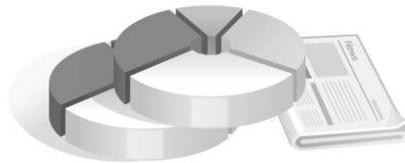


# 통계와 정책 표준교재





# 제 출 문

## 통계교육원장 귀하

『통계와 정책 표준교재 개발』에 대한 최종보고서를 다음과 같이 제출합니다.

2008년 11월

연구 기관 : 한국행정학회

연구책임자 : 이석훈(충남대학교 정보통계학과 교수)  
참여연구원 : 김성철(명지대학교 행정학과 교수)  
박진우(수원대학교 통계정보학과 교수)  
유홍림(단국대학교 행정학과 교수)  
전명식(고려대학교 통계학과 교수)  
조재근(경성대학교 정보통계학과 교수)





# CONTENTS

CHAPTER

1

## 서론 .....1

1.1 국정과 통계 .....3

1.2 두 가지 정책 사례 .....4

1.3 이 책의 특징과 구성 .....6

■ 요약 .....9

CHAPTER

2

## 정책과정과 통계 .....11

2.1 정책과정의 제 단계 .....13

2.1.1 정책입안 단계 .....13

2.1.2 정책집행 단계 .....18

2.1.3 정책평가 단계 .....20

2.2 정책과정에서 통계의 역할 .....23

2.2.1 정책입안 단계에서 통계의 역할 .....24

2.2.2 정책집행 단계에서 통계의 역할 .....24

2.2.3 정책평가 단계에서 통계의 역할 .....25

2.3 통계학사 속에서 살펴본 정책과 통계 .....26

2.3.1 등불 대신 통계로 병원을 밝힌 여인 .....26

2.3.2 역사 속의 정책과 통계 .....29

2.3.3 영국의 빈곤정책과 부스의 통계조사 .....31

2.3.4 통계에 대한 저항과 통계의 한계 .....35

2.4 정책기반으로서의 통계 .....36

2.4.1 대표적 정책기반 통계 사례 : 인구주택총조사 .....39

2.4.2 통계 부실로 인한 정책상 문제점 .....46

2.4.3 정확하고 타당한 통계를 활용하지 못하는 원인 분석 .....49

2.4.4 정책기반 통계를 작성할 때 고려해야 할 사항 .....53

■ 요약 .....58



CHAPTER

3

통계와 정책의 사례 연구 .....59

- 3.1 주택 관련 통계 사례 .....62
- 3.2 사교육비실태조사 사례 .....75
- 3.3 손상감시체계 구축 사례 .....87
- 3.4 성과지표 설정 사례 .....100
- 3.5 도서관예산안 작성 연습과제 .....104
  - 3.5.1 연습과제 개요 .....104
  - 3.5.2 예산편성 지침 .....104
  - 3.5.3 과제 풀이 .....112
- 3.6 통계 DB 활용 연습 .....121
- 3.7 정책과정에서 고려해야 할 통계 사례 .....128
- 요약 .....141

CHAPTER

4

정책과정과 자료 .....143

- 4.1 통계적 사고 .....145
  - 4.1.1 통계적 사고의 필요성 .....145
  - 4.1.2 통계적 사고가 필요한 경우들 .....154
- 4.2 정책과정에서의 자료수집 .....156
- 4.3 자료의 표현 .....158
- 4.4 자주 잘못 사용되는 통계 용어들 .....184
  - 4.4.1 지수의 정의와 3요소 .....184
  - 4.4.2 변동률(증감률) .....185
  - 4.4.3 진척률 .....188
  - 4.4.4 비, 비율, 율 .....189
  - 4.4.5 퍼센트와 퍼센트 포인트 .....192
- 요약 .....194



CHAPTER  
**5**

**정책과정에서 나타나는 쟁점들 .....195**

5.1 TOBANKRUPT 은행의 영업계수는 비정상인가? .....197

5.2 신뢰수준 95%에서 표본오차  $\pm 3.1\%$ 포인트는? .....200

5.3 정책 만족도가 지역 간에 다르게 나타났는가? .....202

    5.3.1 남녀 간 월소득의 차이가 해결되었는가? .....203

    5.3.2 지역 간 만족도의 차이가 해결되었는가? .....206

5.4 근무 만족도가 높은 기관이 평가도 잘 받는가? .....209

    5.4.1 산점도와 상관계수 .....210

    5.4.2 관계의 표현과 회귀분석 .....211

5.5 교육효과는 언제부터 나타났는가? .....216

    5.5.1 시계열 자료 .....216

    5.5.2 이동평균법 .....217

    5.5.3 시계열 자료 토의 .....222

5.6 알아두면 유용한 기법들 .....226

    5.6.1 주민 만족도 조사의 발전 방향 제시 .....226

    5.6.2 자료 활용의 효용성 제고 방안 .....230

    5.6.3 지역 특성의 시각화 방안 .....235

    5.6.4 정책지지자 성향분석 방안 .....238

■ 요약 .....241

**참고문헌 .....242**

**부록 .....245**



표  
목  
록

<표 2-1> 정책집행에 있어서 종합모형 .....19

<표 2-2> 성과지표의 미시적·거시적 특성 .....22

<표 2-3> 부문별 및 작성방법별 국가승인통계 수 .....38

<표 2-4> 각 정부기관별 활용내용 .....41

<표 2-5> 주요 표본조사들 .....42

<표 2-6> 인구주택총조사의 활용사례 .....43

<표 2-7> 근대적 총조사 역사, 명칭 및 특징 .....44

<표 2-8> 예측교통량 비교 .....48

<표 2-9> 산업별 전수조사대상 사업체 조사 현황 .....51

<표 3-1> 2005년 주택총조사 결과 주택유형별 주택호수 .....63

<표 3-2> 부동산 정책의 효과 비교 .....74

<표 3-3> 사교육비실태조사를 위한 표본 수 .....80

<표 3-4> 사교육비실태조사의 조사항목 .....80

<표 3-5> 학교급별 사교육비 통계 .....82

<표 3-6> 특성별 사교육비 현황 : 교과목별 사교육비 .....83

<표 3-7> 특성별 사교육비 현황 : 지역별 사교육비 .....83

<표 3-8> 특성별 사교육비 현황 : 학생 성적별 사교육비 .....83

<표 3-9> 특성별 사교육비 현황 : 부모 교육수준별 사교육비 .....83

<표 3-10> 연도별 사망자 수 및 손상 사망자 수 .....94

<표 3-11> 퇴원환자 수 추정값 .....95

<표 3-12> 손상요인별 환자 수 추정 .....96

<표 3-13> 응급실 손상환자 통계 .....96

<표 3-14> 연령별 응급실 손상환자 통계 .....97

<표 3-15> 손상 원인별, 장소별 손상발생 수 .....99

<표 3-16> 복지 분야 성과지표 설정사례 .....100

<표 3-17> 양식1 : 2008 회계연도 태평시 1구 도서관 예산 .....106

<표 3-18> 양식2 : 2008 회계연도 태평시 2구 도서관 예산 .....107

<표 3-19> 양식3 : 2008 회계연도 태평시 3구 도서관 예산 .....108

<표 3-20> 양식4 : 2008 회계연도 태평시 4구 도서관 예산 .....109

<표 3-21> 양식5 : 2008 회계연도 태평시 5구 도서관 예산 .....110

<표 3-22> 구별 인구 변화 .....111

<표 3-23> 구별 10세~29세 인구 변화 .....111

<표 3-24> 2008년 도서관 예산 총액 .....112

<표 3-25> 2009년 도서관 예산 총액 .....115

<표 3-26> 2009년 1구 도서관 예산 예시 .....118

<표 3-27> 2009년 2/3구 도서관 예산 예시 .....119

<표 3-28> 2009년 4/5구 도서관 예산 예시 .....120

<표 3-29> 검색을 통해 최종 정리한 결과 .....125

<표 3-30> 필요 통계지표 .....129

<표 3-31> 매출채권규모 파악과 보험료율 산정을 위한 지표들 .....131

<표 3-32> 통계지표 .....131

<표 3-33> 통계지표 구비여부 검토에 대한 소관부처의 의견 .....133

<표 3-34> 필요 통계지표 .....133

<표 3-35> 사이버대학 운영 .....134

<표 3-36> 고등교육기관으로의 전환 기준 .....135

<표 3-37> 통계지표 .....136

<표 3-38> 대학입학전형관련 권한 .....136

<표 3-39> 가족친화 사회 환경 조성사업 .....137



〈표 4-1〉 2005년도 한국의 월평균 납세액 자료(일부) .....	145
〈표 4-2〉 2005/2006년도의 교육정도별 경제활동인구의 고용률 .....	150
〈표 4-3〉 2004년도 서적 및 인쇄물 관련 가구당 월평균 지출액 .....	152
〈표 4-4〉 기술통계량 .....	159
〈표 4-5〉 품목별지수와 가중치 .....	161
〈표 4-6〉 2005~2007년도의 한국의 일인당 국민소득, 소비자물가지수, 불변화폐액 .....	161
〈표 4-7〉 고위공직자 재산가액 현황 .....	163
〈표 4-8〉 응답자 분포 .....	164
〈표 4-9〉 4인가구의 월생활비 자료 .....	166
〈표 4-10〉 이상치 판별 규칙 .....	167
〈표 4-11〉 주요 도시의 오존(O3)오염도 .....	169
〈표 4-12〉 21개국의 경제 분야별 일자리 비율과 표준점수 .....	172
〈표 4-13〉 TOBANKRUPT 은행의 영업계수와 정상운영 50개 은행 영업계수 .....	173
〈표 4-14〉 정책지지여부의 도수분포표 .....	174
〈표 4-15〉 30대 미혼인 사람들의 한 달 생활비 .....	174
〈표 4-16〉 30대 미혼인 사람들의 한 달 생활비의 도수분포표 .....	175
〈표 4-17〉 4인가구 가구주의 생활만족도 .....	176
〈표 4-18〉 생활만족도의 도수분포표 .....	177
〈표 4-19〉 기관 만족도 .....	178
〈표 4-20〉 특정 광물의 생산량 자료 .....	186
〈표 4-21〉 우리나라의 실업률 .....	187
〈표 4-22〉 진척률 .....	188
〈표 5-1〉 특정 업종에 근무하는 남녀 근로자들의 월소득 .....	203
〈표 5-2〉 특정 업종에 근무하는 남녀 근로자들의 월소득의 평균과 표준편차 .....	203
〈표 5-3〉 성별에 따른 특정 정책의 지지현황 .....	205
〈표 5-4〉 4개 지역의 정책 만족도 .....	207
〈표 5-5〉 일본의 합계출산율 .....	213
〈표 5-6〉 민원인 불평건수 현황 .....	218
〈표 5-7〉 5항 이동평균 계산을 수행한 결과 .....	219
〈표 5-8〉 연도별 성과달성도 .....	222
〈표 5-9〉 가정용 전력 판매량 .....	223
〈표 5-10〉 지역 정책에 대한 주민의 만족도 조사 문항 .....	226
〈표 5-11〉 조사 자료(1) .....	227
〈표 5-12〉 조사 자료(2) .....	227
〈표 5-13〉 2001년 사회지표 조사 자료 구조 .....	231
〈표 5-14〉 2002년 사회지표 조사 자료 구조 .....	231
〈표 5-15〉 2003년 사회지표 조사 자료 구조 .....	232
〈표 5-16〉 데이터 퓨전에 의한 최종 자료 .....	233
〈표 5-17〉 2007년도 품목별 소비자물가지수 및 주성분점수 .....	235
〈표 5-18〉 5개 과목의 성적 .....	237
〈표 5-19〉 표준점수 .....	237
〈표 5-20〉 예상 가구특성 변수 .....	240



그림  
목록

〈그림 2-1〉 등불을 들고 있는 나이팅게일의 동상 .....26

〈그림 2-2〉 나이팅게일이 1857년에 개발한 coxcombs .....28

〈그림 2-3〉 구스타프 도레가 그린 런던의 빈곤 .....33

〈그림 2-4〉 2005년 인구주택총조사 홍보물 .....39

〈그림 2-5〉 국가통계포털의 자료검색결과 화면 .....45

〈그림 3-1〉 주택유형별 주택호수(2005년 주택총조사) .....63

〈그림 3-2〉 연도에 따른 주택유형별 주택수의 변화 .....64

〈그림 3-3〉 우리나라의 주택보급률의 변화 .....65

〈그림 3-4〉 주요 도시별 인구 천 명당 주택 수 .....66

〈그림 3-5〉 주요 국가들의 인구 천 명당 주택 수 .....67

〈그림 3-6〉 우리나라의 토지가격 및 주택가격 변동률 .....68

〈그림 3-7〉 1987년의 주택가격동향조사 통계표 일부 .....69

〈그림 3-8〉 2008년의 주택가격동향조사 아파트 매매가격지수 일부그림 .....71

〈그림 3-9〉 우리나라 주택가격지수의 변화 .....72

〈그림 3-10〉 투기지역 지정과 관련된 신문기사 일부 .....73

〈그림 3-11〉 우리나라 전체 사교육비 규모의 추정치 .....78

〈그림 3-12〉 국가통계포털 사이트의 사교육비실태조사 관련 화면 .....81

〈그림 3-13〉 통계메타자료 화면 .....86

〈그림 3-14〉 손상 피라미드 .....89

〈그림 3-15〉 퇴원환자조사표 일부 .....93

〈그림 3-16〉 연도별 사망자 수와 손상 사망자 수 .....95

〈그림 3-17〉 2008년 항목별 예산액 .....113

〈그림 3-18〉 2008년과 2009년 예산총액 비교 .....114

〈그림 3-19〉 2008년과 2009년 항목별 예산액 비교 .....115

〈그림 3-20〉 2008년과 2009년 구별 예산액 비교 .....116

〈그림 3-21〉 관심지역의 연도별 인구 .....117

〈그림 3-22〉 관심지역의 연도별 도서관 이용자수 .....117

〈그림 3-23〉 국가통계포털 웹사이트 화면(<http://www.kosis.kr>) .....122

〈그림 3-24〉 주제별통계 중 인구·가구 조회 .....123

〈그림 3-25〉 주제별통계 중 인구·가구 조회 결과 .....123

〈그림 3-26〉 조회범위 설정을 위한 화면 .....124

〈그림 3-27〉 국가통계포털의 검색 결과 화면 .....124

〈그림 3-28〉 연도별 주택보급률과 인구 천 명당 주택 수 .....125

〈그림 3-29〉 e-나라지표의 인구 천 명당 주택 수 메뉴 .....126

〈그림 4-1〉 생활비의 상자그림(1) .....166

〈그림 4-2〉 생활비의 상자그림(2) .....167

〈그림 4-3〉 정책지지여부의 막대그래프 .....174

〈그림 4-4〉 정책지지여부의 원그래프 .....174

〈그림 4-5〉 한 달 생활비의 막대그래프 .....175

〈그림 4-6〉 한 달 생활비의 원그래프 .....175

〈그림 4-7〉 생활만족도에 대한 히스토그램 .....177

〈그림 4-8〉 기상자료의 즐기와 잎 그림 .....178

〈그림 4-9〉 절단된 막대그래프 .....179

〈그림 4-10〉 절단되지 않은 막대그래프 .....179

〈그림 4-11〉  $y$ 축 눈금을 조정한 막대그래프(1) .....179

〈그림 4-12〉  $y$ 축 눈금을 조정한 막대그래프(2) .....179

〈그림 4-13〉 너비를 조정한 막대그래프 .....180

〈그림 4-14〉 원그래프의 오류 .....180

〈그림 4-15〉 농업 종사인구비율의 히스토그램 .....181



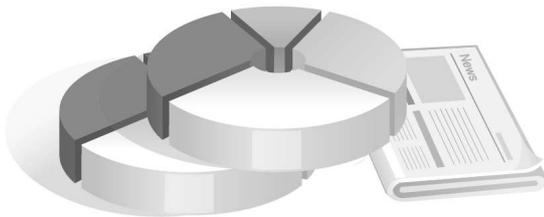
〈그림 4-16〉 서비스업 종사인구비율의 히스토그램 .....	181
〈그림 4-17〉 공업 종사인구비율의 히스토그램 .....	182
〈그림 4-18〉 업종별 종사인구비율의 상자그림 .....	182
〈그림 5-1〉 정규분포의 일반적인 형태 .....	197
〈그림 5-2〉 t-test 결과 .....	204
〈그림 5-3〉 z-test 결과 .....	205
〈그림 5-4〉 분산분석 출력 결과 .....	208
〈그림 5-5〉 평가점수와 만족도의 산점도 .....	210
〈그림 5-6〉 몸무게와 혈압의 산점도 .....	212
〈그림 5-7〉 일본의 합계출산율의 꺾은선그래프 .....	214
〈그림 5-8〉 합계 출산율과 추정치와의 비교 .....	214
〈그림 5-9〉 민원인 불평건수에 대한 꺾은선그래프 .....	218
〈그림 5-10〉 불평건수와 5일 이동평균 .....	220
〈그림 5-11〉 순환변동 .....	220
〈그림 5-12〉 연도별 성과달성도 .....	222
〈그림 5-13〉 가정용 전력 판매량 .....	224
〈그림 5-14〉 월별 가정용 전력 판매량 .....	224
〈그림 5-15〉 사회지표 조사의 데이터 퓨전 .....	232
〈그림 5-16〉 자료 결합(Data Matching) .....	234
〈그림 5-17〉 Biplot .....	236
〈그림 5-18〉 의사결정나무 .....	238



CHAPTER 1

# 서론

- 1.1 국정과 통계
- 1.2 두 가지 정책 사례
- 1.3 이 책의 특징과 구성



## 개요

새로운 정책을 만들고, 집행하고, 평가하는 전 과정에서 통계가 어떠한 역할을 해 왔는지 그 역사적 배경을 살펴보고, 국내에서 이슈가 되었던 새만금과 대운하 사례를 통해 국가 정책에서 통계의 필요성과 중요성을 알아본다.

## 주요내용

### 1.1 국정과 통계

통계는 과거와 현재를 파악하고 미래를 설계하기 위한 가장 유력한 자료인 한편, 갈등 조정을 위한 설득력 있는 도구 중 하나이다.

### 1.2 두 가지 정책 사례

국가에서 정책을 수립하고 실행, 평가할 때 적절한 통계를 얻어 이를 세밀히 활용하는 것이 얼마나 중요한지 새만금과 대운하 사례를 통해 토의한다.

### 1.3 이 책의 특징과 구성

세부적인 분석 방법보다는 정책과정에서 통계의 폭넓은 활용 가능성을 인식하고 통계에 한 걸음 다가설 수 있게 돕고자 한다.



## 1.1 국정과 통계

동서양을 막론하고 통계는 아주 옛날부터 국가가 주도적으로 생산하고 소비해왔다. 물론 ‘어떤 통계를 생산하여 어디에 활용할 것인가?’ 라는 문제는 그 나라의 정치체제에 따라 결정될 수밖에 없는 문제였다. 예컨대 몇 백 년 전 “짐이 곧 국가”였던 시대에는 국민들의 복리가 정복 전쟁보다 앞설 수 없었으므로 통계 또한 전쟁에 필요한 자금과 병력 확보를 위한 도구에 지나지 않았다. 대략 19세기 이전까지만 하더라도 왕이나 지배자들의 필요에 따라 생산되던 통계는 그 종류가 매우 단순했던 만큼, 정부 내에 별도의 기구를 설치하여 통계 업무를 맡기는 경우도 거의 없었다. 것처럼 통계 담당 공무원이나 부서도 제대로 없이 작성된 통계란 아무래도 부실할 수밖에 없었다. 하지만 그렇게 얻은 정보가 종종 국가적인 비밀로 보호되었던 데서 알 수 있듯이 통계는 이미 오래 전부터 꽤 값나가는 지식으로 인정받고 있었다.

그러다가 18세기 끝 무렵이 되자 절대적인 것처럼 보였던 지배 권력이 프랑스로부터 급속히 무너지는 사태가 벌어졌다. 이를 계기로 사람들은 기존의 지배 질서를 뒤흔들 수 있는 요인으로 ‘사회’라는 것을 점점 중요하게 인식하기 시작하게 되었고 사회 변화에 따라 적절한 사회정책을 세우는 문제가 시급한 당면 과제로 떠올랐다. 그런데 사회의 변화라는 것은 마치 지하 깊숙이에서 도도하게 흐르는 거대한 마그마와 같아서 쉽사리 파악할 수가 없었다.

서구에서 정기적인 인구조사를 비롯한 여러 근대적인 통계제도가 마련된 것이 바로 그 무렵, 즉 19세기 초부터였다. 이후 각국에서는 통계 업무를 위한 조직을 정부 내에 만들어 점차적으로 통계의 종류도 늘이고 전문적인 통계 조사 과정도 개발하기 시작하였다. 그리하여 인구나 그 구성의 변화뿐만 아니라 빈곤, 범죄, 자살, 교육, 공중 보건, 노동자들의 생활상 등이 전까지 전혀 작성되지 않거나 주목받지 못한 사회의 여러 모습들이 점점 체계적으로 조사되고 연구되기 시작하였다. 서구에서 이처럼 통계가 크게 발달하기 시작한 시기는 일본이나 식민지 조선에서 ‘국세조사’라는 이름으로 정기적인 인구조사가 실시되기 약 1세기 전에 해당한다.

이런 과정에서 생산되고 누적된 통계는 마치 자연과학 연구자들이 실험과 관찰을 통해 자연법칙을 찾듯 사회연구자들이 사회의 법칙을 찾는 유력한 방법으로 부상하였다. 그 결과 19세기 중반 이후 사람들은 통계를 이용함으로써 비로소 국가와 사회에 대한 과학적인 연구,

즉 ‘사회과학’이 가능해졌다고 생각하였다. 뿐만 아니라 통계는 19세기 말 이후부터는 수학자들 덕분에 조사 중심의 사회통계를 넘어 불확실한 상황에서 추론, 예측하는데 탁월한 능력을 발휘하는 정밀한 학문으로 성장하게 되었다. 또한 20세기 후반 컴퓨터의 발달 역시 통계가 비약적으로 성장하는데 큰 역할을 하였다.

오늘날 통계는 과거와 현재를 파악하고 미래를 설계하기 위한 가장 유력한 자료로서, 또한 갈등을 조정할 수 있는 가장 설득력 있는 도구로서 자리 잡게 되었다. 한편 복지국가로 통칭 하였던 20세기 국가는 그 역할이 이전과 비교할 수 없을 만큼 확대되었고 그만큼 국가 정책 역시 더욱 정밀해져야 했다. 이에 따라 어느 나라에서든 정부 부처에서 정책을 수립하고 수립한 정책을 평가할 때 적절한 통계를 얻어 이를 세밀하게 활용하는 것이 점점 중요해졌다. 여기서 최근 우리나라에서 논란이 된 국책사업 두 가지를 예로 들어보자.



## 1.2 두 가지 정책 사례

우리나라 정부가 수행한 대규모 국책사업 중 오랜 기간 동안 가장 많은 논란과 갈등을 불러일으켰던 사업으로 새만금간척종합개발사업을 들 수 있을 것이다. 이 사업은 전라북도 군산, 김제, 부안 앞바다를 연결하는 방조제 33km를 세우고, 그 안에 땅 28,300 ha, 호수 11,800 ha를 만드는 사업이다. 이 사업은 1991년에 본격적으로 시작된 이래 환경생태계의 파괴를 염려하는 수많은 사람들의 반대에 부딪혀 여러 차례 공사가 중단되고 법정다툼을 벌이는 등 갖은 우여곡절을 겪어야 했다. 2007년 대법원의 최종 판결로 가까스로 고비를 넘기는 했지만 이로 인한 그동안의 국력소모는 말할 수 없이 컸다고 할 것이다.

한편 제17대 대통령선거에서 당시 이명박 후보의 핵심 공약은 한반도대운하사업이었다. 선거기간 내내 이 사안은 공약수준임에도 불구하고 이 후보 지지층의 많은 찬성 논리에 대하여 반대론자들은 환경, 건설기간, 비용, 실효성, 필요성 등을 두고 많은 이의를 제기하였다. 2007년 12월에 대통령으로 당선되어 이명박 정부가 출범하고 나서도 여전히 뜨거운 쟁점 사항이 되고 있다.

대형 토목사업이면서 환경에 관한 찬반 논란이 많았다는 점에서 새만금사업과 대운하사업

은 서로 유사한 면이 많다. 그렇다면 왜 새만금사업은 정책 수립 후에 논란이 발생하였고, 대운하사업은 정책 수립 단계에서부터 쟁점 사항이 되었을까?

가장 큰 요인으로는 그동안 우리나라에서 있었던 정치, 사회 환경의 변화를 들 수 있을 것이다. 새만금사업이 시작될 당시만 하더라도 정책과정에서 합리적이고 객관적인 증거들을 제시하여 시민들을 충분히 납득시키는 과정 없이도 정책을 수립하여 시행할 수 있었다. 그러나 이후 약 20년 동안 민주화가 크게 진전되고 시민들의 환경보호의식도 높아진 결과, 정책 당국이 충분한 증거에 의해 시민들을 납득시키지 않은 채 일방적으로 사업을 추진할 수 없게 된 것이다.

이 사례에서 볼 수 있듯 사회가 민주화, 선진화되어가면서 합리적인 정책과정의 중요성은 더욱 강조되고 있다. 서로 이해가 상충하는 다양한 구성원들의 의견을 조정해가면서 국가적, 사회적 발전을 도모할 수 있는 정책을 수립하고 추진해 나가기 위해서는 구성원들을 설득시킬 수 있는 객관적이고 합리적인 증거를 제시하는 것이 필요하다. 오늘날 세계적으로 정책분야에서 증거기반 정책(evidence-based policy)이 중요한 화두로 부상하는 것도 그 결과일 것이다. 증거기반 정책이란 “정책의 개발이나 수행의 본질에 관한 유용한 증거를 활용하여 정책이나 프로그램, 프로젝트 등을 위한 의사결정을 내리도록 하는 방법”이라고 정의된다. 1999년 영국 Cabinet Office가 발간한 『Modernizing Government』라는 책에서 정부의 정책들은 증거에 기반을 두어야 한다고 언급한 것을 기화로 이 개념은 최근 여러 나라로 빠르게 전파되고 있으며, 2005년부터는 『Evidence & Policy』라는 학술저널도 발간되기 시작하였다.

우리나라에서도 지식 정보화 사회로의 급속한 이행이라는 변화와 맞물려 정책분야에서 통계에 대한 새로운 수요가 과거 어느 때보다 많이 나타나고 있다. 우리는 의식적으로든 무의식적으로든 최적의 의사결정을 내리는 데 통계의 영향을 받고 있으며 행정업무를 수행하면서도 여러 형태의 통계를 이용하고 있다. 이는 통계가 단순히 사실 확인을 위한 도구에 머물지 않고 사실이나 현상이 어떠한 원인에 의한 것인지 규명하고 이러한 원인이 일반화, 규칙화될 수 있는지 밝히는 역할을 수행하기 때문이다. 따라서 통계는 개인, 기업, 연구기관, 국가 등이 합리적이며 효율적인 의사결정을 하는 데 필수불가결한 역할을 맡고 있다. 공공부문에서도 정책의 수립, 집행, 사후관리 등 증거를 기반으로 하는 정책관리가 강조되면서 가장 직접적이고 확실한 기반인 통계를 필요로 하게 되었다.

이러한 추세에 발맞추어 2007년 통계법과 시행령을 개정한 바 있는데, 개정된 통계법 시행령에는 “정책과 제도의 집행·평가에 적합한 통계의 구비 여부 등에 대한 평가”(제33조와 34조) 관련 조항이 수록되어 있다. 이에 따라 통계청에서도 통계기반 정책관리제도를 도입하고 국가통계 통합정보시스템 구축사업을 대대적으로 추진함으로써 정책을 위한 통계기반 구

측에 많은 노력을 기울이고 있다. 그 결과 ‘e-나라지표’와 같은 사이트가 구축되었으며, 정부의 각 부처마다 정책에 필요한 통계정보들을 체계적으로 관리, 운영하는 모습이 점점 보편화되고 있다.

그러나 여전히 정책당국자는 일반인이든 통계를 매우 어렵고 복잡한 것으로 인식하고 혼란스럽게 사용하고 있는 것도 사실이다. 어떤 사람은 통계로 표현된 수치를 마치 사실 그 자체인 것처럼 오해하여 과신하기도 하고, 또 어떤 사람은 통계를 귀에 걸면 귀걸이 코에 걸면 코걸이 식의 장식품 내지는 속임수로 여기기도 한다. 또한 통계를 작성하는 생산자들 역시 신뢰하기 어려운 자료들을 바탕으로 부실한 통계를 만들어 공표하는 경우도 있다. 우리나라에서 대표적인 통계는 국가에서 인증한 승인통계(이하 국가통계)이다. 국가통계는 국가에서 승인하는 절차를 거치므로 정확성, 타당성을 확보한 객관적인 정보라고 생각되지만 실상은 그렇지 못한 경우도 적지 않다. 실제 2007년 6월 감사원에서 발표한 「국가 주요통계 작성 및 활용실태」 감사결과 보고서에 따르면 국가승인통계들에도 문제점이 많다는 사실이 드러난 바 있다.

한편 통계 생산자가 아닌 이용자들은 어떠한 통계가 얼마나 생산되고 있는지 잘 알지 못하는데다 각각의 통계가 가지고 있는 제약이나 한계 등에 대해서도 이해가 부족한 편이므로 통계를 오용하거나 남용하기도 한다. 그들은 또한 자신이 인지하는 느낌과 실제 통계수치 사이에 차이가 생길 경우 통계를 불신해 버리는 경향을 보이기도 한다.



### 1.3 이 책의 특징과 구성

앞에서 언급했듯이 정부의 역할이 확대되고 정책이 정밀해지면서 그에 필요한 통계 또한 점점 전문화되고 복잡해질 수밖에 없다. 그렇지만 정책 담당자나 이용자들의 통계에 대한 이해를 높여주는 동시에 정책과정에 필요한 통계를 작성하고 활용하도록 해주는 체계적인 지침은 여전히 부족한 실정이다. 그 결과 통계의 중요성을 알면서도 그것을 제대로 이해하고 활용하기란 쉽지 않은 일이 되었다. 통계는 주변에 넘쳐나지만 당면 과제에 정말 적절한 통계를 어디서 찾아 어떻게 이용할지 알지 못하는 경우가 적지 않은 것이다. 이 책에서는 통계에

대한 사전 지식이 없는 공무원들을 대상으로 통계와 정책을 설명하려고 하였다. 통계를 이해하는 것은 복잡하고 어려운 수학 문제를 푸는 것과는 다르다. 이 책은 기존의 많은 통계학 교재와는 전혀 다른 내용을 담고 있으며 정책을 소개하는 기존의 행정학 교재와도 역시 다른 모습이다. 이 책의 구성은 다음과 같다.

먼저 2장에서 우리는 정책과정의 단계들을 체계적으로 알아보고 과정별로 통계가 어떻게 활용될 수 있는지 살펴볼 것이다. 아울러 지난 역사 속에서 통계가 의료 정책과 빈곤 정책을 마련하는데 어떤 역할을 하였는지 나이팅게일과 찰스 부스를 사례로 들어 살펴본 다음, 가장 대표적 통계인 인구주택총조사에 대해 알아볼 것이다. 이어서 통계 부실이 불러일으키는 정책상의 문제점들을 구체적인 사례를 중심으로 검토하고 정책의 기반이 되는 올바른 통계를 작성하기 위해 고려해야 할 기본 사항들도 살펴볼 것이다.

이어지는 3장에서는 몇 가지 실제 사례들을 구체적으로 분석하여 바람직한 정책기반 통계를 구축하기 위해 필요한 사항들을 짚어볼 것이다. 먼저 통계가 미비하거나 부실하여 정책과정에서 문제가 생긴 사례들을 돌아볼 것인데 우리나라 부동산 정책 분야의 주택가격지수, 교육 정책 분야의 사교육비실태조사, 보건 정책 분야의 손상감시체계 구축 등의 사례를 알아볼 것이다. 또한 정부 사업들에서 실제 성과목표 및 성과지표를 설정한 사례들을 통해 바람직한 성과지표가 지녀야 할 특성들을 검토할 것이다. 또한 3장에서는 가상적인 예산 기획 예제를 통해 간단한 통계를 작성해 보고 그 유용성을 경험해 볼 것이다. 또한 국가통계포털에서 제공되는 통계 데이터베이스를 활용하여 개인적으로 지표를 직접 만들어 보기도 할 것이다. 마지막으로 통계청 통계심사과에서 다룬 사례를 통하여 정책의 수립, 집행, 평가에 필요한 통계가 존재하는지, 만일 없다면 개발계획을 세웠는지 검토하는 과정도 알아볼 것이다.

이어서 4장에서는 통계표를 읽을 때 흔히 생기는 오류나 해석할 때 생기는 오류를 실제 사례를 통하여 살펴보면서 통계적인 사고방식이란 어떤 것인지 공부할 것이다. 또한 수집된 자료를 통계적으로 표현하는 기법과 통계지표를 가공, 활용할 때 필요한 사항들도 알아볼 것이다.

마지막 5장에서는 통계학의 기본이 되는 정규분포와 신뢰수준에 대해 살펴보고 어떤 것들을 비교하거나 서로 간의 관계를 알아보기 위한 통계적 방법들도 알아볼 것이다. 또한 과거의 자료를 바탕으로 미래를 예측하는 방법에 대해서도 살펴볼 것이다. 통계적 방법을 공부하는 이 장에서 우리는 각 방법들의 이론을 공부하기보다는 사례를 중심으로 각 분석법의 용도와 목적을 쉽고 간략하게 알아볼 것이다. 마지막 부록에는 통계청 통계심사과에서 담당하는 통계기반 정책관리제도에 대한 자료가 첨부되어 있다.

대략적인 내용 소개에서 알 수 있듯이 이 책은 통계학 이론을 공부하거나 자료분석 방법을

자세히 공부하는 책이 아니다. 또한 정책과정만을 소개하는 행정학 책도 아니다. 이 책의 목적은 독자가 세부적인 자료분석 방법보다는 정책과정에서 통계의 폭넓은 활용 가능성을 인식하고 통계에 한 걸음 다가설 수 있게 돕는 것이다. 자료를 기반으로 정책 문제를 파악하고 해결책을 모색하는 통계적 사고, 이것이 이 책의 지향점이다. 끝으로 이 책에 나오는 정부 기관명이나 기타 고유명사는 2008년 이전에 사용된 명칭을 그대로 사용하였음을 유의하기 바란다.

## ■ ■ 요약

\_ 통계는 과거와 현재를 파악하고 미래를 설계하기 위한 가장 유력한 자료이자 갈등을 조정할 수 있는 가장 설득력 있는 도구이다.

\_ 새만금과 대운하 사례에서 보듯 사회가 민주화, 선진화되어 감에 따라 정책의 수립과 평가에서 객관적인 증거에 바탕을 둔 정책관리의 필요성이 매우 높아지고 있다.

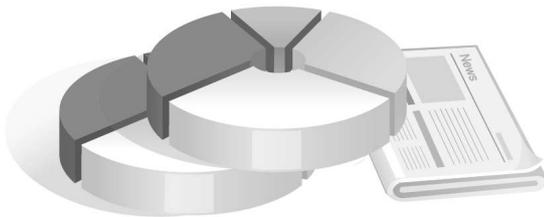
\_ 이 책의 목적은 독자가 세부적인 자료분석 방법보다는 정책과정에서 통계의 폭넓은 활용 가능성을 인식하고 통계에 한 걸음 다가설 수 있게 돕는 것이다. 이 책의 지향점은 자료를 기반으로 정책 문제를 파악하고 해결책을 모색하는 통계적 사고에 있다.



CHAPTER 2

# 정책과정과 통계

- 2.1 정책과정의 세 단계
- 2.2 정책과정에서 통계의 역할
- 2.3 통계학사 속에서 살펴본 정책과 통계
- 2.4 정책기반으로서의 통계



## 개요

정책 단계별 내용과 특성을 대체적으로 살펴본 후, 각 정책 단계별로 통계가 어떤 방법으로 활용될 수 있는지를 논의한데 이어 정책과 통계가 접목된 대표적인 역사적 사례들을 소개한다. 그 다음으로는 우리나라의 국가승인통계에 대해 소개하는 한편, 통계가 정책기반으로 활용되는 예와 통계의 미비나 부실로 인해 생기는 문제점들을 간단하게 살펴본다. 마지막으로 정책기반으로서 활용될 통계의 작성을 위해 고려해야 할 사항들을 정리한다.

## 주요내용

### 2.1 정책과정의 제 단계

정책의 입안, 집행, 평가 각 단계별 특성을 소개한다.

#### 2.1.1 정책입안 단계

#### 2.1.2 정책집행 단계

#### 2.1.3 정책평가 단계

### 2.2 정책과정에서 통계의 역할

정책의 입안, 집행, 평가 각 단계에서 통계의 역할을 살펴본다.

#### 2.2.1 정책입안 단계에서 통계의 역할

#### 2.2.2 정책집행 단계에서 통계의 역할

#### 2.2.3 정책평가 단계에서 통계의 역할

### 2.3 통계학사 속에서 살펴본 정책과 통계

의료정책에서 나이팅게일, 빈곤정책에서 부스의 사례 등을 소개한다.

#### 2.3.1 등불 대신 통계로 병원을 밝힌 여인

#### 2.3.2 역사 속의 정책과 통계

#### 2.3.3 영국의 빈곤정책과 부스의 통계조사

#### 2.3.4 통계에 대한 저항과 통계의 한계

### 2.4 정책기반으로서의 통계

#### 2.4.1 대표적 정책기반 통계 사례 : 인구주택총조사

#### 2.4.2 통계 부실로 인한 정책상 문제점

(1) 통계의 미비로 인한 문제 : 보건복지부의 「의료급여 수급권자를 대상으로 한 본인부담금 제도」 사례

(2) 통계의 미활용으로 인한 문제 : 건설교통부의 「국도확장사업 수요 재검증 사업」 사례

(3) 통계의 오용으로 인한 문제 : 농림부의 「농축임업인의 영유아에 대한 양육비 지원 사업」 사례

#### 2.4.3 정확하고 타당한 통계를 활용하지 못하는 원인 분석

(1) 기관 간 행정자료의 공유 미흡 : 경찰청의 「도로 교통 정책」 사례

(2) 부적합한 조사의 범위와 대상 : 과학기술부의 「여성과학기술인의 육성 및 지원에 관한 기본계획 수립」 사례

(3) 대표성이 결여된 표본조사 : 산업자원부의 「산업인력 양성 관련 정책」 사례

(4) 전체집계에서 일부 유형의 통계 누락 : 재정경제부의 「해외직접투자통계」 사례

(5) 동일 대상 통계를 부처별로 다른 기준에서 작성 발표 : 산업자원부와 건설교통부의 「산업입지 정책수립」 사례

#### 2.4.4 정책기반 통계를 작성할 때 고려해야 할 사항



## 2.1 정책과정의 제 단계

정책의 정의에 대해서는 입장과 관점에 따라 학자들마다 다양한 의견을 제시하고 있지만 정책이란 대체적으로 “정부기관이 공공 목표의 달성을 위해 마련한 행동방안 또는 지침”이라고 할 수 있겠다. 공공 목표의 달성이라는 정책의 목적에는 사회 가치, 인간 가치, 미래 가치, 그리고 구체적인 행동과 다양한 사회의 관계에서 나타나는 불확실성과 대표성이 모두 들어있다. 정책과정은 정책의 입안, 집행, 평가 등으로 구분될 수 있으며, 사회 문제를 해결하기 위한 방안을 제시하기 위한 이론적 접근과 실증적 접근이라는 두 축을 중심으로 진행된다. 실증적 접근의 중요한 도구 가운데 하나인 통계와 통계적 기법을 정책과정과 연계시켜 살펴보는 의의도 여기에 있다.

적절한 정책을 수립하고 집행하며 그 결과를 평가하는 것은 효율적인 국정운영을 위해 필수적이며 매우 중요하다. 그런데 적절한 정책의 수립은 사회현실에 대한 정확한 정보의 토대 위에서 가능하며 그 정책의 성과를 평가하기 위해서는 계량화된 성과지표를 활용해야 한다. 또한 성과평가를 통해서 얻어진 정책의 효과와 부작용을 새로운 정책수립에 환류시킴으로써 정책의 품질을 높이고 국정운영의 효율을 향상시킬 수 있다.

### 2.1.1 정책입안 단계

많은 연구에서 볼 수 있듯이 초기의 정책 연구는 사회 문제에 대해 어떤 정책을 수립해야 하는가보다는 정책의 입안, 집행, 평가에 주로 관심을 두었다. 하지만 오늘날은 정책수립에 앞서 사회적 이슈가 무엇인지를 미리 파악하여 이를 정책 대상으로 규정하는 과정을 중요시하는 경향을 보이고 있다. 따라서 정책 대상(또는 정책 의제; policy agenda)이 되는 사회적 이슈의 채택 과정과 정책 대상으로 규정하는 과정까지도 정책입안 단계에 포함한다.

사회적 쟁점이 공론화되려면 우선 그 문제가 많은 사람들의 관심을 끌어 대다수의 사람들

에 의해 정책 당국의 조치가 필요하다고 생각되어야 한다. 그런데 대부분의 경우 정부가 사회문제를 인지하게 되는 것은 사회집단의 정치적 압력과 요구보다는 '사회지표(social indicator)'가 보여주는 사회의 상태 또는 조건 때문이다. 대표적인 지표로 각종 경제지표, 범죄율, 특정 질병의 발생 빈도 등이 있는데 최근 우리 사회의 인구감소와 고령화 추세도 그 좋은 예로 볼 수 있다.

### (1) 사회 문제의 특징

정책학 분야에서는 정책 문제가 될 수 있는 사회적인 문제의 일반적 특성을 대체로 다음과 같이 정리한다.

① **상호의존성과 복잡성** : 한 영역에서의 정책문제(보건 복지, 노동)는 종종 다른 영역의 정책문제(경제, 산업)에 영향을 미치므로 현실적으로 정책문제들은 하나의 독립적인 실체가 될 수 없다. 따라서 문제에 개입된 요소나 변수가 아주 많아서 쉽게 그 구성 법칙을 알아내기가 어렵고 또한 그 변수들의 관계가 시간과 상황에 따라 가변적이며 서로 얽혀 상호작용하고 있는 경우가 많다.

② **주관성과 모호성** : 비록 문제 자체는 객관적인 것일지라도 그 문제를 해석하는 방법은 주관적이다. 또한 많은 사람들의 이해가 교차되기 때문에 정책의 목표가 모호하게 되며, 정책을 정의하는 개념 또한 포괄적이거나 사용하는 개념이 사람에 따라 달라진다.

③ **불확실성** : 정책 형성은 인과관계에 대해 충분히 알지 못하여 앞으로 나타날 결과를 예측하지 못하는 불확실한 상태에서 시작하게 된다.

### (2) 사회문제의 인식

사회문제나 공공의 요구가 공공의 관심을 끌게 된 결과, 정책전문가들에 의하여 이슈로 제기되고 정책 의제로 채택되어 정책으로 형성되는 과정은 다음과 같은 특성을 띤다.

① **정책 의제 간의 경쟁** : 정책 의제로 채택되기까지 여러 사회문제나 이슈 간의 경쟁은 피할 수 없다. 사회문제로 제기된 것이 모두 정책 의제로 채택되지는 않는다. 정책결정이 될 때까지는 지속적인 경쟁 상황에 놓이게 된다.

② **정책의제과정의 정치성** : 정부의 공식 의제로 채택되어도 사회의 여러 문제가 끊임없이

제기되어 관련 이익집단들과 정책참여자들에 의해서 지속적으로 영향을 받는다.

③ **사회문제의 재정의를** : 사회문제가 정책 의제로 채택되는 과정에서 사회의 여러 관련 세력들은 각자에게 유리한 입장에서 사회문제를 정의하게 되는데 초기에 제기된 이슈가 변질되거나 도태되기도 한다.

### (3) 정책 의제 설정 모형

학자들은 정책 의제가 설정되는 데는 일반적으로 다음과 같은 세 가지 모형이 있다고 주장한다. 다음의 모형들은 그 사회가 처해있는 상황이나 문화적, 사회 구조적 특성이나 정치구조에 따라 설명력에 있어서 차이가 있을 수 있다.

① **외부주도모형(Outside Initiative Model)** : 외부주도모형은 정부 밖의 개인이나 집단에 의하여 사회문제가 제기되고 일반대중에게 확산되어 정부의 정책 의제로 채택되는 경우이다. 다원론과 이익집단이론이 그 이론적 배경이라고 볼 수 있다.

② **동원모형(Mobilization Model)** : 정부기관 내에서 사회문제를 인식하고 대중에게로 확산시키는 경우로서 엘리트이론에 그 근거를 두고 있다.

③ **내부접근모형(Inside Access Model)** : 정부기관 내에서 사회문제가 제기되지만 대중에게로 확산되지는 않고 내부에서 처리하는 경우를 말한다.

### (4) 목표설정의 역할

정책의 목표를 정확하고 구체적으로 설정하게 되면 다음과 같은 여러 면에서 유익하다.

① **정책 결정의 기준** : 정책 목표는 여러 가지 정책 대안 중에서 최선의 것을 선택하는 기준으로 이용된다.

② **정책 집행의 기준** : 정책 입안을 할 때 집행 과정에 필요한 정책 대안의 내용이나 기타 필요한 활동에 대한 지침을 모두 결정할 수는 없지만 정책 집행 과정에서의 결정이나 활동에 길잡이 역할을 한다.

③ **정책 평가의 기준** : 정책 평가는 정책이 바람직한 것인지 사후적으로 평가하는 것이기 때문에 구체적이고 정확하게 설정된 정책 목표는 정책 평가의 중요한 기준 중의 하나가 된다.

## (5) 대안의 탐색 및 비교 평가

대안형성과정은 사회문제를 파악하고, 정책목표를 설정하며, 수단을 확보하고 비교분석하는 과정을 말한다. 사회문제를 파악할 때에는 다음 다섯 가지 측면을 고려해야 한다. 첫째, 사회문제가 발생한 원인을 규명한다. 둘째, 사회문제가 미치는 직·간접적 영향을 파악한다. 셋째, 얼마나 많은 사람들이 사회문제를 인지하고 있는지를 파악한다. 넷째, 사회문제를 해결할 수 있는지 파악하고 비용과 자원조달방법을 파악한다. 마지막으로 사회문제가 어느 정도로 정치·사회적 중요성을 띠고 있는지도 살펴야 한다.

① **대안의 개발** : 정책목표가 뚜렷이 밝혀진 후, 정책 대안을 만들기 위해서는 문제, 상황, 미래에 대한 정보를 수집하여 종합적으로 검토 분석해야 한다. 정책 대안을 개발하는 방법에는 그 문제와 관련된 과거의 정책이나 현존 정책 및 외국이나 다른 지역의 정책 등을 검토하는 방법이 있다. 또 사회과학적 지식이나 이론적 모형으로부터 대안을 추론해내는 방법이 있다. 이 이외에도 대안개발자의 주관적 판단 하에 직관을 통해서 정책대안을 개발하는 방법도 가능하다.

② **대안의 결과 예측** : 정책 대안들중에서 최선의 대안을 선택하기 위해서는 정책 대안들 각각이 집행 또는 실현되었을 경우에 나타날 결과들을 예측하여야 한다. 예측이란 정책문제에 관한 사전 정보를 기초로 미래의 바람직한 상태에 관한 실제적인 정보를 산출한 것이다. 정책의 목표를 수립하기 위해서 미래를 예측하는 기법으로는 유추, 경향성 분석, 마르코프 모형, 회귀분석, 델파이 기법, 경로분석 등이 있다.

③ **대안의 비교 평가** : 정책 대안들 간의 우선순위를 판단할 수 있는 기준에 따라 각 정책 대안들을 평가하고 최선의 정책 대안을 선택한다. 기술적 실현가능성은 기술적 문제점은 없는지를 검토할 뿐만 아니라 집행의 가능성도 함께 살펴보아야 한다. 대안의 비교분석기법으로는 비용편익분석, 모의실험기법, 결정분석, 선형계획기법 등이 있다.

## (6) 대안형성과정에서의 문제점

① **예측능력의 한계** : 예측능력을 제고시키기 위해서는 행정정보관리체계를 구축할 필요가 있으며, 공통의 데이터베이스를 구축하여 자료교환 및 협조가 요구된다.

② **계량화의 문제** : 대안개발에 있어서 계량화 문제는 정책분야에 따라서 그 양상이 다르게 나타나지만 사람을 대상으로 질적인 측면의 효과를 분석하는 데는 더욱 어려움이 많다. 정책

목표를 보다 명백히 규명할 필요가 있으며, 측정에 필요한 자료 및 척도의 개발도 요구된다.

③ **비교 기준의 공통척도 결핍** : 대안비교에 있어서 공동 척도가 없는 경우가 많은데 가치기준들이 다양화되기 때문으로 볼 수 있으며, 주관적 판단이 작용하는 사회적 가치들이 충돌하게 되면 더욱 어려워진다.

④ **비용과 시간의 문제** : 시간과 비용이 한정되어 있는 상황에서 정책 대안을 탐색하고 개발할 경우, 비용 대비 효과를 고려해야 한다.

### (7) 정책결정에 영향을 미치는 요인

정책결정에 영향을 미치는 요인들은 다양하지만, 주로 정책결정자의 정책적 성향, 정책대안의 존재여부, 다른 정책과의 연관성, 정치·경제·사회적 상황 등을 주요 요인으로 제시할 수 있다.

① **정책결정자의 성향** : 정책결정자의 정책의지 정도는 정책 형성과 결정에 큰 영향을 미친다. 정책결정자의 공식적 권한뿐만 아니라 참모들의 정책성향 및 능력, 의지 정도도 요인에 포함된다.

② **정책대안의 준비정도** : 정책대안이 준비되어 있는 경우에는 상대적으로 결정이 쉽게 이루어질 수 있지만 그렇지 않다면 대안의 탐색과 개발이 선행되어야 하는데 준비정도가 미비한 경우 적기에 정책결정이 이루어지기는 쉽지 않다.

③ **다른 정책과의 연관성** : 정책결정에 있어서 또 한 가지 고려되어야 하는 것이 다른 정책과의 관계성이다. 예컨대, 사회복지정책과 국방정책 사이에는 상쇄관계가 있을 수 있다는 주장이 있는데 국방예산이 증가하면 사회복지 예산이 감소하는 경향이 있다는 것이다.

④ **정치·경제·사회적 상황** : 정책을 산출하는 데는 정치체계가 정책결정에 영향을 미칠 뿐만 아니라 경제적, 사회적 상황도 함께 고려해야 한다.

### (8) 정책참여자와 정책결정에 관한 이론

정책에 있어서 참여의 의미는 국민들이 선호하는 이익이나 가치를 정책과정에 반영하기 위

하여 직·간접적으로 영향력을 행사하거나 관여, 개입하는 과정을 말한다. 정책결정에 있어서 참여자들은 법 제도적으로 정책과정에 참여할 수 있는 권한을 가진 공식적 참여자들(의회, 대통령, 사법기관, 행정기관)과 비공식참여자(정당, 이익집단, 일반 국민(선거, 여론), 대중 매체, 전문가)로 구분할 수 있다.

정책결정에 관한 이론적 모형들은 정책결정이 어떻게 이뤄지는가를 설명하기 위해서 제시된 모델로서 합리모형, 만족모형, 점증모형, 혼합주사모형, 최적모형, 쓰레기통모형 등이 있다.

## 2.1.2 정책집행 단계

정책집행은 결정된 정책을 구체화시켜 나가는 과정으로써 정책목표에 부합되도록 구체적인 활동을 제시하는 정책의 실현 단계로 볼 수 있다. 구체적 활동은 정책수단을 말하는데 목표의 달성에 관한 활동들이 단순히 이루어지는 것은 아니며 목표가 반드시 달성되는 것도 아니다. 이 과정에서도 여러 참여자들 간의 상호작용(협상과 타협)을 통해서 원래의 목표가 왜곡되기도 하고 시간이 지연되는 경우도 빈번히 일어난다.

### (1) 정책집행의 이론적 모형

정책집행과 관련된 이론적 모형을 여기서는 종합모형으로 <표 2-1>에 제시하였다. 먼저 정책적 변수와 비정책적 변수로 구분한 하향적 집행요인을 설정하였다. 정책적 변수는 정책목표의 타당성과 구체성, 정책에 소요되는 예산의 충분성, 정책 집행권한 및 책임소재의 명백한 규정정도, 서비스 전달체계에 대한 조정 및 감독, 서비스 전달 방법의 명백한 규정 등을 세부요인으로 들 수 있다. 비정책적 변수로는 최고의사결정권자의 충분한 관심과 지지를 포함시켰다.

상향적 집행요인으로는 일선공무원의 전문성 및 재량권, 상급기관과 경쟁기관과의 협력, 클라이언트의 적극적인 참여, 정책 관련 단체와 주민과의 협력 등을 포함시켰다. 또 다른 학자들은 전술한 5가지의 세부요인 이외에도 일선 집행 공무원의 사고방식, 지방정부와의 상호관계 등도 정책집행에 영향을 미친다고 주장하였다.

〈표 2-1〉 정책집행에 있어서 종합모형

분석요인	분석내용	
하향적 집행요인	정책적 변수	- 정책 목표의 타당성과 구체성
		- 정책에 소요되는 예산의 충분성
		- 정책 집행 권한 및 책임소재의 명백한 규정
		- 서비스 전달 체계에 대한 조정 및 감독
	비정책적 변수	- 서비스 전달 방법의 명백한 규정
		- 최고의사결정권자의 관심과 지지정도
상향적 집행요인		- 일선공무원의 전문성
		- 일선공무원의 재량권
		- 상급기관과 경쟁기관과의 협력
		- 클라이언트의 적극적 참여
		- 이익집단 및 관련 단체와 주민과의 협력정도
정책 환경적 요인		- 정치적 상황의 변화
		- 경제적 상황의 변화
		- 사회적 상황의 변화

정책 환경적 요인으로는 정치적, 경제적, 사회적 상황의 변화 요인을 설정하였다. 정책 환경적 요인의 대부분이 정책에 간접적인 영향을 미치는 경우가 많고, 정책 결정자나 정책 담당자가 의도적으로 통제하기 어렵다. 그러나 간접적인 영향이라 할지라도 국가와 사회가 처한 상황에 따라 그 중요성이 부각될 수 있다.

## (2) 정책 지침의 작성

정책 지침은 집행자가 따라야 할 것과 정책의 집행에 필요한 사상 등에 대하여 규정한 것을 말한다.

## (3) 자원의 확보와 배분

정책을 실현하기 위해서 필요한 자원을 마련해야 하는데, 여기에는 물적·인적 자원 및 조직, 기구 까지도 포함하게 된다.

#### (4) 집행 점검과 평가

정책이 집행되면서 수행되는 구체적인 활동에 대한 점검과 평가 단계이다. 정책 평가의 개념보다는 정책 집행 분석 및 평가로서 집행 책임자나 그 아래 있는 집행 관리자 등에 의하여 진행되어 이들에게 실현 성과와 실현 활동에 관한 정보를 제공하는 활동이다.

### 2.1.3 정책평가 단계

정책 평가란 한마디로 정책목표를 어느 정도 달성했는지를 가늠하는 단계로서 관련 정보를 수집하여 분석하고 해석하여 그 가치를 판단한다. 협의로는 정책집행의 결과에 대한 평가로 볼 수 있으나 광의로는 정책과정 전반에 관한 평가활동을 의미하기도 한다. 평가대상은 정책집행의 얻어지는 산출(outputs), 결과(results), 성과(performance) 등으로 구분할 수 있다.

#### (1) 평가 기준 설정

정책성과에 대해서 신뢰할 수 있고 타당한 근거 즉, 정보의 제공이 매우 중요하다. 구체적인 평가 기준은 다음과 같다.

- ① **효과성** : 정책목표의 달성도를 말하는데 실제 정책평가에 있어서는 효과성을 판단하기 어려운 경우가 많다. 정책목표를 명백히 규정해야 하는 전제조건이 필요하다.
- ② **능률성** : 투입과 산출의 비율을 의미한다.
- ③ **형평성** : 사회 계층 간의 불평등이 어느 정도 시정되었는가를 평가하는 기준을 의미한다.
- ④ **반응성** : 정책집행의 결과가 어느 정도 정책대상 집단의 요구, 선호, 가치 등을 반영시켰는지 또는 만족시켰는지를 판단하기 위한 평가기준이다.
- ⑤ **적정성** : 정책이 사회문제의 해결과 어느 정도 관련성이 있는지를 살펴보는 기준으로써 정책목표와 수단 간의 관계가 타당한 것인지를 평가하는 기준이다.

## (2) 평가 방식(누가, 언제, 무엇을)

- ① 평가자가 내부인가, 외부인가?
- ② 평가 시기가 사전인가, 중간인가, 사후인가?
- ③ 평가 대상이 과정인가, 결과의 영향인가?

## (3) 정책평가의 유형

정책평가에 대한 사회적 관심이 고조되는 70년대 중반이후 다양한 평가방법이 등장하게 되었다. 일반적으로 평가의 유형은 과정평가, 영향평가, 종합평가 등이 있으나 평가의 관점, 목적, 시기, 대상 등에 따라서 다양하다.

## (4) 정책평가의 절차와 접근방법

정책을 평가하는 방법이 다양하듯이 절차도 다양하다. 전술한 바와 같이 평가주체에 따라 서로 다양하게 전개될 수 있다. 일반적으로 평가절차는 정책의 특성, 평가담당자와 대상기관, 평가의뢰기관 등에 따라 다양하나, 정책관련 이해집단의 파악, 평가목적의 결정, 평가성 사전, 평가에 대한 행정적 합의 도출, 구체적 프로그램 평가, 평가결과의 활용 등의 절차를 거치게 된다.

## (5) 정책평가 접근방법과 평가 설계

과학적 접근방법에 의한 평가는 계량적인 특성과 사실관계의 추론을 수반하게 된다. 단순 인과관계를 규명하는 모형으로는 결정론적 모형, 확률적 모형, 효과모형 등이 있다. 평가연구 설계단계에서는 우선 목표의 구체화, 효과모형의 구성, 평가에 적합한 연구 설계, 측정과 표준화, 자료수집, 자료분석과 해석 등이 포함된다.

평가연구방법은 비실험적 접근과 실험적 접근으로 구분되며, 비실험적 연구로는 단일사례 분석, 단일집단 사전·사후 측정연구설계, 정적 집단비교설계 등을 예로 들 수 있다.

## (6) 성과지표 설정

정부의 주요사업 운영 성과를 관리하기 위해 성과지표를 설정할 때가 있다. 이 경우 적절한 통계의 마련은 필수적인 전제조건이라 할 수 있다. 성과지표는 성과목표의 달성 여부를 알아보기 위한 것으로 계량화뿐만 아니라 정성적 부분까지 포괄할 수 있도록 구체적이고 합

리적으로 개발되어야 한다. 성과지표 설정에 있어서 반드시 갖추어야 할 미시적·거시적 특성에 대해서 정리한 것이 다음의 <표 2-2>에 나와 있다.

〈표 2-2〉 성과지표의 미시적·거시적 특성

성과지표의 특성		내 용		
거시적 요인	종합성	<ul style="list-style-type: none"> <li>○ 조직의 사명, 전략목표, 성과목표 및 성과지표, 자료출처 등 종합적인 자료 포함</li> <li>○ 구체적인 사업추진계획 및 방법과 기대효과, 필요재원, 과거의 목표달성 정보에 대한 자료 포함.</li> <li>○ 이해관계자들의 의견 반영 및 외부요인 명시</li> </ul>	사업선정 및 계획 단계  사업집행 및 성과 단계	
	연계성 (인과관계)	○ 전략목표와 성과목표, 그리고 성과목표와 성과지표간의 연계성		
미시적 요인	성과목표의 명확성	○ 계량화		
		○ 시간적 범위		
		○ 이해관계자들의 의견 반영		
	성과지표의 적절성	특 성	○ 적합성	
			○ 계량성	
○ 민감성				
○ 비편향성				
지표분류	○ 투입지표		○ 효율성 지표 ○ 효과성 지표	
	○ 과정지표			
	○ 산출지표			
	○ 결과지표			

성과지표란 조직의 임무, 전략목표, 성과목표의 달성여부를 측정하는 기준으로서, 성과를 측정할 수 있도록 계량적 혹은 질적으로 나타낸 것을 의미한다. 바람직한 성과지표가 지녀야 할 필수적인 항목으로 첫째, 달성 여부를 판단할 수 있도록 전략목표와 성과목표는 구체적이고 명백하게 세워져야 한다. 전략목표는 해당 기관의 핵심 사업을 포괄할 수 있도록 설정되어야 하며, 중복되지 않아야 한다. 성과목표는 달성여부를 정확히 확인할 수 있도록 구체적이어야 하며, 계량적으로 표현되어야하고 시간적 범위가 명확해야 한다. 그리고 전략목표와 성과목표를 설정하는데 있어서 이해관계자들의 의견을 수렴하여 반영하는 노력이 필요하다.

둘째, 성과지표는 성과목표를 달성하는 과정에서 그 진척정도(적합성)를 나타낼 수 있어야 하고, 성과목표를 전반적으로 반영할 수 있어야 하며, 주관적인 의견이 개입될 수 없도록 계량화로 표현되어야 하고, 외부 상황의 변화(민감성)에 따라 큰 영향을 받지 않아야 한다. 또한 성과지표는 사업성과의 질, 산출, 적시성, 비용, 결과 등의 여러 면을 종합적으로 평가할 수 있도록 설정되어야 하기 때문에 해당기관의 사업집행이 의도되지 않은 방향으로 유도(비편향성)되어서는 안 된다.



## 2.2 정책과정에서 통계의 역할

효율적인 국정운영을 위해서는 공공부문에서 정책을 수립하여 집행하고 사후관리를 할 때 합리적이고 객관적인 증거에 기반을 두는 이른바 증거기반 정책관리가 필수적이다. 정책관리를 위한 증거 중 대표적인 것으로 통계를 들 수 있다. 통계는 정책관리자에게 증거를 기반으로 정책을 수행할 수 있도록 근거를 제공해준다. 통계를 활용하면 현황분석과 정책효과에 대한 성과분석을 수행할 수 있다. OECD 등 국제기구에서도 최근 증거를 기반으로 한 정책관리를 위하여 통계의 활용성을 강조하고 있다.

통계란 사회의 현 상태를 수량적으로 기록한 것으로서 우리의 일상생활에서 넓고 깊게 자리 잡고 있다. 우리는 의식적으로든 무의식적으로든 최적의 의사결정을 내리는 데 통계의 영향을 받고 있으며 행정업무를 수행하면서도 여러 형태의 통계를 이용하고 있다. 이는 통계가 단순히 사실 확인을 위한 도구에 머물지 않고 사실이나 현상이 어떠한 원인에 의한 것인지 규명하고 이러한 원인이 일반화, 규칙화될 수 있는지 밝히는 역할을 수행하기 때문이다. 따라서 통계는 개인, 기업, 연구기관, 국가 등이 합리적이며 효율적인 의사결정을 하는 데 필수불가결한 역할을 맡고 있다. 앞에서 정책의 단계를 입안, 집행, 평가로 구분하여 살펴보았는데, 여기서는 각 정책의 단계별로 통계가 어떤 역할을 할 수 있는지 검토하기로 한다.

## 2.2.1 정책입안 단계에서 통계의 역할

① **사회 문제 인지 단계** : 일반적으로 사회문제가 제기되는 상황에서 구성원들은 상반되거나 다양한 주장에 노출되므로, 이들을 설득시키고자 하는 관계 기관들의 작업이 진행된다. 이때 가장 많이 나타나는 것이 언론기관을 포함한 각 기관의 여론조사 활동일 것이다. 이 경우 여론조사는 각 기관의 입장을 강화하려는 불순한 동기에서 수행될 수도 있고 실제 여론의 위치와 향방을 찾아내려는 본래의 조사목적에 갖고 수행되는 경우도 있다. 또 다른 통계적 접근은 연관성 조사가 될 것이다. 이때에는 관련된 분야의 영향력을 측정 가능한 것으로 바꾸어서 객관적인 수치로 제시하려는 노력이 진행된다.

② **정책 목표 설정 단계** : 정책 목표를 구체화하는 과정에서 목표 설정의 합리성과 현실가능성을 보이기 위한 모의실험과 상이한 목표 간의 비교 또는 시점 간의 비교를 위한 통계적 방법이 사용된다. 또한 설정되는 목표의 평가 기준을 예상하여 각각의 기준이 적용된 후에 나타나는 결과를 분석, 해석하는 데에도 통계적 접근이 요구된다.

③ **대안의 탐색 및 비교 평가** : 입안되는 정책 대안의 집행단계에 대한 모의실험들이 수행되고, 실험결과와 대안들을 비교 분석하는데 통계적 방법이 사용된다.

## 2.2.2 정책집행 단계에서 통계의 역할

① **집행 지침 작성 단계** : 집행정도를 측정할 수 있는 척도를 개발 제시하여야 하고 집행방법을 객관적으로 비교할 수 있는 정량적 지표 개발이 필요하다.

② **자원의 확보와 배분 단계** : 정책집행에 요구되는 예산 및 기타 자원의 예측, 분배 방법을 결정하기 위한 다양한 논리와 그 논리를 정당화하는 실증적 자료 분석이 요구된다.

③ **집행 점검과 평가 단계** : 집행정도의 계량화와 진척률을 정의할 논리를 개발해야 한다. 독자적인 진행 상황을 보여주는 지표를 개발하고 다른 정책과 비교할 수 있는 지표도 개발해야 한다.

### 2.2.3 정책평가 단계에서 통계의 역할

① 평가 기준 설정 및 평가 내용 : 평가 기준의 결정과정에서 측정 가능한 내용을 확정하고, 그에 따라 자료를 수집, 분석할 방법을 논의해야 한다. 또한 다양한 기준에 따른 평가에 대응하여 분석 결과를 비교하기 위한 고급 통계적 기법이 요구된다.

② 평가 방식 : 평가에 영향을 미치는 외부요인인 평가자, 평가 시기, 평가 방법 등에 따른 평가자 일치도, 신뢰도, 평가 방법에 의한 편향, 평가 시기 결정 등을 개별적, 종합적으로 고려하는 데에도 역시 고급통계기법이 필요하다.



## 2.3 통계학사 속에서 살펴본 정책과 통계

### 2.3.1 등불 대신 통계로 병원을 밝힌 여인



〈그림 2-1〉 등불을 들고 있는 나이팅게일의 동상

여자 아이들한테 “커서 뭐가 될래?” 물어보면 “간호사 언니”라는 대답을 종종 듣게 된다. 사실 여성들이 장차 활동하는 데 제약이 많다는 사실을 가르치는데 위인전만큼 좋은 교과서도 없다. 위인전에서 볼 수 있는 인물은 동서와 고금을 가릴 것 없이 온통 남성들뿐인 것이다. 그렇게 본다면 어린 소녀들이 간호사라는 직업에 관심을 갖는 이유가 꼭 병원에서 본 간호사 언니를 닮겠다는 속셈 때문만은 아닐 테다. 거기에는 아마도 위인전에서 만난 나이팅게일이라는 ‘여성 위인’이 작지 않은 역할을 했을 것이다. 퀴리부인처럼 노벨상을 들이나 타는

과학자가 되겠다고 할 수도 없고, 식민지도 아닌 지금 “유관순 언니처럼 되겠다.”고 할 수도 없는 판에, 플로렌스 나이팅게일! 이름도 멋지다. 게다가 위인전에 실린 삽화에서처럼 등불을 밝혀 들고 부상병들을 굽어보는 나이팅게일의 모습은 대단히 인상적이다. 널리 알려졌듯 나이팅게일이 유명해진 것은 크림아전쟁 때의 활약 때문이었고 거의 모든 삽화에서 등불을 들고 있는 이유도 크림아 전쟁에서 간호사로 맹활약을 펼친 그녀를 사람들이 “The Lady with the Lamp” 라는 애칭으로 불렀던 때문이다(이를 우리나라에서는 보통 ‘광명의 천사’라고 옮겼으니 거의 성모 마리아 급으로까지 치켜 올린 셈이다).

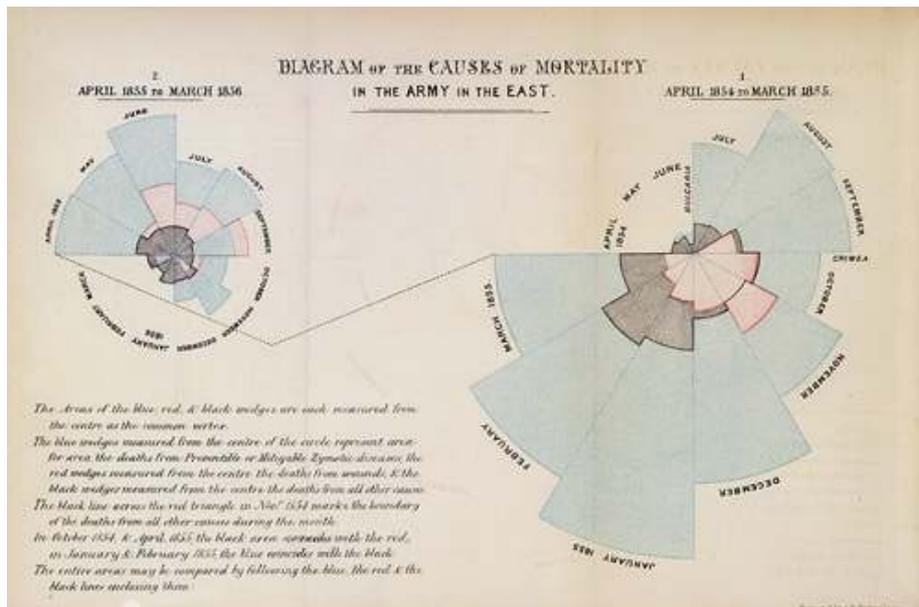
그런데 그녀가 아프고 다친 군인들이 누워있는 병원에서 크게 깨닫게 된 것은, 병사들을 돌보기 위해 열심히 등불을 들고 다니는 것만으로는 병들고 다친 이들을 구할 수 없다는 것이었다. 그 이유는 무엇보다도 군대 병원의 사망률이라는 것이 부상병의 경우를 제외하고도 일반병원에 비해 대단히 높았기 때문이었다. 그녀가 볼 때 군대의 의료시스템에는 개선해야 할 것들이 무척 많았다. 그 중에서도 제대로 된 통계가 작성되지 않고 있다는 점이 특히 문제였다. 당시 영국군의 병원들은 병의 종류나 치료 결과, 사망자 수 등에 대해 주먹구구식의 기록만 남길 뿐 제대로 된 통계라고는 전혀 작성하지 않고 있었다. 제대로 된 통계가 없으니 어디가 문제인지 알 수도 없고 개선책을 찾을 수 없는 것도 당연했다. 그리하여 크림아전쟁에서 돌아온 후 그녀는 등불을 내려놓는 대신 통계라는 체계적인 도구를 들게 되었고, 나이팅게일 덕분에 통계는 19세기 후반 영국의 의료정책을 개선하는데 큰 역할을 하게 되었다.

나이팅게일은 무엇보다도 대단히 열정적인 사람이었다. 그녀는 크림아 전쟁에서 얻은 명성에 만족하지 않고 60%에 이르는 군 병원의 높은 사망률에 주목하였다. 그녀는 그 원인을 밝히기 위해 여러 가지 다른 조건에서 군대의 사망률을 조사하고 그 결과를 민간인의 사망률과 비교하였다. 이러한 통계로부터 영국 국내에 주둔하고 있는 군대에서도 사망률은 민간인 사망률의 두 배에 달한다는 사실이 밝혀졌다. 게다가 이처럼 높은 사망률은 전쟁이 없는 시기에 나온 것이었다. 그녀는 이처럼 놀라운 결과와 함께 그 원인도 조사했는데 그 결과 많은 병사들이 집단생활을 하는 좁은 막사의 환기나 하수처리 등의 위생 상태가 매우 나쁘다는 사실이 밝혀졌다.

