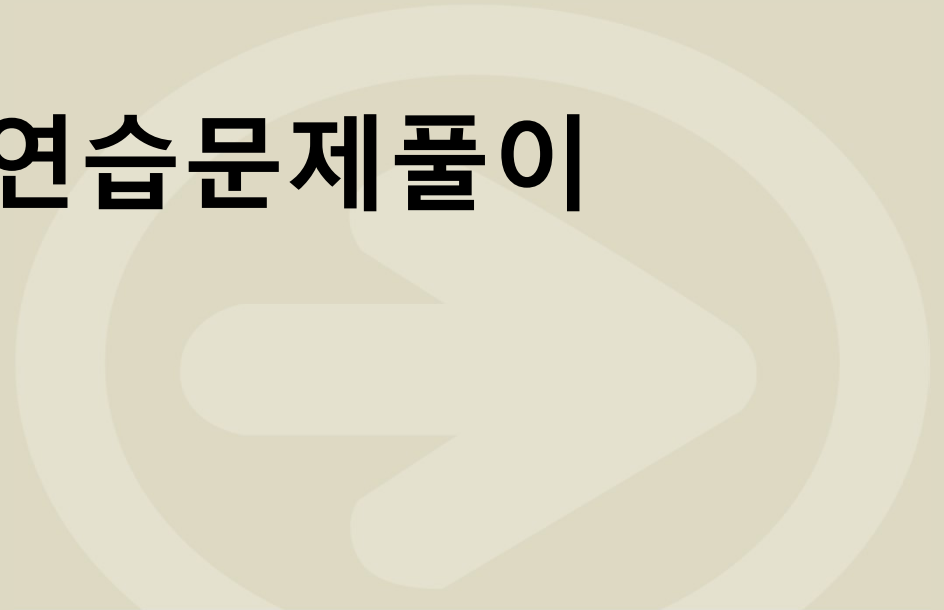


4주2강. 복습 및 연습문제풀이



줄기와 잎 그림

- ❑ **줄기와 잎 그림(Stem-and-Leaf Plot)**은 숫자로 관측된 양적 자료를 정리하는 방법으로 막대그림표와 유사하나 막대그림표에서는 얻을 수 없는 정보인 자료의 최소값, 최대값 그리고 각 구간 내부에 있어서의 자료의 분포에 대한 정확한 정보를 제공해 준다.
- ❑ 막대그림표로 그렸을 때의 분포를 나타내는데 막대그림표가 각 구간 내 자료의 형태에 대한 정보를 제공하지 못하는 데 비하여 줄기와 잎 그림은 자료의 요약에 따른 정보의 손실이 전혀 없다.

줄기와 앞 그림 작성방법

1. 관측값의 숫자 단위(1단위, 10단위, 100단위, ...)를 이용하여 숫자를 두 부분으로 나누어 앞 부분은 줄기로, 그리고 뒷 부분은 앞으로 한다.
2. 줄기의 숫자를 작은 것부터 크기 순서에 따라 열(columnize)로 나열한다.
3. 각 관측값을 그 숫자가 속한 위치의 줄기에 맞추어 앞 부분을 기록한다.
4. 만약 각 줄기에 너무 많은 관측값이 주어지면 각 줄기에 두 줄을 할당하여 첫 줄에는 앞의 0, 1, 2, 3, 4를 기록하고, 둘째 줄에는 앞의 5, 6, 7, 8, 9를 기록한다.
5. 각 줄기 내의 앞의 값들은 작은 것부터 크기 순서로 정리한다.

