된 것에 기인한다. 전체적으로 사업장에서 시작한 가입자의 식 (14)에 의해서 산정된 부과소득  $z_{x+t}^w$ 의 절반에도 미치지 못하는 것으로 나타난다.

〈그림 7〉 가입종별 경과시점별 부과소득

(단위: 만원)



## V. 결론 및 향후 연구과제

본 연구는 보다 현실성이 높은 기여-수급 모형의 개발을 위하여 Markov chain 방법론에 의한 국민연금 기여 모형을 제안하였다. Markov chain 기여 모형의 장점은 국민연금 가입자의 생애 중에 발생하는 사업장과 지역 간의 전이 현상을 반영한 체류시간을 통하여 보험료 부과소득과 기여연수를 현실성 있게 결정할 수 있는 것이다. 국민연금은 확정급여 제도이므로 기여 모형에서 결정된 연령별 부과소득의 현금흐름으로부터 연금급여의 수급 모형은 비교적 용이하게 결정될 수 있다.

Markov chain 기여 모형은 기본적으로 확률적 모형이므로 보험계리적 수급모형과 결합하면 전체적으로 확률적인 국민연금의 기여-수급모형이 완성된다.

본 연구에서 개발된 Markov chain 기여 모형의 한계점과 향후 과제는 주로 Markov chain 모형이 전적으로 의존하는 전이확률들과 관련된다. 첫째, 본 연구는 국민연금 가입자의 가입상태 전이에 대한 Markov chain 가정의 타당성에 대한 통계적 검정이 생략되었다. 둘째, 본 연구는 국민연금 재정추계 모형을 위한 최근의 추정치를 사용하였다. 그러나 노동생명표에서 사용되는 보다 정교한 전이확률의 추정방법(e.g. Schoen and Woodrow, 1980)에 대한 연구가 필요하다. 셋째, 추정된 전이확률이 상당히 장기간에 걸쳐 유지된다는 가정에 대한 검토가 필요하다. 그것은 우리나라에서 지역가입자를 구성하는 주요 근로 형태인 자영자의 비중이 현재 선진국들에 비해 유난히 높아서 향후 점진적으로 감소할 것이라는 사실을 반영하는 노력이 필요하다. 마지막으로 본 연구에서 사용된 Markov chain 방법론은 여성 등 한계근로계층의 전이확률을 추정을 통하여 국민연금의 최소 기여연수 10년을 충족시키지 못하는 가입자들의 규모에 대한 측정에도 활용될 수 있을 것이다.

## 참 고 문 헌

- 국민연금관리공단, 『2005 국민연금통계연보』, 2006.
- 국민연금발전위원회, 『2003국민연금 재정계산 및 제도개선방안』, 2003. 6.
- 김원식, 『개인연금제도의 정착을 위한 정책방안연구』, 한국금융연구원, 1996.
- 마승렬, 「피해자의 특성을 고려한 상실 수익액 산정방법」, 『보험학회지』, 제63호, 한국보험학회, 2002. 12, pp.145~177.
- 문형표, 『국민연금제도의 재정건실화를 위한 구조개선방안』, 한국개발연구원, 1995.
- 박성민, 『국민연금 가입자의 가입기간별 추계』, 국민연금연구원, 연구보고서, 2006.
- 석재은·김용하, 「국민연금의 소득보장효과에 대한 Simulation 분석」, 『사회보장연구』, 제18권, 제1호, 한국사회보장학회, 2002, pp.67~104.
- 전병목, 『국민연금 중장기 재정추계 모형』, 한국조세연구원, 연구보고서, 2002.
- 최기홍, 「우리나라 국민연금 내부수익률의 기여기간에 대한 비교정태 특성」, 『공공경제』, 제10권, 제2호, 한국재정공공경제학회, 2005. 11, pp.105~127.
- Hoem, J. M., "Point estimation of forces of transition in demographic models", *Journal of Royal Statistical Society, Series B*, No.33, 1970, pp.275~289.
- \_\_\_\_\_, "A Markov Chain Model of Working Life Tables", *Scandinavian Actuarial Journal*, Vol.58, No.1, 1977, pp.1~20.
- Holzmann, R., Palacios, R. and Asta Zvinieni, "Implicit Pension Debt: Issues, Measurement and Scope in International Perspective", World Bank, 2001.
- Millimet, D.L., Nieswiadomy, M., H. Ryu, D. Slottje, "Estimating worklife expectancy: an econometric approach", *Journal of Econometrics*, No.113, 2003, pp.83~113.
- Schoen, R. and Woodrow, K., "Labor Force Status Life Tables for the United States, 1972", *Demography*, Vol.17, No.3, August 1980, pp.297~322.
- Smith, S., "New worklife estimates reflect changing profile of labor force", *Monthly Labor Review*, Vol.17, No.3, U.S. Department of Labor, March 1982.
- \_\_\_\_\_, "Tables of Working Life: The Increment-Decrement Model", *BLS Bulletin* 2135, US. Bureau of Labor Statistics, 1982.

