

## 제 4 장

# 지역사회조사

김 영 수 · 이 병 진

### 학습목표

1. 일차자료를 설명할 수 있다.
2. 확률표본추출을 설명할 수 있다.
3. 비확률표본추출을 설명할 수 있다.
4. 이차자료를 설명할 수 있다.
5. 변수를 정의할 수 있다.
6. 양적변수를 설명할 수 있다.
7. 질적변수를 설명할 수 있다.
8. 척도를 분류할 수 있다.
9. 모집단의 특성을 설명할 수 있다.
10. 모집단과 통계량의 분포를 설명할 수 있다.
11. 확률분포를 설명할 수 있다.
12. 통계적 가설검정의 원리를 설명할 수 있다.
13. 유의확률을 설명할 수 있다.
14. t-분포를 이용한 의사결정을 설명할 수 있다.
15. 표본 수의 결정을 설명할 수 있다.

## 1 구강보건통계의 기본 개념

### 1) 구강보건통계의 필요성

국가 혹은 지역사회처럼 다수의 구성원으로 이뤄진 집단의 구강건강상태를 파악하고 관련된 요인을 조사하는 것은 현대 국가의 구강보건사업을 계획하거나 구강보건관리 및 질병관리에 매우 중요한 일이다. 집단의 구강건강상태를 파악하기 위해 일일이 모든 구성원을 조사하는 것은 물리적으로 불가능하다. 이러한 한계를 극복하기 위하여 해당 집단의 일부 구성원을 선택하여 구강건강상태를 조사하고, 이를 통해 집단의 구강건강상태를 유추하는 과정을 구강보건통계라고 할 수 있다. 이와 같이 각종 구강건강 및 구강보건과 관련된 지표를 생산하고 이를 위해 조사방법을 계획하는 모든 과정에서 통계적 방법이 활용되고 있으며, 새로 개발되는 치료방법의 효과나 안전성을 평가하는 과정 모두 통계적 방법을 통해 검증되고 활용되고 있다.

### 2) 통계에 대한 이해

연구하고자 하는 모집단의 특성을 나타내는 것을 ‘모수(parameter)’라고 한다. 일반적으로 많이 활용되는 모수는 ‘평균’, ‘분산’, ‘표준편차’, ‘분율’ 등이 있다. ‘평균’은 측정치의 산술적인 대푯값으로 활용되는 경우가 많고, ‘분산’과 ‘표준편차’는 평균을 중심으로 자료가 흩어진 정도인 산포도를 나타낸다. 예를 들어 ‘우리나라 국민의 치아우식증 경험 치아 수’를 산출한다고 가정하면 모집단은 우리나라 전체 집단이 되고, 목표 모수는 평균이 된다.

통계학은 모집단의 특성을 구명하기 위한 방법으로, 모집단의 특성은 모수를 통해 표현되며, 모집단의 모수는 ‘추정’과 ‘가설검정’을 통해 추론된다. 통계학은 모집단의 모수를 표본의 추정량을 통해 설명하고 이해하고자 하는 것이며, 표본 자체의 특성을 연구하는 것은 아니다.

표본으로부터 얻어진 값을 통해 모수의 특성을 잘 표현해 주는 값을 찾는 과정을 ‘추정’이라고 하며, 하나의 값으

로 표현한 것을 ‘점추정’, 구간으로 나타낸 것을 ‘구간추정’이라고 한다.

‘가설검정’은 모집단의 모수에 대한 주장이나 추측 등의 옳고 그름을 판단하는 것이다. 치의학 연구에서 다루는 많은 가설들은 위와 같은 통계학적 추정과 검정을 통해 답을 얻기 때문에 통계학적 추론은 치의학 연구에서 매우 중요하다.

