2024/03/07, 2024 확률 기초 세미나

# 확률 및 통계학

## - 4장 수학적 기대값 -


손 우 영(wooyoung@pel.sejong.ac.kr)

세종대학교 프로토콜공학연구실

# 목 차

- 보충
- 확률변수의 평균
- 분산과 공분산
- 선형결합된 확률변수의 평균과 분산
- 체비세프 정리

# 목 차

- 보충
- 확률변수의 평균
- 분산과 공분산
- 선형결합된 확률변수의 평균과 분산
- 체비세프 정리

# 보충 (1/10)

- Analytical Method vs. Numerical Method (1/3)
  - Analytical Method
    - 정의
      - 수학적 원리와 이론을 활용하여 문제의 해를 명확한 형태로 도출하는 접근법
    - 장점
      - 명확한 수학적 공식이나 정리를 사용함에 따라 정확한 답을 도출할 수 있음
    - 단점
      - 적합한 수학적 원리와 이론이 존재하는 문제에만 적용할 수 있음
      - 비선형성이나 고차원과 같은 복잡한 특성을 갖는 시스템에 적용하는 데에 어려움이 있으며 많은 시간이 소요됨
      - 고급 수학적 기술과 전문 지식이 필요함

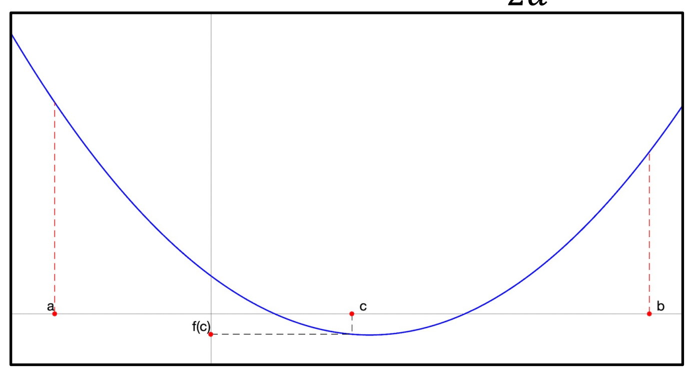
# 보충 (2/10)

- Analytical Method vs. Numerical Method (2/3)
  - Numerical Method
    - 정의
      - Analytical Method로 풀기 어려운 문제를 해결하기 위해 반복적인 계산, 이산화, 알고리즘적 접근을 사용함으로써 수치적 근사값을 계산하는 접근법
    - 장점
      - 복잡하거나 해석적으로 해결하기 어려운 문제에 대해 근사해를 제공하고, 다양한 유형의 문제에 유연하게 적용할 수 있음
    - 단점
      - 근사값을 제공함에 따라 오차를 포함할 수 있고, 이는 최종적으로 제공하는 근사해의 정확성에 영향을 미칠 수 있음

# 보충 (3/10)

- Analytical Method vs. Numerical Method (3/3)
  - 활용 예
    - e.g., 2차 방정식 $2x^2 - 5x + 2 = 0$의 해결
      - Analytical Method
        - 2차 방정식을 풀기 위한 근의 공식 ($x = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$) 사용

1. $x = \frac{-(-5) \pm \sqrt{(-5)^2 - 4(2)(2)}}{2(2)}$
2. $x_1 = 2, \quad x_2 = \frac{1}{2}$




<그림 4.1 이등분법을 활용한 방정식($2x^2 - 5x + 2 = 0$) 해 도출 그래프>

      - Numerical Method
        - 해를 근사화하는 방법 중 하나인 '이등분법' 사용

1. 근이 둘 사이에 놓이도록 초기 구간 $[a, b]$을 설정한 후, 중간점 $c = (a + b) / 2$를 계산함
2. $c$에서 $f(c)$의 부호를 기반으로 함수 $f(c)$를 평가함
3. 근이 어느 구간 ($[a, c]$또는 $[c, b]$)에 있는지 판단함
4. 구간 $[a, b]$를 근을 포함하는 하위 구간으로 업데이트함
5. 원하는 정확도 수준에 도달할 때까지 $1 \sim 4$단계를 반복함
6. 최종 결과는 $Analytical\ Method$와 유사한 근사치($x \fallingdotseq 2$ 및 $x \fallingdotseq \frac{1}{2}$)로 도출됨

# 보충 (4/10)

- 확률변수 $g(X)$의 기대값
  - 정의

    ❖ $g(X)$: 확률변수 $X$에 종속되는 확률변수

    - 확률변수 $X$가 확률분포 $f(x)$를 가질 때 확률변수 $g(X)$의 평균값
  - 공식
    - $X$가 이산형인 경우: $\mu_{g(X)} = E[g(X)] = \sum_x g(x)f(x)$
    - $X$가 연속형인 경우: $\mu_{g(X)} = E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx$
  - 특징
    - 확률변수 $X$를 통해 새로운 확률적 현상이나 특징을 파악하고자 할 때 활용됨

# 보충 (5/10)

- 확률변수 $g(X)$의 분산
  - 정의

❖ $g(X)$: 확률변수 $X$에 종속되는 확률변수

    - 확률변수 $X$가 확률분포 $f(x)$를 가질 때 확률변수 $g(X)$의 값이 평균($\mu_{g(X)}$)으로부터 벗어난 정도의 평균
  - 공식
    - $X$가 이산형인 경우
      :$\sigma^2{}_{g(X)} = E\left\{\left[g(X) - \mu_{g(X)}\right]^2\right\} = \sum_x \left[g(x) - \mu_{g(X)}\right]^2 f(x)$
    - $X$가 연속형인 경우
      :$\sigma^2{}_{g(X)} = E\left\{\left[g(X) - \mu_{g(X)}\right]^2\right\} = \int_{-\infty}^{\infty} \left[g(x) - \mu_{g(X)}\right]^2 f(x)\, dx$

# 보충 (6/10)

- 예제 4.11

**확률변수 $X$의 확률분포**가 다음과 같이 주어졌을 때 $g(X) = 2X + 3$의 분산을 계산하라

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $f(x)$ | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{2}$ | $\frac{1}{8}$ |

- 평균 계산
  - $\mu_{2X+3} = E(2X+3) = \sum_{x=0}^{3}(2x+3)f(x) = 6$

- $\sigma^2$ 계산
  - $\sigma^2{}_{2X+3} = E\{[(2X+3) - \mu_{2X+3}]^2\} = E[(2X+3-6)^2]$

$$= E(4X^2 - 12X + 9) = \sum_{x=0}^{3}(4x^2 - 12x + 9)f(x) = 4$$

# 보충 (7/10)

- 예제 4.11 – 실생활 적용 예시

한 시간 동안 판매되는 음료의 총 개수를 확률변수 $X$라고 할 때, **확률변수 $X$의 확률분포**는 다음과 같다.

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $f(x)$ | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{2}$ | $\frac{1}{8}$ |

판매된 음료 개수에 따라 카페가 얻는 총 수익 (단위: 백 원)은 $g(X) = 2X + 3$라고 할 때, **한 시간 동안 카페 수익의 평균과 분산을 구하라**

- 평균 계산
  - $\mu_{2X+3} = E(2X+3) = \sum_{x=0}^{3}(2x+3)f(x) = 6$

- $\sigma^2$ 계산
  - $\sigma^2{}_{2X+3} = E\{[(2X+3) - \mu_{2X+3}]^2\} = E[(2X+3-6)^2]$

$$= E(4X^2 - 12X + 9) = \sum_{x=0}^{3}(4x^2 - 12x + 9)f(x) = 4$$

# 보충 (8/10)

- 예제 4.5, 4.12

$X$를 **확률밀도함수** $f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2 \\ 0, & \text{otherwise} \end{cases}$ 를 가지는 확률변수라고 할 때, **확률변수 $g(X) = 4X + 3$의 기대값과 분산**을 구하라

- $E[g(X)] = E(4X+3) = \int_{-1}^{2} \frac{(4x+3)x^2}{3} dx = \frac{1}{3}\int_{-1}^{2}(4x^3+3x^2)dx = 8$
- $\sigma^2{}_{4X+3} = E\{[(4X+3) - \mu_{4X+3}]^2\} = E\{[(4X+3)-8]^2\} = E[(4X-5)^2]$

$$= \int_{-1}^{2}(4x-5)^2\frac{x^2}{3}dx = \frac{1}{3}\int_{-1}^{2}(16x^4 - 40x^3 + 25x^2)dx = \frac{51}{5}$$

# 보충 (9/10)

- 예제 4.5, 4.12 – 실생활 적용 예시

확률변수 $X$를 한 달간의 비만율 변화량이라고 할 때, **<u>확률변수 $X$의 밀도함수</u>**가

$$f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2 \\ 0, & \text{otherwise} \end{cases}$$

이다.

**<u>$g(X) = 4X + 3$</u>**이 건강 보험료 변화량(단위: 만원)을 나타낸다고 할 때, 건강 보험료 변화량의 **<u>기대값과 분산</u>**을 구하라

- $E[g(X)] = E(4X + 3) = \int_{-1}^{2} \frac{(4x+3)x^2}{3} dx = \frac{1}{3}\int_{-1}^{2}(4x^3 + 3x^2)dx = 8$
- $\sigma^2{}_{4X+3} = E\{[(4X + 3) - \mu_{4X+3}]^2\} = E\{[(4X + 3) - 8]^2\} = E[(4X - 5)^2]$

$$= \int_{-1}^{2}(4x - 5)^2 \frac{x^2}{3} dx = \frac{1}{3}\int_{-1}^{2}(16x^4 - 40x^3 + 25x^2)dx = \frac{51}{5}$$

# 보충 (10/10)

- 선형 결합
  - 정의
    - 여러 변수들이 각각의 스칼라 계수와 곱해진 후, 그 결과들이 합쳐진 형태
  - 특징
    - 변수 $X_1, X_2, \cdots, X_n$과 각 변수에 대응하는 계수 $a_1, a_2, \cdots, a_n$에 대해 $a_1X_1 + a_2X_2 + \cdots + a_nX_n$으로 표현됨
    - 선형결합된 확률변수의 경우, 두 개 이상의 확률변수가 각각의 스칼라 계수와 곱해진 후 합해져서 새로운 확률변수를 형성하는 것을 의미함
      - e.g., $Z = aX + bY$

# 목 차

- 보충
- 확률변수의 평균
- 분산과 공분산
- 선형결합된 확률변수의 평균과 분산
- 체비셰프 정리

# 확률변수의 평균 (1/16)

- 수학적 기대값(Mathematical Expectation) (1/4)
  - 정의
    - 확률변수가 취할 수 있는 각각의 값에 그 값이 발생할 확률을 곱한 후 모두 더하여 계산된 확률변수의 평균값
  - 표현
    - 확률변수 $X$의 평균(Mean of the Random Variable $X$)
    - $X$의 확률분포의 평균(Mean of the Probability Distribution of $X$)
    - $\mu_X$, $\mu$
  - 공식
    - $X$가 이산형인 경우: $\mu = E(X) = \sum_x xf(x)$
    - $X$가 연속형인 경우: $\mu = E(X) = \int_{-\infty}^{\infty} xf(x)\,dx$

# 확률변수의 평균 (2/16)

- 수학적 기대값(Mathematical Expectation) (2/4)
  - 필요성 (1/3)
    - 산술평균과의 비교 (1/3)
      - 산술평균은 모든 값이 동일한 확률을 가진다고 가정하는 반면, 수학적 기대값은 각 결과에 해당하는 확률을 고려함
      - 수학적 기대값의 경우, 결과가 발생할 가능성에 따라 가중치를 부여함으로써 현실 세계의 불확실성을 더 잘 반영함
      - 산술평균은 실제 관측된 데이터의 평균을 나타내는 반면, 수학적 기대값은 이론적인 확률분포에 따른 값들의 평균적 기대치를 나타냄
        - e.g., 실험에서 얻은 데이터의 산술평균값을 계산하여 수학적 기대값과 비교함으로써, 실험 결과의 타당성을 평가할 수 있음

**산술평균**

$$\frac{x_1 + x_2 + x_3 + \cdots + x_n}{\mathrm{n}}$$
$$= (x_1)\left(\frac{1}{n}\right) + (x_2)\left(\frac{1}{n}\right) + \cdots + (x_n)\left(\frac{1}{n}\right)$$

**$X$가 이산형일 때의 수학적 기대값**

$$\mu = E(X) = \sum_x x f(x)$$
$$= x_1 f(x_1) + x_2 f(x_2) + \cdots$$

# 확률변수의 평균 (3/16)

- 수학적 기대값(Mathematical Expectation) (3/4)
  - 필요성 (2/3)
    - 산술평균과의 비교 (2/3)
      - 산술평균은 극단적인 결과(e.g., 매우 낮은 확률로 발생하는 사건이지만 그 영향력이 큰 경우)를 반영하지 못하지만, 기대값은 각 결과의 확률을 고려하여 이를 반영함
        - e.g., 대규모 재난 발생에 따른 보험금 지급

보험 회사는 다양한 종류의 보험을 판매하고 있다.
- 소규모 사건(e.g., 차량 사고, 작은 화재 등): 발생확률 1%, 보험금 5,000달러
- 대규모 사건(e.g., 대지진): 발생확률 0.001%, 보험금 1,000,000달러

<수학적 기대값>
- 소규모 사건의 기대값: 5,000달러 × 0.01 = 50달러
- 대규모 사건의 기대값: 1,000,000달러 × 0.00001 = 10달러
- 수학적 기대값: 50달러 + 10달러 = 60달러

# 확률변수의 평균 (4/16)

- 수학적 기대값(Mathematical Expectation) (4/4)
  - 필요성 (3/3)
    - 산술평균과의 비교 (3/3)
      - 산술평균은 극단적인 결과(e.g., 매우 낮은 확률로 발생하는 사건이지만 그 영향력이 큰 경우)를 반영하지 못하지만, 기대값은 각 결과의 확률을 고려하여 이를 반영함
        - e.g., 대규모 재난 발생에 따른 보험금 지급

보험 회사는 다양한 종류의 보험을 판매하고 있다.
- 소규모 사건(e.g., 차량 사고, 작은 화재 등): 발생확률 1%, 보험금 5,000달러
- 대규모 사건(e.g., 대지진): 발생확률 0.001%, 보험금 1,000,000달러

<산술평균>
보험금 지급 사건이 총 100회 발생했으며, 그 중 99회는 소규모 사건이고, 1회는 대규모 사건일 시
- 소규모 사건 총 보험금: 99 × 5,000 달러 = 495,000달러
- 대규모 사건 총 보험금: 1 × 1,000,000 달러 = 1,000,000달러
- 산술평균 보험금: 1,495,000달러 / 100 = 14,950달러

# 확률변수의 평균 (5/16)

- 예제 4.1

품질검사원이 7개의 부품으로 구성되어 있는 로트를 검사하려고 한다. 만일 이 로트에 4개의 양호한 부품과 3개의 결함이 있는 부품이 들어 있다고 하면, 검사원이 3개의 부품을 추출하였을 때 나타나는 양호한 부품의 **평균개수**를 구하라.

