


# 경영통계학



원광대학교 경영학부

담당교수: 정호일

# 제 2장

## 기술통계학 I : 자료의 정리 및 표현

---

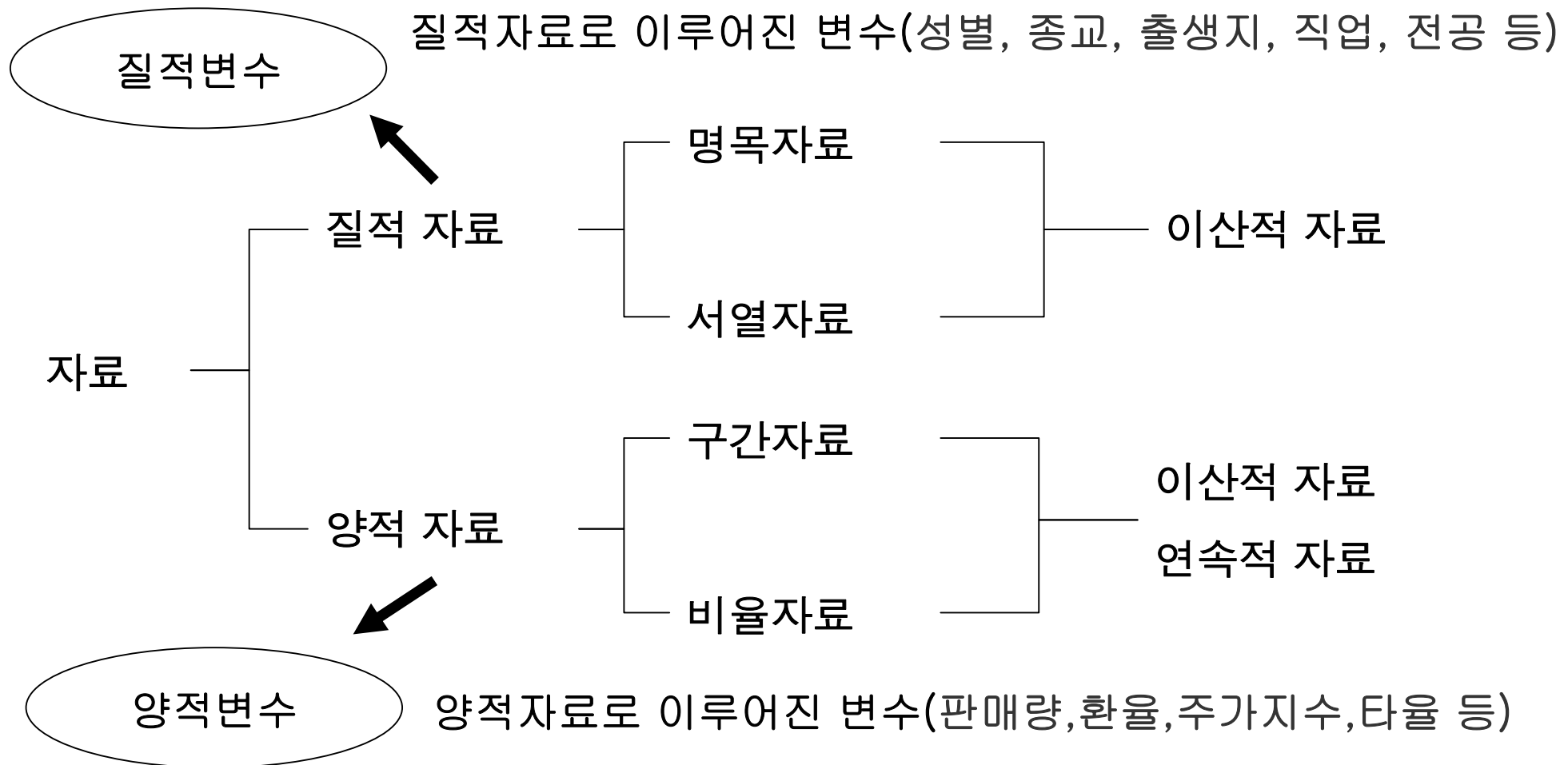
- 자료의 종류
- 측정척도의 형태
- 도수분포표
- 질적자료의 정리
- 양적자료의 정리
- 줄기-입 그림