나이팅게일은 이처럼 통계를 이용하여 군대의 높은 사망률 실태와 그 원인을 찾아낸 데 그치지 않고 병사들의 생활 조건을 개선시킬 정책 마련을 위해서도 끈질기게 노력하였다. 언젠가 그렇듯이 정부기관과 정치권은 완고할 뿐 아니라 사고방식도 유연하지 못했으므로 그녀의 권고에도 불구하고 통계라는 방법을 써서 밝혀낸 현실을 금세 받아들이지 못했다. 당연히 통계에 바탕을 둔 새로운 정책도 빨리 입법화되어 시행되지 못했다. 이러한 상황에서 그녀는 개인적인 인맥과 재산을 동원하여 여론을 움직이고 정치인들의 생각을 바꾸어 나갔다. 평생

에 걸친 이러한 노력은 “나는 살해당한 사람들의 제단에 서있다. 내가 살아있는 한 그들을 죽게 만든 원인과 싸울 것이다.”라는 말에서도 잘 드러난다. 이처럼 남다른 열정을 지녔던 나이팅게일은 여성의 활동을 막던 시대적인 굴레를 뛰어넘었다는 점에서 비범했을 뿐 아니라, i) 직접 현장에서 문제를 발견하고, ii) 문제의 원인을 과학적인 방법으로 찾았으며, iii) 문제를 풀 수 있는 정책 대안을 마련한 뒤, iv) 그 대안을 실현하기 위해 정치적인 노력을 기울이기 까지 했다는 면에서 보기 드물게 비범한 인물이었다.

그녀의 뛰어난 능력을 보여주는 한 가지 사례로 그녀가 만든 독창적인 통계 그래프를 들 수 있다. 그녀는 숫자가 많이 나오는 통계표를 읽기 싫어하는 의회의원들을 설득하는데 이 그림을 자주 이용하였다.



(그림 2-2) 나이팅게일이 1857년에 개발한 coxcombs (또는 나이팅게일의 장미)

각 파이조각들은 월별 군인 사망수를 나타내며, 파이조각 내부의 세 가지 색깔은 서로 다른 원인으로 인한 사망수를 나타낸다. 오른쪽 그림에서 왼쪽 그림으로 갈수록 사망자수가 줄고 있음을 쉽게 알 수 있는데 그 원인은 군 병원의 위생조건 개선에 있었다. 나이팅게일은 숫자로 된 통계를 싫어하던 정치인들을 설득하는데 이 그림을 종종 이용하였다.

또한 그녀는 군대의 생활조건 뿐만 아니라 일반 병원의 상태를 개선하기 위해서도 많은 노력을 기울였다. 여기서도 물론 그녀가 실태를 파악하고 처방을 찾는 길은 무엇보다 우선 통계자료에 있었다. 당시 병원들은 서로 다른 방식으로 일관성 없는 진료 통계를 작성하고 있

있음은 물론이고 같은 런던 시내에 있는 주요 병원들끼리도 병명이나 병의 분류체계가 제각각이었기 때문에 과학적인 의학 연구와 체계적인 병원 행정이 불가능한 상태였다. 그녀는 파르(William Farr)라는 사람과 계속 교류하면서 이러한 문제를 개선하는데 힘썼다. 파르는 오늘날의 통계청에 해당하는 곳에서 일한 공무원이었는데 오랫동안 국가통계 분야에서 활발하게 활동한 인물이었다. 특히 그는 원래 의사로서 보건 통계의 전문가였기 때문에 의료 개혁을 위해 어떻게 통계를 활용해야 할지 잘 알고 있었으므로 나이팅게일과 함께 활동하기에는 최적임자였다. “세상에서 가장 중요한 과학은 통계학”이라고 했던 나이팅게일과 마찬가지로 파르도 통계를 통해 정확하고 객관적인 사실을 파악할 수 있다고 믿었고 그러한 통계가 올바른 국가 정책의 바탕이라고 생각했다. 그는 나이팅게일에게 보낸 편지에서 이런 말도 했다: “우리는 사실만을 원합니다. 통계는 건조하면 건조할수록 좋은 것입니다. 통계는 우리가 읽는 것 가운데 가장 건조한 것이어야만 합니다.” 즉 통계가 건조하면 건조할수록 편견이나 주관적인 이해관계를 떠나 올바른 정책을 만들 수 있도록 문제의 핵심을 알려준다는 것이었다. 한편, 파르는 고위통계공무원으로서 매년 통계보고서를 펴냈는데, 골턴이라는 통계학자는 파르의 통계보고서로부터 ‘poetical side of statistics’ 즉 ‘통계의 시적인 측면’을 본다고까지 말했다. 마치 짧은 구절 속에 정수만을 담아내는 시처럼, 통계는 복잡한 현실을 예리하게 숫자로 요리해서 중요한 사실만을 보여준다는 뜻이었다.

숫자로 쓰인 시!

시인이 언어의 마술사라면, 통계 전문가는 숫자의 마술사라고나 할까?

### 2.3.2 역사 속의 정책과 통계

우리는 대개 중학교 때부터 수학 시간에 확률, 통계를 공부한다. 흰 공과 검은 공의 이합집산으로 이루어진 경우의 수를 헤아리다보면 너무 헛갈려서 확률, 통계란 흡사 마술사의 교묘한 술수와 다를 바 없어 보이기까지 한다. 알고 보면 학생들을 괴롭히는 경우의 수와 같은 주제들은 17세기 중반부터 수학이나 천문학 연구자들이 도박과 천문관측 결과를 분석하기 위해 연구하기 시작하였다. 그러다가 그러한 수학적 이론들이 사회 현상을 설명하는 데에까지 본격적으로 이용된 것은 19세기가 다 끝나갈 무렵의 일이었다.

그러므로 앞에서 살펴본 19세기 중엽에 나온 나이팅게일의 통계는 우리가 수학 시간에 공부하는 확률, 통계 이론과는 먼 것이었다. 나이팅게일이 활용했던 통계는 수학책 속에 있는 수학자들의 연구가 아니고 19세기 초 서구에서 활발하게 일어났던 통계운동의 결과였다. 그 운동의 주역들은 신학적, 형이상학적 이론보다 통계를 통해 사회의 객관적 사실을 과학적으

로 파악할 수 있다고 믿은 사람들이었다. 거의 모든 나라에서 동시에 시작된 총인구조사와 지역별로 이루어진 다양한 사회조사가 그 대표적인 사례라 할 수 있다. 주민들의 구체적인 생활상을 조사했던 사회조사는 급격한 산업재편에 따라 빠른 속도로 변화하는 사회를 파악하기 위해 새로 등장한 방법이었다. 총인구조사가 전국을 대상으로 했던 반면 사회조사는 특히 도시별로 활발하게 이루어졌다. 주지하다시피 산업화의 결과 농업과 같은 전통적인 산업을 영위하기가 점점 어려워지면서 많은 사람들이 도시의 빈민 주거지역으로 흘러들었다. 그곳에서 그들은 실업 또는 반실업 상태에서 극도로 열악한 생활을 이어갈 수밖에 없었다. 한편 노동자와 빈민의 반대쪽에는 산업화 과정에서 막대한 부를 축적한 지배계급이 있었는데 그들은 노동운동을 혹독하게 탄압할 뿐 아니라 노동계급에게 선거권조차 허용하지 않고 있었다. 이러한 현실을 보고 당연히 여기저기에서 혁명세력들이 등장하게 되었다.

당시 사회통계조사를 위해 활동했던 사람들 역시 이러한 현실을 그대로 둘 수는 없다고 생각하였다. 그 결과 그들은 노동자들의 생활상을 정확히 밝혀내려 하였고 그것을 토대로 궁핍한 사람들의 생활조건을 개선하기 위해 노력하였다. 그런데 그러한 활동을 펼친 사람들은 현실의 비참한 모습을 숨김없이 밝히고 그것을 이용하여 혁명을 도모한 것이 아니었다. 그들이 바꾸려 했던 것은 근본적인 소유와 지배구조가 아니라 국가의 법과 정책이었다. 즉 그들은 의회나 정부기관을 상대로 정책과 법안을 제시하는 온건 개량주의자들이었던 것이다. 이처럼 주로 지역의 젊은 지식인, 상공인들이 혁명이 아닌 사회개량을 위해 선택한 방법이 통계조사였기 때문에, 통계는 처음부터 정부 정책과 대단히 밀접한 관계를 가질 수밖에 없었다.

유럽 여러 나라 가운데에서도 그와 같은 사회통계운동이 가장 활발했던 나라는 영국이었다. 사회통계가 산업화와 밀접한 관계를 가졌고 영국이 산업화가 가장 빨랐으므로 그것은 당연한 일이었을 것이다. 영국에서는 1830년대에 맨체스터를 시작으로 도시마다 통계협회들이 생겨 조사활동을 벌이고 보고서를 발간하였다. 특히 통계협회에서는 그 지역 노동자들의 주거와 교육, 종교, 보건의로 등에 대해 조사하였다. 그런데 그들은 몇 개월에 걸쳐 수천 가구를 일일이 방문하여 조사하였기 때문에, 정부의 영성한 자료보다 훨씬 정확한 자료를 얻을 수 있었다. 따라서 통계협회의 조사는 정부의 자료를 대폭 수정하도록 만들었고 그들이 제안한 정책 역시 보다 강한 설득력을 가질 수 있었다. 아무런 학교 교육도 받지 못하던 많은 노동자 계급 자녀들과, 학교에 가더라도 자질이 부족한 교사로부터 형편없는 환경에서 공부해야 했던 아동들이 보다 나은 학교에서 교육을 받게 된 것도 그러한 성과 중 하나였다.

우리의 눈길을 끄는 점은 이러한 활동 가운데 적지 않은 활동들이 정부 주도로 이루어진 것이 아니었다는 점이다. 노동계급의 성장과 사회주의 운동에도 불구하고 사실상 20세기 초 까지도 유럽 각국 정부에서는 노동통계조차 제대로 작성하지 않고 있었다. 물론 사회의 여러

분야 가운데 각국 정부에서 가장 많은 인력과 예산을 투입하여 자료를 생산했던 분야는 노동 통계였다. 급속한 산업화에 따라 산업구조가 바뀌면서 노동자계급의 규모가 커졌고 점점 노동운동이 활발해짐에 따라 각국 정부에서는 노동청과 같은 기구를 만들어 노동자들의 생활상의 변화라든가 실업률을 파악해서 적절한 정책으로 대응하기 위해 많은 노력을 하였다. 하지만 노동통계는 20세기 초반까지도 만족할만한 수준에 오르지 못하고 있었다. 특히 19세기에는 전반적으로 통계에 대한 인식이 낮았으므로 노조 지도부와 기업주들이 모두 필요한 정보를 정부에 제공하기 꺼려하였다. 게다가 정부의 각 부처와 기관에는 각자 생산한 통계를 취합하여 종합적인 정보를 만들어낼 수 있는 기구가 없었다. 그러다 보니 실업률을 산출하기 위한 통계라는 것이 직업소개소에 찾아오는 사람들만을 실업자로 간주하는 식으로 영성하게 작성되었던 것이다. 특히 정부에서는 실업이라는 문제를 단지 범죄나 소모와 같은 사회적인 문제의 원인으로만 인식하였으므로 그에 대한 정책 역시 실업자를 구제하기보다는 억압하는 방향으로 나아갈 수밖에 없었다. 빈곤 역시 그러하였다. 예컨대 사람들은 오랫동안 게으름이나 습관적인 음주와 같은 개개인의 문제 때문에 빈곤이 생긴다고 믿고 있었다. 그런데 그런 이유로 돌리기에 현실의 빈곤이 너무 큰 규모이며, 것처럼 대규모의 빈곤은 저임금과 대량 실업과 같은 경제구조적인 문제 때문에 생긴다는 것을 통계조사가 밝혀내게 되었다. 19세기 말에 있었던 부스(Charles Booth, 1840-1912)의 런던지역 빈곤조사가 바로 그것이었다.

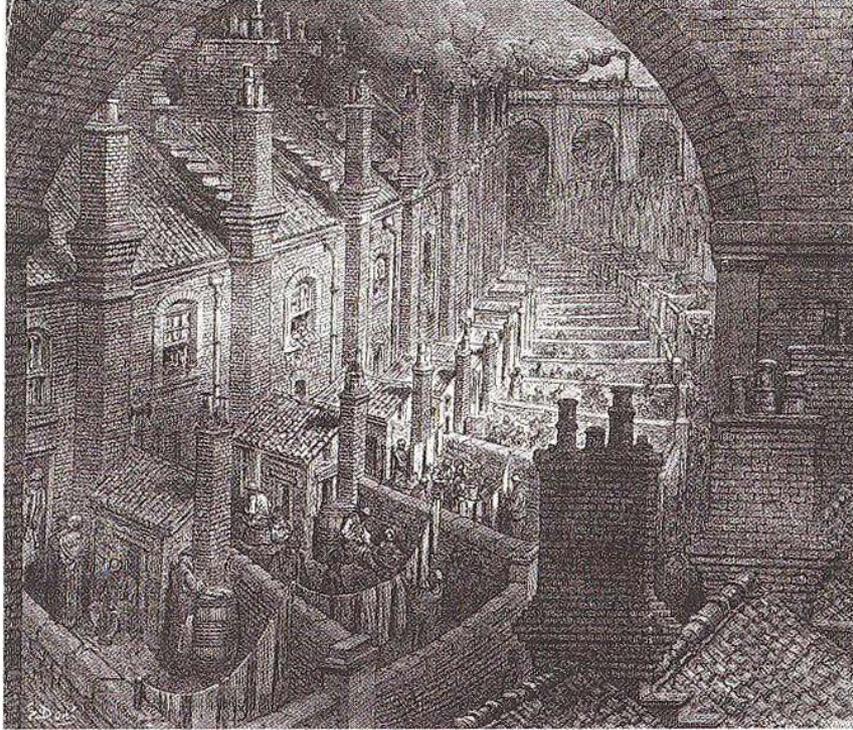
### 2.3.3 영국의 빈곤정책과 부스의 통계조사

영국은 19세기에 세계에서 가장 부유한 나라였다. 그리고 런던은 세계에서 가장 번창하는 도시였다. 하지만 그 엄청난 부에도 불구하고 가난에 찌들어 사는 사람들도 있었다. 특히 런던의 East End 지역은 대표적인 빈민지역이었다. 그런데 당시 영국인들은 누구나 스스로의 힘으로 살아가야 하며 누구나 그럴 능력을 가지고 있다고 생각하였다. 그러므로 빈민은 근면 절약하는 자질을 갖고 있지 못한 사람들로 치부되었다. 즉 가난이란 모두 당사자의 악덕 때문에 생긴다고 여겼던 것이다. 따라서 정부가 할 일이란 대규모의 실질적인 빈민구제사업보다는 치안을 강화하여 빈민들이 일으킬지도 모를 문제를 막는 것이었고 빈민 구제는 자선에 맡길 일이었다.

역사적으로 볼 때 영국 정부가 내놓았던 빈민정책은 모두 빈민을 돕는 데 효과를 발휘하지 못하는 것들이었다. 1563년부터 1601년까지 여러 차례 나온 엘리자베스구빈법을 필두로 영국 정부가 내놓은 구빈법, 그리고 1834년에 실시한 신구빈법에 따른 구빈정책은 빈민구제정책이라기보다는 빈민에게 심한 인간적 모멸감을 주는 정책이었다. 그러므로 그와 같은 정책

들은 큰 효과를 발휘하지 못한 채 빈민문제가 정치적인 문제가 될 조짐까지 나타냈다. 물론 구빈정책의 원칙과 목적에 대해 논의한 보고서와 책들이 많이 나왔지만 가난한 사람들의 실제 삶을 이해한다거나 경제상황이 그 사람들에게 어떤 영향을 미쳤는지 이해하려는 노력은 별로 없었다. 빈민의 생활상을 정확히 알아보려는 통계조사가 없었기 때문에 빈민정책 역시 객관적 현실보다는 관념적인 편견에 따라 만들어지고 실행되었던 것이다.

19세기 말이 가까워지자 눈부신 번영을 구가하던 영국경제는 독일과 미국 등의 경쟁국이 부상하면서 점점 어려움을 겪게 되었다. 이에 따라 저임금 노동자와 빈민의 생활은 더욱 어려워졌다. 광범위한 빈곤 통계조사가 실시된 것은 이러한 사회의 변화를 파악할 필요 때문이었고 19세기가 거의 끝날 무렵이 되어서야 비로소 과학적인 방법으로 설계된 빈곤 통계조사가 부스라는 사람에 의해 이루어지게 되었다. 해운업으로 성공한 사업가였던 부스는 런던 한 가운데에서 언제나 빈민과 중산층 이상 사람들이 서로 지극히 대조적인 삶을 살아가는 것을 보고 빈민 문제는 “모든 문제들의 문제”라고 생각했다. 결국 그는 사회적인 논란이 되고 있는 빈민이 과연 얼마나 되는지 알아보려고 1886년부터 조사를 시작하였다. 원래 정치적으로 보수주의자였던 그의 의도는 정확한 빈민의 수를 헤아려서 빈민이 많다는 사회주의자들의 주장을 반박하려는 것이었다. 그때까지 그는 사업가였을 뿐, 수학을 전공하지 않았음은 물론 통계와도 아무런 관련이 없는 사람이었다.



〈그림 2-3〉 구스타프 도레가 그린 〈런던의 빈곤〉 (1872년경).

그림 위편 연기를 뿜으며 철교 위를 달리고 있는 기차에서 알 수 있듯이 철교 아래에까지 골목조차 볼 수 없을 만큼 뻥뻥이 들어찬 빈민촌을 그린 것이다.

즉 부스가 통계학자나 통계전문가였기 때문에 빈곤 조사를 하게 된 것이 아니었다. 반대로 빈곤이라는 심각한 사회 문제가 눈앞에 있는데 그 문제의 원인과 처방에 대한 무수한 주장들만 난무할 뿐 문제에 대한 객관적이며 기초적인 사실조차 파악되지 않고 있는 현실이 그를 통계조사로 이끌었던 것이다. 통계에 대해 문외한과 다름없던 그는 먼저 영국의 센서스 결과를 연구하였으며 조사를 위해 새롭고 과학적인 조사방법을 창안하였다. 그 결과 그는 나중에 왕립통계협회의 회장 자리에 오르기까지 하였다. 먼저 그는 East End 지역에 대한 조사부터 시작하여 런던 전 지역의 100만 세대, 400만 이상 모든 주민에 대해서까지 조사를 확대하였다. 그는 조사원들을 동원하여 i) 사람들이 일하는 곳과 그곳의 작업 조건, ii) 그들이 거주하는 집과 사는 환경, iii) 그리고 그들의 종교생활에 대해 조사하였다. 그는 조사결과를 1889년 <런던 사람들의 생활과 노동>이라는 제목의 첫 번째 책으로 발표한 데 이어 1903년까지 모두 17권으로 발표하였다. 그가 동원한 방법은 당시로서는 최고의 조사방법이자 최초의 근대적인 방법이었으므로 그의 조사는 발표될 때마다 큰 화제가 되었고 대단한 신뢰를 얻게 되

었다.

부스의 빈곤조사는 요크지방을 대상으로 한 라운트리(Benjamin Seebohm Rowntree, 1871-1954)의 조사로 이어졌다. 라운트리는 부스의 방법을 보다 더 개량한 방법을 이용하였으며 1950년대까지 계속해서 빈곤에 대한 조사 결과를 발표하였다. 그들의 조사결과는 다음과 같이 요약할 수 있다.

- **영국에는 엄청난 규모의 빈곤이 존재한다** : 부스는 런던 주민을 A부터 H까지 8개 등급으로 나누고 A부터 D까지를 빈곤층이라고 불렀다. 그의 조사에 따르면 런던 주민 총 4백20만 9,170명 가운데 최극빈층인 A와 B 등급이 8.4%, 그다음 빈곤층인 C와 D 등급이 22.3%로 나왔다. 즉 런던의 경우 빈곤층이 전체의 30.7%에 이르렀던 것이다. 부스에 이어 1899년 라운트리는 부스의 조사방법보다 더 개선된 방법과 새로 정의한 빈곤선(poverty line)을 써서 요크 지역을 조사하였다. 그는 인구 7만의 소도시인 요크에서는 런던과 같은 대도시에 비해 다른 결과가 나올지 알아보려 하였다. 그런데 라운트리가 얻은 빈곤율은 런던과 비슷한 28%였다. 결국 부스와 라운트리의 조사 결과 영국 전체에 걸쳐 빈곤의 규모가 엄청나다라는 사실이 밝혀졌다. 이 정도의 규모라면 빈곤을 더 이상 개인 탓으로 돌린 채 방치할 수가 없게 된 것이다.

- **빈곤의 원인은 사회경제적 구조에 있다** : 부스와 라운트리는 빈곤의 원인에 대해서도 분석하였다. 두 사람은 공통적으로 저임금과 부정기적인 수입을 빈곤의 가장 큰 원인으로 지적하였다. 즉 아무리 열심히 일을 해도 임금이 너무 낮고 정기적인 수입을 얻을 수 없기 때문에 빈곤을 벗어나기 힘들다는 사실을 밝혔던 것이다. 이에 따라 빈곤은 나태함이라든가 과음과 같은 개인적인 원인에 의해 생기는 것이 아니고 영국 사회 전체의 구조적인 결함 때문에 생긴다는 것이 명백해졌다. 또한 부스와 라운트리의 조사는 종래의 구빈행정이 빈곤문제를 해결하기보다는 악화시켰다는 것도 드러냄으로써 이전까지의 구빈정책에 대한 평가 역할도 하였다.

부스의 조사 이후 사람들은 빈곤은 사회, 경제의 구조적인 현상이므로 빈곤을 일으키는 요인에 대처할 수 있는 유일한 사회제도는 국가뿐이라는 인식을 갖게 되었다. 따라서 19세기 말~20세기 초 영국 빈곤조사의 성과들은 다음과 같이 정리해볼 수 있다.

- 빈곤을 개인의 탓으로 돌리는 19세기의 자유방임적인 빈곤관을 20세기의 복지국가적 빈곤관으로 전환시킴으로써 국가의 빈곤정책을 크게 바꾸어 놓는 계기가 되었다.
- 공공 정책의 수립 과정에서 관념이나 이론에 따른 주장이 아니라, 통계조사로부터 얻은 자료를 근거로 정책을 수립하도록 만들었다.
- 라운트리가 빈곤선을 정의할 때 이용했던 방법은 최저생계비 산출 등을 위해 지금까지도 널리 활용됨으로써 정책 수립의 바탕이 되었다.

### 2.3.4 통계에 대한 저항과 통계의 한계

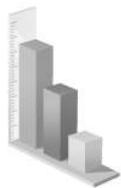
하지만 통계의 역사가 늘 성공과 진보의 역사였던 것은 아니다. 다른 모든 분야들과 마찬가지로 통계 역시 시대의 산물이었다. 가령 자료에 들어있는 정보를 간략하게 요약해서 나타내는 방법 가운데 통계그래픽은 지금도 대단히 중요한 방법이다. 그래픽의 시각적 효과는 숫자들이나 문장으로 된 설명들이 따라올 수 없는 것이다. 그러므로 우리는 통계그래픽이 다른 통계적 분석법보다 먼저 등장했으리라고 생각하기 쉽다. 그런데 우리가 흔히 보는 원그래프나 막대그래프와 같은 단순하고 명료한 그래픽기법들이 등장한 것은 겨우 19세기가 시작될 무렵이었다. 당시는 이미 수 세기동안 많은 분야에서 적지 않은 통계자료들이 축적된 뒤였다. 우리는 여기서 ‘왜 그래픽처럼 쉽고 명료한 방법이 그때까지도 이용되지 않았던가?’라는 의문을 당연히 갖게 된다. 그 이유는 겨우 백여 년 전인 18세기까지만 하더라도 그림은 숫자나 언어와 같이 인간 이성의 인식 대상이 아니라 시각적 인상에 호소하는 것이기 때문에 정확하지 않다고 여겼기 때문이었다.

오늘날 통계학은 과학적인 방법의 대표로 간주된다. 하지만 수많은 사람들이 통계학적인 논리를 받아들이기 어려워하고 있으며, 숫자로 모든 것을 파악하려는 시도에 대해서도 강력하게 저항하고 있다. 통계그래픽의 사례에서 보듯 통계학 역시 오늘날의 시대적인 조건의 산물이다. 따라서 사람들이 통계학에 대해 거부감을 갖고 통계 책을 덮어버리는 것이 그 사람들의 지적인 능력 탓이라기보다는 통계학이 가진 어떤 한계 때문일지도 모른다.

역사적으로 볼 때 19세기는 통계가 널리 활용된 시기인 만큼 통계에 대한 반대 역시 많았던 시기였다. 일례로 19세기 초까지만 하더라도 모든 사람을 하나의 숫자 1로 헤아리는 인구조사에 대해 “고귀한 사람과 하찮은 사람을 어떻게 같이 취급할 수 있는가?”라는 반발이 당연히 나왔었다. 1850년대에 발표된 찰스 디킨스의 소설 <어려운 시절 The Hard Times>에서도 사실만을 추구한다는 통계는 가차 없이 비판받은 바 있다. 디킨스의 주장은 모든 사람을 무

차별적으로 숫자로 환원시켜버린 사회란 인간적인 감정도 윤리도 모두 사라져버린 무서운 세상과 다를 바 없다는 것이었다. 19세기의 의학자들 역시 의학 연구에 통계를 이용하는 것을 환영하기만 했던 것은 아니었다. 도리어 환자 치료는 모두 다른 특성을 가진 개인에 대한 치료인데 많은 환자에 대한 통계로 얻은 결과는 개인들의 특성을 모두 제거해버린 것이기 때문에 특정 환자 치료에는 도움이 되지 못한다는 비판이 19세기말까지도 의학자들 사이에서 상당한 설득력을 얻고 있었다.

앞서 통계와 정책은 서로 밀접한 관계를 갖는다고 하였지만 통계에 바탕을 두고 제안된 정책이 실현되는 속도가 종종 꽤 느렸던 데에는 통계에 대한 인식부족 탓도 크겠지만 통계가 세계를 파악하는 방식에 대해 모두가 동의할 수 없었던 탓도 당연히 컸을 것이다. 즉 사람과 사회에는 단순한 통계숫자로 나타내기 어려운 면들이 적지 않다고 생각하는 사람도 꽤 있었던 것이다. 19세기에만 그러했던 것이 아니다. 지금도 역시 사회과학에서는 통계학이 중심이 된 계량적인 방법으로는 불가능한 연구들이 적지 않으며, 당연히 그런 경우에는 질적인 연구 방법이 뚜렷한 성과를 거두고 있다. 예컨대 동성애자들이 이룬 가정들에 대해 실태를 알아보고 적절한 지원정책을 세우려할 때, 그들을 심층면접하지 않고 몇몇 통계숫자만으로 그들에게 요긴한 지원정책을 만들어내기란 아무래도 어려울 것이기 때문이다.



## 2.4 정책기반으로서의 통계

효율적인 국정운영을 위해서는 합리적인 정책의 수립·집행 및 이에 대한 사후평가가 필수적이라고 하겠다. 합리적인 정책의 수립과 집행은 각종 현상에 대한 정확한 정보를 토대로 했을 때라야 가능하며, 사후 성과평가 역시 정책목표의 달성 정도를 계량화 할 수 있어야만 제대로 이루어 질 수 있다. 이처럼 통계는 사회·경제현상에 대한 정확한 진단을 가능하게 하므로 정책과정의 중요한 기반이 되는 증거로 활용된다.

현재 우리나라에서 정책기반으로 활용할 수 있는 대표적인 통계로는 국가승인통계를 들 수 있다. 2008년 10월 1일 현재 통계청의 승인을 얻어 작성, 공표되는 국가승인통계의 수는 956종이다. 다음의 <표 2-3>에는 부문별, 작성방법별 국가승인통계의 수가 제시되어 있다.

부문별로 보면, 보건/사회/복지 부문의 통계수가 142개로 가장 많은 편이며 작성방법별로는 보고통계가 511개, 조사통계가 379개의 순서이고 가공통계는 66종에 불과하다.

여기서 보고통계란 법령에 의한 개인·단체의 신고, 보고, 신청, 인·허가 등과 같이 다른 행정업무에 수반하여 수집된 자료로부터 작성된 통계를 말한다. 우리나라의 대표적인 보고통계로는 주민등록인구통계, 사망원인통계, 주택건설실적통계 등이 있다. 보고통계는 조사의 어려움이 적고 전수를 대상으로 파악하는 것이므로 세부적인 통계작성이 가능하다는 장점이 있으나, 신고율이나 신고내용의 정확성 등에 따라 통계의 질이 달라지기도 한다.

한편 조사통계는 통계 작성을 주목적으로 조사를 실시하여 얻은 통계로서 모집단의 모든 단위를 조사하는 전수조사와 모집단의 일부 표본만을 조사하는 표본조사로 나눌 수 있다. 우리나라의 대표적인 전수조사로는 인구주택총조사, 농림어업총조사, 사업체총조사 등이 있다. 또한 중요한 표본조사에는 경제활동인구조사, 소비자물가조사, 주택가격조사, 사교육비실태조사 등이 있다. 전수조사는 방대한 규모로 인해 많은 비용과 노력이 필요하지만 모집단의 실상을 정확히 반영할 수 있다는 특징을 가지고 있다. 그에 반해 표본조사는 복잡하고 많은 특성에 관한 정보들을 적은 비용으로 파악할 수 있다는 장점을 갖지만 표본오차가 생길 수밖에 없으며 전문적인 설계 및 추정 이론을 도입하여 검토해야하는 단점도 있다.

보통 조사, 보고 등을 통해 일정 집단에 속하는 개체의 수나 특성의 자료를 직접적으로 집계하여 얻는 통계를 1차 통계라고 한다. 그리고 1차 통계를 통해 얻어진 자료에 일정한 연산을 가하여 얻는 통계를 2차 통계 또는 가공통계라고 부른다. 따라서 가공통계는 1차 통계에 비해 해석적인 특징을 지닌다고 볼 수 있다. 국민계정, 생명표, 추계인구 등이 대표적인 가공통계이다.

〈표 2-3〉 부문별 및 작성방법별 국가승인통계 수

통계 부문	통계 수	작성방법별		
		조사	보고	가공
합계	956	379	511	66
인구	29	4	20	5
고용/임금	54	47	7	-
물가/가계소비	17	17	-	-
보건/사회/복지	142	77	62	3
환경	27	9	17	1
농림/수산	73	50	22	1
광공업/에너지	33	20	9	4
건설/주택/토지	49	15	28	6
교통/정보통신	54	28	25	1
도소매/서비스	17	16	1	-
경기/기업경영	67	47	10	10
국민계정/지역계정	14	-	-	14
재정/금융	29	4	25	-
무역/외환/국제수지	11	4	4	3
교육/문화/과학	68	38	27	3
기타	272	3	254	15

출처 : 통계청 국가통계포털(<http://www.kosis.kr>)

이제 다음 세 가지 질문을 통해 우리나라의 국가통계에 대해 좀 더 알아보기로 하자. 1) 위에 언급된 바와 같은 국가승인통계가 정책의 기반으로 구체적으로 활용되고 있는 대표적인 사례는 무엇인가? 2) 정책기반으로서의 통계 자체가 미비하거나 부실할 경우 어떤 문제점들이 발생할까? 3) 정책과정의 기반으로서 새로운 통계를 작성할 필요가 있을 때에는 어떤 점들을 고려해야 할까? 이러한 세 가지 질문들에 대한 답을 실제적인 사례들을 통해 찾아보고자 한다. 첫째 질문에 대해서는 우리나라의 가장 기본적이고 대표적 통계인 인구주택총조사를, 그리고 둘째 질문에 대해서는 지난 수년간 언론 등에 실제 노출된 사례들을 분석할 것이다. 또한 마지막 질문에 대해서는 통계의 작성과정 및 작성방법에 따른 고려사항을 살펴보고자 한다.

### 2.4.1 대표적 정책기반 통계 사례 : 인구주택총조사

인구주택총조사는 우리나라에서 작성되는 모든 통계 중 가장 대표적인 통계이다. 통계청의 인구주택총조사 웹사이트(<http://www.census.go.kr>)에서는 이 조사를 다음과 같이 저의하고 있다.

“우리나라의 모든 인구와 주택의 총수는 물론 개별 특성까지 파악하여 각종 경제 사회 발전계획의 수립 및 평가와 각종 학술연구, 민간부문의 경영계획수립에 활용하기 위해 실시하는 전국적 규모의 통계조사입니다.”

위 정의에서 이 조사는 각종 경제사회 발전계획의 수립 및 평가에 활용하기 위해 실시된다고 언급하고 있는데, 구체적으로 어떤 기관에서 어디에 이 조사결과를 활용할까? <그림 2-4>에서 한 가지 사례를 볼 수 있다. 이 그림은 지난 2005년 인구주택총조사 홍보용 그림 중 하나로서 “학교가 필요한 곳에 학교를”이라는 표어로부터 이 조사 결과가 새 학교를 어디에 세울 것인지 판단할 기초자료로도 쓰일 것임을 알 수 있다.



<그림 2-4> 2005년 인구주택총조사 홍보물

또한 인구주택총조사 웹사이트를 보면 이 조사결과가 ① 각종 국가정책 및 지방자치단체의 지역발전 계획 수립을 위한 기초자료 ② 다른 표본조사의 모집단 자료 ③ 각종 비교·분석자료 등으로 활용된다고 소개하고 있다.

그 가운데 먼저 국가정책 및 지방자치단체의 지역발전 계획 수립을 위한 기초자료로 이 조사결과가 어떻게 활용되는지 살펴보자. <표 2-4>를 보면 조사의 활용내용과 해당 정부기관이 나와 있다. 우선 이 표에서 알 수 있는 것은 인구주택총조사 결과는 거의 모든 정부기관들이 중요하게 활용하고 있다는 점이다.

쉬운 예로 국방부를 들 수 있다. 국방정책을 세우는 데 있어서 가장 중요한 기초자료 가운데 하나가 미래의 병력자원 규모일 것이다. 우리나라는 세계에서 출산율의 변동이 가장 급격한 편인지라, 연령대별 인구수가 극심하게 변한다. 따라서 장기적인 국방정책 마련을 위해서는 미래의 인구 상황을 나타내는 통계가 필요할 것이다. 또 다른 예로 교육정책을 들 수 있다. 교사 수급계획이나 학교의 설립 및 운영 정책을 마련하기 위해서는 세부 지역별 인구수 자료가 필수적이다. 이와 같이 다양한 분야에서 정책 대상이 되는 집단의 현황을 파악하는 것은 합리적인 정책수립을 위한 기초가 된다.

〈표 2-4〉 각 정부기관별 활용내용

활용내용	이용기관
고용구조 및 인력대책 마련, 고용보험 개발	노동부
주택정책 수립	건설교통부, 재정경제부
국토건설종합계획 수립	건설교통부
노인, 아동, 부녀 복지정책 수립	보건복지부
농어촌 구조개선 대책	농림부, 해양수산부
병력자원 예측	국방부, 병무청
교육정책 수립	교육인적자원부
문화예술정책 수립	문화관광부
대도시 교통망 건설계획 수립	건설교통부, 서울특별시 및 광역시
중소기업 육성시책 수립	중소기업청, 중소기업협동조합중앙회
여성복지 정책 수립	여성부, 여성개발원
주택개발 등 각종 사업계획	지방자치단체
시도별 주택건설계획 수립	대한주택공사

출처 : 통계청 인구주택총조사(<http://www.census.go.kr>)

둘째, 인구주택총조사 결과는 여러 표본조사의 설계를 위한 모집단자료로 활용된다. 인구주택총조사와 같은 총조사를 위해서는 매우 많은 예산과 인력이 소요되므로 5년 또는 10년 주기로 실시되는 것이 일반적이다. 인구주택총조사를 제외한 상당수의 조사들은 총조사가 아닌 표본조사 방식으로 이루어진다. 다음의 <표 2-5>에는 각 기관에 의해 작성되는 주요 표본조사들이 열거되어 있다. 인구주택총조사 결과는 이러한 조사에서 표본설계를 위한 추출틀로 활용된다.

〈표 2-5〉 주요 표본조사들

표본조사명	이용기관
가계조사, 생활시간조사, 가계자산조사, 경제활동인구조사, 인력실태조사, 소비자물가조사, 소비자전망조사 등	통계청
조손가족 실태조사	여성가족부
결혼 및 출산동향조사	저출산 고령사회위원회
정보화실태조사	정보통신부
근로환경조사	노동부
방문객여행실태조사, 고용통계조사, 사회통계조사, 농업기본통계조사 등	지방자치단체
가족실태조사, 장애인실태조사, 국민생활실태조사 등	한국보건사회연구원
중고령자 패널조사	한국노동연구원
소비자 동향조사	한국은행
정신질환자 실태조사	서울대학교병원
산업직업별 고용구조조사	한국고용정보원
주거실태조사	국토연구원

출처 : 통계청 인구주택총조사(<http://www.census.go.kr>)

셋째, 인구주택총조사는 각종 비교분석을 위한 연구의 자료로도 활용되는데, 아래 <표 2-6>에는 그 사례들이 소개되어 있다. 노동통계, GDP, 지역의 인구 및 생산현황, 지역소득 추계 등 주요 가공통계의 작성을 위한 기초자료를 제공하는 데에 그리고 각종 간행물 수록, UN이나 OECD, ESCAP 등 각종 국제기관에 대한 자료 제공 등의 목적으로도 인구주택총조사가 활용된다.

〈표 2-6〉 인구주택총조사의 활용사례

활용내용	이용기관
노동통계비교	노동부, 노동통계연구원
GDP 작성	한국은행
지역인구 및 생산현황 작성	지방자치단체
고용통계비교, 지역소득추계, 소비자물가지수 도시별 가중치 산정 등	통계청

- 대학, 연구소 등 학술 연구 자료로 활용됩니다.
- 장애인구추계 및 생명표 작성을 위한 기초자료로 활용됩니다.
- 각종 간행물에도 수록됩니다.
  - + 교과서 수록(교육과학기술부)
  - + 경제지표, 사회지표(시·도)
  - + 지역경제백서(시·도)
  - + 보건복지통계연보(보건복지가족부)
  - + 시·도 통계연감(각시·도)
  - + 여성사회통계보고서(여성개발원)
- 대학의 강의·연구 자료로 활용됩니다.
- 개인의 논문자료로 활용됩니다.
- 국제비교자료 및 국제기관 제공 자료로 활용됩니다(UN, ESCAP, 각국).

출처 : 통계청 인구주택총조사(<http://www.census.go.kr>)

우리나라의 총조사에 관한 기록을 더듬어 올라가면 삼한시대에 이르게 된다. 이후 시작되어 삼국시대, 고려 및 조선시대에 이르기까지 소위 호구조사(戶口調査)라는 명칭의 조사기록이 발견되고 있다. 근대적 의미의 인구조사인 인구총조사는 1925년에 처음으로 실시된 후, 대략 5년마다 조사가 실시되어 2005년까지 총 17차례의 조사가 있었다. 1925년의 간이국세 조사를 시작으로 근대적인 총조사의 역사와 명칭, 그리고 조사의 특징을 정리한 것이 <표 2-7>에 나와 있다.

〈표 2-7〉 근대적 총조사 역사, 명칭 및 특징

회수	조사기준일	명 칭	
1	1925. 10. 1	간이국세조사	- 최초의 인구센서스
2	1930. 10. 1	조선국세조사	- 최초로 직업 등 경제활동 사항 포함
3	1935. 10. 1	조선국세조사	
4	1940. 5. 1	국세조사	
5	1944. 5. 1	간이국세조사	
6	1949. 5. 1	총인구조사	- 최초로 인구이동 사항 포함
7	1955. 9. 1	간이총인구조사	
8	1960. 12. 1	인구주택 국세조사	- 경제활동 및 출산력 사항의 20% 표본 집계 - 노동력 개념 설정 - 최초로 주택에 관한 조사 실시
9	1966. 10. 1	인구센서스	- 10% 표본조사 병행(경제활동 및 출산력)
10	1970. 10. 1	총인구 및 주택조사	- 10% 표본조사 병행 (경제활동, 출산력, 인구이동 및 일부 주택에 관한 사항)
11	1975. 10. 1	총인구 및 주택조사	- 5% 표본조사 병행 (경제활동, 출산력, 인구이동 및 일부 주택에 관한 사항)
12	1980. 11. 1	인구 및 주택센서스	- 10% 표본조사 병행(경제활동, 출산력, 인구이동)
13	1985. 11. 1	인구 및 주택센서스	- 전항목 전수조사 - 성씨, 본관 및 종교에 관한 조사 실시
14	1990. 11. 1	인구주택총조사	- 광학판독 자료입력 방식(OMR System) 도입
15	1995. 11. 1	인구주택총조사	- 빈집조사표를 이용한 빈집조사 실시
16	2000. 11. 1	인구주택총조사	- 아파트 전수조사구에 대한 응답자 기입방식 도입 - PC 지방분산 자료입력 방식 도입
17	2005. 11. 1	인구주택총조사	- 인터넷 조사방식 도입과 웹 기반 현지 자료입력 방식 도입

출처 : 통계청 인구주택총조사(<http://www.census.go.kr>)

<표 2-7>을 살펴보면 알 수 있듯이, 우리나라의 총조사는 간이국세조사라는 이름으로 시작되어 총인구조사라고 불렀다가 60년대에 조사범위가 확대되면서 인구주택 국세조사가 되었고 오늘날에는 인구주택총조사라고 불리고 있다. 이와 같은 이름의 변화만 보더라도 조사 내용과 방법의 변화를 짐작할 수 있다. 처음에는 단순히 인구에 관한 조사를 하다가 점차 경제, 사회, 문화 등 제반 속성과 주택에 관한 사항까지 조사하게 되었다. 이는 여러 정부부처와 다

양한 민간 부문에서 확대된 조사 결과를 필요로 했기 때문이었을 것이다. 조사내용 뿐 아니라 조사방법과 기술면에서도 10% 표본조사를 병행한다든지 전자식 자료입력 방식 및 인터넷 조사 방법을 도입하는 등 Cyber 공간을 이용하는 조사 및 교육 기법까지 영역을 넓히고 있다.

한편 통계청에서는 이와 같은 인구주택총조사 결과를 인쇄된 보고서와 CD-ROM 등의 형태로 간행하는 이외에 국가통계포털(KOSIS)을 통해 웹으로 자료를 검색할 수도 있도록 하고 있다. 참고로 아래 <그림 2-5>는 국가통계포털의 자료검색 결과 화면이다. 이 그림은 전국의 성별, 연령별 인구수 자료만 나타내고 있지만 실제 국가통계포털에서는 인구주택총조사에서 조사한 모든 특성에 대해 검색할 수 있고, 연도별로도 시계열 자료를 검색할 수 있다. 그러므로 이 결과를 이용하면 특정 소지역에 대한 통계를 별도로 관리할 수도 있다. 보다 구체적으로 3장에서 인구주택총조사 자료를 검색하여 활용하는 사례를 다룰 것이다.

총조사인구 총괄(행정구역/성/연령별)

행정구역별	# 연령별	2005			
		인구	남자	여자	성비
전국	계	47,041,434	23,465,650	23,575,784	99.53
	0 - 4세	2,382,350	1,237,301	1,145,049	108.06
	5 - 9세	3,168,887	1,654,228	1,514,659	109.21
	10 - 14세	3,434,891	1,816,318	1,618,573	112.22
	15 - 19세	3,100,523	1,626,378	1,474,145	110.33
	20 - 24세	3,662,123	1,915,902	1,746,221	109.72
	25 - 29세	3,671,847	1,858,332	1,813,515	102.47
	30 - 34세	4,096,282	2,059,913	2,036,369	101.16
	35 - 39세	4,112,785	2,065,668	2,047,117	100.91
	40 - 44세	4,123,041	2,082,427	2,040,614	102.05
	45 - 49세	3,900,899	1,961,859	1,939,040	101.18
	50 - 54세	2,855,297	1,426,597	1,428,700	99.85
	55 - 59세	2,278,438	1,126,997	1,151,441	97.88
	60 - 64세	1,888,853	897,384	991,469	90.51
	65 - 69세	1,680,067	755,949	924,118	81.80
	70 - 74세	1,252,734	514,241	738,493	69.63
	75 - 79세	766,870	270,632	496,238	54.54

<그림 2-5> 국가통계포털의 자료검색결과 화면

## 2.4.2 통계 부실로 인한 정책상 문제점

정확성·신뢰성을 위시한 통계의 여러 속성을 제대로 갖춘 통계는 불확실한 시대를 살아가는 현대인에게 객관적인 판단 방향을 제공해주는 정보인 동시에, 올바른 정책 수립에 없어서는 안 될 핵심 인프라이자 무형의 사회간접자본이라 할 수 있다. ‘과학적이고 합리적인 행정’이 이루어지기 위해서는 모든 정책의 입안과 결정이 정확하고 신뢰할 만한 통계를 기반으로 이루어져야 하며, 정책의 집행과 평가에서도 역시 그러한 통계가 토대를 이루어야 한다. 만약 해당 분야의 실상을 제대로 반영하지 못하는 부실한 통계를 근거로 단기적 또는 중장기적 대책들이 마련된다거나, 새로운 수요를 가진 정책 형성에 활용할 통계가 없거나 부족하다면, 그리고 만약 기존의 통계를 다른 의도를 가지고 오용·남용한다면, 그로 인한 피해는 고스란히 해당 고객들은 물론 국민 전체에게 돌아갈 것이다. 다시 말해 통계의 부실로 인해 각종 정책들에 문제가 발생하면 예산 낭비는 물론 정책의 실패로까지 이어지고, 심지어는 정부에 대한 신뢰마저 훼손되어 사회적 불안까지도 초래할 수 있을 것이다.

정책과정에서 통계 때문에 발생하는 문제의 유형은 크게 ① 관련 통계가 존재하지 않아 정책형성에 어려움을 제공하는 경우 ② 관련 통계가 존재함에도 불구하고, 그 통계들을 정책에 적절하게 활용하지 않는 경우 ③ 부정확하거나 사업목적에 부적절한 통계를 활용하여 정책의 실효성을 떨어뜨린 경우 등으로 나누어 볼 수 있다. 여기서는 2007년 6월 감사원에서 발표한 「국가 주요통계 작성 및 활용실태」 감사 결과보고서에 소개된 사례들 가운데 일부와 최근 2~3년 동안 각종 매스컴에서 비중 있게 보도한 실제 사례 등을 원용하여 각 경우들을 살펴보기로 한다.

### (1) 통계 미비로 인한 문제 : 보건복지부의 「의료급여 수급권자를 대상으로 한 본인부담금제도」 사례

■ **제도의 개요** : 보건복지부는 2005년 11월 19일 그동안 무료로 병·의원 외래진료를 이용해온 ‘의료급여’ 수급권자들도 앞으로는 본인부담금제에 따라 1000~2000원의 진료비를 내야 한다고 발표하였다. 또한 그러한 대책의 근거로 의료급여 대상자의 1명당 진료비가 건강보험 가입자보다 3.3배나 많다는 자료를 제시하였다.

■ **통계적 측면의 문제점** : 이 자료는 진료비에 중대한 영향을 끼치는 성별, 나이, 중증질환 비율 등 주요 변수들이 전혀 고려되지 않은 잘못된 통계에 근거를 둔 것으로 드러났다. 2005

년 건강보험 통계연보에 따르면, 의료급여 대상자의 경우 진료비가 많이 드는 노인 비율이 25.6%로 건강보험의 8.3%보다 세 배 이상 높았으며, 의료급여 대상자들의 경우 한 달에 수백·수천만 원의 진료비가 드는 중증환자의 비율이 건강보험 가입자보다 훨씬 높다는 사실이 감안되지 않았던 것이다.

■ **조치 사항** : 당시 보건복지부는 이러한 지적에 따라 연령과 중증질환자 수 등을 보정하여 새로 계산한 결과를 발표한 적이 있다(한겨레신문, 2006.12.31일자). 그 결과에 따르면 의료급여 대상자의 평균 진료비가 건강보험 가입자보다 1.48배 높은 것으로 나타났다. 즉 보건복지부의 의료급여 대상자들에 대한 본인부담금제도는 잘못된 통계에 근거를 두고 수립·시행된 것이었다. 따라서 해당 부처에서는 보정된 통계에 맞춰 본인부담금제 등의 의료급여 대책을 다시 세워야만 했다.

## (2) 통계의 미활용으로 인한 문제 : 건설교통부의 「국도확장사업 수요 재검증 사업」 사례

■ **사업관리지침의 내용** : 국도확장사업을 위한 타당성 조사에 필요한 수요예측 기초통계자료(기종점교통량, 교통분석용 네트워크, 통행실태자료)가 통일된 기준 없이 용역 수행기관마다 다르게 작성되어 교통량 예측의 신뢰성이 낮다는 문제가 제기됨에 따라 건설교통부에서는 2003년부터 통일된 기준에 따라 생산하는 국가교통데이터베이스(Korea Transport DataBase; KTDB)를 활용하여 교통량을 예측하도록 하였다. 또한 기획예산처의 「총사업비 관리지침」 제 35조, 제46조 및 제50조의 규정에 따르면, 수요예측 방법론상의 오류 등으로 현저한 수요 감소가 발생할 것이 예상되면 수요예측을 재검증하고 그 결과에 따라 사업계획을 조정하도록 되어 있다.

■ **통계적 측면의 문제점** : 그러나 건설교통부에서는 2006년 10월 현재 「국가교통 DB」를 관련 사업의 타당성 검증에 활용하지 않은 채 2002년 이전에 사업 타당성조사를 마친 국도 42호선 평창-정선구간 등 162개 국도확장사업을 시행하고 있었다.

■ 「국가교통 DB」 활용 결과 : 162개 국도확장사업 중 국도 82호선 우정-장안구간 등 12개 국도확장사업에 대해 설계 당시의 예측교통량과 KTDB를 활용하여 산출한 예측교통량을 비교·검토한 결과, 일부 구간에서 사업계획 당시 예측교통량이 과다 추정된 것으로 나타났으며, 그 구체적인 예의 일부가 아래 <표 2-8>에 제시되어 있다.

〈표 2-8〉 예측교통량 비교

사업명	차로	설계시 예측교통량		KTDB 예측교통량	
		기준연도	대/일	기준연도	대/일
국도 27호 서수-군산구간	4	2001	21,397	2015	7,072
국도 42호 평창-정선구간	4	2009	8,326	2015	5,554

주 : 건설교통부 「도로업무편람」에 따른 국도 4차로 확장기준 - 1일 7,300대

### (3) 통계의 오용으로 인한 문제 : 농림부의 「농축임어업인의 영유아에 대한 양육비 지원 사업」 사례

■ **지원사업의 개요** : 농림부에서는 「농림어업인 삶의 질 향상 및 농산어촌지역개발 촉진에 관한 특별법」 제17조, 제18조 및 제22조의 규정과 「여성농어업인육성법」 제11조의 규정에 따라 2004년부터 농축임어업인의 만 5세 이하 영유아에 대해 양육비를 지원하는 사업을 시행하였다. 그리고 양육비 지원 대상으로는 도시근로자 가구와의 형평성을 고려하여 연간 도시근로자 평균소득(2004년도 통계청 가계조사 37,360천원) 이상의 소득(2006년도 양육비 지원대상 제외 기준이었음)을 올리는 농축임어가는 제외하기로 하였다.

■ **통계적 측면의 문제점** : 개별 가구소득을 알 수 있는 통계자료를 이용하여 도시근로자 가구 평균 소득을 웃도는 경우나 재산이 많은 경우는 지원 대상에서 제외해야 했다. 그러나 농림부는 개별가구의 실제 소득은 고려하지 않은 채, 통계청의 ‘농가경제조사’의 경지규모별 농가소득 통계만을 활용하여 지원 대상을 선정하여 양육비를 지급하였다. 보다 구체적으로 언급한다면, 농지소유규모가 5ha인 농가의 평균소득이 37,350천원으로 도시근로자 가구 평균소득에 상응한다는 이유로 5ha 미만의 농지를 소유한 농가 및 이에 상당하는 임·축·어가들을 지원 대상으로 선정하였다.

해양수산부 역시, 어로어가의 대표업종인 채낚기 어선의 적당 연평균 소득이 36,122천 원 이고, 근해 어선어업의 최대톤수가 90톤 이하라는 사유로 가구별 실제 소득은 고려하지 않은 채 어선이 없는 어가와 무동력선 사용어가 및 90톤 이하 동력선 사용어가 등을 지원 대상으로 결정하고 이를 농림부에 통보하였다.

그 결과 농·축·임·어업 중 둘 이상을 겸업하거나, 영유아의 부모 중 1인이 농·축·임·어업 이외의 직업이 있어 도시근로자가구 평균소득을 웃도는 소득을 올리는 가구에도 소득 및 재산 상황에 관계없이 일률적으로 양육비를 지원하게 되었다.

### 2.4.3 정확하고 타당한 통계를 활용하지 못하는 원인 분석

정부기관이 통계를 작성·공표할 때에는 통계자료의 신뢰성이 확보되어야 하며, 이를 위해서는 정확한 통계수치를 제공해야하고 조사대상을 선정하는 과정에서 모집단 중 일부를 누락시키지 않아야한다. 또한 유사하거나 같은 내용의 통계를 여러 정부기관에서 서로 다르게 발표함으로써 통계 이용자들에게 혼란을 주는 일도 없어야한다. 여기서 부정확하거나 사업 목적에 타당하지 않은 통계를 활용할 수밖에 없는 이유를 원인 별로 분류하고 그 사례들을 살펴보자.

#### (1) 기관 간 행정자료 공유 미흡 : 경찰청의 「도로 교통 정책」 사례

■ **부실한 교통사고 관련 통계** : 경찰이 발표하는 교통사고 사상자 수는 매년 줄어들고 있는 반면 보험사에 접수된 사상자 수는 오히려 큰 폭으로 증가하고 있다. 그 결과 양쪽의 교통사고 사상자 통계가 3~4배가량 차이가 나고 있으며, 그 차이는 갈수록 커지고 있다. 형사 처벌 대상인 10대 중과실 사고 통계의 경우에서도 그 차이는 심각한 수준이라고 한다.

■ **부실 원인** : 경찰은 교통사고 통계가 이처럼 다른 이유로 인력 부족과 업무과다를 들고 있다. 하지만 경찰이 교통사고 수를 줄이기 위해 사상자가 발생한 교통사고 조사를 꺼리고 있는 것 아닌가 하는 오해마저 받고 있는 실정이다. 이러한 통계의 차이는 사실 그동안 보험사와 경찰 사이에 정보를 공유할 법적 근거가 없었기 때문에 경찰로서는 보험사가 만든 통계를 가져올 수 없었다는 이유가 가장 크게 작용하였던 것 같다.

■ **부실한 통계의 활용 결과** : 이처럼 현실과 크게 다른 경찰의 잘못된 교통사고 통계를 기초로 도로교통 정책이 만들어지다 보니, 정책의 실효성은 떨어지고 관련 예산도 효율적으로 집행되지 못하였다.

#### (2) 부적합한 조사의 범위와 대상 : 과학기술부의 「여성과학기술인의 육성 및 지원에 관한 기본계획 수립」 사례

■ **실태조사의 목적** : 과학기술부에서는 「여성과학기술인의 육성 및 지원에 관한 법률」 제6조와, 같은 법의 시행령 제8조 규정에 따라 2005년부터 ‘여성과학기술인력활용에 관한 실태조사’를 하고 있다. 이 실태조사는 여성과학기술인의 활용 현황과 이 법을 시행한 성과를 파

악하여 여성과학기술인 육성과 지원에 관한 기본계획을 수립하고 심의하는 데 활용하기 위한 것이다.

■ **실태조사의 문제점** : 그런데 과학기술부에서는 여성과학기술인의 규모에 대해서는 파악하지 않고 있었다. 그 항목이 여성과학기술인 활용 현황의 기본 조사항목에 들어있지 않다는 이유였다. 또한 그 법 제2조와 시행령 제2조의 규정에서 ‘여성과학기술인’이란 ‘이공계 분야의 학위나 자격을 가지고 연구직·기술직 또는 관련 직종에 종사하거나 종사하고자 하는 여성’이라고 되어 있으므로, 연구기관 외에 기술관련 기관에 근무하는 여성도 조사 대상에 포함되어야 할 것이다. 하지만 실태조사 대상을 구체적으로 규정하고 있는 위 법 시행령 제8조 제1항에서는 기술표준원 등과 같이 시험·검사·검정·인증 등의 업무를 수행하는 기관에서 근무하는 여성과학기술인들은 조사 대상에 포함시키지 않고 있다.

■ **불완전하게 작성된 통계의 결과** : 여성과학기술인의 활용 현황 외에 법 시행에 따른 성과를 종합적으로 파악하지 못할 뿐만 아니라, 전체 여성과학기술인력 현황에 근거를 둔 여성과학기술인력 지원, 육성정책이 수립되지 못할 우려가 있다.

### (3) 대표성이 결여된 표본조사 : 산업자원부의 「산업인력 양성 관련 정책」 사례

■ **실태조사의 목적** : 산업자원부는 「산업기술기반 조성에 관한 법률(2006. 4. 28. 「산업기술혁신촉진법」으로 대체)」 제6조의 규정에 의거하여 산업기술인력 양성 관련 정책에 참고하기 위하여 한국산업기술재단에 위탁하여 ‘2005 산업기술인력 수급동향 실태조사’ 통계를 작성하여 공표(2006. 7. 13.)한 바 있다.

■ **실태조사의 문제점** : 이 재단은 8대 기간산업 중 근로자 수의 비중이 큰 300인 이상 대형사업체 595개는 전수 조사하기로 표본설계를 하고도 반도체 산업의 경우 36개 사업체 중 21개, 조선 산업의 경우는 10개 사업체 중 5개 사업체만 조사하는 등 아래 <표 2-9>와 같이 전체의 48.4%인 288개 사업체만 조사하였다. 또한 실제 조사에서 생략된 사업체의 산업기술인력을 추정하면서 해당 사업체의 근로자 규모를 고려하지 않고 단순히 시·도별로 조사된 ‘사업체수’의 비율에 따라 단순 배수를 적용하여 산업기술인력의 현원과 부족인원 등을 산출하였다.

■ **부실한 실태조사의 결과** : 이렇게 산출된 결과는 현실을 제대로 반영하지 못하므로 실제 산업기술인력 현원과 부족인원이 조사결과와 큰 차이를 보이게 마련이다. 이 같이 표본이 모

집단을 대표하지 못하는 부실조사에 대해 산업자원부는 재조사 또는 보완요구 등을 하지 않은 채 그대로 공표하였으며, 이처럼 신뢰성이 낮은 통계조사 결과를 ‘공과대학 혁신비전과 전략’등 산업기술인력 양성관련 정책 자료로 활용한 바 있다.

〈표 2-9〉 산업별 전수조사대상 사업체 조사 현황

산업별		300~499인		500인 이상		계	
		사업체 수	조사업체 수	사업체 수	조사업체 수	사업체 수	조사업체 수
8대 기간 산업	기계	38	13	34	15	72	28
	반도체	9	5	27	16	36	21
	섬유	45	16	26	15	71	31
	자동차	32	21	50	29	82	50
	전자	79	33	92	42	171	75
	조선	2	2	8	3	10	5
	철강	21	7	22	15	43	22
	화학	58	24	52	32	110	56
소 계		284	121	311	167	595	288 (48.4%)

출처 : 산업자원부 감사자료 재구성

#### (4) 전체집계에서 일부 유형의 통계 누락 : 재정경제부의 「해외직접투자통계」 사례

■ **통계작성의 목적** : 재정경제부에서는 「외국환거래법령」의 규정에 따라 해외직접투자 현황을 파악하여 민간기업의 해외투자전략수립, 정부의 해외투자 및 대외자산에 관한 정책수립 등을 위한 기초정보를 제공할 목적으로 ‘해외직접투자통계’를 작성하여 매 분기 공표하고 있다.

■ **통계작성의 실태** : 「외국환거래법」 제3조 제1항과 외국환거래규정(재정경제부 고시) 제9장에서는 해외직접투자를 금융·보험업 이외의 해외직접투자, 금융·보험업에 대한 해외직접투자, 그리고 국내기업 등 해외지사에 대한 해외직접투자의 세 가지 유형으로 분류하고 있다. 그러나 재정경제부에서는 한국수출입은행이 집계하고 있는 금융·보험업 이외의 해외직접투자 실적만으로 통계를 작성하고, 나머지 두 가지 유형에 대해서는 신고체계와 사후관리기관이 다르다는 이유로 통계작성을 하지 않아 왔다.

■ **문제점** : 이처럼 해당 통계에 대한 정의나 통계수집 방식에 차이가 있기 때문에 집계하

는 기관마다 통계수치가 크게 다를 수밖에 없었다. 그 결과 2005년도 해외직접투자통계에서는 739백 만 달러의 실적이 누락되었고, 통계작성 범위 등을 명시하지 않은 채 「2005년도 해외직접투자 동향」에 관한 보도자료(2006. 1. 31.)를 공표하여 혼란을 초래했다.

#### (5) 동일 대상 통계를 부처별로 다른 기준에서 작성 발표 : 산업자원부와 건설교통부의 「산업입지 정책수립」 사례

■ **통계작성의 목적** : 산업자원부와 건설교통부에서는 효율적인 산업입지 정책수립을 위한 기초 통계자료로 활용하고 잠재적인 투자기업에 대해 입지수요 정보를 제공할 목적으로 매 분기와 매월 전국 산업단지(국가, 지방, 농공단지) 현황에 관한 통계자료를 작성·공표하고 있다.

■ **통계작성의 기준** : 「산업입지 및 개발에 관한 법률」 제2조 제5호 및 「산업집적활성화 및 공장설립에 관한 법률 시행령」 제43조 제1항 제4호 등의 규정에 따르면, 자원비축시설도 산업단지의 산업시설구역에 포함하고 있으므로 자원비축산업단지는 산업단지 통계에 포함시켜야 할 것이다. 또한 산업단지에서 해제된 경우에는 산업단지 통계에서 제외시켜야 하는 것은 당연한 일이라 하겠다.

■ **통계작성의 실태** : 그러나 산업자원부에서는 고정국가산업단지 등 5개의 자원비축산업단지(부지면적 16,813천 m<sup>2</sup>, 총 국가산업단지면적의 3.4%)와 천안영상문화복합단지(부지면적 493천 m<sup>2</sup>, 총 지방산업단지 면적의 0.2%)를 법령을 제대로 모르거나 지방자치단체와의 업무협조가 제대로 되지 않아 국가산업단지와 지방산업단지 통계에서 누락시켰다. 한편 건설교통부의 산업단지현황 통계(2006년 6월 말 기준)에서는 화전지방산업단지 등 4개 지방산업단지와 17개 농공단지를 산업단지현황 통계에서 누락시킨 반면, 2004년 1월과 2006년 1월에 각각 산업단지에서 해제된 이동농공단지(53천 m<sup>2</sup>)와 구성지방산업단지(1,668천 m<sup>2</sup>)는 통계자료에 여전히 포함시켰다.

■ **작성 결과** : 이처럼 정부의 두 부처에서 각각 전체 산업단지들의 모집단 가운데 일부에 대해 통계조사를 제대로 하지 않아 통계의 신뢰성과 효용성을 떨어뜨리게 되었다.

#### 2.4.4 정책기반 통계를 작성할 때 고려해야 할 사항

앞에서 통계가 갖추어지지 않았거나 부실하여 합리적인 정책을 수행하기 어려워진 사례들을 살펴보았는데 이러한 문제점을 개선하기 위해 2008년 6월 개정된 통계법 시행령 제33조에는 “중앙행정기관의 장은 소관 법령의 제정 또는 개정을 통하여 새로운 정책과 제도를 도입하거나 종전의 정책과 제도의 중요 사항을 변경하는 경우에는, 통계에 기반한 정책관리를 위하여 해당 법령의 제정 또는 개정을 통해 도입되거나 변경되는 정책과 제도의 집행·평가에 적합한 통계의 구비 여부 등에 대한 평가를 통계청장에게 요청하여야 한다.”라는 내용이 명시되어 있다. 이는 정책이나 제도와 관련하여 객관성, 일반성, 포괄성, 정책관련성, 신뢰성 등의 기본적인 요건을 갖춘 타당성 있는 기반으로서의 통계를 작성하기 위해서이다. 해당 정책이나 제도의 집행이나 평가에 필요한 통계가 제대로 갖춰져 있지 않을 때에는 새로운 통계를 개발하거나 기존 통계를 보완해야한다. 통계청에서는 통계를 보완, 개선하거나 개발 계획을 세울 때에 검토할 사항들을 다음과 같이 제시하였다.

- ① 통계 개발·개선 계획을 통해 필요한 모든 통계지표의 작성이 가능한가?
- ② 통계 개발·개선 계획은 적절한가?(조사목적, 조사대상, 주요 조사사항, 작성주기 등)
- ③ 개발·개선되는 통계지표의 공표 시기는 적절한가?
- ④ 필요한 통계지표를 생산하고 있는 기존의 통계조사와 중복되지 않는가?
- ⑤ 개발·개선 계획이 다른 기관에서 이루어져야 하는 경우, 해당기관과 충분히 협의되었는가?

여기서 위의 다섯 가지 검토사항들을 만족시킬 수 있는 정책기반 통계를 개발하기 위해 통계 작성과정별로 마땅히 고려해야 할 기본 사항들을 정리해보자.

통계의 개발이나 개선 계획을 세우기 위해 가장 먼저 해야 할 일은 정책목적을 이루기 위해 필요한 모든 통계들을 망라(list up)하는 작업이다. 현재 관련 통계가 작성되고 있느냐 하는 것과는 별도로 기존에 작성되지 않는 것이라고 해도 필요한 통계를 다 열거하는 것이 필요하다.

다음으로는 열거된 통계지표들을 구체적인 내용에 따라 구분한 후, 필요로 하는 통계가 이미 작성되고 있는 공식통계나 행정기록 등으로 존재하는지 파악해야 한다. 이를 통해 새롭게

개발해야 할 통계가 무엇이며, 기존에 작성되는 통계 중 개선이 필요한 것이 무엇인지 구체화시키는 작업도 이루어져야 할 것이다.

기존에 작성되는 공식통계가 없거나 행정기록으로 파악할 수 없어서 부득이하게 새로운 통계를 개발해야 할 경우가 있다. 이 경우 전반적인 통계작성과정을 몇 단계로 구분하여 각 단계별 고려사항을 체계적으로 정리할 필요가 있다. 「통계행정편람」(통계청, 2008)에는 통계작성과정을 크게 ① 조사의 기획 및 설계 ② 자료수집 ③ 자료처리 ④ 자료의 정리 및 공급 단계로 구분하고 있다. 여기서는 이를 조사의 기획 및 설계, 자료수집, 자료처리와 자료 분석 및 공표라는 세 단계로 나누고 각 단계별 주요 고려사항들을 살펴보기로 한다.

### (1) 조사의 기획 및 설계

조사의 기획이란 조사의 목적을 정하고 그 목적을 달성할 수 있는 조사의 구조와 규모를 결정한 다음, 그에 따른 조사 시스템 구축과 인원 및 예산 확보 등 통계생산을 위한 제반 기반을 구축하는 일련의 과정을 포괄한다.

조사목적에 따라 조사의 방향이나 규모, 구조, 방법 등이 달라질 수 있기 때문에 명확하고 구체적인 조사목적의 세우는 일은 매우 중요하다. 조사목적의 정하기 위해서는 먼저 가급적 다양한 이용자들로부터 정확한 통계 수요를 파악하는 작업이 선행되어야 한다. 현실적인 여건을 감안하여 통계 수요 중 어디까지 충족시킬 것인지에 대한 한계를 명확히 하는 것도 필요하다.

조사목적이 정해지고 나면 다음으로 조사의 구조(structure), 규모, 조사사항, 방법 등을 결정해야 한다. 가능한 한 기존의 행정자료 등을 이용하여 비용을 줄여야 하며, 기존 자료를 활용하기 어려운 경우에는 추가적인 조사를 검토해야 한다. 여기서 조사의 구조란 계속조사 여부와 조사의 빈도 등을 의미한다. 조사규모는 전수조사나 표본조사나에 따라 달라진다. 표본조사일 경우, 조사규모는 통계의 작성범위와 밀접한 연관을 지니고 있다. 가령 전국 통계만 작성할 때와 소지역 단위의 통계까지 작성할 때의 조사규모는 매우 다르다. 조사사항을 정할 때에는 표준적인 개념 및 정의와 분류기준을 고려해야 하며, 현실적인 조사가능성을 반영해야 한다. 아울러 모든 항목들을 빠짐없이 망라할 수 있도록 신중을 기해야 한다. 조사방법을 선택하기 위해서는 조사목적, 경비, 시간, 조사여건 등을 감안해야 한다.

조사표 설계, 표본설계, 결과표 양식설계 등도 기획 단계에서 고려해야 할 중요한 사항들이다. 이 부분들은 모두 전문적인 이론에 따라 매 과정 엄밀한 검토가 이루어져야 하는 부분이다. 실제 이런 전문적인 부분들을 제대로 이해하지 못한 채 다른 조사를 흉내 내어 조사를

기획했다가 통계의 신뢰성 문제에 봉착한 사례들이 과거에 비일비재하였다. 예를 들면, 조사표의 표현이 잘못되어 객관성을 잃는 경우, 주요 조사항목을 누락한 채 조사표를 작성한 경우, 제대로 된 표본설계 없이 자의적으로 표본을 추출하여 조사한 경우, 조사의 수준을 반영하지 않은 채 무분별하게 통계표를 작성하는 경우 등등이 그러한 경우들이다.

## (2) 자료수집

통계조사의 기획과 설계가 끝나면 실제 자료를 수집하는 단계에 들어가게 된다. 보고통계의 경우 적절한 절차를 거쳐 행정자료를 수집하게 될 것이고, 조사통계인 경우에는 조사원이 조사대상자로부터 정보를 수집하는 과정을 거치게 될 것이다. 아무리 조사의 기획과 설계가 완벽하게 되었다고 해도 궁극적으로 통계의 품질을 좌우하는 것은 개별 자료의 정확성이라고 할 수 있으므로 정확한 자료를 얻기 위해 세심한 정성을 기울여야 한다.

조사통계에서는 자질 있는 조사원 선발, 조사원 훈련, 체계적인 조사 관리와 통제, 무응답 오차를 비롯한 비표본오차를 최소화할 수 있는 대책 등을 마련해야한다. 그 중에서도 특히 조사원의 자질과 훈련이 중요한데 숙련되지 못하고 무능한 조사원의 조사는 조사 내용을 부실하게 만들 뿐 아니라 응답자의 비협조적 태도를 조장하게 되어 앞으로의 조사에도 악영향을 끼치게 되기 때문이다.

조사원이 실사수행에서 가져야 할 필수적인 자세로 ① 해당 조사에 필요한 충분한 지식과 성의를 가질 것 ② 훌륭한 자질과 태도를 구비할 것 ③ 상황변화에 대응할 수 있는 기술적인 조사방법을 숙지할 것 ④ 조사대상처에 대해 필요한 지식을 가질 것 ⑤ 응답자와 유기적인 유대를 강화할 것 등이 요구된다(통계청, 2008).

흔히 보고통계인 경우에는 자료수집에 문제가 없을 것으로 여기기 쉽지만 실상은 그렇지 않다. 보고통계를 기획하는 담당자는 먼저 보고되는 최초 원자료의 신뢰성에 관심을 기울여야 한다. 가령, 단속실적을 기초로 보고통계를 작성한다고 하자. 일선의 보고기관에서 무성의하게 실적을 보고한다면 이를 기초로 한 통계는 무의미할 것이다. 그러므로 보고통계를 작성하고자 할 때에는 무엇보다 먼저 최초 자료의 신뢰성을 검토해야 한다. 다음으로는 합리적이고 효율적인 보고체계를 확립해야 한다. 예를 들어, 해마다 일정 시기가 되면 여러 기관으로부터의 다양한 통계작성 요구가 일선 행정기관의 통계담당자에게 몰린다. 이때 자칫하면 사실에 기초한 보고가 이루어지지 않고 전년 자료를 그대로 옮기는 등의 부작용이 생길 수 있다. 또한 자료에 대한 적절한 검증절차가 마련되어 있지 않으면 보고자가 임의대로 처리해버릴 수도 있다. 따라서 사전에 미리 보고체계를 잘 검토하여 정확한 자료수집이 가능하도록 하는 것이 반드시 필요하다.

### (3) 자료처리와 자료 분석 및 공표

자료처리는 개별 수집 자료를 디지털 정보로 저장하기까지의 작업을 포함하며 크게 수집된 자료의 검토와 오류수정, 코딩, 입력, 연산 등의 과정을 포함한다. 이때 불필요한 비표본오차가 생기는 것을 방지하기 위해서는 사전에 충분한 검토를 거쳐야 한다. 원자료에 포함된 오류를 발견하고 적절히 수정하는 절차를 일컬어 자료 편집(data editing)이라고 한다. 자료 편집은 논리적 검증과 통계적 검증으로 나눌 수 있는데, 대규모 조사의 경우 자료 편집을 위한 별도의 시스템을 마련하기도 한다. 자료 편집을 위한 합리적인 전략을 마련하여 합리적이면서도 투명한 오류 수정 절차를 확립하는 것은 자료의 품질을 높이기 위해 중요한 일이다.