표 2-3

통계학 과목 수강생의 학기말 성적

65	62	73	85	65	46	36	49	81	76
60	44	43	72	21	33	83	46	64	49
12	74	91	78	60	48	24	62	54	97
69	31	89	96	96	97	86	88	85	61
95	54	85	89	51	77	81	72	47	35

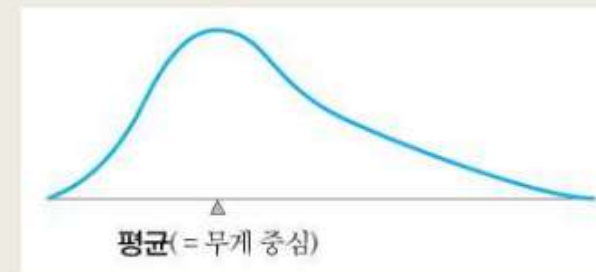
그림 2-11 통계학 과목 수강생 성적의 줄기와 잎 그림

도수	줄기(stem)	잎(leaf)
1	1	2
2	2	1 4
4	3	1 3 5 6
8	4	3 4 6 6 7 8 9 9
3	5	1 4 4
9	6	0 0 1 2 2 4 5 5 9
7	7	2 2 3 4 6 7 8
10	8	1 1 3 5 5 5 6 8 9 9
6	9	1 5 6 6 7 7

자료의 중심을 측정하는 통계량

□ (표본)평균 (mean ; \bar{X})

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$



□ 중위수 (median ; \tilde{X})

- 자료를 크기 순서로 정리했을 때 가운데에 위치하는 관측값
- 소수의 크거나 작은 관측값에 의해 영향을 받지 않음

□ 최빈값 (mode ; M_0)

- 관측 도수가 가장 많은 값

평균의 특징

1. 자료 관측값의 산술평균이다.
2. 각 자료에 있어서 유일하게 구하여진다.
3. 소수의 매우 크거나 작은 값에 의하여 영향을 받는다.
4. 자료를 몇 개의 작은 집단으로 나누었을 때 각 집단의 평균의 평균은 전체 자료를 이용하여 구한 평균과 같다.
5. 수량으로 관측된 자료에만 이용 가능하다.

중위수를 구하는 방법

1. n 이 홀수인 경우

- $\frac{n+1}{2}$ 의 n 개의 가운데 위치이므로 관측값 중에서 순서상 $\frac{n+1}{2}$ 번째 값이 중위수이다.
- 예를 들어 $n = 7$ 인 경우 $\frac{7+1}{2} = 4$ 이므로 관측값을 크기순서로 나열하였을 때, 4번째 순서에 있는 값이 중위수이다.

2. n 이 짝수인 경우

- 짝수인 경우에 있어서 가운데 위치는 $\frac{n}{2} + 0.5$ 가 된다. 이 경우에는 가운데 위치에 가장 가까운 값인 $\frac{n}{2}$ 번째 순서값과 $\frac{n}{2} + 1$ 번째 순서값을 구하여 두 값의 평균을 중위수로 한다. 즉, $\tilde{x} = \frac{1}{2}(x_{\frac{n}{2}} + x_{\frac{n}{2}+1})$ 이다.
- 예를 들어 $n = 10$ 인 경우 가운데 위치는 5.5이며, 중위수는 $\frac{10}{2} = 5$ 번째 순서값과 $\frac{10}{2} + 1 = 6$ 번째 순서값의 평균이 된다.

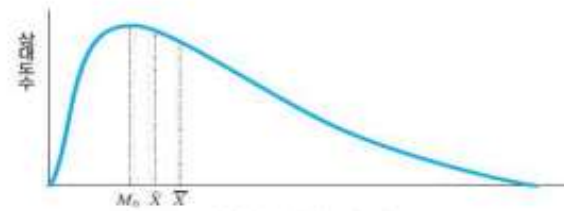
중위수의 특징

1. 중앙위치의 값으로 관측값의 50%가 왼쪽에, 그리고 나머지 50%가 오른쪽에 존재한다.
2. 각 자료에 있어서 유일하게 구하여진다.
3. 소수의 매우 크거나 작은 값에 의하여 영향을 받지 않는다.
4. 자료를 몇 개의 작은 집단으로 나누었을 때 각 집단의 중위수의 중위수는 전체자료를 이용하여 구한 중위수와 항상 일치하지는 않는다.
5. 수량으로 관측된 자료에만 이용 가능하다.