- \_\_\_\_\_, “Worklife Estimates: Effects of Race and Education”, *BLS Bulletin 2254*, US. Bureau of Labor Statistics, 1986.
- Wessels, W.J., “A Test of Increment-Decrement Method of Calculating Worklife Expectancies”, *Journal of Forensic Economics*, 2006 forthcoming.
- Wolfbein, S. and H. Wool, “Table of Working Life: Length of Working life for Men”, *BLS Bulletin 1001*, US. Bureau of Labor Statistics, 1950.

## Abstract

This paper applies the methodology of Markov chain work-life table in modelling the contribution behavior of National Pension insured persons. This model enables us to model the transition pattern between types of coverage, namely workplace based insured or individually insured in the National Pension system. It is well known that there are significant differences in the amount of monthly contributory incomes as well as in the intensity of contribution by types of coverage, which has been major troubles in constructing realistic models of contribution behavior for the insured persons. In theory, this paper proves that the backward recursion formular of Hoem(1977) and the forward recursion formular of BLS in the work-life expectancy literature have identical solution and identified the closed form solution. The solution provides the formular for expected sojourn time at each age since the insured person enters the National Pension system. The expected sojourn time at each age also provides weighted averaged contributory incomes as well as the cumulative insured period net of non-contribution period due to unemployment etc. This Markov chain contribution model combined with actuarial pension benefit model can make a complete probabilistic contribution-pension benefit model for the National Pension system of Korea.

※ Key Words: contribution model, contributory income, insured period, Markov chain, national pension

【부 록 1】

상태간 연령별 전이확률  $p_x^{ij}$ <sup>14)</sup>

구분	W			R			H		
	W	R	H	W	R	H	W	R	H
28	0.868	0.066	0.066	0.156	0.797	0.047	0.172	0.061	0.767
29	0.875	0.062	0.063	0.151	0.804	0.045	0.127	0.054	0.818
30	0.882	0.058	0.060	0.132	0.823	0.045	0.113	0.049	0.838
31	0.888	0.054	0.057	0.115	0.842	0.043	0.117	0.046	0.837
32	0.892	0.053	0.055	0.105	0.853	0.041	0.108	0.046	0.846
33	0.897	0.051	0.052	0.099	0.860	0.041	0.096	0.045	0.859
34	0.900	0.049	0.051	0.093	0.867	0.040	0.087	0.043	0.871
35	0.902	0.048	0.049	0.088	0.872	0.040	0.078	0.043	0.878
36	0.903	0.047	0.049	0.084	0.875	0.041	0.075	0.043	0.882
37	0.904	0.047	0.050	0.081	0.878	0.041	0.073	0.042	0.885
38	0.905	0.046	0.049	0.077	0.881	0.042	0.067	0.041	0.891
39	0.905	0.045	0.049	0.073	0.885	0.042	0.064	0.041	0.895
40	0.905	0.045	0.050	0.070	0.887	0.043	0.062	0.040	0.898
41	0.905	0.045	0.050	0.068	0.888	0.044	0.059	0.039	0.901
42	0.905	0.046	0.049	0.066	0.891	0.043	0.058	0.038	0.903
43	0.906	0.045	0.048	0.064	0.894	0.042	0.057	0.038	0.905
44	0.905	0.046	0.049	0.060	0.898	0.041	0.055	0.038	0.906
45	0.906	0.046	0.049	0.059	0.900	0.041	0.054	0.037	0.909
46	0.905	0.047	0.049	0.057	0.902	0.041	0.054	0.036	0.910

14) 박성민(2006) 추정치.