### 3) 지역사회조사를 위한 자료 수집

#### (1) 일차자료

일차자료(primary data)는 연구 혹은 보건사업을 목적으로 조사자가 조사 대상으로부터 직접 수집한 자료를 뜻한다. 일반적으로 지역사회 인구집단을 대상으로 설문조사, 구강검사와 같은 신체 측정, 타액 혹은 치면세균막과 같은 시료 검사 등을 시행하여 얻은 자료가 이에 해당된다. 일차자료는 조사자의 노력과 역량에 따라 다양한 자료의 형태로 구성되는데 조사하고자 하는 목적과 대상의 크기에 따라 사용하는 방법이 달라질 수 있다. 일차자료는 의사소통을 통한 방법과 관찰방법으로 수집되는데 면접설문, 우편, 전화, 인터넷 등이 의사소통에 의한 방법에 해당한다. 설문조사는 표준화, 구조화된 설문을 이용하여 필요한 정보를 얻는 조사방법으로 누가 기록하는지에 따라 분류되며, 스스로 기록하는 자기기입 설문조사와 조사자가 기록하는 조사로 구분된다. 일반적으로 응답률이 50% 이상이면 긍정적, 60% 이상이면 양호, 70% 이상이면 만족할 만한 수준이라고 평가하고 있다.

#### ① 표본조사

일반적으로 많은 수의 인가로 구성된 집단에서는 모든 모집단 구성원을 대상으로 조사를 진행하는 것이 불가능하다. 이러한 문제를 극복하는 방법은 조사 대상 모집단에서 일부의 조사대상자를 선택하여 조사하는 것이다. 표본조사가 성공적이라면 추출된 표본이 모집단의 특성을 잘

반영하여야 하며 이를 '대표성'이라고 표현한다. 그리고 모집단 전체를 조사하였을 때 얻을 수 있는 결과와 다른 결과가 도출될 수 있으며, 이 둘 간의 차이를 '표본오차'라고 한다. 그렇기 때문에 표본을 추출할 때에는 표본오차를 최소화하고, 모집단의 특성을 잘 반영하는 것이 중요하다. 이렇듯 대표성을 확보한 표본이 추출하기 위해서는 산출한 통계량에 대한 오차 허용 정도, 추출 틀과 단위, 필요한 표본 수, 표본 선정방법, 적절한 모수 추정법 등을 사전에 고려하여야 한다. 표본조사는 모집단 전체를 조사한 결과와 일치하지 않을 수도 있고, 표본을 추출하는 과정에서 발생할 수 있는 오차가 발생할 수 있으나, 조사에 들어가는 비용과 시간을 줄이고 측정의 오차를 줄일 수 있다는 장점이 있다.

## ② 표본추출법

### 1. 확률표본추출

연구자의 의도가 포함되지 않은 확률적 규칙에 의해 무작위로 표본을 추출하는 방법을 말한다. 확률표본추출법으로 추출된 표본은 모집단을 대표할 수 있기 때문에 신뢰구간을 산출하고, 통계적 유의성을 검정할 수 있다.

#### a. 단순무작위 표본추출(simple random sampling)

가장 단순한 확률표본추출법으로 모집단의 모든 구성원

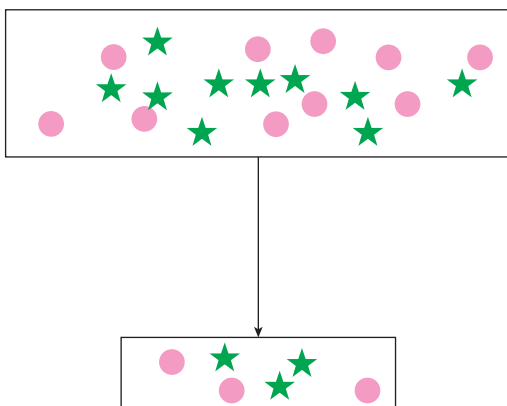


그림 4-1. 단순무작위 표본추출의 개념.

의 표본추출 확률을 동일하게 배정하여 추출하는 방법이다. 대상자에게 일련번호를 부여하여 일정한 순서대로 추출하거나, 난수(random number)를 생성하여 번호에 해당하는 사람을 표본으로 선정하는 방법이다(그림 4-1).