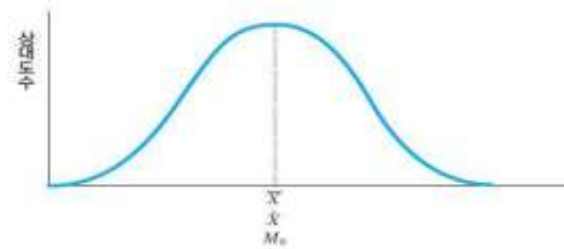
최빈값의 특징

1. 자료에서 관측빈도의 수가 가장 많은 값이다.
2. 각 자료에는 하나 이상의 최빈값이 있을 수 있다.
3. 소수의 극한값에 영향을 받지 않는다.
4. 자료를 몇 개의 작은 집단으로 나누었을 때 각 집단의 최빈값에 의하여 전체의 최빈값을 유도할 수 없다.
5. 양적으로 측정된 자료와 질적으로 측정된 자료 모두에 이용 가능하다.

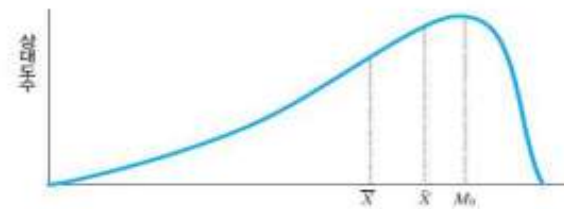
그림 2-12 자료의 분포모양에 따른 평균(\bar{X}), 중위수(\tilde{X}), 최빈값(M_o)의 위치



(a) 오른쪽 긴 꼬리 자료



(b) 대칭인 자료



(c) 왼쪽 긴 꼬리 자료

범위 (range ; R)

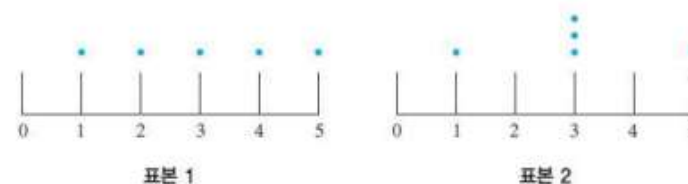
- 관측된 자료들 중에서 가장 큰 값과 가장 작은 값의 차이
- 자료 $\{X_1, X_2, \dots, X_n\}$ 중에서 가장 작은 값 X_i , 가장 큰 값 $X_j \Rightarrow R = X_j - X_i$
- 표에 나타난 두 자료는 평균과 중위수가 모두 3으로 동일하고 범위 $R = 5 - 1 = 4$ 도 서로 동일함을 알 수 있다.
- 그러나 자료 1과 자료 2는 분포의 형태가 전혀 상이함을 알 수 있다. 자료 1 과 자료 2 의 범위는 동일하나 자료 1 이 자료 2에 비하여 흩어진 정도가 더 크다고 판단할 수 있다.
- 범위는 쉽게 구할 수 있는 통계량이나 많은 정보를 제공하지 못함

표 2-4

두 표본의 분포표

	관측값	평균	중위수
자료 1	1, 2, 3, 4, 5	3	3
자료 2	1, 3, 3, 3, 5	3	3

그림 2-14 두 표본의 분포



순서통계량

- ❑ 아래사분위수 (lower quartile; Q_1) : 관측값의 25% 순서에 있는 값
- ❑ 중위수(median) : 관측값의 50% 순서에 있는 값
- ❑ 위사분위수 (upper quartile; Q_3) 는 관측값의 75% 순서에 있는 값
- ❑ **사분위 범위** : 위사분위수 Q_3 와 아래사분위수 Q_1 의 차이 ($Q_3 - Q_1$)을 사분위범위(IQR)라고 하며, 자료의 흩어진 정도를 측정하는 또 다른 통계량으로 이용된다.
- ❑ **중위수와 사분위수의 위치 결정 공식** : n 개의 관측값 $\{X_1, X_2, \dots, X_n\}$ 이 작은 값으로부터 올림 차순으로 정리되어 있을 때 중위수와 사분위수의 위치는 다음 공식에 의해 구한다.
 - 중위수의 위치 = $\frac{n+1}{2}$
 - Q_1 의 위치 = $\frac{[\text{중위수의 위치}] + 1}{2}$
 - Q_3 의 위치 = $[\text{중위수의 위치}] + \frac{[\text{중위수의 위치}] + 1}{2}$

2장 25. 다음 자료는 한 회사의 입사시험에 응시한 32명의 적성시험 성적이다.