- 확률변수 $X$: 추출된 표본에 포함되는 양호한 부품의 수
- $f(x) = \frac{\binom{4}{x}\binom{3}{3-x}}{\binom{7}{3}},\ x = 0, 1, 2, 3$
- $f(0) = \frac{1}{35}, f(1) = \frac{12}{35}, f(2) = \frac{18}{35}, f(3) = \frac{4}{35}$
- $\mu = E(X) = (0)\left(\frac{1}{35}\right) + (1)\left(\frac{12}{35}\right) + (2)\left(\frac{18}{35}\right) + (3)\left(\frac{4}{35}\right) = \frac{12}{7} = 1.7$

# 확률변수의 평균 (6/16)

- 예제 4.2

의료기기 외판원이 어느 날 두 고객을 만나게 되었다. **첫 번째 고객과 거래가 성사될 가능성은 70%**이고 이 경우 $1000을 벌게 되며, **두 번째 고객과는 40%의 거래 성공 가능성**에 성공 시 $1500을 벌게 된다. 각 고객과의 거래 결과는 서로 독립적이라고 할 때, 그가 **기대할 수 있는 성공 보수**는 얼마인가?

- 확률변수 $X$: 의료기기 외판원의 성공 보수

- $f(\$0) = (1 - 0.7)(1 - 0.4) = 0.18$, $f(\$1000) = (0.7)(1 - 0.4) = 0.42$, $f(\$1500) = (1 - 0.7)(0.4) = 0.12$, $f(\$2500) = (0.7)(0.4) = 0.28$

- $E(X) = (\$0)(0.18) + (\$1000)(0.42) + (\$1500)(0.12) + (\$2500)(0.28) = \$1300$

# 확률변수의 평균 (7/16)

- 예제 4.3

어떤 전자장치의 수명(단위: 시간)을 확률변수 $X$라고 하자. **<u>확률밀도함수</u>**가

$$f(x) = \begin{cases} \dfrac{20{,}000}{x^3}, & x > 100 \\ 0, & \text{otherwise} \end{cases}$$

로 주어졌을 때 이 장치의 **<u>기대수명</u>**을 구하라

- $\mu = E(X) = \int_{100}^{\infty} x \frac{20{,}000}{x^3} dx = \int_{100}^{\infty} \frac{20{,}000}{x^2} dx = 200$

- 따라서, 전자장치의 수명은 평균적으로 200시간임

# 확률변수의 평균 (8/16)

- 확률변수 $g(X)$의 기대값 (1/2)
  - 정의
    - 확률변수 $X$가 확률분포 $f(x)$를 가질 때 확률변수 $g(X)$의 평균값

❖ $g(X)$: 확률변수 $X$에 종속되는 확률변수

  - 공식
    - $X$가 이산형인 경우: $\mu_{g(X)} = E[g(X)] = \sum_x g(x)f(x)$
    - $X$가 연속형인 경우: $\mu_{g(X)} = E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx$

# 확률변수의 평균 (9/16)

- 확률변수 $g(X)$의 기대값 (2/2)
  - 특징
    - 확률변수 $X$를 통해 새로운 확률적 현상이나 특징을 파악하고자 할 때 활용됨
      - e.g., 커피숍의 일일 판매량 예측

커피숍을 운영하는 사장은 매일 다양한 수의 고객이 방문하고 각각 다른 양의 커피를 구매한다는 것을 알고 있다. 이 때, 사장은 하루에 판매될 커피의 총량을 예측하고자 한다.

- 하루에 방문하는 고객 수를 확률변수 $X$, 고객이 구매할 커피의 양을 나타내는 함수를 $g(X)$라고 하자. 평균적으로 각 고객이 2잔의 커피를 구매함을 파악했을 때, $g(X) = 2X$이다.
- $\mu_{g(X)} = E[g(X)] = \sum_x g(x)f(x) = \sum_x 2xf(x)$의 식을 통해 커피숍은 일일 판매량을 예측할 수 있음

- 예제 4.4

어느 쾌청한 금요일 오후 4시에서 5시 사이에 세차장에서 서비스를 받는 차의 수를 $X$라고 할 때, **<u>$X$의 확률분포</u>**가 다음과 같다고 하자.

| $x$ | **4** | **5** | **6** | **7** | **8** | **9** |
|---|---|---|---|---|---|---|
| $\mathrm{P}(X=x)$ | $\frac{1}{12}$ | $\frac{1}{12}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |

<u>$\boldsymbol{g(X)=2X-1}$</u>을 종업원이 받는 수당(단위: 달러)이라고 할 때, 이 시간대의 종업원의 **<u>기대수익</u>**을 구하라

- $E[g(X)]=E(2X-1)=\sum_{x=4}^{9}(2x-1)f(x)$

$$= (7)\left(\frac{1}{12}\right)+(9)\left(\frac{1}{12}\right)+(11)\left(\frac{1}{4}\right)+(13)\left(\frac{1}{4}\right)+(15)\left(\frac{1}{16}\right)+(17)\left(\frac{1}{6}\right)$$

$$= \$12.67$$

# 확률변수의 평균 (11/16)

- 예제 4.5

**확률변수 $\boldsymbol{X}$의 밀도함수**가

$$f(x) = \begin{cases} \dfrac{x^2}{3}, & -1 < x < 2 \\ 0, & \text{otherwise} \end{cases}$$

일 때, **$\boldsymbol{g(X) = 4X + 3}$의 기대값**을 구하라

- $E[g(X)] = E(4X + 3)$

$$= \int_{-1}^{2} \frac{(4x+3)x^2}{3} dx$$

$$= \frac{1}{3} \int_{-1}^{2} (4x^3 + 3x^2) dx = 8$$

**실생활 적용 예시**

한 달간의 비만율 변화량을 확률변수 $X$라고 할 때, **확률변수 $\boldsymbol{X}$의 밀도함수**가

$$f(x) = \begin{cases} \dfrac{x^2}{3}, & -1 < x < 2 \\ 0, & \text{otherwise} \end{cases}$$

이다.

**$\boldsymbol{g(X) = 4X + 3}$**가 건강 보험료 변화량(단위: 만원)을 나타낸다고 할 때, 건강 보험료 변화량의 기대값을 구하라

# 확률변수의 평균 (12/16)

- 확률변수 $g(X,Y)$의 기대값 (1/3)
  - 정의
    - 확률변수 $X$와 $Y$가 결합확률분포 $f(x,y)$를 가질 때 확률변수 $g(X,Y)$의 기대값
  - 공식 (1/2)
    - $X$와 $Y$가 이산형인 경우
      : $\mu_{g(X,Y)} = E[g(X,Y)] = \sum_x \sum_y g(x,y)\, f(x,y)$
    - $X$와 $Y$가 연속형인 경우
      : $\mu_{g(X,Y)} = E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f(x,y)\, dx\, dy$

# 확률변수의 평균 (13/16)

- 확률변수 $g(X, Y)$의 기대값 (2/3)
  - 특징

❖$f(x, y)$: 확률변수 $X$에 종속되는 확률변수

    - 두 확률변수 $X, Y$가 결합하여 발생하는 현상을 분석하거나 예측하는 데 사용됨
      - e.g., 가정의 월간 전기요금 예측

가정에서의 다양한 전기 제품을 사용하기 위한 전기의 요금은 사용한 전력량과 전기 단가에 따라 달라진다.

- 하루 전기 사용량을 확률변수 $X$, 전기 단가를 확률변수 $Y$, 하루 전기 사용량과 단가에 따른 월간 전기요금을 계산하는 함수를 $g(X, Y)$라고 하자.
- $f(x, y)$는 하루 전기 사용량과 해당하는 전기 단가의 결합 확률 밀도 함수
- 한 달을 30일로 가정하면, $g(X, Y) = 30XY$이다.
- $\mu_{g(X,\mathrm{Y})} = E[g(X, Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y) f(x, y) dx\, dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 30xyf(x, y) dx\, dy$을 통해 가정의 월간 전기요금을 예측할 수 있음

# 확률변수의 평균 (14/16)

- 예제 4.6

$X$와 $Y$를 다음과 같은 **결합확률분포**를 가지는 확률변수라 할 때, $\boldsymbol{g(X,Y)=XY}$**의 기대값**을 구하라

| $f(x,y)$ | | $x$ | | | 행의 합 |
|---|---|---|---|---|---|
| | | 0 | 1 | 2 | |
| $y$ | 0 | $\frac{3}{28}$ | $\frac{9}{28}$ | $\frac{3}{28}$ | $\frac{15}{28}$ |
| | 1 | $\frac{3}{14}$ | $\frac{3}{14}$ | 0 | $\frac{3}{7}$ |
| | 2 | $\frac{1}{28}$ | 0 | 0 | $\frac{1}{28}$ |
| 열의 합 | | $\frac{5}{14}$ | $\frac{15}{28}$ | $\frac{3}{28}$ | 1 |

- $E[g(X,Y)]=E(XY)$

$$=\sum_{x=0}^{2}\sum_{y=0}^{2}xyf(x,y)=(0)(0)f(0,0)+(0)(1)f(0,1)+(1)(0)f(1,0)$$

$$+(1)(1)f(1,1)+(2)(0)f(2,0)=f(1,1)=\frac{3}{14}$$

# 확률변수의 평균 (15/16)

- 예제 4.7

다음과 같은 **결합밀도함수**에 대하여 $\boldsymbol{E}\left(\frac{\boldsymbol{Y}}{\boldsymbol{X}}\right)$를 구하라

$$f(x,y)=\begin{cases}\dfrac{x(1+3y^2)}{4}, & 0<x<2, 0<y<1\\ 0, & \text{otherwise}\end{cases}$$

- $E\left(\frac{Y}{X}\right)$

$$=\int_0^1\int_0^2\frac{y(1+3y^2)}{4}\,dx\,dy=\int_0^1\frac{y+3y^3}{2}\,dy$$

$$=\frac{5}{8}$$

# 확률변수의 평균 (16/16)

- 확률변수 $g(X, Y)$의 기대값 (3/3)
  - 공식 (2/2)
    - $g(X, Y) = X$이고, $X$의 주변분포가 $g(X)$인 경우
      - $X$와 $Y$가 이산형인 경우: $\mathrm{E(X)} = \sum_x \sum_y xf(x, y) = \sum_x xg(x)$
      - $X$와 $Y$가 연속형인 경우: $\mathrm{E(X)} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf(x, y) dy\, dx = \int_{-\infty}^{\infty} xg(x)\, dx$
    - $g(X, Y) = Y$이고, $Y$의 주변분포가 $h(y)$인 경우
      - $X$와 $Y$가 이산형인 경우: $\mathrm{E(Y)} = \sum_y \sum_x yf(x, y) = \sum_y yh(y)$
      - $X$와 $Y$가 연속형인 경우: $\mathrm{E(Y)} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} yf(x, y) dx\, dy = \int_{-\infty}^{\infty} yh(y)\, dy$

# 목 차

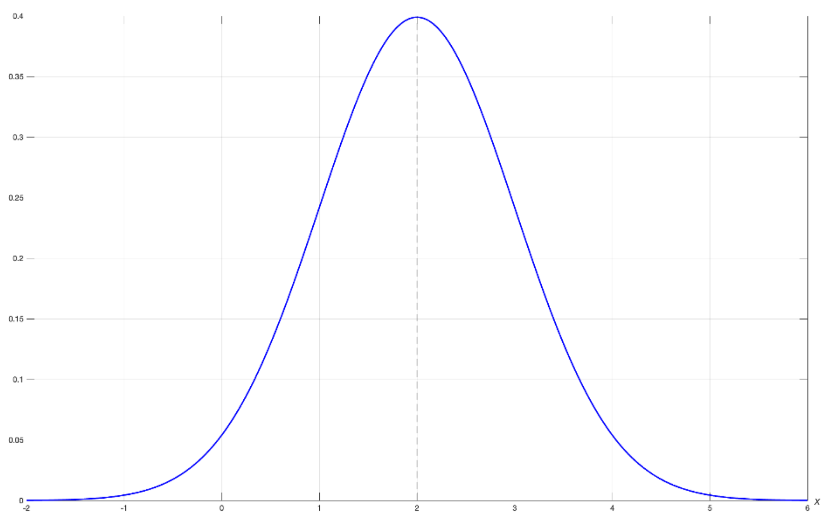
- 보충
- 확률변수의 평균
- 분산과 공분산
- 선형결합된 확률변수의 평균과 분산
- 체비셰프 정리

# 분산과 공분산 (1/25)

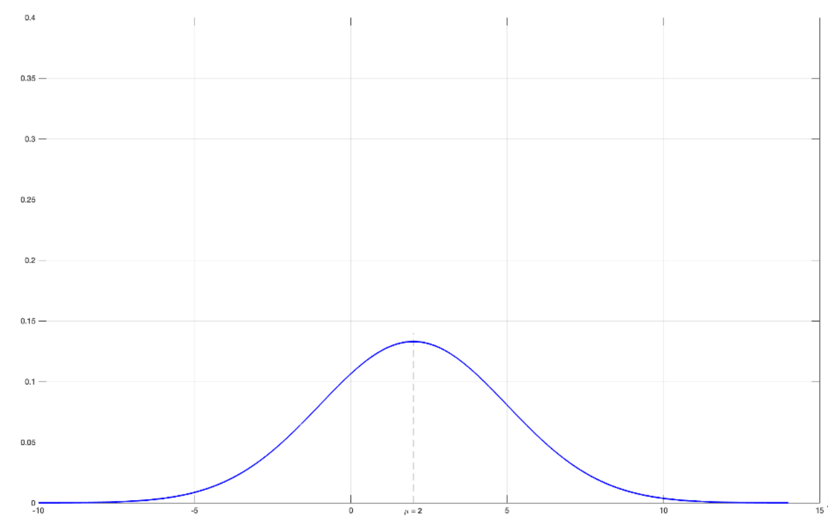
- 분산(Variance) (1/2)
  - 정의
    - 확률변수가 평균값으로부터 얼마나 퍼져 있는지를 측정하는 통계적 척도
  - 공식
    - 모분산($\sigma^2$): $\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$
    - 표본분산($s^2$): $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$

# 분산과 공분산 (2/25)

- 분산(Variance) (2/2)
  - 필요성
    - 분산은 데이터의 변동성을 측정함으로써 일관성, 불확실성, 위험도 등을 평가함
      - e.g., 주식 투자 시



<그림 4.2 분산이 1인 연속형 확률분포>



<그림 4.3 분산이 3인 연속형 확률분포>

# 분산과 공분산 (3/25)

- 확률변수 $X$의 분산 (1/4)
  - 정의
    - 확률변수 $X$가 확률분포 $f(x)$와 평균 $\mu$를 가질 때 확률변수의 값이 평균($\mu$)으로부터 벗어난 정도의 평균
  - 표현
    - $Var(X), {\sigma^2}_X$
  - 공식 (1/3)
    - $X$가 이산형인 경우: $\sigma^2 = E[(X-\mu)^2] = \sum_x (x-\mu)^2 f(x)$
    - $X$가 연속형인 경우: $\sigma^2 = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx$