#### b. 층화무작위 표본추출

층화무작위 표본추출법(stratified random sampling)은 조사 결과에 영향을 줄 있는 변수(예: 성별 나이 등)에 따라 적절한 수의 층(strata)을 구분하고, 각 층에서 추출할 표본 수를 먼저 결정한 뒤, 단순무작위 표본추출 방법에 따라 각 층별로 표본을 추출한다(그림 4-2). 표본이 층분히 크지 않을 때에는 모집단의 구성과 유사해야 하므로 이 방법을 주로 사용한다. 층화추출법은 각 층의 크기에 비례하여 표본의 크기를 결정하는 비례층화추출법과 이와 무관하게 표본을 추출하는 비비례층화추출법으로 구분된다. 비례층화추출은 각 층으로부터 단순무작위 표본추출을 할 때 표본의 크기를 각 층의 크기에 비례하여 설정하며(그림 4-3), 비비례층화추출은 각 층의 크기와 관계없이 표본의 크기를 설정한다.

#### c. 계통표본추출법

모집단의 목록이 잘 정리된 경우 일정한 간격으로 표본

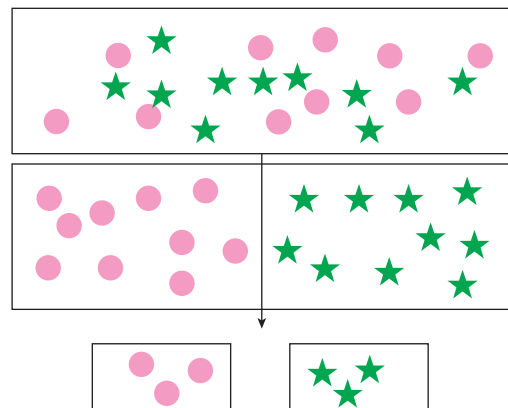


그림 4-2. 층화무작위 표본추출의 개념.

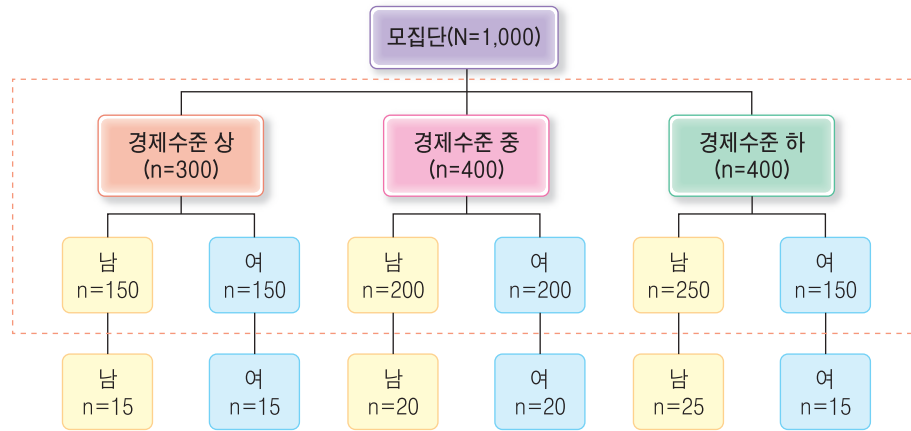


그림 4-3. 비례층화표본추출의 개념.

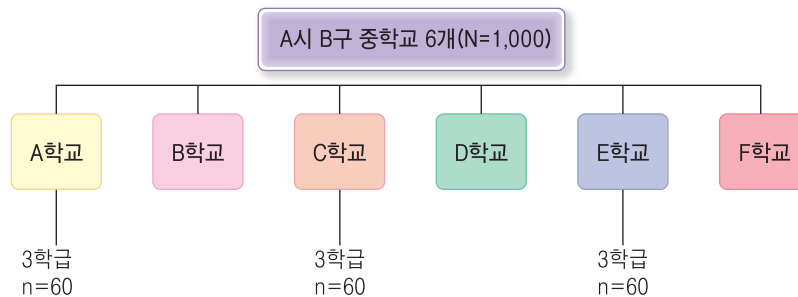


그림 4-4. 집락표본추출의 개념.

을 추출하는 방법이다. 이 방법은 단순무작위 표본추출법보다 추출작업이 쉽고, 선정된 표본들이 고르게 분포되어 있는 경우 표본의 대표성을 확보할 수 있다는 장점이 있다. 하지만 모집단의 목록이 무작위가 아니고, 일정한 경향성이 있는 경우 표본의 대표성이 떨어지므로, 경향성이 있는지 검증해보아야 한다.