# 1. 자료의 종류

---

- 기본용어
- ◆ 자료(data): 사람, 개체, 조건, 상황 등을 나타내는 기본적인 사실과 수치의 집합
- ◆ 기본단위(Elementary unit=요소): 조사대상
- ◆ 변수(Variable=변량): 관찰대상(사람, 기업, 사물 등)의 특성이 시간에 따라(시계열) 또는 특정 시점에서 관찰객체들에 따라(횡단면) 고정되지 않고 변하는 어떤 특성을 말한다.
- ◆ 관측치(Observation): 측정이나 관측을 통해 얻은 변수의 값

# • 자료의 형태



## •단변수 자료와 다변수자료

---

- 단변수 자료란 관측 대상이 되는 기본 단위가 하나의 변수만을 갖는 자료
- 다변수 자료란 관측 대상이 되는 기본 단위가 여러 개의 변수를 갖는 자료

# 2. 측정척도의 형태

## 측정의 정의

- 관찰 대상이 가지고 있는 특성에 대해서 일정한 규칙에 따라 기술적으로 수치를 부여함으로써 계량화하는 것을 의미함
- 개체가 가지고 있는 속성을 관찰하여 변수값으로 표현하는 것을 의미함

## 자료수집 과정

### 관심대상 선정

- 모집단을 결정하고 표본을 선정함

### 관심속성 파악

- 관찰대상으로부터 파악하고자 하는 속성을 정함

### 수집방법 결정

- 설문지, 관찰, 면접법...

### 측정

### 측정값의 기록

ID	성별	변수					브랜드선호도
		키	허리둘레	팔길이	몸무게		
1	1	173	28	83	65	3	
2	1	179	32	84	70	6	
3	1	175	30	85	68	7	
4	1	186	38	89	89	7	
5	1	175	32	86	70	3	
6	1	171	28	83	67	2	
7	1	179	32	87	76	5	
8	1	165	32	78	69	6	
9	1	183	34	87	76	5	
10	1	175	30	85	68	7	
11	1	186	38	89	89	7	
12	1	175	32	86	70	3	
13	1	171	28	83	67	2	
14	1	179	32	87	76	5	
15	1	165	32	78	69	6	
16	1	183	34	87	76	5	
17	1	175	30	85	68	7	
18	1	186	38	89	89	7	
19	1	175	32	86	70	3	
20	1	171	28	83	67	2	
21	1	179	32	87	76	5	
22	1	165	32	78	69	6	
23	1	183	34	87	76	5	
24	1	175	30	85	68	7	
25	1	186	38	89	89	7	
26	1	175	32	86	70	3	
27	1	171	28	83	67	2	
28	1	179	32	87	76	5	
29	1	165	32	78	69	6	
30	1	183	34	87	76	5	
31	1	175	30	85	68	7	
32	1	186	38	89	89	7	
33	1	175	32	86	70	3	
34	1	171	28	83	67	2	
35	1	179	32	87	76	5	
36	1	165	32	78	69	6	
37	1	183	34	87	76	5	
38	1	175	30	85	68	7	
39	1	186	38	89	89	7	
40	1	175	32	86	70	3	
41	1	171	28	83	67	2	
42	1	179	32	87	76	5	
43	1	165	32	78	69	6	
44	1	183	34	87	76	5	
45	1	175	30	85	68	7	
46	1	186	38	89	89	7	
47	1	175	32	86	70	3	
48	1	171	28	83	67	2	
49	1	179	32	87	76	5	
50	1	165	32	78	69	6	

### 3. 측정척도의 이해

---

- 측정되어진 변수의 값이 숫자의 특성을 얼마나 가지고 있는가에 따라서 구분되어지는 개념이다.
- 변수가 가지는 척도의 형태에 따라 자료분석 방법이 결정되어진다.