# 분산과 공분산 (4/25)

- 확률변수 $X$의 분산 (2/4)
  - 공식 (2/3)
    - $\sigma^2 = E[(X-\mu)^2]$
      - $(x-\mu)$는 관측값의 평균으로부터의 편차(Deviation)임에 따라 $\sigma^2$는 편차들에 대한 제곱의 평균을 의미함
      - $x$(관측값)들이 $\mu$(평균)에 가까울수록 분산의 값은 작음
      - 분산의 양의 제곱근 $\sigma$는 $X$의 표준편차(Standard Deviation)임

# 분산과 공분산 (5/25)

- 확률변수 $X$의 분산 (3/4)
  - 공식 (3/3)
    - $\sigma^2 = E(X^2) - \mu^2$

**증명**

$$\sigma^2 = \sum_x (x-\mu)^2 f(x) = \sum_x (x^2 - 2\mu x + \mu^2) f(x)$$
$$= \sum_x x^2 f(x) - 2\mu \sum_x x f(x) + \mu^2 \sum_x f(x)$$

정의에 의해서 $\mu = \sum_x x f(x)$이고, $\sum_x f(x) = 1$이므로

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

# 분산과 공분산 (6/25)

- 확률변수 $X$의 분산 (4/4)
  - 필요성
    - 모분산과의 비교
      - 모분산은 실제 측정된 데이터의 변동성을 반영하는 반면, 확률변수의 분산은 이론적인 확률 모델의 변동성을 나타냄
      - 실제 데이터가 아직 발생하지 않았거나 관찰할 수 없는 경우, 확률변수의 분산을 통해 미래의 불확실성 추정 가능
        - e.g., 기상 예측

**모분산**

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$$

$$= (x_1 - \mu)^2\left(\frac{1}{N}\right) + (x_2 - \mu)^2\left(\frac{1}{N}\right) + \cdots + (x_N - \mu)^2\left(\frac{1}{N}\right)$$

**확률변수 $X$의 분산**

$$\sigma^2 = E[(X - \mu)^2] = \sum_{x}(x - \mu)^2 f(x)$$

$$= (x_1 - \mu)^2 f(x_1) + (x_2 - \mu)^2 f(x_2) + \cdots$$

# 분산과 공분산 (7/25)

- 예제 4.8

A와 B 두 회사에서 어느 날 사업목적으로 사용된 자동차의 수를 확률변수 $X$라고 하자. A 회사에 대한 확률분포와 B 회사에 대한 확률분포가 각각 다음과 같다.

| $x$ | 1 | 2 | 3 |
|---|---|---|---|
| $f(x)$ | 0.3 | 0.4 | 0.3 |

| $x$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $f(x)$ | 0.2 | 0.1 | 0.3 | 0.3 | 0.1 |

<u>**B 회사에 대한 확률분포의 분산이 A 회사의 분산보다 큼을 보여라**</u>

- A 회사
  - $\mu_A = E(X) = (1)(0.3) + (2)(0.4) + (3)(0.3) = 2.0$
  - ${\sigma_A}^2 = \sum_{x=1}^{3}(x-2)^2 f(x)$
    $= (1-2)^2(0.3) + (2-2)^2(0.4) + (3-2)^2(0.3) = 0.6$

- B 회사
  - $\mu_B = E(X) = (0)(0.2) + (1)(0.1) + (2)(0.3) + (3)(0.3) + (4)(0.1) = 2.0$
  - ${\sigma_A}^2 = \sum_{x=0}^{4}(x-2)^2 f(x)$
    $= (0-2)^2(0.2) + (1-2)^2(0.1) + (2-2)^2(0.3) + (3-2)^2(0.3) + (4-2)^2(0.1) = 1.6$

- 예제 4.9

생산라인으로부터 3개의 부품을 추출하여 검사하였을 때 결함이 있는 부품의 수를 확률변수 $X$라고 하자. **$\boldsymbol{X}$의 확률분포**가 다음과 같을 때, '$\boldsymbol{\sigma^2 = E(X^2) - \mu^2}$'**를 이용하여 $\boldsymbol{\sigma^2}$을 계산**하라

| $x$ | **0** | **1** | **2** | **3** |
|---|---|---|---|---|
| $f(x)$ | 0.51 | 0.38 | 0.10 | 0.01 |

- 평균 계산
  - $\mu = (0)(0.51) + (1)(0.38) + (2)(0.10) + (3)(0.01) = 0.61$
- $E(X^2)$ 계산
  - $E(X^2) = (0)(0.51) + (1)(0.38) + (4)(0.10) + (9)(0.01) = 0.87$
- $\sigma^2$ 계산
  - $\sigma^2 = E(X^2) - \mu^2 = 0.87 - (0.61)^2 = 0.4979$

# 분산과 공분산 (9/25)

- 예제 4.10

어느 연쇄점의 주당 콜라의 수요(단위: 1000리터)가 다음과 같은 **확률분포를 가지는 확률변수 $X$**라고 할 때, **$X$의 평균과 분산**을 구하라

$$f(x) = \begin{cases} 2(x-1), & 1 < x < 2 \\ 0, & \text{otherwise} \end{cases}$$

- 평균 계산
  - $\mu = E(X) = 2\int_1^2 x(x-1)dx = \frac{5}{3}$
- $E(X^2)$ 계산
  - $E(X^2) = 2\int_1^2 x^2(x-1)dx = \frac{17}{6}$
- $\sigma^2$ 계산
  - $\sigma^2 = E(X^2) - \mu^2 = \frac{17}{6} - \left(\frac{5}{3}\right)^2 = \frac{1}{18}$

# 분산과 공분산 (10/25)

- 확률변수 $g(X)$의 분산
  - 정의

    ❖ $g(X)$: 확률변수 $X$에 종속되는 확률변수

    - 확률변수 $X$가 확률분포 $f(x)$를 가질 때 확률변수 $g(X)$의 값이 평균($\mu_{g(X)}$)으로부터 벗어난 정도의 평균
  - 공식
    - $X$가 이산형인 경우
      :$\sigma^2{}_{g(X)} = E\left\{\left[g(X) - \mu_{g(X)}\right]^2\right\} = \sum_x \left[g(x) - \mu_{g(X)}\right]^2 f(x)$
    - $X$가 연속형인 경우
      :$\sigma^2{}_{g(X)} = E\left\{\left[g(X) - \mu_{g(X)}\right]^2\right\} = \int_{-\infty}^{\infty} \left[g(x) - \mu_{g(X)}\right]^2 f(x)\, dx$

# • 예제 4.11

**<u>확률변수 $X$의 확률분포</u>**가 다음과 같이 주어졌을 때 <u>$g(X) = 2X + 3$</u>의 분산을 계산하라

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $f(x)$ | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{2}$ | $\frac{1}{8}$ |

- 평균 계산
  - $\mu_{2X+3} = E(2X+3) = \sum_{x=0}^{3}(2x+3)f(x) = 6$
- $\sigma^2$ 계산
  - ${\sigma^2}_{2X+3} = E\{[(2X+3) - \mu_{2X+3}]^2\} = E[(2X+3-6)^2]$

$$= E(4X^2 - 12X + 9) = \sum_{x=0}^{3}(4x^2 - 12x + 9)f(x) = 4$$

- 예제 4.11 – 실생활 적용 예시

한 시간 동안 판매되는 음료의 총 개수를 확률변수 $X$라고 할 때, **확률변수 $X$의 확률분포**는 다음과 같다.

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $f(x)$ | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{2}$ | $\frac{1}{8}$ |

판매된 음료 개수에 따라 카페가 얻는 총 수익 (단위: 백 원)은 $g(X) = 2X + 3$라고 할 때, **한 시간 동안 카페 수익의 평균과 분산을 구하라**

- 평균 계산
  - $\mu_{2X+3} = E(2X+3) = \sum_{x=0}^{3}(2x+3)f(x) = 6$

- $\sigma^2$ 계산
  - $\sigma^2{}_{2X+3} = E\{[(2X+3) - \mu_{2X+3}]^2\} = E[(2X+3-6)^2]$

$$= E(4X^2 - 12X + 9) = \sum_{x=0}^{3}(4x^2 - 12x + 9)f(x) = 4$$

# 분산과 공분산 (13/25)

- 예제 4.12

$X$를 예제 4.5의 **확률밀도함수** $f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2 \\ 0, & \text{otherwise} \end{cases}$ 를 가지는 확률변수라고 할 때, **확률변수 $g(X) = 4X + 3$의 분산**을 구하라

- 예제 4.5를 통해 $\mu_{4X+3} = 8$임을 파악함
- $\sigma^2{}_{4X+3} = E\{[(4X+3) - \mu_{4X+3}]^2\} = E\{[(4X+3) - 8]^2\} = E[(4X-5)^2]$

$$= \int_{-1}^{2} (4x-5)^2 \frac{x^2}{3} dx = \frac{1}{3}\int_{-1}^{2} (16x^4 - 40x^3 + 25x^2) dx = \frac{51}{5}$$

# 분산과 공분산 (14/25)

- 예제 4.12 – 실생활 적용 예시

확률변수 $X$를 한 달간의 비만율 변화량이라고 할 때, **확률변수 $\boldsymbol{X}$의 밀도함수**가

$$f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2 \\ 0, & \text{otherwise} \end{cases}$$

이다.
$\boldsymbol{g(X) = 4X + 3}$가 건강 보험료 변화량(단위: 만원)을 나타낸다고 할 때, 건강 보험료 변화량의 **기대값과 분산**을 구하라

- $E[g(X)] = E(4X+3) = \int_{-1}^{2} \frac{(4x+3)x^2}{3} dx = \frac{1}{3}\int_{-1}^{2}(4x^3+3x^2)dx = 8$
- $\sigma^2{}_{4X+3} = E\{[(4X+3) - \mu_{4X+3}]^2\} = E\{[(4X+3)-8]^2\} = E[(4X-5)^2]$

$$= \int_{-1}^{2}(4x-5)^2\frac{x^2}{3}dx = \frac{1}{3}\int_{-1}^{2}(16x^4 - 40x^3 + 25x^2)dx = \frac{51}{5}$$

# 분산과 공분산 (15/25)

- 공분산(Covariance) (1/5)
  - 정의
    - 두 확률변수 간의 관련성의 척도
  - 표현
    - $\sigma_{XY}, \mathrm{Cov}(X, Y)$
  - 공식 (1/2)
    - $X$와 $Y$가 이산형인 경우
      :$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = \sum_x \sum_y (x - \mu_X)(y - \mu_Y) f(x, y)$
    - $X$와 $Y$가 연속형인 경우
      :$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \big((x - \mu_X)(y - \mu_Y)\big) f(x, y) dx\, dy$

# 분산과 공분산 (16/25)

- 공분산(Covariance) (2/5)
  - 공식 (2/2)
    - $\mu_X$와 $\mu_Y$를 평균으로 하는 두 확률변수 $X$와 $Y$의 공분산인 경우
      : $\sigma_{XY} = E(XY) - \mu_X\mu_Y$

**증명**

$$\sigma_{XY} = \sum_x \sum_y (x - \mu_X)(y - \mu_Y) f(x, y)$$

$$= \sum_x \sum_y xyf(x, y) - \mu_X \sum_x \sum_y yf(x, y) - \mu_Y \sum_x \sum_y xf(x, y) + \mu_X\mu_Y \sum_x \sum_y f(x, y)$$

이때, $\mu_X = \sum_x \sum_y xf(x, y)$, $\mu_Y = \sum_x \sum_y yf(x, y)$이고, $\sum_x \sum_y f(x, y) = 1$이므로

$$\sigma_{XY} = E(XY) - \mu_X\mu_Y - \mu_Y\mu_X + \mu_X\mu_Y = E(XY) - \mu_X\mu_Y$$

# 분산과 공분산 (17/25)

- 공분산(Covariance) (3/5)
  - 필요성
    - 공분산을 통해 두 변수 간의 선형 관계의 방향성을 파악하여 데이터 상호 연관성을 이해하고 한 변수의 변화가 다른 변수에 어떤 영향을 미칠지 예측할 수 있음
  - 공분산을 통한 두 확률변수 간의 관련성 분석 (1/2)
    - $X$의 값이 클 때 $Y$의 값도 크고, $X$의 값이 작을 때 $Y$의 값도 작다면 $(X-\mu_X)(Y-\mu_Y)$는 양의 값을 가짐
    - $X$의 값이 클 때 $Y$의 값이 작고, $X$의 값이 작을 때 $Y$의 값이 크면 $(X-\mu_X)(Y-\mu_Y)$는 음의 값을 가짐

# 분산과 공분산 (18/25)

- 공분산(Covariance) (4/5)

❖ $(X-\mu_X)(Y-\mu_Y)$

  - 공분산을 통한 두 확률변수 간의 관련성 분석 (2/2)

**양의 상관**

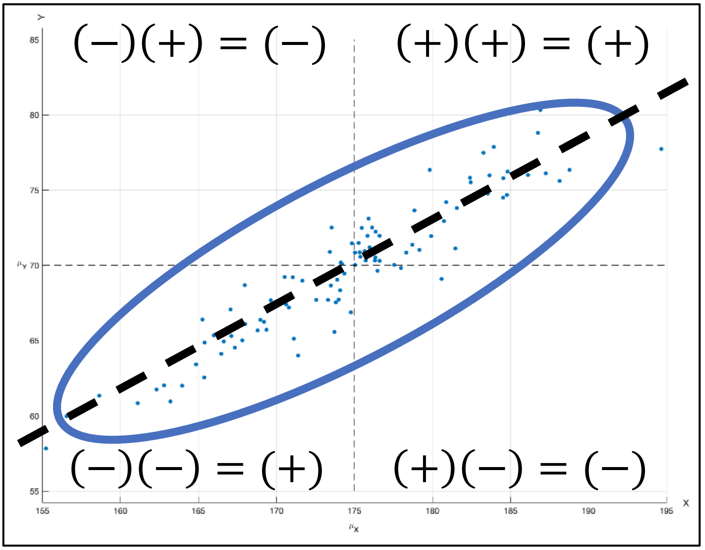
- 온도와 에어컨 사용량
- 칼로리 섭취량과 체중

**독립**

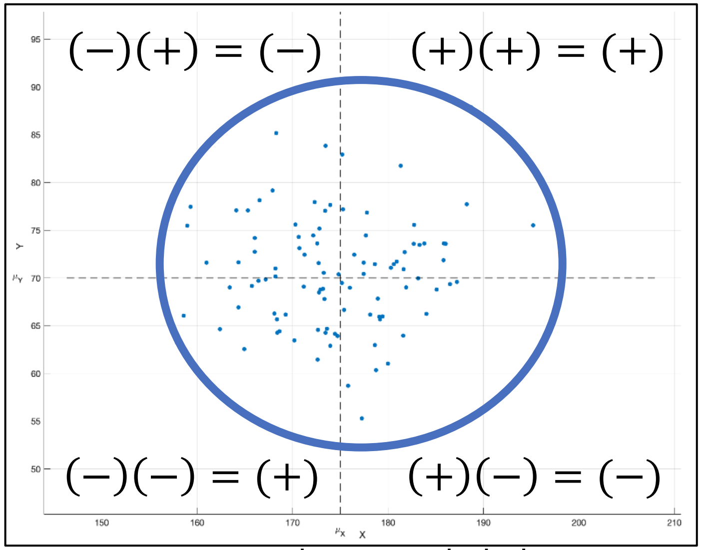
- 어느 날의 최고 기온과 로또 번호
- 티셔츠 색깔과 냉장고 속 당근의 개수