#### d. 집락표본추출

표본이 집락을 중심으로 분포되어 있는 경우 집락 내에서 표본추출을 하거나, 집락을 표본추출하여 집락 내 전수 조사하거나 혹은 집락 내에서 추출한 표본을 조사하는 방법이다(그림 4-4). 이 방법은 대상자가 넓은 지역에 흩어져 있어 모든 대상으로 병합하여 무작위추출하는 것이 비실용적이거나 경비가 많이 들 때 사용할 수 있다. 우리나라의 국민(아동)구강건강상태조사도 집락표본추출방법을 사

용하고 있는데, 전국의 중학교 1학년을 모집단으로 지역별 집락을 중심으로 학교와 학급을 추출하여 12세 아동의 구강건강상태를 조사하고 있다.

#### 7. 비확률표본추출

모집단의 특성이 확률표본추출법으로 표본을 추출하기 어려운 경우 조사자의 의도에 따라 표본을 추출하는 방법으로, 일반적인 역학조사에서는 통계적 검정법을 통해 모수를 추정하기 어렵기 때문에 사용하기 어렵다.

#### ③ 표본조사방법

##### 7. 응답자가 직접 기록하는 방법

###### a. 자기기입식 설문조사

- **개별배포조사:** 피조사자인 응답자가 설문지를 직접 작성한 후 조사자가 회수하는 방법이다. 응답을 결정하는 데

에 충분한 시간이 있어 회수율이 높고, 비용이 적게 들어가기만 응답자가 문해력이 있어야 하고, 잘못된 응답을 바로잡기 어려우며, 타인 혹은 대리응답이 가능하다는 단점이 있다.

- **집단조사:** 응답자들을 정해진 장소에 모은 후 설문지를 배포하고 조사자가 조사와 관련된 내용을 응답자에게 설명하면서 응답자가 직접 기재하는 설문조사 방법이다. 조사대상자가 학교, 기업 등 집단으로 구성된 경우 유용하다. 조사 시간과 비용이 적게 들고, 응답자의 조건이 같으며, 개별 조사로는 어려운 경험이나 특성에 관한 조사가 가능하며, 응답 시 추가적으로 설명이 필요한 사항을 해결할 수 있는 장점이 있으나, 응답자끼리 답변에 영향을 줄 수 있다는 단점이 있다.

#### b. 우편조사

설문지를 우편으로 전달하고 응답자들이 직접 작성한 후 우편으로 반송 받는 방법으로, 조사목적 및 설문자의 정보 등과 조사협조의 안내문이 동봉하는 경우가 많다. 일반적으로 우편으로 회신하는 경우에는 반송봉투와 우표를 함께 보낸다. 이 조사는 대상자에 대한 정보가 봉투에 기재되어 있어 회송상태를 확인할 수 있으나, 회신율이 낮은 경우가 많기 때문에 독촉을 하는 경우도 있다.

#### c. 인터넷조사

인터넷 사용이 대중화되면서 최근에는 인터넷 사이트와 이메일을 이용한 조사방법이 다양하게 개발되었다. 인터넷 조사는 응답자의 이메일 정보를 수집하여 형식을 갖춘 설문 문항을 이메일로 발송하여 작성 후 회신하는 방법과 응답자가 조사자가 개설한 설문사이트를 방문하여 직접 응답을 기록하는 방식이 있다.

### 7. 설문조사자가 기록하는 방법

#### a. 면접조사

조사원이 응답자를 개별적으로 만나 정보를 얻는 방법으로 응답률이 높고 깊이 있는 조사가 가능하다. 과거에는 인쇄된 설문지를 이용하여 조사하는 경우가 많았는데, 최근에는 컴퓨터나 태블릿 등을 이용하여 조사자의 질문에 응답하거나, 조사자 설명 후 태블릿에 직접 입력하는 방법도 활용되고 있다. 면접조사법은 설문을 통해 정보를 수집하는 방법 중에서 가장 많이 활용되는 방법이다.