- 1)명목변수(척도):성별, 지역
- 2)서열변수(척도):학력, 석차
- 3)등간변수(척도):온도, 주가지수
- 4)비율변수(척도):강도, 길이, 무게, 수익률

## 1) 명목척도(nominal scale)

---

- 측정대상의 특성을 확인하거나 분류할 목적으로 수치를 부여하는 경우이다.
- 양적인 크기나 산술적인 계산이 의미가 없음.
- 측정대상을 상호배타적인 집단으로 구분하는데 이용 즉, 모든 측정대상을 어느 한 집단으로 분류할 수 있어야 한다. 동일한 집단 속에 들어있는 대상은 동일한 척도값을 가져야 한다.
- 포괄성(exhaustive) : 모든 분류대상이 전부 포함될 수 있도록 범주를 정함.
- 상호배타성(mutual exclusive) : 하나의 대상이 두 개 이상의 범주에 속하지 않도록 분류



# 명목척도의 예

---

- 성별조사
- 종교조사
- 판매지역
- 상표인지여부
- 직업구분
- 회사유형, 경영형태 등

ex) 당신은 어느 정당을 지지하십니까?

- ① 신한국당 ② 국민회의 ③ 자민련 ④ 국민신당 ⑤ 없음

## 2) 서열척도(ordinal scale)

---

- 측정대상간의 순서관계를 밝혀주는 척도로서, 측정 대상의 특정한 속성을 파악하여 **측정대상간의 대소나 높고 낮음 등의 순위를** 부여해주는 척도이다.
- 양적인 비교를 할 수 있는 정보를 제공하지 못한다.
- 정확하게 정량화 하기 어려운 응답자의 태도, 선호도, 사회계층, 학력, 석차 등의 측정에 이용된다.

# 서열척도의 예

---

- 학년의 구분-학년별 구분, 학업정도의 순위
- 성적순위-등수 별 집단구분, 누가 더 잘하는지 순위구분가능 하지만 순위간의 차이가 어느 정도인지는 알 수 없음
- 집중경향 측정, 가감승제 무의미
  
- ex) 당신은 현재 우리 나라의 경제사정을 어떻게 생각하십니까? ① 매우 호황 ② 호황 ③ 보통이다 ④ 불황 ⑤ 매우 불황

### 3) 등간척도(interval scale)

---

- 속성에 대한 순위를 부여하되 순위사이의 간격이 동일한 척도를 말한다.
- 즉 등간척도에 의해서 얻어진 측정치는 측정대상 양적인 정도의 차이를 나타냄
- 절대적인 원점은 존재하지 않지만 상대적인 원점은 존재한다.
- ※ 조건 : 총망라성, 상호배타성, 대수관계, 부가성