**음의 상관**

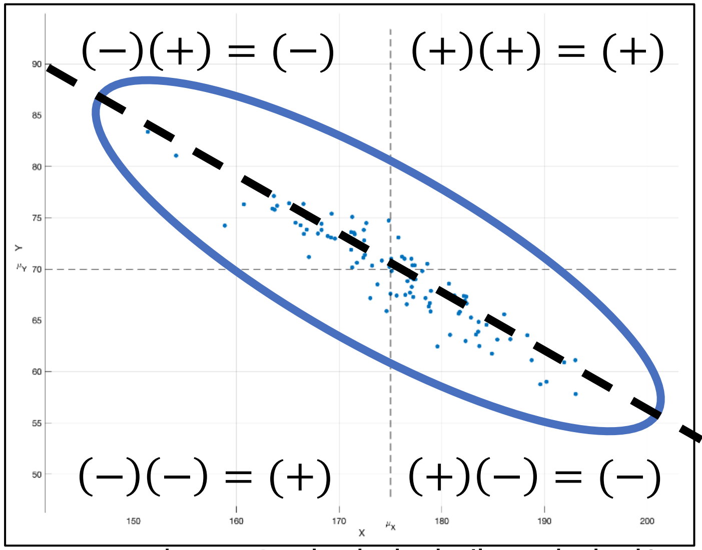
- 제품의 가격과 수요
- 자동차의 속도와 소요 시간




<그림 4.4 양의 선형 관계를 나타내는 변수 $X$와 $Y$의 산점도>




<그림 4.5 독립적인 변수 $X$와 $Y$의 산점도>




<그림 4.6 음의 선형 관계를 나타내는 변수 $X$와 $Y$의 산점도>

# 분산과 공분산 (19/25)

- 예제 4.13

어떤 상자에서 임의로 2개의 볼펜을 꺼낼 때, **<u>청색볼펜의 수 $X$와 적색볼펜의 수 $Y$의 결합확률분포</u>**가 다음과 같다. **<u>$X$와 $Y$의 공분산</u>**을 구하라

| $f(x,y)$ | | $x$ = 0 | $x$ = 1 | $x$ = 2 | $h(y)$ |
|---|---|---|---|---|---|
| $y$ | 0 | $\frac{3}{28}$ | $\frac{9}{28}$ | $\frac{3}{28}$ | $\frac{15}{28}$ |
| | 1 | $\frac{3}{14}$ | $\frac{3}{14}$ | 0 | $\frac{3}{7}$ |
| | 2 | $\frac{1}{28}$ | 0 | 0 | $\frac{1}{28}$ |
| $g(x)$ | | $\frac{5}{14}$ | $\frac{15}{28}$ | $\frac{3}{28}$ | 1 |

- 예제 4.6으로부터 $E(XY) = 3/14$
- $\mu_X = \sum_{x=0}^{2} xg(x) = (0)\left(\frac{5}{14}\right) + (1)\left(\frac{15}{28}\right) + (2)\left(\frac{3}{28}\right) = \frac{3}{4}$
- $\mu_Y = \sum_{y=0}^{2} yh(y) = (0)\left(\frac{15}{28}\right) + (1)\left(\frac{3}{7}\right) + (2)\left(\frac{1}{28}\right) = \frac{1}{2}$
- 따라서, $\sigma_{XY} = E(XY) - \mu_x\mu_Y = \frac{3}{14} - \left(\frac{3}{4}\right)\left(\frac{1}{2}\right) = -\frac{9}{56}$

# 분산과 공분산 (20/25)

- 예제 4.14

마라톤코스를 완주한 남자의 비율 $X$와 여자의 비율 $Y$의 **<u>결합확률분포</u>**가 다음과 같다. **<u>$X$와 $Y$의 공분산</u>**을 구하라

$$f(x,y) = \begin{cases} 8xy, & 0 \le y \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

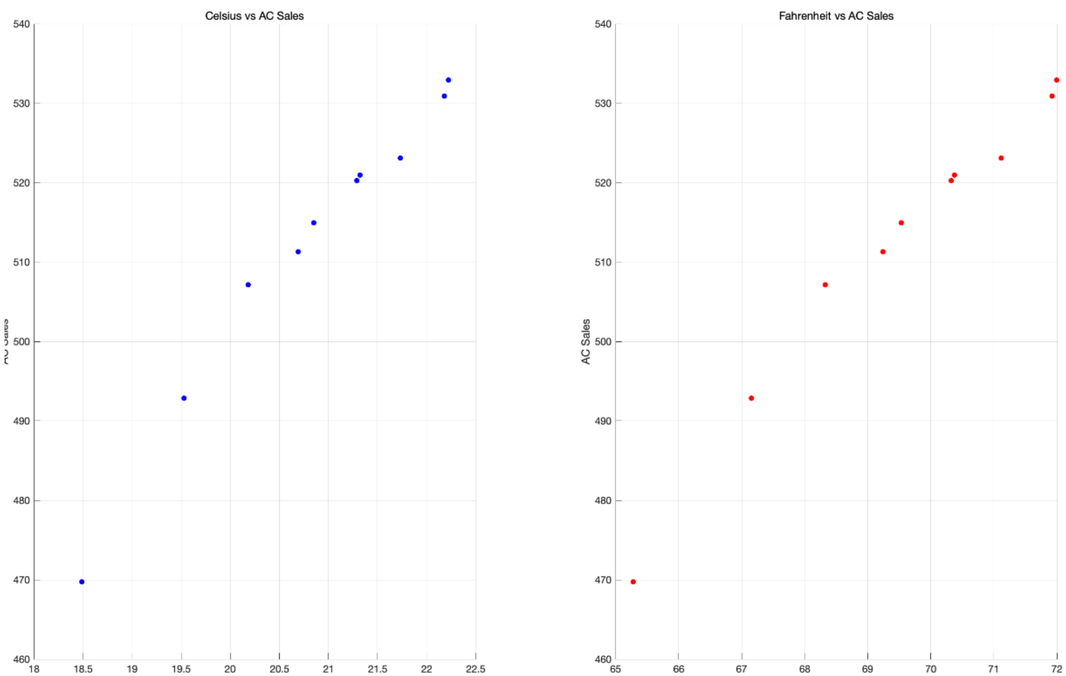
- 주변밀도함수 계산
  - $g(x) = \begin{cases} 4x^3, & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$
  - $h(y) = \begin{cases} 4y(1-y^2), & 0 \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$
- $\mu_x, \mu_Y$ 계산
  - $\mu_X = E(X) = \int_0^1 4x^4 dx = \frac{4}{5}$
  - $\mu_Y = E(Y) = \int_0^1 4y^2(1-y^2)dy = \frac{8}{15}$
- $E(XY)$ 계산
  - $E(XY) = \int_0^1 \int_y^1 8x^2y^2\, dx\, dy = \frac{4}{9}$
- 공분산 계산
  - $\sigma_{XY} = E(XY) - \mu_X\, \mu_Y = \frac{4}{9} - \left(\frac{4}{5}\right)\left(\frac{8}{15}\right) = \frac{4}{225}$

# 분산과 공분산 (21/25)

- 공분산(Covariance) (5/5)

❖$F = \frac{9}{5}C + 32$

  - 한계점
    - 공분산($\sigma_{XY}$) 값은 $X$와 $Y$의 측정단위에 따라 달라지므로 관련성의 강도를 나타내지 못함
      - e.g., 일일 최고 기온과 에어컨 판매량의 공분산




<그림 4.7 섭씨 및 화씨 기온과 에어컨 판매량에 대한 산점도 그래프>

- 섭씨 기온 데이터
  - [18.49, 21.73, 20.69, 22.22, 21.29, 22.18, 20.85, 19.53, 21.32, 20.18] °C
- 화씨로 변환된 기온 데이터
  - [65.28, 71.11, 69.24, 72.00, 70.32, 71.92, 69.53, 67.15, 70.38, 68.32] °F

- 공분산
  - **섭씨 기온과 에어컨 판매량: 188.61**
  - **화씨 기온과 에어컨 판매량: 339.49**

# 분산과 공분산 (22/25)

- 상관계수(Correlation Coefficient) (1/2)
  - 정의
    - 두 확률변수 간의 선형적 관계의 방향과 강도를 나타내는 척도
  - 표현
    - $\rho_{XY}$
  - 공식
    - 확률변수 $X$와 $Y$의 공분산이 $\sigma_{XY}$이고, 표준편차가 각각 $\sigma_X$, $\sigma_Y$일 경우, $X$와 $Y$의 상관계수

    : $\rho_{XY} = \dfrac{\sigma_{XY}}{\sigma_X \sigma_Y}$

- 상관계수(Correlation Coefficient) (2/2)
  - 필요성
    - $X$와 $Y$의 측정 단위에 따라 값이 달라지는 공분산($\sigma_{XY}$)과 달리, 상관계수($\rho_{XY}$)는 단위에 무관하게 관계의 강도를 정량적으로 파악할 수 있음
  - 특징
    - 상관계수 값의 범위는 $-1 \leq \rho_{XY} \leq 1$
    - 공분산($\sigma_{XY}$)이 0이면 상관계수($\rho_{XY}$)값도 0

# 분산과 공분산 (24/25)

- 예제 4.15

예제 4.13에서 ***X*와 *Y*의 상관계수**를 구하라
($X$와 $Y$의 결합확률분포는 <표 4.1>과 같다)

- 예제 4.13으로부터 $\mu_X = \frac{3}{4}, \mu_Y = \frac{1}{2}, \sigma_{XY} = -\frac{9}{56}$
- $E(X^2), E(Y^2)$ 계산
  - $E(X^2) = (0^2)\left(\frac{5}{14}\right) + (1^2)\left(\frac{15}{28}\right) + (2^2)\left(\frac{3}{28}\right) = \frac{27}{28}$
  - $E(Y^2) = (0^2)\left(\frac{15}{28}\right) + (1^2)\left(\frac{3}{7}\right) + (2^2)\left(\frac{1}{28}\right) = \frac{4}{7}$
- 분산 계산
  - $\sigma^2{}_X = \frac{27}{28} - \left(\frac{3}{4}\right)^2 = \frac{45}{112}$
  - $\sigma^2{}_Y = \frac{4}{7} - \left(\frac{1}{2}\right)^2 = \frac{9}{28}$
- 상관계수 계산
  - $\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{-9/56}{\sqrt{(45/112)(9/28)}} = -\frac{1}{\sqrt{5}}$

<표 4.1 예제 4.13 - $X$와 $Y$의 결합확률분포>

| $\boldsymbol{f(x,y)}$ | | $\boldsymbol{x}$ | | | $h(y)$ |
|---|---|---|---|---|---|
| | | 0 | 1 | 2 | |
| $\boldsymbol{y}$ | 0 | $\frac{3}{28}$ | $\frac{9}{28}$ | $\frac{3}{28}$ | $\frac{15}{28}$ |
| | 1 | $\frac{3}{14}$ | $\frac{3}{14}$ | 0 | $\frac{3}{7}$ |
| | 2 | $\frac{1}{28}$ | 0 | 0 | $\frac{1}{28}$ |
| $g(x)$ | | $\frac{5}{14}$ | $\frac{15}{28}$ | $\frac{3}{28}$ | 1 |

# 분산과 공분산 (25/25)

- 예제 4.16

예제 4.14에서 ***X*와 *Y*의 상관계수**를 구하라

($X, Y$의 결합확률분포는 $f(x,y) = \begin{cases} 8xy, & 0 \le y \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$)

- 예제 4.14로부터 $\mu_X = \frac{3}{4}, \mu_Y = \frac{1}{2}, \sigma_{XY} = -\frac{9}{56}$
- $E(X^2), E(Y^2)$ 계산
  - $E(X^2) = \int_0^1 4x^5 dx = \frac{2}{3}$
  - $E(Y^2) = \int_0^1 4y^3(1-y^2)dy = 1 - \frac{2}{3} = \frac{1}{3}$
- 분산 계산
  - $\sigma^2{}_X = \frac{2}{3} - \left(\frac{4}{5}\right)^2 = \frac{2}{75}$
  - $\sigma^2{}_Y = \frac{1}{3} - \left(\frac{8}{15}\right)^2 = \frac{11}{225}$
- 상관계수 계산
  - $\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{4/225}{\sqrt{(2/75)(11/225)}} = \frac{4}{\sqrt{66}}$

# 목 차

- 보충
- 확률변수의 평균
- 분산과 공분산
- 선형결합된 확률변수의 평균과 분산
- 체비셰프 정리

# 선형결합된 확률변수의 평균과 분산 (1/15)

- 선형 결합
  - 정의
    - 여러 변수들이 각각의 스칼라 계수와 곱해진 후, 그 결과들이 합쳐진 형태
  - 특징
    - 변수 $X_1, X_2, \cdots, X_n$과 각 변수에 대응하는 계수 $a_1, a_2, \cdots, a_n$에 대해 $a_1X_1 + a_2X_2 + \cdots + a_nX_n$으로 표현됨
    - 선형결합된 확률변수의 경우, 두 개 이상의 확률변수가 각각의 스칼라 계수와 곱해진 후 합해져서 새로운 확률변수를 형성하는 것을 의미함
      - e.g., $Z = aX + bY$

# 선형결합된 확률변수의 평균과 분산 (2/15)

- 확률변수의 평균을 손쉽게 계산하기 위한 성질 (1/4)
  - $a$와 $b$가 상수이면 $E(aX+b) = aE(X) + b$

**증명**

기대값의 정의에 따라

$$E(aX+b) = \int_{-\infty}^{\infty} (ax+b) f(x)\, dx = a\int_{-\infty}^{\infty} x f(x)\, dx + b\int_{-\infty}^{\infty} f(x)\, dx$$

$\int_{-\infty}^{\infty} x f(x)\, dx = E(X)$이고, $\int_{-\infty}^{\infty} f(x)\, dx = 1$ 이므로

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

이다.

**따름정리 4.1**

$a = 0$ 으로 놓으면 $E(b) = b$가 된다.

**따름정리 4.2**

$b = 0$으로 놓으면 $E(aX) = aE(X)$가 된다.