#### b. 전화조사

전화를 이용하여 조사원이 질문하고 기록하는 방법이다. 다수를 대상으로 한 여론조사나 시급히 조사해야 하는 주제에서 많이 활용되며 질문내용이 짧고 명확한 대상인 경우에 활용 가능하다. 최근에는 유선전화뿐만 아니라 무선전화의 메시지 혹은 애플리케이션을 이용한 조사 방법도 활용되고 있다.

## (2) 이차자료

### ① 이차자료란?

이차자료는 명확하게 설정된 목표를 달성하기 위해 수집된 자료가 아닌 다른 목적으로 수집되거나 보고된 자료 중에서 연구자가 활용하는 자료를 의미한다. 많이 사용되는 이차자료로는 인구, 사망, 건강보험, 의료기관, 보건기관, 행정기관에서 수집하는 자료와 국가적인 건강조사 등이 있다. 이차자료는 연구하고자 하는 건강 관련 사안의 특성에 따라 그 가치가 달라지므로 이차자료를 이용하거나 결과를 해석할 때에는 조사대상 집단에 대한 특성과 자료의 질적 수준을 고려하여야 한다. 최근에는 일차자료와 이차자료 및 다른 인구사회학적 자료를 병합하여 조사를 수행하는 경우도 있다.



## ② 인구자료

치아우식증이나 치주질환의 유병은 조사 연령에 따라 큰 차이가 있기 때문에 조사대상자의 연령별 구성에 따라 조사결과에 오차가 발생할 수 있다. 이러한 오차를 줄이기 위해서는 연령표준화를 거쳐야 하며 표준화를 산출을 위해서는 인구자료가 필요하다. 우리나라에서 많이 활용되는 인구자료는 인구주택총조사 자료와 행정안전부의 주민등록인구 자료 등이 있다(그림 4-5, 4-6). 5년마다 실시되는 인구주택총조사는 통계청이 시행하고 발표하며 통계청 홈페이지에 원자료를 공개하고 있다.

주민등록인구통계는 매일 주민등록부에 등재된 내용 중에서 연령별 성별 인구수를 지역 자치단체 및 행정구역 단위별로 구분하여 행정자치부 홈페이지에 등재한 자료이다.

인구주택총조사 자료는 인구수뿐만 아니라 가구, 주택 등에 대한 다양한 특성을 파악할 수 있는 장점이 있지만 5년마다 수행되어 시의성 있는 조사에서는 오차가 발생할 수 있고, 자료의 결과도 지연되는 단점이 있다.

## ③ 사망자료

사망자료(mortality)는 ‘가족관계의 등록 등에 관한 법률’에 따라 신고하는 사망신고 자료를 종합하여 국가 혹은 지역사회 인구집단의 사망 수준과 원인을 파악하기 위해

생산하는 자료를 말한다. 통계청은 사망신고서에 기재된 사고, 질병 등의 사망 원인과 사망자의 인구특성을 종합하여 ‘사망원인통계보고’를 작성하며 이를 통해 사망률 및 사망 원인에 대한 통계자료를 공개하고 있다. 사망자료는 비교적 정확하게 보고되고 수집되기 때문에 일반 보건학에서는 중요한 지표로 활용되고 있으며, 이를 이용하여 인구의 조기 사망을 예방하기 위한 정책자료로 활용되고 있다.

## ④ 상병자료

상병자료는 조사자가 직접 조사하는 건강관련 자료가 아닌 다른 기관에서 생산된 질병과 부상 및 진료행위에 대한 자료이다. 조사자가 직접 조사하기 어려운 대규모 집단의 정보를 손쉽게 확보할 수 있고, 목적에 따라 가공이 가능하지만 최근 개인정보보호에 대한 인식 및 제도가 강화되면서 자료의 접근성에 한계가 있기도 하다.

## 7. 국민건강보험

우리나라의 국민건강보험제도를 통해 거의 모든 의료기관이 요양급여기관으로 등록되고, 거의 모든 국민이 지정된 진료를 요양급여로 받고 있기 때문에 이에 해당하는 진료행위와 관련된 국민건강보험의 통계자료는 상당히 규모가 크고 정확한 수준이다. 특히 건강보험제도가 정착된 이



그림 4-5. 대한민국 인구주택총조사 홈페이지.



