3.8 기대값, 분산, 공분산(Expectation, Variance and Covariance)

x~P(x)에 대한 어떤 함수 f(x)의 기대값은 아래와 같이 정의함.

- 이산(discrete) 변수인 경우

$$\mathbb{E}_{\mathrm{x}\sim P}[f(x)] = \sum_x P(x)f(x), \tag{3.9}$$

- 연속(continuous) 변수인 경우

$$\mathbb{E}_{\mathrm{x}\sim p}[f(x)] = \int p(x)f(x)dx. \tag{3.10}$$

- 기대값은 선형적(linear):

$$\mathbb{E}_{\mathrm{x}}[\alpha f(x) + \beta g(x)] = \alpha \mathbb{E}_{\mathrm{x}}[f(x)] + \beta \mathbb{E}_{\mathrm{x}}[g(x)], \tag{3.11}$$

분산(Variance):

$$\mathrm{Var}(f(x)) = \mathbb{E}\left[(f(x) - \mathbb{E}[f(x)])^2\right]. \tag{3.12}$$

- 표준 편차(standard deviation)= 분산의 (+) 제곱근.

공분산(Covariance):

- 두 개의 확률 변수가 (선형적으로) 얼마나 연관이 있나를 보고자 할 때 사용

$$\mathrm{Cov}(f(x), g(y)) = \mathbb{E}\left[(f(x) - \mathbb{E}\left[f(x)\right])(g(y) - \mathbb{E}\left[g(y)\right])\right]. \tag{3.13}$$

  - 공분산의 절대값이 크다는 것은 두 개의 변수가 서로 각각의 평균값으로부터 멀어지는 정도가 동일하게 큼을 의미
  - 한쪽 변수가 상승하는 값일 때 다른 쪽도 상승하는 경향을 보이면 공분산은 양(+)의 값이 되고, 한쪽 변수는 상승하는데 다른 쪽은 하강하는 경향을 보이면 공분산은 음(-)의 값을 갖는다.
  - 공분산의 값 자체는 각 변수의 측정 단위에 따라 달라지므로 상관관계만을 따질 때는 각 변수의 영향도를 정규화시켜서 따져야 함(correlation).
  - 두 변수가 서로 독립이면 공분산 값은 0이 됨. (역은 성립하지 않음).
- 공분산 행렬(covariance matrix):

The *covariance matrix* of a random vector $\boldsymbol{x} \in \mathbb{R}^n$ is an $n \times n$ matrix, such that

$$\mathrm{Cov}(\mathbf{x})_{i,j} = \mathrm{Cov}(\mathrm{x}_i, \mathrm{x}_j). \tag{3.14}$$

The diagonal elements of the covariance give the variance:

$$\mathrm{Cov}(\mathrm{x}_i, \mathrm{x}_i) = \mathrm{Var}(\mathrm{x}_i). \tag{3.15}$$