# 선형결합된 확률변수의 평균과 분산 (3/15)

- 예제 4.17

**'$a$와 $b$가 상수이면 $E(aX+b)=aE(X)+b$이다'**를 이산형 확률변수 **$g(X)=2X-1$에 적용**하여 예제 4.4를 다시 풀어 보라

- 예제 4.4

어느 쾌청한 금요일 오후 4시에서 5시 사이에 세차장에서 서비스를 받는 차의 수를 $X$라고 할 때, **$X$의 확률분포**가 다음과 같다고 하자.

| $x$ | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|
| $P(X=x)$ | $\frac{1}{12}$ | $\frac{1}{12}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |

**$g(X)=2X-1$**을 종업원이 받는 수당(단위: 달러)이라고 할 때, 이 시간대의 종업원의 기대수익을 구하라

- $E[g(X)]=E(2X-1)=\sum_{x=4}^{9}(2x-1)f(x)$

$$=(7)\left(\frac{1}{12}\right)+(9)\left(\frac{1}{12}\right)+(11)\left(\frac{1}{4}\right)+(13)\left(\frac{1}{4}\right)+(15)\left(\frac{1}{16}\right)+(17)\left(\frac{1}{6}\right)$$
$$=\$12.67$$

- $E[g(X)]=E(2X-1)=2E(X)-1$
- $\mu=E(X)=\sum_{x=4}^{9}xf(x)=(4)\left(\frac{1}{12}\right)+(5)\left(\frac{1}{12}\right)+(6)\left(\frac{1}{4}\right)+(7)\left(\frac{1}{4}\right)+(8)\left(\frac{1}{6}\right)+(9)\left(\frac{1}{6}\right)=\frac{41}{6}$
- $\mu_{2X-1}=2E(X)-1=(2)\left(\frac{41}{6}\right)-1=\$12.67$

# 선형결합된 확률변수의 평균과 분산 (4/15)

- 예제 4.18

**'$a$와 $b$가 상수이면 $E(aX+b) = aE(X)+b$이다'**를 연속형 확률변수 $g(X) = 4X+3$ **에 적용**하여 예제 4.5를 다시 풀어 보라

- 예제 4.5

확률변수 X의 밀도함수가

$$f(x) = \begin{cases} \frac{x^3}{3}, & -1 < x < 2 \\ 0, & \text{otherwise} \end{cases}$$

일 때, $g(X) = 4X+3$의 기대값을 구하라

- $E[g(X)] = E(4X+3)$

$$= \int_{-1}^{2} \frac{(4x+3)x^2}{3} dx$$

$$= \frac{1}{3}\int_{-1}^{2} (4x^3+3x^2) dx = 8$$

- $E[g(X)] = E(4X+3) = 4E(X)+3$
- $\mu = E(X) = \int_{-1}^{2} x\left(\frac{x^2}{3}\right) dx = \int_{-1}^{2} \frac{x^3}{3} dx = \frac{5}{4}$
- $\mu_{4X+3} = 4E(X)+3 = 4\left(\frac{5}{4}\right)+3 = 8$

# 선형결합된 확률변수의 평균과 분산 (5/15)

- 확률변수의 평균을 손쉽게 계산하기 위한 성질 (2/4)
  - 두 개 이상의 확률변수 $X$의 함수의 합이나 차의 기대값은 각 함수의 기대값의 합이나 차와 동일함
    : $E[g(X) \pm h(X)] = E[g(X)] \pm E[h(X)]$

**증명**

정의에 의하여

$$E[g(X) \pm h(X)] = \int_{-\infty}^{\infty} [g(X) \pm h(X)] f(x)\, dx$$

$$= \int_{-\infty}^{\infty} g(x) f(x) dx \pm \int_{-\infty}^{\infty} h(x) f(x)\, dx = E[g(X)] \pm E[h(X)]$$

# 선형결합된 확률변수의 평균과 분산 (6/15)

- 예제 4.19

$X$가 다음과 같은 확률분포를 가지는 확률변수일 때, **$Y = (X-1)^2$의 기대값**을 구하라

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $f(x)$ | $\frac{1}{3}$ | $\frac{1}{2}$ | 0 | $\frac{1}{6}$ |

- $E(Y) = E[(X-1)^2] = E(X^2 - 2X + 1) = E(X^2) - 2E(X) + E(1)$
- $E(X) = (0)\left(\frac{1}{3}\right) + (1)\left(\frac{1}{2}\right) + (2)(0) + (3)\left(\frac{1}{6}\right) = 1$
- $E(X^2) = (0)\left(\frac{1}{3}\right) + (1)\left(\frac{1}{2}\right) + (4)(0) + (9)\left(\frac{1}{6}\right) = 2$
- 따라서, $E[(X-1)^2] = E(X^2) - 2E(X) + E(1) = 2 - (2)(1) + 1 = 1$

- 예제 4.20

어느 연쇄점에서 판매하는 어떤 음료의 주당 수요(단위: 1000리터)가 **연속형 확률변수 $\boldsymbol{g(X) = X^2 + X - 2}$**로 나타내며, $X$는 다음과 같은 확률분포를 가진다고 할 때, **음료수의 주당 수요의 기대값**을 구하라

$$f(x) = \begin{cases} 2(x-1), & 1 < x < 2 \\ 0, & \text{otherwise} \end{cases}$$

- $E[g(X)] = E(X^2 + X - 2) = E(X^2) + E(X) - E(2)$
- $E(X) = \int_1^2 2x(x-1)\,dx = 2\int_1^2 (x^2 - x)\,dx = \frac{5}{3}$
- $E(X^2) = \int_1^2 2x^2(x-1)\,dx = 2\int_1^2 (x^3 - x^2)\,dx = \frac{17}{6}$
- 따라서, $E(X^2 + X - 2) = E(X^2) + E(X) - E(2) = \frac{17}{6} + \frac{5}{3} - 2 = \frac{5}{2}$
- 단위는 1000리터 이므로, $\frac{5}{2}$(1000리터) = 2500리터

# 선형결합된 확률변수의 평균과 분산 (8/15)

- 확률변수의 평균을 손쉽게 계산하기 위한 성질 (3/4)
  - 확률변수 $X$와 $Y$의 함수들의 합이나 차의 기대값은 각 함수의 기대값의 합이나 차와 동일함
    : $E[g(X,Y) \pm h(X,Y)] = E[g(X,Y)] \pm E[h(X,Y)]$

**증명**

$$E[g(X,Y) \pm h(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[g(x,y) \pm h(x,y)]\,f(x,y)dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)\,dx\,dy \pm \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)f(x,y)\,dx\,dy$$

$$= E[g(X,Y)] \pm E[h(X,Y)]$$

**따름정리 4.3**

$g(X,Y) = g(X)$이고, $h(X,Y) = h(Y)$라 하면, $E[g(X) \pm h(X)] = E[g(X)] \pm E[h(Y)]$

**따름정리 4.4**

$g(X,Y) = X$이고, $h(X,Y) = Y$라 하면, $E(X \pm Y) = E(X) \pm E(Y)$

# 선형결합된 확률변수의 평균과 분산 (9/15)

- 확률변수의 평균을 손쉽게 계산하기 위한 성질 (4/4)
  - 두 확률변수 $X$와 $Y$가 서로 독립인 경우,
    : $E(XY) = E(X)E(Y)$

**증명**

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dx\,dy$$

$X$와 $Y$의 주변분포를 각각 $g(x)$와 $h(y)$라고 하면, $X$와 $Y$는 독립이므로 $f(x,y) = g(x)h(y)$가 된다. 따라서,

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyg(x)h(y)dx\,dy$$

$$= \int_{-\infty}^{\infty} xg(x)\,dx \int_{-\infty}^{\infty} yh(y)\,dy = E(X)E(Y)$$

**따름정리 4.5**

확률변수 $X$와 $Y$가 서로 독립이면 $\sigma_{XY} = 0$이다.

# 선형결합된 확률변수의 평균과 분산 (10/15)

## • 예제 4.21

$X$와 $Y$를 서로 독립이고 다음과 같은 결합확률분포를 가지는 확률변수라 할 때, 정리 4.8의 $\underline{\boldsymbol{E(XY) = E(X)E(Y)}}$가 성립함을 증명하라

$$f(x,y) = \begin{cases} \dfrac{x(1+3y^2)}{4}, & 0 < x < 2, 0 < y < 1 \\ 0, & \text{otherwise} \end{cases}$$

- $E(XY) = \int_0^1 \int_0^2 xyf(x,y)dx\,dy = \int_0^1 \int_0^2 \frac{x^2y(1+3y^2)}{4} dx\,dy = \int_0^1 \frac{x^3y(1+3y^2)}{12} \Big|_{x=0}^{x=2} dy = \int_0^1 \frac{2y(1+3y^2)}{3} dy = \frac{5}{6}$
- $E(X) = \int_0^1 \int_0^2 xf(x,y)dx\,dy = \int_0^1 \int_0^2 \frac{x^2(1+3y^2)}{4} dx\,dy = \int_0^1 \frac{x^3(1+3y^2)}{12} \Big|_{x=0}^{x=2} dy = \int_0^1 \frac{2(1+3y^2)}{3} dy = \frac{4}{3}$
- $E(Y) = \int_0^1 \int_0^2 yf(x,y)dx\,dy = \int_0^1 \int_0^2 \frac{xy(1+3y^2)}{4} dx\,dy = \int_0^1 \frac{x^2y(1+3y^2)}{8} \Big|_{x=0}^{x=2} dy = \int_0^1 \frac{y(1+3y^2)}{2} dy = \frac{5}{8}$
- 따라서, $E(X)E(Y) = \left(\frac{4}{3}\right)\left(\frac{5}{8}\right) = \frac{5}{6} = E(XY)$

# 선형결합된 확률변수의 평균과 분산 (11/15)

- 확률변수의 분산을 손쉽게 계산하기 위한 성질 (1/4)
  - $X$와 $Y$가 결합확률분포 $f(x,y)$를 가지는 확률변수이고, $a,b,c$가 상수일 때
    : $\sigma^2{}_{aX+bY+c} = a^2\sigma^2{}_X + b^2\sigma^2{}_Y + 2ab\sigma_{XY}$

**증명**

정의에 의하여

$$\sigma^2{}_{aX+bY+c} = E\{[(aX+bY+c)-\mu_{aX+bY+c}]^2\}$$

이다.

'$E(aX) = aE(X)$', '$E(X \pm Y) = E(X) \pm E(Y)$'의 성질에 의하여

$$\mu_{aX+bY+c} = E(aX+bY+c) = aE(X) + bE(Y) + c = a\mu_x + b\mu_Y + c$$

가 되므로,

$$\begin{aligned}\sigma^2{}_{aX+bY+c} &= E\{[a(X-\mu_X)+b(Y-\mu_Y)]^2\}\\ &= a^2E[(X-\mu_X)^2] + a^2E[(Y-\mu_Y)^2] + 2abE[(X-\mu_X)(Y-\mu_Y)]\\ &= a^2\sigma^2{}_X + b^2\sigma^2{}_Y + 2ab\sigma_{XY}\end{aligned}$$

가 된다.

# 선형결합된 확률변수의 평균과 분산 (12/15)

- 확률변수의 분산을 손쉽게 계산하기 위한 성질 (2/4)
  - $X$와 $Y$가 결합확률분포 $f(x,y)$를 가지는 확률변수이고, $a, b, c$가 상수일 때
    : $\sigma^2{}_{aX+bY+c} = a^2\sigma^2{}_X + b^2\sigma^2{}_Y + 2ab\sigma_{XY}$

**따름정리 4.6**

$b = 0$으로 놓으면 $\sigma^2{}_{aX+c} = a^2\sigma^2{}_X = a^2\sigma^2$이 된다.

**따름정리 4.7**

$a = 1, b = 0$으로 놓으면, $\sigma^2{}_{X+c} = \sigma^2{}_X = \sigma^2$이 된다.

**따름정리 4.8**

$b = 0, c = 0$으로 놓으면, $\sigma^2{}_{aX} = a^2\sigma^2{}_X = a^2\sigma^2$이 된다.

# 선형결합된 확률변수의 평균과 분산 (13/15)

- 확률변수의 분산을 손쉽게 계산하기 위한 성질 (3/4)
  - $X$와 $Y$가 결합확률분포 $f(x, y)$를 가지는 확률변수이고, $a, b, c$가 상수일 때
    : $\sigma^2{}_{aX+bY+c} = a^2\sigma^2{}_X + b^2\sigma^2{}_Y + 2ab\sigma_{XY}$

**따름정리 4.9**

$X$와 $Y$를 독립인 확률변수라고 하면,

$$\sigma^2{}_{aX+bY} = a^2\sigma^2{}_X + b^2\sigma^2{}_Y$$

이다.

**따름정리 4.10**

$X$와 $Y$를 독립인 확률변수라고 하면,

$$\sigma^2{}_{aX-bY} = a^2\sigma^2{}_X + b^2\sigma^2{}_Y$$

이다.

- 확률변수의 분산을 손쉽게 계산하기 위한 성질 (4/4)
  - $X$와 $Y$가 결합확률분포 $f(x,y)$를 가지는 확률변수이고, $a,b,c$가 상수일 때
    : ${\sigma^2}_{aX+bY+c} = a^2{\sigma^2}_X + b^2{\sigma^2}_Y + 2ab\sigma_{XY}$

**따름정리 4.11**

$X_1, X_2, \cdots, X_n$을 독립인 확률변수라 하면,

$${\sigma^2}_{a_1X_1+a_2X_2+\cdots+a_nX_n} = {a^2}_1{\sigma^2}_{X_1} + {a^2}_2{\sigma^2}_{X_2} + \cdots + {a^2}_n{\sigma^2}_{X_n}$$

이다.