79	97	86	76	93	87	98	68
84	88	81	91	86	87	70	94
77	92	66	85	63	68	98	88
46	72	59	79	36	72	68	82

- (a) 줄기 잎 그림을 그려라.
- (b) 평균, 중위수, 최빈값을 구하여라.
- (c) 범위를 구하여라.
- (d) 아래사분위수와 위사분위수를 구하여라.

3장 2. $S = \{\alpha, \beta, \gamma, \omega, \delta, \varepsilon\}$ 라 할 때, S 의 부분집합을 $A = \{\alpha, \beta, \gamma\}$, $B = \{\alpha, \gamma, \delta\}$, $C = \{\omega, \varepsilon\}$ 라 하자.

(a) $A \cup B, A \cap B, A \cup C, A \cap C$ 를 구하여라.

풀이) $A \cup B = \{\alpha, \beta, \gamma, \delta\}$, $A \cap B = \{\alpha, \gamma\}$, $A \cup C = \{\alpha, \beta, \gamma, \omega, \varepsilon\}$, $A \cap C = \emptyset$

(b) C^c 를 구하여라.

풀이) $C^c = \{\alpha, \beta, \gamma, \delta\}$

(c) $A - B$ 를 구하여라.

풀이) $A - B = \{\beta\}$

(d) 세 부분집합 A, B, C 중 상호 배반집합인 쌍을 구하여라.

풀이) $A \cap B = \{\alpha, \gamma\}$, $A \cap C = \emptyset$, $B \cap C = \emptyset$ 이므로 A 와 C , B 와 C 가 상호배반이다.

3장 12. 5명으로 구성된 시의회에서 교통문제를 전담할 2명의 위원을 선출한다고 한다. 5명의 시의회 의원을 A, B, C, D, E 라 할 때, 다음 질문에 답하여라.

(a) 표본공간 Ω 을 구하여라.

풀이) $\Omega = \{AB, AC, AD, AE, BC, BD, BE, CD, CE, DE\}$

(b) 위원회에 시의원 A 가 반드시 포함될 확률은 얼마인가?

풀이) A 가 포함된 경우가 4개이므로 확률은 $\frac{4}{10} = \frac{2}{5}$

(c) 시의원 B 와 C 가 위원으로 선출될 확률은 얼마인가?

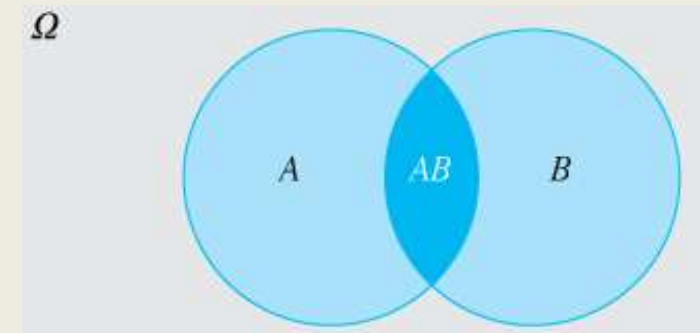
풀이) B 와 C 가 포함된 경우는 1개이므로 확률은 $\frac{1}{10}$

조건부 확률

- 사건 A 가 주어졌을 때 사건 B 의 조건부 확률은 $P(A) \neq 0$ 이라는 전제하에서 다음과 같이 정의한다.

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

- 사건 A 와 B 를 벤 다이어그램으로 그리면 다음과 같다.



- 즉, 조건부 확률은 사건 A 의 크기에 대한 곱사건 AB 의 크기의 비율이라고 할 수 있다. $P(B|A)$ 는 실험에서 나타나는 결과를 A 에 한정할 때 사건 B 가 나타날 가능성을 측정하는 것이다.

3장 17. 주사위 2개를 던지는 실험에서 두 사건 A, B 를 다음과 같이 정의할 때, 다음 확률을 구하여라.