자료의 처리가 끝나면 마지막으로 자료에 대한 분석을 한 후 결과를 공표하게 된다. 자료의 분석 및 공표 단계에서는 정확하고 적절한 분석이라는 측면과 이용자 중심의 효과적인 정보 제공이라는 측면을 동시에 고려해야 한다. 정확하고 적절한 분석에만 방점을 둘 경우 이용자에게 불친절하고 무미건조한 정보를 제공할 수도 있으므로 좋은 정보임에도 불구하고 활용도가 떨어질 가능성이 크다. 반면, 정보를 잘 포장하여 전달하는 일에만 방점을 둘 경우 부실한 정보를 잘 포장하여 내세우는 우를 범할 수가 있다. 따라서 분석의 정확성과 이용자 중심이라는 두 가지 토끼를 다 잡을 수 있도록 하는 것이 필요하다.

정확한 자료 분석을 위해 고려해야 할 점으로는 타당한 분석방법의 사용, 조사목적과 한계를 고려한 합리적인 통계작성, 작성되는 다양한 통계의 일치성 확보, 연관되는 다른 통계정보와의 일관성 유지 등이 있다. 자료의 특성에 합당한 분석방법을 사용해야 하는 것은 당연한 일이지만, 간혹 표본조사 자료인데 가중값을 무시한다든지, 무응답에 대한 고려 없이 분석하는 등 잘못 분석하는 경우가 발생하기 때문이다. 한편 통계를 작성할 때 조사 특성이나 자료의 한계를 고려하지 않아서 문제가 생길 때도 있다. 특히 표본조사통계일 때에는 반드시 작성되는 통계의 표본오차를 구하여 검토한 후 오차의 수준에 따라 통계작성 여부를 결정해야 하는데도 이를 무시한 채 무리하게 세분화된 통계를 작성하는 경우가 적지 않다. 심지어 아예 표본오차 계산 자체를 전혀 고려하지 않는 경우도 있는데 이는 힘들여 수집, 작성한 통계의 신뢰성을 스스로 허물어뜨리는 어리석은 일이라 할 것이다. 적절한 이론적 고려를 통해 올바른 통계를 작성했다고 해도 미리 내적 일치성 및 외부 통계와의 일관성을 충분히 점검하여 이상이 없음을 확인한 후 통계를 공표하는 것이 바람직하다. 간혹 작은 실수를 거르지 못한 채 통계가 공표되어 통계의 신뢰성을 떨어뜨리는 경우가 있다.

이용자 중심의 통계제공 서비스라는 측면에서는, 이용자가 통계정보를 잘 이해할 수 있도록 쉽고 간편하게 전달하는 것과 통계 활용에 있어서 필요한 다양한 정보들을 충분히 제공하

는 것, 이용자가 필요한 정보를 문서, 디지털 DB, 웹 등 다양한 형태로 활용할 수 있도록 접근성을 높이는 것 등을 생각할 수 있다. 구슬이 서 말이라도 꿰어야 보배라고 했듯이 작성기관에서 수집한 1차 자료를 충분히 분석하여 효용성 높은 정보로 가공하려는 노력이 필요하다.

## ■ ■ 요약

\_ 정책과정은 정책의 입안, 집행, 평가로 구분되는데, 중요한 도구 중에 하나인 통계와 통계적 기법을 정책과정의 단계별 특성과 연계시켜 살펴보았다.

\_ 정책 입안 단계에서는 사회문제 인지, 대안의 비교, 집행 단계에서는 집행 정도의 측정, 자원 배분을 위한 자료분석, 진척률 지표, 평가단계에서는 평가기준 마련, 평가방식 결정 등을 위해 통계적 방법이 활용된다.

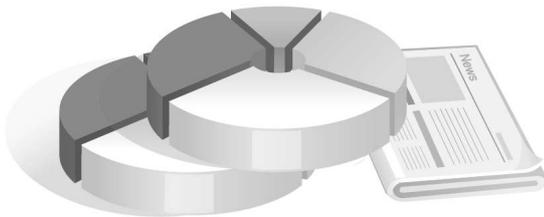
\_ 통계를 이용한 나이팅게일의 의료개혁, 부스의 빈민조사로 인한 빈곤정책의 변화 등은 19세기 역사에서 통계가 정책과정에 영향을 미친 대표적인 사례이다.

\_ 구체적인 정책 분야에서 통계의 미비, 미활용, 오용 등으로 인해 문제가 발생하는 경우가 많다. 그러므로 정책기반이 되는 통계를 작성하기 위해서는 통계의 기획부터 자료 분석 및 공표에 이르는 각 과정에서 고려해야 할 사항들을 파악하여 사전에 점검하는 것이 필요하다.

CHAPTER 3

# 통계와 정책의 사례 연구

- 3.1 주택 관련 통계 사례
- 3.2 사교육비실태조사 사례
- 3.3 손상감시체계 구축 사례
- 3.4 성과지표 설정 사례
- 3.5 도서관예산안 작성 연습과제
- 3.6 통계 DB 활용 연습
- 3.7 정책과정에서 고려해야 할 통계 사례



## 개 요

몇 가지 구체적인 사례를 통해 정책 문제에서 통계가 어떻게 활용되는지 알아 보고 적절한 통계가 갖춰져 있지 않을 때 생기는 문제점과 올바른 통계의 유용성을 소개한다. 주요 사례로 주택 관련 통계, 사교육비 통계, 보건 통계, 성과지표 등을 다루며, 일반인이 일상 업무에서 통계를 유용하게 활용하는 방법과 국가통계포털의 정보를 활용하는 방법도 연습한다. 마지막으로 과거 각 부처의 실제 정책문제 사례를 살펴보고 정책의 집행이나 평가를 위해 필요한 통계나 통계지표를 도출하는 연습도 한다.

## 주요내용

### 3.1 주택 관련 통계 사례

- 【토의1】 주택공급 정책수립을 위한 통계지표
- 【토의문제1】 토지가격과 주택가격의 변동률
- 【토의2】 통계표를 보고 문제점 찾기
- 【토의3】 통계표를 보고 활용방안을 살펴보기

### 3.2 사교육비실태조사 사례

- 【토의1】 사교육비 실태에 대한 통계 생산의 필요성
- 【토의문제1】 다양한 기관에서 작성한 통계를 이용할 때의 문제점
- 【토의문제2】 통계청에서 실시하는 사교육비실태조사의 개요 설명
- 【토의2】 사교육비실태조사 통계의 중요 시사점 및 향후 정책 활용방안
- 【토의문제3】 국가통계포털인 KOSIS를 통한 사교육비실태조사 통계 활용 시 유의점

### 3.3 손상감시체계 구축 사례

- 【토의1】 손상 단계별 통계작성 현황 파악
- 【토의2】 합리적이고 객관적인 조사체계를 구축하는 단계별 과정
- 【토의3】 손상감시체계 통계의 특징
- 【토의4】 손상통계가 정책에 활용된 사례 찾기

### 3.4 성과지표 설정 사례

- 【토의문제1】 사업의 정의와 부합되는 사업명 제시하기
- 【토의1】 복지 분야의 재정사업의 성과목표와 성과지표 제시하기

### 3.5 도서관예산안 작성 연습과제

- 3.5.1 연습과제 개요
- 3.5.2 예산편성 지침
- 3.5.3 과제 풀이

- 【토의1】 태평시 도서관의 2008 회계연도 예산총액은?
- 【토의2】 태평시 도서관의 2009년도 예산안을 제시하기

### 3.6 통계 DB 활용 연습

- 【토의1】 1980년부터 2005년까지의 주택보급률과 우리나라 인구 천 명당 주택수의 변화 정리
- 【토의2】 1990년 이후의 대전시 구별 주택보급률 및 인구 천 명당 주택수 통계 작성 및 비교

### 3.7 정책과정에서 고려해야 할 통계 사례

- 【사례1】 음식점 원산지 표시 정책
- 【사례2】 중소기업매출채권보험 제도와 중소기업 가업승계 지원 정책
- 【사례3】 수도권 이외 지역의 회원제 골프장에 대한 취득세 등 지방세 지원
- 【사례4】 사이버대학 설립·운영정책
- 【토의사례1】 가족친화 사회 환경 조성정책
- 【토의사례2】 농업기계용 면세석유류 공급정책

앞에서 우리는 통계가 갖춰져 있지 않거나 이미 갖춰진 통계를 활용하지 못하거나 잘못 활용하여 생기는 정책과정상의 문제점들에 대해 간단히 살펴보았다. 아울러 우리는 정책의 기반이 되는 통계를 작성하기 위해 고려해야 할 점들을 간략히 정리한 바 있다. 이 장에서는 보다 구체적인 몇 가지 사례를 통해 정책 문제에서 통계가 어떻게 활용되는지 알아보고 적절한 통계가 갖춰져 있지 않을 때 생기는 문제점과 올바른 통계의 유용성을 집중적으로 살펴볼 것이다. 이를 위해 네 가지 분야의 사례들을 살펴보자.

우리나라 사람들 대부분이 연관되어 있고 또 많은 관심을 보이는 분야로 부동산과 교육 문제를 들 수 있다 따라서 주택 관련 통계와 사교육비실태조사 사례를 먼저 살펴볼 것이다. 주택통계 사례에서는 단일지표가 아닌 복합지표의 필요성과 정책 목표에 합당하게 기획된 통계의 중요성을 강조할 것이다. 또 사교육비통계에서는 새로운 통계의 필요성을 제시하고 체계적인 표본조사 체계를 구축할 때 고려할 점들을 점검한다. 다음으로 질병관리본부의 손상감시체계 구축 사례를 살펴보는데 이 사례는 서로 다른 기관에서 작성되는 다양한 통계들을 종합하여 하나의 체계를 구성하는 과정을 보여줄 뿐 아니라 실제 이를 이용하여 구체적인 정책을 세워나가는 과정을 보여주므로 정책과 통계의 관계를 이해하는 데 모범적인 사례라고 할 수 있다. 또 하나의 사례는 성과지표 관련 사례들인데, 바람직한 성과지표의 조건이 무엇인지 실제로 이해하는데 도움이 되는 사례이다.

또한 이 장에서는 예제를 통해 가상적인 예산안을 기안해보고, 예제와 국가통계포털에서 제공하는 통계DB를 검색하여 자기 나름의 통계를 작성하는 방법을 제시한다. 가상 도시의 도서관 예산안 작성 예제를 통해서 일반인이 일상적인 업무에서 통계를 유용하게 활용하는 법을 살펴볼 것이다. 또 통계DB 검색 예제에서는 국가통계포털의 정보를 활용하여 각자 맞춤형 통계를 구축하는 연습을 한다. 이러한 연습을 통해 통계란 전문가들만이 다룰 수 있는 특별한 것이 아니라 누구나 쉽고 유용하게 활용할 수 있는 정책의 기반이라는 것을 느낄 수 있을 것이다. 마지막으로 과거 각 부처에서 실제 있었던 정책문제 사례를 제시한 후, 정책의 집행이나 평가를 위해 필요한 통계나 통계지표를 도출하는 연습을 한다.



### 3.1 주택 관련 통계 사례

부동산 문제는 우리나라에서 국민 대부분의 관심이 집중된 분야라고 해도 과언이 아니다. 우리나라에서 주택은 주거공간이라는 역할만 하는 것이 아니라 ‘부동산 불패’라는 표현에서 보듯이 중요한 재테크 수단이기도 하다. 지난 40여 년 동안 우리나라는 높은 경제성장과 더불어 급속한 인구증가와 도시 팽창을 경험하였다. 그리고 그에 따른 주택공급 확대는 부동산투기로 이어져 적지 않은 사회적 물의를 불러일으켰다.

이제 주택정책과 관련하여 가장 기초가 되는 주택보급률 통계와 주택가격동향지수(작성기관: 국민은행)를 예로 들어, 이 통계들이 실제 정책과정에서 어떻게 활용되어 왔으며 필요에 따라 어떻게 개선되어 왔는지 집중적으로 살펴보자.

#### 토의 1

우리나라 주택 수의 현황을 파악하고, 향후 주택공급 정책 수립을 위한 통계지표를 만들어보자. 또한 단 하나의 지표에만 의지하게 될 때 나타나는 문제점을 살펴보자.

이 질문은 주택정책 담당자에게 있어서 가장 중요하고 기본적인 질문이라고 할 수 있다. 이 질문에 대한 합리적이고 객관적인 해답을 얻기 위해서는 무엇보다 먼저 적절한 주택 관련 통계의 증거를 찾는 것이 필요하다. 즉 우리나라 주택공급 현황을 나타내는 적절한 지표들을 찾고, 가능하다면 다른 나라의 사례도 수집하여 우리나라의 객관적인 위치를 파악해야 한다. 먼저 과거 주택정책 담당부서에서 실제로 해왔던 일을 살펴보고 이 문제를 풀어보자.

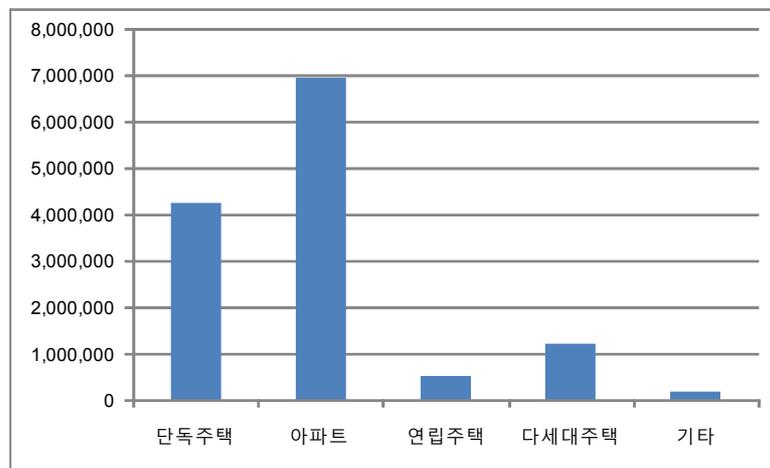
우리나라의 주택 수 현황을 살펴보기 위해서는 통계청에서 5년마다 정기적으로 실시하는 인구주택총조사 결과를 활용하면 된다. 국가통계포털인 KOSIS(<http://www.kosis.kr>)에 접속하면 주택총조사 결과를 구할 수 있다. 다음의 <표 3-1>과 <그림 3-1>은 2005년 주택총조사 결과에 나타난 주택종류별 세대수 현황을 나타내고 있다.

2005년 현재 주택수는 약 1,300만호인데, 아파트가 52.7%, 단독주택이 32.2%, 다세대주택이 9.3%, 연립주택이 4.2%를 차지하고 있다.

〈표 3-1〉 2005년 주택총조사 결과 주택유형별 주택호수

주택유형	단독주택	아파트	연립주택	다세대주택	기타	합계
주택호수	4,263,541	6,962,689	558,513	1,229,208	208,690	13,222,641
비율(%)	32.2	52.7	4.2	9.3	1.6	100.0

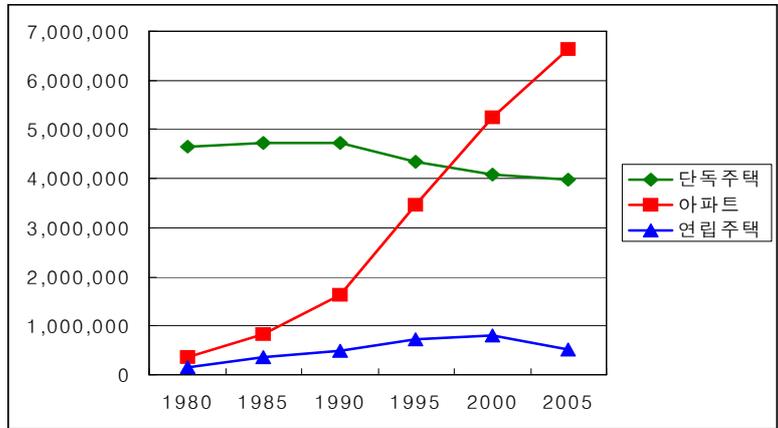
출처 : 통계청 국가통계포털(<http://www.kosis.kr>)



〈그림 3-1〉 주택유형별 주택호수(2005년 주택총조사)

한편 <그림 3-2>는 1980년부터 2005년까지 주택종류별 주택 수의 비율이 어떻게 바뀌어왔는지 보여주는 그림이다. 이 그림을 보면 아파트의 비율이 두드러지게 늘어난 반면 상대적으로 단독주택의 비율이 줄어들고 있음을 알 수 있다. 1990년까지만 해도 우리나라의 가장 대표적인 주택종류가 단독주택이었는데 1995년부터 아파트가 단독주택의 비율을 초과하기 시작하였다. 이 그림을 통해 우리나라의 주거형태가 급격하게 변해가고 있다는 사실을 능히 짐작할 수 있다. 따라서 주택정책에 있어서도 이러한 경향이 반영되어야 할 것이다.

그런데 이와 같은 주택 수 현황만으로는 주택의 부족 여부를 알기 어려우므로 이를 주택공급 정책을 위한 직접적인 지표로 삼기는 어렵다. 그래서 주택수와 가구 수를 동시에 고려하는 주택보급률 통계를 고려하게 된다. 우리나라 주택 공급정책에서 가장 중요하게 활용되어 온 주택보급률 통계를 살펴보자.

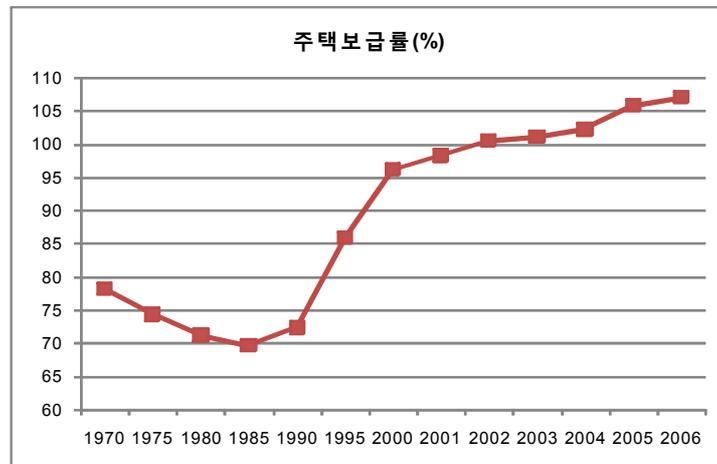


〈그림 3-2〉 연도에 따른 주택유형별 주택수의 변화

지난 1970년대 초 주택문제를 국가적인 계획의 차원에서 해결하려고 시도한 이래 오늘날에 이르기까지 주택정책의 중요한 목표는 주택 건설을 확대하는 것이었다. 주택 공급 확대의 논리는 가구 수에 비해 주택 수가 턱없이 부족하므로 궁극적으로 1가구 1주택을 달성하기 위해서는 주택을 많이 건설해야 한다는 것이었다. 이때 근거로 제시되는 통계가 바로 주택보급률이었다.

주택보급률이란 보통가구 수에 대한 주택 수의 백분율(주택 수 / 보통가구 수 × 100)로 정의된다. 여기서, 보통가구 수란 일반가구 수에서 비혈연가구와 1인 가구를 제외한 가구 수로 정의되며, 주택 수는 인구주택총조사 결과를 기준으로 빈집을 포함하여 산정되는데 다가구주택은 1호로 계산된다. 주택보급률은 특정국가 또는 특정지역에 있어서 주택 재고가 그곳에 거주하고 있는 가구들의 수에 비하여 얼마나 부족한지 또는 여유가 있는지를 총괄적으로 보여주는 양적지표라고 할 수 있다.

<그림 3-3>에서 1970년대 이후의 주택보급률 현황을 볼 수 있다. 주택보급률은 지속적으로 낮아지다가 1980년대 후반에 이르러 최저를 기록하였다. 당시 이러한 주택부족은 부동산 가격 상승과 맞물리게 되어 부동산가격의 급등으로 이어졌다. 따라서 90년대 이후 주택 200만호 건설 등 대대적인 주택공급 확대 정책이 마련되었고, 그로 인해 이후 지속적으로 주택보급률은 높아지게 되었다. 급기야 2002년부터는 전국의 주택보급률이 100%를 넘어서게 되었다. 그러나 지역별로 살펴보면 상대적으로 주택의 수요가 많은 수도권과 일부 광역시의 경우 아직도 주택보급률이 100%에 미치지 못하여 나머지 지방과는 다른 양상을 보여준다.



〈그림 3-3〉 우리나라의 주택보급률의 변화

참고로 주요 선진국들의 주택보급률을 살펴보면, 일본 109.3%('03년), 미국 108.5%('03년), 프랑스 120.5%('99년), 독일 100.6%('02년)로 우리나라보다 대체로 약간 높은 편이다. 현재 정부에서는 장기 주택종합계획('03~'12년)에 따라 연평균 50만호 건설을 통해 2012년에는 주택보급률을 선진국 수준인 116.7%로 높일 계획을 가지고 있다.

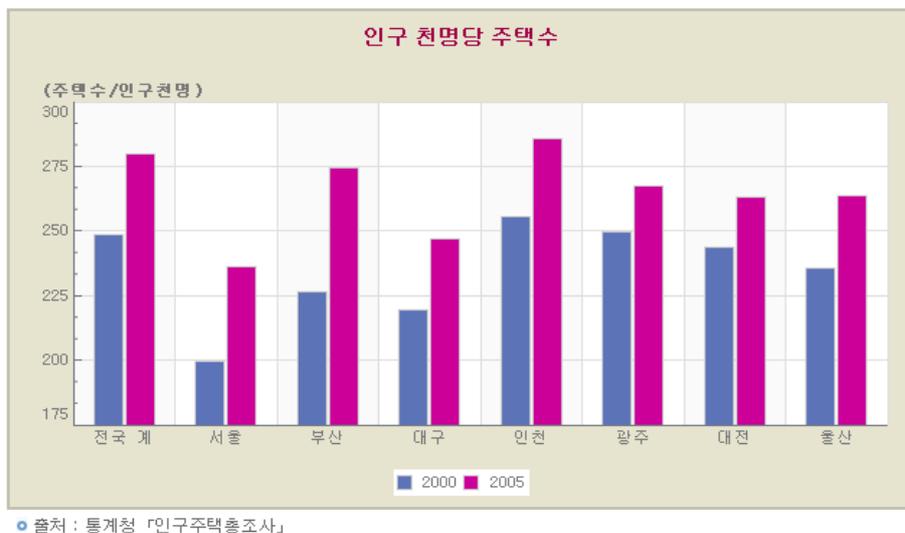
우리나라의 주택보급률이 100%를 넘어서게 되면서 주택보급률과 같은 단일지표를 지향하던 주택정책에 대한 이의가 제기되기 시작했다(진미윤, 한수진, 2003; 건설교통부, 1999). 주택은 많아졌지만, 주택을 자력으로 구입하기 어려운 계층은 여전히 존재한다. 또한 소득계층별로 주택구입능력과 주거소비에는 큰 격차가 존재하는 등 여러 복잡한 양상이 있음에도 불구하고 기존의 주택보급률이라는 지표는 이러한 주거문제의 실상을 제대로 반영하지 못하고 있다는 것이었다. 실제 주택재고의 절대부족문제가 해결된 선진국에서도 지속적인 정책지표로는 주택보급률을 사용하지 않고 분석목적과 대상에 따라 국지적, 부분적, 간헐적인 주택보급률을 계산하여 활용하고 있다. 이에 따라 2005년 행정자치부에서는 정부의 부동산 대책 발표를 앞두고 새로운 기준의 주택보급률을 계산한 결과 63%밖에 되지 않았다고 발표했다. 이 값은 건설교통부에서 발표한 주택보급률 102.2%와 크게 달랐기 때문에 소위 '고무줄 통계' 논란을 불러일으키기도 했다(조선일보 2005년 8월 30일자 만물상 기사 참고).

수십 년 동안 전가의 보도처럼 주택보급률이라는 단일지표를 주택정책의 기초통계로 사용하다가 문제점이 드러나자 갑자기 새로운 통계를 내세웠으니 이러한 논란이 생기는 것은 어쩌면 당연하다고 할 것이다. 이는 정부가 만드는 통계는 정책을 합리화하기 위해 이현령비현령(耳懸鈴鼻懸鈴) 식으로 만든 것이라는 의심을 불러일으키기에 충분하였다. 정책 당국자로서는 평소에 우리나라 주택시장의 현황 및 특성을 제대로 반영할 수 있는 다양한 지표들을

구축하여 여러 신호들을 종합하여 정책을 수립하려는 노력을 기울였어야 했던 것이다.

주택보급률 통계가 가지는 이런 한계를 인식하여 국토해양부에서는 최근 인구 천 명당 주택 수라는 새로운 지표를 개발하여 발표하고 있다. 인구 천 명당 주택 수는 주택보급률과 마찬가지로 주택보급에 대한 양적 지표이지만 가구가 아닌 1인별 주택보급을 측정한다. 종래의 주택보급률이 갖는 한계는 1인 가구나 비혈연가구 등을 제외시킨 점, 다가구주택을 1호로 계산하는 점, 오피스텔을 제외시키는 점 등인데 새로운 지표를 사용하면 이러한 문제를 보완할 수 있다. 그리고 가구 수보다는 인구수를 측정하는 것이 쉬우므로 이 지표로 주택의 양적수급 여건 변화를 보다 정확하게 측정할 수 있다.

<그림 3-4>는 주요 도시의 인구 천 명당 주택 수를 나타내고 있다. 전국 평균 279.7호에 비해, 서울(236.4호), 대구(247.1호), 수도권(254.0호) 지역이 낮은 편이다. 반면, 다른 지역들은 전국의 평균과 비슷하거나 평균을 상회하고 있다. 지역에 따라 주택의 공급이 비교적 큰 차이를 보이므로 향후 주택공급 정책을 수립할 때에는 지역적 특성을 고려하는 것이 반드시 필요할 것이다.



<그림 3-4> 주요 도시별 인구 천 명당 주택 수

다음은 주요 국가별로 비교해보자. 2005년 현재 우리나라의 인구 1000명당 주택 수는 279.7호인데 비해, 일본 423호('03년), 영국 417호('01년), 스페인 462호('01년), 독일 445호('00년) 등으로 나타났다 (<그림 3-5> 참조). 이를 통해 우리나라의 주택 수는 아직 주요 선진국들에 비해 상당히 적은 편임을 알 수 있다.

이상의 예에서 보듯이 하나의 지표는 현상의 한 측면만을 보여준다. 따라서 주택공급 현황

을 파악하기 위해서는 주택보급률과 같은 하나의 지표만 가지고 단면적으로 보는 것보다는 다양한 각도에서 구한 여러 지표들을 복합적으로 고려하여 입체적으로 파악하는 것이 바람직하다. 이때 국제적인 비교를 고려한다면, 국제적으로 통용되는 표준화된 지표, 용어 및 개념을 사용해야 한다.



〈그림 3-5〉 주요 국가들의 인구 천 명당 주택 수



토의 2

〈그림 3-7〉은 1987년 당시 주택은행이 공식적으로 공표하기 시작한 주택가격동향지수 통계표의 일부를 나타내는 그림이다. 이 통계는 전국 약 2,500개 표본주택을 대상으로 부동산중개업소를 통한 가격조사 결과를 근거로 작성된 통계이다. 이때 이 통계가 지니는 문제점이 무엇인지 지적하라.

1. 都市別 地域別 賣買價格指數

(單位:指數, '85.12月末 = 100.0)

地域	總 合				單 獨				聯 立				아 과 트			
	前月	今月	增 減													
	(7.15)	(8.15)	比前月 比前年末													
全 都 市	98.3	98.2	△0.1 △1.0	98.6	98.4	△0.2 △1.6	97.1	97.0	△0.1 △3.0	97.2	97.2	0 △2.8				
6 大 都 市	98.2	98.1	△0.1 △1.9	98.5	98.4	△0.1 △1.8	97.3	97.2	△0.1 △2.8	97.2	97.1	△0.1 △2.9				
서 京	97.5	97.5	0 △2.5	97.9	97.9	0 △2.1	96.2	96.0	△0.2 △4.0	97.2	97.2	0 △2.8				
江 北	98.5	98.5	0 △1.5	98.9	98.9	0 △1.1	94.8	95.0	0.2 △5.0	97.6	97.6	0 △2.4				
江 南	96.6	96.6	0 △3.4	96.3	96.2	△0.1 △3.8	97.2	96.7	△0.5 △3.3	97.1	97.1	0 △2.9				
地方 5 大 都 市	99.2	99.0	△0.2 △1.0	99.5	99.3	△0.2 △0.7	99.5	99.5	0 △0.5	97.3	96.8	△0.5 △3.2				
釜 山	99.4	99.3	△0.1 △0.7	100.2	100.2	0 0.2	98.6	98.6	0 △1.4	97.1	96.4	△0.7 △3.6				
大 邱	99.7	99.7	0 △0.3	99.7	99.7	0 △0.3	101.7	101.7	0 1.7	99.4	99.4	0 △0.6				
仁 川	96.8	95.9	△0.9 △4.1	97.3	96.5	△0.8 △3.5	98.8	98.8	0 △1.2	95.4	94.2	△1.2 △5.8				
光 州	99.1	99.1	0 △0.9	99.1	99.1	0 △0.9	99.0	99.0	0 △1.0	98.2	98.8	0.6 △1.2				
大 田	99.2	99.2	0 △0.8	99.5	99.5	0 △0.5	100.5	100.5	0 0.5	96.6	96.6	0 △3.4				
31 個 中 小 都 市	98.5	98.2	△0.3 △1.8	98.7	98.4	△0.3 △1.6	96.7	96.6	△0.1 △3.4	97.0	97.5	△0.1 △2.5				
京 畿 地 域	95.0	95.2	△0.8 △4.8	96.1	95.0	△1.1 △5.0	95.0	95.1	0.1 △4.9	96.5	96.0	△0.5 △4.0				
水 原	99.7	99.7	0 △0.3	99.8	99.8	0 △0.2	100.0	100.0	0 0	99.5	99.1	△0.4 △0.9				
咸 南	99.8	99.5	△0.3 △0.5	99.9	99.6	△0.3 △0.4	96.7	96.7	0 △3.3	98.4	97.6	△0.8 △2.4				
議 政 府	100.4	100.4	0 0.4	100.5	100.5	0 0.5	100.0	100.0	0 0	100.0	100.0	0 0				
安 養	96.7	96.7	0 △3.3	98.1	98.1	0 △1.9	91.8	91.8	0 △8.2	96.5	96.5	0 △3.5				
富 川	98.2	98.0	△0.2 △2.0	97.4	96.9	△0.5 △3.1	98.2	98.2	0 △1.8	102.0	102.0	0 2.0				
光 明	93.7	92.6	△1.1 △7.4	93.8	92.1	△1.7 △7.9	97.6	98.8	1.2 △1.2	87.9	86.4	△1.5 △13.6				
安 山	89.2	86.8	△2.4 △13.2	89.6	86.6	△3.0 △13.4	77.2	76.3	△0.9 △23.7	88.4	88.4	0 △11.6				
江 原 地 域	100.1	100.0	△0.1 0	100.4	100.3	△0.1 0.3	97.3	97.3	0 △2.7	98.8	98.8	0 △1.2				
春 川	99.9	99.9	0 △0.1	100.0	100.0	0 0	96.5	96.5	0 △3.5	99.1	99.1	0 △0.9				
原 州	99.9	99.9	0 △0.1	100.0	100.0	0 0	99.4	99.4	0 △0.6	99.3	99.3	0 △0.7				
江 陵	101.1	100.8	△0.3 0.8	101.3	101.0	△0.3 1.0	96.3	96.3	0 △3.7	98.1	98.1	0 △1.9				
忠 清 地 域	98.1	98.0	△0.1 △2.0	98.3	98.2	△0.1 1.8	99.6	99.6	0 △0.4	95.0	95.0	0 △5.0				
濟 州	94.9	94.8	△0.1 △5.2	94.8	94.7	△0.1 5.3	100.9	100.9	0 0.9	99.4	99.4	0 △0.6				

〈그림 3-7〉 1987년의 주택가격동향조사 통계표 일부

1980년대 말에 부동산 가격 폭등이 사회적인 이슈가 되었을 때 주택은행에서 공표하는 주택가격동향지수 통계의 신뢰성에 대해 이의가 제기되었다. 당시 서울대 통계연구소가 주택가격동향지수 작성체계를 종합적으로 진단하여 다음과 같은 몇 가지 심각한 결함을 밝혀냈다.

첫째, 전문적이고 과학적인 표본설계를 통해 표본을 추출하지 않아 표본의 대표성에 문제가 있었다. 당시 표본이론 지식이 전혀 없는 실무자들이 조사하기 편한 주택을 표본으로 선정하였던 것이다.

둘째, 표본조사를 통한 추정에 의해 가격동향지수가 생산되는데 추정의 오차 즉, 표본오차에 대한 고려가 전혀 없었다. 위의 <그림 3-7>에 나온 통계표를 보면, 37개 도시 각각에 대해 종합, 단독주택, 연립주택, 아파트의 동향지수가 작성되어있다. 그런데 각 도시별 주택유형에 따른 표본 수가 충분치 않은 관계로 표본오차가 너무 커서 통계로 공표하기에는 부적절한 것들이 많이 있었으나 이런 부분을 도외시한 채 통계를 작성하였다. 한 마디로 믿을 수 없는 수치들을 무분별하게 공표하였던 것이다. 일반적으로 통계이용자들은 공표된 통계를 사실로 받아들이는 경향이 있기 때문에 작성기관에서는 통계적인 정확성을 충분히 검토하여 공표범위를 결정해야 할 텐데도 전문성이 결여된 이런 통계가 작성, 공표되었던 것이다.

참고로, 표본조사 통계의 표본오차는 일반적으로 신뢰구간의 형태로 표현되는 경우가 많은데 이에 대한 자세한 설명은 5장에서 다루기로 한다.

토의 3

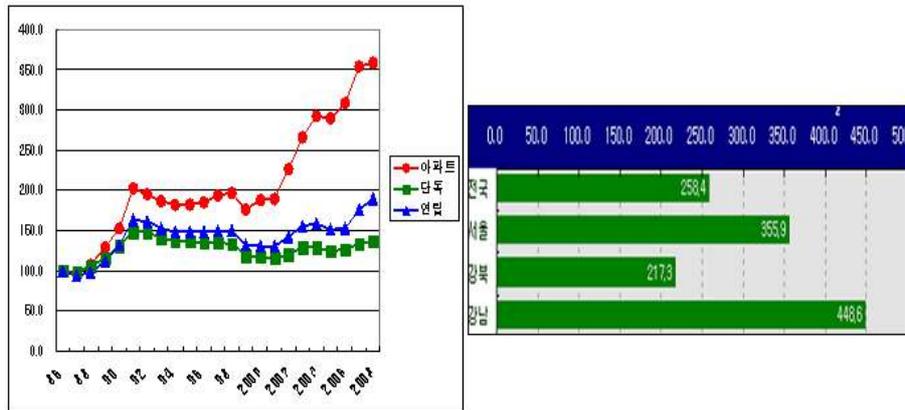
〈그림 3-8〉은 2008년 6월의 주택가격동향지수 중 아파트지수 통계표 일부(국민은행, 2008)를 나타내고 있다. 이 표로부터 전국이나 광역단위 뿐 아니라 시군구별 소지역 통계도 작성되고 있음을 확인할 수 있다. 주택분야의 정책과정에서 이러한 통계가 구체적으로 어떻게 활용되는 지 살펴보자.

구분		아파트						
지역	Classification	금월	전월	전년말	전년동월	증감률(%)		
		This month	Previous month	End of last year	Same month of last year	Changing Rate (%)		
		2008.6	2008.5	2007.12	2007.6	전월비	전년말비	전년동월비
		Over the previous month	Over the end of last year	Over the same month of last year				
전국	Total	103.1	102.6	100.0	99.2	0.5	3.1	4.0
서울	Seoul	105.9	105.3	100.0	98.6	0.5	5.9	7.4
	강북 (14개구) Northern Seoul	110.5	109.6	100.0	96.7	0.9	10.5	14.3
	강북구 Gangbuk-gu	110.2	108.9	100.0	92.5	1.2	10.2	19.1
	광진구 Gwangjin-gu	101.3	101.2	100.0	100.2	0.1	1.3	1.2
	노원구 Nowon-gu	120.9	119.6	100.0	95.0	1.1	20.9	27.3
	도봉구 Dobong-gu	111.3	110.2	100.0	96.4	1.0	11.3	15.5
	동대문구 Dongdaemun-gu	106.8	106.3	100.0	96.5	0.5	6.8	10.7
	마포구 Mapo-gu	102.7	102.1	100.0	99.6	0.5	2.7	3.1
	서대문구 Seodaemun-gu	106.3	105.0	100.0	96.9	1.2	6.3	9.7
	성동구 Seongdong-gu	101.5	101.4	100.0	99.5	0.1	1.5	2.0
	성북구 Seongbuk-gu	108.3	106.8	100.0	97.3	1.3	8.3	11.3
	용산구 Yongsan-gu	102.5	102.2	100.0	98.4	0.3	2.5	4.1
	은평구 Eunpyeong-gu	105.0	103.8	100.0	98.3	1.1	5.0	6.8
	종로구 Jongno-gu	103.6	103.2	100.0	99.2	0.4	3.6	4.5
	중구 Jung-gu	103.0	103.0	100.0	98.3	0.0	3.0	4.8
	중랑구 Jungnang-gu	112.9	111.4	100.0	95.7	1.4	12.9	18.0
	강남 (11개구) Southern Seoul	102.0	101.8	100.0	99.9	0.2	2.0	2.1
	강남구 Gangnam-gu	100.7	101.0	100.0	100.2	-0.3	0.7	0.6
	강동구 Gangdong-gu	99.0	100.1	100.0	100.6	-1.1	-1.0	-1.6

출처 : KB국민은행 전국주택가격동향조사

〈그림 3-8〉 2008년의 주택가격동향조사 아파트 매매가격지수 일부

주택가격지수의 가장 기본적인 역할은 주택가격의 현황을 파악하게 해주는 데 있다. <그림 3-9>의 (1)은 1987년 이후 주택종류별 매매가격지수를 나타내고 있다. 1980년대 말의 가격폭등 이후 1990년대에 접어들면 주택 200만호 건설 등의 영향으로 가격 안정세가 2000년까지 이어진다. 그러나 IMF 외환위기 이후 주택가격이 다시 폭등하는 양상을 보인다. 주택가격의 변동 양상을 주택종류별로 살펴보면 우리나라 주택가격의 폭등은 바로 아파트가 견인하고 있음을 뚜렷이 알 수 있다. 한편 과거 20년 동안의 아파트 가격지수 변동을 살펴보면 전국 평균 가격이 약 2.58배 올랐는데 비해 서울지역은 3.56배가 올랐다. 서울에서도 강북에 비해 강남지역의 주택가격 상승이 두드러져 4.49배가 오른 것으로 나타났다(<그림 3-9>의 (2) 참조).



(1) 주택종류별 매매가격지수의 변화

(2) 20년간 아파트 매매가격지수 상승률

<그림 3-9> 우리나라 주택가격지수의 변화

참고로 효과적인 부동산정책을 수립하기 위해서는 전국적, 일반적 지표 이외에 소지역별, 주택 특성별로 통계를 작성하는 것이 절실하게 필요함을 알 수 있다. 하지만 광역 단위로 작성되던 통계를 소지역 단위로 작성하는 일은 그리 간단한 문제가 아니다. 표본 수를 보면 광역통계를 작성할 때는 3,000개 내외였는데, 소지역통계를 작성하게 되면서 15,000개를 넘게 되었다. 이러한 표본 수의 증가는 일차적으로 조사비용과 시간의 증가를 요구한다.

부동산 정책과정에서 통계가 사용된 또 다른 사례를 들어보자. 2002년 건설교통부와 재정경제부에서는 투기수요를 억제하기 위해 부동산가격이 급등하는 지역에 대해 주택투기과열지구 및 주택투기지구로 지정하는 투기 방지대책을 내놓았다. 이 대책의 골자는 어떤 시군구가 투기지구로 지정될 경우, 해당지역에서는 분양권 전매요건을 강화하고 양도소득세를 실거래가로 과세하는 것이다. 이 경우 과연 어떤 시군구를 투기과열지구나 투기지구로 지정할 것인가

가라는 문제가 따른다. 만일 정부에서 자의대로 투기지역을 지정할 경우 해당지역의 반발이 거셀 것이므로 충분히 수긍할 수 있는 합리적인 기준을 마련하는 것이 절대적으로 필요하다.

다음의 <그림 3-10>에는 주택 투기지역 지정과 관련한 일간신문의 기사 일부가 나와 있다.

경기도 광명시는 이번에 주택 투기지역으로 지정됨에 따라 오는 29일로 예정된공고일 이후 주택을 사고 팔 경우 양도소득세를 기준시가가 마진 실거래가액으로 계산해 신고.납부해야 한다.

주택 투기지역 지정요건은 △전달 주택가격 상승률이 소비자물가 상승률보다 30% 이상 높고 △최근 2개월 평균 상승률이 전국 평균보다 30% 이상 높거나 최근 1년간 상승률이 최근 3년간 전국 평균 상승률 이상이어야 한다.

(서울/연합뉴스)

(<http://hani.co.kr> 2005년 4월 26일 게시 기사)

<그림 3-10> 투기지역 지정과 관련된 신문기사 일부

이 기사를 보면 주택 투기지역 지정요건에 주택가격동향지수가 활용되는 것을 알 수 있다. 2002년에 주택가격동향조사 체계를 대폭 확대 개편하여 구시별 소지역통계 작성이 가능하게 만들었던 목적 중 중요한 것이 바로 이러한 기준을 객관적으로 마련하는 것이었다. 만약 구시별 통계의 오차가 매우 클 경우 이를 기초로 해서 중요한 정책 결정을 내릴 수는 없기 때문에 구시별로 정확도가 높은 통계를 먼저 생산하게 한 것이다.

주택가격지수는 부동산정책의 효과를 파악하는 데에도 유용하게 쓰일 수 있다. 2005년 소위 ‘버블 세븐’지역(강남, 서초, 송파, 분당, 과천, 용인, 평촌, 산본)이 부동산 가격상승의 진원지였는데 이들 지역의 집값 상승을 잡기 위해 소위 ‘8.31 대책’을 시행한 바 있다. 이후 정부에서는 이 대책이 제대로 작동하지 않을 경우에 대비해서 추가적인 대책을 마련해놓고 부동산가격의 추이를 계속 감시하다가 마침내 다음해인 2006년 3월 30일에 추가 대책을 발표하였다.

<표 3-2>는 ‘8.31 대책’ 후 1년, 그리고 ‘3.30 대책’ 후 4개월 동안 버블 세븐 지역별 아파트가격 상승률을 나타낸 것이다. 먼저 ‘8.31대책’ 이후 1년 동안의 상승률을 보면 이 조치가 서울의 강남, 서초, 송파 지역에서는 어느 정도 효과가 있었던 반면 나머지 수도권 지역에서는 제대로 효과를 나타내지 못했음을 알 수 있다. 이러한 통계를 기반으로 정책 당국에서는 추가 대책을 시행하기로 결정하였다. 한편, 후속조치인 ‘3.30 대책’ 후 4개월 동안의 상승률을 보면 7개 지역에서 상승률이 모두 상당히 낮아져 실제적인 효과가 나타나고 있음을 보

여주고 있다.

〈표 3-2〉 부동산 정책의 효과 비교

지역	2005년 8월 30일 ~ 2006년 3월 30일	2006년 3월 31일 ~ 2006년 7월 3일
강남	3.5	-0.3
서초	2.5	-0.4
송파	2.8	-0.6
분당	14.9	-2.7
과천	15.8	3.1
용인	15.9	-3.5
평촌	23.1	1.7
산본	22.9	4.6

출처 : KB국민은행 전국주택가격동향조사



## 3.2 사교육비실태조사 사례

각종 입시와 관련된 과열 과외의 문제는 예나 지금이나 여전한 고민거리이다. 우리나라 교육정책의 역사는 과외와의 전쟁의 역사라는 말이 있을 정도로 교육당국은 과외문제와 씨름해 왔다. 그동안 역대 정부는 과외 문제를 해결하기 위해 다양한 처방을 내놓았으나 근본적인 문제해결은 이루어지지 않아 지금까지도 과외문제는 뜨거운 감자가 되고 있다.

과열과외로 인한 폐해는 말할 필요조차 없을 정도로 심각하다. 국제연합(UN) 아동권리위원회에서는 2003년 “한국의 지나친 조기교육과 입시교육이 아동과 청소년들의 인권을 침해하고 있다”고 지적한 바 있다(국정브리핑 특별기획팀, 2007). 앞서 한국교육재정경제학회에서도 1997년에 낸 <사교육비 문제와 대응방안>이라는 연구보고서를 통해 과열과외의 폐해로 성장기 청소년들의 건전한 성장발달 저해, 학교교육의 비정상화 초래, 가계의 경제적 부담 가중, 사교육시장이라는 지하경제 비대화로 교육재정의 투자왜곡 초래, 사회적 위화감 초래 등을 들었다. 이밖에도 한국교육개발원(2006)은 사교육으로 인해 창의력과 자기주도 학습력 감퇴 증세가 나타남을 실증적으로 보인 바 있다. 이는 우리나라 초중고교 학생들의 학업성취도가 세계적으로 상위 수준인데 반해 대학 경쟁력은 떨어지는 현상으로도 설명된다. 한마디로 우리나라의 비정상적인 과열과외 열풍은 개인이나 사회, 국가적인 낭비를 초래하고 있는 것이다.

2007년 정부는 사교육의 실체를 인정하고 사교육의 뿌리부터 차근차근 되짚어 봄으로써 보다 체계적이고 종합적인 사교육 정책을 마련하기로 하였다. 여기서 통계 기획이라는 관점에서 사교육비 통계조사를 기획, 구축하는 과정에서 검토해야할 점을 살펴보기로 하자.

## 토의 1

우리나라 초·중·고등학생들의 사교육비 실태를 알기 위해서는 많은 예산과 인력이 필요하다. 그럼에도 불구하고 굳이 이 통계를 생산할 필요성에 대한 근거를 제시하라.

비정상적 사교육 열풍으로 인해 발생하는 심각한 문제들을 해결하기 위해 역대 정부에서는 다양한 사교육 관련 정책의 마련에 부심하였다. 매번 사교육 문제가 논란을 일으킬 때마다 정부에서는 갖가지 대책을 내놓았지만 정치적 계산이 가미된 대중요법의 남발로 인해 근본적인 문제를 해결하지 못했다는 지적도 많았다.

사교육이 금지되었던 1980년 이후 2000년까지의 사교육 관련 정책은 초강수로 내린 금지 조치에 대한 부분적인 수정 및 보완의 성격이 짙었다. 그러나 2000년 4월 헌법재판소에서 과외금지조치에 대한 위헌결정이 나면서 사교육 정책은 큰 전환점을 맞이하게 되었다. 학원의 설립·운영 및 과외교습에 관한 법률 개정이 이루어지면서 사교육을 전면적으로 허용하게 된 것이다.

사교육 전면 허용은 사교육 시장을 규제 중심에서 자율시장으로 변화시켰다. 그 결과 서울의 강남 일대는 사교육의 중심지역으로 부상하였고, 그로 인해 부동산 가격 폭등의 진원지가 되는 등 새로운 문제들을 일으키기 시작했다. 학부모의 고학력화, 가계소득 증가, 자녀 수 감소, 맞벌이 부부의 증가, 초등학교 영어교육으로 인한 영어 열풍, 외고나 과학고 등 특수목적 고의 부각 등 우리사회의 급속한 변화와 맞물리면서 사교육 문제는 새로운 양상을 띠게 되었다.

2000년대 이후 새로운 사교육 문제를 풀기 위해 정부에서는 2004년 사교육비 경감종합대책을 발표하였다. 이는 1980년 과외금지조치 이후 지속되던 사교육 정책에 대한 패러다임을 변화시킨 조치라고 할 수 있다. 당시 정부에서 정부의 대책만으로 사교육 수요를 잡기 어렵다는 것을 현실적으로 인정하면서 사교육의 수요를 공교육의 장 안으로 끌어들이겠다는 발상의 전환을 하기 시작했다. 그 결과 EBS 수능강의, 방과 후 학교 등의 정책이 시행된 것이다.

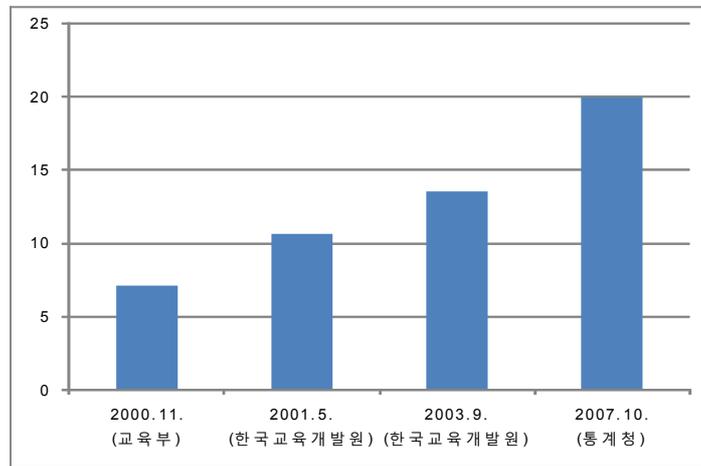
하지만 정부에서 시행한 여러 정책들의 효과를 평가하고 추가적인 대책 마련의 필요성을 가늠하기 위해서는 객관적인 증거로서의 통계가 절대적으로 필요한데도 2007년 이전까지는 정부 차원에서 일관된 기준에 따른 사교육비 조사가 정기적으로 실시되지 않았다. 물론 일부 기관들이 특별한 목적을 위해 일회성 조사를 실시한 적은 있는데 조사마다 기준이나 방법, 규모 등이 다르고 표준화되지 않아 객관적인 증거로 활용하기 어려운 실정이다.

정부차원에서도 정책통계의 중요성이 강조되면서 e-나라지표라는 이름 아래 정부의 각 정책을 평가하고 그 성과를 살펴볼 수 있도록 하고 있다. e-나라지표에는 교육 분야의 다수 정책통계가 소개되어 있으나 교육정책 전반을 포괄하지 못하고 있어서 전반적인 개선 및 새로운 통계 개발이 요청되고 있다.

한국교육개발원(2006)에서는 우리나라 교육 통계의 문제점을 연구한 보고서에서 대부분의 통계가 단순 기술통계(descriptive statistics) 수준에 머무르는 것을 지적하고, 기술통계보다는 정책적 수요를 반영하는 새로운 통계를 구축함으로써 통계자료에 근거한 과학적 교육정책이 수립되고 집행, 평가되도록 해야 한다고 강조한 바 있다. 즉 정책의 성과를 평가하고 정책 수립에 활용할 수 있는 고부가 가치 통계로서 정책통계(policy statistics)의 필요성을 역설한 것이다. 아울러 교육 분야의 연구자, 교사, 공무원 등 전문가들을 대상으로 설문조사를 한 결과, 교육 분야에서 가장 시급하게 작성해야 할 필요가 있는 정책통계가 사교육비에 관한 통계라는 응답이 가장 높은 비율을 차지하였다.

### 토의문제 1

정부에서 공식적인 사교육비 통계를 공표하기 전에도 여러 기관에서 수차례 일회성 표본조사를 통해 사교육비 실태 통계를 작성한 적이 있었다. <그림 3-11>은 우리나라의 사교육비 규모에 관한 서로 다른 조사 결과를 나타내고 있다. 2001년 교육부 조사에서는 우리나라 전체의 사교육비 규모가 7.12조, 2001년과 2003년의 한국교육개발원 조사에서는 각각 10.7조와 13.6조, 2007년 통계청 조사에서는 약 20조인 것으로 추정되었다. 이렇듯 다양한 기관에서 각기 조사, 작성한 통계를 이용할 때의 문제점은 무엇인가?



<그림 3-11> 우리나라 전체 사교육비 규모의 추정치

토의  
내용

## 토의문제 2

통계청에서 실시하는 사교육비실태조사(웹사이트 : <http://pedu.nso.go.kr>)의 개요를 설명하라. 사교육비실태조사는 과학적인 표본조사를 통해 이루어지는데, 조사목적, 법적근거, 조사주기 및 조사규모, 사교육비의 개념 및 조사 범위, 조사항목, 조사체계 등을 설명하라.

- ① 조사목적
- ② 법적근거
- ③ 조사주기
- ④ 표본설계(아래의 내용을 알아보자)

1차 추출단위는 학교, 2차 추출단위는 학부모로 한다.

층화변수로는 지역규모(서울, 광역시, 중소도시, 읍면지역)와 학교급(초·중·일반고, 전문고) 등이 사용되며, 층화확률비례추출방법으로 표본을 추출한다. 표본규모는 최종 통계공표 단위별로 일정한 목표오차 수준을 만족시킬 수 있도록 결정되는데, 최종적으로 전국 초·중·고 272개 학교의 학부모 34,000여명이 표본으로 선정된다. 다음의 <표 3-4>에는 2007년 조사의 표본규모가 나와 있다.

- ⑤ 조사방법
- ⑥ 사교육비의 개념 및 조사범위
- ⑦ 조사항목

조사표 설계 단계에서 핵심적인 부분은 조사목적에 이를 수 있는 효과적인 조사항목을 정하는 것이다. 사교육비실태조사의 주요 조사항목은 공통항목과 실태조사 관련 항목으로 구분되는데, 자세한 내용은 <표 3-4>에 있다.

〈표 3-3〉 사교육비실태조사를 위한 표본 수

학교급별	조사대상 학 교 수	조사대상 학 급 수	조사대상자 수	
			1차	2차
합 계	272	1,038	33,891	33,716
초 등 학 교	74	444	13,833	13,767
중 학 교	82	246	8,577	8,557
고 등 학 교	116	348	11,481	11,392
일반고	88	264	8,955	8,899
전문고	28	84	2,526	2,493

출처 : 통계청 사교육비실태조사(<http://pedu.nso.go.kr>)

〈표 3-4〉 사교육비실태조사의 조사항목

구분	조 사 항 목	
공 통	◦ 학생의 인적사항	- 학생의 인적사항(학교명, 학년, 성별) - 학급 내 성적
	◦ 부모 인적 사항 및 가구소득	- 부모의 인적사항 • 부모의 연령(만 나이), 교육정도(최종학력) - 부모의 경제활동 참여여부 - 월평균 가구 소득 - 월평균 소득 중 조사대상 자녀의 사교육비 지출 비중
실 태 조 사  관 련 항 목	◦ 방과 후 학교 비용 (학교 안)	- 초등 방과 후 보육 프로그램 - 특기·적성 프로그램, 기타
	◦ EBS 교재비 및 어학연수비	- EBS 교육방송(EBS 인터넷 강의 포함) 관련 교육비 - 국내 어학연수비(국내 영어캠프·국내 영어마을 등) - 해외 어학연수비(어학연수 목적)
	◦ 일반교과 및 논술 관련 사교육비	- 일반교과 및 논술 관련 사교육비 • 국어, 영어, 수학, 사회·과학, 제2외국어·한문, 논술 - 일반교과 및 논술 관련 사교육을 받은 이유 - 사교육을 받은 1주당 평균시간
	◦ 예체능 관련 사교육비	- 예체능, 취미·교양 관련 사교육비 • 음악, 체육, 미술, 취미·교양 - 예체능, 취미·교양 관련 사교육을 받은 이유 - 사교육을 받은 1주당 평균시간
	◦ 취업 목적 관련 사교육비	- 취업 목적 관련 사교육비 (공업계, 상업계, 농업계, 전산계, 기타분야 등) - 사교육을 받은 1주당 평균시간

출처 : 통계청 사교육비실태조사(<http://pedu.nso.go.kr>)

토의 2

2007년의 사교육비실태조사 통계를 살펴보고 이 통계에서 볼 수 있는 중요한 시사점을 찾아보라. 또한 이 통계가 향후 정책에 어떻게 활용될 수 있을지 설명하라.

사교육비실태조사 결과는 사교육비실태조사 웹사이트(<http://pedu.nso.go.kr>)나 e-나라지표 사이트(<http://www.index.go.kr>), 또는 국가통계포털사이트 등을 통해 찾아볼 수 있다. 통계청의 사교육비실태조사 웹사이트의 자료실에서는 통계조사 결과를 요약한 보도자료와 조사표 등이 제공되고 있다. e-나라지표의 ‘사회 > 교육 > 교육/인적자원일반’ 분야 중 “우리나라 초·중·고 학생의 사교육 현황”에서는 사교육비실태조사 결과 중 전국 종합 통계와 간단한 요약 정보가 제공된다. 한편, 국가통계포털에서는 속성별 요약통계를 DB형식으로 제공해준다.

<그림 3-12>는 국가통계포털 사이트의 사교육비실태조사 관련 화면을 나타낸 것이다. 이 그림을 보면 제공 가능한 정보의 수준이 어느 정도인지 알 수 있는데 학교급별, 지역별, 학생 성적 순위별, 부모 연령 및 교육정도별, 가구 소득별 사교육비 및 방과 후 학교, EBS교재비, 어학연수 비용 등의 통계가 제공되고 있다.



<그림 3-12> 국가통계포털 사이트의 사교육비실태조사 관련 화면

<표 3-5>는 2007년 사교육비실태조사 결과를 정리한 것이다. 전체 사교육비는 약 20조 원에 이르며 학교급별로는 초등학교가 10조 2천억 원으로 가장 많은 것을 알 수 있다. 한편 학생 1인당 월평균 사교육비는 약 22만 원정도이고 전체 학생 중 사교육을 받는 학생의 비율은 77% 정도이다.

〈표 3-5〉 학교급별 사교육비 통계

	총 사교육비 (억 원)	학생 1인당 연평균(만원)	학생1인당 월평균(만원)		참여율 (%)	참여시간 (주당)
			전 체1)	참여자2)		
전 체	200,400	266.4	22.2	28.8	77.0	7.8
초등학교	102,098	272.2	22.7	25.6	88.8	8.9
중 학교	56,120	281.0	23.4	31.4	74.6	8.9
고등학교	42,181	236.8	19.7	35.0	55.0	4.5
-일반고	38,655	288.3	24.0	38.8	62.0	5.2
-전문고	3,526	80.1	6.7	19.8	33.7	2.5

\* 주1) 전체 학생을 대상으로 산출한 월평균 금액 (사교육을 받지 않은 학생 포함)

주2) 사교육을 받은 학생만을 대상으로 산출한 월평균 금액

출처 : 통계청 사교육비실태조사(<http://pedu.nso.go.kr>)

또 <표 3-6>은 사교육비 현황을 다양한 특성별로 구분하여 정리한 결과이다. 교과목별로는 영어와 수학에 가장 많은 비용이 드는 편이며 지역별로는 서울 학생들의 사교육비가 28.4만원으로 읍면지역 학생들의 12.1만원에 비해 약 2.3배로 나타났다. 학생 성적별로는 상위 10% 이내 학생들의 사교육비가 가장 높으며 성적이 떨어질수록 사교육비가 적게 드는 편이다. 부모의 교육수준별 사교육비 지출 양상을 보면 부모의 교육수준이 높을수록 사교육비가 많이 드는 것을 알 수 있다.

〈표 3-6〉 특성별 사교육비 현황 : 교과목별 사교육비

(단위 : 만원, %)

	전 체	일반교과					예체능, 취미, 교양
			국어	영어	수학	기타	
사교육비	22.2	17.8	2.2	6.8	5.7	3.0	4.3
참 여 율	77.0	68.4	39.3	55.6	58.6	38.6	37.0

출처 : 통계청 사교육비실태조사(http://pedu.nso.go.kr)

〈표 3-7〉 특성별 사교육비 현황 : 지역별 사교육비

(단위 : 만원, %)

	전 국	서 울	광역시	중소1) 도시	읍면 지역
사교육비	22.2	28.4	22.0	22.8	12.1
참 여 율	77.0	80.6	79.0	77.5	66.4

\* 주1) 서울 및 6대 광역시를 제외한 일반 시, 洞지역 소재 학교

출처 : 통계청 사교육비실태조사(http://pedu.nso.go.kr)

〈표 3-8〉 특성별 사교육비 현황 : 학생 성적별 사교육비

(단위 : 만원, %)

	전 체	상위 10% 이 내	11~30%	31~60%	61~80%	하위 20% 이 내
사교육비	22.2	30.0	26.6	21.0	15.9	12.0
참 여 율	77.0	89.3	87.0	77.3	62.3	51.2

출처 : 통계청 사교육비실태조사(http://pedu.nso.go.kr)

〈표 3-9〉 특성별 사교육비 현황 : 부모 교육수준별 사교육비

(단위 : 만원, %)

		전 체	초졸이하	중 졸	고 졸	대졸이상
아버지	사교육비	22.2	6.8	9.9	17.1	29.6
	참 여 율	77.0	36.6	50.2	71.7	88.0
어머니	사교육비	22.2	8.4	10.5	19.3	31.1
	참 여 율	77.0	40.5	50.4	75.0	89.6

출처 : 통계청 사교육비실태조사(http://pedu.nso.go.kr)

---

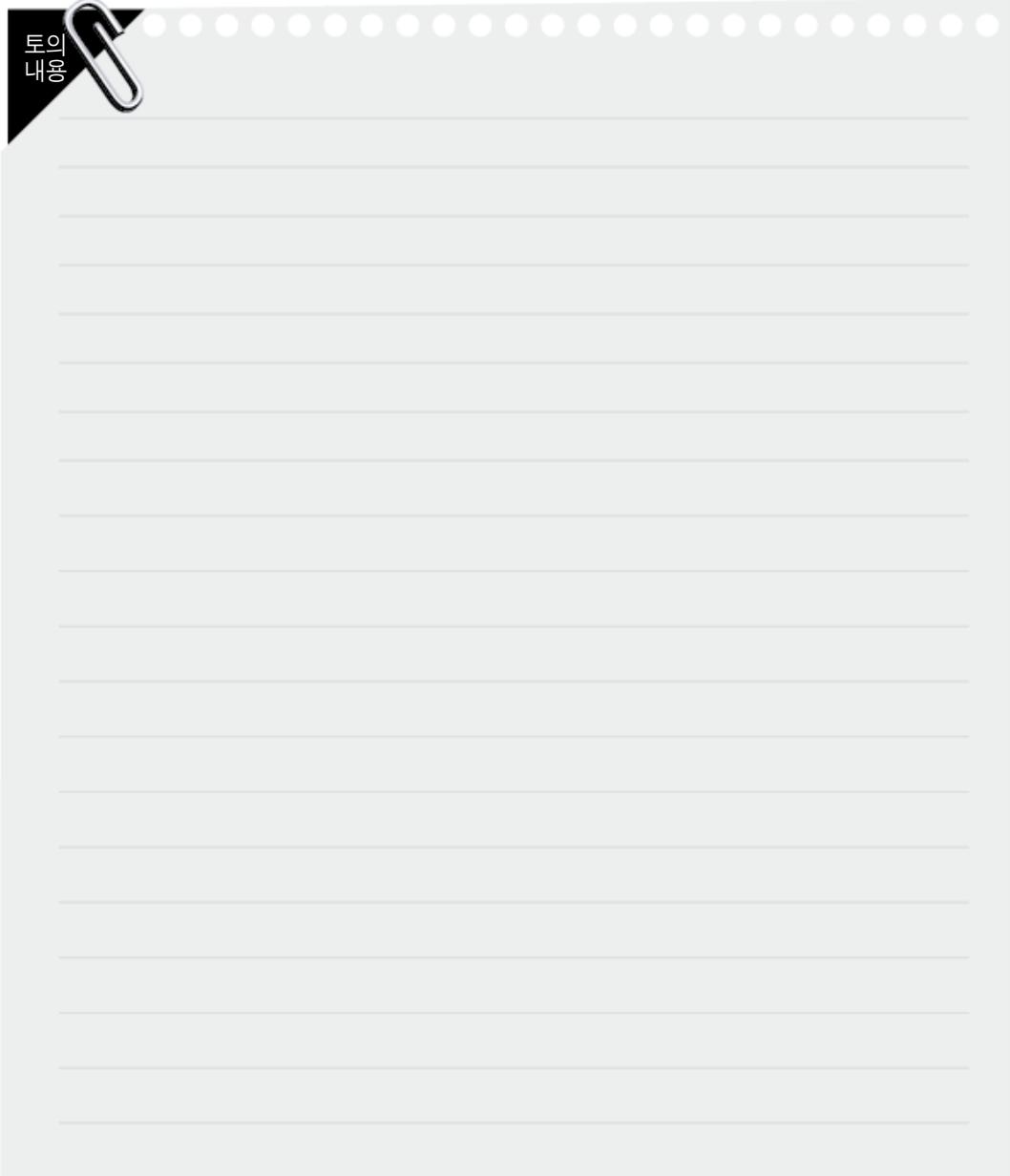
위에서 소개한 사교육비실태조사 결과는 그동안 우리사회에서 막연하게 생각했던 사교육과 관련된 우려가 실제 상황임을 그대로 보여주고 있다. 즉, 서울 및 도시지역과 시골지역 간의 격차, 부모의 교육수준에 따른 사교육 격차가 실재함을 보여준다. 또한 성적이 우수할수록 사교육비가 늘어난다는 사실은 우리나라 사교육이 공교육을 따라가지 못하는 부진한 학생들을 위해서가 아니라 더 좋은 상급학교에 진학하기 위한 입시 위주의 목적으로 활용된다는 점을 시사하고 있다.

사교육비실태조사는 2007년부터 시작되었기 때문에 현재 시계열통계가 제공되고 있지 못한데, 앞으로 조사가 계속되어 지속적인 통계가 구축될 경우 사교육비 문제를 둘러싼 각종 억측이나 비과학적 주장이 사라질 뿐 아니라 물가지수처럼 사교육 관련 지표가 일목요연하게 정리되어 각종 정책들의 효과를 계량적으로 평가할 수 있게 될 것이다.

토의문제 3

국가통계포털인 KOSIS를 통해 사교육비실태조사 통계를 활용하는 데 있어서 주의해야 할 점은 무엇인가?

토의  
내용



참고로 국가통계포털(KOSIS)에서 통계정보를 검색하고 나면 결과 창 윗부분에 통계메타정보라는 단추가 나타난다. 이것을 검색하면 <그림 3-13>과 같이 해당 통계와 관련하여 꼭 알아야 할 중요한 참고사항들을 확인할 수 있다. 이 메타정보 중에는 표본오차 관련 사항도 들어 있다.

The screenshot shows the KOSIS search interface. At the top, there is a search bar with the text '통합검색' (Integrated Search) and buttons for '검색' (Search) and '상세검색' (Detailed Search). Below the search bar are several navigation tabs: '항목 시계열 조회' (Item Time Series Query), '메타자료 비교' (Meta-data Comparison), 'KOSIS 바로가기' (KOSIS Direct Access), '자료 내려 받기' (Download Data), and '인쇄' (Print). The main content area shows search criteria: '통계명: 사교육비실태조사' (Statistic Name: Private Education Expenses Survey) and '기간: 2007년 ~ 2007년' (Period: 2007 ~ 2007). A '2007년' (2007) filter is active. The central part of the page is a table with columns for '조사개요' (Survey Overview), '조사방법론' (Survey Methodology), '자료제공' (Data Provision), '조사표' (Survey Instrument), '응여 해설' (Response Explanation), and '기타' (Other). Each column contains a list of metadata items with radio buttons. For example, under '조사개요', there are items like '통계명', '통계종류', '법적근거', and '조사목적'. Below the table, there is a section for '조사개요' (Survey Overview) with a sub-section for '통계명' (Statistic Name) listing '사교육비실태조사' (Private Education Expenses Survey) and '통계종류' (Statistic Type).

<그림 3-13> 통계메타자료 화면



### 3.3 손상감시체계 구축 사례

어느 국가, 어느 문화에서든 건강은 교육과 더불어 개인의 행복 추구를 위해 중요한 요소이다. 한 개인의 건강은 개인 뿐 아니라 가족에게 있어서도 미래의 발전과 경제적 안정을 위한 역량이 되며 경제적 생산성을 위한 인적자본의 초석이기도 하다.

과거에는 건강 문제를 상당 부분 개인의 책임으로 인식했었다면 최근에는 국민 건강증진 및 질병예방을 위해 국가가 정책을 수립하여 시행해야한다는 요구가 매우 높아지고 있다. 또한 건강을 국가 인적자본의 필수요소로 인식하는 발상의 전환과 함께 건강투자(Health Investment) 전략을 국가적 과제로 채택하고 있기도 하다(한국보건사회연구원, 2007).

보건통계는 국민의 건강수준과 의료 이용 실태를 파악하는 도구로 현대 복지국가에서 국가 정책의 우선순위를 정하고 정책 방향을 설정하는 중요한 기초자료가 되고 있다. 그런 까닭에 선진 각국에서는 이미 수십 년 전부터 과학적인 보건통계 생산을 위한 체계를 구축해왔으며, OECD를 비롯한 국제기구들 또한 국가 간 보건통계의 교류와 생산이 촉진될 수 있도록 다각적인 노력을 기울이고 있다(한국보건사회연구원, 2002와 2007).

여기서는 질병관리본부에서 수행한 국가 손상감시체계 구축 사례를 통해 국민의 질환 예방 관리 및 보건정책 수립에 통계가 어떻게 활용될 수 있는 지 살펴보기로 한다.

## 토의 1

손상(injury)이란 의도적 또는 비의도적 사고로 발생하는 신체나 정신에 미치는 건강상의 해로운 결과를 뜻하는 말이다. 과학적 근거에 기반을 둔 손상예방 정책을 기획하고 평가하기 위해서는 사회 안전 분야를 총괄하는 국가단위의 감시체계가 필요하다. 이를 위해 손상의 단계별 통계작성 현황을 파악하라 (여기서는 손상감시체계가 구축되기 전이던 2005년의 현황을 기준으로 하기로 함).

산업화 및 도시화로 인해 질병구조가 과거 전염병 위주로부터 점차 만성퇴행성 질환과 손상 위주의 질환구조로 변화되고 있다. 그 중 손상은 모든 연령대에서 주요 사망원인이며 특히 가장 활발히 활동하게 되는 청장년층의 사망원인 1위를 차지하고 있다. 손상은 개인의 노동력 상실 및 그로 인한 경제적 손실을 초래하게 되어 삶의 질을 떨어뜨릴 뿐 아니라 국가적으로도 손실을 가져오는 요인이다. 하지만 역학적 모형을 통해 손상의 관련 요인을 찾을 수 있으므로 미래의 위험을 줄이기 위한 예방책을 효과적으로 마련하려면 역학적 감시가 반드시 필요하다. 손상을 효과적으로 예방하면 의료비 절감과 건강수명 연장을 통해 노동생산성을 향상시킬 수 있으며 양극화 시대의 사회적 안전에도 큰 도움이 된다. 그러므로 과학적 근거에 기반을 둔 손상예방 정책을 기획하고 평가하기 위해서는 사회 안전 분야를 총괄하는 국가단위의 감시체계를 구축하는 것이 필요하다(질병관리본부, 2007).

감시체계(surveillance system)란 “공공보건정책을 기획·실행·평가하기 위해 기본적인 자료들을 지속적이고 체계적으로 수집, 분석하며 이를 필요로 하는 사람에게 적시에 배포하는 것과 최종적으로 수집, 분석한 자료를 질병의 예방과 관리를 위해 사용하는 것”이다. 그러므로 손상감시체계는 손상에 관련된 자료를 지속적, 체계적으로 수집, 분석함으로써 손상의 유병분포와 발생 현황을 파악하고 손상의 원인과 의료서비스 이용 등에 관한 신뢰성 있는 지표를 산출하는 체계라고 할 수 있다. 또 관련 전문가나 정책결정자는 손상감시체계를 통해 수집된 자료와 산출되는 지표를 통해 손상현황을 모니터링 함으로써 손상 예방 및 안전 증진 정책을 효과적으로 수립, 실행할 수 있게 된다.

손상감시체계의 틀을 도식적으로 요약하면, 다음의 <그림 3-14>와 같이 손상, 입원, 응급실, 입원, 사망의 단계별 피라미드 형태로 나타낼 수 있다. 손상을 당한 환자들은 증상이 경미할 경우 집안에서 약품을 이용하여 처치할 것이다. 그러나 그보다 증상이 심할 때에는 병원을 이용하게 되는데, 본격적인 손상감시의 대상이 된다. 손상감시체계를 세우기 위해서는 응급실, 입원, 사망의 각 단계별 손상 자료를 확보할 수 있는 방안을 마련해야 한다.