# 등간척도의 예

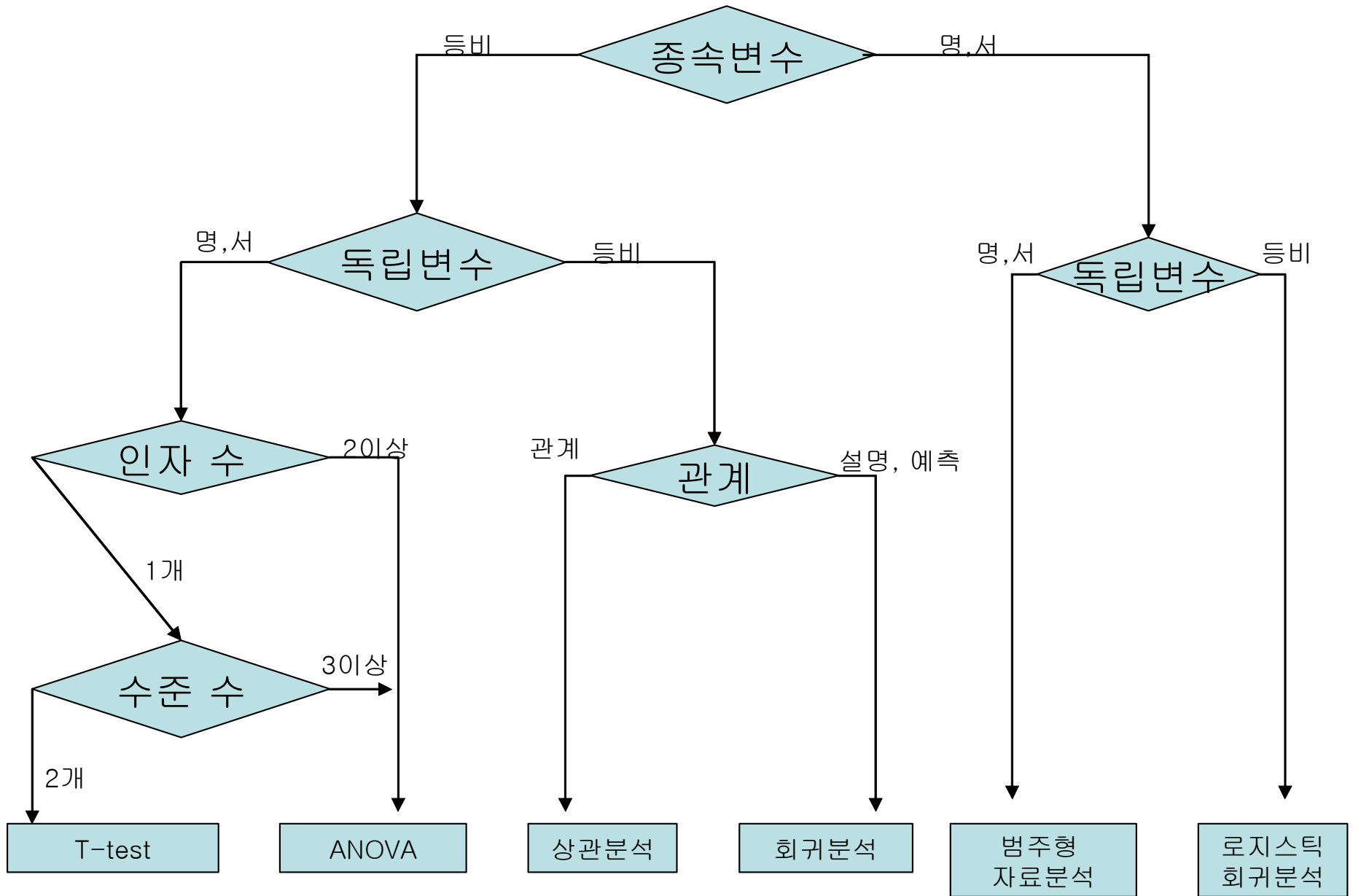
---

- 수치들간의 차이가 어느 정도인지 알 수 있다.
- 온도 40도와 80도 80도가 40도의 2배인가? 40도 더 높은 것이지 2배 더 높은 것은 아니다.
- ex) 귀하의 지난해 인사고과 평점은 어디에 속하십니까? ① 4.0 ② 3.0 ③ 2.0 ④ 1.0

## 4) 비율척도(ratio scale)

---

- 등간척도가 갖는 특성에 추가적으로 측정값 사이의 비율 계산이 가능한 척도이다.
- 비율계산이 가능한 이유는 측정하고자 하는 속성이 전혀 존재하지 않는 절대영점이 존재하기 때문이다.
- ※ 조건 : 총망라성, 상호배타성, 대소관계, 부가성, 두 개 척도지수간 동일 비율
- 계산된 비율이 제의미를 지닌다.(사칙연산이 의미있음)
- 판매량, 구매자수, 구매확률, 무게 등



## •자료의 비교

비율(범주+서열+등간격+절대영점)

구간(범주+서열+등간격+상대영점)

서열(범주+서열)

명목(범주)



# 3. 도수분포표

---

- 기본개념

도수분포표(Frequency distribution table)란 수집된 자료를 몇 개의 계급으로 나누고 각 계급에 해당되는 관측치의 도수를 기록하는 통계표이다.

-질적자료와 양적자료 모두 표현할 수 있음.

## •도수분포표의 그래프

---

- 질적 자료를 나타내는 그래프: 막대그래프와 원 그래프
- 양적 자료를 나타내는 그래프: 히스토그램, 꺾은선 그래프, 누적백분율곡선이 있다.

# 4. 질적자료의 정리

---

- 도수분포표

상대도수(relative frequency)란 각 계급의 도수(관측치)가 총도수에서 차지하는 비율을 말한다.

