

# 선수 지식 - 통계

## 분산과 표준편차

분산과 표준편차 | 딥러닝의 기초가 되는 확률 개념 알아보기

강사 나동빈

# 선수 지식 - 통계

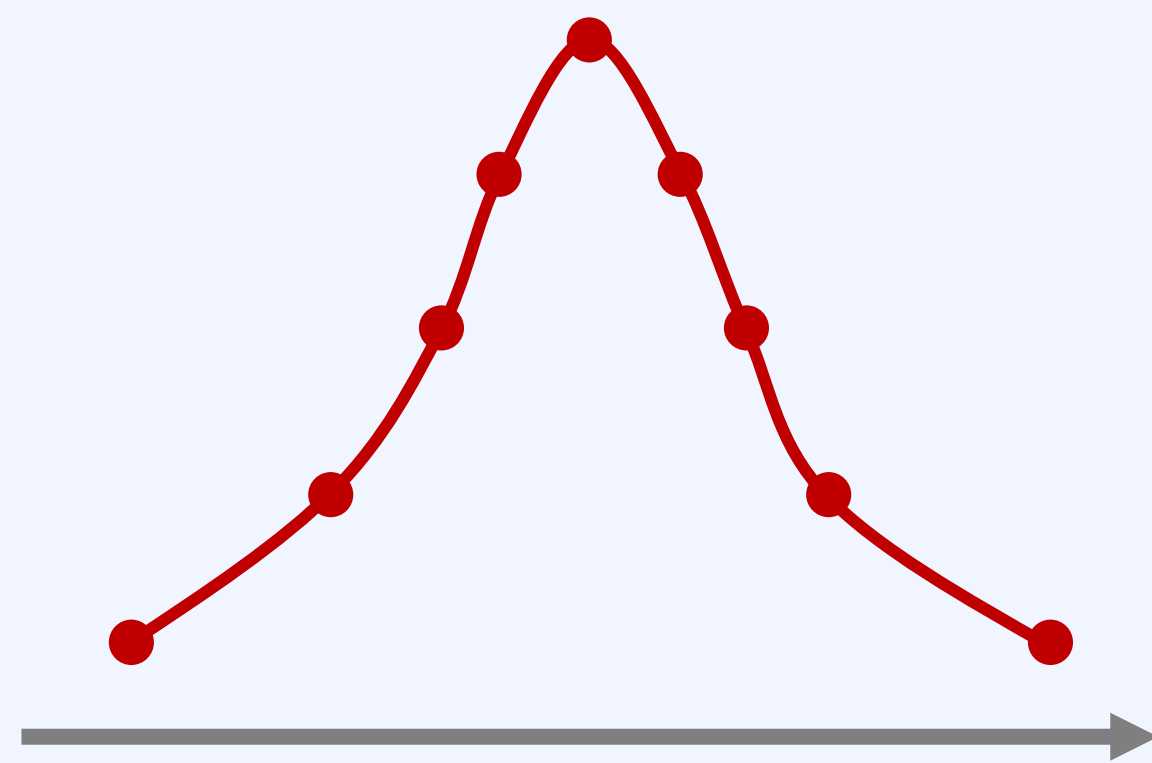
분산과 표준편차

## 분산(Variance)

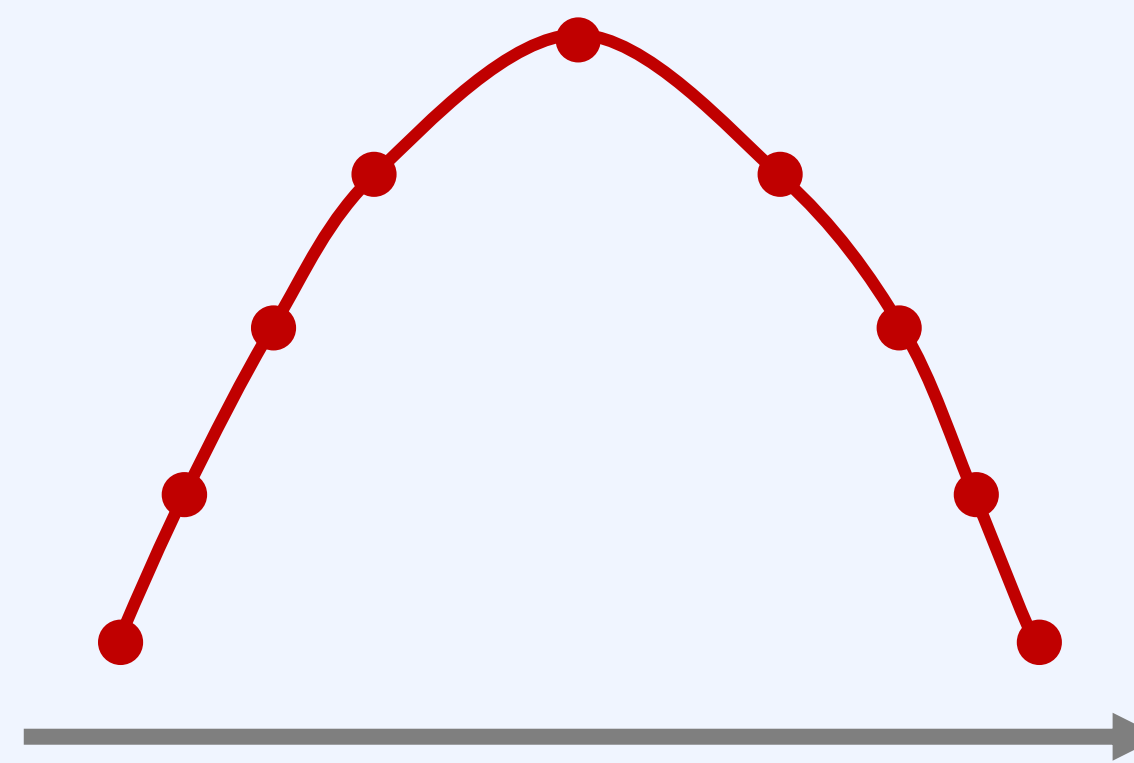
- 평균과 관측치에 대하여 편차(difference) 제공의 평균 값을 의미한다.
- $N$ 개의 데이터의 평균(mean) 값  $\mu$ 가 주어졌을 때, 분산은 다음의 공식으로 계산할 수 있다.
- **분산(variance) 공식:**  $\frac{1}{N} \sum_{i=1}^N |x_i - \mu|^2$

## 분산(Variance)을 사용하는 이유

- 편차는 평균과의 차이이므로, 편차를 모두 더하면 단순히 0이 되어버린다.
- 그래서 편차를 단순히 더하지 않고, 제곱해서 더한다.
- 분산이 작을 때: 각 데이터가 평균에 가까울수록 (평균에 근접할수록)
- 분산이 클 때: 각 데이터가 평균에서 멀수록 (퍼져있을 수록)



작은 분산



큰 분산

## 표준편차(Standard Deviation)

- 표준편차(standard deviation)는 분산(variance)의 양의 제곱근이다.
- 분산은 편차의 제곱을 사용하기 때문에 값이 커지는 경향이 있다.  
→ 제곱하여 커졌던 값을 다시 원래의 값의 크기로 “표준화”할 필요가 있다.