〈그림 3-14〉 손상 피라미드

① **사망** : 손상으로 인한 사망을 감시하기 위해서는 질병관리본부에서 별도의 감시체계를 구축할 필요 없이 기존에 통계청에서 매년 작성, 공표하는 사망원인통계를 활용할 수 있다. 통계청의 사망원인통계에 따르면 손상은 전체 사망원인 중 약 12% 내외를 차지한다. 손상은 40대 이하의 조기사망과 장애의 주요 원인이 되어 생산력 손실 및 막대한 보건의료재원의 소요를 야기하고 있다.

② **입원** : 손상으로 인해 사망에 이르는 극단적인 경우는 일부에 불과하다. 치명상이 아닌 중증의 손상을 입는 환자들의 경우 입원으로 이어진다. 따라서 손상으로 인한 입원환자의 정보를 얻기 위해서는 별도의 조사가 필요하다. 미국의 경우 National Center for Health Statistics(NCHS)가 중심이 되어 보건의료 관련 자료 수집을 위한 다양한 조사가 운영되고 있는데, 각종 질환 현황을 파악하기 위해 의료기관을 대상으로 표본조사를 실시하는 퇴원환자조사(National Hospital Discharge Survey)가 대표적이다. 우리나라에서는 이와 같은 조사체계를 구축할 필요가 있다.

③ **응급실** : 응급실은 모든 종류의 손상이 모이는 곳인데 특히 중증의 손상은 거의 대부분 응급실을 경유하게 되므로 손상감시를 위해 가장 중요한 곳이라고 할 수 있다. 사망원인통계나 퇴원환자조사의 손상 자료만으로는 손상감시체계가 완성될 수 없고 거기에 응급실을 경유하는 손상을 덧붙여야만 비로소 기본적인 감시체계가 이루어질 수 있는 것이다.

중앙응급의료센터에서는 2004년 응급환자진료정보망을 구축한 바 있으므로, 질병관리본부는 이 정보망의 데이터베이스로부터 응급실 방문 손상환자의 현황, 손상발생 기전 및 의도성 등의 정보를 추출함으로써 응급실 관련 손상 정보를 얻을 수 있다. 이 경우 별도의 조사가 필요 없이 기존의 정보망에서 필요한 자료만 추출하여 통계를 작성하면 되므로 매우 편리하다. 현재 116개 권역 및 지역 응급의료센터에서 연간 약 20만 건 가량의 손상자료가 수집되고

---

있다(질병관리본부, 2007).

위의 세 단계를 종합하면, 우리나라 손상감시체계를 구축하기 위해서는 추가적으로 손상으로 입원하는 환자에 대한 통계를 생산할 수 있는 방안을 마련하는 것이 필요하다는 결론에 이르게 된다.

## 토의 2

손상 입원환자 현황을 파악할 수 있는 조사체계를 구축하고자 한다. 합리적이고 객관적인 조사 체계를 구축하는 과정을 단계별로 설명하라.

우리나라의 질병관리본부에서도 손상 및 만성질환자에 대한 통계 작성을 위해 철저한 사전 준비를 거쳤으며 2005년부터 퇴원환자조사 체계를 구축하여 실시하고 있다. 이 예는 하나의 표본조사가 어떤 과정을 거쳐 구축되는지를 보여주는 좋은 사례이다. 새로운 통계조사를 구축할 때에는 통상적으로 ① 사례연구 ② 조사기획 ③ 표본설계 ④ 조사표설계 등의 단계를 거치게 된다. 여기서도 이 단계에 따라 손상통계조사의 구축 과정을 설명한다.

## ① 사례연구

이미 손상감시체계를 구축하여 운영하고 있는 선진국들인 미국, 캐나다, 호주의 사례를 비교하여 각각의 장·단점을 집중적으로 검토하였다. 그 결과, 우리나라 실정을 감안한다면 미국의 퇴원환자조사가 구체적으로 참조할 수 있는 좋은 모델이 됨을 확인하였다.

미국의 퇴원환자조사는, 1965년 이후 매년 시행되는 병원 대상의 조사로, 단기요양병원을 대상으로 의료기관의 규모, 위치, 소유권, 환자의 진단명, 외과수술 및 처치, 재원일수, 환자의 특성 및 치료비 출처 등에 대한 항목을 조사한다. 이 조사는 500여개 병원, 약 30만 건의 의무기록 자료를 대상으로 이루어진다.

## ② 조사기획

조사체계를 구축하기 위해 우선적으로 조사대상을 정하고 그에 따른 모집단을 규정해야 한다. 조사대상으로는 의무기관 즉 전국의 병·의원을 생각할 수 있는데, 병·의원의 종류가 매우 다양하므로 현실적으로 조사가 가능한 범위가 어느 정도인지를 충분히 감안해야 한다. 질병관리본부에서는, 우리나라 보건의료기관을 대표할 수 있도록 단과병원, 요양병원, 노인전문병원, 보훈병원, 국군병원, 재활병원은 제외하고 100병상 이상의 급성기 일반병원(general hospital)을 대상으로 종합병원, 병원, 보건의료원을 포함한 의료기관을 조사대상으로 정하였다.

조사방법은, 조사대상 병원을 대표하는 표본병원을 추출한 후, 표본병원의 퇴원환자기록 중 일부를 추출하여 조사하는 방식을 택하기로 하였다. 의무기록이 전산화된 표본병원에 대

해서는 의무기록사의 협조를 받아 데이터베이스 자료 중 일부를 추출하며, 비전산화 병원에 대해서는 조사원을 직접 파견하여 퇴원기록 파일 중 일부를 직접 추출하여 조사하는 방식을 사용한다.

조사는 매년 1회 실시하지만 1년 동안의 퇴원환자기록 전체를 대상으로 계통추출 한다. 따라서 조사대상 기간은 1년이 된다.

### ③ 표본설계

표본설계에서 고려해야 할 중요한 요소로 대표성(representativeness), 정확성(accuracy), 효율성(efficiency)을 들 수 있다. 대표성이란 표본이 모집단을 잘 대표할 수 있어야 함을 뜻하고, 정확성이란 표본조사를 통해 작성되는 추정의 정도(precision)가 높아야 함을 의미하며, 효율성이란 가능한 한 적은 표본으로 목표정도를 달성할 수 있게 하는 성질을 말한다.

질병관리본부에서는 이러한 요구를 충족시키기 위해 표본이론 전문가에게 의뢰하여 표본을 설계하였다. 그 결과 우리나라 보건의료기관을 대표할 수 있도록 단과병원, 요양병원, 노인전문병원, 보훈병원, 국군병원, 재활병원은 제외하고 100병상 이상의 급성기 일반병원(general hospital)을 대상으로 종합병원, 병원, 보건의료원을 포함하여 시도별, 병상규모별 층화확률추출법에 의해 170여개 병원이 추출되었다. 질병별로 정도 높은 전국 추정이 이루어질 수 있도록 표본규모가 결정된 것이다.

표본으로 선정된 병원의 퇴원환자기록 중 약 10% 정도를 계통추출법에 의해 추출하도록 하였다. 결과적으로 약 20만 건 내외의 기록이 최종 표본 자료로 수집될 것이다.

### ④ 조사표 설계

의무기관을 대상으로 퇴원환자기록을 추출하는 조사를 하기 위해서는 적절한 조사표가 마련되어야 한다. 이를 위해 사전에 우리나라 의료기관별 퇴원환자에 대한 의무기록 정보를 파악하는 작업을 선행한 바 있다(한국보건사회연구원, 2002). 그 결과 의무기록은 의료기관의 규모와 특성에 따라 사용하는 서식의 종류와 서식에 기재되는 항목들이 상이한 것을 알 수 있었다. 의무기록이 전산화된 병원이 있는가 하면 그렇지 못한 병원도 있고, 전산화된 병원이라고 해도 기록양식이 표준화되어 있지 않은 상태임을 발견할 수 있었다. 이런 실태 파악과 더불어 선진국들의 표준적인 퇴원환자조사표 등을 참고하여 먼저 현재 우리나라 실정에 적합한 조사표를 마련하였다.

<그림 3-15>는 현재 질병관리본부의 웹사이트(<http://www.cdc.go.kr>)에 소개된 퇴원환자조사표 일부를 나타낸 그림이다. 조사내용은 환자 정보, 의료기관 정보, 환자 내원 정보, 환자의 질환과 치료 정보 등인데, 손상환자의 경우 추가적으로 손상항목과 사고 유형이나 기전을

문는 항목들이 포함된다.

1. 의료기관기호	일련번호	조사자(성명)	조사일시
			/
기본조사(모든 퇴원 환자) 가. 환자신상정보			
2. 환자등록번호	3. 성별	4. 나이	5. 출생일
	<input type="checkbox"/> 1) 남자 <input type="checkbox"/> 2) 여자 <input type="checkbox"/> 3) 양성 <input type="checkbox"/> 4) 불명	__세	YYYYMMDD(2003.1.10이후)
6. 주소(우편번호)	7. 진료비지불방법		
<input type="checkbox"/> 0) 무료 <input type="checkbox"/> 1) 국민건강보험 <input type="checkbox"/> 2) 의료급여 1종	<input type="checkbox"/> 3) 의료급여 2종 <input type="checkbox"/> 4) 산재보험 <input type="checkbox"/> 5) 자동차보험	<input type="checkbox"/> 6) 공상 <input type="checkbox"/> 7) 일반 <input type="checkbox"/> 8) 기타 <input type="checkbox"/> 9) 불명	
나. 내원정보			
8. 입원일	9. 퇴원일	10. 입원경로	
YYYYMMDD	YYYYMMDD	<input type="checkbox"/> 1) 응급 <input type="checkbox"/> 2) 외래	<input type="checkbox"/> 8)기타(분만장,신생아실) <input type="checkbox"/> 9)불명
다. 질환 및 치료정보			
진단구분	질병코드	진단명	13. 외인코드
11. 주진단코드	1)		
	2)		
	3)		
	4)		
12. 부진단코드	5)		
	6)		
	7)		
	8)		
	9)		
	10)		
14. 주수술코드	주수술명	15. 주수술일	
		YYYYMMDD	
16. 부수술 및 처치코드	부수술명 및 처치명		
1)			
2)			
3)			

〈그림 3-15〉 퇴원환자조사표 일부

### 토의 3

손상감시체계를 통해 작성되는 손상통계를 찾아보고 그 통계에서 관찰되는 특징들을 설명하라.

현재 우리나라에서는 손상으로 인한 사망자 수, 입원 환자 수, 응급실 방문 환자 수에 대한 통계작성 체계가 구축되어 2006년부터 정기적으로 통계가 공표되고 있다. 여기서는 각 단계별로 공표되는 통계들을 간단히 소개한다.

참고로, 손상의 각 단계별로 통계의 특성이 서로 다르다는 점을 주목할 필요가 있다. 즉, 사망자 수와 응급실 방문 환자 수 통계는 전수조사를 통해 얻는 것으로서 보고통계의 특성을 지니는 반면, 입원 환자 수나 외래 내원 환자 수 통계는 표본조사를 통해 얻는다. 전수조사를 통해 계산된 통계는 표본오차를 수반하지 않지만, 표본조사를 통해 얻어진 통계는 참값이 아닌 추측값으로서 표본오차를 수반한다. 그러므로 전수조사 통계는 값 자체를 그대로 받아들일 수 있지만, 표본조사 통계는 오차를 감안하여 받아들일 필요가 있다.

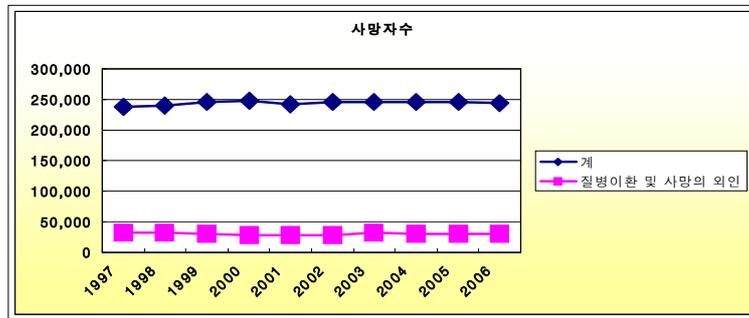
#### ① 손상 사망통계

<표 3-10>과 <그림 3-16>에는 과거 수년 간 우리나라의 사망자 수와 손상 사망자 수가 나타나 있다. 2006년의 경우, 총 사망자 수가 약 24만 4천 명인데 그 중 손상으로 인한 사망자 수는 3만 명 정도로 약 12.3%를 차지하고 있다.

〈표 3-10〉 연도별 사망자 수 및 손상 사망자 수

연도	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
전체	238,714	240,254	246,539	247,346	242,730	246,515	245,817	245,771	245,511	243,934
손상	32,459 (14%)	32,011 (13%)	29,947 (12%)	28,874 (12%)	28,736 (12%)	28,834 (12%)	31,558 (13%)	30,567 (12%)	30,957 (13%)	29,615 (12%)

출처 : 통계청 사망원인통계



〈그림 3-16〉 연도별 사망자 수와 손상 사망자 수

② 손상 퇴원환자 통계

<표 3-11>은 퇴원환자 조사 결과로 얻어지는 대표적인 통계표이다. 2005년 한 해 동안 병원에 입원한 전체 환자의 수는 약 467만 명이며, 그 중 손상 환자는 90여만 명으로 전체 퇴원환자 중 약 19.3%에 이르는 것을 알 수 있다. 사망원인 중 손상에 의한 사망의 비율이 12%인 것에 비하면 손상 환자의 비율이 좀 더 높은 편이다.

〈표 3-11〉 퇴원환자 수 추정값

구 분	전체	남자	여자
전체 퇴원환자(명)	4,674,816	2,405,415	2,269,401
손상 퇴원환자(명)	900,772	548,861	351,911
손상퇴원환자 비율(%)	19.3	22.8	15.5
손상퇴원율(10만 명당)	9,711.3	9,943.5	9,476.7

출처 : 제2차 퇴원환자조사(2006)

한편 <표 3-12>에는 손상 원인별로 환자 수를 추정한 결과가 나와 있다. 표를 살펴보면, 운수사고(40.3%)와 추락(28.6%), 부딪힘(10.1%) 등이 주된 원인임을 알 수 있다.

〈표 3-12〉 손상요인별 환자 수 추정

손상기전	전체		남자		여자	
	퇴원환자수	%	퇴원환자수	%	퇴원환자수	%
운수사고	332,056	40.3	199,907	39.6	132,149	41.3
성폭행	182	0.02	-	-	182	0.1
추락	236,145	28.6	126,275	25.0	109,870	34.4
부딪힘(사람, 물체)	83,146	10.1	63,923	12.7	19,223	6.0
자상	67,758	8.2	53,289	10.6	14,469	4.5
총상	79	0.01	58	0.01	21	0.01
불, 화염, 열	15,820	1.9	8,585	1.7	7,235	2.3
질식	1,616	0.2	1,002	0.2	614	0.2
익수	609	0.1	404	0.1	205	0.1
중독	22,721	2.8	9,585	1.9	13,136	4.1
기타	47,776	5.8	32,057	6.4	15,719	4.9
미상	16,519	2.0	9,592	1.9	6,927	2.2
계	834,437	100.0	504,677	100.0	319,750	100.0

출처 : 제2차 퇴원환자조사(2006)

### ③ 응급실방문 손상환자 통계

응급실방문 손상환자에 대한 분석결과는 월간 『손상감시정보』를 통해 공표되고 있다. <표 3-13>은 2007년의 응급실 손상환자 통계를 요약한 표인데 약 19만 명의 응급환자 중 26%에 해당되는 50만 명가량이 손상환자였음을 알 수 있다. 한편, 월별로는 7월에서 12월 사이에 응급실 방문환자가 많았다.

〈표 3-13〉 응급실 손상환자 통계

	1	2	3	4	5	6	7	8	9	10	11	12	계
응급실 방문환자	101,154	92,518	101,235	111,124	115,815	111,797	215,307	227,281	235,488	203,825	208,404	223,489	1,947,437
손상환자 (%)	22,161 (21.9)	22,566 (24.4)	26,824 (26.5)	27,917 (25.1)	30,815 (24.9)	30,815 (27.6)	58,886 (27.3)	62,212 (27.4)	61,641 (26.2)	57,231 (28.1)	53,022 (25.4)	53,379 (23.9)	505,491 (26.0)

출처 : 국가응급환자진료정보망(07.1.~07.12.)

<표 3-14>에는 월별, 성별, 연령별, 기전별 응급실 손상환자 비율이 나와 있다. 남자가 여자보다 많은 편이고, 10세 미만의 어린이 손상 비율이 가장 높다. 기전별로는 운수사고, 미끄러짐, 기타 둔상 등의 비율이 높음을 알 수 있다.

〈표 3-14〉 연령별 응급실 손상환자 통계

구 분	1	2	3	4	5	6	7	8	9	10	11	12	누계
성별													
남	60.8	61.1	62.2	62.8	63.2	63.0	62.7	61.7	62.0	62.3	62.2	61.8	62.2
여	39.2	38.9	37.8	37.2	36.8	37.0	37.3	38.3	38.0	37.7	37.8	38.2	37.8
연령(세)													
0~9	21.4	22.3	22.7	23.2	23.1	23.7	23.1	22.1	20.5	21.5	21.0	20.7	21.9
10~19	8.4	8.8	11.3	11.8	12.2	12.1	11.0	10.0	10.7	11.3	11.5	9.9	10.8
20~29	16.3	16.1	15.2	14.6	14.5	14.4	15.5	15.7	15.4	14.5	15.2	16.9	15.4
30~39	15.5	15.4	14.9	14.6	14.3	14.5	15.6	15.6	15.8	15.0	15.3	16.4	15.4
40~49	15.4	15.4	15.0	14.5	14.3	14.9	14.8	15.1	15.2	15.0	14.7	14.9	14.9
50~59	10.2	9.8	9.5	9.5	9.5	9.3	9.3	9.8	10.3	10.1	10.0	10.0	9.8
60~69	6.4	6.2	6.0	6.1	6.2	6.0	5.6	6.2	6.5	6.5	6.3	5.8	6.1
≥70	6.4	5.9	5.4	5.8	5.8	5.2	5.1	5.5	5.6	6.1	5.9	5.5	5.6
손상기전													
운수사고	21.2	21.8	20.3	20.3	20.7	20.7	19.2	19.9	19.9	20.7	21.0	19.7	22.0
추락	4.0	4.0	4.2	4.6	5.0	5.3	4.7	4.1	3.8	4.7	4.3	3.6	4.7
미끄러짐	19.2	18.2	16.8	16.8	17.2	16.6	16.9	17.2	16.7	18.2	18.4	19.5	19.2
기타둔상	17.2	18.3	20.1	19.1	19.3	19.6	17.6	17.0	18.2	18.3	18.0	18.7	19.8
관통상	6.7	7.9	8.5	8.3	7.9	8.2	7.5	8.4	7.7	7.2	6.9	6.8	8.3
기계	1.5	1.4	1.6	1.5	1.5	1.7	1.5	1.5	1.4	1.5	1.5	1.3	1.6
화염, 고온체	3.8	3.6	3.4	3.2	2.7	3.2	2.9	3.1	2.9	2.8	3.0	3.4	3.4
익수	0.0	0.1	0.1	0.1	0.1	0.1	0.2	0.3	0.1	0.1	0.1	0.1	0.1
중독	2.2	2.4	2.1	2.3	2.3	2.2	2.1	2.1	2.1	1.8	1.6	1.6	2.2
질식	0.5	0.6	0.5	0.5	0.3	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.3
기타요인	13.2	14.1	11.9	12.7	14.5	15.0	16.9	18.4	19.3	17.5	18.5	18.3	17.1
미상	10.6	7.7	10.6	10.6	8.6	7.2	10.6	7.9	7.8	7.0	6.6	6.8	8.1
계(명)	22,161	22,566	26,824	27,917	28,837	30,815	58,886	62,212	61,641	57,231	53,022	53,379	505,491

출처 : 국가응급환자진료정보망(07.1.~07.12.)

#### 토의 4

손상통계가 구체적으로 정책에 활용된 사례를 찾아보라.

우리나라에서는 2000년대 중반이 지난 최근에 기초적인 손상감시체계가 구축되었기 때문에 손상 감시의 연륜이 풍부하다고는 할 수 없다. 그러나 짧은 연륜에 비해 일단 감시체계 구축으로 인한 기본적인 통계가 생산된 후부터는 매우 적극적이고 활발하게 손상예방을 위한 정책이 수립, 시행되고 있다고 할 수 있다. 질병관리본부에서는 손상감시를 위한 웹사이트 (<http://injury.cdc.go.kr>)를 구축하여 손상정보와 대책을 체계적으로 마련하고 있다. 뿐만 아니라 손상포럼 개최, 월간 『손상감시정보』 발간 등 다양한 활동을 벌이고 있다. 또한 현재 운영 중인 손상감시체계 자료에 기반을 두고 손상의 주요 원인별로 예방을 위한 안전수칙을 개발하여 보급할 계획을 마련하고 있다. 그 일례로 2007년 8월에는 취학 전 어린이 손상 예방을 위한 「안전한 우리 집 : 가정 손상을 예방하는 방법」이라는 안전수칙 자료집을 개발, 보급한 바 있는데 이에 대해 살펴보기로 한다.

어린이 손상은 예후가 나쁠 경우 장애나 후유증이 동반되어 삶의 질을 저하시키게 되며, 사회·경제적으로 부담이 되어 국가적으로 큰 손실을 일으킨다. 2007년의 정부 보도 자료에 의하면 어린이 손상으로 인한 사회적 손실 추산액은 1년에 2조 136억 원이라고 한다.

손상통계에 의하면, 취학 전 어린이 손상의 71.2%가 가정에서 발생하며, 가정에서도 공간에 따라 손상유형이 다르게 나타났다. <표 3-15>를 보면, 방이나 거실에서는 둔상, 부엌에서는 화상, 욕실에서는 미끄러짐, 계단에서는 구름이 가장 자주 일어난다는 것을 알 수 있다. 또한 어린이 손상은 어린이가 집에 혼자 있을 때 일어나기 쉬우며 하루 중 오후 7시에서 11시 사이에 많이 발생하는 편이다.

〈표 3-15〉 손상 원인별, 장소별 손상발생 수

건(%)

구 분	방·침실	거실	부엌	욕실·화장실	계단
추락	146(22.4)	75(12.8)	9(5.5)	6(6.3)	0(0.0)
미끄러짐	62(9.5)	65(11.2)	3(1.8)	53(55.2)	1(2.4)
계단에서 구름	0(0.0)	0(0.0)	0(0.0)	0(0.0)	32(78.0)
둔상	238(36.4)	245(42.3)	41(24.8)	24(25.0)	7(17.1)
끼임	48(7.4)	24(4.1)	4(2.4)	2(2.1)	0(0.0)
이물질 흡입	49(7.5)	49(8.5)	24(14.5)	2(2.1)	0(0.0)
관통	53(8.1)	52(9.0)	16(9.7)	3(3.1)	0(0.0)
화상	11(1.7)	34(5.9)	66(40.0)	5(5.20)	0(0.0)
감전	1(0.2)	1(0.2)	0(0.0)	0(0.0)	0(0.0)
물에 빠짐	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)
기타	45(6.9)	35(6.0)	2(1.2)	1(1.0)	1(2.4)
계	653(100.0)	579(100.0)	165(100.0)	96(100.0)	41(100.0)

출처 : 응급실 손상환자 표본심층조사('06.8-'07.5), 질병관리본부

질병관리본부에서는 이러한 통계에 근거를 두고 가정 내 안전을 위한 다섯 가지 기본 원칙을 마련하였는데 그 내용은 ① 어린이를 혼자 집에 두지 말 것 ② 문은 항상 꼭 닫아둘 것 ③ 집안 높은 곳에서 뛰지 말 것 ④ 위험한 물건은 아이 손이 닿지 않는 곳에 둘 것 ⑤ 어린이가 위험한 곳에 들어가지 못하게 할 것 등이다. 뿐만 아니라 손상 종류별 예방법과 가정 내 장소별 예방법 등도 제시하였다.

어린이 손상 예방을 위한 안전수칙 마련과 더불어 정책당국에서는 각종 마스크를 통한 홍보, 응급의료기관을 통한 자료집 배포, 관련 기관 및 민간기관을 통한 자료집 배포 및 홍보, 웹사이트를 통한 홍보 등을 수행하고 있다. 이러한 예방대책의 효과는 지속적인 조사를 통해 생산될 손상감시통계에서 나타나게 될 것이다.



### 3.4 성과지표 설정 사례

정부가 지원하는 재정사업은 성과측정의 목표가 분명하고 성과지표를 제시할 수 있는 사업 단위로 구성되어야 한다. 일반적으로 사업이라고 명명할 경우에 그 의미는 “특정한 목표를 달성하기 위해서 관련된 세부 활동을 사전에 계획하고 그에 맞추어 실행함으로써 의도한 결과가 달성되도록 고안된 것” 이라고 할 수 있다. <표 3-16>에는 복지 분야 5개 사업의 사업명, 성과목표, 성과지표(지표 특성) 등이 사례로 제시되어 있다.

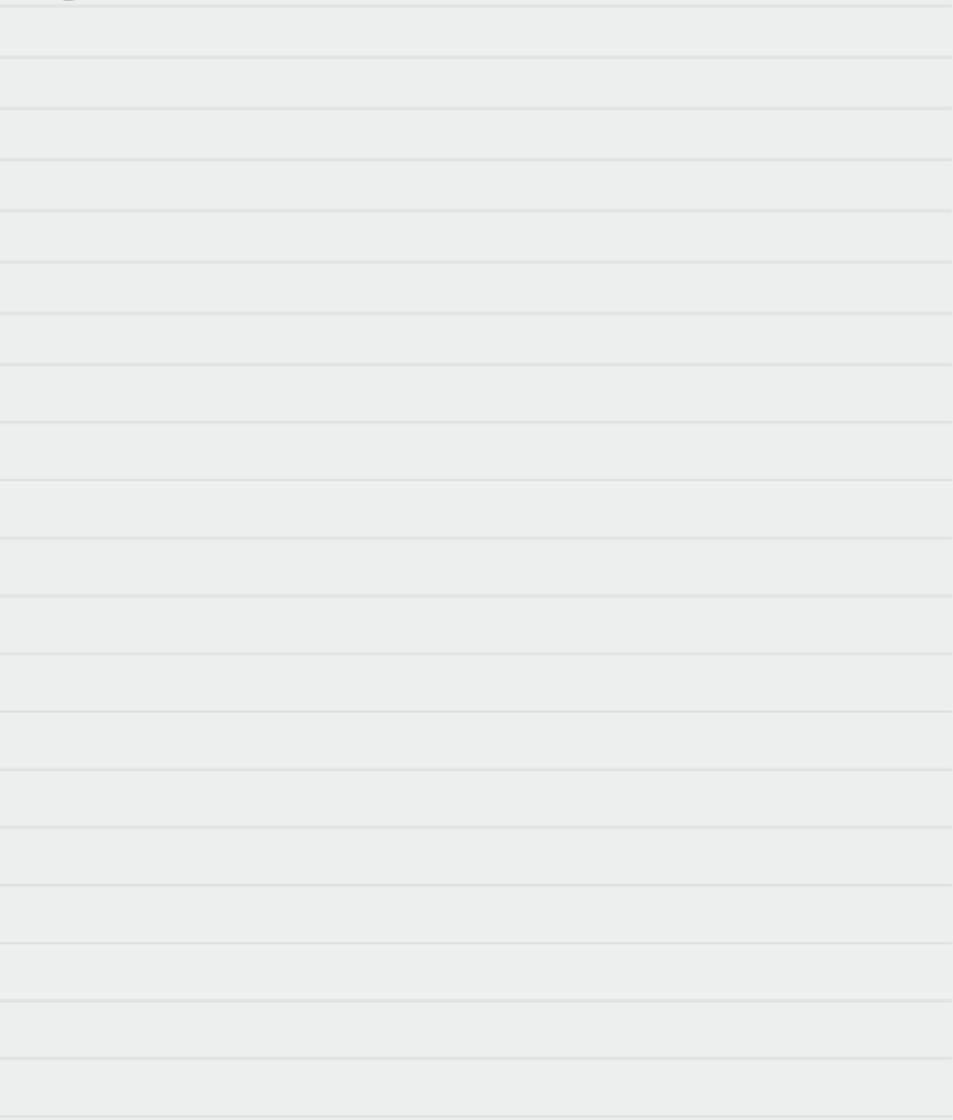
<표 3-16> 복지 분야 성과지표 설정사례

사업명	성과목표	성과지표	지표성격
□ 금연홍보사업	○ 효과적인 금연광고를 통해 흡연을 예방하고 금연분위기를 확산	- 금연광고 인지율(%)	만족도 조사
□ 노인취업지원센터 지원	○ 구직희망 노인에 대한 노인적합형 일자리 취업, 알선 지원 과 사업내 실화 도모	- 노인취업실적(천명) - 장기취업실적(천명) - 단기취업실적(천명) - 장기취업성공률(%) - 단기취업성공률(%) - 참여자만족도(%)	산출 산출 산출 결과 결과 만족도 조사
□ 원폭피해자복지회관 운영	○ 원폭피해자 건강관리 및 복지증진	- 건강진료비지원 - 진료보조비지원 - 장례비지원 - 원폭피해자요양관리 - 2세 건강검진	산출 산출 산출 산출 산출
□ 장애인복지사업 및 진흥회 지원	○ 장애인에 대한 인식 개선, 장애인의 사회적 활동 능력 배양 및 재활 도모 ○ 국가 장애인복지시책 수행 및 정책 홍보	- 장애인주거환경개선조사(건수) - 재활체육프로그램 운영(명수) - 편의증진인식개선교육(명수) - 인식개선 홍보(횟수) - 복지연구 실적(건수)	산출 산출 산출 산출 산출
□ 결핵관리사업	○ 면역인구 확대 ○ 환자발견사업의 지속적인 추진 ○ 보건교육 강화 ○ 국가결핵정보 감시체계의 효율적 활용	- 신생아 BCG 접종률(%) - 도말양성 결핵환자 발견율(%) - 민간병의원 결핵환자 신고율(%) - 결핵초치료 성공률(%)	결과 결과 결과 결과

토의문제 1

〈표 3-16〉에서 사업의 정의와 부합되는 사업명을 고르고 그 이유를 설명하라. 사업명을 수정해야 하는 경우는 수정된 사업명을 제시하라.

토의  
내용



## 토의 1

<표 3-16>에서는 복지 분야의 재정사업에 대해서 성과목표와 성과지표가 제시되어 있다. 이들 간의 연계성, 즉 인과관계가 올바르게 설정되어 있는지, 또 성과지표의 특성이 적절히 반영되어 있는지를 검토하여 5개 사업 중 3개 사업을 선택하여 수정목표와 성과지표를 제시하라.

① 정책평가부문에서는 사업의 전략목표와 성과목표를 구분하여 설명하였으나, <표 3-16>에는 전략목표가 제시되어 있지 않으므로 성과목표와 성과지표가 적절히 설명되어 있는지에 대해 알아본다. 성과목표로서 적합한지를 검증하는 기준으로 앞에서 계량성, 기간의 구체성, 이해관계자의 의견반영 등을 제시하였는데 여기서는 계량성을 중심으로 살펴볼 것이다. 이때 구체적인 사업기간이 제시되어 있고 이해관계자의 의견도 충분히 반영한 것으로 가정한다(제1장의 성과지표부문 참조).

② 금연홍보사업의 경우 “효과적인 금연광고를 통해 흡연예방 및 금연분위기 확산”을 성과목표로 제시하였다. 제1장에서 제시한 성과지표의 특성과 부합되도록 성과목표를 제시해본다면 “효과적인 금연홍보를 통한 흡연을 저하”로 바꿀 수 있을 것이다. 이 사업의 구체적인 계획은 홍보활동인데 이를 통해서 얻는 것은 결국 흡연율을 낮추는 것이다. 그러므로 성과측정을 통해 흡연자를 줄이거나 흡연을 예방해서 총 흡연자수가 줄었다는 것을 확인해야 하는 것이다. 장기적 관점에서는 기간을 제시하여 흡연율을 얼마만큼 저하시킬 것인지 구체적으로 제시할 수도 있다. 또 최근 여성과 청소년 흡연율이 급증하고 있으므로 여성과 청소년 흡연율에 대한 세부 성과지표를 제시할 수도 있다. 예를 들어 성과목표는 ‘2015년 까지 청소년과 여성 흡연율을 각각 ○○%, ○○%로 낮춘다’ 라고 설정하고, 당해 연도 성과지표로는 흡연율을 몇 %로 설정할 것인지 제시할 수도 있다. 이와 같이 성과목표는 객관적으로 측정가능하게 계량화시켜 표현해야 하며, 사업기간도 함께 명시할 필요가 있다. 지표를 설정함에 있어서 사업을 수행하는 부서나 민간단체 등은 사업과 관련된 이해관계자(보건복지부, 사업수혜자, 여타 협회 및 관련단체, 연구소 등)들과도 상호 공감할 수 있어야 한다.

③ 성과지표의 타당성 검증은 1장에서 논의한 바와 같이 적합성, 계량성, 민감성, 비편향성 등으로 확인할 수 있다. 성과지표의 적합성이란 성과목표를 달성하는 과정에서 그 달성정도를 파악할 수 있어야 한다는 것을 의미하며, 계량성은 개인의 주관적 의견이 개입될 수 없도

록 객관적인 수치로 표현하는 것을 말한다. 또한 설정된 성과지표는 사업의 외부 환경변화에 큰 영향을 받지 않아야 하며, 추진단체의 사업집행이 의도되지 않는 방향으로 유도되어서는 안 된다.

④ 노인취업지원사업의 경우 “구직희망 노인에 대한 노인적합형 일자리 취업, 알선 지원 및 사업내실화 도모”를 성과지표로 제시하였는데 성과지표를 “노인적합형 일자리 취업,알선을 통한 취업률 제고”라고 재설정할 수 있다. 이 사업의 구체적인 계획은 취업알선활동으로서 이를 통해 얻는 것은 결국 취업률을 높이는 것이다. 성과측정의 목표는 이 사업을 통해 노인 취업자가 얼마만큼 늘었는지 확인할 수 있는 것이어야 한다. 또 취업률에 대한 통계자료가 장·단기로 구분된다면 단기취업과 장기취업에 대한 구체적 성과지표를 기간과 함께 제시할 수 있다. 예를 들어 ‘2015년 까지 노인 단기취업률 및 장기취업률을 각각 ○○%, ○○%로 제고’라고 설정할 수 있다.

⑤ 결핵관리사업의 경우 초기 성과목표로 “면역인구확대, 환자발견사업의 지속적 추진, 보건교육강화, 결핵감시체계의 효율적 활용”등을 제시하였으나 이는 올바른 성과목표라고 볼 수 없다. 우선 면역인구확대라는 목표는 “2015년까지 신생아 BCG 접종률을 95%로 제고”라고 수정할 수 있고, 성과지표는 해당 회계 연도의 신생아 BCG 접종률을 몇 % 올릴 것인지 구체적으로 제시할 수 있다. ‘도말양성 결핵환자 발견율’, ‘민간병원 결핵환자 신고율’, ‘결핵초기 성공률’도 같은 식으로 수정할 수 있다.

⑥ <표 3-16>에는 산출과 만족도 조사를 성과지표로 제시한 경우가 많은데 이러한 것들은 목표설정이 불분명하여 지표로서의 적합성이 떨어지거나 계량화하기가 어렵다. 정부예산으로 무엇을 얼마만큼 개선시킬 것인지 불분명하거나 상황을 어디까지 개선시킬 것인지 명확하지 않기 때문이다. 장애인복지지원사업의 경우 성과목표가 광범위하여 성과목표설정의 기준을 제시하기 어렵고, 너무 많은 성과목표를 제시하여 목표달성정도를 파악하기도 힘들다. 따라서 전략목표를 2개 혹은 3개로 구체화하여 성과목표를 설정하되 성과지표와의 연계성을 높여야 한다. 원폭피해자복지사업도 같은 맥락에서 각종 지원 수단을 구체화하여 성과목표를 설정하고 당해 연도의 성과지표를 제시하여야 한다.



## 3.5 도서관예산안 작성 연습과제

### 3.5.1 연습과제 개요

이 연습과제의 목적은 제시된 가상의 자치단체인 태평시 도서관 예산 자료를 분석하되 기초적인 통계지식을 예산정책에 활용해보는 것이다. 최근 정부예산운영의 초점은 효율성과 효과성을 높이는 데에 있다. 효율성을 높이기 위해서는 정부사업에 대한 단위 원가를 계산하게 하는 등 다양한 방법이 있겠지만 이 과제에서는 제시된 가상의 2008 회계연도 실행예산과 관련 자료를 분석하여 예산을 삭감해야 한다.

다음의 예산편성지침에 따라 실행예산 대비 5% 삭감한 2009 회계연도 예산안을 제시해보자. 지침에 없는 내용은 가상 시나리오를 통해 예산삭감의 정당성을 제시해야 한다. 우리나라의 경우 중앙이든 지방이든 의회에 제출된 예산안에서 1% 내외 정도만 삭감되는 실정이므로 이와 같은 편성이 갖는 의미는 매우 크다고 볼 수 있다.

### 3.5.2 예산편성 지침

〈 2008년 태평시 도서관 예산 공문〉

2008년 2월

받는 사람: 태평시 예산실 도서관 예산담당관

보내는 사람: 태평시 예산실장

내 용: 통합될 태평시 도서관 예산에 대한 예산편성지침

(1) 양식 1에서 5까지는 2008년 태평시 5개 도서관의 집행예산 내역이다. 단일시스템 안에 의해서 태평시에 있는 도서관을 다음과 같이 조정한다. 1구 도서관은 기존과 같이 존속시키며, 지역적으로 인접해 있는 4개 도서관을 2개로 개편하여 예산안을 마련한다(2구와 3구의 도서관들을 하나로 통합하여 그 명칭을 2/3구 도서관으로, 4구와 5구는 4/5구 도서관으로 통합하여, 중앙도서관을 포함한 총 3개 도서관으로 개편하여 운영한다).

(2) 태평시가 단일시스템으로 도서관을 운영하게 되면 이전보다 예산을 5%이상 절감하여야 한다. 예산절감 요인에 대한 자세한 보고서는 추가로 그 내역서를 제출한다.

(3) 1구 도서관의 전임직 직원을 해임할 수 없다. 2구에서 5구는 통합에 따라서 중복된 전임직에 대해서만 직위해임하며, 대기발령을 위해서 대기발령자의 직위와 수를 보고한다(물론, 시간직 직원의 해임이나 기타 운영비 삭감은 예산담당자가 새로 개편되는 도서관의 사정에 따라 조정한다. 각 도서관의 예산요구서에 의해 예산안이 마련되어야 하나 편의상 2008년 예산내용에 의거하여 예산담당관의 재량으로 편성한다).

(4) 분류된 직급의 연봉과 시간급은 그 직급 중 2008년 실행예산에서 가장 높은 연봉과 시간급을 주는 도서관에 준하여 통일한다.

(5) 1구를 제외한 도서관장은 '區도서관장'으로 재임명되며(대기발령자는 제외), 1구의 부관장은 1구 도서관장으로 발령한다(2009년 회계연도부터 구도서관장의 연봉은 2008년 예산에서 1구 부관장의 연봉에 준하여 책정한다).

(6) 2008년도 1구 도서관장은 새로운 직위인 '市도서관장'으로 임명되며 새로 개편되는 3개 도서관을 총괄 관리하는 책임을 진다. 市도서관장의 연봉은 40,000,000원으로 책정한다.

(7) 3개 도서관 예산을 합산한 예산안만을 제시하고, 3개 도서관의 개별 예산안은 추가로 첨부하되 변경 및 삭감 내용은 그 근거를 분명히 하여 정당성을 제시하여야 한다.

(표 3-17) 양식1 : 2008 회계연도 태평시 1구 도서관 예산

(단위: 천원)

품 목			연봉/시간급	인원수/총시간	비 용
전문직	전입직	도서관장	40,000	1	40,000
		부관장	30,000	1	30,000
		참고열람실사서	25,000	3	75,000
		보조사서	23,000	5	115,000
전문직	시간직	참고열람사서(6개월)	15,500	2	31,000
		보조사서(6개월)	13,375	2	26,750
기타직	전입직	행정보조원	22,000	1	22,000
		도서대출원	20,000	5	100,000
		비서	19,500	1	19,500
		관리직원	16,500	3	49,500
		급사	15,000	2	30,000
	기타직	시간직	도서대출원	6/시간당	7,020
		관리직원	5.75/시간당	2,000	11,500
		급사	4.85/시간당	13,260	64,311
기 타 운영비	기 타 내 용	도서구입비			275,000
		비품비			110,000
		공공요금			96,700
		관리비			32,600
		기타			43,000
		▷ 이동도서관 운영비보조(₩21,000) ▷ 초, 중등, 고등학교 행사 지원비(₩11,000) ▷ 여름방학 어린이 행사비(₩6,500) ▷ 여행경비(₩4,500)			

※ 시간급의 연봉은 시간급 × 총시간으로 계산된다.

예를 들어 시간직 도서대출원의 연봉은 6,000원 × 7,020시간 = 42,120,000원이다.

※ 기타 프로그램에 대한 자세한 설명은 제시되어 있지 않으므로 예산담당자가 나름대로 시나리오를 구성하여 예산변경에 대한 정당성을 제시한다.

〈표 3-18〉 양식2 : 2008 회계연도 태평시 2구 도서관 예산

(단위: 천원)

품 목		연봉/시간급	인원수/총시간	비 용
전문직	전임직	25,000	1	25,000
	시간직			
기타직	전임직			
	시간직	5.45/시간당	780	4,251
기 타 운영비				2,500
				750
				3,600
				1,260
	기 타 내 용			

(표 3-19) 양식3 : 2008 회계연도 태평시 3구 도서관 예산

(단위: 천원)

품 목		연봉/시간급	인원수/총시간	비 용
전문직	전입직	도서관장 부관장	1	27,500
		참고열람실사서 보조사서	1	23,000
	시간직	참고열람사서(6개월) 보조사서(6개월)		
기타직	전입직	행정보조원	1	14,000
		도서대출원 비서 관리직원 급사	1	14,500
	시간직	도서대출원 관리직원 급사	1,040	4,680
기 타 운영비		도서구입비		12,500
		비품비 공공요금 관리비 기타		2,450 4,200 3,500
	기 타 내 용			

〈표 3-20〉 양식4 : 2008 회계연도 태평시 4구 도서관 예산

(단위: 천원)

품 목		연봉/시간급	인원수/총시간	비 용	
전문직	전임직	도서관장	30,000	1	32,500
		부관장			
		참고열람실사서	24,750	1	24,750
		보조사서	22,500	4	90,000
	시간직	참고열람사서(6개월)	14,500	1	14,500
		보조사서(6개월)	13,750	1	13,750
기타직	전임직	행정보조원	19,500	1	19,000
		도서대출원 비서 관리직원 급사	15,400	1	15,400
	시간직	도서대출원 관리직원 급사	5.25/시간당	3,300	17,325
기 타 운영비		도서구입비			124,000
		비품비			34,000
		공공요금			3,780
		관리비			3,600
		기타			12,500
	기 타 내 용	▷ 어린이를 위한 토요일 행사와 일요일 외국영화 상영비(W12,500)			

〈표 3-21〉 양식5 : 2008 회계연도 태평시 5구 도서관 예산

(단위: 천원)

품 목			연봉/시간급	인원수/총시간	비 용
전문직	전임직	도서관장	27,000	1	27,000
		부관장 참고열람실사서 보조사서	22,500	1	22,500
	시간직	참고열람사서(6개월) 보조사서(6개월)			
기타직	전임직	행정보조원	16,500	1	16,500
		도서대출원 비서 관리직원 급사	15,000	1	15,000
	시간직	도서대출원 관리직원 급사	5.75/시간당 4.75/시간당	1,560 2,300	8,970 10,925
기 타 운영비		도서구입비 비품비 공공요금 관리비 기타			35,700 4,300 3,800 5,625
	기 타 내 용				

〈표 3-22〉 구별 인구 변화

(단위: 천명)

년도 구	1985	1990	1995	2000	2005
1구	350	478	531	586	641
2구	83	70	61	57	52
3구	102	100	137	144	132
4구	142	115	112	116	112
5구	68	52	43	38	32

〈표 3-23〉 구별 10세~29세 인구 변화

(단위: 천명)

년도 구	2004년	2005년	2006년	2007년	2008년
1구	39	42	41	42	51
2구	4	4	4	4	3
3구	10	10	10	10	10
4구	8	8	8	8	9
5구	3	2	2	2	2

### 3.5.3 과제 풀이

토의 1

태평시 도서관의 2008 회계연도 예산총액은 얼마였는지 계산하라.

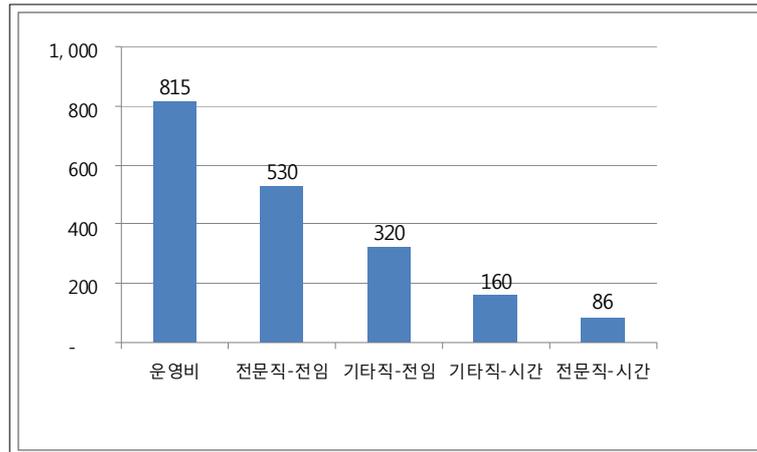
2008년 태평시 5개 도서관의 구체적인 예산 내역을 정리한 결과가 <표 3-24>에 나와 있다. 태평시의 지난해 예산총액은 약 19.11억 원이었다.

<표 3-24> 2008년 도서관 예산 총액

(단위 : 천원)

품목		비용 총합	
전문직	전임직	도서관장	149,500
		부관장	30,000
		참고열람실사서	99,750
		보조사서	250,500
	시간직	참고열람서	45,500
		보조사서	40,500
기타직	전임직	행정보조원	22,000
		도서대출원	150,000
		비서	19,500
		관리직원	98,651
		급사	30,000
	시간직	도서대출원	51,090
		관리직원	22,425
		급사	86,316
기타 운영비	도서구입비	449,700	
	비품비	151,500	
	공공요금	112,080	
	관리비	46,585	
	기타	55,500	
총 계		1,911,097	

전체 예산 중 항목별로 보면 인건비가 약 11억, 운영비가 약 8.15억을 차지한다. 인건비를 구체적으로 보면, 전문직 전임(5.3억), 기타직 전임(3.2억), 기타직 시간(1.6억), 기타직 전임(0.86억) 등으로 나타났다 (<그림 3-17> 참조).

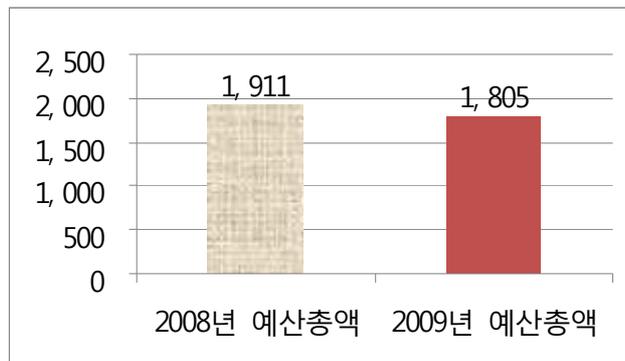


<그림 3-17> 2008년 항목별 예산액

토의 2

2009년도 예산안을 제시하되, 지방의회 의원들이 이 예산안을 쉽게 이해할 수 있도록 나타내보라. 아울러 제시하는 예산안의 타당성을 입증할 수 있는 근거자료를 제시하라.

2009년의 태평시 도서관들의 예산안을 정리한 결과가 <표 3-25>에 나와 있다. 2구와 3구 도서관을 2/3구 도서관으로, 4구와 5구를 4/5구 도서관으로 통합하여 2009년 예산안을 작성한 결과, 2009년의 예산총액은 약 18억으로 2008년의 19.1억에 비해 5% 가량 절감된 것으로 나타났다(<그림 3-18> 참조).



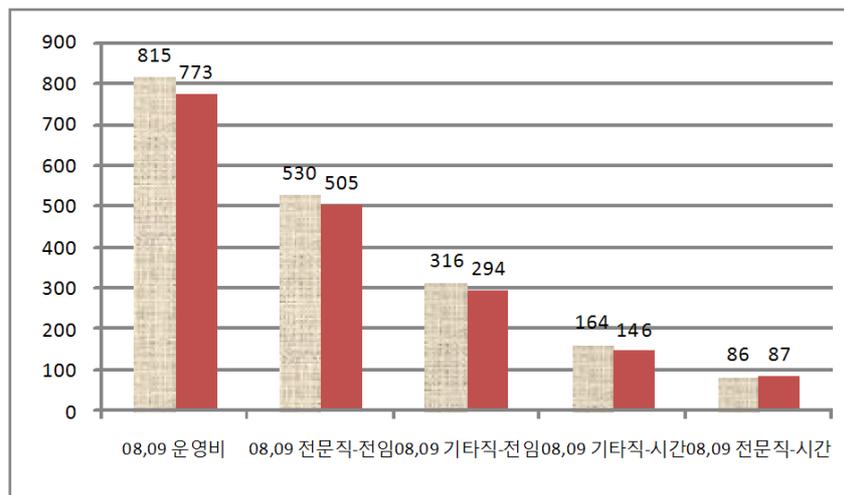
<그림 3-18> 2008년과 2009년 예산총액 비교

2009년과 2008년의 예산을 항목별로 비교한 그림이 <그림 3-19>에 나와 있는데, 도서관 통합으로 인해 운영비가 일부 절감되고 4명의 대기 발령자(도서관장, 보조사서, 도서대출원, 관리직원)로 인해 인건비가 일부 절감되는 것으로 나타났다.

〈표 3-25〉 2009년 도서관 예산 총액

(단위: 천원)

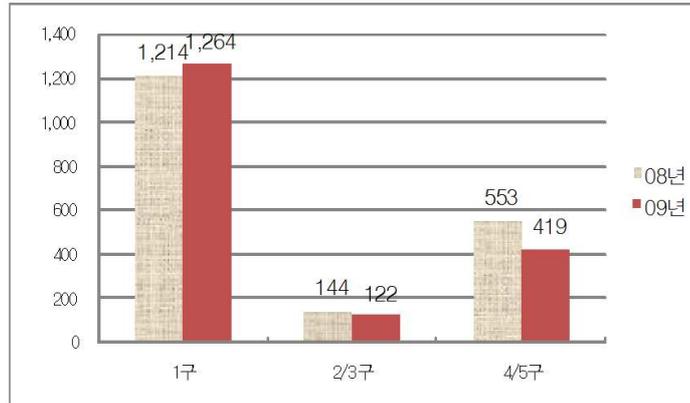
품 목		연봉/시간급	인원수/총시간	비 용	
전문직	전임직	시도서관장	45,000	1	45,000
		도서관장(1구)	40,000	1	40,000
		도서관장	30,000	2	60,000
		부관장	30,000	1	30,000
		참고열람실사서	25,000	4	100,000
		보조사서	23,000	10	230,000
	시간직	참고열람사서	15,500	3	46,500
	보조사서	13,375	3	40,125	
기타직	전임직	행정보조원	22,000	1	22,000
		도서대출원	20,000	7	140,000
		비서	19,500	1	19,500
		관리직원	16,500	5	82,500
		급사	15,000	2	30,000
	시간직	도서대출원	6	7,020	42,120
		관리직원	5.75	2,000	11,500
	급사	5.25	17,600	92,400	
기 타 운영비		도서구입비			419,000
		비품비			148,050
		공공요금			108,580
		관리비			42,080
		기타			55,500
	기 타 내 용	▷ 이동도서관 운영비보조(₩21,000) ▷ 초, 중등, 고등학교 행사 지원비(₩11,000) ▷ 여름방학 어린이 행사비(₩6,500) ▷ 여행경비(₩4,500) ▷ 어린이를 위한 토요일 행사와 일요일 외국영화 상영비(₩12,500)			
		총 계			



〈그림 3-19〉 2008년과 2009년 항목별 예산액 비교

한편 두 가지 예산을 도서관별로 비교한 그림이 <그림 3-20>에 나와 있다. 1구는 오히려 예산이 늘어나게 되는데 그 이유는 새롭게 시도서관장 직책이 신설되었기 때문이다. 하지

만 나머지 두 도서관은 2008년에 비해 예산이 다소 줄어든 것을 알 수 있는데 특히 4/5구 도서관의 예산이 상대적으로 많이 줄어 든 편이다.

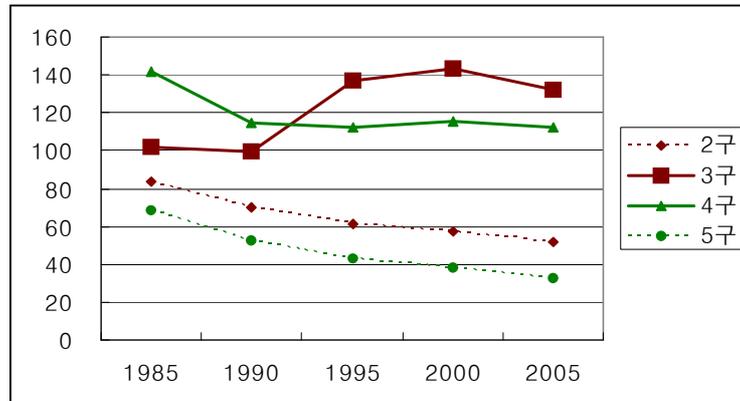


〈그림 3-20〉 2008년과 2009년 구별 예산액 비교

위 예산안의 가장 큰 특징으로는, 예산 절감을 위해 인위적인 시도를 하는 대신 일부 도서관들을 통합함으로써 자연스럽게 예산을 절감할 수 있게 한 점이다. 예산 절감을 하되 전임 직이나 시간직 인원의 연봉과 시간급을 내리지는 않았고, 통합이 되는 도서관의 경우 동일 직급의 최고 연봉을 채택함으로써 오히려 대우를 향상시켰다. 또한 통합으로 인한 논란을 최소화하기 위해 인위적인 감원은 시도하지 않았다. 사라지는 두 도서관의 관장 중 1명은 1구 도서관의 부관장에 보임하고 나머지 1명은 대기발령을 내린다. 또한 일부 중복되는 보직 담당자 3명에 대해서도 대기발령을 내리도록 한다. 도서구입비는 통합 전 금액이 큰 도서관의 예산을 채택하며, 그 밖의 운영비는 남은 도서관의 기존 운영비에 사라지는 도서관 운영비 50%를 추가하도록 한다.

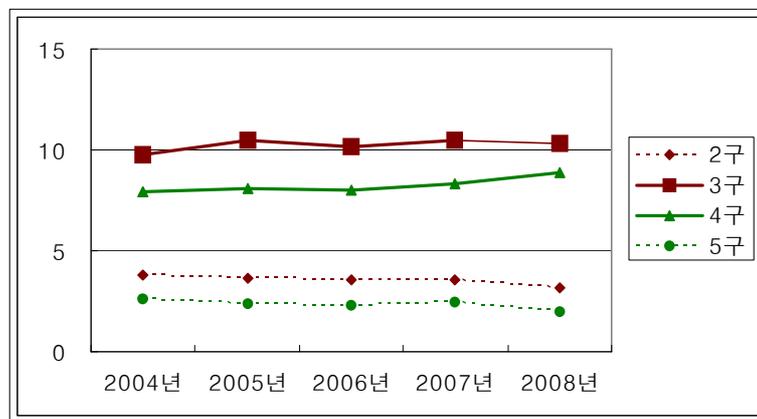
통합대상이 되는 2구와 3구, 4구와 5구 도서관 중 어느 것을 그대로 두고 어느 것을 정리해야 할 것인가를 합리적으로 결정하는 것이 필요하다. 이를 위해서 태평시 2, 3, 4, 5구의 인구수 현황과 10세~29세 인구 현황 및 각 도서관별 이용자수의 변화를 살펴보았다. 인구 외에 10세~29세 인구를 따로 고려하는 이유는 도서관의 주된 이용 연령대가 이들이기 때문이다.

〈그림 3-21〉은 구별 인구 변화를 나타낸다. 실선으로 표시된 3구와 4구는 통합 과정에서 도서관이 계속 남게 되는 구이고, 점선으로 표시된 2구와 5구는 도서관이 정리될 구이다. 3구와 4구는 각각 인구가 10만을 초과하는 구인데 비해 2구와 5구는 인구가 6만에 못 미칠 뿐 아니라 1985년 이후 지속적으로 인구가 줄어드는 곳이다.



〈그림 3-21〉 관심지역의 연도별 인구

<그림 3-22>는 과거 5년간 구별 도서관 이용자 수 현황을 나타낸다. 이 그림을 보면 3구와 4구의 경우 이용자 수가 8천 명이 넘고 동시에 매년 이용자 수가 비슷하거나 다소 늘어나는 추세인 반면, 2구와 5구의 경우 이용자 수가 각각 5천명 미만이며 매년 조금씩 줄어들고 있다.



〈그림 3-22〉 관심지역의 연도별 도서관 이용자수

구별 인구나 이용자 수 등을 고려한다면, 2구와 3구 도서관을 통합할 때 2구 도서관을 정리하는 것이 합리적이며, 4구와 5구 도서관을 통합할 때는 5구 도서관을 정리하는 것이 자연스럽다고 할 것이다.

각 구별 세부적인 예산안은 다음의 <표 3-26>에서 <표 3-28>까지 나와 있다.

〈표 3-26〉 2009년 1구 도서관 예산 예시

(단위 : 천원)

품목		연봉/시급	인원/총시간	비용	
전문직	전임직	市 도서관장	45,000	1	45,000
		1구 도서관장	40,000	1	40,000
		부관장	30,000	1	30,000
		참고열람실사서	25,000	3	75,000
		보조사서	23,000	5	115,000
	시간직	참고열람사서(6개월)	15,500	2	31,000
		보조사서(6개월)	13,375	2	26,750
기타직	전임직	행정보조원	22,000	1	22,000
		도서대출원	20,000	5	100,000
		비서	19,500	1	19,500
		관리직원	16,500	3	49,500
		급사	15,000	2	30,000
	시간직	도서대출원	6	7,020	42,120
		관리직원	5.75	2,000	11,500
		급사	5.25	13,260	69,615
기타 운영비	도서구입비			275,000	
	비품비			110,000	
	공공요금			96,700	
	관리비			32,600	
	기타			43,000	
기타내용	이동도서관 운영비보조(W21,000)				
	초, 중등, 고등학교 행사 지원비(W11,000)				
	여름방학 어린이 행사비(W6,500)				
	여행경비(W4,500)				
총 계				1,264,285	

[착안 사항]

1. 1구 도서관의 사항들을 2~5구 도서관 통합의 기준으로 적용한다. 2~5구 도서관 전임직원의 임금수준을 1구 도서관에 준하여 책정한다.
2. 당해 도서관의 관장을 시(市) 도서관장으로, 부관장을 도서관장으로 각각 승진 인사 한다.
3. 부관장은 2/3구 도서관 통합 때 발생하는 도서관장 중복으로 인한 대기발령자인 기존 2구 도서관장을 보직 임용한다.

〈표 3-27〉 2009년 2/3구 도서관 예산 예시

(단위 : 천원)

품목		연봉/시급	인원/총시간	비용	
전문직	전임직	도서관장	30,000	1	30,000
		부관장			
		참고열람실사서			
		보조사서	23,000	1	23,000
	시간직	참고열람서(6개월)			
		보조사서(6개월)			
기타직	전임직	행정보조원			
		도서대출원	20,000	1	20,000
		비서			
		관리직원	16,500	1	16,500
		급사			
	시간직	도서대출원			
		관리직원			
		급사	5.25	1,040	5,460
기타 운영비	도서구입비			13,750	
	비품비			2,825	
	공공요금			6,000	
	관리비			4,130	
	기타				
총 계				121,665	

## 〔착안 사항〕

1. 기존의 2구 도서관과 3구 도서관을 통합하여 2/3구 도서관을 운영한다.
  - 1) 인구와 이용객수가 많은 3구를 중심으로 통합한다.
2. 통합 때 중복되는 2구 도서관장은 1구 도서관의 부관장으로 인사 조치한다.
3. 통합 운영 때 직원의 임금에 대해서는 1구 도서관의 기준에 준하여 책정한다.
4. 시의 예산안 5% 삭감지침을 고려하여 예산안 편성에 반영한다.
  - 1) 당해 도서관의 동일 직무에 대해 전임직과 시간직의 중복 보직이 있는 경우 시간직 직원을 해임한다.
    - 관리직원의 경우 전임직과 중복되는 시간직 관리직원 해임.
  - 2) 시간직 해임에 따른 인력부족은 임금이 인상된 전임 직원들의 업무량과 효율성 확대에 대체한다.
  - 3) 3구로의 통합에 따라 중복이 발생하는 기타운영비는 통합 전 2구 기타운영비의 50%를 반영하여 증액 책정.
    - 예) 2구 도서관 기타운영비의 50%와 3구 기타운영비의 100%를 합쳐 2/3구 기타운영비로 책정한다.

〈표 3-28〉 2009년 4/5구 도서관 예산 예시

(단위 : 천원)

품목		연봉/시급	인원/총시간	비용	
전문직	전임직	도서관장	30,000	1	30,000
		부관장			
		참고열람실사서	25,000	1	25,000
		보조사서	23,000	4	92,000
	시간직	참고열람서(6개월)	15,500	1	15,500
		보조사서(6개월)	13,375	1	13,375
기타직	전임직	행정보조원			
		도서대출원	20,000	1	20,000
		비서			
		관리직원	16,500	1	16,500
		급사			
	시간직	도서대출원			
		관리직원			
		급사	5.25	3,300	17,325
기타 운영비	도서구입비			130,250	
	비품비			35,225	
	공공요금			5,880	
	관리비			5,350	
	기타			12,500	
기타내용	어린이를 위한 토요일 행사와 일요일 외국영화 상영비(W12,500)				
총 계				418,905	

[착안 사항]

1. 기존 4구도서관과 5구 도서관을 통합하여 4/5구 도서관을 운영한다.
  - 1) 인구와 이용객수가 많은 4구를 중심으로 통합한다.
2. 통합 때 중복되는 5구 도서관장은 직위해제하여 대기발령한다.
3. 통합운영 때 직원의 임금에 대해서는 1구 도서관의 기준에 준하게 책정한다.
4. 시의 예산안 5% 삭감지침을 고려하여 예산안 편성에 반영한다.
  - 1) 5구의 전문직(전임)-보조사서, 기타직(전임)-도서대출원, 기타직(전임)-관리직원을 직위해제하여 대기발령한다.
  - 2) 4구로 통합에 따른 중복이 발생하는 기타운영비는 통합 전 5구 기타운영비의 50%를 반영하여 증액 책정.  
 예) 2008년도 5구 도서관 기타운영비의 50%와 4구 기타운영비의 100%를 합쳐서 4/5구 기타운영비로 책정한다.



## 3.6 통계 DB 활용 연습

과거에는 일반인들이 통계정보를 찾아 활용하는 것이 쉽지 않았다. 가령 어느 도시의 인구를 알아보려면 도서관에 가서 직접 좁쌀만한 글씨로 뿔뿔하게 인쇄된 방대한 보고서를 찾아야만 했다. 오늘날 그런 정보는 도서관에 가는 대신 인터넷을 이용하여 누구든 손쉽게 검색해서 활용할 수 있다. 통계를 작성하는 각 정부부처나 연구기관, 민간기관 등에서는 웹사이트를 통해 다양한 통계정보들을 제공하고 있다. 한편, 정부에서는 주요 통계정보들을 정책담당자나 연구자 및 일반인들이 손쉽게 활용할 수 있도록 국가통계포털(<http://www.kosis.kr>)과 e-나라지표(<http://www.index.go.kr>) 등의 사이트를 개설하여 운영하고 있기도 하다. 국가통계포털에서는 국가에서 공표하는 각종 공식통계들을 열람할 수 있을 뿐 아니라 자료를 내려 받을 수도 있다. 그리고 e-나라지표에서는 국정지표가 되는 다양한 통계들과 그 의미를 분석한 정보들을 얻을 수 있다. e-나라지표 웹사이트에 실린 소개말을 일부 인용하면 다음과 같다.

“21세기 지식정보화 사회를 맞아 세계를 선도적으로 이끌 주역으로 도약하기 위한 첫걸음은 우리의 현주소를 정확히 짚어보고 평가하는 것입니다. 통계는 과거에 대한 평가, 현 위치 진단 그리고 미래 상황에 대한 예측을 가능하게 하는 가장 객관적이고 명확한 척도입니다. e-나라지표 시스템은 국정 전 분야에 걸친 각종 국정통계에 대해 정확한 통계와 추이, 생생한 통계의 의미 분석 등을 제공함으로써 국민, 전문가, 정책결정자들 모두에게 국정을 알기 쉽게 설명하고, 모니터링할 수 있도록 할 뿐만 아니라 필요한 국정통계정보를 손쉽게 접근할 수 있도록 하고자 함에 있습니다.”

통계정보들을 손쉽게 검색하여 사용할 수 있는 환경이 이처럼 마련되어 있지만 이에 대해 알고 있는 사람도 드물뿐더러, 개인이 직접 정보를 검색하여 이용할 업무를 내는 사람도 많지 않다. 부뚜막의 소금도 집어넣어야 짜듯이 인터넷에 있는 많은 통계정보 역시 잘 활용해야 비로소 그 진가가 발휘되는 법이다. 따라서 이 절에서는 몇 가지 문제를 가지고 직접 인터넷 검색을 통해 유용한 통계정보를 정리하는 연습을 해보자.

토의 1

1980년부터 2005년까지 우리나라의 주택보급률과 인구 천 명당 주택 수의 변화를 정리해보자.

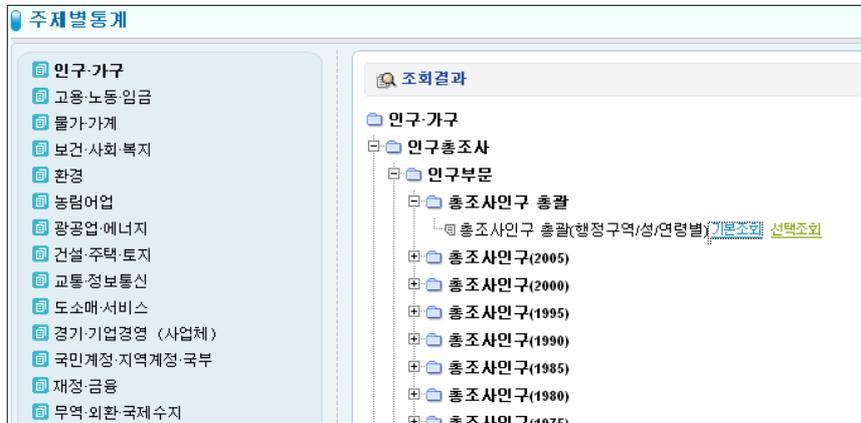
주택보급률이란 주택 수(빈집 제외)를 보통가구 수로 나누어 구하며, 인구 천 명당 주택 수는 인구수를 주택 수로 나누어 구한다. 따라서 1980년에서 2005년까지의 주택보급률과 인구 천 명당 주택 수를 구하려면, 인구, 가구 및 주택 수에 대한 통계정보를 찾아야 하는데 국가통계포털 사이트에 접속하면 이 모든 통계들을 검색할 수 있다.

<그림 3-23>은 국가통계포털 웹사이트 화면이다. 인구와 가구 수를 검색하기 위해서는 주제별통계에서 ‘인구·가구’를 클릭해야 하며, 주택 수는 ‘건설·주택·토지’를 선택해야 한다.



<그림 3-23> 국가통계포털 웹사이트 화면(http://www.kosis.kr)

먼저 인구를 검색해보자. 인구 통계는 인구총조사에 수록되어 있으므로 주제별통계에서 ‘인구·가구’를 클릭한 후 인구 부문에서 인구총조사를 선택하면, <그림 3-24>의 화면이 나타난다. 여기서는 여러 해의 결과를 얻는 것이 목적이므로 제공되는 메뉴 중 특정 년도의 인구총조사 결과보다 총조사인구 총괄을 선택하는 것이 효과적이다.



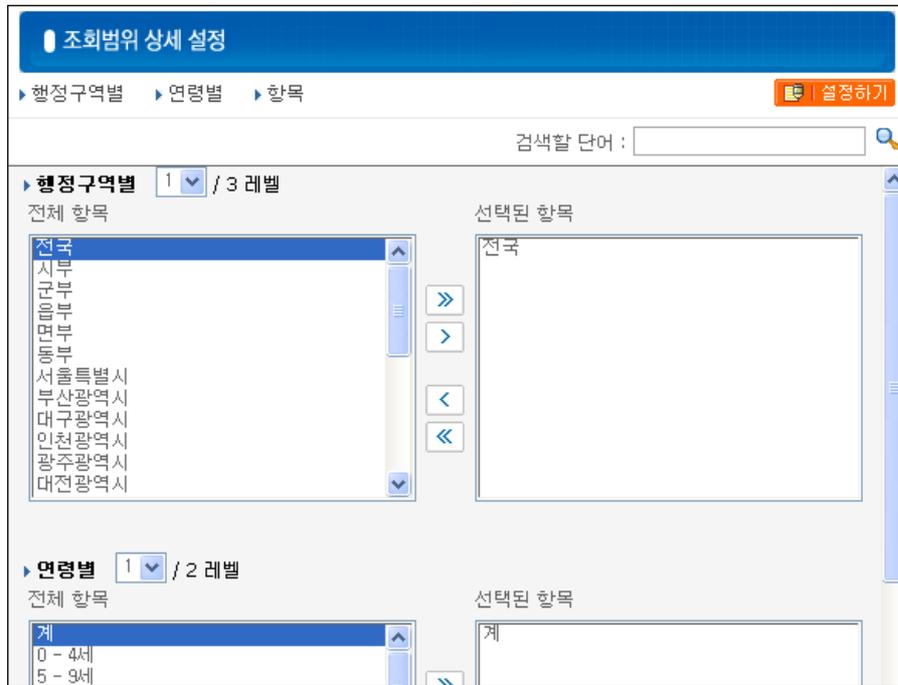
<그림 3-24> 주제별통계 중 인구·가구 조회

총조사인구 총괄의 기본조회를 클릭하면 <그림 3-25>의 초기화면이 나타난다. 이 화면은 제공되는 통계 자료의 기본양식이다. 기본양식에서는 제공 가능한 항목들이 모두 망라되어 있기 때문에 불필요하게 방대하며, 최근의 결과인 2005년 통계만을 나타내고 있다. 먼저 그림 왼쪽 상단의 조회기간 설정 칸에서 1980년부터 2005년까지를 설정한 후, 그 위에 있는 “조회범위 상세 설정하기”라는 단추를 클릭하여 목적에 적합한 항목만을 지정하면 된다.

행정구역별	# 연령별	2005			
		인구	남자	여자	성비
=전국	계	47,041,434	23,465,650	23,575,784	99.53
	0 - 4세	2,382,350	1,237,301	1,145,049	108.06
	5 - 9세	3,168,887	1,654,228	1,514,659	109.21
	10 - 14세	3,434,891	1,816,318	1,618,573	112.22

<그림 3-25> 주제별통계 중 인구·가구 조회 결과

<그림 3-26>은 조회범위 상세 설정을 위해 나타나는 화면이다. 왼쪽의 전체항목 창에 나타난 항목 가운데 필요한 것만 선택하여 오른쪽 선택된 항목 창에 보낸 후 ‘설정하기’를 클릭한다.



〈그림 3-26〉 조회범위 설정을 위한 화면

위에서 설명한 과정을 거치면 최종적으로 <그림 3-27>과 같은 결과를 얻게 된다. 이 그림의 오른쪽 상단을 보면 이 결과를 엑셀의 파일로 내려 받게 하는 ‘자료받기’ 기능이 있다. 이것을 활용하여 결과를 별도의 엑셀 파일로 저장해둔다.

