

3. 수치형 자료의 요약

이산형 자료 요약

관측된 수치형 자료가 셀 수 있는 경우 ▶ 이산형 자료 요약

관측값의 종류 수



범주형 자료 요약 기법



연속형 자료 요약 기법

연속형 자료 요약

수치형 자료가 연속적으로 관측 ▶ 연속형 자료 요약

관측값의 종류 수

↑
많음

연속형 자료 요약 기법

↓
적음

점도표, 도수분포표, 히스토그램,
상대도수다각형, 줄기-잎 그림

점도표 (dot diagram)

관측값의 개수가 상대적으로 적은 경우(20 또는 25이하) 사용

- 자료 전체의 개요를 파악 가능
- 모든 자료를 나타낼 수 있도록 줄 위에 각 관측값에 해당되는 점을 찍어 표시

연속형 자료의 경우 **중복된**
정보를 판단하기 어려움



자료를 크기에 따라 묶어서
분석하는 것이 **효율적**

도수분포표(Frequency Table)

- 각 관측값에 대한 도수를 측정하여 도수분포표 작성
- 연속형 자료의 경우 다수의 구간(계급)으로 나누고
각 구간마다 관측값의 개수(도수)로 작성

계급(Class)	계급구간	계급구간의 폭
위에서 나눈 구간	각 계급에 포함되는 값의 범위	계급구간의 크기

도수분포표 작성 순서

1. 자료의 범위

자료에서 최대값과 최소값을 찾아 자료의 범위를 구함

2. 계급의 폭

계급의 개수를 분포의 경향이 잘 드러날 수 있도록 정함

3. 계급구간

모든 관측 값을 포함하도록 계급구간의 경계점을 구함

4. 도수

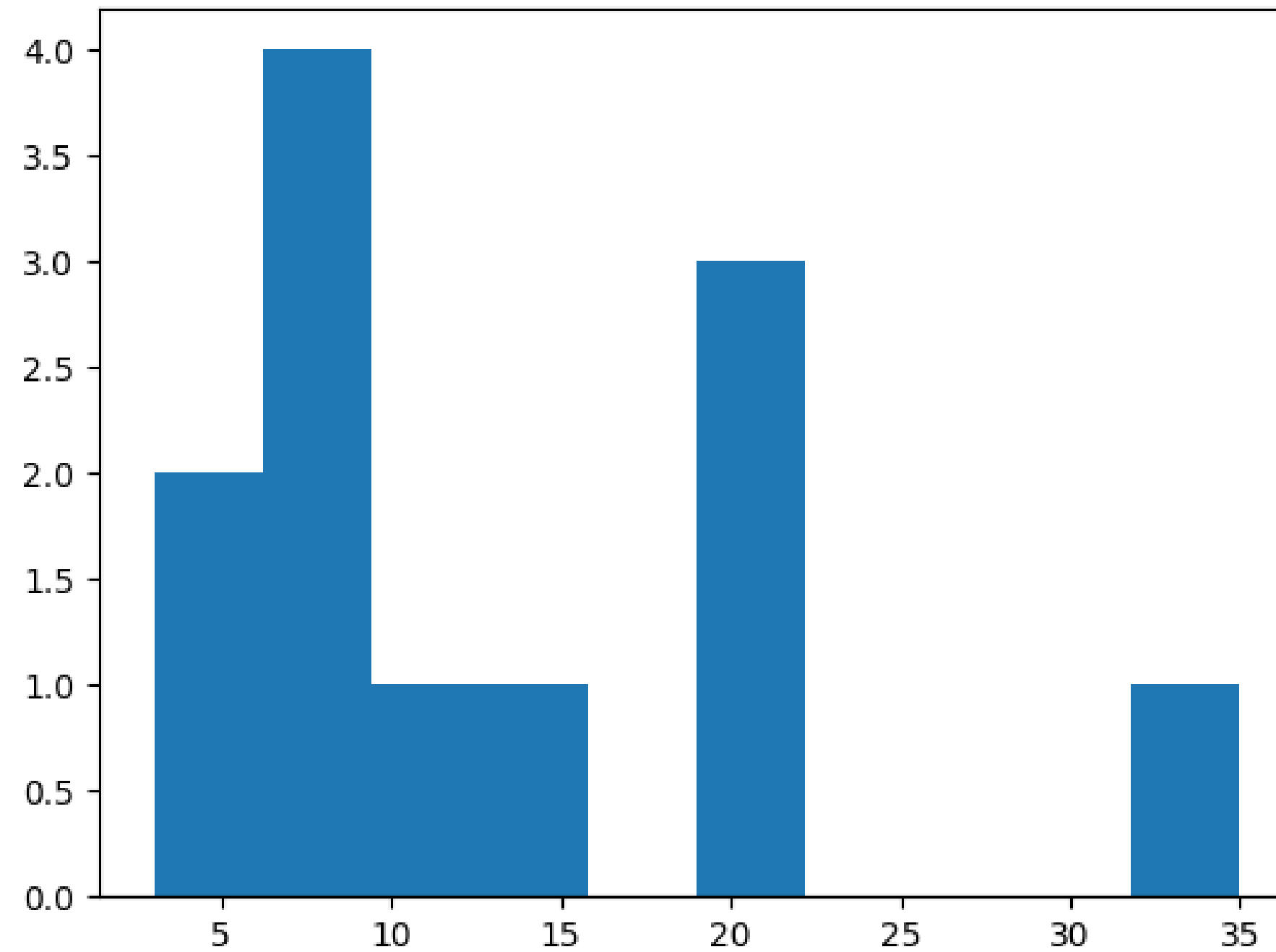
각 계급구간에 속하는 관측값의 개수를 세어 계급의 도수를 더함

5. 상대도수

각 계급의 도수를 전체 관측값의 개수로 나눠 계급의 상대도수를 구함

히스토그램(Histogram)