$$\text{상대도수} = \frac{f_i}{n}$$

$$\text{상대도수(백분율)} = \frac{f_i}{n} \times 100$$

도수=빈도

$f_i$  : 어떤 계급  $i$ 의 도수

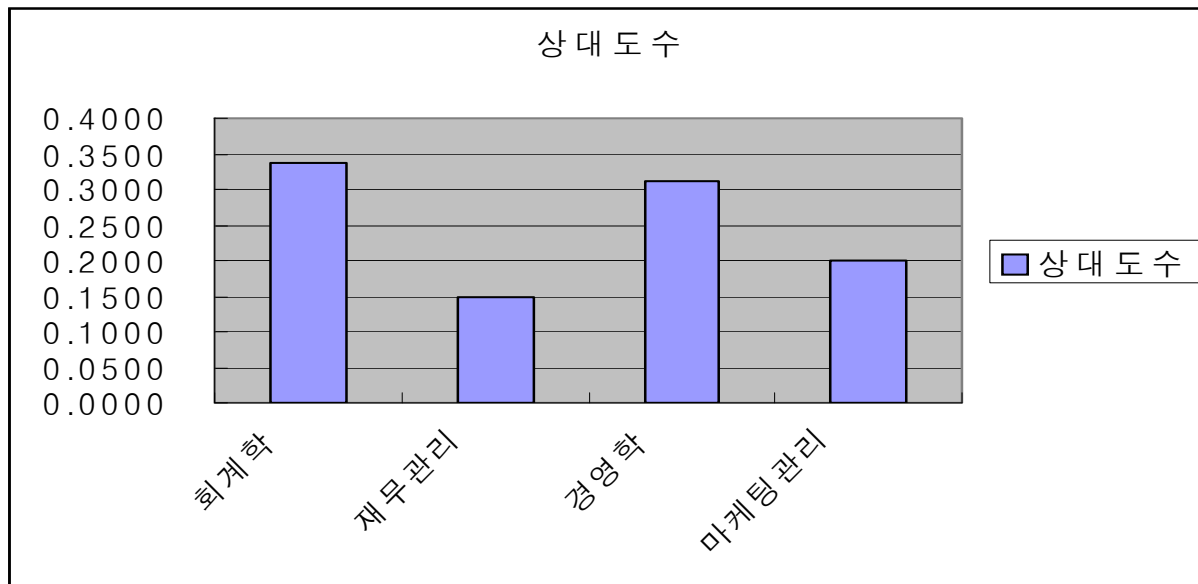
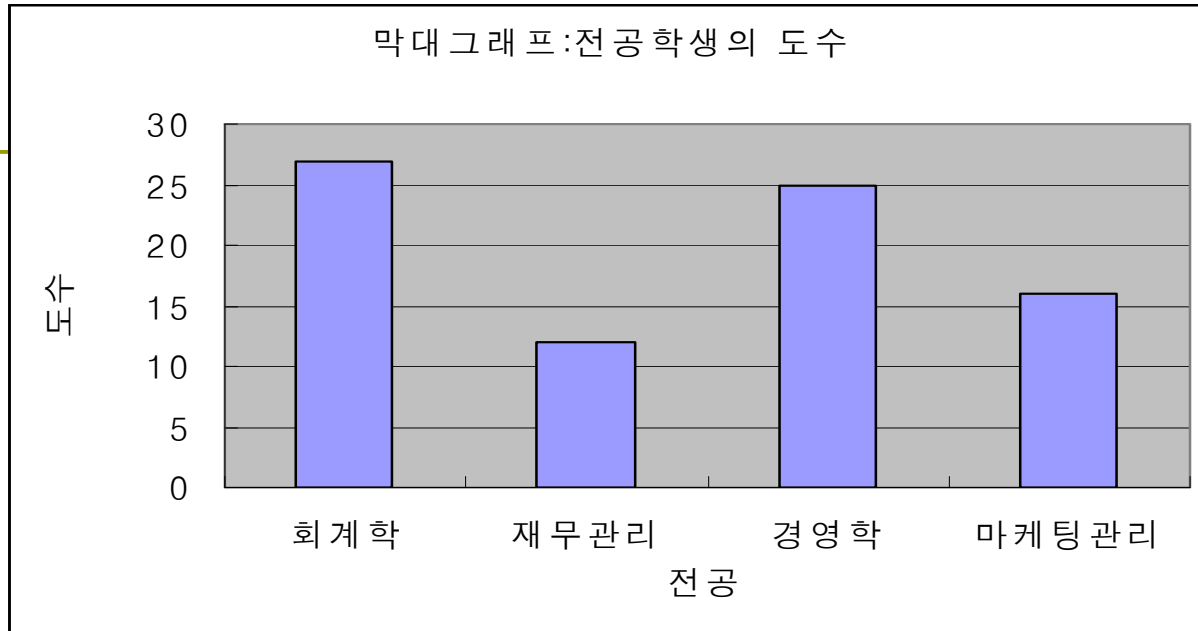
$n$  : 총도수