■ 위치 : 주체별 > 인구·가구 > 인구총조사 > 인구부문 > 총조사인구 총괄

통계메타 조회 HTML 형식으로 보기

▶ 분류 설정 **조회범위 상세설정하기** ▶ 분류의 항목의 조회범위를 재설정하여 원하는 상세자료를 볼수 있습니다.

▶ 조회 기간 부정기 1970 ~ 2005 수록기간 (1925 ~ 2005)

▶ 소수점 한글보기  계층별 컬럼 구분

자료받기 자료분석 표 구성 기본 통계

자료다시 불러오기

표측 행정구역별 # 연령별 표두 부정기 항목 숨김

**총조사인구 총괄(행정구역/성/연령별)**

행정 구역별	# 연령별	1970	1975	1980	1985	1990	1995
		인구	인구	인구	인구	인구	인구
전국	계	31,435,252	34,678,972	37,406,815	40,419,652	43,390,374	44,553,710

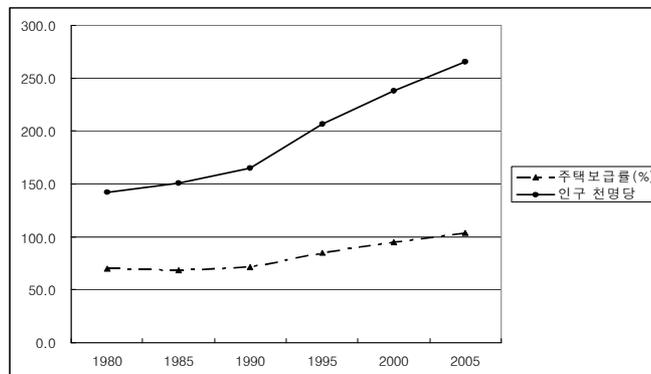
〈그림 3-27〉 국가통계포털의 검색 결과 화면

가구, 보통가구 및 주택 수에 대한 검색도 인구와 비슷한 단계를 거치면 된다. 이렇게 구한 자료를 기초로 주택보급률과 인구 천 명당 주택 수를 계산할 수 있다. <표 3-29>는 이상의 결과를 최종적으로 모아 정리한 표이다.

〈표 3-29〉 검색을 통해 최종 정리한 결과

연도	1980	1985	1990	1995	2000	2005
인구	37,407	40,420	43,390	44,554	45,985	47,041
가구	7,969	9,571	11,355	12,958	14,312	15,887
보통가구	7,586	8,910	10,333	11,316	12,087	12,716
주택	5,319	6,104	7,357	9,570	11,472	13,223
주택보급률(%)	70.1	68.5	71.2	84.6	94.9	104.0
인구 천 명당 주택 수	142.2	151.0	165.0	206.6	238.3	265.6

<그림 3-28>은 위의 자료를 이용하여 연도별 주택보급률과 인구 천 명당 주택 수의 변화를 각각 나타낸 것이다. 이 그림을 보면 주택보급률이나 인구 천 명당 주택 수 모두 1990년 이후 두드러지게 증가하기 시작했음을 관찰할 수 있는데 이는 1990년대 초 주택 200만호 건설 등의 영향을 받은 결과이다. 주택보급률은 2000년을 지나면서 비로소 100%를 넘어서게 되었다.



〈그림 3-28〉 연도별 주택보급률과 인구 천 명당 주택 수

주택보급률의 경우 %로 표시되므로 지표의 크기에 대해 어느 정도 감을 잡을 수가 있다. 그러나 인구 천 명당 주택 수의 경우 과연 제시된 숫자만을 보고 현재 우리나라의 위치가 어떤 수준인지 파악하기 어렵다. 이럴 때에는 다른 나라의 통계와 비교해보면 우리나라의 위치를 가늠해볼 수 있을 것이다. 하지만 전문가가 아닌 일반인이 각 나라의 통계정보들을 구하는 것은 그리 쉽지 않은 일이다. 정부에서 운영하는 e-나라지표 사이트에는 국가의 주요 지표들을 제공하는 동시에 관련된 주요국들의 지표도 소개하고 있다.

<그림 3-29>는 e-나라지표 웹사이트에서 '분야별지표>경제>부동산>주택>인구 천 명당 주택 수' 메뉴를 선택하는 화면을 나타내고 있다. 이를 클릭하면 인구 천 명당 주택 수라는 지표에 대한 상세한 해설 및 통계표를 얻을 수 있을 뿐 아니라 주요 국가들의 관련 지표 통계도 구할 수 있다. 다만 e-나라지표에서 제공되는 통계는 주로 국가 단위의 통계인 까닭에 보다 세부적인 지역별 비교를 하기는 어렵다는 한계가 있다.



(그림 3-29) e-나라지표의 인구 천 명당 주택 수 메뉴

토의 2

1990년 이후의 대전시 구별 주택보급률 및 인구 천 명당 주택 수 통계를 작성하고 비교하라.

앞의 「토의 1」에서는 주택보급률 및 인구 천 명당 주택 수 통계를 전국 단위로만 고려 했는데, 이번에는 통계의 작성단위를 대전시의 구별로 보다 세분화한 지역통계를 작성해야 한다.



## 3.7 정책과정에서 고려해야 할 통계 사례

다음에 소개되는 사례들은 정부의 각 부처에서 과거 실제로 검토되었던 문제들이다. 제시되는 각 정책들의 내용을 검토하고, 이들 정책의 집행이나 평가를 위해 필요하다고 생각되는 통계지표를 제안해야 한다. 먼저 각 문제별로 기존의 정책내용을 소개한 후 변경된 내용과 그로 인해 기대되는 효과를 제시한다. 다음으로 토론을 통해 새로운 정책 변경과 관련하여 필요한 통계지표들을 열거하고 각 지표들의 작성 가능성을 점검해야 한다.

### 사례 1 음식점 원산지 표시 정책

음식점 원산지표시제도는 적절한 품질관리를 통해 농산물의 상품성을 높이고 공정한 거래를 유도함으로써 농업인의 소득을 높이고 소비자도 보호하는 제도이다. 다음은 이 제도를 위한 정책들의 내용이다.

(1) (기존) 원산지 표시관리 및 단속을 농림수산물부와 보건복지가족부가 단계별로 나누어 담당

① 국내 생산 수입된 농축산물 및 그 가공품은 「농산물품질관리법」에 따라 농림수산물부(국립농산물품질관리원)가 담당

② 음식점에서 조리·판매되는 농축산물은 「식품위생법」에 따라 보건복지가족부(식품의약품안전청)와 지자체가 담당

(2) (신규) 음식점에서 조리·판매되는 농축산물의 관리 및 단속을 국립농산물품질관리원에서도 수행할 수 있도록 동 법령에 근거규정 마련

(3) (기대효과) 농 축산물의 생산(수입)부터 최종소비처(식당)까지 일관된 원산지표시 관리로 공정거래 질서를 확립하고 생산농가 및 소비자 보호

## (4) 필요한 통계

① 이 제도를 효율적으로 추진하기 위해 정책대상에 대한 기초적인 통계가 필요하고 이를 위해서는 음식점 원산지표시 대상 사업체 현황, 대상 품목의 수입 현황 통계를 관리할 필요

② 소비자의 선택권 보장, 축산농가 보호 등 이 제도의 근본취지를 달성하기 위해 제도의 준수현황 등에 대한 판단이 필요하고 이를 위해서는 음식점 원산지표시 단속인력 현황 및 위반단속 현황에 대한 통계를 관리할 필요

〈표 3-30〉 필요 통계지표

필요 판단	필요 통계지표	주요 항목	비고(출처 등)
대상 사업체현황 파악	음식점 원산지표시 대상사업체 현황	지역별 영업구분별 규모별 사업체수 등	접객업 현황 (보건복지가족부 등)
대상 수입농산물 현황	음식점 원산지표시 대상 품목 수입 현황	국가, 품목, 수입량 및 금액 등	검역정보시스템 (국립수의과학검역원) 무역통계(관세청)
단속인력 현황파악	음식점 원산지표시 단속인력 현황	특수사법경찰관리(농관원 직원), 식약청, 지자체 등	행정보고
제도 준수현황 파악	음식점 원산지표시 위반 단속현황	점검 및 위반 업소 수, 위반품목, 위반내용 등	개발예정 (실적자료)

## 사례 2 중소기업매출채권보험 제도와 중소기업 가업승계 지원 정책

### (1) 중소기업매출채권보험 가입대상 확대(안 제8항 신설)

- ① 기존 제도 내용으로 종전의 어음보험을 매출채권보험으로 확대하여 보험가입대상을 어음 외에 매출채권에까지 확대('04년) : 가입한도는 업체당 10억 원까지(매출채권보험 및 어음보험 포함), 상품별 한도는 매출채권보험 10억 원, 어음 8억 원
- ② 변경 정책 내용으로 중소기업매출채권보험 가입대상 확대 : 기존 매출액 150억 원 이하 기업만 가입이 가능하였으나, 매출액 150억 이상 중소기업도 이 제도상의 보험을 이용할 수 있도록 함. 약 11,000여 개 중소기업이 추가로 가입대상에 포함(가입대상 확대/변경)
- ③ 기대효과 : 매출채권보험 가입대상 확대로 중소기업의 거래 안정망 확보

### (2) 중소기업 가업승계 지원 정책 도입(안 제54조의2 신설)

- ① 가업승계 중소기업의 경쟁력 강화 및 고용안정을 위한 세제지원, 지원센터 운영 등 기업승계 원활화를 지원
  - 정책 대상 : CEO 55세 이상으로서 업력 15년 이상 중소기업
  - 지원 내용 : 가업승계 비용 완화를 위한 세제지원, 컨설팅 지원, 후계자 양성 등
- ② 기대효과 : 가업승계 중소기업의 경영안정과 지속적 성장에 기여

### (3) 필요한 통계

- ① 제도 실행과 관련한 매출채권의 규모, 보험료율 산출을 위한 관련 통계자료와 정책 평가를 위한 매출채권 가입중소기업 현황 파악이 필요하고 이를 위해 매출채권규모, 어음부도율, 매출채권보험 인수규모 및 사고율과 매출채권보험 가입 중소기업 현황에 대한 통계지표가 필요함. 특히 매출채권규모 파악과 보험료율 산정을 위해 지표들의 일부는 아래와 같이 산업별, 규모별 등으로 세분화하여 작성할 필요가 있음

〈표 3-31〉 매출채권규모 파악과 보험료를 산정할 위한 지표들

구분	중소기업 업체수	부도율	매출채권 (받을 어음)	매출액	보험료	대손상각비
산업별						
기업 규모별						
매출액 구간별						
지역별						

※ 산업별 : 제조업, 제조관련 서비스업, 제조관련 도매업, 지식기반서비스업  
 기업 규모별: 10인 미만, 10-19인, 20-49인, 50-99인, 100-199인, 200-299인,  
 300인 이상 매출액 구간별: 분석대상 기업의 매출액을 8-10 구간으로 나눔.

〈표 3-32〉 통계지표

필요 판단	통계지표	자료출처
보험인수대상 파악	매출채권규모	신용보증기금
보험료를 산출	어음부도율, 매출채권인수규모 및 사고율	신용보증 기금 및 한국은행
제도 성과측정	매출채권보험 가입 중소기업 현황	신용보증기금
제도 개선	매출채권, 부도율, 보험료 등에 대한 산업별, 규모별 등 상세 자료	없음 자료 * 신용보증기금을 통해서 기초자료 확보 가능

### 사례 3 수도권 이외 지역의 회원제 골프장에 대한 취득세 등 지방세 지원

#### (1) 수도권 이외 지역 회원제 골프장에 대한 세제 지원

##### ① 취득세 중과세 폐지(골프장용 토지·건축물·토지상 입목 대상)

<기존> 취득물건의 가액 또는 연부금액의 1,000분의 100 (10%)

<변경> 취득물건의 가액 또는 연부금액의 1,000분의 20 (2%)

##### ② 재산세 세율 인하(골프장용 토지·건축물 대상)

<기존> 과세표준액(시가표준액)의 1,000분의 40 (4%)

<변경> 과세표준액(시가표준액)의 1,000분의 20 (2%)

##### ③ 일몰제 적용

- 취득세 : 2008. 9. 1. ~ 2010. 12. 31.

- 재산세 : 2008. 6. 1. ~ 2009. 12. 31.

#### (2) 기대효과

##### ① 지방 회원제 골프장의 세부담을 완화하여 가격경쟁력을 높임

##### ② 골프비용 인하로 해외골프 여행객 감소 및 외화 절약

##### ③ 지역경제 활성화 및 고용효과 창출

⇒ 시행성과에 따라 지속 여부 및 수도권 확대 여부 판단(일몰제로 운영)

#### (3) 필요한 통계

① 필요 통계지표 검토 - 지방세 통계관련 : 취득세 및 재산세인하가 세수에 미치는 영향 판단 필요

② 골프장 이용업 활성화 관련 - 수도권 이외 지역 회원제 골프장에 대한 세제지원 성과를 평가하기 위해서는 아래 지표가 필요 (회원제 골프장의 증감현황, 골프장이용료 증감 및 내장객 현황, 해외골프여행객 감소 및 외화절약과 관련한 지표, 종사자수 증감현황, 골프장이용업 산업전반의 활성화 판단을 위한 지표 등)

③ 통계지표 구비여부 검토 - 소관부처의 의견

〈표 3-33〉 통계지표 구비여부 검토에 대한 소관부처의 의견

구 분	소관부처의견
통계의 필요성 여부	•제도 운영의 효율성 및 추진 실적을 평가하기 위한 통계지표 관리 필요
통계의 구비 여부	•지방세 관련 통계 구비

(검토의견) 소관부처에서는 취득세 및 재산세 인하로 인한 효과를 판단하는 지방세 통계관련 지표만 제시하였으나, 골프장 이용업 활성화 등과 관련해서는 아래 필요지표를 정책의 집행·평가에 활용할 필요가 있다고 판단됨

〈표 3-34〉 필요 통계지표

필요 판단	필요 통계지표	출 처
지방세 증감 판단	취득세 및 재산세 징수현황	지방세표작성보고 (행정안전부_국가승인통계)
골프장수 증감 판단	전국 골프장 현황 (회원/대중)	행정자료 (문화체육관광부)
골프장 내장객 증감 판단	골프장 내장객 현황 (시도별/연도별)	내부행정자료 (한국골프장경영협회)
골프장 이용료 증감 판단	품목별소비자 물가지수 (골프장 이용료: 시도별/연도별)	소비자물가지조사 (통계청_국가승인통계)
해외골프여행자 증감 판단	해외여행자 중 골프여행자 비중	국민해외여행실태조사 (한국관광공사_국가승인통계)
내국인출국자수 증감 및 관광수지 추이	내국인출국자수 현황 및 관광수지 현황	한국관광통계 (한국관광공사_국가승인통계)
골프장 이용업 고용증감 판단	골프장이용업 종사자수 현황 (시도별/연도별)	사업체기초통계조사 (시도, 통계청_국가승인 통계)
골프장 이용업 활성화 판단	서비스업생산지수(골프장 이용업 : 월별)	서비스업동태조사 (통계청_국가승인통계)

사례 4 사이버대학 설립 · 운영정책

(1) 사이버대학 운영의 이원화

- ① 평생교육 시설로서의 사이버대학 : 평생교육기관으로 학위 부여가 불가능함(평생교육법 적용)
- ② 고등교육 시설로서의 사이버대학 : 고등교육기관으로 학위 부여가 가능함(고등교육법 적용)

〈표 3-35〉 사이버대학 운영

구 분	평생교육시설	고등교육시설
설립주체	- 재단법인, 학교법인	- 학교법인
혜택	- 학위부여 불가능	- 학위부여 가능
	- 대학원설립 불가능	- 대학원설립 가능
	- 재정지원	- 재정·행정지원 및 산학협력지원
규제	- 회계 보고 의무 없음	- 회계 보고 의무화 및 관리감독 강화

(2) 고등교육기관으로의 전환 기준 마련

- ① 기존 사이버대학을 운영하는 재단법인인 경우 고등교육기관으로의 사이버대학 운영을 위해서는 학교법인으로 전환 신청 필요
- ② 사이버대학 및 특수대학원의 설립기준 마련
  - 교사·교원·수익성 기본 재산의 확보 기준 제시
  - 고등교육기관으로서의 사이버대학 기준 세부 내용
  - 교사기준 : 입학정원 기준으로 일정 면적 이상 확보

〈표 3-36〉 고등교육기관으로의 전환 기준

구분	교육 기본 시설	지원 시설
교사범위	대학본부 및 행정실, 교수실, 강의실, PC실습실, 세미나실, 콘텐츠개발실	서버 및 통신장비관리실, 시스템운영실
기준면적	입학정원에 따른 시설의 기준면적 - 1,000명 미만 : 990㎡ - 1,000명 ~ 2,000명 미만 : 1,485㎡ - 2,000명 ~ 3,000명 미만 : 1,980㎡ - 3,000명 이상 : 2,475㎡	

- 교원기준 : 전공별 학생 수 200명당 전임교원 및 조교 각각 1인 확보
- 재산기준 : 연간 운영 수익총액(1)의 50% 이상의 수익성기본재산(최소 35억 원(2) 이상) 확보
  - ※ 1) 학생수 × 학생 1인당 평균 연간수업료 × 1.5(대학의 수익용 기본재산 평균 확보율)
  - 2) 학생수 1,000명 기준

### (3) 필요한 통계

① 고등교육기관으로서의 사이버대학 설립·운영 정책의 집행과 평가를 위해서는 사이버대학의 현황과 그 성과에 대한 통계지표를 파악할 필요가 있음. 이를 위해 사이버대학의 교원 확보현황, 수익용 기본재산 확보현황, 교사(校舍) 확보현황, 사이버대학 등록률, 졸업자 취업률에 대한 통계지표가 필요

② 사이버대학 설립 및 운영정책의 효율적 운영에 필요한 지표는 아래와 같이 구비되어 있으나 정책의 성과평가가 지표로 “사이버대학 졸업자 취업률”이 구비되지 않은 것으로 판단됨

〈표 3-37〉 통계지표

필요 판단	통계지표	구비여부(자료출처)
사이버대학(원) 설립운영 기반여건	교원 확보현황	구비(교육과학기술부 행정자료)
	교사(校舍) 확보현황	"
	수익용 기본재산 확보현황	"
사이버대학 입학정원 증감 및 신규설립 필요성 판단	사이버대학 신입생 등록률	"
사이버대학 인력수급 전망 판단	사이버대학 졸업자 취업률	구비되지 않음

대학교육의 대중화 및 세계화, 지식정보화 등 사회적 변화에 맞춰 고등교육이 학문성을 지향하는 전통적인 정규고등교육의 기능과 직업교육 및 평생교육의 기능을 동시에 수행할 수 있도록 하기 위한 『고등교육정책(대학입시제도)』으로 그 내용은 다음과 같다.

③ 대학입학전형관련 권한 이양

[현행] 교육과학기술부장관 → [변경] 한국대학교육협의회장

〈표 3-38〉 대학입학전형관련 권한

내 용	비 고
- 대학입학전형 기본계획 수립 및 공표 권한	안 제32조
- 대학별 고사에 대한 심의 권한	안 제35조
- 학생 선발일정 수립 및 공표 권한	안 제41조
- 대입 지원방법 위반자 심의 및 처리 권한	안 제42조

④ 대학별 대학입학전형 시행계획 심의 및 시정요구권한 명시(안 제33조)

협의회는 대학별 대학입학전형시행계획을 심의하여 위법하고 부당한 경우 시정 요구 및 이행부 공표

[현행] 시정요구 불이행시 재원의 지원·보조 삭감, 연구비 등 중단 → [변경] 제재수단 규정 삭제

다만, 입시부정 등 중대한 비리사실이 드러날 경우 정부가 계속 제재

⑤ 입학지원방법 관련 규정 통합 및 정비(안 제42조) : 대학, 전문대학, 산업대학의 유사한 입학지원방법 통합·정비

## 토의사례 1

## 가족친화 사회 환경 조성정책

## (1) 추진방향

① 국가와 지자체에 대하여 가족친화 사회 환경 조성을 위한 탄력적 근무제도 등 각종 제도를 포함한 기본계획 수립·시행 의무화

② 사업주에 대하여 가족친화 사회 환경 조성을 위한 노력 의무화 → 복지부는 가족친화 인증기관 지정 등 가족친화 사회 환경 조성사업을 시행하고, 가족친화지수 작성 등 이행 상황을 평가

## (2) 가족친화 사회 환경 조성사업

① 가족친화 사회 환경\* 조성의 기본(시행)계획 수립·시행

(\* 가족친화 제도가 운영되고 있는 직장 환경)

② 가족친화 마을환경의 조성 촉진

③ 가족친화기업 인증 및 지원현황

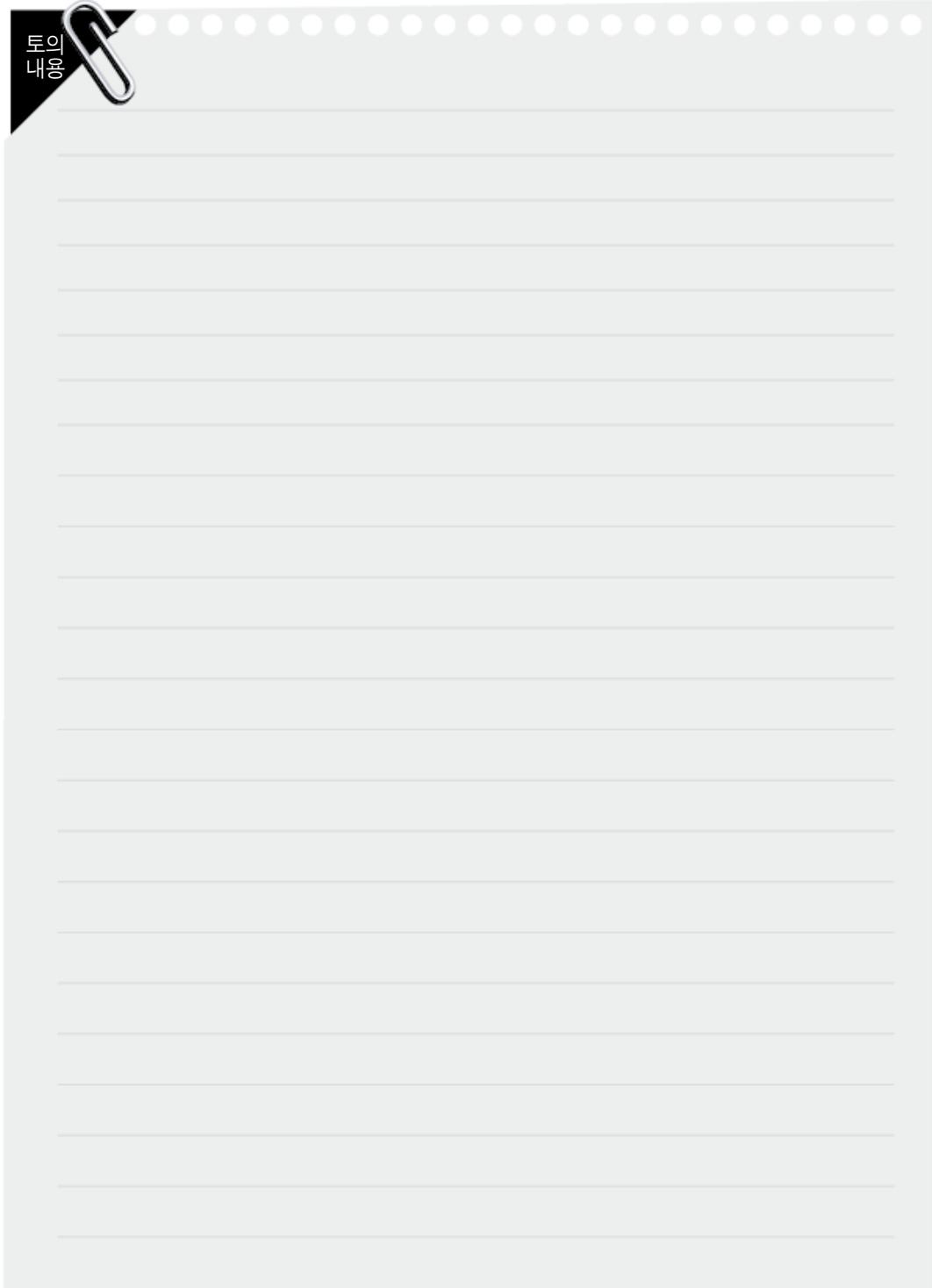
④ 가족친화 인증기관 및 가족친화지원센터 지정

⑤ 가족친화 사회 환경 조성 실태조사 실시

⑥ 가족친화지수 작성 및 공포

〈표 3-39〉 가족친화 사회 환경 조성사업

구 분	세 부 내 용
탄력적 근무제도	시차출퇴근제, 재택근무제, 시간제 근무 등
자녀의 출산양육 및 교육 지원제도	배우자 출산휴가제, 육아휴직제, 직장보육 지원, 자녀 교육지원 프로그램 등
부양가족 지원제도	부모 돌봄 서비스, 가족간호휴직제 등
근로자 지원제도	근로자 건강·교육·상담프로그램 등
가족관계 증진제도	자녀 방학 중 휴가제, 근로자 가족초청 행사, 정시 퇴근제, 육아데이, 가정의 날 등 운영지원
가족여가문화 촉진제도	가족단위 문화체험 지원
가족친화 사회공헌제도	한 부모 가족, 조손가족, 다문화가족, 독거노인, 장애인 가족 등 소외계층 지원제도



## 토의사례 2

## 농업기계용 면세석유류 공급정책

## (1) 면세유 공급대상 확대(영 제14조) : 개정

① 기존 정책에서 면세유 공급대상 농민의 범위는 개인농가, 영농조합법인, 농업회사법인, 조합이었음

② 신규 정책에서는 2개 이상 단위농협이 출자하여 설립한 '조합공동사업법인('07년 26개)'에게도 면세유를 공급하여 벼 등을 건조할 수 있게 지원

## (2) 면세유류 구입카드 도입(영 제16조 및 제20조) : 개정

① 기존 정책은 직전년도에 면세석유류를 10kl 이상 공급받은 농민은 직불카드 또는 신용카드를 사용할 의무, 그 외 농민은 종이 면세유류구입권 사용

② 신규 정책은 면세유 부정유통을 방지하고 면세유 업무를 전산화하기 위하여 면세유구입권을 면세유류 구입카드(직불카드 또는 신용카드)로 전환 - '08.7.1.부터 전 농민에게 면세유류 구입카드를 사용하도록 하고, 면세유류 구입카드 등의 사용지역을 경작지, 주소지 사구 등으로 제한

## (3) 석유판매업자의 환급절차 규정(영 제15조의 2) : 신설

① 석유판매업자가 면세유를 공급한 날의 다음달 15일까지 관할세무서장에게 감면세액을 환급 신청하도록 절차 마련

## (4) 면세유류 판매업자 지정절차 규정(영 제20조의2) : 신설

① 면세유류기관으로부터 면세유 판매업자로 지정 받을 수 있는 절차 규정

## (5) 면세유 공급명세의 웹사이트 공개(영 제20조의 3) : 신설

① 농협 웹사이트에 면세유 공급 관련 성명, 주소, 석유제품별 공급량, 농기계 보유현황을 공개할 수 있는 근거 마련



투입  
내용

## ■ ■ 요약

\_ 정책수립을 위한 통계지표를 마련할 때 가능한 한 현상을 다차원적으로 이해할 수 있는 복합지표를 마련하는 것이 바람직하다.

\_ 표본조사에 의해 통계를 작성해야 할 경우 조사방법의 결정, 표본설계, 추정, 조사관리 등 여러 부분에서 통계학적 이론에 입각하여 조사를 기획해야 한다. 기획과정에서 세심하고 전문적으로 고려하지 않았을 경우 막상 통계가 작성된다 해도 신뢰성, 정확성이 의심되므로 객관적인 증거로 활용되기 어렵다.

\_ 합리적으로 작성된 주택가격지수는 부동산 가격 폭등에 따른 투기지역 지정, 부동산 정책의 효과 파악 등에 유용하게 활용된다.

\_ 사교육비 관련 조사가 사교육비실태조사로 체계화됨으로써 사교육비 현황 및 특성을 파악할 수 있을 뿐 아니라 새로운 교육정책이 사교육비에 미치는 영향을 파악할 수 있게 되었다.

\_ 손상감시체계를 구축하기 위해서는 서로 다른 자료원으로부터 생성되는 데이터 베이스를 결합할 필요가 생긴다. 새로운 조사가 필요한 부분을 파악하여 통계를 작성함으로써 전체 통계를 작성할 수 있다.

\_ 사업단위 성과목표와 성과지표를 세울 때 올바른 인과관계를 설정하고 성과지표의 특성을 적절히 반영하는 것이 필요하다.

\_ 예산안 작성과 같은 일상적인 업무에서 간단한 그래프나 통계를 효과적으로 사용하여 새로운 제안의 타당성을 효과적으로 나타낼 수 있다.

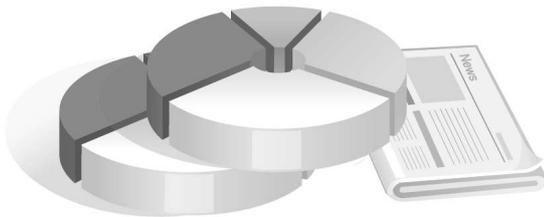
\_ 국가통계포털이나 e-나라지표를 활용하면 많은 유용한 통계정보들을 쉽게 검색할 수 있다. 이를 이용하여 각자 나름대로 필요에 맞는 통계지표들을 마련할 수 있다.



CHAPTER 4

# 정책과정과 자료

- 4.1. 통계적 사고
- 4.2 정책과정에서의 자료수집
- 4.3 자료의 표현
- 4.4 자주 잘못 사용되는 통계 용어들



## 개요

정책과정에 통계를 활용할 때 필요한 통계적 사고방식들을 사례 중심으로 소개하고, 자료를 효과적으로 표현하는 통계적 방법들을 소개한다. 아울러 잘못 사용하기 쉬운 통계용어들에 대해서도 설명한다.

## 주요내용

### 4.1 통계적 사고

#### 4.1.1 통계적 사고의 필요성

【사례1】 우리나라 최상위 집단의 납세액은 전문가 집단 납세액의 77%에 불과한가?

【사례2】 이혼한 여성은 이혼한 남성보다 훨씬 빨리 재혼한다?

【토의사례1】 저학력자들의 고용률은 끝없이 추락하는가?

【토의사례2】 가구당 도서구입비는 '0' ?

#### 4.1.2 통계적 사고가 필요한 경우들

### 4.2 정책과정에서의 자료수집

【사례1】 멜라민 부작용 조사

【사례2】 선거 예측조사의 역사 한 토막

### 4.3 자료의 표현

【사례1】 17대 국회의원의 재산액

【사례2】 고위공직자 재산가액 현황

【토의사례1】 우리나라 7개 대도시의 오존 오염도

【토의문제1】 21개국의 업종별 종사인구비율의 특징

### 4.4 자주 잘못 사용되는 통계 용어들

#### 4.4.1 지수의 정의와 3요소

#### 4.4.2 변동률(증감률)

【토의사례1】 특정 광물의 생산량

【토의사례2】 실업률

#### 4.4.3 진척률

【토의문제1】 진척률

#### 4.4.4 비, 비율, 율

【사례1】 이혼율에 대한 사회적 혼란을 잠재우려면?

#### 4.4.5 퍼센트와 퍼센트 포인트



## 4.1 통계적 사고

### 4.1.1 통계적 사고의 필요성

어떤 문제에 대해 통계적으로 사고한다는 것은 우선 자료에 기반을 두고 생각하는 것이다. 통계적 사고는 물리적인 법칙에 따라 결정되는 현상이 아니라 다양성과 불확실성을 포함하고 있는 현상을 다룰 때 필요하며, 바로 그 때문에 자료에 바탕을 두어야 한다. 통계적 사고방식은 불확실성과 다양성을 이해하고 표현함으로써 최선의 합리적인 의사결정을 수행할 수 있게 해준다.

정책 과정에서도 역시 문제 인식에서부터 정책 대안의 비교, 평가, 선택 나아가 선택된 정책의 집행과 평가에 이르기까지 적절한 자료를 구하여 그 자료를 정확히 이해하는 것이 중요하다. 문제에 대한 타당한 설명과 예측은 그러한 이해로부터 나올 수 있다. 여기서 자료를 잘못 읽은 몇 가지 사례를 통해 보다 정밀한 통계적인 사고의 필요성을 살펴보자.

**사례 1** 우리나라 최상위 집단의 납세액은 전문가 집단 납세액의 77%에 불과한가?

〈표 4-1〉 2005년도 한국의 월평균 납세액 자료(일부)

(단위 : 천원)

구분		전가구	근로자 가구	근로자 외 가구
의회의원, 고위임직원 및 관리자	조 세	171	264	135
	(가구비중)	(100)	(30)	(70)
전문가	조 세	223	243	126
	(가구비중)	(100)	(84)	(16)

출처 : 통계청(<http://nso.go.kr>)

이 자료를 근거로 하여 어느 기관에서 다음과 같은 주장을 하였다. 이 주장에 대한 당신의 의견을 기술해 보시오.

[ 기관의 주장 ]

통계청의 직업그룹 분류에 따르면 의회의원, 고위임직원 및 관리자는 전체 9개 직업군 가운데 최상위에 해당된다고 한다. 여기에는 입법부, 사법부, 행정부의 1급 이상 공무원, 기업체의 고위직 임직원, 국회의원, 지방의원, 구청장 그리고 부시장급 이상의 지자체 고위직 등이 속한다. 그런데 이들 그룹이 올 들어 9월까지 납부한 세금은 월평균 17만 1,201원으로 전문가 그룹이 낸 월평균 세금 22만 2,827원의 76.8%에 불과하다.

(1) <표 4-1>에 나타난 정보를 토의해 보자.

- ① 이 기관에서는 통계청 직업그룹 분류에 따른 최상위 집단과 전문가 집단의 월납세액을 비교하고 있다. 그런데 이 표를 해석하는 과정에서 ‘의회의원, 고위임직원 및 관리자’를 최상위 그룹으로 일컫는 것은 적절하지 않다. 당시 통계청이 해명자료를 통해 밝혔듯 「최상위 계층」은 소득수준에 따라 구분할 때의 범주로서 직업분류에 따라 구분했을 때의 범주가 아닌 것이다.
- ② 의회의원, 고위임직원 및 관리자 그룹의 월평균 납세액은 171,000원이고, 전문가 그룹의 월평균 납세액은 223,000원이다.
- ③ 의회의원, 고위임직원 및 관리자 그룹의 30%는 근로자가구인 반면, 전문가 그룹의 경우 84%가 근로자이다. 즉 의회의원, 고위임직원 및 관리자 그룹의 70%가 근로자가 아닌 반면, 전문가 그룹은 대부분이 근로자 가구이다.
- ④ 근로자 가구의 월평균 납세액만 비교하면 전체 가구에서 나타난 것과는 달리 의회의원, 고위임직원 및 관리자 그룹이  $(\frac{264}{243}-1) = 8.9\%$  더 많이 냈다.
- ⑤ 근로자가 아닌 가구에서도 마찬가지이다. 의회의원, 고위임직원 및 관리자 그룹이  $(\frac{135}{126}-1) = 7.1\%$  더 많이 냈다.

**(2) 모순된 사실을 설명해 보자.**

이 기관이 표로부터 얻고자 하는 주요 정보는 두 집단(의회의원, 고위임직원 등으로 이루어진 관리자 그룹과 전문가 그룹)의 비교결과이다. 이러한 주제를 다루는 분야를 통계학에서는 「두 집단 비교」라고 부르는데 통계학에서 상당히 중요한 주제이다. 그런데 두 집단을 비교하기 위해서 가장 중요한 것은 두 집단을 구별하는 특성(여기서는 ‘의회의원, 고위임직원 및 관리자’와 ‘전문가’라는 직업)를 제외하고는 납세액과 관련될 가능성이 있는 나머지 어떤 특성에서도 두 집단이 비슷해야 한다는 점이다. 이 사례는 그렇지 못했을 경우 두 집단 비교가 온전하게 이루어지지 않는 사례로서 무엇보다 두 집단의 근로자 가구 비율이 서로 크게 다르다. 그 결과 근로자 가구 여부를 고려하면 국회의원, 고위임직원 및 관리자 그룹이 전문가 그룹보다 월평균 더 많은 세금을 내고 있다. 앞서 전문가 그룹이 세금을 더 많이 내고 있는 것처럼 보인 것은 근로자 가구 여부를 고려하지 않았기 때문이었던 것이다.

사례 2

이혼한 여성은 이혼한 남성보다 훨씬 빨리 재혼한다?

2007년 통계청에서는 이혼과 재혼 연령에 대해 다음과 같은 발표를 한 적이 있다.

지난해 기준 우리나라 남자의 평균 이혼 연령은 42.6세, 평균 재혼 연령은 44.4세로 나타났고, 여성의 경우 평균 이혼 연령은 39.3세, 재혼 연령은 39.7세로 나타났다.

이러한 발표에 대해 어느 기관에서 다음과 같은 주장을 내놓았다.

[ 기관의 주장 ]

이혼에서 재혼까지 남자 22개월, 여자 5개월이 걸린다. 이혼의 아픔을 잊고 다시 새 짝을 찾기까지 걸리는 시간이 남자는 평균 21.6개월(1.8년), 여자는 4.8개월(0.4년) 가량인 것으로 나타났다.

(1) 발표내용에서 얻은 정보를 토의해 보자

- ① 평균 이혼/재혼 연령의 의미는 2006년 한 해 동안 혼인상태에 있다가(사별이나 이혼상태에 있다가) 이혼(재혼)한 남자와 여자들의 평균 연령이다.
- ② 남자의 경우, 평균 재혼 연령 44.4세와 평균 이혼 연령 42.6세의 차이인 1.8년은 재혼하는 사람들의 평균 연령과 이혼하는 사람들의 평균 연령을 단순히 뺀 값일 뿐이다. 여기서 '단순'하다는 것은 2006년 이라는 한 해에 이혼한 사람들 집단과 그해에 재혼한 사람들 집단의 평균 연령 비교 이상의 의미가 없다는 것이다.

(2) 이혼에서 재혼까지의 기간이 남자는 평균 1.8년(22개월), 여자는 평균 0.4년(5개월)이라는 해석이 올바른지 토의해보자.

- ① 평균이 0.4년으로 나오려면 0.4년보다 짧은 기간에 재혼하는 사람들도 어느 정도 있어야 하는데, 그런 경우는 우리 사회에서 쉽게 볼 수 없다. 따라서 통계적으로 생각하지 않더라도 이 해석에 문제가 있음을 알 수 있다.
- ② 이런 질문을 던져보자.
  - i “2006년도에 이혼한 사람들 집단과 2006년도에 재혼한 사람들 집단에 모두 속하는

사람들이 얼마나 될까?”

ii “2006년도에 재혼한 사람들이 언제쯤 사별이나 이혼을 경험했는지 알 수 있을까?”

질문 i에 대해서는 추측이지만 아마도 이혼한 사람들 중 그렇게 많지 않은 일부만이 재혼한 사람들 집단에 포함될 것이다. 따라서 두 집단은 서로 다른 사람들로 구성되어 있을 것이다. 질문 ii는 이들을 추적 조사하여 이혼이나 사별시점을 알아야만 답할 수 있다.

그럼에도 불구하고 이 기관의 주장에서는 그러한 차이를 무시하고 마치 이혼한 사람으로 이루어진 집단과 재혼한 사람들의 집단이 같은 것처럼 간주하고 있다. 그런 주장을 하기 위해서는 이혼한 사람 개개인을 추적하여 재혼까지의 기간을 알아야 할 것이다. 이 문제도 역시 비교의 문제이다. 그런데 여기서 비교는 동일인에게 발생한 두 개의 사건, 이혼 또는 사별과 재혼의 발생시점을 비교하는 문제이다. 통계학에서는 이를 쌍체비교(대응비교, 짝을 이룬 자료 분석)라고 부른다. 이 기관에서는 서로 다른 개체들로 구성된 두 집단(이혼자 집단과 재혼자 집단)의 평균을 비교했음에도 불구하고, 해석은 쌍체비교를 한 것처럼 잘못 해석하였다.

올바른 해석을 하기 위해 이혼상태에 있던 사람 중에서 2006년도에 재혼한 남녀의 혼인해소 일자를 통계청 자료에서 추적 조사하여 이혼일자와 재혼일자의 차이를 평균 낸 결과, 남자는 평균 3.7년, 여자는 3.6년으로 나타났다.

이와 같이 다양한 사례들로 이루어지는 집단을 비교하는 것은 통계학에서 매우 중요한 주제인데 이때 비교되고 있는 집단들을 혼동하면 실제와 매우 다른 주장을 하게 된다.

〈표 4-2〉 2005/2006년도의 교육정도별 경제활동인구의 고용률

(단위 : %)

학력	2005년도 고용률	2006년도 고용률
초졸이하	43.7	43.1
중졸	44.4	43.7
고졸	62.7	62.7
대졸이상	75.2	75.4
전문대졸	74.3	75.3
대학교졸이상	75.6	75.5
전체	59.7	59.7

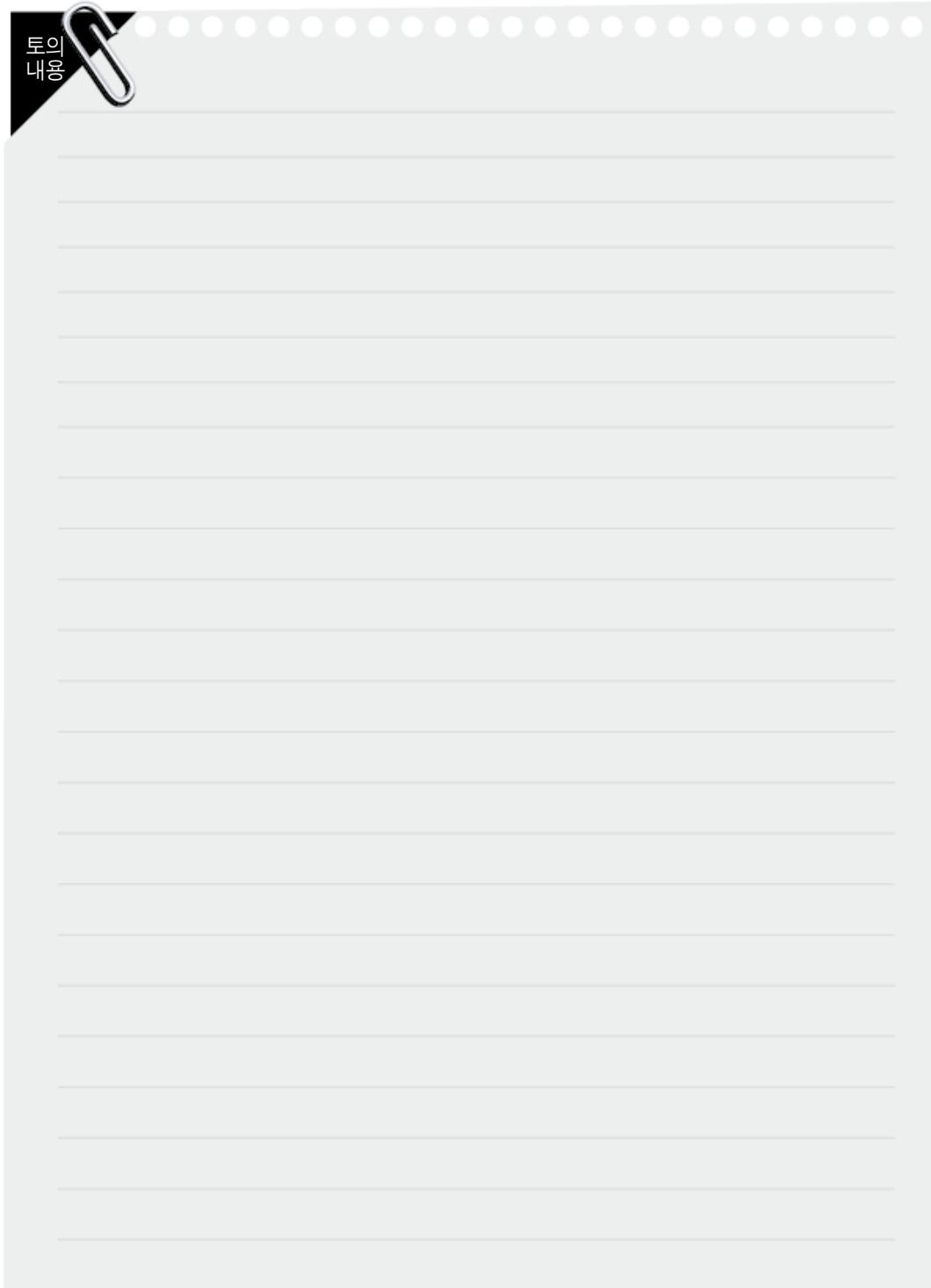
출처 : 통계청(<http://nso.go.kr>)

이 자료를 근거로 하여 어느 기관에서 다음과 같이 주장하였다. 이 주장에 대한 당신의 의견을 기술해 보시오.

[ 기관의 주장 ]

2006년 현재 초등학교나 중학교를 졸업하는 데 그친 저학력자들의 고용률은 각각 43.1%(초졸이하), 43.7%(중졸)로 1년 전보다 0.6%포인트, 0.7%포인트씩 하락한 것으로 나타나서 초·중등학교 이하 저학력자들의 고용률이 관련 통계 작성 이후 최저치로 떨어졌다. 고용률은 생산가능인구(15세이상) 가운데 취업자가 차지하는 비율로서 한 나라 경제의 고용 기동률과 노동공급 현황을 나타낸다. 결국 15세 이상 인구 중 중졸이하 저학력자들의 60% 정도가 직장을 구할 의사가 없거나 구직을 아예 포기하거나 또는 실업상태에 빠져 있다는 뜻으로서, 저학력자들의 고용률은 끝없이 추락하고 있다고 하겠다.

투  
의  
내  
용



This form area is designed for handwritten input. It features a series of horizontal lines for writing, a perforated top edge, and a paperclip icon. The text '투의내용' (Tui-ui Naeyong) is positioned in the top left corner, indicating the intended use for recording meeting content.

2004년 통계청 자료 중 가구당 월평균 지출액에 대한 다음과 같은 표가 있다.

〈표 4-3〉 2004년도 서적 및 인쇄물 관련 가구당 월평균 지출액

년도	구 분	서적 및 인쇄물	구분		
			일간신문	서적	잡지 기타 인쇄물
2004년도	지출액(원)	10,931	2,755	7,750	426
	(구성비, %)	(100.0)	(25.2)	(70.9)	(3.9)
2005년도	지출액(원)	10,590	2,459	7,667	464
	(구성비, %)	(100.0)	(23.2)	(72.4)	(4.4)
2004년 3분기	지출액(원)	10,148			
	(구성비, %)	(100.0)			
2005년 3분기	지출액(원)	10,397			
	(구성비, %)	(100.0)			

출처 : 통계청(<http://nso.go.kr>)

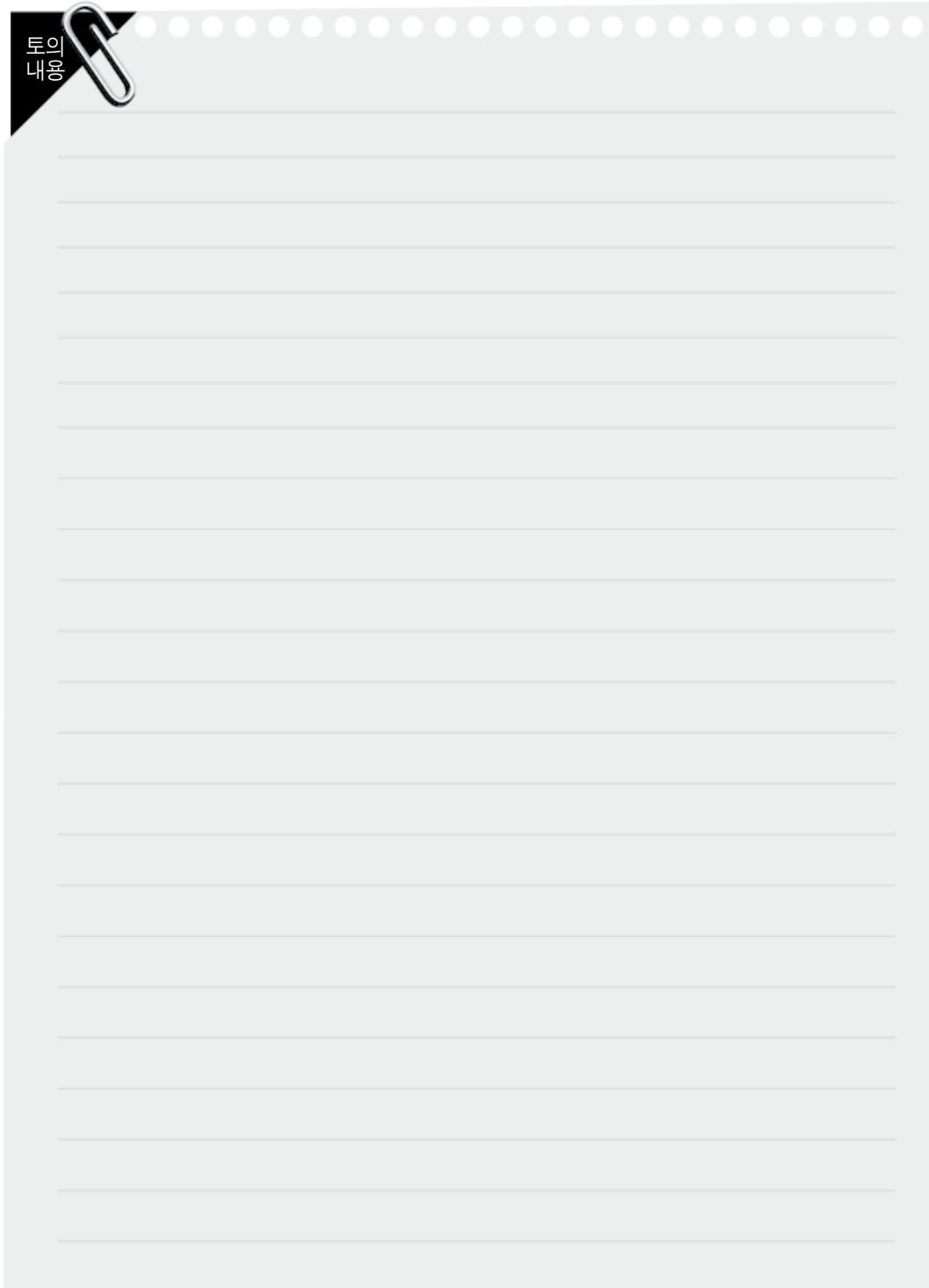
이 자료를 근거로 하여 어느 기관에서는 다음과 같이 주장하였다.

[ 기관의 주장 ]

아·미용 장신구비 지출액은 월평균 5만 9,611원으로서 서적 및 인쇄물 지출액의 5.7배에 이르는 규모이다. 서적 및 인쇄물 지출액은 신문과 잡지뿐 아니라 자녀들의 동화나 교양서적 구입비까지 포함하고 있다. 신문구독료가 월 1만 2,000원에 이른다는 점을 감안하면 성인들의 책 구입비는 사실상 「제로(0)」 수준으로 추정된다.

이 주장에 대해 당신의 의견을 기술해 보시오.

투  
의  
내  
용



## 4.1.2 통계적 사고가 필요한 경우들

### (1) 다양성의 인식과 정확한 표현

일상생활에서 우리는 다른 사람들이 자신과 같은 생각을 가질 것이라고 여긴다든지 또는 음식점에서 식사를 주문할 때 상대방이 선호하는 음식이 나와 다를 수 있다는 생각을 하지 못한다든지 또는 나의 지지 여부로 상대방을 판단한다든지 또는 다름과 틀림을 구분하지 못하는 등 다양성을 제대로 인식하지 못할 경우가 많다. 또한 “xx대학교 취업률 1위”와 같은 광고 문구를 별다른 느낌 없이 받아들이기도 한다. 통계적 사고라는 기준으로 보았을 때 이 광고에만 하더라도 따져볼 것들이 적지 않다. 먼저 “1위”라는 말 속에는 어느 집단에서인지, 예컨대 “전국 4년제 대학 중에서인지”, “전국 2년제 대학 중에서인지”, “특수 전문대학 중에서인지”, “어느 지역에 있는 대학들 중에서인지”가 명시되지 않는 경우가 많다. 또한 “취업률”이라는 것도 “취업”을 어떻게 정의하고, 졸업생들의 취업 여부를 어떻게 조사하였는지 확인해볼 필요가 있다. 또 대학의 취업률을 누가 어떻게 비교를 하였는지도 중요한 문제이다. 혹시 XX대학교 자신이 비교한 것은 아닐지 알아보아야 한다. 만약 취업에 대한 정의가 다르다면(예컨대 3개월 이상 근무한 것을 취업이라고 볼지, 6개월 이상 근무한 것을 취업이라고 볼지, 인턴도 취업으로 간주할지 등등) 이 비교는 아무런 의미가 없는 것이 될 수도 있고 도리어 “1위”라는 허상 때문에 청소년들에게 피해를 입힐 수도 있다.

다른 사례로 여러 토론장에서 자주 나타나는 “국민 대부분이 반대” 한다는 표현 역시 통계적 사고와 거리가 먼 모호한 주장이다. 대부분이란 얼마를 뜻할까? ‘대부분’이라는 표현을 60%, 70% 등의 정량적인 표현과 함께 사용하는 것이 보다 바람직한 통계적 사고방식일 것이다.

### (2) 다양성의 해석과 표본자료

타올 4할의 타자 A와 타율은 6할이지만 신인인 타자 B를 타율만으로 서로 비교하는 것과 마찬가지로 비교들을 우리 주위에서 많이 보게 된다. 최근 어느 기사에 따르면 “인터넷 조사에 응답하는 비율을 비교해보면 가구주가 고졸 출신인 가구가 가구주가 박사 출신인 가구보다 20% 포인트 앞선다”고 한다. 여기서 비율(percent)에 근거를 둔 정보를 해석할 때 유의할 점이 있다. 총 몇 타석 중에서 6할이며, 총 몇 가구가 응답한 결과인가를 살펴야 한다는 것이

다. 예컨대 극단적으로 타자 A는 1000 타석 중 4할이고, 타자 B는 10 타석 중 6할이라면, 또는 인터넷 조사 응답율이 가구주가 고졸인 3000가구 중 40%였고 가구주가 박사인 50가구 중 20%가 되어서 두 비율이 차이가 났다면 이 수치들을 단순히 비교해서 해석하는 것은 중요한 판단의 오류를 범할 수 있다는 것이다.

한편 어떤 기관에서 민원인을 대상으로 만점이 100점인 만족도 조사를 한 결과를 보도하면서 “금년도 평균이 작년에 비하여 1점 상승에 그쳤다”고 했다면 이 보도를 해석하기란 상당히 어렵다. 먼저 이 보도에서 민원인의 평균만족도가 무엇을 의미하는지부터 분명하지 않다. 그 기관이 접촉하였던 민원인과 접촉할 수도 있는 가능 민원인들까지를 모두 포함한 소위 민원인 모집단의 평균만족도를 의미하는가? 아니면 그 조사에 표본으로 추출된 200명의 평균만족도를 의미하겠는가? 아마도 문맥에서, 그리고 실망을 유도하는 듯한 문구에서 보면 전자 즉, 민원인 모집단의 평균만족도를 의미한다고 볼 수 있겠다.

그렇다면 평균 1점 차이는 어떻게 얻었을까? 특별히 그 기관과 접촉한 민원인을 모두 조사한 것이 아니라면 표본(예컨대 작년 150명, 금년 200명)에서 얻은 평균값의 차이일 것이다. 이제 우리의 질문은 평균 1점 상승이 기관 입장에서 볼 때 실망스러운 정도로 낮은 수준인가 하는 문제이다. 만일 단지 한 민원인의 만족도 점수가 작년에 한 민원인의 만족도 점수에 비해 2점 차이라면 그 정도 차이는 거의 무시할만하다고 볼 것이다(사실 이것도 그렇게 간단한 이야기는 아니지만 만족도 조사를 하면 민원인에 따라 몇 점씩 차이가 나는 것은 일반적인 현상이라고 보기로 한다면 받아들일 수 있다). 하지만 150~200명으로 된 두 집단에서 얻은 평균의 차이가 2점이었다면 전체 민원인들의 작년도 평균만족도와 금년도 평균만족도의 차이를 작은 차이로 보아야 하는 것인지, 상당히 큰 차이로 보아야 하는 것인지 판단하기란 결코 쉬운 일이 아니다.

따라서 200여명의 표본에서 얻은 평균값을 마치 한 개인으로부터 관찰한 값을 대하는 것처럼 큰 값, 작은 값으로 판단하는 것은 위험하므로 200여명의 표본에서 얻은 결과를 전체 집단에 적용하려고 할 때에는 특별히 유의해야 한다.



## 4.2 정책과정에서의 자료수집

통계 분석을 위한 자료를 수집하는 방법에는 실험(experiment)과 조사(survey)가 있다. 일반적으로 자연과학, 의학, 공학 등의 분야에서는 실험을 통해 자료를 얻지만 사회과학에서는 대부분 실험이 불가능하기 때문에 조사를 통해 자료를 얻는 것이 보통이다. 조사에는 전수조사와 표본조사가 있으며, 그 중 전수조사는 인구주택총조사와 같이 모집단의 모든 단위들을 다 조사하는 방법으로 흔히 센서스라고도 불린다. 반면 표본조사는 표본으로 선택된 단위들만 조사하는 방법으로 전수조사에 비해 경제적이고 효율적인 조사방법이다.

물론 표본추출은 조사뿐만 아니라 효율적인 실험을 위해서도 유용한 방법인데 모집단을 잘 대표하는 표본을 얻기 위해서는 체계적이고 과학적인 표본설계가 필요하다. 만일 사용자의 편리에 따라 표본을 추출한다면 어떤 문제점들이 생길지 간단한 사례를 통해 살펴보기로 하자.

### 사례 1 멜라민 부작용 조사

2008년 문제가 되었던 멜라민의 부작용을 조사하기 위해 실험용 흰 쥐를 대상으로 실험을 실시하고자 한다. 연구자는 흰 쥐 열 마리를 상자 속에 가두어 놓고 휘저은 뒤 먼저 다섯 마리를 잡아 멜라민이 없는 사료를 주고, 남은 다섯 마리에는 멜라민이 포함된 사료를 투여한 후 진통 지속시간을 관측하기로 하였다. 이 실험의 문제점을 지적하라.

쥐를 다섯 마리 잡는 과정에서 상대적으로 둔한 쥐들이 먼저 잡히기 쉬우므로 이 실험은 공정한 실험이라고 할 수 없다. 공정한 실험을 하려면 먼저 각 쥐에게 1에서 10까지 숫자가 기록된 꼬리표를 달아 준 후, 난수표를 이용하거나 엑셀의 표본추출 기능을 이용하여 숫자 다섯 개를 추출한다. 뽑힌 숫자에 해당하는 쥐에게는 멜라민이 없는 사료를 주고 나머지 쥐들에게는 멜라민이 포함된 사료를 투여한다.

### 사례 2 선거 예측조사의 역사 한 토막

1936년 미국 대통령 선거를 앞두고 리터러리 다이제스트(Literary Digest)라는 한 잡지사에서 대규모 여론조사를 실시하였다. 조사방법은 그 잡지구독자 명부와 전국의 전화번호부, 그리고 자동차 등록대장에 기록된 성인들 약 1000만 명에게 공화당의 랜돈(Landon) 후보와 민주당의 루즈벨트(Roosevelt) 후보에 대한 지지여부를 묻는 우편조사였다. 응답을 보내온 230만 명 중 57%가 랜돈 후보를, 43%가 루즈벨트 후보를 지지하였으므로 그 잡지사에서는 랜돈 후보가 낙승할 것으로 예상하였다. 그러나 실제 선거의 개표결과, 예상과는 정반대로 루즈벨트 후보가 압도적인 지지(62.5%)를 얻어 승리하였다. 리터러리 다이제스트사의 예측조사가 실패한 이유를 설명하라.

이 조사에 대해서는 두 가지 문제점을 지적할 수 있다. 당시 랜돈 후보는 상류층이 압도적으로 지지했던 반면 루즈벨트 후보는 서민층이 압도적으로 지지하고 있었다. 전체 유권자 가운데에는 서민층 유권자가 상류층의 유권자에 비해 월등하게 많았지만 잡시사가 뽑은 표본에서는 상류층의 비율이 서민층 비율에 비해 높았다. 당시 전화를 소유하거나 자동차를 소유한 사람들은 상류계층이었던 것이다.

예측 실패의 또 다른 요인은 응답률에 있었다. 이 조사의 응답률은 23%에 불과했는데, 상대적으로 랜돈 후보 지지자들의 응답률이 루즈벨트 후보의 응답률보다 높아 추정의 편향(bias)이 초래되었던 것이다.



## 4.3 자료의 표현

숫자로만 된 많은 통계자료들은 우리가 잘 꿰어서 목걸이나 팔찌로 만들어야 할 구슬과 같다. 산만한 자료 자체로는 그 전체적인 모습과 자료들 사이의 관계를 파악할 수 없으므로 자료로부터 필요한 정보를 얻기 위해서는 우선 자료를 요약하고 정리하는 방법을 이용해야 한다. 통계학에서 많이 쓰이는 요약, 정리 방법에는 수치를 이용하여 자료의 특성을 표현하는 방법과 표나 그래프를 통해 전체 자료의 분포형태를 나타내는 방법이 있다. 먼저 수치를 통해 특성을 표현하는 경우에는 각 수치들의 특성을 정확히 이해하는 것이 중요하다. 그리고 표나 그래프를 이용하는 경우에는 자료의 전체적인 형태는 물론 전체적인 형태에서 벗어나는 부분에 대해서도 주목할 필요가 있다.

통계분석에서 모든 자료는 숫자로 표시된다. 이때 그 자료가 갖는 값이 숫자로서 의미가 있는지 여부에 따라 자료를 다음과 같이 둘로 나눌 수 있다.

- 질적 자료 : 자료의 값이 숫자로서의 의미를 갖지 않고 범주만을 나타내는 자료.
- 양적 자료 : 자료의 값이 숫자로서의 의미를 갖는 자료.

질적 자료의 예로는 정책지지여부를 조사한 자료를 들 수 있다. 정책에 대한 찬성을 1로, 반대를 2로 표시한 조사 자료가 있다고 할 때, 여기서 1과 2는 수로서의 의미는 없고 다만 찬성과 반대를 대신하는 것이다. 그밖에 종교, 지지 정당 자료 등도 질적 자료이다. 그리고 생활수준을 상, 중, 하로 표현하거나 연령을 10대, 20대, 30대 등으로 표현한다면, 이런 자료도 그 사람이 속한 범주를 나타낼 뿐 직접적인 수를 나타내지 않으므로 질적 자료가 된다. 한편 양적 자료의 예로는 특정 지역의 주민만족도를 알아보기 위해서 100점 만점으로 조사한 자료를 들 수 있는데, 이때 주민만족도 값은 수의 의미를 갖는다.

한편 질적 자료 가운데 종교라든가 지지 정당은 순수하게 범주로서의 의미만 갖는데 이런 자료를 명목형 자료(nominal data)라고 한다. 반면 상, 중, 하로 수집된 생활수준과 10대, 20대, 30대 등으로 수집된 연령은 절대적으로 비교할 수 있는 수로서의 의미는 없지만 범주 사이에 '크기 순서(또는 높고, 낮고)'를 갖고 있다. 이런 질적 자료를 순서형 자료(ordinal data)라고 한다.

자료의 양이 클 때 자료를 요약하지 않고 그 특성을 파악하는 것은 매우 어려운 일이다. 가령, “통계청 발표에 의하면 2006년도 도시근로자가구의 월 평균소득은 3,443.4천원이며, 상위 20%의 평균소득은 5,237.5천원이다” 라는 보도를 접했다고 하자. 여기서, “평균소득”과 같이 자료를 요약한 수치를 기술통계량(descriptive statistics)이라 하며, 중요한 기술통계량 들에는 다음과 같은 것들이 있다.

〈표 4-4〉 기술통계량

대표값	산술평균, 기하평균, 조화평균, 중위수, 최빈수
산포도	범위, 사분위수범위, 표준편차, 분산, 변동계수
기타	다섯숫자요약, 표준화 점수

### (1) 자료의 가운데 위치는 어디인가?

자료의 중심 위치를 알아보기 위해 자료를 요약하는 대표값에는 평균, 중위수, 최빈값 등이 있는데 어떤 자료에서나 항상 최선인 대표값은 없다. 그러므로 각 대표값의 특징을 이해 하여 자료에 따라 적절한 대표값을 골라 이용하는 것이 바람직하다.

#### 가. 평균

평균에는 산술평균 이외에 절사평균, 가중평균, 기하평균, 조화평균 등 여러 가지가 있다.

##### ① 산술평균

가장 널리 사용되는 대표값인데 종종 신중하게 사용할 필요가 있다.

### 사례 1 17대 국회의원의 재산액

2004년 4월 15일에 치러진 17대 국회의원 선거 다음날, 거의 모든 신문은 다음과 같은 내용의 기사를 실었다.

「국회의원 당선자 299명의 나이는 평균 51세이고 평균 재산액은 21억 700만원」

이 보도에 대하여 토론해 보자.

21억 700만원은 당선자 299명이 신고한 재산의 산술평균이다. 그런데 어떤 한 사람을 제외하고 298명의 평균을 구하면 12억 5,024만원이 되어 평균값이 9억 원 가까이 떨어진다. 그 단 한 사람은 누구일까? 2567억 8,320만원을 신고한 정몽준 당선자였다. 그의 재산은 나머지 298명과 매우 다른 값으로서 평균값에 너무 큰 영향을 미치기 때문에 21억은 299명 전체의 재산신고액을 대표하는 값으로는 부적절한 값이다. 이런 이유 때문에 다른 대표값들이 쓰이는데, 그중 하나가 절사평균이다.

#### ② 절사평균

체조나 피겨스케이팅 경기의 심판들이 내릴지도 모르는 편파적인 판정을 방지하기 위하여 사용된다. 절사평균은 전체 자료에서 상, 하위 10%(또는 5%, 1%)에 해당하는 값들을 제거한 나머지 값들만 가지고 평균을 낸 것이다.

앞의 17대 국회의원 당선자의 재산 자료로부터 5% 절사평균을 구한다면 다음과 같다. 17대 국회의원 당선자의 재산에서 하위 5%를 제거하고, 상위 5%를 제거한 다음에 평균을 계산한 값으로 9억 8,842만원 정도가 된다.

#### ③ 가중평균

관측치들의 중요도나 영향력에 해당하는 가중치를 각 값에 곱한 뒤 평균을 계산한 것이다. 물가지수가 그 대표적인 사례이다. 현재 통계청에서는 2005년을 기준년도(100.00)로 2005년 도시가계의 월평균 소비 지출액 185만원 중에서 1/10,000이상 지출되는 489개 품목을 선정하여 소비자물가지수를 산출한다. 우리나라를 비롯한 대부분의 국가에서는 지수작성의 신속성과 경제성을 고려하여 라스파이레스지수(Laspeyres's index)를 사용하는데, 이를 간단한 예로 보면 다음과 같다.

아래 표에 품목별지수와 가중치가 나와 있다. 여기서 품목별지수는 현재가격을 2005년도 가격으로 나눈 값에 100을 곱한 것이고 가중치는 소비지출금액이다.

〈표 4-5〉 품목별지수와 가중치

품목	2005년도 가격	2005년도 거래량	현재 가격	품목별 지수	가중치
쌀	1000/kg	10kg	1000/kg	100	10,000
사과	100/개	5개	200/개	200	500
기름	200/l	5l	600/l	300	1,000

물가지수를 품목별 지수들의 산술평균으로 구하면  $(100+200+300)/3=200$ 이므로 2005년에 비해 물가가 2배 상승한 셈이 된다. 그러나 이는 각 구성품목의 상대적 중요도를 무시한 값이다. 따라서 각 품목의 소비지출금액을 가중치로 품목별 지수들의 가중평균을 구하면

$$(100 \times 10,000 + 200 \times 500 + 300 \times 1,000) / (10,000 + 500 + 1,000) = 1.22$$

가 되므로 물가수준이 2005년 대비 1.22배 상승했다고 할 수 있다. 그런데 금액으로 표시된 통계에서는 현재 금액을 기준 시점의 금액으로 환산할 필요가 있다. 이 경우 소비지물가지수를 이용한 불변화폐액이 쓰이고 있는데 다음과 같이 구한다.

$$\text{불변화폐액} = (\text{현재 화폐액} / \text{현재 소비자물가지수}) \times 100$$

즉 불변화폐액은 일정한 구매력을 지닌 화폐액을 뜻한다. 2005년부터 2007년도까지 한국의 일인당 국민소득, 소비자물가지수, 불변화폐액은 아래 표와 같다.

〈표 4-6〉 2005~2007년도의 한국의 일인당 국민소득, 소비자물가지수, 불변화폐액

(단위 : 만원)

년도	일인당 국민소득	소비자물가지수 2005=100	불변화폐액
2005	1,681	100	1,681
2006	1,758	102.2	1,720
2007	1,863	104.8	1,778

일인당 국민소득이 2005년 1681만원에서 2007년 1863만원으로 182만원 올랐지만 이는 단순한 금액의 증가일 뿐 실제 구매력이 늘어난 것을 의미하지는 않는다. 그 동안의 물가상승을 감안한 불변화폐액을 살펴보면 2005년에 비해 2007년의 실제구매력의 증가액은 단지 1778만원-1681만원=97만원으로 국민소득 증가액의 절반정도임을 알 수 있다.

④ 그 이외에도 인구의 변동률이나 물가변동률과 같은 비율의 대표값으로 자주 이용되는 기하평균이 있고 조화평균도 있다.

## 나. 중위수

## 사례 2 고위공직자 재산가액 현황

전체 자료들을 크기순으로 나열했을 때 중앙에 위치하는 자료를 말하며 중앙값이라고도 부른다. <표 4-7>은 정부공직자윤리위원회가 2008년 5월 7일 고위공직자의 재산등록 현황에 근거해 공개한 고위공직자 73명의 재산가액 현황 자료이다.

이 자료의 중위수를 구하려면 먼저 자료를 오름차순으로 정렬해야 한다. 그 가운데 중위수는 37번째 값으로 1,181,001천원이다. 그런데 이 자료의 산술평균은 1,766,963천원으로 그 차이가 414,037천원이다. 두 값의 차이에 대해 토의해 보자.

〈표 4-7〉 고위공직자 재산가액 현황

(단위 : 천원)

2,124,649	4,149,142	2,219,388	9,731,559	772,946	2,003,161	416,519	1,485,640
501,308	4,077,191	256,353	1,325,317	3,493,172	650,177	832,739	630,869
999,548	4,751,041	184,260	1,005,737	2,845,854	6,250,937	1,096,834	221,283
1,127,537	701,174	884,650	740,000	1,497,949	1,462,305	774,981	
189,739	5,932,923	1,479,604	2,542,533	5,429,133	803,606	3,860,932	
709,312	903,648	2,077,355	7,248,971	812,671	1,533,240	1,315,029	
1,453,984	1,740,529	323,034	2,599,160	1,181,001	923,340	1,261,680	
2,147,767	3,179,364	1,814,765	2,467,329	1,106,685	800,223	1,189,092	
312,234	255,733	507,006	698,023	1,334,014	493,956	2,194,494	
1,624,862	1,091,848	567,876	795,513	1,266,561	890,108	715,194	

## 다. 최빈값

전체 자료 중에서 가장 도수가 높은 값을 말한다. 사람들에게 어떤 지역에 거주하기 원하는 이유를 다음의 항목에서 한자리만 선택하라고 질문해서 얻은 응답의 분포가 <표 4-8>과 같다고 해보자.

〈표 4-8〉 응답자 분포

(단위 : 명)

선택항목	도수
교육시설이 좋다	300
의료시설이 좋다	200
복지시설이 좋다	100
자연환경이 좋다	50
편의시설이 좋다	50
교통이 좋다	50
기타	50

이 결과로부터 “교육시설”이라는 응답이 최빈값임을 알 수 있다. 지역단체장은 이를 중요한 정보로 활용할 수 있을 것이다.