구분	W			R			H		
	W	R	H	W	R	H	W	R	H
47	0.903	0.048	0.049	0.055	0.905	0.040	0.052	0.034	0.913
48	0.901	0.050	0.049	0.053	0.908	0.039	0.050	0.033	0.916
49	0.901	0.051	0.048	0.052	0.910	0.038	0.049	0.034	0.917
50	0.897	0.054	0.049	0.051	0.912	0.037	0.047	0.032	0.919
51	0.895	0.056	0.048	0.051	0.913	0.035	0.046	0.031	0.922
52	0.894	0.057	0.049	0.050	0.915	0.035	0.046	0.029	0.924
53	0.888	0.062	0.050	0.049	0.917	0.035	0.044	0.028	0.927
54	0.862	0.074	0.063	0.049	0.908	0.043	0.042	0.025	0.930
55	0.856	0.082	0.062	0.049	0.912	0.039	0.043	0.025	0.930
56	0.841	0.093	0.064	0.048	0.915	0.037	0.043	0.025	0.930
57	0.815	0.112	0.071	0.048	0.916	0.036	0.040	0.022	0.935
58	0.825	0.105	0.069	0.047	0.917	0.035	0.037	0.019	0.942
59	0.825	0.105	0.069	0.047	0.917	0.035	0.037	0.019	0.942



【부 록 2】

상태간 연령별 다단계 전이확률  ${}_n P_x^{ij}$  <sup>15)</sup>

구분	W			R			H		
	W	R	H	W	R	H	W	R	H
28	0.868	0.066	0.066	0.156	0.797	0.047	0.172	0.061	0.767
29	0.778	0.110	0.112	0.262	0.653	0.084	0.258	0.101	0.641
30	0.713	0.141	0.146	0.327	0.557	0.116	0.313	0.129	0.557
31	0.667	0.164	0.169	0.368	0.492	0.140	0.358	0.152	0.490
32	0.630	0.183	0.186	0.395	0.446	0.159	0.388	0.171	0.441
33	0.602	0.198	0.200	0.414	0.411	0.175	0.408	0.187	0.406
34	0.577	0.210	0.213	0.426	0.384	0.190	0.419	0.199	0.381
35	0.556	0.220	0.224	0.433	0.363	0.203	0.426	0.210	0.364
36	0.538	0.229	0.234	0.437	0.348	0.216	0.430	0.220	0.350
37	0.521	0.236	0.243	0.439	0.335	0.227	0.432	0.228	0.340
38	0.506	0.242	0.252	0.438	0.324	0.238	0.431	0.235	0.334
39	0.492	0.247	0.261	0.435	0.317	0.248	0.428	0.241	0.330
40	0.479	0.252	0.269	0.431	0.311	0.258	0.425	0.247	0.328
41	0.466	0.256	0.278	0.427	0.305	0.268	0.421	0.251	0.328
42	0.455	0.260	0.285	0.422	0.302	0.276	0.417	0.255	0.328
43	0.445	0.264	0.291	0.418	0.299	0.283	0.413	0.260	0.328
44	0.435	0.269	0.296	0.412	0.299	0.289	0.407	0.265	0.328
45	0.426	0.273	0.302	0.406	0.299	0.295	0.402	0.269	0.329
46	0.417	0.277	0.306	0.400	0.299	0.301	0.397	0.273	0.330

15) <부록 1>로부터 Chapman-Kolmogorov 재귀식으로 구하였음.

구분	W			R			H		
	W	R	H	W	R	H	W	R	H
47	0.407	0.281	0.312	0.393	0.300	0.306	0.390	0.278	0.332
48	0.398	0.286	0.316	0.386	0.303	0.312	0.383	0.283	0.334
49	0.389	0.291	0.320	0.379	0.306	0.316	0.376	0.288	0.335
50	0.379	0.297	0.324	0.370	0.309	0.321	0.368	0.294	0.338
51	0.370	0.302	0.328	0.362	0.313	0.325	0.361	0.300	0.340
52	0.360	0.307	0.332	0.354	0.317	0.329	0.353	0.305	0.342
53	0.350	0.313	0.337	0.345	0.321	0.334	0.343	0.311	0.346
54	0.331	0.319	0.350	0.327	0.326	0.347	0.326	0.316	0.358
55	0.314	0.327	0.359	0.311	0.333	0.356	0.310	0.324	0.366
56	0.296	0.337	0.367	0.293	0.342	0.364	0.293	0.335	0.373
57	0.273	0.351	0.377	0.271	0.355	0.375	0.270	0.348	0.382
58	0.256	0.357	0.387	0.254	0.361	0.385	0.253	0.355	0.392
59	0.242	0.362	0.396	0.241	0.365	0.394	0.240	0.360	0.400