- 편차가 3점인 경우 제곱하면 9점이 되는데, 다시 제곱근을 취하면 3점 차이라는 것을 알 수 있다.

- **표준 편차 공식:**  $\sqrt{\frac{1}{N} \sum_{i=1}^N |x_i - \mu|^2}$

## 분산 및 표준편차 계산 예시

- 계산 예시 1)

학생 번호	1번	2번	3번	4번	5번	평균	분산	표준 편차
성적	56	93	88	72	65	74.80	192.56	13.88

- 계산 예시 2)

학생 번호	1번	2번	3번	4번	5번	평균	분산	표준 편차
성적	41	100	90	63	80	74.80	434.96	20.86

## 파이썬(Python)을 활용한 분산과 표준편차 계산

- 파이썬을 이용해 분산 및 표준편차를 계산할 수 있다.

```
import math

arr = [56, 93, 88, 72, 65]

mean = 0
for x in arr:
    mean += x / len(arr)

variance = 0
for x in arr:
    variance += ((x - mean) ** 2) / len(arr)

std = math.sqrt(variance)

print(f"평균: {mean:.2f}")
print(f"분산: {variance:.2f}")
print(f"표준 편차: {std:.2f}")
```

## [실행 결과]

평균: 74.80  
분산: 192.56  
표준 편차: 13.88