## (2) 자료들이 얼마나 흩어져있나? (자료들이 얼마나 다양한가?)

자료의 중심 위치를 아는 것도 유용하지만 자료들이 얼마나 흩어져 있는지 모른다면 그릇된 해석을 할 수도 있다. 예를 들어, “도시가계조사에 의하면 어느 해 3/4분기 근로자 가구의 경상소득은 평균이 2,530,633원이다.”라는 통계청의 발표에는 상위 10%의 평균소득, 또는 200만원에서 300만원 사이의 가계소득을 갖는 근로자 가구의 비율 등이 들어 있지 않다. 보통 소득 분포는 오른쪽, 즉 고소득 방향으로 긴 꼬리를 갖고 넓게 퍼져 있기 때문에 평균소득만으로는 가계소득의 분포를 제대로 표현할 수 없다.

따라서 자료를 요약할 때에는 자료의 중심 위치를 나타내는 대표값 이외에 자료가 얼마나 다양한 값들을 갖는지 알아보는 방법으로 분산, 표준편차, 변동계수, 범위 등이 함께 사용된다.

### 가. 분산과 표준편차

크기가  $n$ 인 양적인 자료들이 평균으로부터 떨어진 정도를 나타내는 것으로 표준편차 (standard deviation)가 있다. 분산은 표준편차를 제곱한 값이다. 자료들이 평균으로부터 떨어져 양쪽으로 넓게 퍼져 있으면 분산과 표준편차가 커진다. 평균은 통상  $\bar{x}$ 라고 표기하고 표준편차는  $s$ 라고 표기한다.

#### 나. 변동계수

자료의 흩어진 정도를 때로는 다음과 같이 정의된 변동계수를 사용하여 표현하기도 한다. 즉 평균에 대한 표준편차의 비이다.

$$\text{변동계수(coefficient of variation)} = CV = \frac{\text{표준편차}}{\text{평균}} = \frac{s}{\bar{x}}$$

이렇게 정의된 변동계수는 자료의 측정단위에 의존하지 않으며 상대적인 산포의 측도로 여러 개의 자료를 비교할 때 활용될 수 있다.

#### 다. 범위

자료의 흩어진 정도를 재는 비교적 간단한 측도로서 자료에서 가장 큰 최대값과 가장 작은 최소값의 차이를 이용하여 아래와 같이 정의된 범위(range)를 사용하기도 한다.

$$\text{범위(range)} = \text{최대값} - \text{최소값}$$

이와 같은 정의에 의하여, 자료의 흩어진 정도가 심할 때는 자연히 범위도 커지며, 범위는 자료의 흩어진 정도를 나타내는 간단한 측도가 된다. 그러나 주어진 자료에 특별히 크거나 작은 값이 있는 경우 범위는 이들의 직접적인 영향을 받는다.

#### 라. 다섯숫자요약

최대값, 최소값, 중앙값, 상위 25%값, 하위 25%값을 이용해 자료를 표현하는 방법을 다섯 숫자요약(Five Number Summary)이라고 한다. 다음은 어느 지역의 4인 가구가 사용한 월 생활비 자료이다.

〈표 4-9〉 4인가구의 월생활비 자료

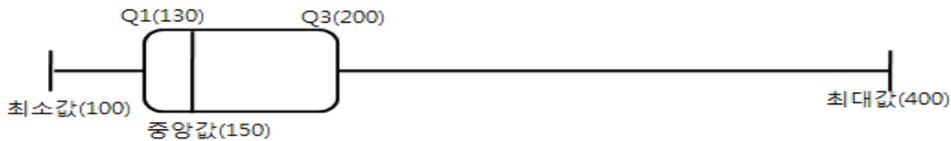
(단위 : 만원)

100	107	110	112	118	122	124	130	132	136
140	144	148	149	149	151	164	168	172	176
180	184	200	205	219	225	235	245	255	400

이 자료의 다섯숫자요약은 다음과 같다.

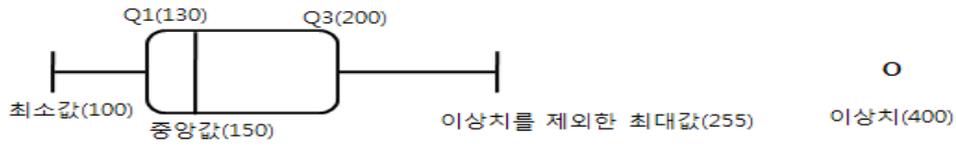
최소값 : 100,                      1사분위수(Q1) : 130,                      중앙값 : 150,  
 3사분위수(Q3) : 200,                      최대값 : 400.

이 다섯숫자요약으로부터 4인 가구의 생활비 분포를 추측할 수 있다. 먼저 하위 25%는 100만원에서 130만원 사이로서 폭이 30만원에 지나지 않는다. 하지만 상위 25%는 200만원에서 400만원 사이에 해당하므로 그 폭이 200만원에 달한다. 그러므로 전체 생활비의 분포는 오른쪽으로 긴 꼬리를 가질 것이다. 다음과 같은 상자그림을 이용하면 다섯숫자요약을 알기 쉽게 나타낼 수 있다.



〈그림 4-1〉 생활비의 상자그림(1)

그런데 자료를 좀 더 살펴보면 오른쪽 긴 꼬리는 최대값 400만원 때문에 생긴 것임을 알 수 있다. 만일 이 값을 제외한다면 꼬리는 훨씬 짧아질 것이다. 이처럼 나머지 값들에 비해 유달리 큰 값(또는 작은 값)을 이상치라고 부르는데 이상치를 두드러지게 표현해서 상자그림을 다시 그리면 다음과 같다.



〈그림 4-2〉 생활비의 상자그림(2)

#### 마. 이상치 판단하기

이제 다섯숫자요약을 이용하여 이상치를 판별하는 규칙에 대해 알아보자. 우선 식을 먼저 보도록 하자.

〈표 4-10〉 이상치 판별 규칙

이상치	$Q1 - (1.5 \times IQR)$ 보다 작은 값
	$Q3 + (1.5 \times IQR)$ 보다 큰 값
극단이상치	$Q1 - (3.0 \times IQR)$ 보다 작은 값
	$Q3 + (3.0 \times IQR)$ 보다 큰 값

‘ $Q3-Q1$ ’을 IQR(InterQuartile Range)라고 하며 사분위범위라고 부른다. 사분위범위는 상위 25%값에서 하위 25%값을 뺀 값으로서 그냥 범위가 최소, 최대 두 값만 이용하는 한계점을 극복하기 위해 만든 또 다른 측도이다. 이렇게 구한 사분위범위를 이용하여 Q1과 Q3로부터 사분위범위의 1.5배 밖에 있는 값을 ‘이상치’라고 부르고, 사분위범위의 3배 밖에 있는 것을 ‘극단이상치’라고 부른다. 생활비 자료에 이상치가 존재하는지 살펴보자.

100	107	110	112	118	122	124	130	132	136
140	144	148	149	149	151	164	168	172	176
180	184	200	205	219	225	235	245	255	400

사분위범위  $Q3 - Q1 = 200 - 130 = 70$

이상치  $130 - 1.5 \times 70 = 130 - 105 = 15$

$200 + 1.5 \times 70 = 200 + 105 = 305$

극단이상치  $130 - 3 \times 70 = 130 - 210 = -80$

$200 + 3 \times 70 = 200 + 210 = 410$

위의 이상치에 대한 기준에 비추어 보면 400은 이상치가 되고, 그 외의 다른 점은 모두 이상치의 범위에 속하지 않는다.

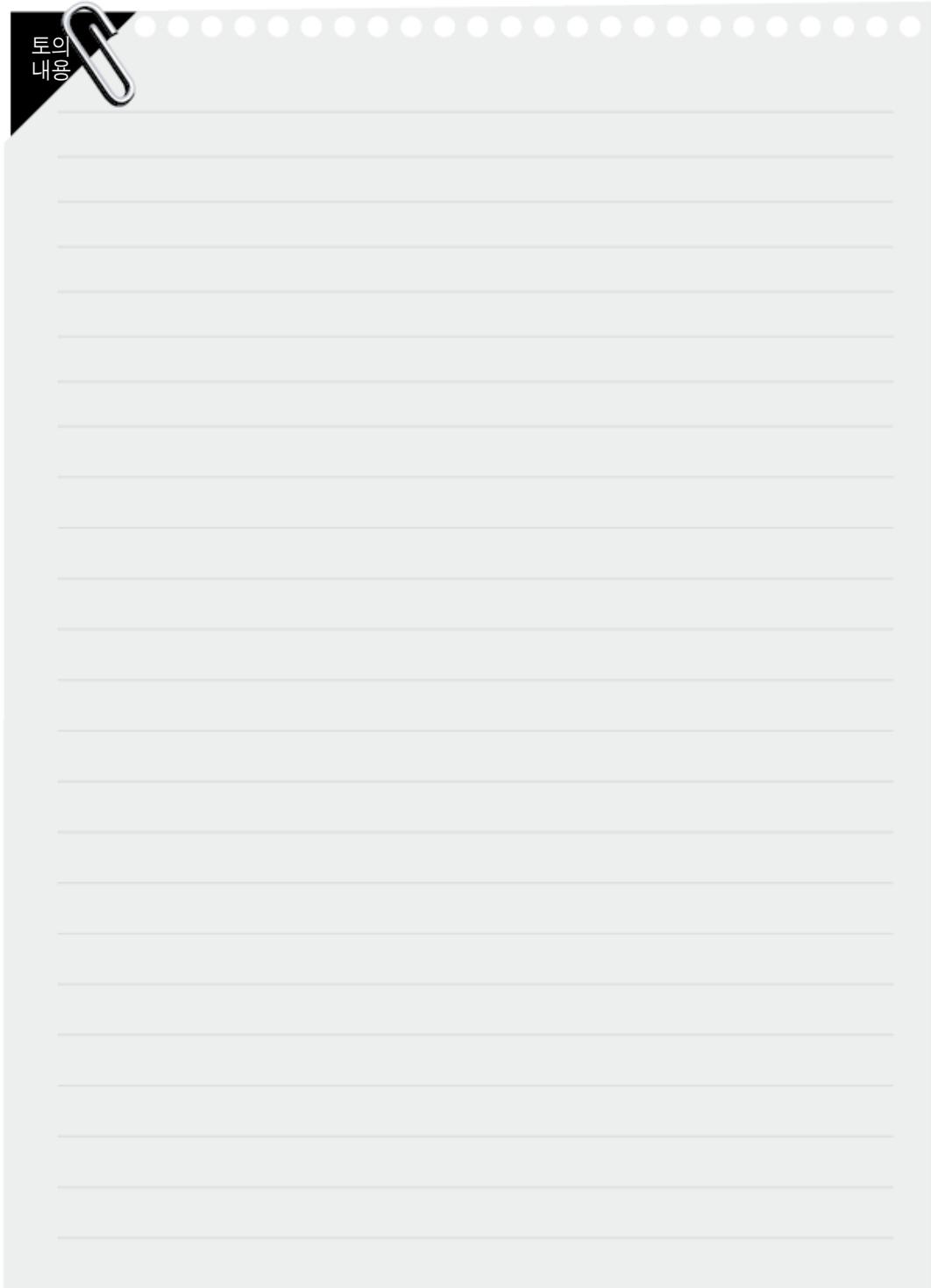
### 토의사례 1 우리나라 7개 대도시의 오존 오염도

다음은 어느 해 우리나라 7개 대도시의 월 오존 오염도 자료이다. 이들 자료로부터 지역별 평균, 표준편차, 변동계수, 범위를 계산하였다. 서울과 인천을 비교해보자. 표준편차, 범위를 보면 서울의 오존 오염도는 인천에 비하여 월별 차이가 작은 것으로 나타났다. 반면에 변동계수를 보면 서울이 0.3685로 인천 0.2895보다 크게 나타나서 더 많이 변동한 듯하다. 이유는 무엇일까? 서울의 평균 오염도가 인천에 비해서 낮는데 변동계수는 이를 고려하였기 때문이다.

〈표 4-11〉 주요 도시의 오존(O<sub>3</sub>)오염도

(단위 : ppm)

	서울	부산	대구	인천	광주	대전	울산
1월	0.009	0.019	0.015	0.016	0.016	0.015	0.017
2월	0.015	0.023	0.021	0.022	0.019	0.020	0.022
3월	0.021	0.029	0.026	0.028	0.026	0.030	0.029
4월	0.023	0.032	0.029	0.030	0.032	0.035	0.030
5월	0.025	0.033	0.034	0.032	0.037	0.036	0.032
6월	0.023	0.026	0.033	0.029	0.034	0.034	0.027
7월	0.021	0.021	0.024	0.022	0.027	0.022	0.021
8월	0.019	0.020	0.017	0.020	0.019	0.012	0.018
9월	0.014	0.020	0.016	0.018	0.020	0.014	0.017
10월	0.011	0.020	0.013	0.017	0.017	0.012	0.018
11월	0.009	0.014	0.011	0.015	0.011	0.010	0.015
12월	0.009	0.017	0.012	0.014	0.011	0.010	0.017
평균	0.017	0.023	0.021	0.022	0.022	0.021	0.022
표준편차	0.007	0.006	0.008	0.0063	0.009	0.010	0.006
변동계수	0.369	0.261	0.3901	0.290	0.388	0.493	0.273
범위	0.016	0.019	0.023	0.018	0.026	0.026	0.017



### (3) 표준화

어떤 관측치의 상대적 위치를 나타내는 척도로 자주 사용되는 것이 표준점수이다. 표준점수(standard score) 또는 표준화점수는 각 관측치에서 평균을 뺀 값을 표준편차로 나눈 것이다. 즉,

$$\text{표준점수} = (\text{관측치} - \text{평균}) / \text{표준편차}$$

이다. 이러한 표준점수들의 평균은 0이고 표준편차는 1이 된다.

위의 정의에서 알 수 있듯이 대단히 큰 표준점수는 해당 관측치가 다른 관측치들에 비해 더 크다는 것을 의미하며, 반면에 작은 표준점수는 그 관측치가 다른 관측치에 비해 작다는 것을 의미한다. 표준점수는 각 관측치의 위치가 평균을 중심으로 몇 표준편차 위 또는 아래에 있다는 것을 의미하므로 전체 자료 내에서의 위치를 나타내는 데 효율적으로 활용될 수 있다. 만약 표준점수가 0에 가깝다면 그 관측치는 자료의 평균에 가깝다고 할 수 있다. 그런데, 관측치의 단위는 표준편차의 단위와 같으므로 각각 분자, 분모에서 상쇄되기 때문에 표준점수는 단위가 없는 순수한 수치이다. 아래 표는 어느 특정한 해에 조사된 21개국의 경제 분야별(농업, 공업, 서비스업) 일자리 비율과 표준점수를 나타내고 있다.

〈표 4-12〉 21개국의 경제 분야별 일자리 비율과 표준점수

국가	전체 일자리 비율			표준점수		
	농업	공업	서비스업	농업	공업	서비스업
미국	8	38	54	-0.93	-0.46	2.30
캐나다	13	43	45	-0.66	-0.02	1.35
스웨덴	14	53	33	-0.60	0.86	0.08
스위스	11	56	33	-0.77	1.13	0.08
룩셈부르크	15	51	34	-0.54	0.69	0.19
영국	4	56	40	0.16	1.13	0.82
덴마크	18	45	37	-0.38	0.16	0.50
서독	15	60	25	-0.54	1.48	-0.76
프랑스	20	44	36	-0.24	0.07	0.40
벨기에	6	52	42	-1.05	0.78	1.03
노르웨이	20	49	32	-0.27	0.51	-0.03
아이슬란드	25	47	29	0.01	0.34	-0.34
네덜란드	11	49	40	-0.77	0.51	0.82
오스트리아	23	47	30	-0.10	0.34	-0.24
아일랜드	36	30	34	0.63	-1.16	0.19
이탈리아	27	46	28	0.12	0.25	-0.45
일본	33	35	32	0.46	-0.72	-0.03
그리스	56	24	20	1.74	-1.69	-1.29
스페인	42	37	21	0.96	-0.55	-1.19
포르투갈	44	33	23	1.07	-0.90	-0.98
터키	79	12	9	3.02	-2.75	-2.45
평균	24.8	43.2	32.2	0.0	0.0	0.0
표준편차	18.0	11.4	9.47	1.0	1.0	1.0

터키의 경우 농업부분의 일자리 비율은 21개 국가의 평균을 중심으로 표준편차의 3.02배 위에 있는 반면 공업 및 서비스업의 일자리 비율은 각각의 평균을 중심으로 표준편차의 2.75, 2.45배 아래에 있다는 것을 알 수 있다. 또 아이슬란드의 경우는 농업, 공업, 서비스업의 표준점수가 각각 0.01, 0.34, -0.34이므로 일자리 비율이 평균에 매우 가까이 있음을 알 수 있다. 그리고 미국의 경우는 공업과 농업이 차지하는 일자리 비율은 매우 낮은 반면 서비스업이 차지하는 일자리 비율은 매우 높은 편임을 알 수 있다.

또 다른 예를 다음 표를 통하여 보자. 아래 표는 ABC 국가의 TOBANKRUPT 은행 영업계수와 정상적으로 운영되는 50개 은행의 연도별 영업계수(operating ratio(%))의 평균과 표준편차이다(영업계수 산출식은 영업비용 ÷ 영업수익 × 100로서 은행의 영업 능력의 정도를 나타낸다. 이 계수가 100이상이면 채산이 맞지 않고 있다는 것이 된다).

(표 4-13) TOBANKRUPT 은행의 영업계수와 정상운영 50개 은행 영업계수

연도	TOBANKRUPT 은행의 영업계수	정상 은행의 영업계수 평균	정상 은행의 영업계수 표준편차	TOBANKRUPT 은행의 표준점수
2003	84.06	78.27	5.40	1.07
2004	86.14	79.54	5.30	1.25
2005	91.19	81.04	5.80	1.75
2006	93.77	82.08	6.10	1.92
2007	95.60	84.90	5.16	2.07

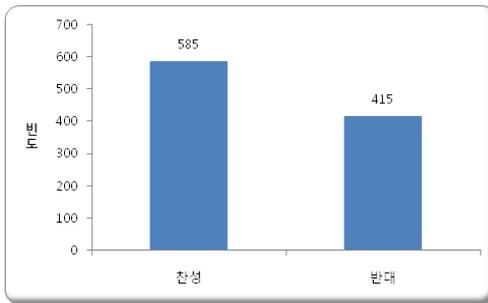
TOBANKRUPT 은행의 경우 영업계수는 2003년부터 계속 증가하고 있어서 100에 가까워지고 있으며 표준점수도 점점 높아지는 것을 볼 수 있다. 그런데 영업계수가 TOBANKRUPT 은행의 실태정보만 나타내는 반면 표준점수는 다음과 같이 상대적 실태정보를 제공한다. 실제로 50개 은행의 연도별 영업계수가 정규분포를 따른다면, 2005년 TOBANKRUPT 은행의 표준점수는 1.75인데, 이 값은 표준정규분포의 극단 상위 5%에 해당하는 1.645보다 큰 값이다. 따라서 이 은행은 영업계수 면에서 정상 운영되는 은행들 중 상위 5% 이내에 포함된다.

#### (4) 표와 그래프

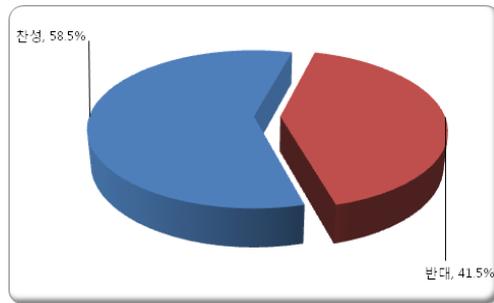
예를 통해서 시작해 보자. 다음 가상자료는 국내 성인 중에서 임의로 1,000명을 뽑아 특정 정책에 대한 지지여부를 조사한 결과이다. 1,000명 중 찬성이 585명, 반대가 415명으로 나타났다 하자. 이 결과를 효과적으로 전달하는 방법으로 <표 4-14>와 <그림 4-3>, <그림 4-4>와 같은 방법들이 있을 것이다.

〈표 4-14〉 정책지지여부의 도수분포표

정책지지여부	도수(명)	백분율(%)
찬성	585	58.5
반대	415	41.5
합 계	1,000	100.0



〈그림 4-3〉 정책지지여부의 막대그래프



〈그림 4-4〉 정책지지여부의 원그래프

그렇다면 다음과 같은 예는 어떻게 표현해야 하는가?

30대 미혼인 사람들이 한 달 생활비로 얼마나 지출하는지 알아보기 위해 20명을 임의로 뽑아 조사한 결과가 <표 4-15>라고 하자.

〈표 4-15〉 30대 미혼인 사람들의 한 달 생활비

(단위 : 만원)

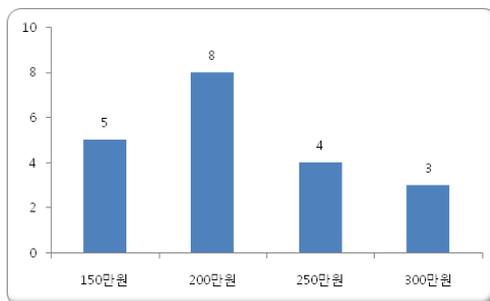
150	200	250	250	150	200	150	200	150	300
200	300	200	200	150	200	300	250	200	250

이 자료는 양적 자료로서, 한 달 생활비는 연속형 변수이다. 그렇지만 서로 다른 종류의 숫자 값이 4가지(150, 200, 250, 300)뿐이므로 이것도 정책지지여부와 마찬가지로 아래의 <표 4-16>의 도수분포표를 작성한다.

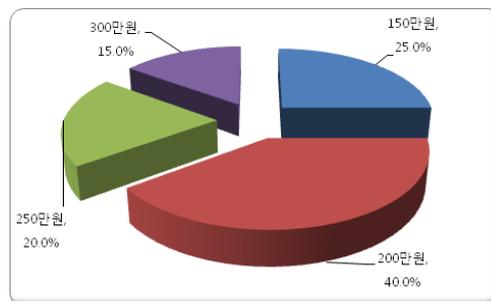
〈표 4-16〉 30대 미혼인 사람들의 한 달 생활비의 도수분포표

한 달 생활비	도수	누적도수	상대도수
150	5	5	0.25
200	8	13	0.40
250	4	17	0.20
300	3	20	0.15
합 계	20		1.00

〈표 4-16〉을 살펴보면 두 번째 행의 값 200에서의 도수 8은 조사대상자 20명 중 200만원을 한 달 생활비로 쓴다고 대답한 응답자의 수이고, 누적도수 13은 이 자료에서 200만원 이하의 한 달 생활비를 쓰는 응답자 수를 의미이다. 200의 상대도수 0.4는 이 자료에서 한 달 생활비로 200만원을 쓰는 응답자의 비율을 나타낸다. 이 도수분포표를 이용하여  $X$ 축을 관찰값,  $Y$ 축을 도수로 하여 막대그래프를 그리면 〈그림 4-5〉와 같이 된다. 물론 〈그림 4-6〉의 원그래프, 또는 띠그래프 등의 다양한 그래프도 가능할 것이다.



〈그림 4-5〉 한 달 생활비의 막대그래프



〈그림 4-6〉 한 달 생활비의 원그래프

〈표 4-17〉은 가상자료로서 대전광역시에 거주하는 4인가구의 가구주 중에서 임의로 100명을 뽑아 대전시에 대한 생활만족도를 조사한 결과라고 하자. 생활만족도는 0~100점 사이의 값을 가진다.

〈표 4-17〉 4인가구 가구주의 생활만족도

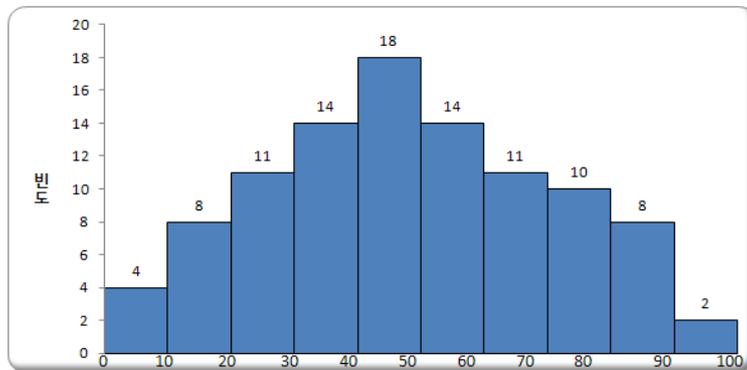
35	32	31	30	25	41	7	17	18	20
8	9	13	14	62	63	63	65	67	66
69	75	15	15	22	23	24	33	34	35
35	36	37	26	27	39	39	38	29	29
27	45	46	47	45	81	80	79	14	17
82	8	43	44	49	48	47	49	44	42
49	50	50	51	52	53	52	49	48	49
54	53	56	56	59	59	60	61	76	77
78	79	80	22	83	84	35	47	85	83
66	77	75	76	77	60	56	52	92	93

<표 4-17>의 자료도 정책지지여부와 한 달 생활비와 같이 각 값이 몇 개인지 빈도와 백분율을 구하여 빈도표를 만들고, 빈도표를 바탕으로 막대그래프, 원그래프로 표현하면 될까? 생활만족도 자료는 정책에 대한 지지자 수 자료나 한 달 생활비 자료와 비교할 때 자료의 구성이 다르다. 정책지지여부는 찬성과 반대로 두 가지의 값만 가지고, 한 달 생활비는 150만원, 200만원, 250만원, 300만원의 네 가지의 값만 가지고 있다. 하지만 생활만족도는 7점부터 93점 사이의 다양한 값을 가지며, 그 값들이 모두 63가지나 된다. 이러한 자료를 정책지지여부와 한 달 생활비처럼 표현하면 전체적인 특징을 발견하기가 어려워진다. 이 자료에 대해서는 서로 다른 숫자 값이 상당히 많으므로 구간을 나누는 것이 좋을 것 같다. 먼저 적당한 계급 구간을 정해야 하는데 구간을 몇 개로 하는 것이 좋을까? 여기서는 최소값이 7점이고 최대값이 93점이므로 구간을 10점대로 구분하자. 그러면 결과는 <표 4-18>과 같다.

〈표 4-18〉 생활만족도의 도수분포표

생활만족도	도수	누적도수	상대도수
0~9	4	4	0.04
10~19	8	12	0.08
20~29	11	23	0.11
30~39	14	37	0.14
40~49	18	55	0.18
50~59	14	69	0.14
60~69	11	80	0.11
70~79	10	90	0.10
80~89	8	98	0.08
90~100	2	100	0.02
합 계	100		1.00

그리고 이 자료에 맞는 그래프는 막대그래프나 원그래프가 아니라 아래의 그래프이다.



〈그림 4-7〉 생활만족도에 대한 히스토그램

누적 도수	도수	줄기	잎
4	4	0	7 8 8 9
12	8	1	3 4 4 5 5 7 7 8
23	11	2	0 2 2 3 4 5 6 7 7 9 9
37	14	3	0 1 2 3 4 5 5 5 5 6 7 8 9 9
(18)	18	4	1 2 3 4 4 5 5 6 7 7 7 8 8 9 9 9 9 9
45	14	5	0 0 1 2 2 2 3 3 4 6 6 6 9 9
31	11	6	0 0 1 2 3 3 5 6 6 7 9
20	10	7	5 5 6 6 7 7 7 8 9 9
10	8	8	0 0 1 2 3 3 4 5
2	2	9	2 3

줄기 단위 : 10

〈그림 4-8〉 가상자료의 줄기와 잎 그림

### (5) 그래프의 오용

그래프는 그 시각적 효과 덕분에 어떤 복잡한 수식을 동원한 분석보다 자료를 훌륭하게 표현해주는 경우가 많다. 그런가 하면, 보도기사 등에서는 주장을 강조하거나 심지어 그릇된 주장을 하기 위해 그래프를 이용하기도 한다. 통계그래프에는 막대그래프, 원그래프, 꺾은선 그래프, 산점도 등의 다양한 그래프가 존재하며, 그 중에서도 막대그래프, 꺾은선그래프, 원그래프를 사용할 때 잘못 사용하는 경우를 많이 볼 수 있다. 여기서는 막대그래프와 원그래프에서 어떻게 오용이 일어나는지 간략히 살펴보기로 한다.

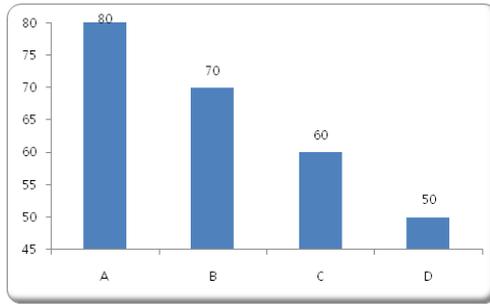
#### ① 막대그래프

다음은 4개 공공기관에 대한 이용자의 만족도 자료이다.

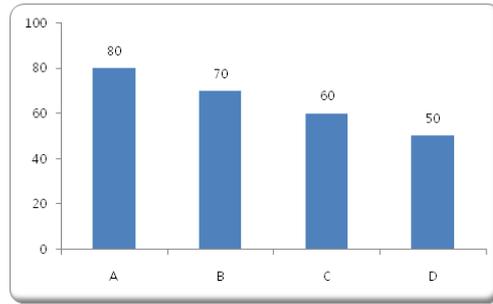
〈표 4-19〉 기관 만족도

기관	만족도
A	80
B	70
C	60
D	50

<표 4-19>의 자료를 막대그래프로 표현하는 과정에서 <그림 4-9>처럼 막대의 일부를 절단(특히 아래 부분을)하면 실제 차이보다 훨씬 더 많은 차이가 존재하는 것처럼 만들 수 있다. 바람직한 것은 <그림 4-10>처럼 절단되지 않도록 표현하는 것이다.

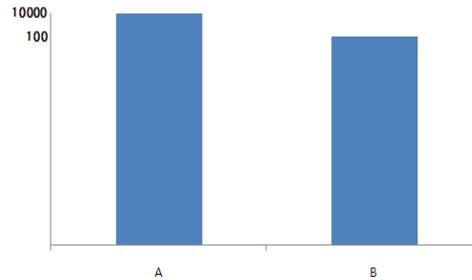
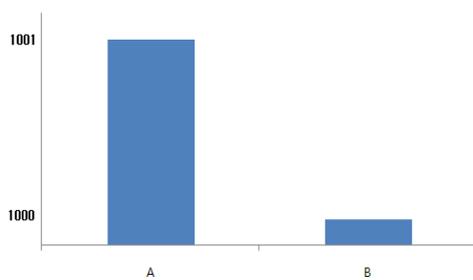


〈그림 4-9〉 절단된 막대그래프

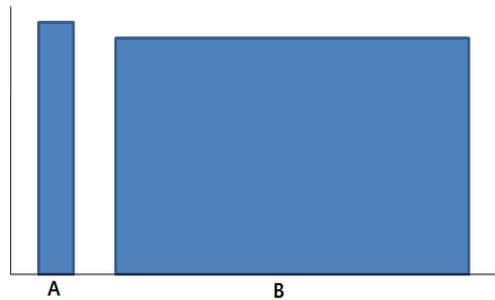


〈그림 4-10〉 절단되지 않은 막대그래프

그리고 <그림 4-11>과 <그림 4-12> 역시  $y$ 축을 조정해서 생기는 그래프 오용의 유형이다. <그림 4-11>은 A와 B 사이의 차이가 1임에도 불구하고 그래프에서는 A가 B보다 몇 배나 큰 것처럼 표현 표현되었다. <그림 4-12>는 <그림 4-11>과 반대로 A와 B 사이의 차이가 큰데도 불구하고 그래프에서는 A와 B 사이의 차이가 별로 없는 것처럼 표현되었다.

〈그림 4-11〉  $y$ 축 눈금을 조정한 막대그래프(1)    〈그림 4-12〉  $y$ 축 눈금을 조정한 막대그래프(2)

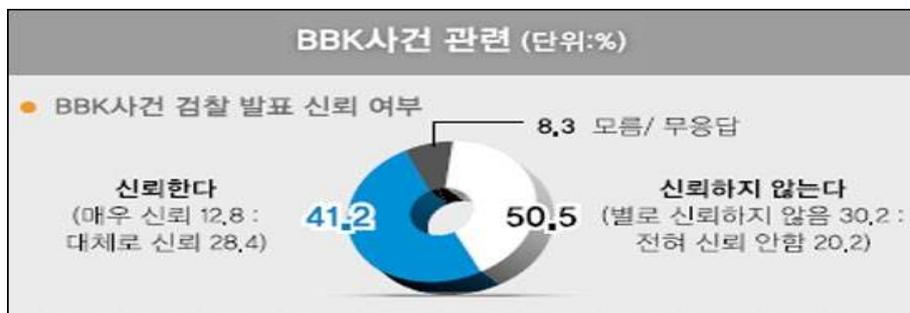
막대그래프를 오용하는 또 다른 유형으로 막대의 너비를 다르게 만들어 잘못된 정보를 전달하는 경우도 있다. 막대그래프에서 정보는 막대의 높이로 전달되는데 막대 A의 너비보다 막대 B의 너비를 몇 배 크게 만들면 B가 A보다 훨씬 더 많은 것처럼 보인다.



〈그림 4-13〉 너비를 조정한 막대그래프

## ② 원그래프

원그래프는 각 항목이 전체에서 얼마나 차지하고 있는지를 알려주는 그래프이다. 하지만 원그래프에서 원 조각이 과장되게 표현되면 정보를 잘못 전달할 수 있다. <그림 4-14>는 BBK사건에 대한 검찰발표의 신뢰여부를 묻는 여론조사 결과이다. 그런데 결과를 원그래프로 표현하는 과정에서 '신뢰한다'에 해당하는 41.2%의 원조각이 '신뢰하지 않는다'에 해당하는 50.5%보다 더 크게 표현되었다. 결과적으로 이 원그래프는 조사결과와는 다르게 약 50%가 검찰발표를 신뢰하는 것처럼 전달하고 있다.

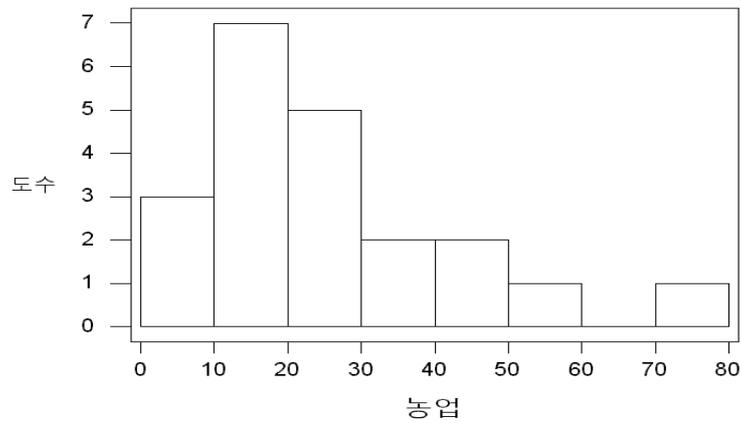


출처 : [http://cafe.naver.com/bloodbird.cafe?iframe\\_url=/ArticleRead.nhn%3Farticleid=11716](http://cafe.naver.com/bloodbird.cafe?iframe_url=/ArticleRead.nhn%3Farticleid=11716)

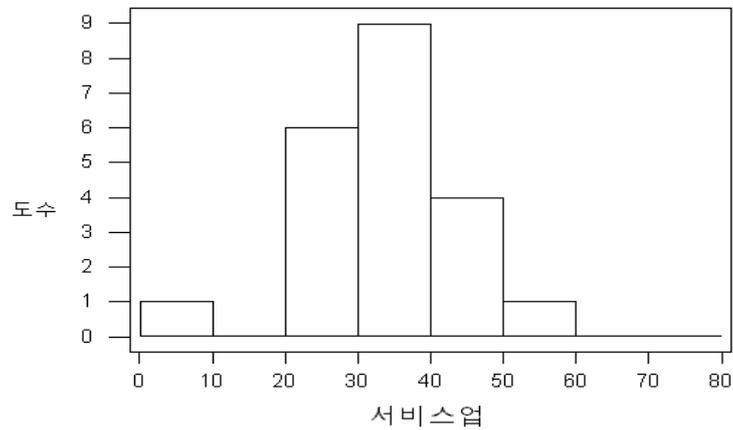
〈그림 4-14〉 원그래프의 오류

## 토의문제 1

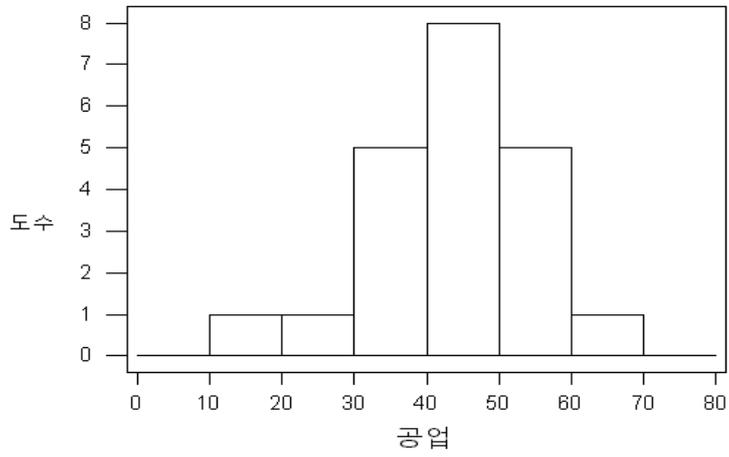
다음은 앞 절에서 다룬 21개국의 업종별 종사인구비율을 히스토그램과 상자그림으로 나타낸 것이다. 그래프를 보고 특성을 토의해 보자.



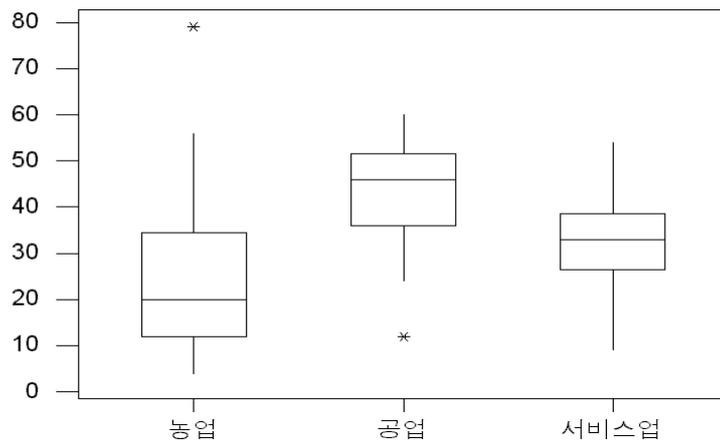
〈그림 4-15〉 농업 종사인구비율의 히스토그램



〈그림 4-16〉 서비스업 종사인구비율의 히스토그램

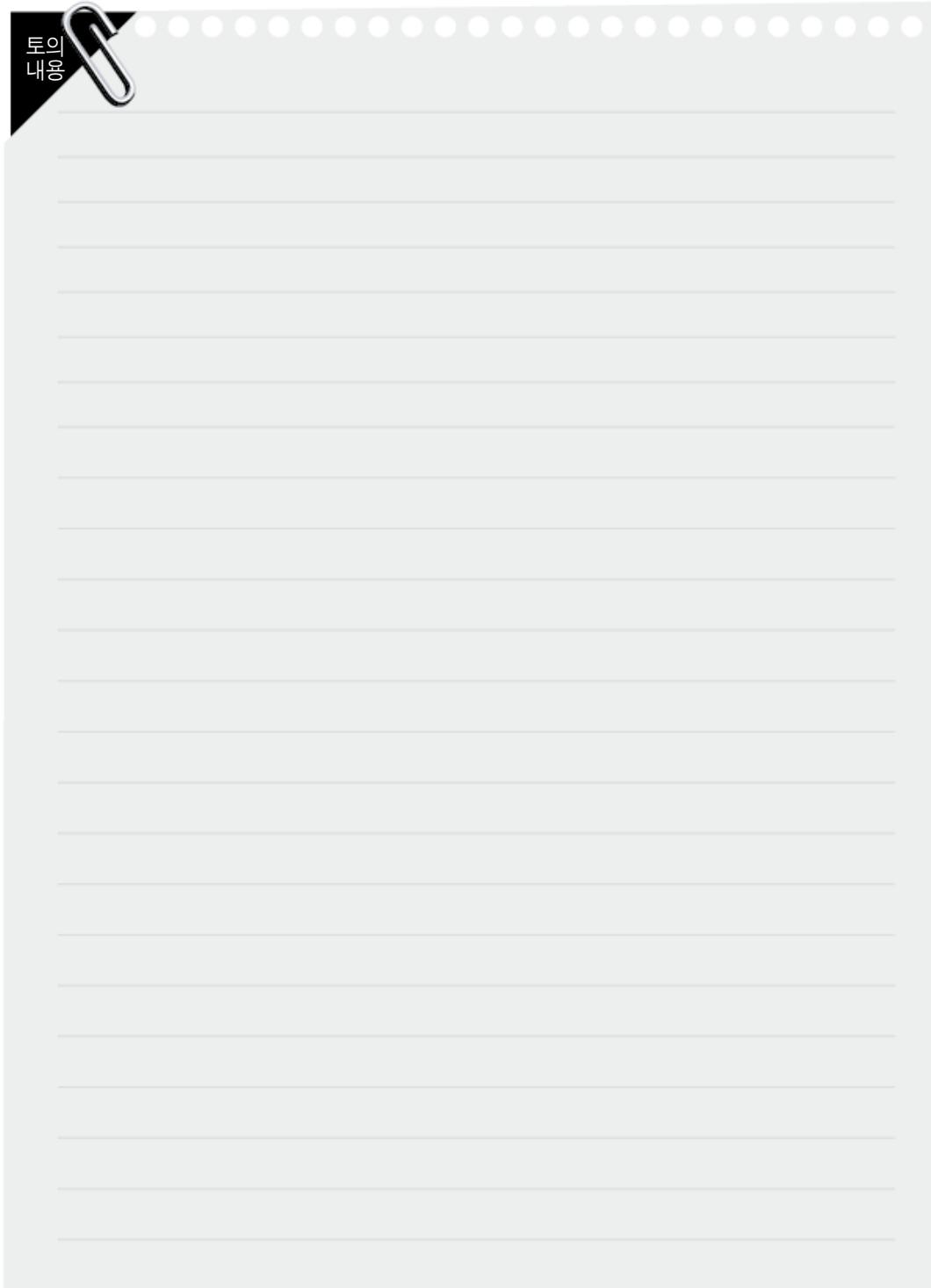


〈그림 4-17〉 공업 종사인구비율의 히스토그램



〈그림 4-18〉 업종별 종사인구비율의 상자그림

투  
의  
내  
용



This area is a large, light gray rectangular field with a perforated top edge, resembling a sheet of paper. A paperclip is attached to the top left corner. The text '투의내용' is written vertically in the top left corner. The rest of the area is blank, intended for handwritten notes.



## 4.4 자주 잘못 사용되는 통계 용어들

### 4.4.1 지수의 정의와 3요소

지수란 물가와 같이 시간의 흐름에 따라 통계치가 어떻게 변화되었는지를 계량적으로 파악할 수 있게 만든 수치로서 통상 비교 기준이 되는 시점의 수치를 100으로 하여 산출한다. 예를 들어 유제품의 물가 지수가 105라는 의미는 기준 시점에 비해 유제품의 가격이 5% 상승하였음을 나타낸다.

한국은행에서 펴낸 “알기 쉬운 경제 지표 해설”에서는 지수를 계산하기 위해 기준 시점, 가중치, 그리고 지수식이 필요하다고 하였는데, 이를 지수의 3요소라고 한다.

#### (1) 요소1 : 기준 시점

지수의 기준 시점은 코스피지수(KOSPI)처럼(1980년) 단일 시점을 선택하는 경우도 있지만, 산업생산지수, 물가지수 등 대부분의 통계에서는 1개월 또는 1년간의 기간을 선택하여 그 기간 중의 평균 수치를 100으로 하여 기준을 정한다.

#### (2) 요소2 : 가중치

종합 지수를 계산할 때는 개별구성 항목의 지수를 산술 평균하지 않고 각 항목의 중요도를 고려하여 가중 평균을 취하는데 이때 고려되는 중요도가 가중치이다. 예컨대 물가지수의 경우에는 상품 거래 금액 또는 소비 지출 금액을, 산업 생산 지수의 경우에는 부가가치액을 가중치로 하고 있다.

#### (3) 요소3 : 지수식

지수를 산출하기 위한 지수식으로는 라스파이레스(Laspeyres), 파셰식(Paasche) 및 피셔식(Fisher)등이 있는데, 우리나라를 비롯한 대부분의 나라에서는 물가 지수를 산출할 때 계산상의 편의를 위해 본래의 식과 수리적으로 같은 값을 갖도록 변형시킨 라스파이레스 수정식을 사용한다. 따라서 본 교재에서는 라스파이레스식만을 간략하게 소개한다.

라스파이레스식은 1864년 독일의 통계학자인 라스파이레스가 창안한 식으로 기준 시점 ( $T=0$ )과 같은 수량( $Q_0$ )의 상품을 비교시점( $T=t$ )의 가격( $P_t$ )으로 구입했을 때 그 구입금액을 기준 시점의 구입 금액으로 나누어 물가의 움직임을 측정하는 방식이다. 이때 기준 시점의 구입 수량( $Q_0$ )이 가중치로 고정되는데 이는 상품 구입 수량이 크게 변화하지 않는다는 것을 전제로 한 것이다. 보다 자세한 내용을 알고 싶은 독자들은 한국은행이 펴낸 “알기 쉬운 경제 지표 해설”을 참고하기 바란다.

#### 4.4.2 변동률(증감률)

시계열 지표(시간에 따라 생산되는 지표)를 비교할 때 사용되는 개념으로 기준 시점에 따라 전월비, 전년 동월비, 전년 대비 증감률 등이 있다. 그리고 포인트 변동률은 두 시점에서의 변동률 차이를 나타낼 때 쓰인다. 이들 용어는 이미 우리 주위에서 많이 만나는 것들이므로 다음의 토의사례들을 통하여 정리해 보자.

토의사례 1    특정 광물의 생산량

다음의 자료로부터 위의 개념을 토의해 보자.

〈표 4-20〉 특정 광물의 생산량 자료

년도	분기	생산량
2006년	1사분기	2,173
	2사분기	920
	3사분기	328
	4사분기	2,734
2007년	1사분기	1,953
	2사분기	1,009
	3사분기	720
	4사분기	2,420
2008년	1사분기	1,640
	2사분기	562
	3사분기	1,032
	4사분기	1,747

(1) 2008년 4분기 생산량의 전분기 대비, 전년 동기 대비 변동률을 구해보자.

(2) 2008년도 생산량의 전년 대비 변동률을 구해보자.

토의사례 2    실업률

다음은 2006년 6월부터 2008년 6월까지 한국은행에서 발표한 실업률이다.

〈표 4-21〉 우리나라의 실업률

(단위: %)

년도	1월	2월	3월	4월	5월	6월	7월	8월	9월	10월	11월	12월
2006	3.7	4.1	3.9	3.5	3.2	3.4	3.4	3.4	3.2	3.3	3.2	3.3
2007	3.6	3.7	3.5	3.4	3.2	3.2	3.2	3.1	3.0	3.0	3.0	3.1
2008	3.3	3.5	3.4	3.2	3.0	3.1						

출처 : 한국은행

- (1) 2008년 6월의 실업률은 전년 동월 대비 얼마나 변동하였는가?
- (2) 2008년 6월의 실업률은 전월 대비 얼마나 변동하였는가?
- (3) 2008년 6월의 실업률은 지난달(5월) 실업률에 비해서 얼마나 높아졌는가?

### 4.4.3 진척률

진척률은 목표 달성도라고도 하는데 이 역시 일반적으로

$$\left( \frac{\text{실적치}}{\text{목표치}} \right) \times 100$$

와 같이 비율로 나타내기도 하지만, 때로는 관점에 따라 여러 가지 변형된 형태로 나타내기도 한다. 따라서 진척률에 관한 자료를 볼 때에는 어떤 관점에서 사용하였는지 유의하여야 한다.

#### 토의문제 1

대표적인 예는 건설 현장에서 나타나는 것으로 금년 7월초에 동시에 착공하여 금년 12월 말에 완공하는 공사 현장과, 금년 말에 공정의 50%를 마치려고 하는 공사 현장이 있다고 하자. 다음은 지난 3개월의 실적이다.

〈표 4-22〉 진척률

	7월 말	8월 말	9월 말	12월 말
공사현장 A	30%	40%	50%	100%
공사현장 B	20%	25%	30%	50%

(1)“시점을 7월말로 했을 때 어느 공사현장이 더 많이 진척했는가?”라는 질문에 답하려 할 때 우리는 다음과 같은 두 개의 관점을 제시할 수 있다.

- ① 관점1 : 7월 현재 두 현장의 실적만을 비교한다면 공사현장 A가 30%로서 20% 진척에 그친 공사현장 B에 비해 더 진척했다고 볼 수 있다.
- ② 관점2 : 그러나 12월 목표치와 함께 본다면 공사현장 A는 12월 목표치 100%에 대

하여 30% 달성한 것이고, 공사현장 B는 12월 목표치 대비 50% 달성하였으므로  $\frac{20}{50} \times 100 = 40\%$ 가 되어 공사현장 B가 더 많이 진척되었다고 할 수 있다.

(2)“8월, 9월 상황을 비교한다면 어느 공사현장이 얼마나 더 많이 진척했다고 하겠는가?”라는 질문에 답해보자.

- ① 관점1 : 공사현장 A는 8, 9월 두 달간  $50\% - 30\% = 20\%$ , 공사현장 B는  $30\% - 20\% = 10\%$ 로서 공사현장 A가  $20\% - 10\% = 10\%$  더 많이 진척시켰다고 할 수 있겠다.
- ② 관점2 : 반면에 앞서서와 같이 12월 목표치를 고려하면 공사현장 A가 20%를 진척하고, 현장 B는 8, 9월에 10%를 진척함으로써 12월 목표치 50%의 20% ( $\frac{10}{50} \times 100$ )를 진척한 것이 된다. 따라서 두 공사현장의 진척 상황은 같다고 할 수 있다.
- ③ 관점3 : 공사현장 A는 7월말 시점에서 공정의 70%가 남았는데, 8, 9월에 20%를 진척시켰으므로,  $\frac{20\%}{70\%} = \frac{2}{7}$ 를 진척했고, 공사현장 B는 7월말 시점에서 공정의 30%가 남았는데, 8, 9월에 10% 진척했으므로  $\frac{10\%}{30\%} = \frac{1}{3}$ 을 진척했다. 따라서 이런 관점에서 본다면 ( $\frac{1}{3} > \frac{2}{7}$ )이므로 공사현장 B가 공사현장 A보다 더 많이 진척했다고 말할 수 있겠다.

#### 4.4.4 비, 비율, 율

우리가 자주 접하는 지표관련 자료에서 통계청에서 정의하고 있는 조이혼율은 1년간 신고된 총 이혼건수를 그 해의 연앙인구(7월 1일 기준 인구)로 나눈 수치를 1,000분비로 나타낸

것이고, 출생성비는 여아 출생자수 100명당 남아 출생자수를 나타내고, 30세 이상에서 35세 이하의 미혼자 중 남자비율은 30세 이상이고 35세 이하인 총 미혼자 중에서 남자 미혼자가 차지하는 정도를 나타낸다.

여기서 비, 비율, 율의 개념이 나오고 있는데, 이들 용어를 별다른 생각 없이 쓰다보면 그 개념들이 갖고 있는 단순한 의미에 비하여 심각한 실수를 범할 수도 있다. 따라서 비, 비율, 율에 대해서 좀 더 살펴보자.

▷ 비(ratio) : 두 개 숫자의 상대적 크기를 나타내는 것으로 A에 대한 B의 비는  $B \div A$ 로 계산된다.

(예) 성비 = (남자 $\div$ 여자) $\times$ 100

▷ 비율(proportion) : 비의 특별한 형태로 분모에 분자가 포함된다.

(예) 미혼자 중 남자비율 = 미혼남자  $\div$  (미혼남자+미혼여자)

▷ 율(rate) : 특정 기간 동안에 발생된 사건을 그 사건의 위험에 노출된 총횟수(건수, 사람수)로 나눈 것으로 보통 천분율, 즉 천 회당 건( $\%_{00}$ )으로 표현한다.

(예) 조이혼율( $\%_{00}$ )=특정연도 인혼건수 $\div$ 당해연도 연앙인구 $\times$ 1,000

2006년 이동통신보유율( $\%_{000}$ )=이동통신보유자 $\div$ 총인구 $\times$ 10,000

### 사례 1 이혼율에 대한 사회적 혼란을 잠재우려면?

다음은 잘못 사용된 이혼통계 사례로서 통계청 웹사이트에서 인용한 것이다.

#### (1) 이혼비를 이혼율로 잘못 사용한 예

특정연도에 혼인한 부부의 수를 분모로, 그 해에 이혼한 부부의 수를 분자로 하여 산출한 값(이혼비)을 이혼율로 사용한 경우

- ① "2002년에 혼인한 부부의 혼인건수(306,600건)와 이혼한 부부의 이혼건수(145,300건)를 비교해 보면, 하루 평균 840쌍이 결혼하고 398쌍이 이혼해 결혼대비 이혼율이 47.4%이다."

※ 이혼율은 그해 이혼건수를 그해의 연앙인구로 나누어 천분율( $\text{‰}$ )로 나타냄

▷ 위의 사례는 2002년에 일어난 이혼건수(145,300건)를 분자로 그해 일어난 혼인건수(306,600건)를 분모로 하여 백분율로 나타낸 경우임

▷ 따라서 이혼율이라고 할 수 없으며 47.4%는 혼인(건)에 대한 이혼(건)을 단순 비교한 이혼비에 해당함

- ② "2002년 한 해 동안 하루 평균 199쌍이 결혼하고 80쌍이 이혼한 것으로 나타났고, 이는 결혼 2쌍 가운데 1쌍에 가까운 40.2%가 이혼한 것이다."

☞ 사례 1과 같이 이혼 80쌍을 분자로 혼인 199쌍을 분모로 하여 백분율로 나타낸 것으로 40.2%는 혼인에 대한 이혼비에 해당하며 이혼율로 부를 수 없음

- ③ "2004년 1월말 현재시점에 혼인경력자의 총혼인횟수가 28,156,405건이고, 총이혼횟수는 2,623,659건이므로 이혼율은 9.3%가 된다."

☞ 이는 전산화된 호적정보를 이용해 계산한 것으로 1월말 시점의 현재인구 중에서 혼인경력자의 총혼인횟수를 분모로, 이혼경력자의 총이혼횟수를 분자로 하여 백분율(9.3%)로 나타낸 것임.

▷ 특정 기간(통상 1년)동안을 대상으로 한 인구학적 이혼율 산정방식에 부합하지 않아 이혼율이라고 할 수 없으며, 총누적 혼인(건)에 대한 총누적 이혼(건)의 비, 즉 이혼비에 해당함

- ▷ 인구통계통계에서 이혼율은 특정 기간(통상 1년) 동안에 발생한 이혼건수를 그해 연앙인구로 나누어 천분율로 나타냄

## (2) 이혼율 산정방식을 잘못 사용한 예

특정 연도에 이혼한 부부의 이혼건수를 분자로, 특정 연도 말의 기준으로 한 혼인부부의 수(연말기준 유배우인구)를 분모로 하여 산출한 값을 유배우이혼율이라고 제시한 경우

- ① "2002년 말 현재 혼인한 부부(유배우인구)가 11,011,902쌍이고, 2002년 중에 이혼한 부부(이혼건수)는 145,300쌍(건)이므로 2002년 우리나라의 유배우이혼율은 1.3%이다"
  - ☞ 유배우이혼율 산정에서 유배우인구(혼인부부의 수)는 연도 말이 아닌 연앙(7월 1일 기준)인구를 분모로 사용함. 또한 이혼율은 통상 백분율이 아닌 천분율로 나타냄
  - ▷ 2002년의 유배우 연앙인구가 11,011,902쌍(22,023,804명)이라고 가정한다면 유배우이혼율(유배우자 1천 명당 이혼건수)은 6.6임

## 4.4.5 퍼센트와 퍼센트 포인트

시간의 흐름에 따른 숫자의 변화가 관심의 대상이 되는 경우가 많다. 그런데 그 관심의 대상이 되는 숫자가 퍼센트(%)로 표시된 것일 때에는 이 퍼센트의 변화를 퍼센트 포인트로 표현한다. 실업률이나 시장점유율, 이자율 등의 변화가 여기에 해당된다. 퍼센트를 직접 비교할 때, 만약 기준이 같다면 퍼센트를 보통의 숫자와 마찬가지로 서로 더하거나 뺄 수 있다. 이때 두 퍼센트의 차이(혹은 변화)를 퍼센트 포인트라고 한다. 이처럼 퍼센트와 퍼센트 포인트는 간단한 개념인데도 불구하고 많은 사람들이 혼동하고 있으며, 신문기사에서도 %포인트를 그냥 %로 잘못 사용하는 경우를 흔하게 볼 수 있다.

대표적인 예로 일본 중앙은행이 엔고의 행진을 막기 위해서 재할인율을 1.75%에서 1.0%로 인하한다고 발표하자 우리나라의 한 일간지는 “재할인율 0.75% 인하 - 일, 엔고 긴급대책 발표”라는 제목으로 1면 머리기사를 실었다. 0.75%포인트를 그냥 0.75%라고 혼동하여 제목을 뽑은 것이다. 이렇게 되면 독자들은  $1.75\% \times \frac{0.75}{100} = 0.013125\%$ 로서 1.736875%로 인하한 것으로 받아들여지게 된다.

다른 예로 실업률 계산의 기준이 되는 노동가능인구는 단기적으로 거의 비슷하므로 실업률

의 변화는 그 차이를 직접 계산해서 퍼센트 포인트로 표현한다. 가상적으로 지난 5년 동안에 실업률이 2.1%에서 3.2%로 증가했다고 하자. 이 퍼센트의 변화는 다음과 같이 두 가지 방법으로 표시할 수 있다. “실업률이 지난 5년간 1.1%포인트 증가했다” 혹은 “실업률이 지난 5년간 52% 증가했다” 이 두 가지 표현은 모두 정확한 표현이다. 그러나 우리에게 주는 인상은 물론 크게 다르다. 둘 중 어느 표현을 선택하는가는 말하는 사람이 듣는 사람에게 어떤 인상을 심어주기를 원하느냐에 달려 있다. 첫 번째 방법은 실업률이 약간만 올랐다는 인상을 주는 반면 두 번째 방법은 기준이 같을 경우 퍼센트의 퍼센트를 구할 수 있으므로 실업률이 52% 증가했다고 표현한 것이다. 3.2%에서 2.1%를 뺀 1.1%를 2.1%로 나누면 0.52, 즉 52% 높아졌다는 계산이 나온다. 실업이 크게 증가했다며 정부의 경제정책을 비판하고 싶은 사람은 아마 두 번째 표현을 선호할 것이 당연하다. 이 표현은 5년 전에 비해 실업자가 될 확률이 52% 증가했다는 것을 의미한다.

## ■ ■ 요약

\_ 정책 과정에서 문제에 대한 인식, 정책 대안을 비교, 평가, 선택하는 과정, 그리고 선택된 정책의 집행과 평가에 이르기까지 적절한 자료를 구하여 그 자료를 정확히 이해하는 것이 중요하다. 거기서 문제에 대한 타당한 설명과 예측이 나올 수 있기 때문이다.

\_ 통계적 문제를 크게 다양성에 대한 인식, 다양성에 대한 정확한 표현, 다양성에 대한 해석, 표본과 전체 집단의 관계에 대한 이해로 단순화할 수 있다.

\_ 대표적인 자료 수집방법으로 실험과 조사를 들 수 있다. 객관적이고 타당한 증거가 될 자료를 얻기 위해서는 합리적이고 과학적인 설계가 필요하다.

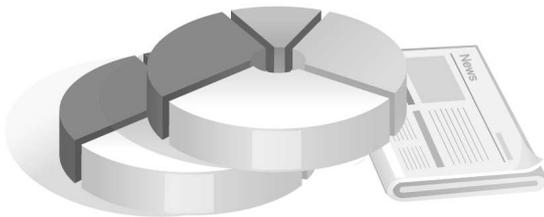
\_ 숫자로 된 통계자료를 효과적으로 정리하는 방법에는 수치 요약법과 표와 그래프를 이용하는 요약법이 있다. 각각의 요약법마다 몇 가지 주요한 고려사항이 있다.

\_ 흔히 접하는 통계들 중에는 지수, 비율, 비 등의 형태로 된 것들이 많다. 그런데 이러한 용어들의 차이를 제대로 이해하지 못해 혼란을 초래하는 경우가 종종 생긴다. 각 통계의 정확한 의미를 이해하는 안목이 필요하다.

## CHAPTER 5

# 정책과정에서 나타나는 쟁점들

- 5.1 TOBANKRUPT 은행의 영업계수는 비정상인가?
- 5.2 신뢰수준 95%에서 표본오차  $\pm 3.1\%$ 포인트는?
- 5.3 정책 만족도가 지역 간에 다르게 나타났는가?
- 5.4 근무 만족도가 높은 기관이 평가도 잘 받는가?
- 5.5 교육효과는 언제부터 나타났는가?
- 5.6 알아두면 유용한 기법들



## 개요

정책과정에서 자주 나타나는 4가지 쟁점들과 이 쟁점들을 다루는데 필요한 개념을 소개하고, 쟁점들을 해결하는 방안들을 설명한다.  
또한 정규분포를 설명하고, 통계적 의사결정이 가질 수밖에 없는 불확실성을 계량적으로 표현하는 신뢰수준과 유의수준을 여러 상황에서 토의한다. 아울러 통계적 비교, 관계, 예측에 관한 개념적 설명과 종래의 통계학 교재에서 잘 다루지 않으나 특히 정책과정에서 유용하게 쓰일 수 있으리라 생각되는 4가지의 통계적 기법과 특수 기법 4가지를 소개한다.

## 주요내용

### 5.1 TOBANKRUPT 은행의 영업계수는 비정상인가?

### 5.2 신뢰수준 95%에서 표본오차 $\pm 3.1\%$ 포인트는?

【사례1】 정책 P의 지지율은 80%쯤 되는가?

### 5.3 정책 만족도가 지역 간에 다르게 나타났는가?

#### 5.3.1 남녀 간 월소득의 차이가 해결되었는가?

【사례1】 남녀 간 월소득의 차이는?

【사례2】 정책에 대한 지지율이 남녀 간에 차이가 있는가?

#### 5.3.2 지역 간 만족도의 차이가 해결되었는가?

【사례3】 정책 만족도가 지역에 따라 차이가 나는가?

### 5.4 근무 만족도가 높은 기관이 평가도 잘 받는가?

#### 5.4.1 산점도와 상관계수

【사례1】 평가점수와 만족도의 관계

#### 5.4.2 관계의 표현과 회귀분석

【사례2】 1983년에 예측해 본 2007년의 출산율은?

### 5.5 교육효과는 언제부터 나타났는가?

#### 5.5.1 시계열 자료

#### 5.5.2 이동평균법

【사례1】 친절교육 효과는 몇 주차부터 나타나기 시작했는가?

#### 5.5.3 시계열 자료 토의

【토의사례1】 성과목표치 설정하기

【토의사례2】 가정용 전력 판매량 예측하기

### 5.6 알아두면 유용한 기법들

#### 5.6.1 주민 만족도 조사의 발전 방향 제시

【사례1】 만족도 조사문항에 가중치 주기

【사례2】 전공 선택에 미치는 요인의 가중치는?

#### 5.6.2 자료 활용의 효용성 제고 방안

【사례3】 경상남도 3개년 사회지표조사 자료 결합하기

#### 5.6.3 지역 특성의 시각화 방안

【사례4】 16개 지역의 품목별 소비자물가지수 종합하기

【토의사례1】 학업능력수준 종합하기

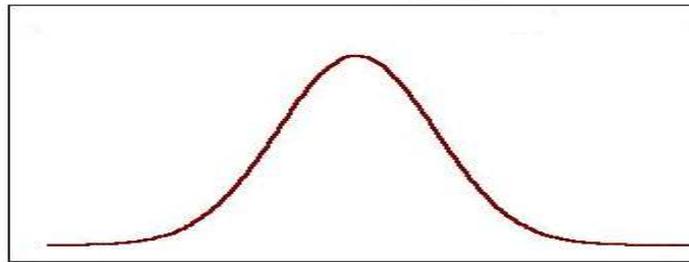
#### 5.6.4 정책지지자 성향분석 방안

【사례5】 어떤 가구들이 인터넷으로 응답하고 있는가?(어떤 사람들이 이 정책을 반대하고 있는가?)



## 5.1 TOBANKRUPT 은행의 영업계수는 비정상인가?

이 절에서는 4.3절에서 다룬 표준화 개념을 완성하기 위하여 반드시 필요한 지식 중 하나인 정규분포를 소개하고, 이를 활용하는 사례를 토의한다. 독일의 수학자 가우스(Gauss, 1777~1855)의 이름을 따서 가우스 분포(Gauss distribution)라고도 불리는 정규분포(normal distribution)는 많은 학문 분야에서 기본 확률모형 또는 근사적인 확률모형으로 쓰이고 있다. 일찍이 통계학 발달의 초기 단계에는 자료의 히스토그램을 그렸을 때 그 모양이 아래와 같은 곡선 형태와 비슷하지 않으면 자료수집 과정이 잘못된 것이라고 믿었던 적도 있었다.



〈그림 5-1〉 정규분포의 일반적인 형태

정규분포는 중심을 나타내는 평균  $\mu$ 와 흩어짐의 정도를 의미하는 표준편차  $\sigma$ 에 의해 결정되는 분포이다. 어떤 확률현상이 정규분포를 따른다는 것은 그 현상을 관찰하여 얻은 결과들을 히스토그램으로 그리면 <그림 5-1>과 같은 종모양(확률밀도함수)에 가깝다는 뜻이다. 어떤 확률현상이 평균이  $\mu$ 이고 표준편차가  $\sigma$ 인 정규분포를 따른다면 다음이 성립한다.

- ▷ 이 현상의 한 관찰값이 ( $\mu - 1\sigma$ ,  $\mu + 1\sigma$ )에서 나타날 확률이 0.6826
- ▷ 이 현상의 한 관찰값이 ( $\mu - 2\sigma$ ,  $\mu + 2\sigma$ )에서 나타날 확률이 0.9544
- ▷ 이 현상의 한 관찰값이 ( $\mu - 3\sigma$ ,  $\mu + 3\sigma$ )에서 나타날 확률이 0.9974

예를 들어보면 “어느 지역 4인 가구의 생활비가 평균 170만원, 표준편차 17만원인 정규분포를 따른다고 가정했다”는 말은 이 지역 가구 중 약 68%의 생활비가 153만원 ( $170 - 1 \times 17 = 153$ )에서 187만원( $170 + 1 \times 17 = 187$ ) 사이라고 생각하겠다는 것이다. 또한 170만원보다 2배의 표준편차 이상 더 쓰는 사람들, 그러니까 204만원 ( $170 + 2 \times 17 = 204$ )이상 생활비를 쓰는 가구가 전체의 약 2.28%( $\frac{1-0.9544}{2}$ ) 라고 생각하겠다는 말이다.

앞서 4장에서 개체의 관찰값과 평균의 차이를 표준편차로 나눈 값을 표준점수라고 소개하였다. 이렇게 구해진 표준화 값들은 원래 관찰값이 신장 자료이든 체중 자료이든 관계없이 평균 0, 표준편차 1인 값들이 된다.

앞의 예에서 평균이 100이고 표준편차가 10인 정규분포를 따른다고 가정한 현상에서 개체들의 표준화 값은 자동적으로 평균이 0이고 표준편차가 1인 형태의 정규분포, 즉 표준정규분포(standard normal distribution)를 따르게 된다. 대부분의 통계학 책의 부록으로 붙어있는 표준정규분포표가 이 분포에 대한 것이다. 예를 들어 원래값이 120 이상일 확률은 표준정규분포에서 120의 표준화 값인  $\frac{120-100}{10} = 2$  이상인 확률이 되고, 그 값은 컴퓨터가 없어도 정규분포표를 이용하면 쉽게 구할 수 있다. 표준화 값을 구하는 식은 다음과 같다.

$$\text{표준화 값} : Z = \frac{\text{관찰값} - \mu}{\sigma}$$

따라서  $Z$ 는 다음과 같은 특징을 갖게 된다.

$$P(-1 < Z < 1) = 0.6826$$

$$P(-2 < Z < 2) = 0.9544$$

$$P(-3 < Z < 3) = 0.9974$$

표준화 값  $Z$ 는 자료값과 평균과의 차이가 그 집단의 표준편차의 몇 배에 해당하는지 알려주므로 어떤 값에 해당하는 표준화 값을 통해 그 값이 집단에서 갖는 상대적 위치를 알 수 있다.

인터넷에서 「**표준화 점수**」를 치면 곧바로 볼 수 있는 제목들이 “등급은 높는데 표준화 점수가 뭐냐?”, “표준화 점수를 이용하여 지방자치단체 웹사이트의 실증적 평가를 수행” 등등

이다. 여기서 표준화에 대해 조금 더 알아보자. 어떤 학생이 중간시험에서 100점 만점에 90점을 받았다. 이 학생의 성적은 우수한 성적인가? 물론 절대평가라면 우수한 성적이라고 할 수 있지만, 만일 상대평가라면 90이라는 점수만으로는 이 학생의 성적을 판단할 수가 없을 것이다. 다른 학생들의 성적이 어떤가에 따라 높은 점수일 수도 있고, 낮은 점수일 수도 있기 때문이다.

그렇다면 어떤 정보가 더 필요한가? 평균보다 얼마나 높은가를 알면 될까? 그렇다면 중간고사 성적은 평균보다 20점 높고 기말고사 성적은 평균보다 10점 높다면, 어떤 시험에서 더 좋은 성적을 얻은 것인가? 평균과의 차이만으로 이 학생의 성적을 판단할 수 있는가? 만일 중간고사 성적의 표준편차가 20점이고, 기말고사 성적의 표준편차가 5점이라면 어떻겠는가? 중간고사에서의 20점 차이는 중간고사 성적 표준편차의 한 배에 해당하는 점수인 반면, 기말고사에서 10점 차이는 기말고사 성적 표준편차의 두 배에 해당하는 점수이다. 따라서 중간고사에서의 20점 차이가 기말고사 10점 차이보다 더 작은 차이가 된다.

두 점수의 표준화 값을 구해보면 중간고사 성적의  $Z$ 값은 1이고, 기말고사 성적의  $Z$ 값은 2이므로, 중간고사 성적은 상위 15.87%에 해당하고 기말고사 성적은 상위 2.28%에 해당함을 알 수 있다. 따라서 평균보다 10점이 높은 기말고사 성적이 평균보다 20점 높은 중간고사 성적보다 훨씬 더 높은 성적인 셈이다.

이렇게 표준화 값은 자료의 상대적 위치를 알려주기 때문에 단위는 같지만 중심이나 퍼진 정도가 다른 자료나, 무게와 부피처럼 단위가 다른 자료를 비교할 수 있게 해준다. 즉, 표준화 값  $Z = \frac{X - \mu}{\sigma} = k$ 는 임의의 개체가 갖는 값  $X$ 와 그가 속한 집단의 평균( $\mu$ )과의 차이가 표준편차( $\sigma$ )의  $k$ 배임을 나타낸다. 정규분포에 관한 설명은 쉽게 접할 수 있으므로 사례나 연습문제는 생략하도록 한다.



## 5.2 신뢰수준 95%에서 표본오차 $\pm 3.1\%$ 포인트는?

사례 1 정책 P의 지지율은 80%쯤 되는가?

### [ 보도 내용 ]

150만 명이 거주하는 특정 지역에서 정책 P에 대하여 여론조사를 실시한 결과 80%가 지지하는 것으로 나타났다. 이 조사는 단순임의추출법으로 표본 1,000명을 추출하여 면접조사로 실시되었고, 신뢰수준 95%에서 표본오차는  $\pm 3.1\%$  포인트이다.

보도내용처럼 조사결과에서 나타난 지지율 80%를 이용해서 이 지역 주민의 지지율을 추측하는 것에 대해서 여러분의 마음은 어떤가? 아래와 같은 생각들이 떠오를 수 있을 것이다.

- ① 불안하다. 150만 명의 의견을 1,000명의 의견으로 말하는 것이니까.
- ② 1,000명은 너무 적은 것 아닌가? 전체가 150만 명이나 되는데...
- ③ 글썄요. 지난번 신문에서 보니까 국민여론조사를 하는데도 1,000명 정도 밖에 안하던데...
- ④ 1,000명으로부터 나온 지지율 80%를 전체 주민의 지지율로 발표해도 되겠는가?