그림 4-6. 행정안전부 주민등록인구통계 홈페이지.

후 오랫동안 누적되어 온 진료행위 및 진료비 통계자료는 최근 빅데이터로 많이 활용되고 있다. 다만 진료행위에 등록된 상병자료는 건강보험 청구과정의 특성상 실제 질병상태를 반영하는지 여부를 판단하기 힘들기 때문에 각 질병별 세부자료로 활용하기는 적절하지 않다.

#### 나. 사업장, 학교 및 기관

학교나 직장 등 각급 기관에서 정기적으로 시행하는 구강검진 자료나 질병 발생, 부상에 관한 자료들은 해당 기관이나 직종의 특성을 이해하는 데 도움이 될 수 있다. 그러나 해당 기관의 자료는 정밀한 조사에 활용하기에는 정확도가 부족할 수 있고 일반화하기 어려운 단점이 있다. 하지만 일상적으로 잘 발생하지 않는 질환이나 직업적·환경적 요인과 관련된 상병을 조사하는 데에 유용하게 활용될 수 있다.

#### 다. 의료기관

의료기관에서 등록·관리하고 있는 환자의 인구학적 특성과 질병 및 진료관련 기록들은 특히 임상적 연구를 진행할 경우에 많이 활용된다. 특히 진단명 및 처치의 종류와 경과 등이 자세하게 기록되어 있기 때문에 세밀한 임상연구의 좋은 정보가 된다. 그러나 이러한 의료기록은 의료윤리심의위원회의 승인을 거치거나 환자의 동의를 통해 확보될 수 있기 때문에 자료가 한정적이고 인구 전체로 일반화하기 어렵기 때문에 특정한 목적이 있는 경우에만 제한적으로 사용하는 것이 좋다.

#### 르. 질병 신고

우리나라의 질병관리본부에서는 법정감염병이나 국민건강의 주요 문제가 되는 질환을 신고하도록 제도화되어 있다. 이렇게 신고되는 질환에 대해서는 질병관리본부의 통계자료가 유용하게 활용될 수 있다.

#### 미. 질병등록

질병관리본부와 건강보험공단에서는 암과 같은 특정 질환(희귀난치병 등)이 발병하여 치료하는 경우에는 질병 발생을 등록하고 환자에게는 치료 혜택을 주는 사업을 진행하고 있다. 구강암과 같은 암등록사업을 통해 암의 발병률, 유병률 및 생존율 등 질환의 특성을 이해하는 데 도움이 되는 자료를 확보할 수 있다.

#### 비. 건강조사

우리나라 정부에서는 국가의 정책수립을 위한 기초 자료를 확보하기 위해 여러 종류의 건강조사를 시행하고 있다. 이 조사를 통해서도 지역사회 건강수준을 파악하는 자료를 확보할 수 있다.

##### 가. 국민구강건강실태조사

국민(아동)구강건강실태조사는 ‘구강보건법’을 근거로 2000년부터 본격적으로 시행된 구강건강관련 국가조사이다. 국민구강건강실태조사는 특히 치아우식증에 취약한 아동연령을 대상으로 집중적으로 시행되고 있으며 이를 통해 구강보건정책 수립에 필요한 지표를 생산하고 있다. 국민구강건강실태조사의 원자료를 이용하면 지역별, 연령별, 성별 치아우식증 발생, 구강건강인식 등에 대한 자료를 확보할 수 있다. 국민구강건강실태조사에 대한 자세한 내용은 제5장에 서술되어 있다(그림 4-7).