$A : \{\text{두 주사위의 합이 짝수인 경우}\}, B : \{\text{첫째 주사위의 값이 3 이하인 경우}\}$

$\Omega = \{(1, 1), (1, 2), (1, 3), \dots, (6, 6)\}$ 모두 36가지 경우

(a) $P(A), P(B), P(B^C)$

풀이) $P(A) = \frac{1}{2}$: 합이 짝수인 경우는 18가지

$P(B) = \frac{1}{2}$: 첫 번째 주사위가 3이하인 경우는 18가지,

$$P(B^C) = \frac{1}{2}$$

(b) $P(A \cup B), P(AB)$

풀이) 첫 번째 주사위가 3이하인 경우는 18가지,

첫 번째 주사위가 4이상이고 합이 짝수인 경우는 9가지

첫 번째 주사위가 3이하이고 합이 짝수인 경우는 9가지

$A \cup B$ 는 $18+9=27$ 가지

$$\therefore P(A \cup B) = \frac{27}{36} = \frac{3}{4}, P(AB) = \frac{9}{36} = \frac{1}{4}$$

3장 17. 주사위 2개를 던지는 실험에서 두 사건 A, B 를 다음과 같이 정의할 때, 다음 확률을 구하여라.

$A : \{\text{두 주사위의 합이 짝수인 경우}\}, B : \{\text{첫째 주사위의 값이 3 이하인 경우}\}$

$\Omega = \{(1, 1), (1, 2), (1, 3), \dots, (6, 6)\}$ 모두 36가지 경우

(c) $P(A | B), P(B | A)$

풀이) $P(A | B) = \frac{P(AB)}{P(B)} = \frac{1/4}{1/2} = \frac{1}{2}, P(B | A) = \frac{P(AB)}{P(A)} = \frac{1/4}{1/2} = \frac{1}{2}$

(d) $P(AB^C), P(A^C B)$

풀이) $P(AB^C) = P(AB) = \frac{1}{4}, P(A^C B) = P(AB) = \frac{1}{4}$

이 경우는 AB^C 와 AB 가 서로 대칭이며, $A^C B$ 와 AB 도 또한 서로 대칭이다.

3장 21. 어린이 들을 대상으로 하는 한 TV광고에 대해 광고에 대한 연령별 이해도를 측정한 결과가 다음 표와 같다.

이해도 \ 나이	5~7	8~10	11~12
이해함	55%	40%	15%
이해 못함	45%	60%	85%

위의 TV광고를 여섯 살과 아홉 살인 두 어린이에게 보여준 후 각 어린이가 이 광고를 이해하는지 여부를 조사한 실험결과에 대한 다음 확률을 구하여라.

(a) 6살 어린이가 이 광고를 이해할 확률은 얼마인가?

풀이) 0.55

(b) 두 어린이 모두 광고를 이해할 확률은 얼마인가?

풀이) $0.55 \times 0.4 = 0.22$

(c) 적어도 한 어린이가 이 광고를 이해할 확률은 얼마인가?

풀이) $1 - P(\text{모두 이해 못함}) = 1 - 0.45 \times 0.6 = 1 - 0.27 = 0.73$

3장 24. 거짓말 탐지기 조사에서 피조사자가 정직하게 말하는 경우에는 10%의 양성반응이 나타나고, 거짓말을 하는 경우에는 95%의 양성반응이 나타난다. 한 범죄에 두 사람의 혐의자가 있으며, 두 사람 중 한 사람이 범인이나 두 사람 모두 범행을 부인한다고 할 때, 다음 확률을 구하여라.