떠오른 생각 ①, ②, ③은 전체 주민의 수 150만 명을 상당히 의식하는 것인데 이 점은 전체 주민의 크기가 추출하려는 표본의 크기보다 훨씬 크면 그렇게 신경 쓰지 않아도 된다. 먼저 ④에 대하여 토의해 보자.

바로 이런 염려 때문에 굳이 “신뢰수준 95%에서 표본오차  $\pm 3.1\%$  포인트”라는 말을 하게 되는 것이다. “표본오차  $\pm 3.1\%$  포인트”라는 말은 전체 주민의 지지율이  $80\% \pm 3.1\%$ 인 76.9%와 83.1% 사이에 있을 것이라고 본다는 뜻이고, 이때 신뢰수준 95%라는 말은 이와

같은 조사를 만약 100회 반복하여 발표를 한다면 그 중에 95번 정도는 맞다는 것을 의미한다. 그런데 이 신뢰수준을 조금 다르게 생각해 본다면 양극단을 추구하는 상반된 이익단체의 불만을 해결할 수 있다고 본다. 우리가 얻은 실제 표본비율 0.8(80%)을 해석하는 두 가지 상반된 입장을 생각해보자.

① 표본을 추출할 때에는 어떻게 하든 전체 주민을 잘 대표하는 표본이 되도록 할 것이다. 그러나 혹시 우리가 얻은 표본에 우연히도 반대하는 사람들이 많이 포함되어 있었다면 우리가 얻은 0.8(80%)은 전체 주민의 지지율  $P$ 보다 작을 것이다. 그렇다고 해서 우리가 표본으로부터 얻은 표본비율 0.8(80%)이 우리가 얻을 수 있었던 가능한 표본비율 값들 중에서 “제일 작은 값”이라고 생각할 필요까지는 없지 않을까? 벼락이 친다고 해서 벼락을 내가 맞을 거라고 생각하지 않는 것처럼 말이다.

② 반대의 경우를 생각해 보자. 혹시 우리가 얻은 표본에 우연히도 지지하는 사람들이 많이 포함되어 있었다면 우리가 얻은 0.8(80%)은 모비율  $P$ 보다 큰 값일 것이다. 그렇다고 해서 0.8(80%)은 가능한 표본비율 값들 중에서 제일 큰 값이라고까지 생각하는 것은 너무 극단적이지 않을까?

성격적으로 극단적인 사람들은 자기에게 극단적인 일이 일어난다고 생각하는 성향이 있다. 시험점수를 받으면 왕자병에 걸린 사람은 자기가 “1등 아닌가?” 생각하고 소심병에 걸린 사람은 자기가 “꼴찌 아닌가?”라고 걱정하는 경우를 생각해 보면 이해가 될 것이다. 따라서 우리가 얻은 0.8(80%)이  $P$ 에서 떨어져 있을 수 있다는 것을 인정하려면 우리가 극단적이라고 말할 만한 상식적인 기준이 어느 정도인지를 정하고 말을 해야 한다. 여기서 신뢰수준 95%란 통계적으로 양극단 2.5%, 즉 상위 2.5%보다 큰 값이거나 하위 2.5%보다 작은 값을 극단으로 보겠다는 뜻이다.



## 5.3 정책 만족도가 지역 간에 다르게 나타났는가?

어느 지역에서 사무직과 생산직 근로자들의 월평균 임금을 표본조사한 결과 사무직 근로자의 월평균 임금은 152만원이고 생산직 근로자의 월평균 임금은 138만원인 것으로 나타났다고 하자. 이를 근거로 사무직과 생산직 근로자의 임금은 차이가 난다고 결론을 내려도 될까? 혹시 우연히도 사무직에서는 임금이 높은 근로자들이 표본에 많이 뽑히고 생산직에서는 그 반대로 임금이 낮은 근로자들이 많이 추출되었을 수도 있지 않은가?

위의 예와 같이 여러 집단들 간의 평균을 비교하는 문제는 여러 분야에서 매우 자주 다루어지는 주제이다. 가령, 수도권과 지방 고등학생들 간의 학력 비교, 학력 수준별 근로자들의 임금 격차 비교, 자동차 연비 비교, 서로 다른 다이어트 프로그램의 효과 비교, 대학생들의 전공별 만족도 비교 등 여러 분야에서 이와 같은 문제가 생길 수 있다.

통계학에서 서로 다른 모집단들의 평균을 비교하는 문제는 크게 두 집단 간의 비교 문제와 셋 이상의 여러 집단 간의 비교 문제로 구분된다. 두 집단의 비교 문제에서는  $z$ -검정 또는  $t$ -검정이라는 통계적 기법을 주로 사용하며, 셋 이상 여러 집단의 비교 문제는 분산분석법 (ANalysis Of VAriance: ANOVA)을 사용한다.

몇 개의 집단을 비교하든지 상관없이 통계적 비교의 문제에서는 분석에 들어가기에 앞서 자료수집 단계의 객관성, 공정성이 중요하다는 점을 다시 한 번 지적할 필요가 있다. 여러 집단에서 각각 자료를 수집할 때에는 집단이라는 특성 이외의 다른 조건들은 모두 비슷하게 되도록 조건을 통제(control)해야 한다.

### 5.3.1 남녀 간 월소득의 차이가 해결되었는가?

#### 사례 1 남녀 간 월소득의 차이는?

특정 업종에 근무하는 경력이 1년 된 남녀 근로자들의 월소득이 다음과 같이 조사되었다고 하자.

〈표 5-1〉 특정 업종에 근무하는 남녀 근로자들의 월소득

(단위 : 만원)

남자	141 143 145 147 149 151 167 155 157 159 179 163 165 175
여자	169 173 153 177 161 181 183 185 187 189 191 193 195 197 199 171

이 자료를 이용하여 월소득의 평균과 표준편차를 성별에 따라 구해보면 다음과 같다.

〈표 5-2〉 특정 업종에 근무하는 남녀 근로자들의 월소득의 평균과 표준편차

성별	빈도	평균	표준편차
남자	14	156.9	11.81
여자	16	181.5	13.28

전체 30명의 자료를 통해 나타난 남자와 여자의 월소득을 살펴보면 남자는 평균 156.9만원이고 여자는 평균이 181.5만원으로 평균적으로 여자가 남자에 비해 24.6만원만큼 더 많이 받는 것으로 나타났다. 반면에 표준편차를 보면 남자는 11.81만원이고 여자는 13.28만원으로 여자가 남자에 비해 개개인의 월소득 차이가 조금 더 큰 것으로 나타났다. 자료에서 나타난 평균차이 24.6만원으로부터 우리는 남자와 여자의 평균 월소득이 차이가 난다고 일반적인 결론을 내릴 수 있겠는가? 이러한 질문에 대해 답하는 분야를 통계학에서는 가설검정이라고 한다.

지금까지 논의한 내용만 이해하더라도 단순히 24.6만원의 평균 차이를 갖고 바로 여자가 남자보다 일반적으로 월소득이 높다고 할 수 없다는 것을 알 것이다. 대표적으로 알려져 있는 통계적 방법으로는 두 집단 비교 t검정이 있는데 대부분의 일반 통계학 교재에 상세히 기술되어 있다.

<그림 5-2>는 이 자료를 엑셀을 이용하여 분석한 결과이다. 분석결과는 유의확률을 나타내는 ' $P(T \leq t)$  양측검정'이 0.000으로 유의수준 0.05보다 작기 때문에 통계적으로 남녀 간에 소득차이가 있다는 결론을 내린다.

t-검정: 등분산 가정 두 집단		
	변수 1	변수 2
평균	156.86	181.50
분산	139.36	176.27
관측수	14	16
공동(Pooled) 분산	159.13	
가설 평균차	0	
자유도	28	
t 통계량	-5.3380	
P(T<=t) 단측 검정	0.0000	
t 기각치 단측 검정	1.7011	
P(T<=t) 양측 검정	0.0000	
t 기각치 양측 검정	2.0484	

<그림 5-2> t-test 결과

사례 2 정책에 대한 지지율이 남녀 간에 차이가 있는가?

다음은 성별에 따라 특정 정책에 대한 지지율이 다른지 알기 위해서 남자 60명, 여자 40명을 적절한 표본추출 이론에 입각하여 추출하여 조사한 결과이다.

〈표 5-3〉 성별에 따른 특정 정책의 지지현황

성별	지지여부		합계
	찬성	반대	
남자	40	20	60
여자	20	20	40
합계	60	40	100

성별에 따라 지지율의 차이가 있다고 할 수 있겠는가? 자료에서만 보면 남자가 약 67%, 여자가 50%로서 남자가 여자보다 높은 지지율을 나타내지만 그렇다고 이것을 전체 국민(또는 주민)에게 단순히 적용할 수는 없다. 왜냐하면 이 결과는 특정의 남자 60명, 여자 40명에게서 나타난 것이기 때문이다. 만약 이들 말고 다른 남자 60명, 여자 40명을 표본 추출하였다면 그 결과는 아마도 거의 확실하게 달랐을 것이다. 바로 이러한 상황에서 의사결정을 해야만 하기 때문에 이를 통계적 의사결정이라고 하는 것이다.

z-검정: 평균에 대한 두 집단		
	변수 1	변수 2
평균	0.67	0.50
기지의 분산	0.23	0.26
관측수	60	40
가설 평균차	0	
z 통계량	1.6396	
P(Z ≤ z) 단측 검정	0.0505	
z 기각치 단측 검정	1.6449	
P(Z ≤ z) 양측 검정	0.1011	
z 기각치 양측 검정	1.9600	

〈그림 5-3〉 z-test 결과

이러한 자료를 분석하는 기법을 두 집단의 모비율 비교 검정이라고 한다. 관심 있는 독자는 일반 통계학 교재를 참고하기 바란다.

<그림 5-3>은 이 자료를 엑셀을 이용하여 분석한 결과이다. 유의확률이 0.1011이므로 유의수준 0.05보다 크기 때문에 통계적으로는 남녀 간에 지지율 차이가 뚜렷하지 않다는 결론을 내린다.

### 5.3.2 지역 간 만족도의 차이가 해결되었는가?

앞 절에서는 두 집단을 비교하는 문제를 다루었다. 그런데 두 모집단이 아닌 셋 이상의 모집단 평균을 비교해야 하는 경우도 많이 생긴다. 이때에는  $t$ -검정이 아닌 분산분석이라는 통계기법을 사용하게 된다.

분산분석은 처음에는 실험 자료를 분석하기 위해 고안된 방법이다.  $t$ -검정은 두 모평균의 비교를 할 수 있는데 반해 분산분석이라는 통계분석 방법은 보다 일반화된 비교 방법으로서 셋 이상의 모평균 비교를 할 때 쓸 수 있다. 분산분석이란 측정값들이 평균을 중심으로 얼마나 떨어져 있는지를 나타내는 변동(variation)을 분석하는 방법인데, 우리가 관심을 갖는 요인이 관측값의 변동에 얼마만큼 영향을 미치는가 파악한 후 적절한 통계적 기준으로 판정하는 방법이다. 따라서 분산분석에서 사용되는 용어들은 실험과 연관되어 있다. 하지만 분산분석법 자체는 조사 자료를 분석할 때에도 유용하게 쓰이고 있다.

**사례 3** 정책 만족도가 지역에 따라 차이가 나는가?

네 지역 주민들의 특정 정책에 대한 만족도를 비교하기 위하여 적절한 표본추출 방법에 따라 5명씩 추출하여 조사한 결과가 다음과 같다고 하자.

〈표 5-4〉 4개 지역의 정책 만족도

지역 1	지역 2	지역 3	지역 4
17.8	11.2	11.8	14.9
16.2	11.4	11.0	10.8
17.5	15.8	10.0	12.8
17.4	10.0	9.2	10.7
15.0	10.4	9.2	10.7

이런 상황에서 연구자가 주로 갖게 되는 관심사항은 “지역 간에 정책에 대한 만족도가 다르게 나타나는가?” 일 것이다. 이것을 다르게 표현하면 다음 중 하나의 입장을 선택하는 것이다.

- 지역 간 정책 만족도에는 차이가 없다.
- 지역 간 정책 만족도에는 차이가 있다.

먼저 이 자료를 분산분석이라는 통계적 기법으로 엑셀을 이용하여 분석한 결과를 보자.

분산 분석: 일원 배치법						
요약표						
인자의 수준	관측수	합	평균	분산		
지역 1	5	83.9	16.78	1.36		
지역 2	5	58.8	11.76	5.43		
지역 3	5	51.2	10.24	1.31		
지역 4	5	59.9	11.98	3.47		
분산 분석						
변동의 요인	제곱합	자유도	제곱 평균	F 비	P-값	F 기각치
처리	120.498	3	40.166	13.8923	0.0001	3.2389
잔차	46.260	16	2.891			
계	166.758	19				

〈그림 5-4〉 분산분석 출력 결과

우리는 두 집단 비교에서와 같은 어려움을 갖게 된다. 자료에서 보면 4개 지역의 평균만족도는 16.78(지역1)부터 11.98(지역4)까지 다양하게 나타났는데, 이것을 일반적으로 각 지역 주민 전체의 평균만족도 차이로 해석할 수 있느냐 하는 것이다. 역시 어려움은 우리가 모든 주민을 조사한 것이 아니고 각 지역에서 5명의 표본만을 추출하여 조사했다는 것이다. 이 어려움을 다르게 표현하면, 4개 지역의 전체 만족도가 모두 같다고 하더라도 이들 지역에서 표본으로 뽑힌 5명의 평균값 사이에는 우연히 이런 정도의 차이가 날 수 있을지도 모른다는 것이다. 이러한 어려움을 논리적으로 잘 정리한 것이 분산분석이라는 통계적 기법인데 위의 엑셀 결과표의 P-값이라는 것을 보면 “0.0001” 이라고 되어 있다. 이 값이 바로 위에서 설명한 “우연의 정도”로 해석하면 된다.

즉, 4개 지역의 전체 만족도가 모두 같을 때, 각 지역에서 임의로 뽑힌 5명으로부터 얻게 되는 평균만족도가 자료와 같은 차이를 우연히 나타낼 가능성이 0.0001(약 0.01%)라는 것이다. 이것은 일반적인 상황에서는 상당히 희귀한 일일 것이다. 여기서부터는 여러분이 토의를 해 보는 것이 좋을 것이다.



## 5.4 근무 만족도가 높은 기관이 평가도 잘 받는가?

표본으로부터 자료를 수집할 때 관심변수가 하나 또는 두 개에 불과한 경우도 있지만 설문 조사 등과 같이 한 대상으로부터 수많은 변수들을 관측하는 경우도 많다. 이 때 개개의 변수들에 대해 자료를 정리하고 추론하는 것 뿐 아니라 한걸음 더 나아가 변수들 상호간의 관계에 대해 관심을 기울이는 것은 자연스러운 일이다. 학생들의 키와 몸무게 사이의 관계, 교육 수준과 월평균소득 사이의 관계, 유권자들의 연령과 정치적 성향과의 관계 등은 흔히 관심을 갖고 살펴보게 되는 관계들이다. 이런 경우 우리는 한 변수의 값이 변화하면 다른 변수의 값도 영향을 받을지 알아보고 다른 변수의 값을 예측하는 일에도 관심을 갖게 된다.

변수들의 연관성을 파악하기 위하여 다양한 통계적 방법들이 개발되어 사용되고 있는데 변수들의 특성에 따라 그 성격은 달라진다. 즉, 질적 변수나 양적 변수나에 따라 적용되는 통계적 방법이 다른데 5.4.1절에서는 양적 변수들에 대해 다루고 5.4.2절에서는 질적 변수들의 관계에 대해 알아보자.

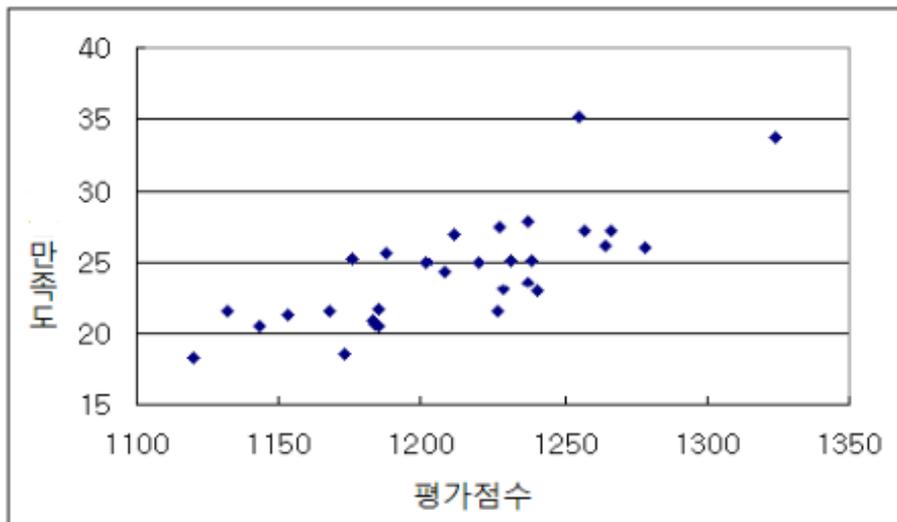
두 양적변수들 사이의 연관관계를 파악하는 순서를 생각해보면 첫 단계로는 자료를 그림으로 나타내어 직관적으로 변수들 사이의 관련성을 파악하는 것이다. 그림을 통해 변수 간 관계의 방향과 정도 등을 짐작할 수 있다. 다음 단계에서는 변수들 사이의 관계의 방향이나 정도를 계량화된 척도로 나타내는 것이 필요하다. 마지막 단계는 변수들 사이에 관계를 나타내는 통계적 모형을 가정한 후 모형에 대해 추론하는 단계이다. 양적변수들 간의 관계를 나타내기 위한 그림으로 산점도를 널리 사용한다. 또한 변수들 사이의 관계를 표현하기 위한 계량적인 척도로는 상관계수가 널리 사용된다. 변수들 사이에 관계에 대해 통계적 모형을 가정하고 추론하는 방법은 회귀분석이다.

## 5.4.1 산점도와 상관계수

### (1) 산점도

#### 사례 1 평가점수와 만족도의 관계

다음의 <그림 5-5>는 어느 기관에서 30명의 직원들을 대상으로 조사한 만족도(40점 만점)와 그들이 받은 평가점수(1,500점 만점) 사이의 관계를 나타내는 산점도이다.



<그림 5-5> 평가점수와 만족도의 산점도

그림을 보면 평가점수가 높을수록 만족도가 커지는 경향을 보이는데, 이와 같은 경우 두 변수는 양(+)의 연관관계를 지닌다고 한다. 또한 그림의 점들이 하나의 직선을 중심으로 어느 정도 집중되어 있는지도 살펴볼 수 있다. 만일 점들이 가상의 직선을 중심으로 집중되어 있다면 연관의 정도가 강한 것이고, 점들이 직선을 중심으로 넓게 흩어져 있으면 연관의 정도는 약한 것이다. 이와 같이 산점도는 두 변수간의 연관성의 방향과 정도를 파악할 수 있는 그림이다.

그 밖에도 산점도를 통해 자료 중 이상점(outlier)이나 영향점(influential point)과 같이 전반적인 패턴에서 벗어나는 값이 존재하는지 손쉽게 파악할 수 있다. 그러므로 변수들 사이의 관계를 파악하기 위해서는 바로 수치 측도에 대한 분석에 들어갈 것이 아니라 사전에 산점도를 그려 자료의 특성을 미리 탐색하는 것이 반드시 필요하다.

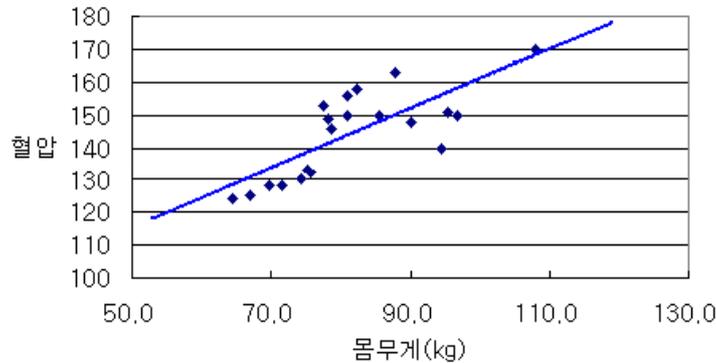
## (2) 상관계수

두 변수간의 관계의 방향과 정도를 재기 위한 측도로 가장 널리 사용되는 것이 바로 피어슨의 상관계수(Pearson's correlation coefficient)인데 일반적으로 상관계수라고 하면 피어슨의 상관계수를 일컬으며  $r$ 로 표시한다.

상관계수  $r$ 은  $-1$ 에서  $+1$  사이의 값을 갖는다. 상관계수의 부호는 두 변수간의 관계의 방향을 나타낸다. 상관계수가 양수이면 두 변수는 양의 관계이다. 즉, 한 변수의 값이 증가하면 다른 변수의 값도 증가하는 관계를 말한다. 그리고 이러한 관계는 키와 몸무게의 관계, 근무연수와 연봉과의 관계, 광고액과 매출액의 관계 등에서 나타난다. 상관계수가 음수이면 두 변수는 음의 관계이다. 즉, 한 변수의 값이 증가하면 다른 변수의 값은 감소하는 관계를 말한다. 그리고 이러한 관계는 살충제를 뿌린 양과 병충해의 수의 관계, 엔진의 가격과 소음의 크기와의 관계 등에서 볼 수 있다. 한편 상관계수  $r$ 의 절대적인 크기는 관계의 강도를 나타낸다. 이 값이 1에 가까울수록 두 변수의 상관의 강도는 강한 반면, 0에 가까울수록 상관의 강도는 약하다.

## 5.4.2 관계의 표현과 회귀분석

<그림 5-6>은 성인남자의 몸무게와 최대혈압 자료를 산점도로 그린 것이다. 참고로 몸무게와 최대혈압의 상관계수는 0.7644로 나타났다. 이러한 산점도를 대할 때 그림에 나온 바와 같이 좌표상의 점들을 대표할 만한 직선을 자연스럽게 생각할 수 있다. 회귀분석(regression analysis)이란 두 변수가 서로 함수관계를 갖는다고 가정하고 자료를 가장 잘 나타내는 함수를 추론하는 통계적 방법이라고 할 수 있다.



〈그림 5-6〉 몸무게와 혈압의 산점도

두 변수 사이의 관계를 함수식으로 표현할 수 있다면 여러 가지로 유용하다. 한 변수값의 변화가 다른 변수값에 어떤 영향을 미치게 될지 설명할 수 있으며, 더 나아가서 한 변수값이 주어질 때 다른 변수값이 얼마가 될지 예측할 수도 있을 것이다. 자료로부터 유용한 정보를 최대한 추출하는 것이 통계학의 목표라고 본다면 회귀분석은 그런 목표에 충실한 통계적 방법이라고 할 수 있다.

회귀분석에서 다른 변수에 영향을 미치는 변수를 설명변수(explanatory variable) 또는 독립변수(independent variable)라고 하며, 영향을 받는 변수를 반응변수(response variable) 또는 종속변수(dependent variable)라고 부른다. 위에서 소개한 예의 경우, 몸무게가 설명변수이고 최대혈압은 반응변수라고 할 수 있다. 회귀분석에서 반응변수를 설명하기 위해 하나의 설명변수를 사용하는 것을 단순회귀(simple regression), 두 개 이상의 설명변수를 이용하여 반응변수를 나타내는 것을 다중회귀(multiple regression)라고 한다.

회귀분석은 통계학의 여명기인 19세기 영국의 골턴(Francis Galton)이 부모의 키와 자녀 키의 관계를 설명하는 과정에서 태동된 통계적 방법인데, 부모의 키가 평균키보다 훨씬 크면 자녀의 키 역시 평균키보다 크기는 하지만 평균으로 돌아가는 경향을 보인다는 뜻에서 ‘평균으로의 회귀(回歸, regression)’라는 용어를 사용했다고 한다.

회귀분석의 과정을 크게 몇 가지 단계로 구분하자면, 첫째 회귀계수를 추정하여 회귀직선을 찾는 과정, 둘째 회귀직선의 유의성을 검정하는 과정, 셋째 회귀직선의 정밀성을 평가하는 과정, 넷째 모형에 필요한 가정을 검토하는 과정, 마지막으로 적절한 회귀모형을 제시하고 설명하는 과정으로 나눌 수 있다. 관심이 있는 독자는 통계학 개론 수준의 교재나 회귀분석에 관한 전문 교재를 참고하기 바라고, 여기서는 국내 정책과 관련된 사례 하나를 소개한다.

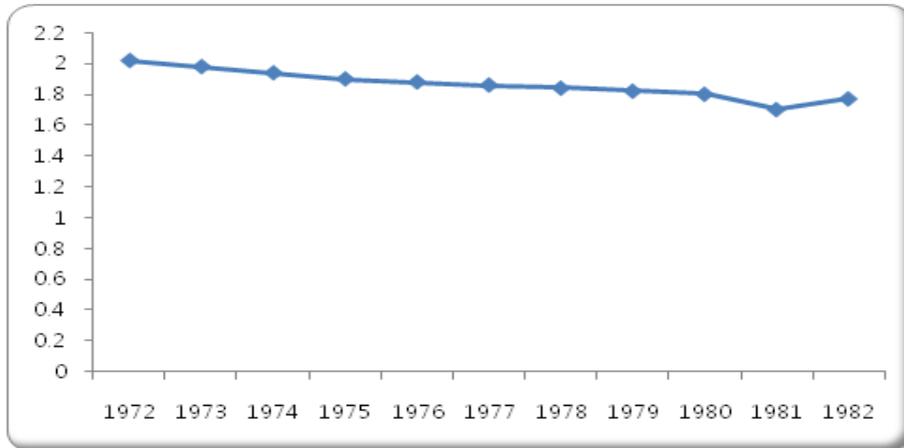
사례 2 1983년에 예측해 본 2007년의 출산율은?

2000년 중반 이후 우리 사회의 화두 중 하나는 “저출산” 문제이다. 인구정책에 관심을 갖고 있는 많은 사람들은 1983년을 기억한다. 왜냐하면 1983년에 우리나라는 소위 인구대체수준이라고 불리는 합계출산율 2.08을 기록하였기 때문이다. 따라서 이 시점에서 당시 인구증가억제를 기조로 하는 정책에 대한 논의가 있었다고 본다. 박원란(2008) 등의 연구자들은 1983년 당시에 사용할 수 있는 자료들 중 이미 합계출산율이 감소하고 있었던 일본의 1972년부터 1982년 합계출산율을 회귀모형을 통하여 분석하고 일본의 감소추세를 우리나라가 따라간다는 가정 하에 1983년 시점에서 2008년까지의 합계출산율을 예측하여 보았다.

다음 <표 5-5>와 <그림 5-7>은 당시의 일본의 1972년부터 1982년까지의 합계출산율 자료와 꺾은선그래프이다.

<표 5-5> 일본의 합계출산율

년도	일본
1972	2.02
1973	1.98
1974	1.94
1975	1.90
1976	1.88
1977	1.86
1978	1.84
1979	1.82
1980	1.80
1981	1.70
1982	1.77

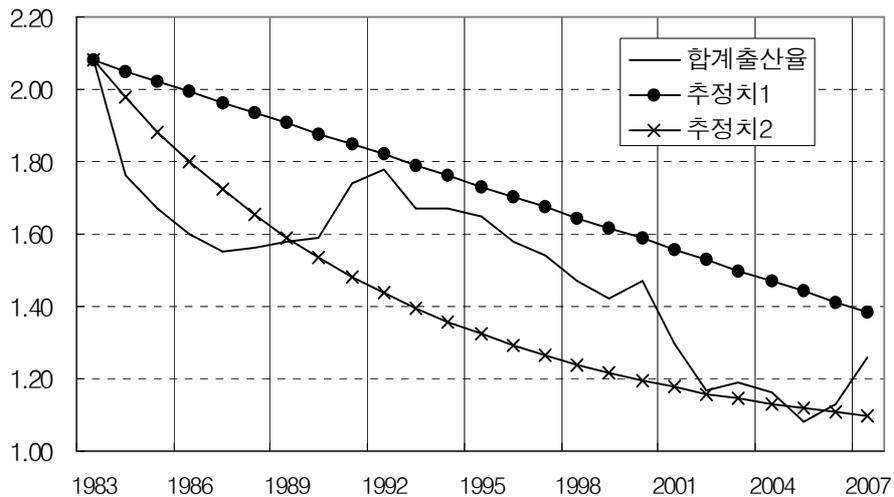


〈그림 5-7〉 일본의 합계출산율의 꺾은선그래프

이 자료를 회귀분석하여 회귀식을 다음과 같이 얻었다.

$$\text{일본의 합계출산율} = -0.03 \times \text{년} + 55.60$$

1년에 합계출산율이 0.03씩 낮아진다는 것을 알았다. 따라서 우리나라의 합계출산율이 1983년부터 매년 0.03씩 낮아진다고 보고, 2007년까지 예측해 보면 <그림 5-8>의 그래프와 같이 나타난다.



〈그림 5-8〉 합계 출산율과 추정치와의 비교

여기서 실선은 1983년부터 2007년까지 우리나라의 실제 출산율이고, 추정치1은 일본의 추세를 적용한 것이며 추정치2는 우리나라의 1972년부터 1982년 자료를 다소 복잡한 비선형 회귀모형을 사용하여 추정해 본 것이다. 이 그림에서 우리는 1983년 시점에서 사용가능했던 자료로부터 우리나라 합계출산율을 예측하는 것이 어느 정도 가능했었음을 알 수 있는데, 이러한 분석이 당시 정책결정에 많이 반영되었다면 하는 아쉬움이 있다.



## 5.5 교육효과는 언제부터 나타났는가?

### 5.5.1 시계열 자료

#### (1) 시계열(時系列) 자료

시계열 자료란 시간의 흐름에 따라 변화하는 관찰치들을 일정한 시간 간격으로 배열해 놓은 자료로서 경영관리자나 사회과학 연구자들이 시간에 따라 변화하는 경향이나 예측에 관심을 가질 때 이용한다. 예로 월별 판매고, 연도별 GNP, 월별·연도별 물가지수, 분기별·연도별 산업 생산지수 등이 여기에 해당한다.

#### (2) 시계열의 변동

##### 가. 추세변동(trend)

경제변동 중에서 장기간에 걸친 성장·정체·후퇴 등 변동경향을 나타내는 움직임

##### 나. 순환변동(cyclic variation)

상하의 움직임을 반복하는 변동 중 계절변동에 따라 설명되는 부분을 제외한 것을 말하며, 경제학에서는 이를 경기변동이라고 한다.

##### 다. 기간변동(periodic variation)

12개월 주기의 계절변동이라고 볼 수 있다. 기간을 주기로 나누어 보는 변동

##### 라. 불규칙변동(irregular variation)

우발적 요인에 의해 일어나는 불규칙변동

### (3) 시계열의 두 가지 유형

#### 가. 일정한 간격을 가진 각 시점에서 변량들의 값을 관찰하는 경우

도심지 오염정도를 매일 6시에 관측하는 것.

#### 나. 일정 시간 간격에 걸쳐서 한 변량을 총체적으로 집계하는 경우

월별 판매량 같이 판매량을 한 달에 걸쳐서 집계

#### 다. 이산 시계열 VS 연속 시계열

관찰치들이 시간적으로 분리된 시점에 속하면 이산 시계열, 시간상 연속적으로 측정되는 경우면 연속 시계열(일정 기간의 온도변화, 뇌파 등)

## 5.5.2 이동평균법

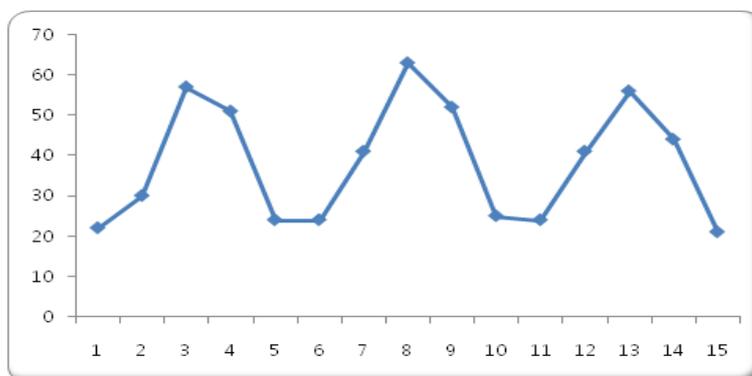
시계열에서 불규칙 변동을 완화시키거나 평준화시켜서 완만한 시계열의 움직임을 나타내는 구성요소를 구하는 통계적 과정을 평활법이라고 하는데, 평활법 중에서 가장 널리 사용되는 방법으로 한 집단의 관찰치들을 연속적으로 평균을 취하여 시계열의 변동을 줄여나가는 방법을 이동평균법이라고 한다. 이것은 연속적으로 놓인 일정한 수인  $k$ 개의 관찰치들의 평균을 취하여 구하는 방법이다. 즉, 관측치들  $Y_1, Y_2, \dots, Y_k$  에서 첫 번째 평균을 구하고,  $Y_2, Y_3, \dots, Y_{k+1}$  에서 두 번째 평균을 구하고 이러한 과정을 반복해 나간다, 여기서 단위 시간인  $k$ 는 시계열 순환 변동의 주기(週期)와 같게 두거나 그 배수로 두는데 일반적으로는 홀수로 둔다.

**사례 1** 친절교육 효과는 몇 주차부터 나타나기 시작했는가?

다음은 어떤 기관이 민원인의 불평이 늘고 있다는 언론의 지적에 대응하기 위하여 일별(주 5일) 불평건수를 조사한 자료이다. 조사를 시작하고 2주가 지난 후 특별한 친절교육을 실시한 바 있다.

〈표 5-6〉 민원인 불평건수 현황

주순	요일	t	민원인 불평건수
1	월	1	22
	화	2	30
	수	3	57
	목	4	51
	금	5	24
2	월	6	24
	화	7	41
	수	8	63
	목	9	52
	금	10	25
3	월	11	24
	화	12	41
	수	13	56
	목	14	44
	금	15	21



〈그림 5-9〉 민원인 불평건수에 대한 꺾은선그래프

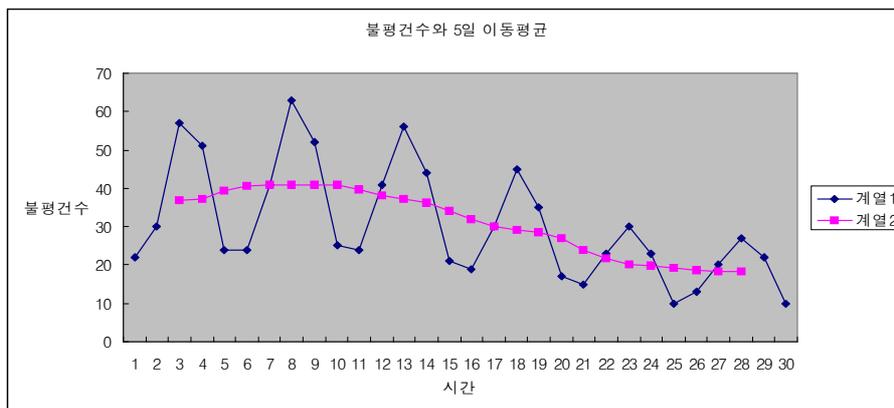
그래프를 보면 불평건수가 줄어드는 현상이 나타나는데 그래도 오르락내리락 하는 변동현상이 보여서 교육효과가 있다고 발표, 보고하기에 충분하지 않다. 그런데 자료나 그래프를 세밀하게 보면 오르락내리락 하는 현상이 규칙적임을 알 수 있다. 즉, 주중에는 불평건수가 많다가 주초, 주말로 가면 작아지는 것이다.

따라서 이런 규칙적인 변동을 의미 있는 순환변동으로 받아들이고, 일단 불규칙적으로 변동하는 것을 제거하기 위해 이동평균법을 다음과 같이 수행했다.

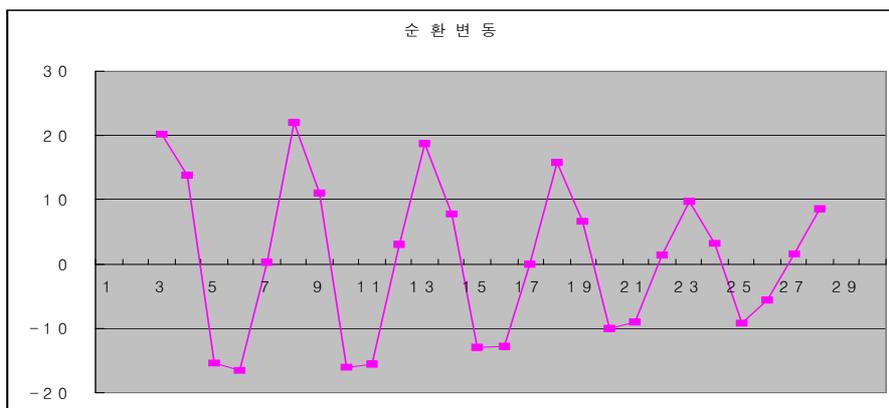
〈표 5-7〉 5항 이동평균 계산을 수행한 결과

주순	요일	t	민원인 불평건수(1)	5항이동총계(2)	5항이동평균치(3)	순환변동(4) (1)-(3)
1	월	1	22			
	화	2	30			
	수	3	57	184	36.8	20.2
	목	4	51	186	37.2	13.8
	금	5	24	197	39.4	-15.4
2	월	6	24	203	40.6	-16.6
	화	7	41	204	40.8	0.2
	수	8	63	205	41	22
	목	9	52	205	41	11
	금	10	25	205	41	-16
3	월	11	24	198	39.6	-15.6
	화	12	41	190	38	3
	수	13	56	186	37.2	18.8
	목	14	44	181	36.2	7.8
	금	15	21	170	34	-13
4	월	16	19	159	31.8	-12.8
	화	17	30	150	30	0
	수	18	45	146	29.2	15.8
	목	19	35	142	28.4	6.6
	금	20	17	135	27	-10
5	월	21	15	120	24	-9
	화	22	23	108	21.6	1.4
	수	23	30	101	20.2	9.8
	목	24	23	99	19.8	3.2
	금	25	10	96	19.2	-9.2
6	월	26	13	93	18.6	-5.6
	화	27	20	92	18.4	1.6
	수	28	27	92	18.4	8.6
	목	29	22			
	금	30	10			

예제 자료에 대해 알아보면 (1)열에 관찰치들이 있으며, (2)열에는 5항 이동총계가 들어 있다. 5항 이동총계이기 때문에 처음 관찰치부터 다섯 개 관찰치의 합계  $22+30+57+51+24=184$ 를 계열의 중앙인  $t=3$ 에 적는다. 이와 같은 방법으로 두 번째는  $30+57+51+24+24=186$ 을  $t=4$  위치에 적는다. 이런 방법으로 반복적으로 (2)열을 완성시키고 같은 방법으로 (2)열을 5로 나누면 (3)열의 이동평균이 나오게 된다. 주의할 것은  $t$ 가 1, 2, 29, 30인 경우에는 이동평균 값이 없다는 것이다. 즉, 처음 2일 간과 마지막 2일 간에는 이동평균 값이 존재하지 않는다. 일반적으로  $k$ 값이 홀수 일 때  $k$ 항 이동 평균에서 처음과 마지막  $(k-1)/2$ 개 시점은 이동평균 값이 존재하지 않는다. (4)열에는 (1)열과 (3)열의 차가 나타나 있는데 관찰치에서 이동평균 값을 뺀 값으로 순환변동이라 불린다. 다음은 예제 자료를 통하여 불평건수와 5일 이동평균, 순환변동의 그래프를 그려 보도록 하겠다.



〈그림 5-10〉 불평건수와 5일 이동평균



〈그림 5-11〉 순환변동

<그림 5-10>에서 계열1은 실제 자료의 그래프이고, 계열2는 5항 이동평균 값의 그래프이다. 시간(x축) 10까지는 친절교육 이전이며 그 후가 친절교육 이후라고 하는데 그래프 상에서 불평건수의 5항 이동평균이 친절교육 이후 곧 하향하고 있다. 이러한 이동평균의 하향은 그 원인이 친절교육 때문일 수도 있고, 이러한 요인과 여타의 요인들의 결합에 기인할 수도 있으므로, 하향의 원인을 규명하려면 좀 더 많은 정보가 필요할 것이다. <그림 5-11>의 순환변동은 주중에 높은 값을 보인다. 이러한 현상은 주중에 방문자가 많기 때문에 일어나는 것인지 또는 다른 요인들의 반응 결과인지는 더 조사를 해보아야 확실해 질 것이다.

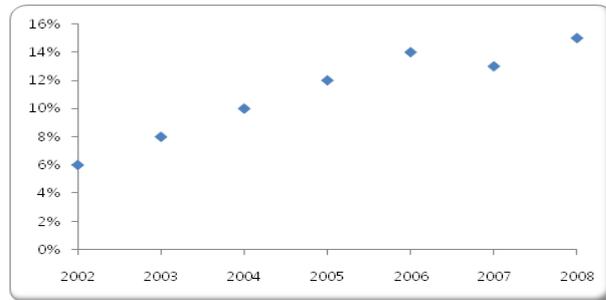
### 5.5.3 시계열 자료 토의

#### 토의사례 1 성과목표치 설정하기

다음은 지난 7년간의 성과달성도이다. 이 자료를 바탕으로 2009년도 성과목표를 설정하는 토의를 해 보자.

〈표 5-8〉 연도별 성과달성도

년도	성과달성도(%)
2002	6
2003	8
2004	10
2005	12
2006	14
2007	13
2008	15



〈그림 5-12〉 연도별 성과달성도

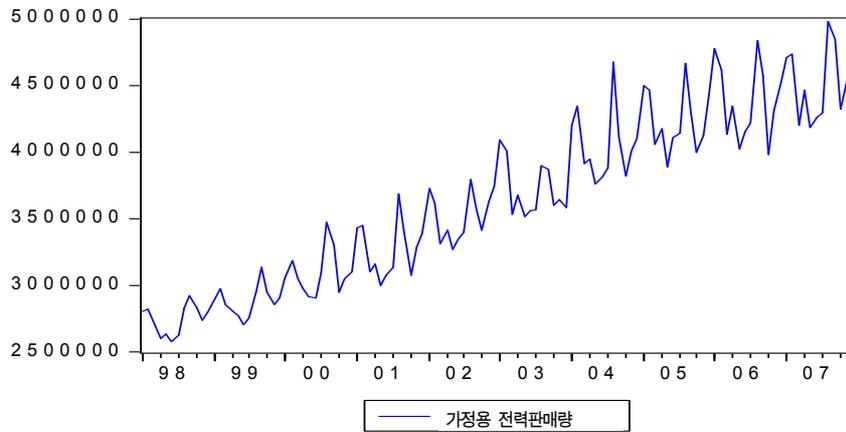
## 토의사례 2

## 가정용 전력 판매량 예측하기

다음은 1998년부터 2007년까지의 매월 가정용 전력 판매량을 보여주는 자료와 그래프이다.

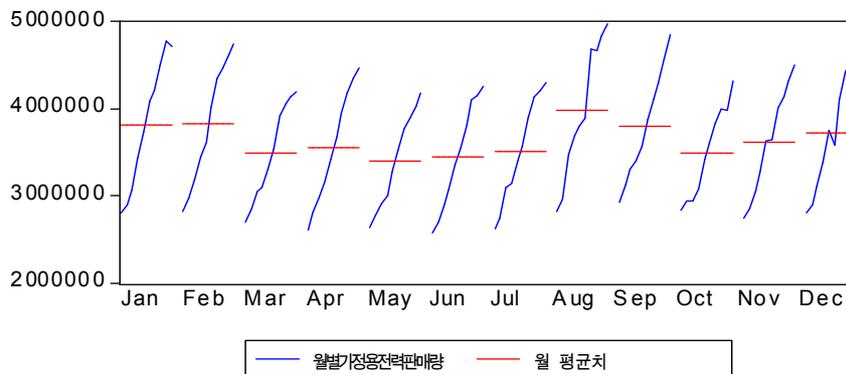
〈표 5-9〉 가정용 전력 판매량

년월	$z_t$	년월	$z_t$	년월	$z_t$	년월	$z_t$
1998M01	2805573	2001M01	3433121	2004M01	4204620	2007M01	4706885
1998M02	2820949	2001M02	3450056	2004M02	4348752	2007M02	4736309
1998M03	2700132	2001M03	3104060	2004M03	3916366	2007M03	4202932
1998M04	2605285	2001M04	3162088	2004M04	3952916	2007M04	4462705
1998M05	2637033	2001M05	3004257	2004M05	3766155	2007M05	4183592
1998M06	2579128	2001M06	3079253	2004M06	3816530	2007M06	4261529
1998M07	2632137	2001M07	3136408	2004M07	3883205	2007M07	4300542
1998M08	2830602	2001M08	3685891	2004M08	4675245	2007M08	4980564
1998M09	2924749	2001M09	3396831	2004M09	4121334	2007M09	4844204
1998M10	2831967	2001M10	3080612	2004M10	3823664	2007M10	4317981
1998M11	2744199	2001M11	3289316	2004M11	4006960	2007M11	4504220
1998M12	2800845	2001M12	3389336	2004M12	4099569	2007M12	4672524
1999M01	2897210	2002M01	3730599	2005M01	4499794		
1999M02	2973971	2002M02	3617588	2005M02	4468249		
1999M03	2857901	2002M03	3315747	2005M03	4061280		
1999M04	2804804	2002M04	3415416	2005M04	4178241		
1999M05	2773756	2002M05	3274314	2005M05	3894252		
1999M06	2707198	2002M06	3351252	2005M06	4108001		
1999M07	2754615	2002M07	3398289	2005M07	4141834		
1999M08	2964056	2002M08	3801275	2005M08	4666902		
1999M09	3134212	2002M09	3576592	2005M09	4297548		
1999M10	2949241	2002M10	3415671	2005M10	3999118		
1999M11	2856047	2002M11	3632047	2005M11	4127784		
1999M12	2907670	2002M12	3749509	2005M12	4430033		
2000M01	3065085	2003M01	4089067	2006M01	4778844		
2000M02	3191737	2003M02	4005757	2006M02	4621615		
2000M03	3054528	2003M03	3535938	2006M03	4134511		
2000M04	2976712	2003M04	3681920	2006M04	4349790		
2000M05	2916591	2003M05	3519326	2006M05	4027015		
2000M06	2906733	2003M06	3562875	2006M06	4151785		
2000M07	3096030	2003M07	3571995	2006M07	4217473		
2000M08	3474141	2003M08	3897428	2006M08	4840413		
2000M09	3309683	2003M09	3875935	2006M09	4575148		
2000M10	2950355	2003M10	3603975	2006M10	3984903		
2000M11	3053447	2003M11	3646259	2006M11	4312013		
2000M12	3107270	2003M12	3581607	2006M12	4528385		



〈그림 5-13〉 가정용 전력 판매량

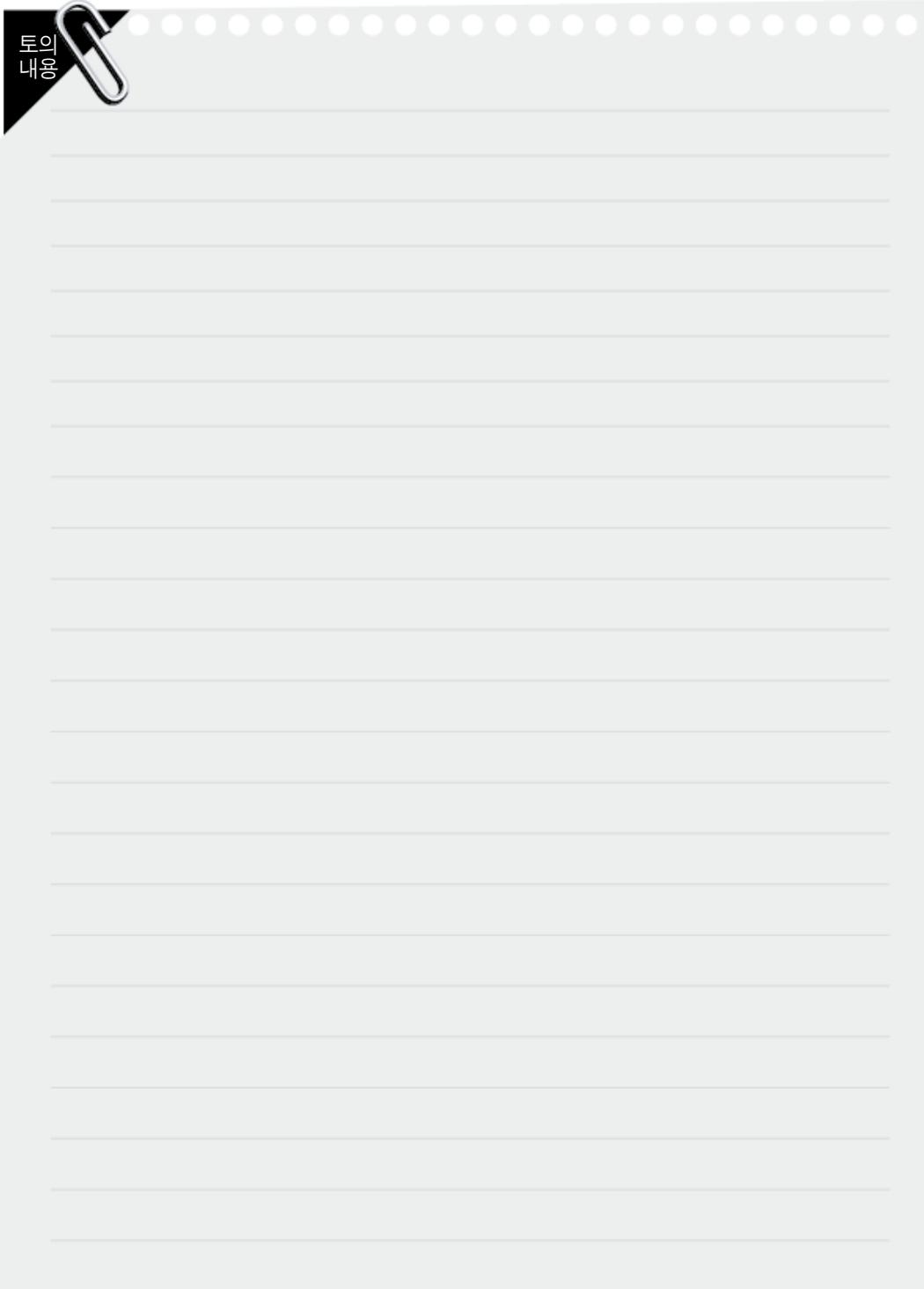
이 그래프를 월별로 나타내면 다음과 같다.



〈그림 5-14〉 월별 가정용 전력 판매량

2008년도 월별 판매량을 예측할 수 있는 아이디어를 토의하고, 만약 2008년도 총판매량이 60,000,000MWh이라면, 월별로 어떻게 예측할 수 있는지 아이디어를 토의해보자.

투의  
내용





## 5.6 알아두면 유용한 기법들

이 절에서는 종래의 통계학개론 수준의 교재에서 많이 다루고 있지는 않지만 정책과정에서 중요하게 활용될 수 있는 기법들을 활용될 수 있는 상황과 기본개념을 중심으로 소개하고자 한다.

### 5.6.1 주민 만족도 조사의 발전 방향 제시

#### 사례 1 만족도 조사문항에 가중치 주기

다음의 단순한 예를 통하여 필요한 상황을 이해해보자. 다음은 어느 지역 정책에 대한 주민의 만족도 조사문항과 4명의 응답자 A, B, C, D의 응답결과이다.

〈표 5-10〉 지역 정책에 대한 주민의 만족도 조사 문항

- |  |
|--|
| 1. 경제생활에 도움이 되십니까?<br>① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 아주 그렇다 |
| 2. 문화생활에 도움이 되십니까?<br>① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 아주 그렇다 |
| 3. 자녀교육에 도움이 되십니까?<br>① 전혀 그렇지 않다 ② 그렇지 않다 ③ 보통이다 ④ 그렇다 ⑤ 아주 그렇다 |

〈표 5-11〉 조사 자료(1)

	문항1 (경제생활)	문항2 (문화생활)	문항3 (자녀교육)	합계
A	1	3	5	9
B	3	3	3	9
C	5	3	1	9
D	2	5	2	9

4명의 총합은 모두 9로 평균 3점의 만족도를 보여주고 있다. 그런데 이러한 분석에서 경제생활, 문화생활, 자녀교육에 대한 만족도는 모두 동일하게 취급된다. 즉, 경제생활 만족도 5점과 문화생활 만족도의 5점을 같은 것으로 생각한다는 것이다. 그러나 조사목적에 따라서는 경제생활 만족도를 문화생활 만족도보다 더 중요하게 생각하고 자녀교육 만족도를 경제생활 만족도보다 더 중요하게 생각해야 할 경우도 있다. 만약 경제생활을 문화생활보다 2배 더 중요하게 생각하고, 자녀교육이 문화생활보다 4배, 경제생활보다 2배 더 중요하다고 생각한다면 위의 조사 자료는 다음과 같이 바뀐다.

〈표 5-12〉 조사 자료(2)

	2×문항1 (경제생활)	1×문항2 (문화생활)	4×문항3 (자녀교육)	합계
A	2	3	20	25
B	6	3	12	21
C	10	3	4	17
D	4	5	8	17

따라서 지역 정책에 대한 만족도는 응답자 중 A가 가장 높고, 그 다음이 B 그리고 C, D는 같은 값이 된다. 보다 현실적인 또 다른 예로, 우리가 음식점을 선택할 때 맛도 좋고, 가격도 좋고, 분위기도 좋은 곳을 찾지만 이렇게 세 가지를 모두 만족시키는 곳은 여간해서 찾기 어려운 법이다. 이러한 경우에 맛, 가격, 분위기라는 기준에서 어느 기준을 얼마나 더 중요하게 생각하느냐에 따라서 답이 결정된다. 이러한 상황에서 가장 상식적이고, 또 일반적으로 사용하는 방법이 가중치를 이용하는 것인데 가중치를 구하는 것은 쉬운 방법이 아니다. 그 이유는 우선 실제 복잡한 상황에서는 가중치를 적절히 정하기가 어렵고, 둘째는 그 가중치가 의사결정판단을 정확히 반영해 주는지에 대한 객관적 근거가 없기 때문이다. 즉, 의사결정자의

판단 자체가 오류가능성이 높고 일관성 없거나 시간에 따라 변할 수 있다는 것이다.

이러한 가중치 도출방법으로는 5점 척도, 7점 척도 등으로 알려진 척도표시법과 합계 100 점을 가지고 각 속성에 분배해 주되 보다 중요한 속성에 많은 점수를 부여하는 점수할당법, 그리고 각 대안의 속성 값들을 각 변수의 계수 값으로 보아 가중치를 묻는 것이 아니라 이러한 속성 값이라면 총체적으로 몇 점짜리일지 생각하기 때문에 총체적 판단법이라고도 불리는 다중회귀분석법, 두 속성 중 한 속성에서는 손해를 보고 다른 속성에서 이득을 보아 종합적으로 특실이 제로가 되게 하는(무차별적 교환) 방법 등과 함께 교환분석법, 그리고 Analytic Hierarchy Process라고 불리는 '계층적 분석 과정/방법'이라는 것이 있는데 이는 의사결정의 전 과정을 여러 단계로 나눈 후 단계별로 분석 해결하여 최적의 의사결정에 이르는 방법이다. AHP의 단계와 그 특징을 다음의 예를 통해서 알아보자.

### 사례 2 전공 선택에 미치는 요인의 가중치는?

X, Y, Z 세 가지의 속성이 있을 때 그 중요도가 X:Y가 2:1, X:Z가 4:1, Y:Z가 2:1이라면 당연히 X:Y:Z는 4:2:1이 될 것이다. 가격이나 무게와 같이 정량적인 수치화가 직접적으로 중요도로 바뀐다면 쌍별로 비교하든 전체를 한꺼번에 비교하든 그 결과는 정확하게 일치될 것이다.

그러나 인간의 주관적 판단이 개입되면 이야기가 전혀 달라진다. X(적성), Y(흥미), Z(의미)의 세 가지 속성을 따져서 전공분야를 선택하려 하는 청년이 있다고 하자. 이 청년이 만일 적성(X)대 흥미(Y)를 1:1, 적성(X)대 의미(Z)를 3:1, 흥미(Y)대 의미(Z)를 1:3으로 중요하다고 판단했다면 과연 X:Y:Z는 얼마일까? 그런데 이 청년은

적성 = 의미

적성 > 의미

흥미 < 의미

라는 일관성 없는 판단을 했기 때문에 답을 구할 수 없을 것이다. 이런 경우 판단이 논리적으로 틀렸으니 일관성 있게 다시 할 것을 요구하게 되는데 다시 한다고 해도 쉽지 않은 문제일 것이다. 이런 경우에 AHP를 활용하면 의사 결정하는 과정에서 의사결정자의 논리적 일관성이 유지되고 있는지 여부를 비일관성지수(Inconsistency Index)를 통하여 확인할 수 있게 된다.

## 5.6.2 자료 활용의 효용성 제고 방안

### 사례 3 경상남도 3개년 사회지표조사 자료 결합하기

경상남도에서 조사되고 있는 사회지표조사의 경우, 조사 항목이 많아 3년 주기로 매년 설문 문항을 다르게 하고 있기 때문에 환경 관련 분석을 할 때 2001년과 2002년 및 2003년 조사된 사회지표 조사의 환경자료를 각각 분석하였다. 박희창, 조광현(2008)은 2001년과 2002년 및 2003년 조사된 사회지표조사의 환경 관련 문항을 통합적으로 분석을 할 수 없는 문제점이 있다는 사실에 주목하여 2001년과 2002년 및 2003년의 사회지표조사를 결합하여 종합적으로 분석할 수 있는 방안을 연구한 바 있다. 그리하여 그들은 데이터 퓨전기법을 통하여 3개년 자료를 통합한 고부가가치 자료를 생성하고 추가된 정보를 이용함으로써 통계 분석의 질을 향상시켰다. 이들이 수행한 방법의 제 단계는 다음과 같이 요약된다.

#### (단계 1) 동질성 검정

데이터 퓨전에 앞서 2001년과 2002년 및 2003년에 조사된 사회지표 조사에 대한 조사 대상 모집단들의 분포 형태가 동일한지 동질성 검정을 실시한다. 동질성 검정 결과 조사 대상 모집단들의 분포가 동일한 경우 데이터 퓨전을 실시한다.

#### (단계 2) 데이터 퓨전

2001년과 2002년 및 2003년에 조사된 사회지표 조사 자료에서는 2001년과 2002년 자료의 데이터 퓨전, 2001년과 2003년 자료의 데이터 퓨전, 2002년과 2003년 자료의 데이터 퓨전 등 총 3번의 데이터 퓨전을 실시한다.

#### (단계 3) 데이터 통합

(단계 2)에서 실시된 총 3회의 데이터 퓨전 결과를 통합하여 최종 퓨전된 데이터 파일을 생성한다.

데이터 퓨전에 사용한 2001년, 2002년, 2003년 자료의 속성은 <표 5-13>, <표 5-14>, <표 5-15>와 같다. 각 자료는 약 10,000건이며 2001년 사회지표 조사는 환경관련 6문항과 인구통계학적 속성 6문항으로 구성되어 있고, 2002년 사회지표 조사는 환경관련 3문항과 인구통계학적 속성 6문항으로 구성되어 있으며, 2003년 사회지표 조사는 환경관련 2문항과 인구통계학적 속성 6문항만으로 구성되어 있다.

〈표 5-13〉 2001년 사회지표 조사 자료 구조

변수 구분	변수명	변수형
환경관련문항	주관적 상수도 환경오염도	연속형
	주관적 하수도 환경오염도	연속형
	주관적 소음진동 환경오염도	연속형
	주관적 악취 환경오염도	연속형
	주관적 대기 환경오염도	연속형
	주관적 토양 환경오염도	연속형
인구통계학적 속성 문항	연령	연속형
	주관적 사회계층	범주형
	학력	범주형
	성별	범주형
	결혼유무	범주형
	거주지역	범주형

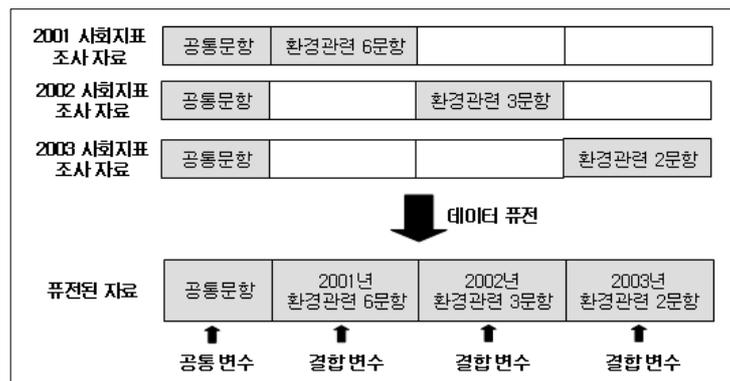
〈표 5-14〉 2002년 사회지표 조사 자료 구조

변수 구분	변수명	변수형
환경관련문항	쓰레기 분리수거의 참여 정도	연속형
	녹색 제품의 구입 여부	범주형
	수돗물 음용수 적정 여부	범주형
인구통계학적 속성 문항	연령	연속형
	주관적 사회계층	범주형
	학력	범주형
	성별	범주형
	결혼유무	범주형
	거주지역	범주형

〈표 5-15〉 2003년 사회지표 조사 자료 구조

변수 구분	변수명	변수형
환경관련문항	자녀의 환경오염 저감 행동 유무	범주형
	산림의 중요성 체감도	연속형
인구통계학적 속성 문항	연령	연속형
	주관적 사회계층	범주형
	학력	범주형
	성별	범주형
	결혼유무	범주형
	거주지역	범주형

2001년, 2002년, 2003년 조사된 사회지표 조사 자료에 대한 데이터 퓨전 방법은 <그림 15>와 같다.



〈그림 5-15〉 사회지표 조사의 데이터 퓨전

세 개의 조사 자료가 공통으로 갖고 있는 변수에는 개인 식별 가능한 변수가 없기 때문에 데이터 퓨전 시스템에 의하여 퓨전을 실시한다. 데이터 퓨전을 할 때, 공통 변수는 인구통계학적 속성 문항인 연령, 주관적 사회계층, 학력, 성별, 결혼유무, 거주지역이고 결합할 변수는 2001년, 2002년, 2003년 사회지표 조사에서의 환경관련 문항이 된다. 데이터 퓨전에 의한 최종 자료는 <표 5-16>과 같다.

〈표 5-16〉 데이터 퓨전에 의한 최종 자료

변수 구분	변수명	변수형
환경관련문항	지역의 상수도 환경오염도	연속형
	지역의 하수도 환경오염도	연속형
	지역의 소음진동 환경오염도	연속형
	지역의 악취 환경오염도	연속형
	지역의 대기 환경오염도	연속형
	지역의 토양 환경오염도	연속형
	쓰레기 분리수거의 참여 정도	연속형
	녹색 제품의 구입 여부	범주형
	수돗물 음용수 적정 여부	범주형
	자녀의 환경오염 저감 행동 유무	범주형
	산림의 중요성 체감도	연속형
인구통계학적 속성 문항	연령	연속형
	주관적 사회계층	범주형
	학력	범주형
	성별	범주형
	결혼유무	범주형
총 데이터 건수	거주지역	범주형
	29,876 건	

<표 5-16>에서 보는 바와 같이 데이터 퓨전에 의한 최종 자료의 레코드 수는 총 29,876 건이다. 변수는 인구통계학적 속성 문항인 연령, 주관적 사회계층, 학력, 성별, 결혼유무, 거주 지역의 6개 문항과 2001년의 환경관련 6문항, 2002년의 환경관련 3문항, 2003년 환경관련 2문항 등 총 17개 문항으로 구성되어 있다.

이들이 적용한 데이터 퓨전기법이 바로 자료결합으로 그 기본 개념은 다음과 같다 서로 다른 두개의 파일 A와 B에 몇 개의 변수는 공통적으로 있고 일부는 A에만 또는 B에만 있다고 가정하자. 즉, A와 B에 공통으로 있는 변수를 X라고 하고 A에만 있는 변수를 Y, B에만 있는 변수를 Z라고 하자. X, Y, Z 세 변수를 분석하고자 한다면 (X, Y, Z)로 구성된 파일을 만들기 위하여 A, B파일을 결합하여야 한다(<그림 5-16>).

좀 더 구체적으로 살펴보자. 개인이 자료를 구하는 경우에는 한계가 있기 마련이다. 이때 정부에서 발표한 조사 자료나 공개적으로 발표되는 자료 또는 인터넷에 공개되는 자료를 이용하면 개인이 가진 한계점을 극복할 수 있다. 즉, 개인이 수집한 자료를 공식 통계 자료(Public Use File)와 결합하면 개인의 자료로는 불가능한 통계 분석이 가능하거나, 어려운 통계 분석을 쉽게 수행할 수 있을 것이다. 이처럼 별개의 데이터 파일을 결합하여 하나의 완전한 데이터 파일을 만드는 방법을 ‘자료 결합(Data Matching)’이라고 한다. 특히 통계적 방법을 이용한 자료 결합을 ‘통계적 결합(Statistical Matching)’이라고 한다.

파일 A	X	Y	
파일 B	X		Z

결합파일	X	Y	Z
------	---	---	---

〈그림 5-16〉 자료 결합(Data Matching)

## 5.6.3 지역 특성의 시각화 방안

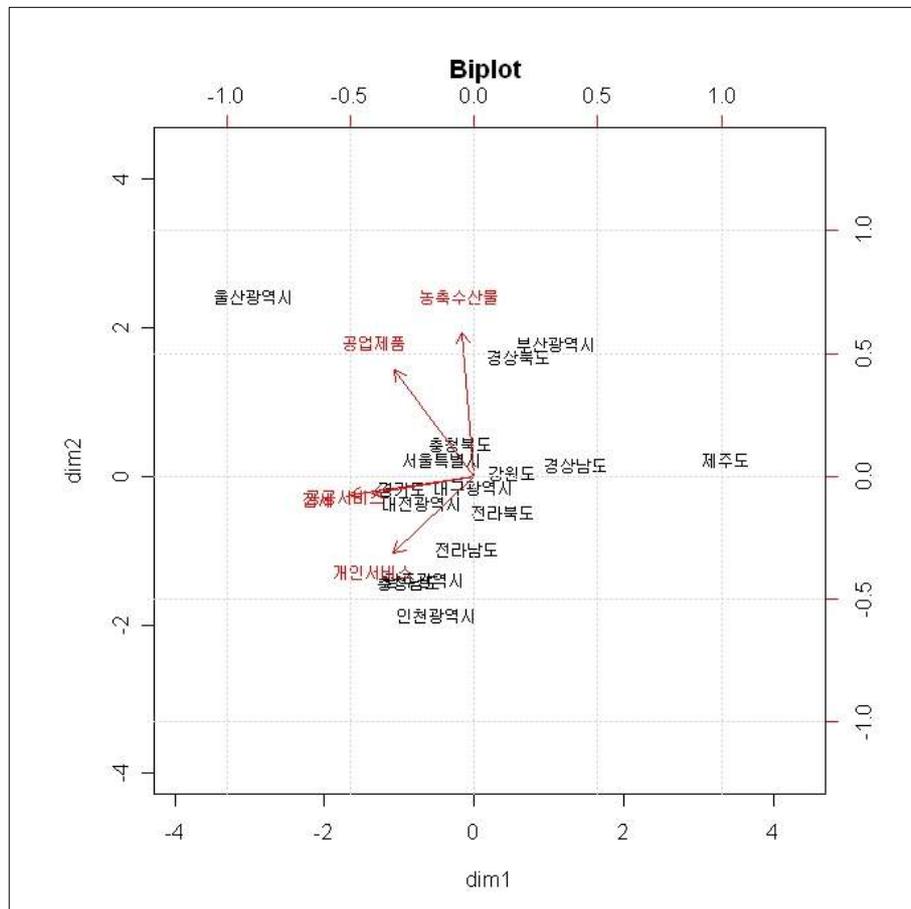
## 사례 4 16개 지역의 품목별 소비자물가지수 종합하기

<표 5-17>는 16개 지역의 2007년도 품목별 소비자물가지수를 나타낸다.

(표 5-17) 2007년도 품목별 소비자물가지수 및 주성분점수

도시	농축 수산물	공업 제품	집세	공공 서비스	개인 서비스	제1주성분 점수	제2주성분 점수
서울특별시	101.9	103.9	102.7	106.8	106.0	-0.42304	0.2127
부산광역시	103.2	104.2	99.9	106.0	105.9	1.09995	1.7796
대구광역시	101.7	103.7	102.3	107.0	105.9	-0.00552	-0.1460
인천광역시	98.6	104.0	101.4	107.1	106.7	-0.49450	-1.8685
광주광역시	100.3	103.8	102.4	106.7	106.8	-0.65859	-1.3977
대전광역시	101.2	104.2	100.8	106.8	107.1	-0.69905	-0.3599
울산광역시	103.8	104.9	103.8	107.2	106.5	-2.93967	2.4204
경기도	101.8	103.9	102.9	107.0	106.4	-0.95783	-0.1807
강원도	101.9	103.8	102.3	105.5	106.2	0.51906	0.0489
충청북도	103.4	103.9	101.8	105.7	107.0	-0.18156	0.4338
충청남도	101.8	103.6	103.4	105.6	107.5	-0.87466	-1.4237
전라북도	102.0	103.5	101.5	107.0	106.2	0.38925	-0.4909
전라남도	100.3	103.8	101.7	107.0	106.3	-0.08963	-0.9945
경상북도	103.7	104.0	101.0	106.3	105.9	0.59963	1.6028
경상남도	101.4	103.8	100.1	106.5	105.8	1.35835	0.1470
제주도	101.4	103.7	99.3	103.9	105.9	3.35781	0.2166

만약 지역 간 특징을 한 눈에 비교하면서 볼 수 있도록 시각화를 하는 것이 필요하다면 어떤 기법을 사용할 수 있을까? 가장 쉬운 방법은 16개 지역에서 각각의 평균 소비자물가지수, 즉 5개 품목의 소비자물가지수의 평균을 구하여 순위를 만드는 일일 것이다. 그런데 만약 누군가가 다음과 같은 그래프를 보여주면서 설명을 아래와 같이한다면 얼마나 유익할 것인가!



〈그림 5-17〉 Biplot

<그림 5-17>은 dim1은 전반적인 물가를 dim2는 농축산물 및 공업제품의 물가와 집세 및 서비스 분야의 물가 차이를 나타내는 것으로 왼쪽에 있는 울산광역시, 경기도 등은 물가가 높은 편이며, 반대로 오른쪽에 있는 제주도, 경상남도 등은 물가가 낮은 편임을 보여주고 그리고 위쪽에 위치한 울산광역시, 부산광역시, 경상북도 등은 농축수산물과 공업제품의 물가가 높은 편이며, 충청남도, 인천광역시, 광주광역시 등은 개인서비스분야의 물가가 높은 편임을 나타낸다.

이와 같은 상황에서 사용되는 대표적인 통계적 기법이 주성분 분석이라고 하는 기법으로 대부분의 통계소프트웨어를 사용하면 쉽게 적용할 수 있다. 추가적인 토의를 위하여 다음의 간단한 자료를 제시한다.

토의사례 1      **학업능력수준 종합하기**

이 자료는 10명의 학생으로부터 5개 과목의 점수를 얻고 이를 바탕으로 4장에서 설명한 표준점수를 <표 5-19>와 같이 구하여 과목별로 학생들의 상대적 위치를 파악하고 각자의 평균 표준점수를 구하였다. 이 자료로부터 학생들의 특징을 토의해 보자.