##### 나. 국민건강영양조사

국민건강영양조사는 ‘국민건강증진법’에 근거하여 우리나라 국민의 건강 및 영양 상태에 대한 통계를 생산하여 국가 보건정책의 기초 자료 및 평가 자료로 활용하고 국제기구의 국가 건강지표로 활용하기 위해 실시하는 조사이다. 국민건강영양조사는 매년 1만여 명의 국민을 대상으로 건강설문조사, 검진조사 및 영양조사를 각 지역별로 방문하여 시행하고 있다. 국민건강영양조사의 검진항목에는 구강

2015년도  
국민구강건강실태조사

### 구강검사 기록지

보건복지부  
MINISTRY OF HEALTH & WELFARE

연령	1=만 12세 2=만 5세	학교번호				일련번호			성별	남=1 여=2	조사자	
이름		태어난해(양력)	20	년	검사일	월	일	조사자의 전체 검사자수				

17 16

$\frac{15}{55}$   $\frac{14}{54}$

$\frac{13}{53}$   $\frac{12}{52}$   $\frac{11}{51}$

$\frac{21}{61}$   $\frac{22}{62}$   $\frac{23}{63}$

$\frac{24}{64}$   $\frac{25}{65}$

26 27

47 46

$\frac{45}{85}$   $\frac{44}{84}$

$\frac{43}{83}$   $\frac{42}{82}$   $\frac{41}{81}$

$\frac{31}{71}$   $\frac{32}{72}$   $\frac{33}{73}$

$\frac{34}{74}$   $\frac{35}{75}$

36 37

치아상태	치석유무	치은염	치아반점도(만 12세)
<p>0 = 건전치면 1 = 우식치면 3 = 우식경험치면 4 = 우식경험상실치면 5 = 우식비경험상실치면 6 = 전색치면 7 = 우식비경험치면 8 = 미명출치면 9 = 기록불가치면</p>	<div style="text-align: center;"> <p>0 = 치석 없음 1 = 치석 있음</p> <p>(하악 전치부 설면)</p> </div>	<div style="text-align: center;"> <p>0 = 치은염 없음 1 = 치은염 있음*</p> <p>* Loe and Silness Gingival Index의 3점에 해당하는 현저한 발적과 부종이 존재하는 경우</p> <p>(상·하악 전치부 순면)</p> </div>	<div style="text-align: center;"> <p>0 = 정상치아 1 = 반점의문치아 2 = 경미도 반점치아 3 = 경도 반점치아 4 = 중등도 반점치아 5 = 고도 반점치아 7 = 불소이외의 원인으로 인한 이상 9 = 기록불가</p> </div>

그림 4-7. 국민구강건강실태조사 구강검사 결과기록지.

검사가 포함되어 있어 조사 현장에서 치아우식증 및 치주 질환에 대한 조사를 수행하고 있으며 영유아 및 성인 인구의 구강질환 발생 상태를 평가하는 대표적인 지표들을 생산하고 있다(그림 4-8).

**c. 지역사회건강조사**

지역사회건강조사는 시군구 단위 지역 주민의 건강상태와 위험요인을 조사하여 건강수준을 평가하기 위한 조사이다. 각 지역별로 약 900명 수준의 표본을 추출하여 주요 만성질환과 의료이용 및 건강관리 행태를 조사하고 있다. 이 조사에서는 구강건강상태를 구강건강 관련 인식과 칫솔질, 치과의료 이용 및 치과요를 받지 못한 이유를 설문으로 조사하고 있다. 지역사회건강조사로는 구강질환의 유병률을 정밀하게 측정하지 못하지만, 보건관리 단위 지역별로 구강건강관리에 대한 지표를 생산할 수 있고, 다



그림 4-8. 국민건강영양조사 홈페이지.



른 만성질환관리와의 관련성을 평가하는 데 유용하게 사용될 수 있다.

#### d. 청소년건강행태온라인조사

청소년건강행태온라인조사는 ‘국민건강증진법’에 근거하여 실시되는 온라인조사로, 중학교 및 고등학교 학생 8만

명을 표본추출하여 건강행동(흡연, 음주, 신체활동, 식생활 등)과 건강에 대한 인식 등을 설문으로 조사하고 있다. 청소년건강행태온라인조사는 익명으로 인터넷에 응답하는 방식을 사용하고 있으며, 구강건강에 대하여는 칫솔질 횟수, 치아홈메우기 수진, 구강증상 경험 등에 대한 설문이 포함되어 있다.

