■ 모분산(population variance)

- 표본분산
  - 표본크기: n
  - 표본이 가질 수 있는 값 $\{x_1, x_2, \ldots, x_k\}$
  - n_i: 표본 중 x_i 값을 가지는 표본의 수

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{k} n_i (x_i - \overline{x})^2 = \frac{n}{n-1}\sum_{i=1}^{k}(x_i - \overline{x})^2 p_i$$

  - $p_i = n_i / n$

- n을 계속 크게 하면
  - $p_i \Rightarrow f(x_i)$
  - $\overline{x} \Rightarrow \mu$ ⇨ 표본분산 ⇒ 모분산
  - $n/(n-1) \Rightarrow 1$

- 모분산을 σ^2로 표시

$$s^2 = \frac{n}{n-1}\sum_{i=1}^{k}(x_i - \overline{x})^2 p_i \quad \rightarrow \quad \sigma^2 = \sum_{i=1}^{k}(x_i - \mu)^2 f(x_i)$$

- 확률변수 X의 분산을 $Var(X)$로 표시

$$Var(X) = \sum_x (x-\mu)^2 f(x) = E((X-\mu)^2)$$

  - $g(X) = (X-\mu)^2$ 의 기댓값
  - $Var(X) = \sum_x (x-\mu)^2 f(x)$

$$= E(X^2) - \mu^2 = E(X^2) - E(X)^2$$

- 표준편차: $\sigma = \sqrt{\sigma^2} = SD(X)$

SOOKMYUNG WOMEN'S UNIVERSITY K-MOOC

⦿ 동전 세 개를 던지기: 앞면의 수 X

- 평균: $\mu = 1.5$
- 분산: $Var(X) = E(X^2) - E(X)^2$
  - $E(X^2) = 3$
  - $Var(X) = 3 - 1.5^2 = 0.75$
- 표준편차: $\sigma = \sqrt{0.75} = 0.866$

⊙ 이산균일분포

x	1	2	3	4
$f(x)$	$1/4$	$1/4$	$1/4$	$1/4$

- 2.5를 중심으로 대칭이므로 $E(X)=5/2=2.5$
- $E(X^2)=30/4$
- $\sigma^2 = E(X^2)-E(X)^2 = \frac{30}{4}-\left(\frac{5}{2}\right)^2 = \frac{5}{4}=1.25$
- $\sigma = \sqrt{\sigma^2} = \sqrt{1.25}=1.118$

● 연속확률변수 X의 분산

$$Var(X) = \int (x-\mu)^2 f(x)dx = \int x^2 f(x)dx - \left(\int x f(x)dx\right)^2$$

⦿ 돌림판

$$f(x) = 1/12, \quad 0 < x \le 12$$

○ $E(X) = 6, \quad E(X^2) = 48$

○ $Var(X) = 12 \Rightarrow SD(X) = \sqrt{12} = 3.464$

- $Var(aX+b) = a^2 Var(X)$

  - $E(aX+b) = aE(X) + b = a\mu + b$

    $\Rightarrow \quad \sum (ax+b-a\mu-b)^2 f(x) = a^2 \sum (x-\mu)^2 f(x)$

  - 위치의 변화를 주는 상수 b는 분산에 영향을 주지 않음
  - 분산은 측정단위의 제곱이기 때문에 a의 제곱을 곱함
  - $SD(aX+b) = \sqrt{Var(aX+b)} = \sqrt{a^2 Var(X)} = |a| SD(X)$

◉ $X \sim U(0,1)$: Uniform distribution

- 구간 (0,1)에서 균등하게 분포 ⇒ 균일(균등)분포
- $E(X) = 1/2$
- $E(X^2) = \int_0^1 x^2 dx = \frac{1}{3} \quad \Rightarrow \quad Var(X) = \frac{1}{12}$
- $Y = 12X$ ⇨ 돌림판
  - $Var(12X) = 12^2 Var(X) = 12$
- $W \sim U(-1,1)$
  - $W = 2X - 1 \Rightarrow Var(W) = 4Var(X) = \frac{1}{3}$

■ 요약

- 분산(σ^2): $Var(X) = E((X-\mu)^2) = \begin{cases} \sum (x-\mu)^2 f(x) \\ \int (x-\mu)^2 f(x)\,dx \end{cases}$
- $Var(X) = E(X^2) - E(X)^2$
- 표준편차(σ): $SD(X) = \sqrt{Var(X)}$
- 상수 a, b에 대해, $Var(aX+b) = a^2 Var(X)$
- $SD(aX+b) = |a| SD(X)$