# 선형결합된 확률변수의 평균과 분산 (15/15)

- 예제 4.22

확률변수 $X$와 $Y$의 분산이 각각 $\sigma^2{}_X = 2, \sigma^2{}_Y = 4$이고, 공분산이 $\sigma_{XY} = -2$일 때, **확률변수 $Z = 3X - 4Y + 8$의 분산**을 구하라

- $\sigma^2{}_Z = \sigma^2{}_{3X-4Y+8}$
  $= \sigma^2{}_{3X-4Y} = 9\sigma^2{}_X + 16\sigma^2{}_Y - 24\sigma_{XY}$
  $= (9)(2) + (16)(4) - (24)(-2) = 130$

- 예제 4.23

어떤 화학제품의 무더기 속에 섞여 있는 두 종류의 불순물을 $X, Y$라고 할 때, $X$와 $Y$는 서로 독립이고 분산이 각각 $\sigma^2{}_X = 2, \sigma^2{}_Y = 3$이라고 한다. **확률변수 $Z = 3X - 2Y + 5$의 분산**을 구하라

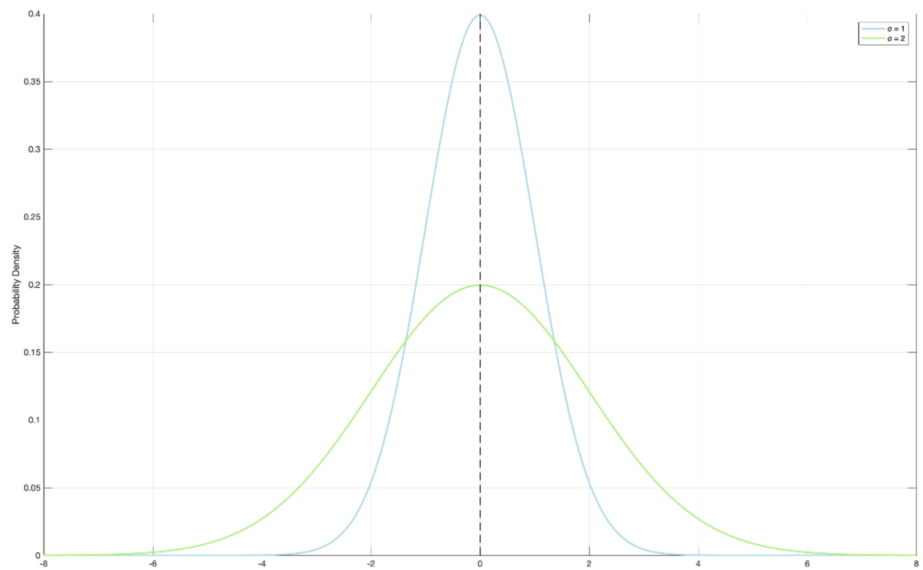
- $\sigma^2{}_Z = \sigma^2{}_{3X-2Y+5}$
  $= \sigma^2{}_{3X-2Y} = 9\sigma^2{}_X + 4\sigma^2{}_Y$
  $= (9)(2) + (4)(3) = 30$

# 목 차

- 보충
- 확률변수의 평균
- 분산과 공분산
- 선형결합된 확률변수의 평균과 분산
- 체비셰프 정리

# 체비셰프 정리(1/5)

- 개요
  - 표준편차가 큰 연속형 분포의 경우 변동성이 커지므로 면적이 더 퍼져 있게 됨
  - 표준편차($\sigma$)의 값이 작으면 대부분의 면적이 평균($\mu$)에 근접함
  - 러시아의 수학자 체비셰프(Chebyshev)는 평균을 중심으로 대칭인 두 값 사이의 면적이 표준편차와 관계되어 있음을 발견함




<그림 4.8 분산이 1, 2인 정규분포>

# 체비셰프 정리(2/5)

- 체비셰프 정리 (1/3)
  - 정의
    - 확률변수 $X$가 평균으로부터 표준편차의 $k$배 범위 내의 값을 취할 확률은 적어도 $1 - \frac{1}{k^2}$
  - 공식
    - $P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$
  - 필요성
    - 데이터의 분포를 정확히 알기 어려운 경우에도 데이터의 흩어짐 정도를 파악할 수 있음

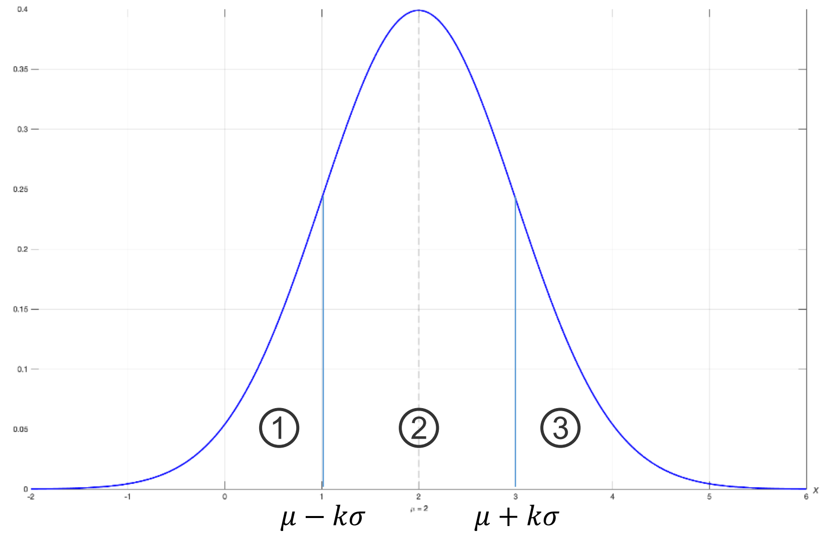
- 체비셰프 정리 (2/3)

**증명**

$$\sigma^2 = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$$

$$= \int_{-\infty}^{\mu-k\sigma} (x-\mu)^2 f(x)\,dx + \int_{\mu-k\sigma}^{\mu+k\sigma} (x-\mu)^2 f(x)\,dx + \int_{\mu+k\sigma}^{\infty} (x-\mu)^2 f(x)\,dx$$

$$\geq \int_{-\infty}^{\mu-k\sigma} (x-\mu)^2 f(x)\,dx + \int_{\mu+k\sigma}^{\infty} (x-\mu)^2 f(x)\,dx$$




<그림 4.2 분산이 1인 연속형 확률분포>

# 체비셰프 정리(4/5)

- ## 체비셰프 정리 (3/3)

**증명**

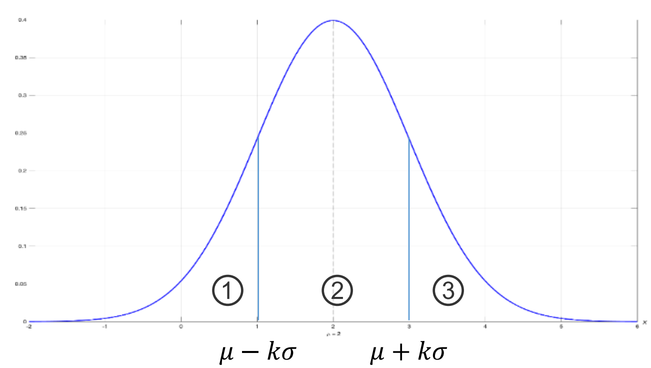
$x \geq \mu + k\sigma$이거나 $x \leq \mu - k\sigma$이면 $(x-\mu)^2 \geq k^2\sigma^2$이 된다.
따라서,

$$\sigma^2 \geq \int_{-\infty}^{\mu-k\sigma} k^2\sigma^2 f(x)\,dx + \int_{\mu+k\sigma}^{\infty} k^2\sigma^2 f(x)\,dx$$

$$\int_{-\infty}^{\mu-k\sigma} f(x)\,dx + \int_{\mu+k\sigma}^{\infty} f(x)\,dx \leq \frac{1}{k^2}$$

이다. 그러므로,

$$P(\mu - k\sigma < X < \mu + k\sigma) = \int_{\mu-k\sigma}^{\mu+k\sigma} f(x)\,dx \geq 1 - \frac{1}{k^2}$$




<그림 4.2 분산이 1인 연속형 확률분포>

# 체비세프 정리(5/5)

- 예제 4.24

**확률변수 $\boldsymbol{X}$의 평균은 $\boldsymbol{\mu = 8}$이고 분산은 $\boldsymbol{\sigma^2 = 9}$**이며, **확률분포는 알려져 있지 않다**고 할 때 다음을 구하라
(a) $P(-4 < X < 20)$
(b) $P(|X - 8| \geq 6)$

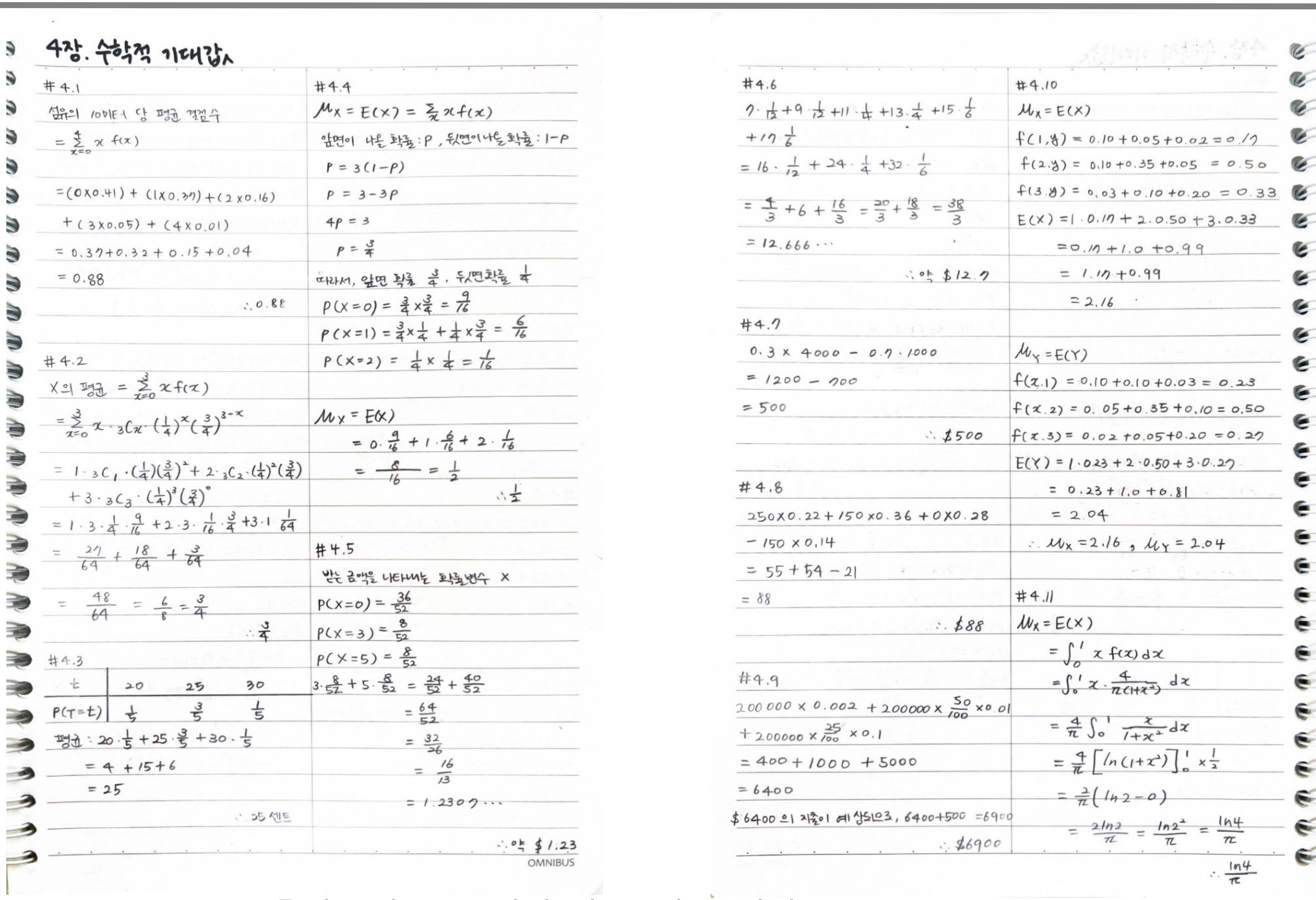
- (a) 계산
  - $P(-4 < X < 20) = P[8 - (4)(3) < X < 8 + (4)(3)] \geq \frac{15}{16}$
- (b) 계산
  - $P(|X - 8| \geq 6) = 1 - P(|X - 8| < 6) = 1 - P(-6 < X - 8 < 6)$

$$= 1 - P[8 - (2)(3) < X < 8 + (2)(3)] \leq \frac{1}{4}$$

# Thanks!

손 우 영 (wooyoung@pel.sejong.ac.kr)

# 부록 #1 – 연습문제 (4.1~4.11)



4장. 수학적 기대값

#4.1

섬유의 10미터 당 평균 결점수

$= \sum_{x=0}^{4} x f(x)$

$= (0\times0.41) + (1\times0.37) + (2\times0.16)$

$+ (3\times0.05) + (4\times0.01)$

$= 0.37 + 0.32 + 0.15 + 0.04$

$= 0.88$

$\therefore 0.88$

#4.2

X의 평균 $= \sum_{x=0}^{3} x f(x)$

$= \sum_{x=0}^{3} x \cdot {}_3C_x \cdot \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{3-x}$

$= 1 \cdot {}_3C_1 \cdot \left(\frac{1}{4}\right)\left(\frac{3}{4}\right)^2 + 2 \cdot {}_3C_2 \cdot \left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)$

$+ 3 \cdot {}_3C_3 \cdot \left(\frac{1}{4}\right)^3 \left(\frac{3}{4}\right)^0$

$= 1 \cdot 3 \cdot \frac{1}{4} \cdot \frac{9}{16} + 2 \cdot 3 \cdot \frac{1}{16} \cdot \frac{3}{4} + 3 \cdot 1 \cdot \frac{1}{64}$

$= \frac{27}{64} + \frac{18}{64} + \frac{3}{64}$

$= \frac{48}{64} = \frac{6}{8} = \frac{3}{4}$

$\therefore \frac{3}{4}$

#4.3

| $t$ | 20 | 25 | 30 |
|---|---|---|---|
| $P(T=t)$ | $\frac{1}{5}$ | $\frac{3}{5}$ | $\frac{1}{5}$ |

평균 : $20 \cdot \frac{1}{5} + 25 \cdot \frac{3}{5} + 30 \cdot \frac{1}{5}$

$= 4 + 15 + 6$

$= 25$

∴ 25 센트

#4.4

$\mu_X = E(X) = \sum_x x f(x)$

앞면이 나올 확률 : $p$ , 뒷면이 나올 확률 : $1-p$

$p = 3(1-p)$

$p = 3 - 3p$

$4p = 3$

$p = \frac{3}{4}$

따라서, 앞면 확률 $\frac{3}{4}$ , 뒷면확률 $\frac{1}{4}$

$P(X=0) = \frac{3}{4} \times \frac{3}{4} = \frac{9}{16}$

$P(X=1) = \frac{3}{4} \times \frac{1}{4} + \frac{1}{4} \times \frac{3}{4} = \frac{6}{16}$

$P(X=2) = \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$

$\mu_X = E(X)$

$= 0 \cdot \frac{9}{16} + 1 \cdot \frac{6}{16} + 2 \cdot \frac{1}{16}$

$= \frac{8}{16} = \frac{1}{2}$

$\therefore \frac{1}{2}$

#4.5

받는 금액을 나타내는 확률변수 X

$P(X=0) = \frac{36}{52}$

$P(X=3) = \frac{8}{52}$

$P(X=5) = \frac{8}{52}$

$3 \cdot \frac{8}{52} + 5 \cdot \frac{8}{52} = \frac{24}{52} + \frac{40}{52}$

$= \frac{64}{52}$

$= \frac{32}{26}$

$= \frac{16}{13}$

$= 1.2307\cdots$

∴ 약 \$1.23

OMNIBUS

#4.6

$7 \cdot \frac{1}{12} + 9 \cdot \frac{1}{12} + 11 \cdot \frac{1}{4} + 13 \cdot \frac{1}{4} + 15 \cdot \frac{1}{6}$

