6-1 확률분포와 기대값

제주권역 대학 e-러닝 지원센터

기대값 (expected value : 평균)

기대값은 확률분포에서 **분포의 무게중심**을 말하며, 확률값을 가중치로 하는 확률변수의 가능한 값에 대한 가중평균이라고 할 수 있다

확률변수 ***X*** 의 기대값은 ***E***(***X***)로 표현하며 각각 다음과 같이 구한다

1. 이산형 확률변수

이산형 확률변수 ***X*** 의 가능한 값이 ($x_1, x_2, \ldots, x_n$) 이며,

$$P(X = x_i) = P_i,\ i=1,2,\ldots,n$$ 일 때

X 의 기대값 ***E***(***X***) 는 다음과 같다

$$E(X) = \sum_{i=1}^{n} x_i P_i$$

2. 연속형 확률변수

연속형 확률변수 X의 확률밀도함수를 $f(x)$라 하면 X의 기대값 $E(X)$는 다음과 같다

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)dx$$

예1 이산형 확률변수의 기대값 계산

두 개의 주사위를 던지는 실험에서 확률변수 X를 '두 주사위 눈금의 합'이라고 정의하면 표본공간은 다음과 같다

첫번째 / 두번째	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

X의 가능한 값은 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12이며, 위 표본공간에서 X의 확률분포는 다음과 같다.

X	2	3	4	5	6	7	8	9	10	11	12
확률	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

위 확률분포표를 이용하여 확률변수 X의 기대값을 계산해 보면 다음과 같다

$$E(X) = 2\times\frac{1}{36} + 3\times\frac{2}{36} + 4\times\frac{3}{36} + 5\times\frac{4}{36} + 6\times\frac{5}{36} + 7\times\frac{6}{36}$$
$$+8\times\frac{5}{36} + 9\times\frac{4}{36} + 10\times\frac{3}{36} + 11\times\frac{2}{36} + 12\times\frac{1}{36}$$
$$= 7$$

기대값의 특성

X, Y 를 확률변수, ***a, b*** 를 상수라고 할 때, 기대값은 항상 다음 조건을 만족한다.

1. $E(a) = a$

   즉, 상수의 기대값은 상수 그 자체이다.