(a) 두 혐의자 모두 양성반응을 나타낼 확률

풀이) 한 사람은 정직하고 다른 사람은 거짓말을 하는데도 둘 다 양성반응을 나타내므로
확률은 $0.1 \times 0.95 = 0.095$

(b) 범인은 양성반응이 나타나고, 범인이 아닌 사람은 음성반응이 나타낼 확률

풀이) 범인이 거짓말을 하고 양성, 범인이 아닌 사람은 진실을 말하고 음성인 경우이므로
확률은 $0.95 \times 0.9 = 0.855$

(c) 거짓말 탐지기가 전혀 반대로 반응을 나타낼 확률(즉, 범인은 음성, 범인이 아니면 양성)

풀이) 범인이 거짓말을 하고 음성, 범인이 아닌 사람은 진실을 말하고 양성인 경우이므로
확률은 $0.05 \times 0.1 = 0.005$

(d) 두 사람 중 최소한 한 사람이 양성반응을 나타낼 확률

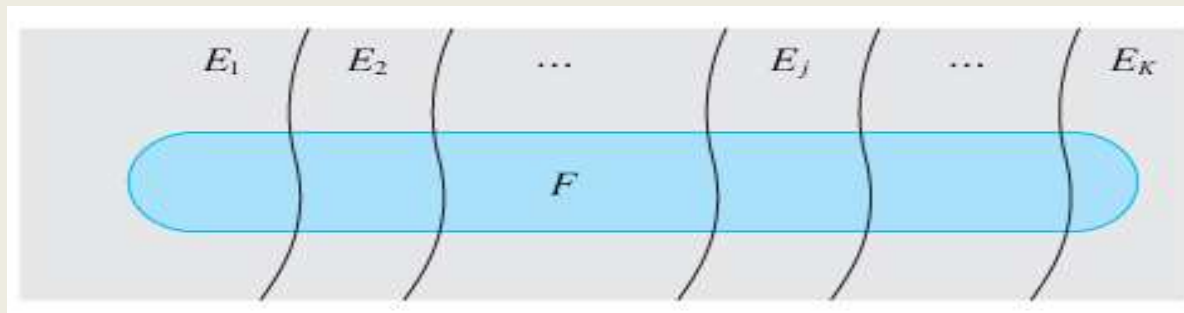
풀이) $1 - P(\text{두 사람 모두 음성}) = 1 - 0.9 \times 0.05 = 1 - 0.045 = 0.955$

베이지 정리 (Bayes' Theorem)

- 표본공간 Ω 가 k 개의 사건 열 E_1, E_2, \dots, E_k 에 의하여 분할(partition)된다고 한다. 다른 사건 F 가 일어났을 때 이 사건이 E_j 에서 일어날 확률은 다음과 같이 계산한다.

$$P(E_j|F) = \frac{P(FE_j)}{P(F)} = \frac{P(FE_j)}{\sum_{i=1}^k P(FE_i)} = \frac{P(F|E_j)P(E_j)}{\sum_{i=1}^k P(F|E_i)P(E_i)}$$

- 여기에서 분할(partition)이란 E_1, E_2, \dots, E_k 가 상호배반이며, $E_1 \cup E_2 \cup \dots \cup E_k = \Omega$ 임을 의미한다.



3장 25. AIDS 항체에 대한 반응검사를 할 경우 AIDS에 걸린 사람들 중에서는 90%가 양성반응을 나타내며, AIDS에 걸리지 않은 사람도 10%의 사람이 양성반응을 나타낸다고 한다. 전 국민 중 1%의 사람들이 AIDS에 감염되었다고 할 때, AIDS 반응 검사에서 양성반응을 나타낸 사람이 실제로 AIDS에 걸렸을 확률은 얼마인가?

풀이) 감염된 사람 집단을 A , 감염되지 않은 사람 집단을 A^C , 양성인 사건의 $+$ 라 할 때,

$$P(+ | A) = 0.9$$

$$P(+ | A^C) = 0.1$$

$$P(A) = 0.01$$

$$P(A^C) = 0.99$$

$$\begin{aligned} P(A | +) &= \frac{P(+, A)}{P(+)} = \frac{P(+, A)}{P(+, A) + P(+, A^C)} \\ &= \frac{P(+ | A)P(A)}{P(+ | A)P(A) + P(+ | A^C)P(A^C)} \\ &= \frac{0.9 \times 0.01}{0.9 \times 0.01 + 0.1 \times 0.99} = \frac{0.009}{0.108} = \frac{1}{12} \end{aligned}$$

끝~~❤❤