$+ 17 \frac{1}{6}$

$= 16 \cdot \frac{1}{12} + 24 \cdot \frac{1}{4} + 32 \cdot \frac{1}{6}$

$= \frac{4}{3} + 6 + \frac{16}{3} = \frac{20}{3} + \frac{18}{3} = \frac{38}{3}$

$= 12.666\cdots$

∴ 약 \$12.7

#4.7

$0.3 \times 4000 - 0.7 \cdot 1000$

$= 1200 - 700$

$= 500$

∴ \$500

#4.8

$250 \times 0.22 + 150 \times 0.36 + 0 \times 0.28$

$- 150 \times 0.14$

$= 55 + 54 - 21$

$= 88$

∴ \$88

#4.9

$200000 \times 0.002 + 200000 \times \frac{50}{100} \times 0.01$

$+ 200000 \times \frac{25}{100} \times 0.1$

$= 400 + 1000 + 5000$

$= 6400$

\$6400의 지출이 예상되므로, 6400+500 =6900

∴ \$6900

#4.10

$\mu_X = E(X)$

$f(1,y) = 0.10 + 0.05 + 0.02 = 0.17$

$f(2,y) = 0.10 + 0.35 + 0.05 = 0.50$

$f(3,y) = 0.03 + 0.10 + 0.20 = 0.33$

$E(X) = 1 \cdot 0.17 + 2 \cdot 0.50 + 3 \cdot 0.33$

$= 0.17 + 1.0 + 0.99$

$= 1.17 + 0.99$

$= 2.16$

$\mu_Y = E(Y)$

$f(x,1) = 0.10 + 0.10 + 0.03 = 0.23$

$f(x,2) = 0.05 + 0.35 + 0.10 = 0.50$

$f(x,3) = 0.02 + 0.05 + 0.20 = 0.27$

$E(Y) = 1 \cdot 0.23 + 2 \cdot 0.50 + 3 \cdot 0.27$

$= 0.23 + 1.0 + 0.81$

$= 2.04$

$\therefore \mu_X = 2.16 , \mu_Y = 2.04$

#4.11

$\mu_X = E(X)$

$= \int_0^1 x f(x)\,dx$

$= \int_0^1 x \cdot \frac{4}{\pi(1+x^2)}\,dx$

$= \frac{4}{\pi} \int_0^1 \frac{x}{1+x^2}\,dx$

$= \frac{4}{\pi} \left[ \ln(1+x^2) \right]_0^1 \times \frac{1}{2}$

$= \frac{2}{\pi}(\ln 2 - 0)$

$= \frac{2\ln 2}{\pi} = \frac{\ln 2^2}{\pi} = \frac{\ln 4}{\pi}$

$\therefore \frac{\ln 4}{\pi}$

<「이공학도를 위한 확률 및 통계학」 p.131~132>

# 부록 #2 – 연습문제 (4.12~4.23)

#4.12

$\mu_X = E(X)$

$= \int_0^1 x f(x) dx$

$= \int_0^1 x \cdot 2(1-x) dx$

$= 2\int_0^1 x - x^2 dx$

$= 2\left[\frac{1}{2}x^2 - \frac{1}{3}x^3\right]_0^1$

$= 2\left(\frac{1}{2} - \frac{1}{3}\right) = 2 \cdot \frac{1}{6} = \frac{1}{3}$

$5000 \times \frac{1}{3} = 1666.666\cdots$

∴약 \$1666.7

#4.13

$\int_0^1 x \cdot x\, dx + \int_1^2 x(2-x)\, dx$

$= \int_0^1 x^2 dx + \int_1^2 2x - x^2 dx$

$= \left[\frac{1}{3}x^3\right]_0^1 + \left[x^2 - \frac{1}{3}x^3\right]_1^2$

$= \frac{1}{3} + 4 - \frac{8}{3} - 1 + \frac{1}{3}$

$= 3 - \frac{6}{3} = 1$

∴ 100시간

#4.14

$\int_0^1 x \cdot \frac{2(x+2)}{5} dx$

$= \frac{2}{5}\int_0^1 x^2 + 2x\, dx$

$= \frac{2}{5}\left[\frac{1}{3}x^3 + x^2\right]_0^1 = \frac{2}{5}\left(\frac{1}{3}+1\right)$

$= \frac{2}{5} \cdot \frac{4}{3} = \frac{8}{15}$

$\frac{8}{15}$

#4.15

균일분포에서 각점은 동일한 확률을 가지므로,
기대값은 원점에 대한 대칭성 때문에
0이됨

∴0

#4.16

$X_1 = 1,\ X_2 = 0$인 경우

: $\frac{20}{1000} \times \frac{980}{999}$

$X_1 = 0,\ X_2 = 1$인 경우

: $\frac{980}{1000} \times \frac{20}{999}$

$\frac{98\cancel{0}}{\cancel{1000}} \times \frac{\cancel{20}}{999} \times 2$ (5)

$= \frac{196}{4995}$

∴ $\frac{196}{4995}$

#4.17

$\mu_{g(X)} = E[g(X)]$

$= \sum_x g(x) f(x)$

$= \sum_x (2x+1)^2 f(x)$

$= 25 \cdot \left(\frac{1}{6}\right) + 169 \cdot \frac{1}{2} + 361 \cdot \frac{1}{3}$

$= \frac{25 + 507 + 722}{6}$

$= \frac{1254}{6} = 209$

∴209

#4.18

$1 \cdot \frac{27}{64} + 4 \cdot \frac{9}{64} + 9 \cdot \frac{1}{64}$

$= \frac{27 + 36 + 9}{64} = \frac{72}{64} = \frac{9}{8}$

∴ $\frac{9}{8}$

#4.19

X 대당 소요금액 : $1200X - 50X^2$

$E(X) = \sum_{x=0}^{3} (1200x - 50x^2) f(x)$

$= 1150 \times \frac{3}{10} + 2200 \times \frac{2}{5}$

$+ 3150 \times \frac{1}{5}$

$= 345 + 880 + 630$

$= 1855$

∴ \$1855

#4.20

$\mu_{g(X)} = E[g(X)]$

$= \int_0^\infty e^{2x/3} \cdot e^{-x} dx$

$= \int_0^\infty e^{-\frac{x}{3}} dx$

$= \left[-3e^{-\frac{x}{3}}\right]_0^\infty$

$= 0 + 3 = 3$

∴ 3

#4.21

$\int_0^1 x^2 \cdot 2(1-x) dx$

$= 2\int_0^1 x^2 - x^3 dx$

$= 2\left[\frac{1}{3}x^3 - \frac{1}{4}x^4\right]_0^1$

$= 2 \cdot \frac{1}{12} = \frac{1}{6}$

∴약 \$833.33

#4.22

$\int_0^\infty (x+4) \cdot \frac{32}{(x+4)^3} dx$

$= \int_0^\infty \frac{32}{(x+4)^2} dx$

$= -32\left[\frac{1}{x+4}\right]_0^\infty$

$= 32 \cdot \frac{1}{4} = 8$

∴ 8일

#4.23

(a) $E_{g(X,Y)} = \sum_x \sum_y x y^2 f(x,y)$

$= 2 \cdot 1 \cdot f(2,1) + 2 \cdot 9 \cdot f(2,3)$

$+ 2 \cdot 25 \cdot f(2,5) + 4 \cdot 1 \cdot f(4,1)$

$+ 4 \cdot 9 \cdot f(4,3) + 4 \cdot 25 \cdot f(4,5)$

$= 2 \cdot 0.10 + 18 \cdot 0.20 + 50 \cdot 0.10$

$+ 4 \cdot 0.15 + 36 \cdot 0.30 + 100 \cdot 0.15$

$= 0.2 + 3.6 + 5 + 0.6 + 10.8$

$+ 15$

$= 35.2$

∴ 35.2

(b)

① $\mu_X$

$f(2,y) = 0.4$

$f(4,y) = 0.6$

$2 \times 0.4 + 4 \times 0.6 = 0.8 + 2.4 = 3.2$

OMNIBUS

<「이공학도를 위한 확률 및 통계학」 p.132~133>

# 부록 #3 – 연습문제 (4.24~4.28)

② $\mu_Y$

$f(x,1) = 0.25$

$f(x,3) = 0.50$

$f(x,5) = 0.25$

$1\times0.25 + 3\times0.50 + 5\times0.25$

$= 0.25 + 1.5 + 1.25$

$= 3.00$

$\therefore$ (a) 35.2

(b) $\mu_X = 3.20$, $\mu_Y = 3.00$

#4.24

(a)

$E(X^2Y - 2XY)$

$= \sum_x \sum_y (x^2y - 2xy) f(x,y)$

$= -f(1,1) - 2f(1,2)$

$= -\frac{16}{55} - \frac{12}{55} = -\frac{28}{55}$

(b)

① $\mu_X$

$f(0,y) = \frac{14}{55}$

$f(1,y) = \frac{28}{55}$

$f(2,y) = \frac{12}{55}$

$f(3,y) = \frac{1}{55}$

$\mu_X = \frac{28+24+3}{55} = \frac{55}{55}$

$= 1$

② $\mu_Y$

$f(x,0) = \frac{14}{55}$

$f(x,1) = \frac{28}{55}$

$f(x,2) = \frac{12}{55}$

$f(x,3) = \frac{1}{55}$

$\mu_Y = \frac{28+24+3}{55} = 1$

$\therefore \mu_X - \mu_Y = 0$

(생각해보니. f(x.y)가 대칭이니까 당연한것;;;)

$\therefore$ (a) $-\frac{28}{55}$, (b) 0

#4.25

킹카드의 평균장수

: $1\cdot\frac{28}{55} + 2\cdot\frac{12}{55} + 3\cdot\frac{1}{55} = 1$

잭 카드의 평균장수

$1\cdot\frac{28}{55} + 2\cdot\frac{12}{55} + 3\cdot\frac{1}{55} = 1$

$1+1 = 2$

$\therefore$ 2장

#4.26

$\int_0^1 \int_0^1 \sqrt{x^2+y^2} \cdot 4xy \, dx dy$

$\int_0^1 \sqrt{x^2+y^2} \cdot 4xy \, dx$

→ 치환 적분 사용!

$u = x^2+y^2$

$du = 2x\,dx$

$\int_0^1 \sqrt{x^2+y^2} \cdot 4xy \, dx$

$= \int \sqrt{u} \cdot \cancel{4}^{2}\cancel{x}y \cdot \frac{1}{\cancel{2x}} du$

$= \int 2y\sqrt{u}\,du$

$u = x^2+y^2$ 이므로, $x$가 $0 \to 1$ 일때

$u$는 $y^2 \to 1+y^2$

$\int_{y^2}^{1+y^2} 2y\sqrt{u}\,du$

$= 2y \int_{y^2}^{1+y^2} \sqrt{u}\,du$

$= 2y \left[\frac{2}{3}u^{\frac{3}{2}}\right]_{y^2}^{1+y^2}$

$= \frac{4}{3}y\left((1+y^2)^{\frac{3}{2}} - (y^2)^{\frac{3}{2}}\right)$

$= \frac{4}{3}y\left((y^2+1)^{\frac{3}{2}} - y^3\right)$

④ : $\int_0^1 \frac{4}{3}y\left((y^2+1)^{\frac{3}{2}} - (y^2)^{\frac{3}{2}}\right)dy$

$= \frac{4}{3}\int_0^1 y\cdot(y^2+1)^{\frac{3}{2}}dy - \frac{4}{3}\int_0^1 y\cdot y^3\,dy$

$= \frac{4}{3}\left[(y^2+1)^{\frac{5}{2}}\cdot\frac{1}{\cancel{2}}\cdot\frac{\cancel{2}}{5}\right]_0^1 - \frac{4}{3}\left[\frac{1}{5}y^5\right]_0^1$

$= \frac{4}{3}\left(2^{\frac{5}{2}}\cdot\frac{1}{5} - 1\cdot\frac{1}{5}\right) - \frac{4}{3}$

$= \frac{4}{3}\cdot2^{\frac{5}{2}}\cdot\frac{1}{5} - \frac{4}{15} - \frac{20}{15}$

$= \frac{4}{15}\cdot2^{\frac{5}{2}} - \frac{24}{15}$

$= \frac{4}{15}\cdot2^{\frac{5}{2}} - \frac{8}{5}$

$\therefore \frac{4}{15}\cdot2^{\frac{5}{2}} - \frac{8}{5}$

#4.27

exp( ) ⇒ 지수함수(exponential function)

$\int_0^\infty x \cdot \frac{1}{2000} \cdot e^{-\frac{x}{2000}} dx$

$= \frac{1}{2000}\int_0^\infty x \cdot e^{-\frac{x}{2000}} dx$

$= \frac{1}{2000}\left[x \cdot e^{-\frac{x}{2000}} \cdot (-2000)\right]_0^\infty$

$- \frac{1}{2000}\int_0^\infty (-2000) \cdot e^{-\frac{x}{2000}} dx$

$= \int_0^\infty e^{-\frac{x}{2000}} dx$

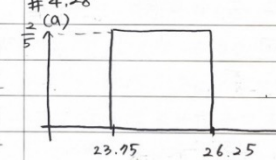
$= \left[-2000 \cdot e^{-\frac{x}{2000}}\right]_0^\infty$

$= 2000$

$\therefore$ 2000시간

#4.28

(a)




(b)

균일분포이므로..