## •도수분포표의 그래프

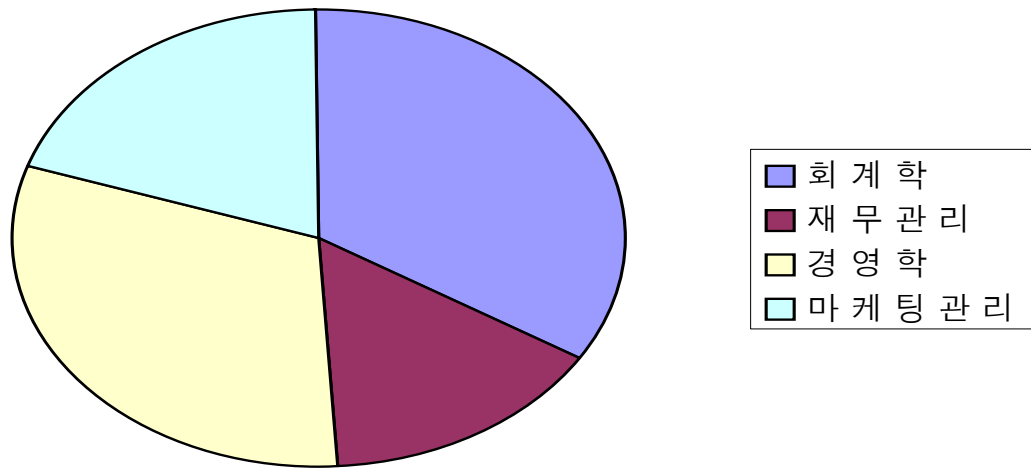
질적자료에 대해 도수분포표를 작성하면 이를 이용하여 히스토그램과 원그래프를 그릴 수 있다.

[표 2-3]

전공	도수	상대도수	상대도수(%)
회계학	27	0.3375	33.75
재무관리	12	0.1500	15.00
경영학	25	0.3125	31.25
마케팅관리	16	0.2000	20.00
합계	80	1.000	100.00



전공 학생의 원그래프



# 5. 양적 자료의 정리

---

## •도수분포표:이산적 자료

- 이산적 자료의 도수분포표 작성방법은 질적자료의 경우와 동일하나 수치가 어떤 의미를 가지기 때문에 계급은 수치적 순서를 지켜야 한다.

- 이산적 자료를 이용해서는 상대도수 뿐만 아니라 질적자료에서 구할 수 없는 누적도수와 누적상대도수를 계산할 수 있다.

- 누적도수는 어느 특정계급의 이하 또는 이상에 해당되는 도수를 누적하여 구한다.

- 누적상대도수는 상대도수를 누적하여 구한다.

