

01. 대푯값

1. 대푯값의 뜻

대푯값 : 자료 전체의 특징을 대표적으로 나타내는 값을 **대푯값**이라 한다.

2. 대푯값의 종류

(1) 평균 : 전체 변량의 총합을 변량의 개수로 나눈 값을 **평균**이라고 한다.

(※대푯값으로 가장 많이 쓰이는 것이 평균이다.)

(2) 중앙값 : 각 변량을 크기순으로 나열할 때, **중앙에 오는 값**을 **중앙값**이라 한다.

① 자료의 개수가 홀수이면 중앙에 있는 자료의 값이 중앙값이다.

② 자료의 개수가 짝수이면 중앙에 있는 두 값의 평균이 중앙값이다.

(예) 자료가 2, 3, 6, 8, 9인 경우의 중앙값은 6이다.

자료가 2, 3, 6, 8, 9, 10인 경우의 중앙값은

$$\frac{6+8}{2}=7 \text{ 이다.}$$

(3) 최빈값 : 각 변량 중에서 가장 자주 나타나는 값, 즉 **도수가 가장 큰 값**을 **최빈값**이라 한다.

① 자료의 값 중에서 도수가 가장 큰 값이 한 개 이상 있으면 그 값이 모두 최빈값이다.

② 각 자료의 값의 도수가 모두 같으면 최빈값은 없다.

(예) 자료가 1, 1, 1, 2, 3, 3, 3, 4, 5인 경우에 최빈값은 1과 3이다.

자료가 2, 3, 5, 7인 경우에 최빈값은 없다.

02. 분산과 표준편차

1. 산포도

(1) 산포도 : 대푯값을 중심으로 **자료가 흩어져 있는 정도**를 하나의 수로 나타낸 값을 **산포도**로라 한다.

(2) 산포도가 크면 자료들이 대푯값으로부터 멀리 흩어져 있고, 산포도가 작으면 자료들이 대푯값 주위에 밀집되어 있다.

2. 분산과 표준편차

(1) 편차 : 어떤 자료의 **변량에서 평균을 뺀 값**을 **편차**라고 한다.

$$(\text{편차}) = (\text{변량}) - (\text{평균})$$

① 편차의 총합은 항상 0이다.

② 평균보다 큰 변량의 편차는 양수이고, 평균보다 작은 변량의 편차는 음수이다.

(2) 분산 : 각 편차의 제곱의 합을 전체 변량의 개수로 나눈 값, 즉 **편차의 제곱의 평균**을 **분산**이라 한다.

$$(\text{분산}) = \frac{\{(\text{편차})^2 \text{의 총합}\}}{\{(\text{변량}) \text{의 개수}\}}$$

분산이 클수록 그 자료의 분포 상태는 평균을 중심으로 흩어져 있는 정도가 더 심하다고 할 수 있다. 반대로 분산이 작을 수록 자료가 평균 가까이에 밀집되어 있으므로 자료의 분포 상태가 고르다고 할 수 있다.

(3) 표준편차 : **분산의 음이 아닌 제곱근**을 **표준편차**라 한다.

$$(\text{표준편차}) = \sqrt{(\text{분산})}$$

3. 도수분포표에서의 평균과 분산, 표준편차

다음 도수분포표와 같이 계급값 $x_1, x_2, x_3, \dots, x_n$ 의 도수가 각각 $f_1, f_2, f_3, \dots, f_n$ 일 때,

계급값	도수
x_1	f_1
x_2	f_2
x_3	f_3
$\vdots$	$\vdots$
x_n	f_n
합계	N

$$\begin{aligned} (1) (\text{평균}) &= \frac{\{(\text{계급값}) \times (\text{도수}) \text{의 총합}\}}{\{(\text{도수}) \text{의 총합}\}} \\ &= \frac{x_1 f_1 + x_2 f_2 + x_3 f_3 + \dots + x_n f_n}{N} \end{aligned}$$

$$\begin{aligned} (2) (\text{분산}) &= \frac{\{(\text{편차})^2 \times (\text{도수}) \text{의 총합}\}}{\{(\text{도수}) \text{의 총합}\}} \\ &= \frac{(x_1 - m)^2 f_1 + (x_2 - m)^2 f_2 + \dots + (x_n - m)^2 f_n}{N} \\ &\quad (\text{단, } m \text{은 평균}) \end{aligned}$$

$$(3) (\text{표준편차}) = \sqrt{(\text{분산})}$$

1. 다음 변량들의 중앙값과 최빈값, 평균을 각각 구하여라.