Y 축 : 빈도



X 축 : 계급

히스토그램(Histogram)

```
plt.hist()
```

연속형 자료의 도수분포표를 기반으로 각 계급을 범주처럼 사용
범주형 자료의 **막대그래프**와 같은 방식으로 그림

도수 비교



범주

막대그래프

연속

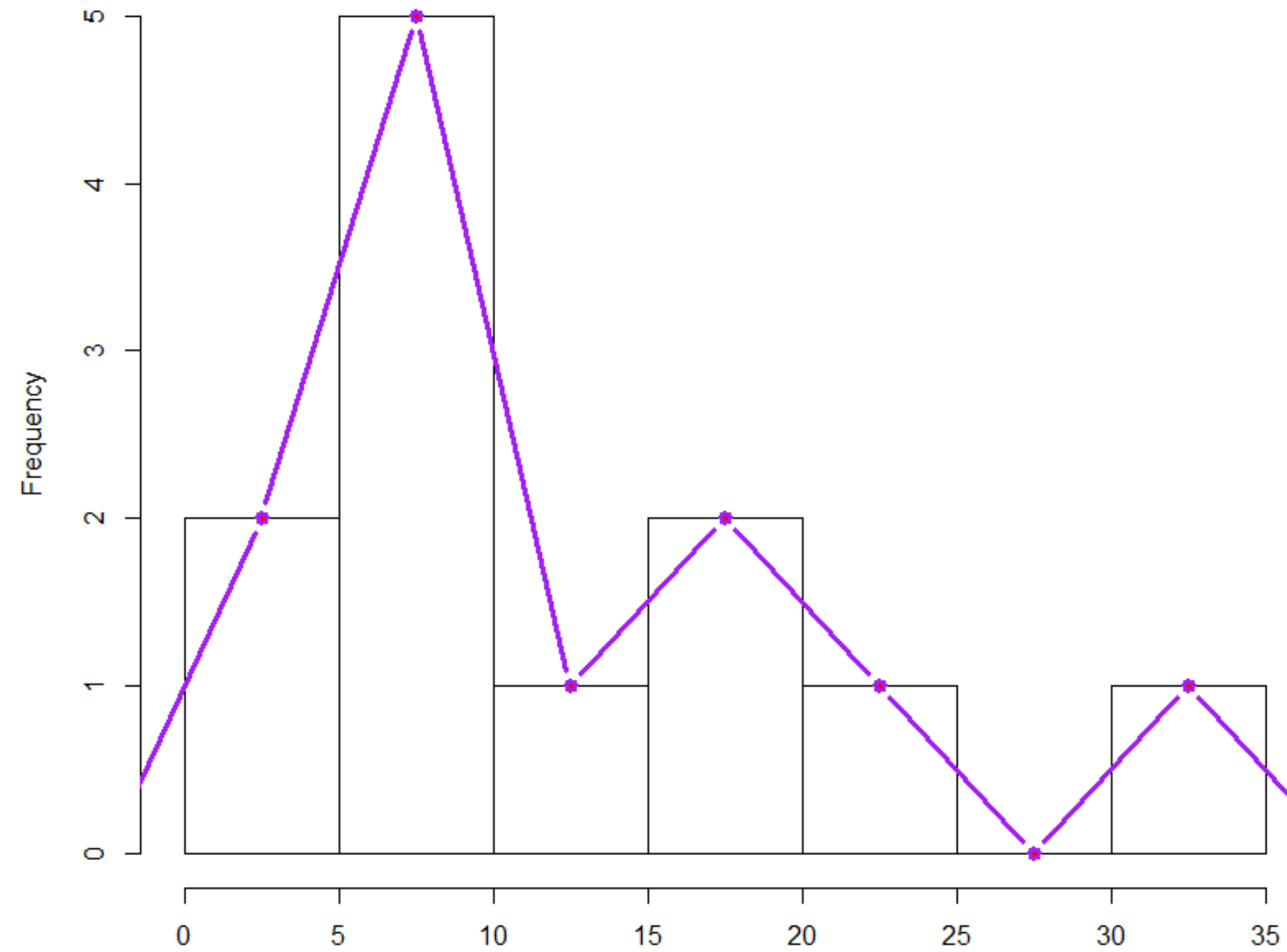
히스토그램

히스토그램의 특징

- 자료의 분포를 알 수 있음
- 계급구간과 막대의 높이로 그림
- 히스토그램의 전체 면적은 1
- 각 계급(막대)간 빈 공간이 없어야 함
- 모든 계급구간의 폭이 같으면
도수, 상대도수를 막대 높이로 사용

도수다각형

Y 축 : 빈도



X 축 : 구간

도수다각형의 특징

- 각 계급구간의 중앙에 점을 찍어 직선으로 연결함
- 관측값의 집중된 위치, 정도, 치우친 정도, 꼬리의 두터움 등 분포의 상태를 쉽게 파악
- 관측값의 변화에 따라 도수 또는 상대도수의 변화를 잘 나타냄