## •도수분포표의 그래프:이산적자료

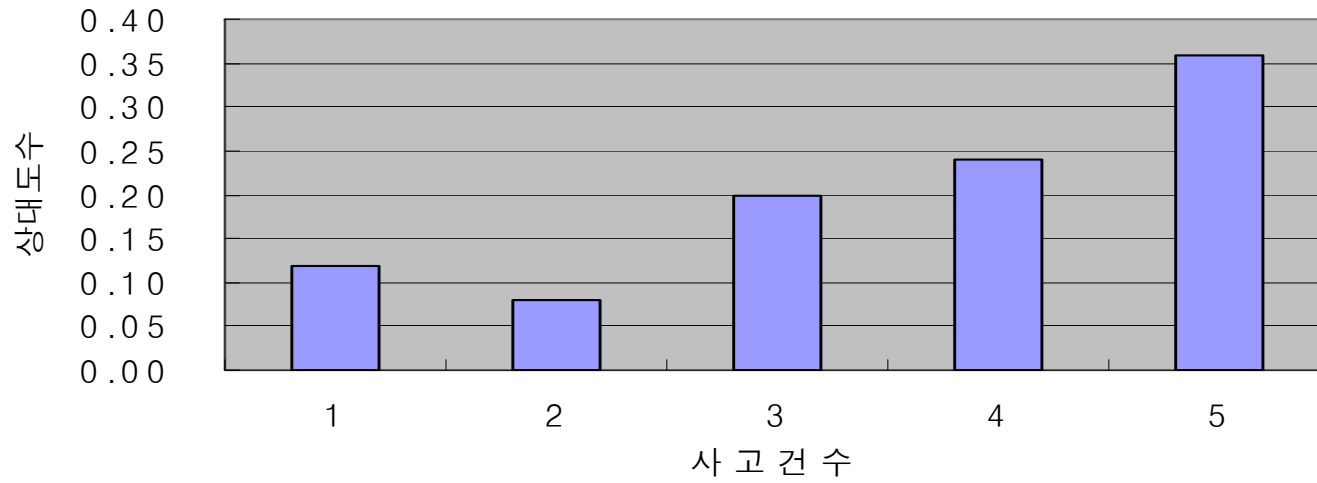
---

[표 2-6]

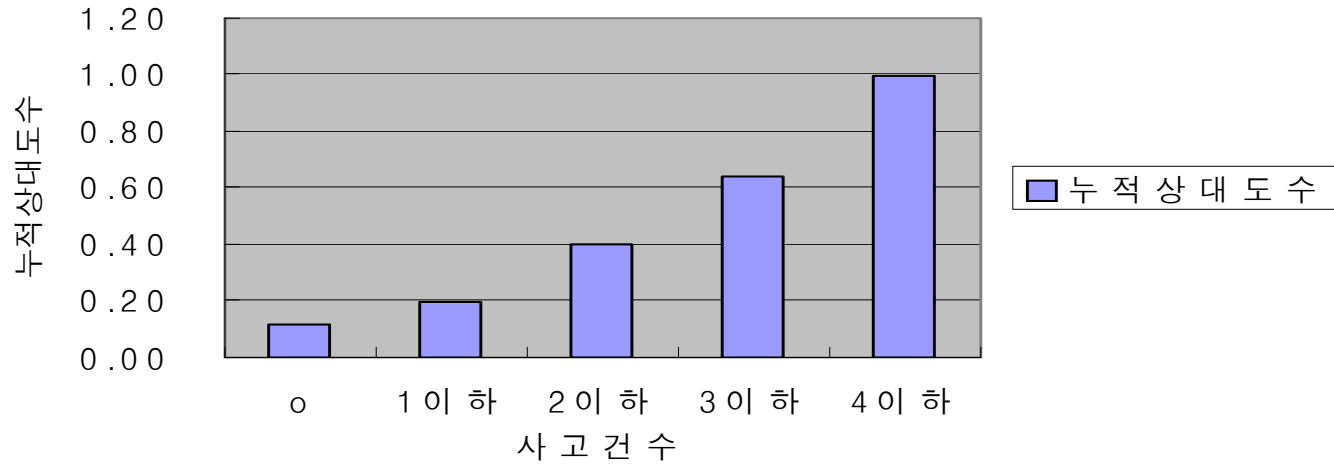
사고건수	도수	상대도수	누적도수	누적상대도수
0	3	0.12	3	0.12
1	2	0.08	5	0.20
2	5	0.20	10	0.40
3	6	0.24	16	0.64
4	9	0.36	25	1.00
합계	25	1.00	59	2.36



막대 그래프 : 사고건수의 상대도수



막대 그래프 : 누적상대도수



## •도수분포표: 연속적자료

---

1. 자료의 범위 결정(최대값-최소값)
2. 계급의 수 결정: 주관적으로 결정(보통 5~15개 이내)
3. 계급구간 (폭)결정:가능한 정수이면서 홀수의 폭이 되도록-반올림 활용)-극단적 수치는 개방구간을 활용
4. 계급의 한계(경계값) 결정-계급의 한계가 중복되지 않도록 정확한 한계(exact limit)를 결정
5. 각 계급의 도수, 상대도수, 누적도수, 누적 상대도수를 구한다.

$$\text{계급구간} = \frac{\text{범위}}{\text{계급의 수}} = \frac{\text{자료의 최대치} - \text{자료의 최소치}}{\text{계급의 수}}$$

## •도수분포표의 그래프:연속적자료

---

연속적자료를 이용하여 도수분포표를 만들고 다시 그래프로 표현하는 데는 다음 3가지를 그릴 수 있다.

- 히스토그램
  - 꺾은선 그래프
  - 누적백분율곡선
- 을 그릴 수 있다.