- (1) 5, 3, 6, 4, 1, 3
- (2) 22, 16, 18, 20, 22, 25
- (3) 14, 17, 16, 20, 17, 18, 17, 16
- (4) 3, 4, 6, 7, 4, 7, 6
- (5) 11, 5, 3, 7, 1, 9

2. 다음 변량들의 평균이 []안의 수와 같을 때, x 의 값을 구하여라.

- (1) 4, 5, 6, x [6]
- (2) 7, 3, x , 4, 5 [5]
- (3) 86, 90, 92, x , 97 [90]
- (4) 16, 22, 19, 25, x [20]

3. 세 변량 x, y, z 의 평균이 4일 때, 다음 변량들의 평균을 구하여라.

- (1) 2, $x, y, z, 6$
- (2) $x, y, z, 8, 4, 12$
- (3) $3x+1, 3y+1, 3z+1$

4. 다음 각 변량들의 평균, 분산, 표준편차를 각각 구하여라.

- (1) 14, 20, 16, 18, 17
- (2) 47, 48, 50, 51, 54
- (3) 1, 2, 3, 4, 5, 6, 7
- (4) 25, 44, 38, 35, 52, 40, 53

[정답]

1. (1) 중앙값 : 3.5, 최빈값 : 3, 평균 : $\frac{22}{6}$
- (2) 중앙값 : 21, 최빈값 : 22, 평균 : 20.5
- (3) 중앙값 : 17, 최빈값 : 17, 평균 : $\frac{135}{8}$
- (4) 중앙값 : 6, 최빈값 : 4, 6, 7, 평균 : $\frac{37}{7}$
- (5) 중앙값 : 6, 최빈값 : 없다, 평균 : 6
2. (1) 9, (2) 6, (3) 85, (4) 18
3. (1) 4, (2) 6, (3) 13
4. (1) 평균 : 17, 분산 : 4, 표준편차 : 2
- (2) 평균 : 50, 분산 : 6, 표준편차 : $\sqrt{6}$
- (3) 평균 : 4, 분산 : 4, 표준편차 : 2
- (4) 평균 : 41, 분산 : $\frac{576}{7}$, 표준편차 : $\frac{24\sqrt{7}}{7}$

5. 다음 주어진 자료들의 평균이 []안의 수와 같을 때, 이 자료들의 분산을 구하여라.

- (1) 4, x , 8, 2 [5]
 (2) 10, 3, x , 6, 9 [7]
 (3) 7, 16, 14, 10, x [13]

6. 다음 주어진 표에서 a 의 값과 평균, 분산을 각각 구하여라.

(1)

계급값	1	3	5	합계
도수	3	4	a	10

(2)

계급값	6	7	8	9	10	합계
도수	2	1	3	3	a	10

(3)

계급값	4	6	8	10	12	합계
도수	a	4	2	1	1	10

(4)

계급값	1	3	5	7	9	합계
도수	6	12	a	8	8	40

7. 다음 도수분포표에서 표준편차를 구하여라.

(1)

계급	도수
60이상~70미만	1
70~80	4
80~90	3
90~100	2
합계	10

(2)

계급	도수
0이상~10미만	2
10~20	9
20~30	6
30~40	3
합계	20

(3)

계급	도수
0이상~30미만	1
30~60	7
60~90	5
90~120	5
120~150	2
합계	20

[정답]

5. (1) 5, (2) 6, (3) 16
 6. (1) a 값: 3, 평균: 3, 분산: 2.4
 (2) a 값: 1, 평균: 8, 분산: 1.6
 (3) a 값: 2, 평균: 7, 분산: 5.8
 (4) a 값: 6, 평균: 5, 분산: 7.6
 7. (1) $2\sqrt{21}$ (2) $5\sqrt{3}$ (3) $6\sqrt{30}$