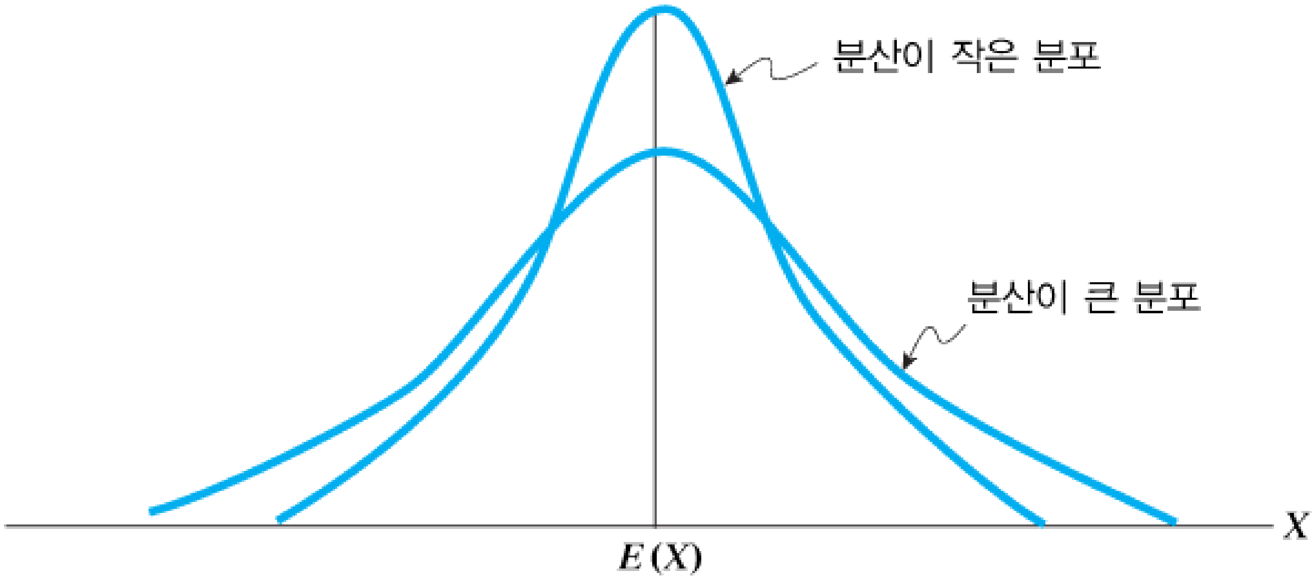
2. $E(aX + b) = aE(X) + b$

3. $E(aX + bY) = aE(X) + bE(Y)$

상수 a와 b를 제외할 때 E(X+Y)=E(X)+E(Y)가 되는데, 이는 **두 확률변수의 합의 기대값은 각 확률변수의 기대값의 합과 같음을 의미**한다.

분산(variance)

평균이 확률분포의 무게중심인데 비하여 분산은 **확률분포의 흩어진 정도를 측정**하는 것으로 평균이 같은 경우에도 분산의 크기에 따라서 분포의 모양이 달라진다.



확률변수 X의 기대값 $E(X) = \mu$ 라고 할 때, 분산(variance)은 X와 μ의 편차의 제곱, 즉 $(X - \mu)^2$ 의 기대값으로 $Var(X)$ 또는 σ_X^2으로 표현되며,

$$\sigma_X^2 = \mathrm{E}(X - \mu)^2 = E(X^2) - \mu^2$$

로 구한다. 단, $E(X^2) = \sum_{i=1}^{n} x_i^2 P_i$

여기에서 확률변수 X를 나타낼 필요가 없는 경우에는 σ_X^2에서 x 를 없애고 σ^2으로 표현한다. 분산의 양의 제곱근을 표준편차(standard deviation)라고 하며 σ 로 표현한다. 즉,

$$\sigma = \sqrt{\sigma^2}$$

이다.

예2 이산형 확률변수의 분산 계산

이산형 확률변수 X의 확률분포가 다음과 같을 때 X의 평균과 분산을 구해보자.

X	0	1	2	3
확률	1/8	3/8	3/8	1/8

$$\mu = \mathrm{E(X)} = 0\times\frac{1}{8} + 1\times\frac{3}{8} + 2\times\frac{3}{8} + 3\times\frac{1}{8} = 1.5$$

$$\mathrm{E}(X^2) = 0\times\frac{1}{8} + 1\times\frac{3}{8} + 4\times\frac{3}{8} + 9\times\frac{1}{8} = 3$$

$$\sigma^2 = \mathrm{E}(X^2) - \mu^2 = 3 - 1.5^2 = 0.75$$

확률변수의 X의 분산 σ_x^2 는 다음과 같이 정의된다.

$$\sigma_X^2 = \mathrm{E}[(X - \mu_X)^2] = \sum_x (X - \mu_X)^2 P_X(x) = E(X^2) - \mu_X^2$$

표준편차 σ_X 는 분산의 제곱근이다.

분산의 특성

1. $Var(a) = 0$
2. $Var(aX) = a^2 Var(X)$
3. $Var(X + a) = Var(X)$
4. 만약 X와 Y가 독립이면 $Var(X \pm Y) = Var(X) + Var(Y)$
5. 만약 X와 Y가 독립이 아니면 $Var(aX \pm bY) = a^2 Var(X) + b^2 Var(Y)$ $\pm$2abCov(X,Y)

참고 기대 및 분산연산자를 이용한 자료의 선형변환

(1) 상수(c)를 더할 때

$$W = X + c$$

평균	$E(W) = E(X+c) = E(X) + c$
분산	$Var(W) = Var(X+c) = Var(X)$

(2) 상수 곱했을때

$$W = cX$$

평균	$E(W) = E(cX) = cE(X)$
분산	$Var(W) = Var(cX) = c^2 Var(X)$

(3) 상수 더하고 곱했을때

$$W = cX + d$$

평균	$E(W) = E(cX+d) = cE(X) + d$
분산	$Var(W) = Var(cX+d) = c^2 Var(X)$