## 2 변수와 척도

### 1) 변수

#### (1) 변수의 정의

변수(variable)는 분류체계나 측정도구에 의해 수치로 표현될 수 있는 현상의 특성이라고 정의할 수 있다. 이에 대해 자료(data)란 어떤 현상을 설명하기 위하여 수치적 특성이나 사실적 특징으로 나타낸 것으로 변수를 수치로 표현한 것이라 할 수 있다. 어느 기업체의 입사 서류·면접에 서라면 인사팀이 요구하는 외국어 듣기, 말하기 능력이 ‘현상’에 해당하고, 사원들의 외국어 듣기와 말하기 능력이 하나의 변수(외국어 실력)에 해당하며, 사원들의 외국어 듣기와 말하기 시험 성적(TOEIC 점수, TEPS 등급, JPT 성적, HSK 성적 등)이 자료에 해당한다.

예를 통해 우리의 이해의 폭을 넓혀 보도록 하자.

다음 표 4-1은 어느 결혼정보회사가 보유하고 있는 배우자가 될 수 있는 대상자들에 관한 정보이다. 우리 시대를 사는 청년들의 ‘속마음’을 통해 ‘변수’라는 말의 의미를 생각해 보자. 여기에서 연수입이나 부모의 재산 액수, 나이, 주거주택의 평수와 같이 수치적 특성을 가지는 변수를 양적변수(quantitative variable)라 부르고, 성별이나 직업, 외모 수준, 주거 형태, 주거 지역, 부모의 생존 여부 및 사회적 위치, 그리고 본인의 출신 학교 등과 같이 비수치적인 특성을 나타낼 때 사용하는 변수를 질적변수(qualitative variable)라고 부른다.

#### (2) 양적변수

무게, 면적 등과 같이 수치적 특성을 가지는 변수를 양적변수(quantitative variable)라고 하며, 양적변수가 갖는 자료를 양적자료(quantitative data)라 한다. 양적변수는 연속변수(continuous variable: 두 관찰 값 사이에 무한한 수의 값이 존재할 수 있는 변수. 예: 키, 몸무게)와 이산

변수(discrete variable: 두 관찰 값 사이가 분리되어 있는 변수. 예: 사무실 안의 사람 수, 우식경험치아 수, 휴가철의 자동차사고의 발생건수)로 나뉜다. 우리가 몸무게를 표시할 때, 70kg이라고 표시하는 것은 정확히 70kg라고 표시한다고 주장해도 70.002kg일 수도 있고, 69.997kg일 수도 있는 것이어서 이는 유효숫자를 소수점 이하 몇 자리로 하느냐에 달려 있는 것이고, 소수점 이하의 숫자는 유효숫자의 제한에 따라 늘거나 줄 수가 있는 것이 연속변수의 또 다른 특징이라 할 수 있으며, 이에 반해 이산변수는 도로의 교통표지판에 나오는 어제의 교통사고 건수나 사망자, 부상자 수 등의 숫자는 1명, 15명 등과 같이 유효숫자의 범위와 관계없이 명확하게 셀 수 있는 정수로 끝나는 구조로 이루어져 있는 것이 특징이라 할 수 있다. 사망자 수가 1.5명이라든지 버스 안의 탑승자 수가 40.5명이라든지 하는 표현은 있을 수 없다는 의미이다.

표 4-1에서 연 수입, 나이, 주거주택의 평수, 부모 재산 등이 연속변수이며, 대상자들의 형제 수와 가족구성원 수, 주거지역별 대상자들의 분포 수, 본인과 부모의 직업별 분포를 숫자로 표현 시 이산변수가 된다.

#### (3) 질적변수

비수치적인 사실적 특성을 나타내는 변수로, 질적변수(qualitative variable)가 갖는 자료를 질적자료(qualitative data)라 하며, 범주형 자료(categorical data)라고도 한다(박스 4-1).

### 2) 척도의 분류

척도란 일정한 규칙을 가지고 관찰대상을 측정하기 위해서 그 속성을 일련의 기호 또는 숫자로 나타내는 것을 말한다. 척도는 명목척도, 순서척도, 등간척도, 비율척도로 분류한다.