〈표 5-18〉 5개 과목의 성적

ID	국어( $x_1$ )	영어( $x_2$ )	제2외국어( $x_3$ )	수학( $x_4$ )	과학( $x_5$ )
1	3	33	73	8	12
2	3	30	59	28	20
3	35	83	91	32	34
4	35	83	85	33	32
5	15	40	55	68	52
6	3	53	76	10	8
7	68	83	85	48	50
8	15	47	77	76	76
9	46	60	83	83	68
10	98	83	91	80	72

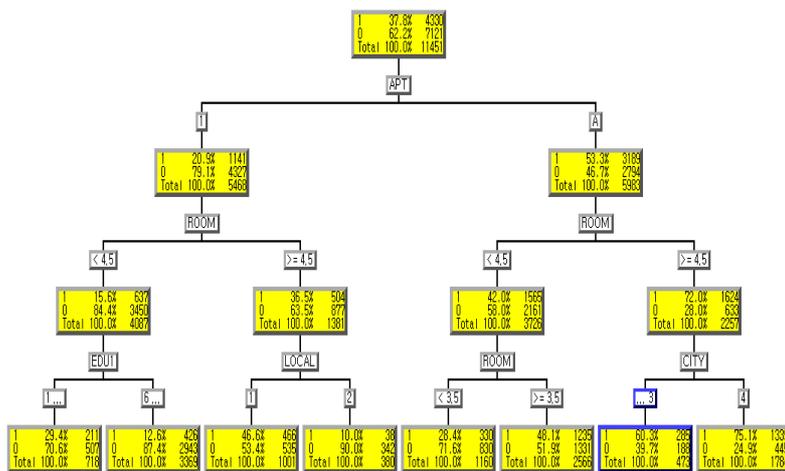
〈표 5-19〉 표준점수

ID	국어( $z_1$ )	영어( $z_2$ )	제2외국어( $z_3$ )	수학( $z_4$ )	과학( $z_5$ )	평균점수
1	-0.922	-1.204	-0.364	-1.352	-1.219	-1.012
2	-0.922	-1.340	-1.495	-0.651	-0.898	-1.061
3	0.092	1.068	1.091	-0.511	-0.337	0.280
4	0.092	1.068	0.606	-0.476	-0.417	0.174
5	-0.542	-0.886	-1.818	0.749	0.385	-0.422
6	-0.922	-0.295	-0.121	-1.282	-1.379	-0.800
7	1.137	1.068	0.606	0.049	0.305	0.633
8	-0.542	-0.568	-0.040	1.030	1.347	0.245
9	0.440	0.023	0.444	1.275	1.026	0.642
10	2.088	1.068	1.091	1.170	1.187	1.321
평균	0	0	0	0	0	0
분산	1.000	1.000	1.000	1.000	1.000	0.633

## 5.6.4 정책지지자 성향분석 방안

**사례 5** 어떤 가구들이 인터넷으로 응답하고 있는가?(어떤 사람들이 이 정책을 반대하고 있는가?)

5년마다 시행되는 인구주택총조사는 우리나라의 인구나 가구 및 주택에 관한 모든 사항을 조사하여 국가기본통계를 생산하는 조사이다. 그런데 인구주택총조사는 1인 가구의 증가, 생활 침해에 따른 응답거부, 조사비용의 증가 등의 문제를 안고 있어서 자기기입식 조사인 인터넷조사와 우편조사의 응답률을 사전에 예측하여 적정 수의 조사원을 배치함으로써 조사원의 인건비를 절감하는 비용문제를 해소하려는 노력을 하고 있다. 이러한 상황에서 인구주택총조사의 조사원 규모를 예측하는 근거로 사용하기 위하여 김혜경(2008) 등은 2010년 인구주택총조사에서 인터넷이나 우편으로 응답하는 가구가 어느 정도 되는지 측정하는 방안을 연구하였다. 그들은 이 연구에서 먼저 어떠한 가구들이 인터넷과 메일로 응답하는지 알아보기 위하여 시험조사에 응답한 가구 자료를 활용하여 의사결정나무라고 부르는 다음과 같은 나무형태를 구하였다.



〈그림 5-18〉 의사결정나무

<그림 5-18>을 설명하면 국내 가구 중에서 임의로 추출된 11,451가구를 대상으로 인터넷조사와 우편조사에 응답한 가구는 1로 표시하였고 응답하지 않은 가구는 0으로 표시하였다. 즉, 11,451 가구 중 37.8%인 4,330가구가 인터넷조사와 우편조사에 응답한 것을 알 수 있다. 표본으로 추출된 11,451가구 중에서 인터넷조사나 우편조사에 응답한 4,330가구와 인터넷조사나 우편조사에 응답하지 않은 7,121가구의 주요한 특징이 무엇인지를 알아보기 위하여 의사결정나무를 이용한다. 응답한 가구와 그렇지 않은 가구를 구분할 수 있는 기준들을 찾아서 그 기준에 근거하여 응답한 가구와 응답하지 않은 가구를 세분화하는 과정을 거치면 <그림 5-18>처럼 제일 하단에는 8개 그룹으로 나누어진다.

어떻게 하여 8개의 그룹으로 나누어졌는지를 살펴보자. 인터넷이나 우편조사에서 가구 응답률이 가장 낮은 그룹은 지역이 면이고 거처하는 집의 방 수가 5개 이상이며 아파트가 아닌 가구들로 380가구 중 38가구인 10%가 응답한 것으로 나타났다. 인터넷이나 우편조사에서 가구 응답률이 두 번째로 낮은 그룹은 거처하는 집의 방 수가 4개 이하이며 아파트가 아니고 가구주의 교육정도가 중졸 이상인 가구들로 3,369가구 중 426가구인 12.6%가 응답하는 것으로 나타났다. 응답률이 세 번째로 낮은 그룹은 거처하는 집의 방 수가 3개 이하인 아파트로 1,160가구 중 330가구인 28.4%가 응답하는 것으로 나타났다. 응답률이 네 번째로 낮은 그룹은 거처하는 집의 방 수가 4개 이하이며 아파트가 아니고 가구주의 교육정도가 초등학교 졸업 이하인 가구들로 718가구 중 211가구인 29.4%가 응답하는 것으로 나타났다. 응답률이 다섯 번째로 낮은 그룹은 지역이 동이고 거처하는 집의 방 수가 5개 이상, 아파트가 아닌 1,001가구 중 466가구인 46.6%가 응답하는 것으로 나타났다. 응답률이 세 번째로 높은 그룹은 거처하는 집의 방 수가 4개인 아파트로 2,566가구 중 1,235가구인 48.1%가 응답하는 것으로 나타났다. 응답률이 두 번째로 높은 그룹 지역이 중·소도시이상이며 거처하는 집의 방 수가 5개 이상인 아파트인 가구들로 473가구 중 285가구인 60.3%가 응답하는 것으로 나타났다. 인터넷이나 우편조사에서 가구 응답률이 가장 높은 그룹은 지역이 특별·광역시이며 거처하는 집의 방 수가 5개 이상인 아파트인 가구로서 1,784가구 중 1,339가구인 75.1%가 응답하는 것으로 나타났다.

위와 같이 만들어진 나무구조가 갖는 특성을 간략하게 직관적으로 설명한다면 다음과 같다.

〈표 5-20〉 예상 가구특성 변수

변 수	변 수 설 명
전수/표본	표본, 전수
방의 수	거처하는 집의 방 수
가구주 나이	가구주 나이
특성	아파트 여부
도시화정도	면, 중소도시, 특·광역시
지역구분	동, 면
교육정도1	미취학~대학원
교육정도2	졸업, 재학 등의 상태
가구원수	가구원수
타지주택보유여부	소유, 무소유
주인가구여부	주인, 대표가구, 세입가구
점유형태1	주거전용, 영업전용
점유형태2	자가, 세입 등 분류
건축연도	건축연도

첫 번째로 임의로 추출된 11,451가구를 가장 잘 구분해주는 변수들 중에서 거처형태, 즉 아파트인지 아닌지의 특성이 인터넷조사나 우편조사에 응답한 가구와 방문 응답가구를 가장 잘 설명하는 변수라는 사실이다. 즉, 아파트에 거주하는 가구들은 53.3%가 인터넷조사나 우편조사에 응답하고 아파트가 아닌 곳에 거주하는 가구들은 20.9%가 인터넷조사나 우편조사에 응답한다. 이와 같이 11,451가구를 두 집단으로 나누었을 때가 응답률이 가장 차이가 많이 난다는 것이다. 동일한 방식으로 아파트에 거주하는 5,983가구들은 거처하는 집의 방 수가 5이상인 가구들과 4이하인 가구들로 나누어서 볼 때 인터넷조사나 우편조사 응답률의 차이가 72%-42%=30%포인트로 커지게 된다. 다시 말하면 전체 11,451가구를 놓고 볼 때, 37.8%가 인터넷조사나 우편조사에 응답하였지만 거처하는 집의 방 수가 5이상인 아파트에 거주하는 가구들은 72%정도가 인터넷조사나 우편조사에 응답한다고 말할 수 있게 되어 인터넷조사나 우편조사에 응답하는 경향이 높은 가구의 특성을 알 수 있게 된다. 이런 기법은 여론조사결과를 분석할 때 많이 사용되는데 특별히 특정 정책을 지지하는 집단의 특성 같은 것을 찾고자 할 때 유용하다. 이 기법 역시 여러 통계 패키지에서 제공하고 있기 때문에 쉽게 활용할 수가 있다.

## ■ ■ 요약

- \_ 정규분포의 특징과 표준정규분포에 대해 알아보고 표준화 개념도 정리하였다
- \_ 신뢰수준과 오차의 한계가 무엇인지 통계적 정의와 개념을 살펴보았다.
- \_ 세 집단 이상의 평균을 비교하는 통계적 기법에 대해 알아보았다.
- \_ 관심 변수가 여러 개인 경우, 각 변수들의 연관성을 파악하기 위한 방법인 상관계수와 관계의 형태를 알아보려는 회귀분석의 용도와 목적을 토의하였다.
- \_ 시계열 자료를 식별하기 위한 변동의 분류 체계를 알아보고, 평활법 중에서 가장 일반적인 이동평균법을 토의하였다.
- \_ 정책과정에서 중요하게 활용될 수 있는 기법들인 AHP기법, 자료 결합, 주성분 분석, 의사결정나무 등에 대해 사례를 중심으로 토의하였다.

## [ 참고 문헌 ]

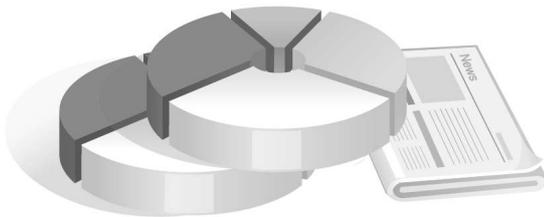
- (1) 감사원, 2007.06, 「감사결과 처분요구서 : 국가 주요통계 작성 및 활용 실태」, 감사원
- (2) 강선미외, 2007, 「국가통계활용매뉴얼」, 통계청 연구보고서
- (3) 국민은행, 2008, 「전국주택가격동향조사」, 국민은행
- (4) 국정브리핑 특별기획팀, 2007, 「대한민국 부동산 40년」, 한스미디어
- (5) 국정브리핑 특별기획팀, 2007, 「대한민국 교육 40년」, 한스미디어
- (6) 김기영, 전명식, 1989, 「SAS 주성분분석」, 자유아카데미
- (7) 김응환, 이석훈, 2007, 「통계교육」, 경문사
- (8) 김혜경, 이재원, 박희창, 한만규, 신지은, 이석훈, 2008, 「인구주택총조사에서 조사구별 자기 기입식 조사 응답 가구수 추정방안」, 한국자료분석학회 추계 발표 논문집 387-394
- (9) 노화준, 1996, 「정책학 원론」, 박영사.
- (10) 류지성, 2007, 「정책학」, 대영문화사.
- (11) 박원란 박희창 정지원 이석훈, 2008, 「1983년도 인구정책에 관한 소고」, 한국자료분석학회 2008년도 추계 발표 논문집 395-401
- (12) 박진우, 2005, 「통계학의 길잡이」, 교우사.
- (13) 박현정, 강성국, 2005, 「정책품질 향상을 위한 교육통계 혁신 방향」, 교육정책연구
- (14) 박희창, 조광현, 2008, 「통계조사에서의 퓨전된 자료에 대한 하이브리드 데이터마이닝의 적용방안」, 응용통계연구 21권 3호 399-412
- (15) 배득중, 2005, 「신재무행정: 정부혁신과 재정개혁」, 박영사.
- (16) 서울대학교통계연구소, 1990, 「전국 도시주택가격 동향조사 연구용역 최종보고서」
- (17) 이기재, 박진우, 박홍래, 1991, 「전국 도시주택가격동향조사를 위한 표본설계 연구」, 응용통계연구4권 2호 137-148.
- (18) 이부일, 신지은, 박영옥, 이석훈, 2008, 「엑셀을 활용한 통계자료분석 기초」, 경문사
- (19) 이석훈, 2006, 「통계적 사고방식」, 통계청
- (20) 장대홍, 1995, 「우리나라 언론매체에 나타나는 그래픽의 오용실태조사와 통계적, 제도적 해결방안에 대한 연구」, 응용통계연구 제8권 2호
- (21) 정정길, 2000, 「정책학 원론」, 박영사
- (22) 진미윤, 한수진, 2003, “주택보급률 지표의 한계점과 개선방안”, 대한건축학회논문집 계획계 19권 2호 133-140.
- (23) 질병관리본부, 2007, 「손상감시월보」
- (24) 질병관리본부, 2007, 「질병관리백서」
- (25) 차경엽, 정제련, 2006, 「효율적인 정책운을 위한 정부통계의 개선방안 연구」, 감사원 평가연구원.
- (26) 최종석, 박석윤, 이낙영, 박래현, 1995, 통계학 개론, 정익사
- (27) 통계청, 2007, 「경제활동인구통계 요약품질보고서」
- (28) 통계청, 2007, 「광공업동태조사에 대한 품질보고서」
- (29) 통계청, 2007, 「서비스업활동지수 품질요약보고서」
- (30) 통계청, 2008, 「2007 한국의 사회지표」, 통계청
- (31) 한국교육개발원, 2006, 「교육·인적자원 정책통계 인프라 구축 방안 연구」
- (32) 한국보건사회연구원, 2007, 「2007년도 OECD 보건생산통계」
- (33) 한국보건사회연구원, 2007, 「건강증진이 경제에 미치는 효과에 관한 연구」
- (34) 한국은행, 2006, 「알기쉬운 경제지표해설」, 한국은행
- (35) 한국주택은행 조사부, 1986, 「도시주택가격동향조사」
- (36) 한국행정연구원, 2006, 「정부혁신 성과지표 개발」

- 
- (37) 황윤원, 김성철, 2005, 「재무행정론」, 법문사
  - (38) Cabinet Office, 1999, 「Modernising Government」, White Paper.
  - (39) Davies, P. T., 2004, 「Is Evidence-based Government Possible?」, Jerry Lee Lecture, Washington D.C.
  - (40) Ken Young and Annette Boaz, 2005, 「Evidence & Policy」
  - (41) Morton, S. C., and Rolph, J. E., 2000, 「Public Policy and Statistics」, Springer
  - (42) 국가기록원 웹사이트(<http://www.archives.go.kr>)
  - (43) 국가통계포털 웹사이트(<http://www.kosis.kr>)
  - (44) 국민일보 웹사이트(<http://www.kukinews.com>)
  - (45) 내일신문 웹사이트(<http://www.naeil.com>)
  - (46) 네이버 카페(<http://cafe.naver.com/bloodbird>)
  - (47) 뉴시스 웹사이트(<http://www.newsis.com>)
  - (48) 동아일보 웹사이트(<http://www.donga.com>)
  - (49) 매일경제 웹사이트(<http://www.mk.co.kr>)
  - (50) 문화일보 웹사이트(<http://www.munhwa.com>)
  - (51) 손상감시 웹사이트(<http://injury.cdc.go.kr>)
  - (52) 인구주택총조사 웹사이트(<http://census.go.kr>)
  - (53) 중앙일보 웹사이트(<http://www.joins.com>)
  - (54) 조선일보 웹사이트(<http://www.chosun.com>)
  - (55) 통계메타DB시스템(<http://meta.nso.go.kr>)
  - (56) 통계청 웹사이트(<http://nso.go.kr>)
  - (57) 한겨레신문 웹사이트(<http://www.hani.co.kr>)
  - (58) e-나라지표 웹사이트(<http://www.index.go.kr>)



부록

# 통계기반 정책관리제도의 소개



## I. 통계기반 정책관리제도 개요

### 1. 도입배경 및 의의

지식정보사회로의 진전과 함께 정부, 기업, 가정 등 모든 사회주체들의 의사결정에 정확한 통계의 뒷받침이 필요하다는 공감대가 형성되고, 이에 따라 각계각층에서 통계 수요가 크게 증가하고 있다. 특히, 사회 각계에 영향을 미치는 정책을 수행하는 정부로서는 주요 정책의 도입 및 개선 시 이에 필요한 통계를 작성 또는 개선하려는 노력을 함께 기울여야 한다. 그럼에도 불구하고, 각 부처의 통계에 대한 관심부족으로 필요한 통계가 취약하고, 이는 정책 비효율과 국민들의 정책에 대한 불신으로 이어질 우려가 있다.

이러한 인식하에서 2007년 1월부터 통계청 주도하에 기획재정부, 보건복지가족부 등 13개 주요 통계작성기관이 함께 대책을 마련, 2007년 6월 사회통계발전을 위한 국가통계혁신계획을 국무회의에 보고, 확정하였으며, 동 계획의 핵심 과제로서 「통계적 정책전생애 관리제도(Statistical Policy Life-Cycle Assessment)」의 도입을 결정하였다.

제도의 궁극적인 취지는 통계와 정책의 연계를 강화하기 위한 것이다. 이를 위해서 각 부처가 법령을 제개정하는 경우, 이를 통해 도입되는 정책에 필요한 통계를 각 부처가 자체적으로 구비하도록 의무화하고, 통계청이 해당 정책의 집행·평가에 필요한 통계지표의 구비여부 및 통계 개선·개발계획의 타당성을 평가하고, 조정하도록 제도의 골자를 설계하였다.

동 제도의 운영절차는 일차적으로 법령 제개정 부처가 관련 정책의 집행·평가에 적합한 통계의 구비여부, 필요시 통계 개발·개선 계획 등을 검토한 평가 요청서를 작성, 통계청에 제출하고, 통계청은 이를 토대로 타당성을 검토, 조정하되, 부처와 통계청간 이견이 있는 경우는 해당 법령안의 국무회의 상정 시, 통계청 의견을 함께 제출한다.

이러한 통계기반 정책관리제도의 성격은 각 부처가 필요한 통계를 스스로 작성하는 분산형 통계 시스템의 골격은 유지하되, 중앙통계기관인 통계청이 정책에 필요한 통계지표의 개선·개발을 유도함으로써 정책부처를 지원하는 통계건설링 서비스라 할 수 있다.

2007년 10월 통계법 시행령을 개정, 제33조(정책통계기반평가의 절차 및 방법)와 제34조(정책통계기반평가를 위한 자문)에 근거를 마련하고, 2008년 1월 1일부터 시행하고 있다.

## 2. 평가대상 기관 및 법령

통계기반 정책관리의 평가 대상기관은 37개 중앙행정기관이다. 법령 제·개정에는 국회도 수행하지만 국회 입법절차에 행정부가 관여하는 것은 불가능하므로 제외하고 있다. 이러한 의원입법은 시행령을 평가함으로써 필요한 통계지표를 빠짐없이 갖추도록 하고 있다.

대상법령은 원칙적으로 모든 제·개정 법률 및 시행령(대통령령)이다. 다만, 법령의 내용이 명백히 통계지표가 불필요하거나 국가안보 등 대외적으로 공개하기가 어려운 법령은 통계법 시행령 제33조 제1항에 단서조항을 마련하여 제외한다. 한편, 총리령 등 시행규칙은 주로 절차적 사안을 규정하므로 일단 제외하였는바, 향후 제도 운영결과 및 성과를 보아가면서 포함 여부를 검토할 예정이다.

2008년 3월 현재 통계기반 정책관리의 대상 법령 수는 1,895개이다. 이는 총 법령 4,038개 가운데 제외법령 740개 및 시행규칙 1,403개를 제외한 것이다.

### 【제외 법령 기준】

- 헌법상 설치기관의 소관법령
- 국가안보, 행정절차, 행정조직, 민사·상사·형사, 소송절차, 재판, 행정 등에 관한 법령
- 국호·국기·연호, 상훈, 전례, 국경(정부기념)일, 의전 등에 관한 법령
- 행정기관 내부를 규율하는 사무 또는 업무분장, 수당, 감사, 통계, 문서 및 관인, 신분증, 복제에 관한 법령
- 법령시행(일), 서식 및 효력 등에 관한 법령
- 기타 통계청장이 통계기반 정책관리가 적절하지 아니하다고 판단한 법령

중앙행정기관별 소관법령 현황('08.3.기준)

기관명	총법령	관리법령	시행규칙	제외법령	심사대상
<b>합 계</b>	<b>4,038</b>	<b>2,635</b>	<b>1,403</b>	<b>740</b>	<b>1,895</b>
교육과학기술부	293	207	86	61	146
행정안전부	383	325	58	171	154
소방방재청	68	44	24	5	39
경찰청	64	44	20	14	30
법무부	341	253	88	253	0
국민권익위원회	4	4	0	4	0
법제처	9	5	4	4	1
통일부	23	17	6	14	3
외교통상부	77	48	29	19	29
국무총리실	31	10	21	4	6
병무청	7	5	2	3	2
국가보훈처	24	15	9	6	9
기타	46	38	8	20	18
문화체육관광부	140	100	40	12	88
문화재청	13	9	4	2	7
방송통신위원회	6	6	0	3	3
여성부	27	17	10	3	14
보건복지가족부	261	140	121	13	141
식품의약품안전청	17	16	1	1	15
환경부	133	92	41	11	81
국토해양부	374	213	161	12	201
해양경찰청	13	10	3	1	9
기상청	8	5	3	1	4
행복청	4	3	1	1	2
기획재정부	496	264	232	34	230
국세청	68	66	2	1	65
관세청	29	25	4	1	24
조달청	14	12	2	1	11
통계청	11	7	4	2	5
공정거래위원회	16	14	2	1	13
금융위원회	2	1	1	1	0
농림수산식품부	389	201	188	18	183
농촌진흥청	10	9	1	1	8
산림청	42	39	3	2	37
지식경제부	391	232	159	25	207
특허청	38	24	14	3	21
중기청	50	36	14	2	34
노동부	116	79	37	10	69

## II. 통계기반 정책관리의 평가 체계 및 절차

### 1. 평가 체계

동 제도의 일차적인 과제는 제·개정 법령에 필요한 통계를 검토하고, 미흡한 통계에 대한 개발·개선 계획을 수립하는 것이다. 이를 위해서는 제·개정 법령마다 심도 있는 검토가 필요하지만, 통계가 불필요한 법령 개정 사항에 대하여도 동일한 수준의 검토를 하는 것은 낭비이다. 이에 따라 개정 법령이 통계가 필요한지 여부를 먼저 검토하는 예비평가 제도를 운영하고 있다. 예비평가 결과에 따라 통계가 필요한 법령으로 판단되면, 통계 필요성, 구비여부, 개발계획 등에 대한 실질평가를 수행하게 된다.

우선, 예비평가는 제·개정되는 법령에 통계기반 정책관리가 필요한 정책과 제도가 포함되어 있는지 여부를 평가한다. 평가결과는 크게 ① 실질평가 면제 법령, ② 평가 제외 법령, ③ 실질평가 대상, ④ 향후 시행령 입안 시 평가법령으로 구분된다.

통계기반 정책관리 실질평가 면제법령은 관련법령 전체적으로는 통계적 기반이 필요한 정책이나 제도가 포함되어 있으나 개정되는 법령에 한해서 통계적 기반이 필요없는 경우이다.

통계기반 정책관리 제외법령은 법령전체가 절차, 조직, 서식이나 효력 등만을 규정하고 있어 통계적 기반이 필요없는 경우이다(국립대학치과대학설치법령, 한센인 피해사건 진상규명 및 피해자 생활지원 등에 관한 법령 등).

실질평가 대상법령은 소관부처는 통계가 필요없다는 의견을 제시하였으나 제·개정 법령을 통해 도입·변경되는 정책과 제도가 통계에 기반한 정책관리를 할 필요성이 있는 경우이다. 법령소관기관은 다시 실질평가를 요청하여야 한다.

향후 시행령 입안 시 평가법령은 통계는 필요하지만, 법률이 아직 구체적이지 않아서 검토가 곤란하므로 향후 그 내용이 구체화되는 시행령 개정 시 검토하겠다는 의미이다.

다음으로 실질평가는 제개정 법령이 통계가 필요한 법령이 그 대상이다. 실제에 있어서는 예비평가 결과에 따라 실질평가 대상법령으로 결정된 법령 또는 각 부처가 예비평가 요청 없이 바로 실질평가를 요청한 법령이 대상이 된다.

평가의견 유형에는 ① 평가요청 제외법령, ② 원안동의, ③ 사용권고, ④ 개발·개선권고가 있다.

평가요청 제외법령은 법령소관기관이 예비평가의 요청을 생략하고 실질평가를 요청한 법령으로서, 통계청 평가결과 통계가 필요치 않은 경우이다.

원안동의를 법령소관기관이 해당 정책이나 제도의 집행·평가를 위해 제시한 통계지표 및 통계개발·개선계획에 대해서 통계청의 이견이 없는 경우이다. 이는 당초 이견이 있었으나 평가과정에서 소관기관과 합의가 된 경우도 포함한다.

사용권고는 법령소관기관이 해당 정책이나 제도의 집행·평가를 위해 제시한 통계지표보다 더 적합한 통계지표가 있어 이의 사용을 권고한 경우이다.

개발·개선권고는 법령소관기관이 해당 정책이나 제도의 집행·평가를 위해 제시한 통계지표나 통계개발·개선계획이 부적절하여 이에 대한 수정·보완을 협의하였으나 합의가 되지 않은 경우이다. 개발·개선권고인 경우 법령소관기관은 법령안을 국무회의에 상정 시 통계청 의견을 첨부하여야 한다.

## 2. 평가 절차 및 방식

법령을 제·개정하는 중앙행정기관은 전문적인 예비평가 또는 실질평가 요청서를 작성하여 통계청에 평가를 요청하여야 한다. 한편 법령소관 기관과 해당 정책을 추진하는 중앙행정기관이 다른 경우는 정책 소관기관이 평가요청을 하여야 한다.

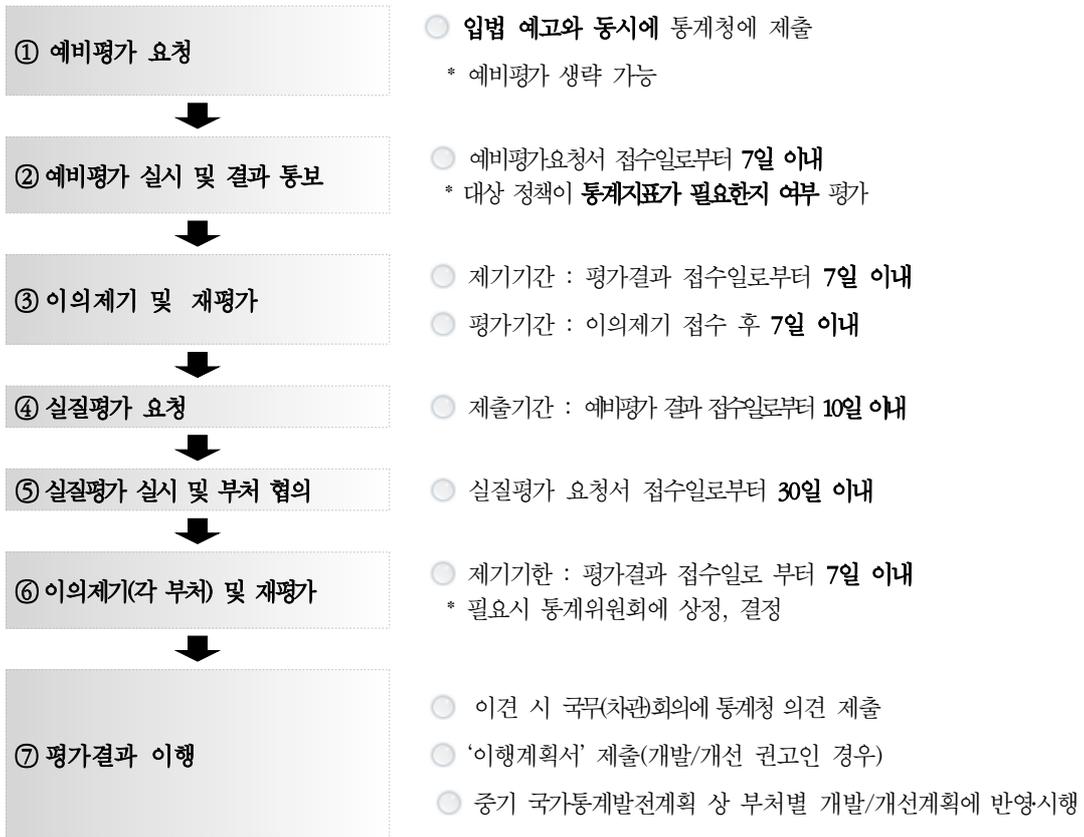
평가요청을 할 때 평가요청 기관의 담당부서는 해당기관 통계책임관의 검토를 받아 평가요청서를 제출하여야 한다.

제출 서류는 평가요청서(예비평가요청서 혹은 실질평가요청서), 법령 전문 및 제개정 내용, 산구조문 대비표, 제개정 설명자료 등이다. 필요 시 통계청의 담당직원이 평가를 위해 필요한 타 관련자료를 대상기관에 요구할 수 있다.

평가요청시기는 제개정 법령안에 대한 관계기관 협의 시 또는 입법예고 후 즉시이다. 한편, 예비평가를 요청하여 실질평가 대상으로 결정된 경우는 예비평가결과서를 통보받은 날로부터 10일 이내에 실질평가요청을 해야한다.

중앙행정기관은 통계청의 예비평가나 실질평가의 결과에 대해 이의가 있는 경우 이의제기를 할 수 있으며, 이 때 통계청은 전문성과 객관성을 확보하기 위하여 전문가에게 자문을 의뢰하거나 통계위원회에 상정하여 처리할 수 있다.

【평가 절차 흐름도】



### III. 통계기반 정책관리 평가 사항

#### 1. 예비평가

예비평가는 다음 2가지 사항을 평가한다.

- ① 법령의 제·개정을 통하여 도입·변경되는 정책·제도는 무엇인가?
- ② 각 정책·제도의 집행·평가가 통계에 기반할 필요성이 있는가?

상기의 두 가지 사항 등을 검토하여 해당 제·개정 법령이 제외법령인지, 실질평가면제 법령인지, 시행령 입안 시 평가 법령인지, 실질평가 대상법령인지를 판단한다.

#### 2. 실질평가

실질평가는 통계지표, 통계 개발·개선 계획으로 구분하여 평가한다.

우선, 통계지표에 대하여는 5개의 사항을 검토하는데, 수치성, 객관성, 일반성, 비교가능성의 기본요건과, 정책이나 제도에 관련해서 포괄성, 정책관련성, 신뢰성 등의 요건을 갖추고 있는지를 검토한다.

- ① 정책·제도별로 통계지표가 모두 제시되었는가?
- ② 제시된 통계지표만으로 정책·제도의 집행·평가가 충분한가?
- ③ 통계지표가 부적절한 경우, 대체할만한 다른 통계지표는 있는가?
- ④ 제시된 통계지표가 기존 통계조사 또는 행정보고를 통해 작성 가능한가?
- ⑤ 통계지표가 행정보고를 통해 작성되는 경우, 안정적이고 신뢰성있게 작성할 수 있는 체계가 갖추어져 있는가?

통계 개발·개선 관련하여는 각 부처가 작성하여 제출하는 요청서의 유형에 따라서 두 가지로 경우에 따라 검토하는데, 각각의 경우마다 개발·개선 계획의 조사목적, 조사대상, 조사항목, 공표일정 및 주기의 타당성에 대하여 검토한다.

[중앙행정기관이 통계 개선·개발 계획을 제시한 경우]

- ① 통계 개발·개선계획을 통해 필요한 모든 통계지표의 작성이 가능한가?
- ② 통계 개발·개선계획은 적절한가?
  - 조사목적, 조사대상, 주요 조사사항, 작성주기 등
- ③ 개발·개선되는 통계지표의 공표시기는 적절한가?
- ④ 필요한 통계지표를 생산하고 있는 기존의 통계조사와 중복되지는 않는가?
- ⑤ 개발·개선계획이 타 기관에서 이루어져야 하는 경우 해당기관과 충분히 협의되었는가?

[중앙행정기관이 통계 개선·개발 계획을 제시하지 않은 경우]

- ① 상기 평가기준을 토대로 개발·개선계획을 마련, 해당기관과 협의하여 개발·개선계획이 합의되었는가?

## IV. 평가결과에 대한 후속조치 및 향후과제

### 1. 평가결과에 대한 이행확보 및 후속조치

동 제도상 평가과정에서 통계를 개발·개선키로 한 경우 각 부처는 이를 이행해야 한다.

특히 개발·개선 권고 법령에 대해서 각 부처는 개발·개선 권고 의견을 접수한 날로부터 30일 이내 이행계획서를 제출해야 한다(통계법 제12조 통계의 작성·보급에 관한 사무에 대한 개선요구 등, 제21조 통계의 작성 권고, 동법 시행령 제16조 통계의 작성·보급에 관한 사무에 대한 개선 등의 이행).

아울러 개발·개선 권고에 따른 통계 개발·개선 계획을 포함하여 통계청과 부처간 협의한 통계 개발·개선 계획 가운데 국가승인통계로 포함되는 것이 타당한 계획은 중기국가통계발전계획(매년 보완하는 연동계획으로써 통계청 주도하에 모든 중앙행정기관이 함께 참여하여 작성, 반기별로 이행실적을 점검함)상 각 기관별 개발·개선 계획에 반영하여 반기별로 점검한다.

### 2. 향후과제

통계기반 정책관리제도의 조기 정착을 위해 몇 가지 풀어야 할 과제가 있다.

첫째는 대상 중앙행정기관을 대상으로 동 제도운영에 대한 폭넓은 홍보를 통해 적극적인 참여를 유도하는 일이다. 2008년은 제도시행 원년으로 대상 중앙행정기관의 이해도가 낮아 동 제도에 대한 참여가 미흡할 것으로 예상된다. 기관회의나 통계책임관제도 등을 활용한 고위급 대상의 홍보활동, 업무관련자를 대상으로 한 설명회 등을 통해 동 제도의 취지를 충분히 이해시켜 각 기관의 적극적 참여를 이끌어내야 한다.

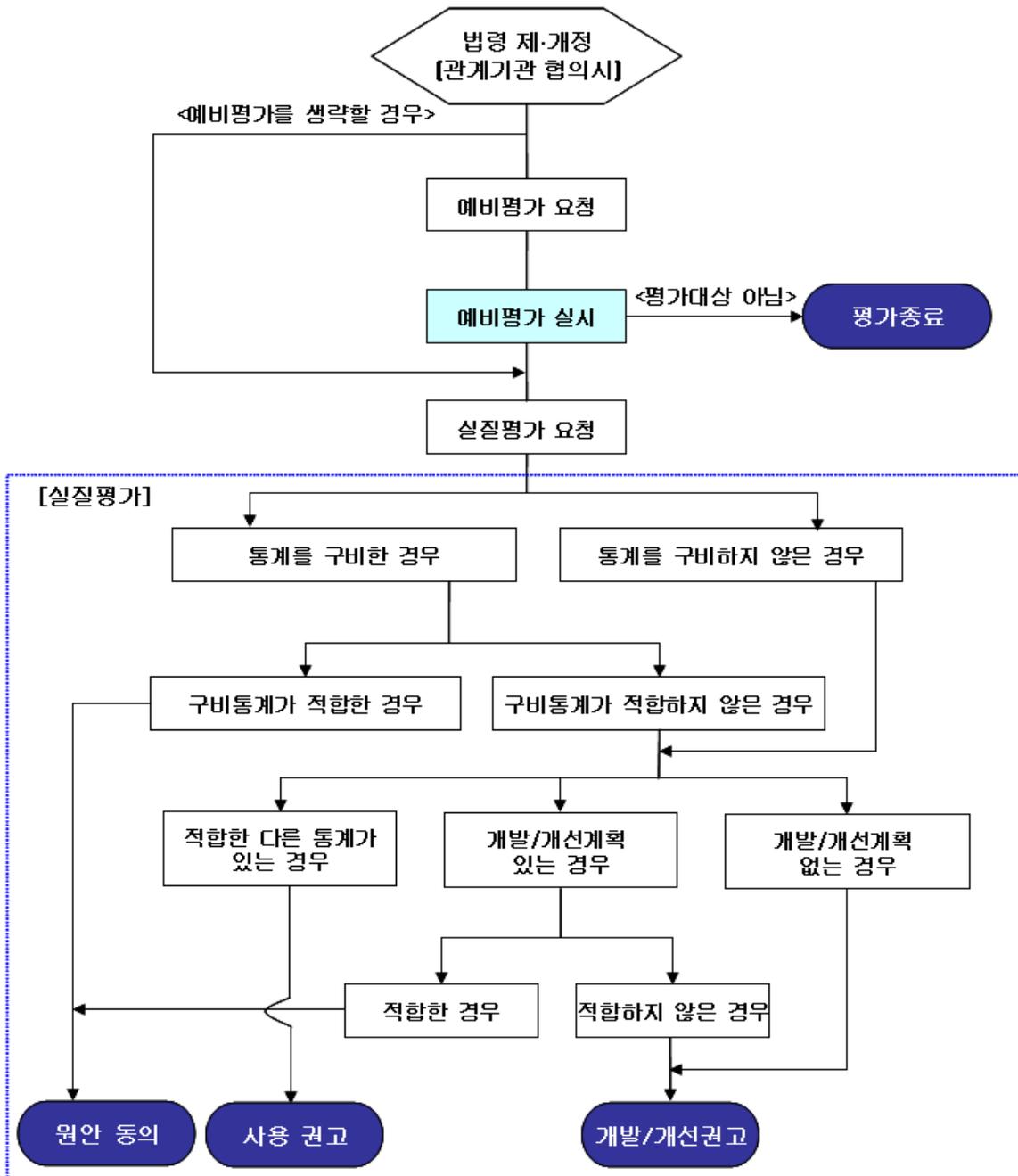
둘째는 통계기반 정책관리제도의 기반을 구축하는 일이다. 알기 쉽고 일관성을 갖춘 운영지침 및 사례중심 운영매뉴얼을 작성·보급하여 동 제도를 처음 접하는 업무관련자도 쉽게 수행할 수 있도록 여건을 조성해야 한다.

---

셋째는 평가업무 담당인력의 역량강화이다. 중앙행정기관이 입안하는 법령에서 해당 정책을 파악하고 이에 필요한 통계지표를 검토하는 것은 매우 어려운 일이다. 따라서 관련 정책이나 제도에 대한 자료를 수집·분석하고 연구활동 등을 통해 정보를 공유하며 전문가의 도움을 받는 등, 전문성을 배양하는 일은 무엇보다 중요하다.

이외에도 정책 분야별 전문 자문인력의 발굴, 외국의 유사한 사례수집·연구 등의 검토도 이루어져야 할 것이다.

## V. 평가단계별 업무흐름도



## 〈참고자료 : 식품위생법 시행령 일부개정령안〉

### 1. 개정이유

소비자 알권리를 보장하고 유통질서를 확립하기 위하여 음식점에서 조리·판매하는 식품의 원산지 표시대상 영업자의 범위 및 표시 대상 품목을 확대하고, 식품관리의 안전성을 제고하기 위하여 식품의 제조·가공단계부터 판매단계까지의 각 단계별 정보를 기록·관리하는 식품이력추적관리제도를 도입한 「식품위생법」이 개정(법률 제8779호, 2007. 12. 21 공포, 2008. 6. 22 시행)됨에 따라 이 법에서 대통령령으로 위임된 원산지 표시대상 영업자 및 품목의 범위와 원산지 표시기준 및 식품이력추적관리 신고의무 위반자에 대한 과태료 기준 등 필요한 사항을 정하려는 것임.

### 2. 주요내용

가. 원산지 표시대상 영업자와 쌀·김치류의 원산지 및 육류의 원산지등 표시 대상 품목의 범위를 정함(안 제2조의2)

(1) 원산지 등을 표시하여야 하는 영업자를 식품접객업 중 일반음식점영업 신고를 한 자 중 영업장의 면적이 100제곱미터 이상인 자로 확대함

(2) 원산지등을 표시하여야 하는 품목의 범위를 정함

(가) 쌀 : 쌀(기타 곡류를 혼합한 경우 포함)의 원형을 유지하여 조리한 후 밥류(떡류, 죽류, 면류 및 식혜 제외)로 제공하는 것

(나) 김치 : 배추를 주원료로 하여 절임, 양념혼합과정 등을 거쳐 그대로 또는 발효시킨 것이거나 이를 가공한 것으로서 식사류와 함께 반찬으로 제공되는 김치

(다) 쇠고기 : 쇠고기를 구이용, 탕용, 찜용, 튀김용 및 생식용(육회)으로 조리하여 판매하는 것

(라) 돼지고기 : 돼지고기를 구이용, 탕용, 찜용 및 튀김용으로 조리하여 판매하는 것

(마) 닭고기 : 닭고기를 구이용, 탕용, 찜용 및 튀김용으로 조리하여 판매하는 것

(3) 원산지 등 표시대상 영업자 및 품목을 확대함으로써 소비자의 알권리를 강화하고 유통질서 확립에 기여

나. 음식점에서 쌀·김치류 및 육류의 원산지 등 표시를 하지 아니한 영업자에 대하여 과태료 부과(안 별표 2 제1의2호)

(1) 쌀, 배추김치, 돼지고기 및 닭고기의 원산지를 표시하지 아니한 자는 100만원의 과태료를 부과하도록 함

(2) 원산지 표시 위반자에 대하여 과태료를 부과하여 처분의 실효성 확보로 동 제도의 조기정착에 기여

다. 식품이력추적관리 변경된 등록사항 미신고시 과태료부과(안 별표 2 제10의2호 신설)

(1) 식품이력추적관리 등록은 영업자 자율이나, 이미 등록한 자에게는 그 등록한 사항의 변경 사항을 신고할 의무를 부여하고 이를 지키지 아니한 경우 과태료 부과

(2) 식품이력추적관리 등록사항이 변경된 경우 1개월 이내에 신고하지 아니한 영업자에 대하여 30만원 과태료 부과

(3) 식품이력추적관리 등록 사항을 변경할 경우 신속정확하게 신고하도록 함으로써 식품이력 추적관리 제도의 효율적 운영에 기여

### 3. 참고사항

가. 관계법령 : 생략

나. 예산조치 : 별도조치 필요 없음

다. 합 의 :

라. 기 타 : 신·구조문대비표, 별첨

대통령령 제 호

식품위생법 시행령 일부개정령안

식품위생법 시행령 일부를 다음과 같이 개정한다.

제2조의2를 다음과 같이 한다.

제2조의2(쌀김치류의 원산지 및 육류의 원산지등 표시) ①법 제10조의3제1항에서 “대통령령이 정하는 영업을 영위하는 자”라 함은 제7조제8호나목에 따른 일반음식점영업을 영위하는 자를 말한다.

②법 제10조의3제1항제1호에서 “대통령령으로 정하는 것”이라 함은 쌀(기타 곡류를 혼합한 경우 포함)의 원형을 유지하여 조리한 후 밥류(떡류, 죽류, 면류 및 식혜 제외)로 제공하는 것을 말한다.

③법 제10조의3제1항제2호에서 “대통령령으로 정하는 것”이라 함은 배추를 주원료로 하여 절임, 양념혼합과정 등을 거쳐 그대로 또는 발효시킨 것이거나 이를 가공한 것으로서 식사류와 함께 반찬으로 제공되는 김치(이하 “배추김치”라고 한다)를 말한다.

④법 제10조의3제1항제3호에서 “대통령령으로 정하는 것”이라 함은 다음 각 호와 같다.

1. 쇠고기 : 쇠고기를 구이용, 탕용, 찜용, 튀김용 및 생식용(육회)으로 조리하여 판매하는 것
2. 돼지고기 : 돼지고기를 구이용, 탕용, 찜용 및 튀김용으로 조리하여 판매하는 것
3. 닭고기 : 닭고기를 구이용, 탕용, 찜용 및 튀김용으로 조리하여 판매하는 것

제13조의2제7호중 “「지방자치법」 제3조제3항 또는 제4항의 규정에 의한 동일 읍·면·동”을 “동일 시(제주특별자치도의 경우에는 행정시를 말한다)·군·구(자치구를 말한다)”로 한다.

별표2 제1호사목중 “식품접객조리판매등에 대한 미생물권장규격”을 “식품접객업소의 조리판매 등에 대한 기준 및 규격”으로 한다.

별표2 제1의3호를 다음과 같이 한다.

1의3. 법 제10조의3에 위반하여 원산지등을 표시하지 아니하고 식품을 조리판매한 영업자	법 제78조제1항제2호	500만원
가. 쇠고기의 원산지 및 종류 전부를 표시하지 아니한 자		300만원
나. 쇠고기의 종류를 표시하고, 원산지를 표시하지 아니한 자		100만원
다. 쇠고기의 원산지를 표시하고, 종류를 표시하지 아니한 자		100만원
라. 쌀의 원산지를 표시하지 아니한 자		100만원
마. 배추김치의 원산지를 표시하지 아니한 자		100만원
바. 돼지고기의 원산지를 표시하지 아니한 자		100만원
사. 닭고기의 원산지를 표시하지 아니한 자		100만원

별표2 제10의2호를 제10의3호로 하고, 같은 표에 제10의2호를 다음과 같이 신설한다.

10의2. 법 제32조의3제3항을 위반하여 식품이력추적관리의 등록사항이 변경된 경우 1개월 이내에 신고하지 아니한 영업자	법 제78조제3항	30만원
---	-----------	------

#### 부 칙

① (시행일) 이 영은 2008년 6월 22일부터 시행한다. 다만, 제2조의2제3항, 같은 조 제4항제2호 및 제3호, 별표 2 제1의3호사목부터 사목까지의 개정규정은 2008년 12월 22일부터 시행한다.

② (과태료 처분기준에 관한 경과조치) 이 영 시행 전의 위반행위에 대한 과태료 처분기준의 적용에 있어서는 종전의 규정에 따른다.

신·구조문대비표

현 행	개 정 안
<p>제2조의2(원산지 표시 대상 영업자) ①법 제10조의3제1항에서 "대통령령으로 정하는 영업자"라 함은 제7조제8호 나목에 따른 일반음식점영업 중 영업장의 면적(영업신고서에 기재된 면적을 말한다)이 300제곱미터 이상인 영업자를 말한다.</p> <p>②법 제10조의3제2항에 따라 원산지등을 표시하여야 하는 식육은 쇠고기로 하되, 쇠고기의 생육 또는 양념육을 구입용으로 조리하여 판매하는 경우에 한한다.</p> <p>〈신 설〉</p> <p>〈신 설〉</p>	<p>제2조의2(쌀·김치류의 원산지 및 육류의 원산지등 표시) ①법 제10조의3제1항에서 "대통령령이 정하는 영업"을 영위하는 자"라 함은 제7조제8호나목에 따른 일반음식점영업을 영위하는 자를 말한다.</p> <p>②법 제10조의3제1항제1호에서 "대통령령으로 정하는 것"이라 함은 쌀(기타 곡류를 혼합한 경우 포함)의 원형을 유지하여 조리한 후 밥류(떡류, 죽류, 면류 및 식혜 제외)로 제공하는 것을 말한다.</p> <p>③법 제10조의3제1항제2호에서 "대통령령으로 정하는 것"이라 함은 배추를 주원료로 하여 절임, 양념혼합과정 등을 거쳐 그대로 또는 발효시킨 것이거나 이를 가공한 것으로서 식사료와 함께 반찬으로 제공되는 김치(이하 "배추김치"라고 한다)를 말한다.</p> <p>④법 제10조의3제1항제3호에서 "대통령령으로 정하는 것"이라 함은 다음 각 호와 같다.</p> <p>1. 쇠고기 : 쇠고기를 구입용, 탕용, 찜용, 튀김용 및 생식용(육회)으로 조리하여 판매하는 것</p> <p>2. 돼지고기 : 돼지고기를 구입용, 탕용, 찜용 및 튀김용으로 조리하여 판매하는 것</p> <p>3. 닭고기 : 닭고기를 구입용, 탕용, 찜용 및 튀김용으로 조리하여 판매하는 것</p>
<p>제13조의2(신고를 하여야 하는 변경사항) 법 제22조제5항 후단의 규정에 의하여 신고를 하여야 하는 변경사항은 다음 각호와 같다.</p> <p>1. ~ 6. (생략)</p> <p>7. 「지방자치법」 제3조제3항 또는 제4항의 규정에 의한 동일 읍·면·동에서 식품자동판매기의 설치대수를 증감하고자 하는 경우(제7조제5호나목(5)의 규정에 의한 식품자동판매기영업을 하는 자에 한한다)</p>	<p>제13조의2(신고를 하여야 하는 변경사항) -----</p> <p>-----</p> <p>1. ~ 6. (현행과 같음)</p> <p>7. 동일 시(제주특별자치도의 경우에는 행정시를 말한다)·군·구(자치구를 말한다)-----</p> <p>-----</p>

## VI. 통계기반 정책관리 사례

### 1 평가요청서 및 참고자료 접수

〈평가요청서〉

※관리번호	통계기반 정책관리 실질평가 요청서			
법령명	식품위생법 시행령 일부 개정안			
관련법령명	식품위생법 및 같은 법 시행규칙			
구분	제정		개정	
	법률	대통령령	법률	대통령령
				○
정책(제도)명	식품 원산지 표시제도 및 식품이력 추적관리제도			
소관기관	기관명	보건복지가족부		
	총괄부서 (검토부서)	부서명(과 혹은 팀명)	통계담당관	
		담당자 작성명/ 전화번호	주무관 ○○○/02-2023-7167	
	사업부서 (집행부서)	부서명(과 혹은 팀명)	식품정책과	
담당자 작성명/ 전화번호		행정사무관 ○○○/02-2023-7785		
입법일정(예정)	관계기관 협의	대상기관	전부처	
		협의기간	2008. 4. 8.부터 2008. 4. 11.까지(4일간)	
	입법예고	2008. 4. 15.부터 2008. 5. 5.까지(21일간)		
첨부자료	필수자료	법령안 전문 및 개정내용, 신구조문 대조표, 통계개발/개선계획		
	기타자료			
실질평가에 대한 법령 소관기관 의견	이번 개정의 식품 원산지 표시제도 확대 및 식품이력추적제도 도입에 관련된 사항은 통계를 기반하여 관리할 필요성이 있다고 판단됨			

<p>정책(제도)의 주요내용</p>	<ul style="list-style-type: none"> <li>□ 정책(제도) 목적 및 필요성             <ul style="list-style-type: none"> <li>○ 원산지 표시에 따른 소비자의 선택권과 알 권리를 보장하여 부정확한 원산지 정보에 따른 소비자 피해 방지 및 정당한 영업자의 불이익 방지</li> <li>○ 식품안전사고 발생시 해당 식품을 추적하여 신속한 원인규명 및 회수 등 식품안전을 확보할 수 있도록 식품이력추적관리제도 도입 필요성 대두</li> </ul> </li> <li>□ 정책내용             <ul style="list-style-type: none"> <li>○ 의의                 <ul style="list-style-type: none"> <li>- 원산지 표시 의무화 확대와 식품이력추적관리제도는 소비자에 대한 올바른 정보 제공과 유통질서 확립, 식품관리의 안전성을 제고할 수 있음</li> </ul> </li> <li>○ 도입근거                 <ul style="list-style-type: none"> <li>- 「식품위생법」이 개정(법률 제8779호, 2007. 12. 21 공포, 2008. 6. 22 시행)됨에 따라 이 법에서 대통령령으로 위임된 원산지 표시대상 영업자 및 품목의 범위와 원산지 표시기준, 식품이력추적관리제도 도입에 관하여 규정</li> </ul> </li> <li>○ 추진방향                 <ul style="list-style-type: none"> <li>- 원산지 표시 대상 업소를 일반음식점 영업장의 면적이 100제곱미터 이상인 자로 확대하고, 표시대상 식품도 쇠고기 뿐만 아니라 쌀, 배추김치, 돼지고기, 닭고기 로 확대함</li> <li>- 식품을 제조·가공단계부터 판매단계까지 각 단계별로 정보를 기록·관리하여 해당 식품의 안전성 등에 문제가 발생할 경우 해당 식품을 추적하여 원인규명 및 필요한 조치를 할 수 있도록 관리</li> </ul> </li> <li>○ 세부내용                 <ul style="list-style-type: none"> <li><b>1) 식육 원산지 표시제도</b> <ul style="list-style-type: none"> <li>- 표시의무대상 : 쌀, 배추김치*, 쇠고기, 돼지고기, 닭고기                             <ul style="list-style-type: none"> <li>* 수입한 배추 : 원산지 표시, 제조가공한 배추김치 : 수입국 표시</li> </ul> </li> <li>- 시설규모 : 100㎡ 이상 일반음식점</li> <li>- 의무품목의 범위                             <ul style="list-style-type: none"> <li>① (2008년 6월 22일 : 시행시기) 쇠고기 탕용·튀김용·찜용·생식용(육회)의 원산지와 종류를 표시, 밥류로 제공하는 쌀의 원산지도 표시                                     <ul style="list-style-type: none"> <li>⇒ 쇠고기 구이용은 2007년 1월 1일부터 원산지 및 종류를 표시</li> </ul> </li> <li>② (2008년 12월 22일 : 시행시기) 돼지고기와 닭고기 구이용·탕용·튀김용·찜용, 배추김치에 대해서도 원산지를 표시</li> </ul> </li> <li>- 의무 위반시 처벌조항                             <ul style="list-style-type: none"> <li>·원산지 및 종류를 표시하지 아니한 영업자에 대하여 영업정지 및 과태료 부과</li> </ul> </li> </ul> </li> </ul> </li> </ul> </li></ul>
-------------------------	---

## 2) 식품이력 추적관리제도

- 정의

·식품을 제조·가공단계부터 판매단계까지 각 단계별로 정보를 기록·관리하여 해당 식품의 안전성 등에 문제가 발생할 경우 해당 식품을 추적하여 원인규명 및 필요한 조치를 할 수 있도록 관리하는 제도

- 대상 : 식품이력추적관리를 희망하는 자

- 등록사항

①국내식품 : 영업소 명칭 또는 상호, 제품명, 유통기한 및 품질유지기한, 보존 및 보관방법

②수입식품 : 영업소 명칭 또는 상호, 제품명, 원산지(국가), 제조회사 또는 수출회사

- 자금지원 대상

·식품이력추적관리 시스템 구축·운영에 필요한 장비 구입

·식품이력추적관리를 위한 프로그램 개발 비용

·그 밖의 식품이력추적관리에 필요하다고 정한 사업

- 등록기준 준수하지 아니한 경우 벌칙조항

□ 기대효과

○ 소비자의 선택권과 알 권리를 보장

○ 부정확한 원산지 정보에 따른 소비자 피해 방지

○ 정당한 영업자의 불이익 방지

○ 식품의 유통질서 확립

○ 식품관리의 안전성을 제고

<p>정책(제도)의 집행/평가단계별 통계지표</p>	<ul style="list-style-type: none"> <li>□ 정책집행 및 평가             <ul style="list-style-type: none"> <li>○ 통계 지표 : ① 원산지 표시 위반현황</li> <li>○ 구비 여부                 <ul style="list-style-type: none"> <li>- 통계지표 개발예정</li> </ul> </li> <li>○ 유용성                 <ul style="list-style-type: none"> <li>- 소비자의 권리보호와 주요 식품의 원산지 표시제도의 근본취지를 달성하기 위하여 준수 여부현황에 대한 판단에 필요함</li> </ul> </li> <li>○ 정보 출처                 <ul style="list-style-type: none"> <li>- 식품의약품안전청</li> </ul> </li> <li>○ 개발·개선 여부                 <ul style="list-style-type: none"> <li>- 원산지 표시 위반 업소 현황</li> <li>- 품목별 원산지 표시 위반현황</li> </ul> </li> </ul> </li> <li>○ 통계 지표 : ② 식품이력추적관리 등록현황</li> <li>○ 구비 여부             <ul style="list-style-type: none"> <li>- 통계지표 개발예정</li> </ul> </li> <li>○ 유용성             <ul style="list-style-type: none"> <li>- 식품이력추적제도 도입에 따라 그 운영실태를 파악하기 위한 기초자료로써 활용</li> </ul> </li> <li>○ 정보 출처             <ul style="list-style-type: none"> <li>- 식품의약품안전청</li> </ul> </li> <li>○ 개발·개선 여부             <ul style="list-style-type: none"> <li>- 식품이력추적관리 등록업체 수, 품목별 등록업체 수, 등록업체 지원 현황</li> </ul> </li> </ul>
--------------------------------------	---

통계개선/ 개발계획	<p>&lt;개발/개선 계획 1 : 원산지 표시 위반현황&gt;</p> <ul style="list-style-type: none"> <li>□ 조사(보고)목적 <ul style="list-style-type: none"> <li>- 원산지 표시 위반 현황을 파악하여 원산지 표시관련 정책의 기초자료 활용</li> </ul> </li> <li>□ 조사대상(보고체계) <ul style="list-style-type: none"> <li>- 시·군·구→식약청→보건복지부</li> </ul> </li> <li>□ 조사(보고)사항 <ul style="list-style-type: none"> <li>○ 원산지 표시 위반 사항 <ul style="list-style-type: none"> <li>- 원산지 표시 위반 업소 현황</li> <li>- 품목별 원산지 표시 위반현황</li> </ul> </li> </ul> </li> <li>□ 작성(보고)주기 <ul style="list-style-type: none"> <li>- 매 분기별</li> </ul> </li> <li>□ 추진일정 <ul style="list-style-type: none"> <li>○ 식약청 전산망 구축(3개월 정도 전산개발기간 필요) <ul style="list-style-type: none"> <li>→ update(계속된 자료 축적) → 국가승인통계 절차</li> </ul> </li> <li>※ 식약청이 감시 업무 과정에서 획득 되는 정보로서, 매 분기마다 통계 작성 예정</li> </ul> </li> </ul>				
	<p>&lt;개발/개선 계획 2 : 식품이력추적관리 등록현황&gt;</p> <ul style="list-style-type: none"> <li>□ 조사(보고)목적 <ul style="list-style-type: none"> <li>- 식품이력추적제도 도입에 따라 그 운영실태를 파악하기 위한 기초자료로 활용</li> </ul> </li> <li>□ 조사대상(보고체계) <ul style="list-style-type: none"> <li>- 식약청→보건복지부</li> </ul> </li> <li>□ 조사(보고)사항 <ul style="list-style-type: none"> <li>○ 식품이력추적관리 등록사항 <ul style="list-style-type: none"> <li>- 국내식품/수입식품별 등록업체 수</li> <li>- 식품이력추적관리 월별 등록 실적</li> <li>- 품목별 등록업체 수</li> <li>- 등록업체 지원 현황</li> </ul> </li> </ul> </li> <li>□ 작성(보고)주기 : 1년</li> <li>□ 추진일정 <ul style="list-style-type: none"> <li>○ 식약청 전산망 구축(6개월 정도 전산개발기간 필요) <ul style="list-style-type: none"> <li>→ update(계속된 자료 축적) → 국가승인통계 절차</li> </ul> </li> <li>※ 희망업체에서 식약청에 등록하는 과정에서 획득 되는 정보로서, 1년마다 통계 작성 예정</li> </ul> </li> </ul>				
작성자	기관명	부서명	직급	성명	전화번호
	보건복지부				

\* 「※」 란은 기재하지 말 것

## 2 검토결과보고서 작성

### 실질평가 요청에 대한 검토결과 보고

#### 1. 일반사항

※관리번호	입법예고기간	2008. 4. 15 ~ 2008. 5 6	
2008 - 221			
법령명	식품위생법 시행령 일부 개정령안		
관계법령	식품위생법 및 같은 법 시행규칙		
정책(제도)명	식품 원산지 표시제도		
정책(제도)의 통계기반 여부에 대한 부처의견	이번 개정의 식품 원산지 표시제도 확대 및 식품이력추적제도 도입에 관련된 사항은 통계를 기반하여 관리할 필요성이 있다고 판단됨.		
소관기관	기관명	보건복지가족부	
	협의부서	부서명(과/팀명)	식품정책과
		담당자 작성명/ 전화번호	행정사무관 ○○○ / 02-2023-7785

#### 2. 관련법령 개관 및 개정 사유

##### 정책(제도)의 필요성

○ 식품으로 인한 위생상의 위해를 방지하고 식품영양의 질적 향상을 도모하며 식품에 관한 올바른 정보를 제공함으로써 국민보건 증진에 이바지

## □ 정책(제도)내용

- 식품 및 식품첨가물 제도가공판매의 기준과 규격에 관한 사항
- 식품판매 또는 영업상 사용기구·용기·포장 및 표시에 관한 사항
- 식품검사 및 식품위생검사기관에 관한 사항
- 영업 및 위생, 식품이력추적관리제도에 관한 사항
- 식품위생단체(동업자조합, 식품공업협회)의 설립 및 운영에 관한 사항
- 조리사 및 영양사 면허취득 및 관리에 관한 사항
- 동 법령 위반자에 대한 시정명령, 허가취소 등에 관한 사항

## □ 금번 개정이유

- 「식품위생법」이 개정(법률 제8779호, 2007. 12. 21 공포, 2008. 6. 22 시행)됨에 따라 이 법에서 대통령령으로 위임된 원산지 표시대상 영업자 및 품목의 범위와 원산지 표시기준, 식품이력추적관리제도 도입에 관하여 규정

### 3. 법령 개정 관련 정책의 주요내용

#### □ 식육 원산지 표시제도

- 표시의무대상 : 쌀, 배추김치\*, 쇠고기, 돼지고기, 닭고기  
\* 수입한 배추 : 원산지 표시, 제조가공한 배추김치 : 수입국 표시

- 시설규모 : 100㎡ 이상 일반음식점
- 의무품목의 범위

- (2008년 6월 22일 : 시행시기) 쇠고기 탕용·튀김용·찜용·생식용(육회)의 원산지와 종류를 표시, 밥류로 제공하는 쌀의 원산지도 표시

⇒ 쇠고기 구이용은 2007년 1월 1일부터 원산지 및 종류를 표시

- (2008년 12월 22일 : 시행시기) 돼지고기와 닭고기 구이용·탕용·튀김용·찜용, 배추김치에

대해서도 원산지를 표시

○ 의무 위반시 처벌조항

- 원산지 및 종류를 표시하지 아니한 영업자에 대하여 영업정지 및 과태료 부과

□ 식품이력 추적관리제도

○ 정의

- 식품을 제조·가공단계부터 판매단계까지 각 단계별로 정보를 기록·관리하여 해당 식품의 안전성 등에 문제가 발생할 경우 해당 식품을 추적하여 원인규명 및 필요한 조치를 할 수 있도록 관리하는 제도

※ 식품이력추적제도

- 식품의 생산부터 소비까지 식품이력정보를 소비자에게 제공하여 알권리를 확보하고, 식품사고 발생시 신속한 원인규명과 회수를 통해 식품안전을 확보하기 위해 RFID\* 칩을 제품에 부착·사용하여 식품의 원료 구입, 제조 가공, 유통, 판매 등 전 과정을 추적관리하는 제도를 도입(식품위생법 개정, 2007. 12. 21일 공포)

\* RFID (Radio Frequency Identification: 무선인식) : 안테나와 칩으로 구성된 종이 두께의 얇은 태그를 해당제품에 부착 후 판독기를 통하여 정보를 인식하고 확인하는 기술

- 대상 : 식품이력추적관리를 희망하는 자

○ 등록사항

- 국내식품 : 영업소 명칭 또는 상호, 제품명, 유통기한 및 품질유지기한, 보존 및 보관방법
- 수입식품 : 영업소 명칭 또는 상호, 제품명, 원산지(국가), 제조회사 또는 수출회사

○ 자금지원 대상

- 식품이력추적관리 시스템 구축운영에 필요한 장비 구입
- 식품이력추적관리를 위한 프로그램 개발 비용
- 그 밖의 식품이력추적관리에 필요하다고 정한 사업

- 등록기준 준수하지 아니한 경우 벌칙조항

#### 4. 필요 통계지표 및 구비여부 검토

##### 원산지 표시 대상 확대

○ 소비자의 권리보호와 주요 식품의 원산지 표시제도의 근본취지를 달성하기 위하여 준수 여부현황에 대한 판단이 반드시 필요하며,

- 이를 위하여 원산지 표시 위반업소 및 품목별 원산지 표시 위반현황, 일반음식점업 사업장 건물면적 규모별 사업체현황에 대한 통계를 관리할 필요가 있음

##### 식품이력추적제도 도입

○ 식품이력추적제도 도입에 따라 그 집행상황을 파악하기 위한 기초자료로써 식품이력추적관리 등록업체 현황 및 등록업체에 대한 지원현황을 관리할 필요성이 있음

## 【 필요 통계지표 정리】

필요 판단	필요 통계지표	조사 사항
원산지 표시제도의 준수여부 파악	· 원산지 표시 위반현황	· 원산지표시 위반업소 현황 · 품목별 원산지표시 위반현황
식품이력추적관리 운영실태 파악	· 식품이력추적관리 등록현황	· 국내/수입식품별 등록업체 현황 · 월별 등록 실적현황 · 품목별 등록업체 현황
	· 식품이력추적관리 등록업체 지원현황	· 등록업체에 대한 지원실적
원산지 표시대상 일반음식점업 현황 파악	· 일반음식점업 건물면적 규모별 사업체현황	· 사업장 건물면적 규모별 사업체현황*(참고지표, 통계청)

\* 출처 : 도소매업통계조사 보고서(연간), 서비스업총조사 보고서(5년)

## 5. 소관부처 요청서 검토

 통계지표 관련

○ (소관부처 의견) 보건복지가족부는 ① 원산지 표시 위반 현황 ② 식품이력추적관리 등록현황 및 등록업체 지원현황에 대한 통계지표의 필요성을 제시하였음

○ (검토 의견) 동 정책의 집행에 필요한 통계지표를 모두 제시한 것으로 판단됨

※ 다만, 원산지 표시대상 일반음식점업 현황 파악을 위해 도소매업통계조사 및 서비스업총조사 결과를 참고할 것을 제시함

□ 통계개발/개선 계획 관련

○ (소관부처 의견) 식품의약품안전청 전산망을 구축하여 원산지 표시 위반 현황, 식품이력추적 관리 등록현황 및 등록업체에 대한 지원현황을 작성할 계획임

**【통계지표 작성계획】**

□ 원산지 표시 위반현황

- ◆ 조사대상(보고체계) : 시군구 → 식약청 → 보건복지부
  - ◆ 조사(보고)사항
    - 원산지 표시 위반업소 현황
    - 품목별 원산지 표시 위반현황
  - ◆ 작성주기: 매분기
  - ◆ 추진일정
    - 자료조사 → 식약청 전산망 구축 (전산개발 예정기간 : 3개월) → 국가승인통계 여부 협의
- ※ 식약청이 감시 업무 과정에서 획득 되는 정보로서 매 분기별로 통계 작성 예정

□ 식품이력추적관리 등록현황

- ◆ 조사대상(보고체계) : 식약청 → 보건복지가족부
  - ◆ 조사(보고)사항
    - 국내식품/수입식품별 등록업체 수
    - 식품이력추적관리 월별 등록 실적
    - 품목별 등록업체현황
    - 등록업체 지원 현황
  - ◆ 작성주기: 1년
  - ◆ 추진일정
    - 자료조사 → 식약청 전산망 구축 (전산개발 예정기간 : 3개월) → 국가승인통계 신청여부 협의
- ※ 희망업체에서 식약청에 등록하는 과정에서 획득 되는 정보로서 1년마다 통계 작성 예정

○ (검토 의견) 동 개발계획의 조사대상, 조사사항, 조사주기, 공표시기 등에 대하여 이견 없음

- 보건복지가족부에서 제시한 통계지표는 행정정보를 통해서 필요 통계지표의 작성 및 신뢰성 확보가 가능할 것으로 판단됨

## 6. 평가 의견 및 조치 계획 : 원안동의

① 원산지 표시 위반 현황 ② 식품이력추적관리 등록현황 및 등록업체 지원현황에 대한 통계지표가 필요하며

○ 소관부처에서 제시한 통계개발/개선계획을 통하여 동 제도의 집행평가에 적절한 통계가 구비될 것으로 판단됨

※ 다만, 원산지 표시대상 일반음식점업에 대한 일반현황을 파악하기 위해 도소매업통계조사 및 서비스업총조사 결과를 참고할 것을 제시

### 3 최종평가결과 통보서

※ 관리번호	2008 - 221	
통계기반 정책관리 최종평가 결과 통보서		
<b>I. 대상법령</b>		
법령명	식품위생법 시행령 일부 개정령안	
소관기관/부서	보건복지가족부/식품정책과	
정책(제도)명	식품 원산지 표시제도	
<b>II. 종합 평가의견 : 「원안동의」</b>		
<p><input type="checkbox"/> ① 원산지 표시 위반 현황 ② 식품이력추적관리 등록현황 및 등록업체 지원현황에 대한 통계지표가 필요하며</p> <p>○ 소관부처에서 제시한 통계개발/개선계획을 통하여 동 제도의 집행·평가에 적절한 통계가 구비될 것으로 판단됨</p> <p>※ 다만, 원산지 표시대상 일반음식점업에 대한 일반현황을 파악하기 위해 도소매업통계조사 및 서비스업총조사 결과를 참고할 것을 제시함</p> <p><b>【필요 통계지표 정리】</b></p>		
<b>필요 판단</b>	<b>필요 통계지표</b>	<b>조사 사항</b>
원산지 표시제도의 준수여부 파악	· 원산지 표시 위반현황	· 원산지표시 위반업소 현황 · 품목별 원산지표시 위반현황
식품이력추적관리 운영실태 파악	· 식품이력추적관리 등록현황	· 국내/수입식품별 등록업체 현황
	· 식품이력추적관리 등록업체 지원현황	· 월별 등록 실적현황 · 품목별 등록업체 현황 · 등록업체에 대한 지원실적
원산지 표시대상 일반음식점업 현황 파악	· 일반음식점업 건물면적 규모별 사업체현황	· 사업장 건물면적 규모별 사업체현황 지표, 통계청)
* 출처 : 도소매업통계조사 보고서(연간), 서비스업총조사 보고서(5년)		

Ⅲ. 평가항목		
평가항목	평가착안사항	평가결과
통계지표의 적합성 평가	■ 정책(제도)별로 통계지표가 제시되었는가?	제시됨
	■ 제시된 통계지표만으로 정책(제도)의 집행 / 평가가 충분한가?	충분함
	■ 통계지표가 부적절한 경우 대체할만한 다른 통계지표는 있는가?	해당없음
	■ 제시된 통계지표가 기존 통계조사 또는 행정보고를 통해 작성 가능한가?	가능함
	■ 통계지표가 행정보고를 통해 작성되는 경우, 안정적이고 신뢰성 있게 작성할 수 있는 체계가 갖추어져 있는가?	갖추어져 있음
통계개선/ 개발계획의 타당성 평가	■ 통계개발/개선계획을 통해 필요한 모든 통계지표의 작성이 가능한가?	가능함
	■ 통계개발/개선계획은 적절한가? - 조사목적, 조사대상, 주요 조사사항, 작성주기 등	가능함
	■ 개발/개선된 통계지표의 공표시기는 적절한가?	적절함
	■ 필요한 통계지표를 생산하고 있는 기존의 통계조사와 중복되지는 않는가?	해당없음
	■ 개발/개선계획이 타 기관에서 이루어져야 하는 경우 해당기관과 충분히 협의되었는가?	해당없음
	■ 중앙행정기관이 통계개발/개선계획을 제시하지 않은 경우에 상기 평가기준을 토대로 개발/개선계획을 마련, 해당기관과 협의하여 개발·개선 계획이 합의되었는가?	해당없음
<p>위와 같이 통계법 시행령 제33조(정책통계기반평가의 절차 및 방법)에 의한 평가결과를 통보 하오니, 평가결과가 “개선권고”인 경우에는 평가의견을 소관 통계 개발/개선계획에 적극 반영하시기 바랍니다.</p> <p style="text-align: center;">2008. 5. 27.</p> <p style="text-align: center;"><b>통 계 청 장</b></p>		

## 통계와 정책 표준교재

---

발 행 : 2008년 12월 26일

인 쇄 : 2008년 12월 23일

펴낸이 : 문창용

펴낸곳 : 통계교육원

총괄지휘 : 이명호

기 획 : 김정란 · 김명옥

주 소 : 대전 유성구 과학로 60번지

전 화 : 042-601-4430, 4427

홈페이지 : <http://sti.nso.go.kr>