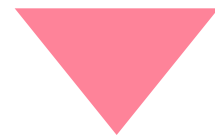
도수다각형과 히스토그램

히스토그램

옆으로 나열하여
자료 비교

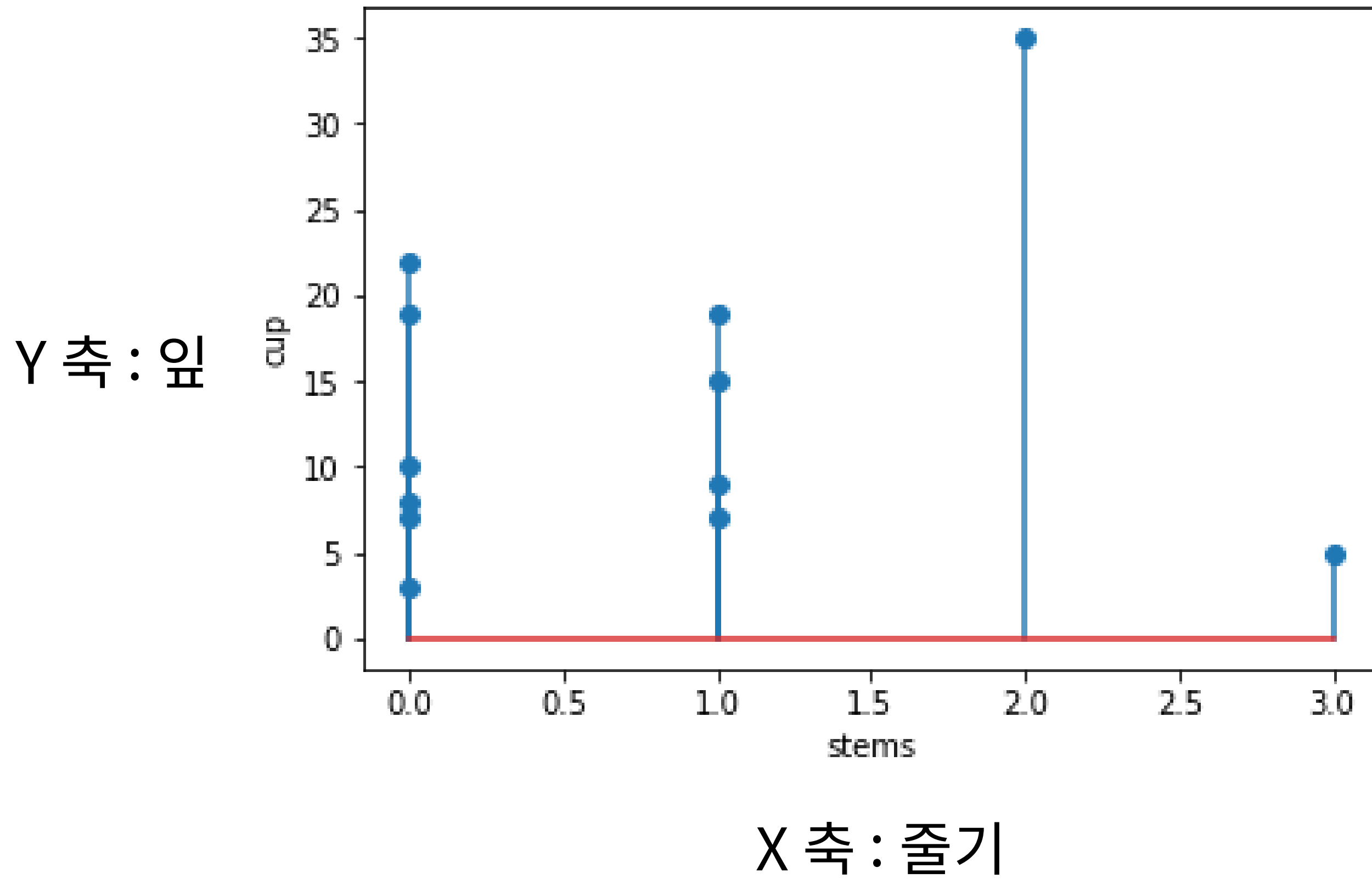
도수다각형

꺾은선으로 표시하여
자료 비교



여러 자료를 비교하기 위해서는
도수다각형이 히스토그램보다 알맞음

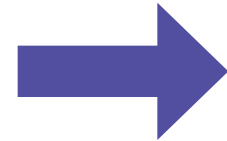
줄기-잎 그림



줄기-잎 그림 자료 정리 방법

예) 관측값의 자릿수가 최대 2자리 일때

관측값 35



앞 단위 : 3



뒷 단위 : 5

1. 관측값을 보고 앞 단위와 뒷 단위를 정함
2. 앞 단위를 줄기로 하여 세로로 배열하고 수직선을 그림
3. 뒷 단위를 잎으로 하여 관측값을 앞 단위 오른쪽에 오름차순 기입

줄기-잎 그림 자료 정리 예시

예) 22,7,19,3,10,8,19,7,15,9,35,5

줄기 | 잎

-- | --

0 | 3 5 7 7 8 9

1 | 0 5 9 9

2 | 2

3 | 5

줄기-잎 그림 함수

```
plt.stem(줄기, 관측 값)
```

자료의 분포를 시각적으로 쉽게 파악

각 관측값도 유지 가능

함수 사용시에 줄기 값을 따로 지정해줘야 함
: 줄기를 데이터마다 다르게 설정할 수 있기 때문

줄기-잎 그림 장단점

장점

- 관측값을 보여주므로 최댓값, 최솟값 등의 위치 파악 쉬움
- 순서대로 배열된 관측값의 장점과 히스토그램의 장점을 모두 가지고 있음
- 그리기 쉬움

단점

- 관측값의 개수가 많은 경우 제한된 공간에 그리기 불가능
- 관측값이 지나치게 흩어져 있으면 부적절