$\frac{23.75+26.25}{2} = 25$

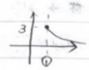
OMNIBUS

<「이공학도를 위한 확률 및 통계학」 p.132~134>

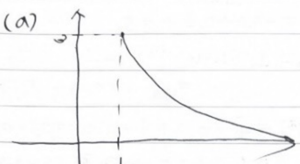
# 부록 #4 – 연습문제 (4.29~4.35)

#4.29

$f(x) = \begin{cases} \frac{3}{x^4} & (x>1) \\ 0 & (\text{다른곳에서}) \end{cases}$




(a)




(b)

$\int_1^\infty x \cdot 3x^{-4}\,dx$

$= \int_1^\infty 3x^{-3}\,dx$

$= \left[-\frac{3}{2}x^{-2}\right]_1^\infty$

$= \frac{3}{2}$

#4.30

$\int_0^\infty y \cdot \frac{1}{4} e^{-\frac{y}{4}}\,dy$

$= \frac{1}{4}\left[y \cdot (-4)e^{-\frac{y}{4}}\right]_0^\infty - \frac{1}{4}\int_0^\infty (-4) \cdot e^{-\frac{y}{4}}\,dy$

$= \int_0^\infty e^{-\frac{y}{4}}\,dy$

$= \left[-4e^{-\frac{y}{4}}\right]_0^\infty$

$= 4$

$\therefore 4$

#4.31

(a) $\int_0^1 y\ 5(1-y)^4\,dy$

$= 5\int_0^1 y \cdot (1-y)^4\,dy$

$= 5\left[-y \cdot \frac{1}{5}(1-y)^5\right]_0^1$

$-5\int_0^1 -\frac{1}{5}(1-y)^5\,dy$

$= \int_0^1 (1-y)^5\,dy$

$= \left[-\frac{1}{6}(1-y)^6\right]_0^1$

$= \frac{1}{6}$

(b)

$\int_{\frac{1}{6}}^1 5(1-y)^4\,dy$

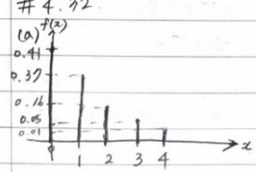
$= \left[-(1-y)^5\right]_{\frac{1}{6}}^1$

$= (1-\frac{1}{6})^5 = (\frac{5}{6})^5$

$\therefore$ (a) $\frac{1}{6}$ (b) $(\frac{5}{6})^5$

#4.32

(a)




(b)

$\mu_X = E(X) = \sum_x x f(x)$

$= 0\times0.41 + 1\times0.37 + 2\times0.16 + 3\times0.05$

$+ 4\times0.01$

$= 0.37 + 0.32 + 0.15 + 0.04$

$= 0.69 + 0.19$

$= 0.88$

(C)

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

$= 0^2\times0.41 + 1^2\times0.37 + 2^2\times0.16$

$+ 3^2\times0.05 + 4^2\times0.01$

$= 0.37 + 0.64 + 0.45 + 0.16$

$= 1.01 + 0.61$

$= 1.62$

#4.33

$\sigma^2 = E[(X-\mu)^2] = \sum_x (x-\mu)^2 f(x)$

$\mu = 0.3\times4000 - 0.7\times1000$

$= 1200 - 700$

$= 500$

⑦ $(4000-500)^2\times0.3$

$+(-1000-500)^2\times0.7$

$= 3500^2\times0.3 + 1500^2\times0.7$

$= 3,675,000 + 1,575,000$

$= 5,250,000$

$\therefore$ \$ 5,250,000

#4.34

$\mu = E(X)$

$= -2\times0.3 + 3\times0.2 + 5\times0.5$

$= -0.6 + 0.6 + 2.5$

$= 2.5$

$\frac{25}{125}$

$E(X^2) = 4\times0.3 + 9\times0.2 + 25\times0.5$

$= 1.2 + 1.8 + 12.5$

$= 15.5$

$\sigma^2 = E(X^2) - 2.5$

$= 15.5 - 2.5$

$= 13$

$\sigma = \sqrt{13}$

$\therefore \sqrt{13}$

#4.35

$\mu = 0.02 + 0.75 + 1.6 + 1.5 + 0.24$

$= 0.77 + 3.1 + 0.24$

$= 3.87 + 0.24$

$= 4.11$

$E(X^2) = 4\times0.01 + 9\times0.25 + 16\times0.4$

$+ 25\times0.3 + 36\times0.04$

$= 0.04 + 2.25 + 6.4 + 7.5$

$+ 1.44$

$= 2.29 + 13.9 + 1.44$

$= 17.63$

$17.63 - 4.11^2 = 17.63 - 16.8921$

$= 0.7379$

$\therefore$ 약 0.74

OMNIBUS

<「이공학도를 위한 확률 및 통계학」 p.134, p.144>

# 부록 #5 – 연습문제 (4.36~4.40)

\#4.36

평균: $0.3+2\times0.2+3\times0.1$

$=0.3+0.4+0.3$

$=1$

분산: $E(X^2)-\mu^2$

$=E(X^2)-1$

$E(X^2)=1\times0.3+4\times0.2+9\times0.1$

$=0.3+0.8+0.9$

$=2$

따라서, 분산은 $2-1=1$

$\therefore$ 평균: 1, 분산: 1

\#4.37

$\mu=\int_0^1 2(1-x)x\,dx$

$=2\int_0^1 x-x^2\,dx$

$=2\left[\frac{1}{2}x^2-\frac{1}{3}x^3\right]_0^1$

$=2\cdot\frac{1}{6}=\boxed{\frac{1}{3}}$

$E(X^2)=\int_0^1 2(1-x)\cdot x^2\,dx$

$=2\int_0^1 x^2-x^3\,dx$

$=2\left[\frac{1}{3}x^3-\frac{1}{4}x^4\right]_0^1$

$=2\cdot\frac{1}{12}$

$=\frac{1}{6}$

$\frac{1}{6}-\left(\frac{1}{3}\right)^2=\frac{1}{6}-\frac{1}{9}=\frac{1}{18}$

$\therefore\frac{1}{18}$

√(실제 분산은 $\frac{1}{18}(5000)^2$)

\#4.38

$\mu=\int_0^1\frac{2(x+2)}{5}\cdot x\,dx$

$=\frac{2}{5}\int_0^1 x^2+2x\,dx$

$=\frac{2}{5}\left[\frac{1}{3}x^3+x^2\right]_0^1$

$=\frac{2}{5}\left(\frac{1}{3}+1\right)=\frac{2}{5}\cdot\frac{4}{3}$

$=\frac{8}{15}$

$E(X^2)=\int_0^1\frac{2(x+3)}{5}\cdot x^2\,dx$

$=\frac{2}{5}\int_0^1 x^3+3x^2\,dx$

$=\frac{2}{5}\left[\frac{1}{4}x^4+x^3\right]_0^1$

$=\frac{2}{5}\left(\frac{1}{4}+1\right)$

$=\frac{2}{5}\cdot\frac{5}{4}$

$=\frac{1}{2}$

⑦ $\frac{1}{2}-\left(\frac{8}{15}\right)^2$

$=\frac{1}{2}-\frac{64}{225}$

$=\frac{225-128}{450}$

$=\frac{97}{450}$

$\therefore\frac{97}{450}$

OMNIBUS

\#4.39

$\mu=\int_0^1 x\cdot x\,dx+\int_1^2(2-x)\cdot x\,dx$

$=\int_0^1 x^2\,dx+\int_1^2 2x-x^2\,dx$

$=\left[\frac{1}{3}x^3\right]_0^1+\left[x^2-\frac{1}{3}x^3\right]_1^2$

$=\frac{1}{3}+4-\frac{8}{3}-1+\frac{1}{3}$

$=3-\frac{6}{3}=1$

$E(X^2)=\int_0^1 x\cdot x^2\,dx+\int_1^2(2-x)\,x^2\,dx$

$=\int_0^1 x^3\,dx+\int_1^2 2x^2-x^3\,dx$

$=\left[\frac{1}{4}x^4\right]_0^1+\left[\frac{2}{3}x^3-\frac{1}{4}x^4\right]_1^2$

$=\frac{1}{4}+\frac{2}{3}\cdot2^3-4-\frac{2}{3}+\frac{1}{4}$

$=\frac{1}{2}+\frac{16}{3}-\frac{2}{3}-4$

$=\frac{1}{2}+\frac{14}{3}-4$

$=\frac{3+28-24}{6}$

$=\frac{7}{6}$

⑦. $\frac{7}{6}-1^2=\frac{1}{6}$

\#4.40

$\sigma^2_{g(X)}=E\{[g(X)-\mu_{g(X)}]^2\}$

$=\int_{-\infty}^{\infty}[g(x)-\mu_{g(X)}]^2f(x)\,dx$

$\mu_{g(X)}=\int_0^1(3x^2+4)\cdot\frac{2(x+2)}{5}\,dx$

$=\frac{2}{5}\int_0^1(3x^2+4)(x+2)\,dx$

$=\frac{2}{5}\int_0^1 3x^3+6x^2+4x+8\,dx$

$=\frac{2}{5}\left[\frac{3}{4}x^4+2x^3+2x^2+8x\right]_0^1$

$=\frac{2}{5}\left(\frac{3}{4}+12\right)$

$=\frac{\cancel{2}}{5}\times\frac{51}{\cancel{4}\,2}$

$=\frac{51}{10}$

따라서, $\sigma^2_{g(X)}$

$=\int_0^1\left\{(3x^2+4)-\frac{51}{10}\right\}^2\cdot\frac{2(x+2)}{5}\,dx$

$=\int_0^1\left(3x^2-\frac{11}{10}\right)^2\cdot\frac{2(x+2)}{5}\,dx$

$=\frac{2}{5}\int_0^1\left(9x^4-\frac{33}{5}x^2+\frac{121}{100}\right)\cdot(x+2)\,dx$

$=\frac{2}{5}\int_0^1 9x^5+18x^4-\frac{33}{5}x^3-\frac{66}{5}x^2+\frac{121}{100}x+\frac{121}{50}\,dx$

$=\frac{2}{5}\left[\frac{9}{6}x^6+\frac{18}{5}x^5-\frac{33}{20}x^4-\frac{22}{5}x^3+\frac{121}{200}x^2+\frac{121}{50}x\right]_0^1$

$=\frac{2}{5}\left(\frac{3}{2}+\frac{18}{5}-\frac{33}{20}-\frac{22}{5}+\frac{121}{200}+\frac{121}{50}\right)$

$=\frac{2}{5}\cdot\frac{300+720-330-880+121+484}{200}$

<「이공학도를 위한 확률 및 통계학」 p.144>

# 부록 #6 – 연습문제 (4.41~4.44)

$= \frac{\cancel{6}}{\cancel{8}} \cdot \frac{\cancel{415}\ 83}{\cancel{200}\ 100}$

$= \frac{83}{100}$

$\therefore 0.83$

#4.41

$\begin{array}{r} 208 \\ 12 \\ \hline 196 \end{array}$

$\mu_{g(x)} = 209$

$\sigma^2_{g(x)} = E\{[g(X) - \mu_{g(x)}]^2\}$

$= \sum_x \{g(x) - 209\}^2 f(x)$

$= \sum_x \{(2x+1)^2 - 209\}^2 f(x)$

$x = -3$

$: (25 - 209)^2 \cdot f(-3)$

$= (-184)^2 \cdot \frac{1}{6}$

$x = 6$

$: (169 - 209)^2 \cdot f(6)$

$= 40^2 \cdot \frac{1}{2}$

$x = 9$

$(19^2 - 209)^2 \cdot f(9)$

$= (361 - 209)^2 \cdot f(9)$

$= 152^2 \cdot \frac{1}{3}$

따라서, $\sigma^2_{g(x)} = 14144$ 이며,

표준편차는 $\sqrt{14144} = 118.929$

$\therefore$ 약 118.9

#4.42

$\mu_{g(x)} = \frac{1}{6}$

$\sigma^2_{g(x)} = E\{[g(X) - \mu_{g(x)}]^2\}$

$= \int_0^1 (g(x) - \mu_{g(x)})^2 \cdot f(x)\,dx$

$= \int_0^1 (x^2 - \frac{1}{6})^2 \cdot f(x)\,dx$

$= \int_0^1 (x^4 - \frac{1}{3}x^2 + \frac{1}{36}) \cdot 2(1-x)\,dx$

$= 2\int_0^1 (x^4 - \frac{1}{3}x^2 + \frac{1}{36})(1-x)\,dx$

$= 2\int_0^1 x^4 - x^5 + \frac{1}{3}x^2 + \frac{1}{3}x^3 + \frac{1}{36} - \frac{1}{36}x\,dx$

$= 2\left[\frac{1}{5}x^5 - \frac{1}{6}x^6 + \frac{1}{9}x^3 + \frac{1}{12}x^4 + \frac{1}{36}x - \frac{1}{72}x^2\right]_0^1$

$= 2(\frac{1}{5} - \frac{1}{6} + \frac{1}{9} + \frac{1}{12} + \frac{1}{36} - \frac{1}{72})$

$= \frac{2}{5} - \frac{1}{3} + \frac{2}{9} + \frac{1}{6} + \frac{1}{18} - \frac{1}{36}$

$= \frac{2}{5} + \frac{-12+8+6+2-1}{36}$

$= \frac{2}{5} + \frac{3}{36}$

$= \frac{2}{5} + \frac{1}{12}$

$= \frac{24+5}{60}$

$= \frac{29}{60}$

$500\cancel{0} \times \frac{29}{6\cancel{0}} \fallingdotseq 2416.67$

$\therefore \$2416.67$

#4.43

$E(Y) = \mu_Y$

$= E(3X-2)$

$= \int_0^\infty (3x-2) \cdot \frac{1}{4}e^{-\frac{x}{4}}\,dx$

$= \frac{1}{4}\int_0^\infty (3x-2) \cdot e^{-\frac{x}{4}}\,dx$

$= \frac{1}{4}\left[(3x-2) \cdot e^{-\frac{x}{4}} \cdot (-4) - \int_0^\infty 3 \cdot e^{-\frac{x}{4}} \cdot (-4)\,dx\right]_0^\infty$

$= \frac{1}{4}\left[(3x-2) \cdot e^{-\frac{x}{4}} \cdot (-4)\right]_0^\infty$

$- \frac{1}{4} \cdot 3 \cdot (-4) \int_0^\infty e^{-\frac{x}{4}}\,dx$

$= \frac{1}{4}(0 - (-2) \cdot (-4))$

$+ 3\left[-4 \cdot e^{-\frac{x}{4}}\right]_0^\infty$

$= -2 + 3 \cdot (0 + 4)$

$= -2 + 12$

$= 10$

$\frac{1}{4}e^{-\frac{x}{4}} \xrightarrow{\int} -e^{-\frac{x}{4}}$

$\sigma^2_Y = E[(Y - \mu_Y)^2]$

$= E[(3X - 2 - 10)^2]$

$= E[(3X - 12)^2]$

$= \int_0^\infty (3x-12)^2 \cdot \frac{1}{4}e^{-\frac{x}{4}}\,dx$

$= \left[(3x-12)^2 \cdot e^{-\frac{1}{4}} \times (-1)\right]_0^\infty$

$- \int_0^\infty 2(3x-12) \cdot 3 \cdot (-e^{-\frac{x}{4}})\,dx$

$= (0 + 144) + 6\int_0^\infty (3x-12) \cdot e^{-\frac{x}{4}}\,dx$

$= 144 + 6\left[(3x-12) \cdot (-e^{-\frac{x}{4}})\right]_0^\infty$

$- 6\int_0^\infty 3\,(-e^{-\frac{x}{4}})\,dx$

$= 144 + 6(0 - 12)$

$+ 18\int_0^\infty e^{-\frac{x}{4}}\,dx$

$= 144 - 72 + 18\left[-4 \cdot e^{-\frac{x}{4}}\right]_0^\infty$

$= 72 + 18(0 + 4)$

$= 72 + 72$

$= 144$

$\therefore \mu_Y = 10,\ \sigma_Y^2 = 144$

#4.44

$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)]$

$= \sum_x \sum_y (x - \mu_X)(y - \mu_Y) f(x, y)$

$= E(XY) - \mu_X \mu_Y$

$\mu_X = \sum_{x=0}^{3} x \cdot g(x)$

$= 0 \cdot \frac{5}{70} + 1 \cdot \frac{30}{70} + 2 \cdot \frac{30}{70} + 3 \cdot \frac{5}{70}$

$= \frac{30+60+15}{70} = \frac{105}{70} = \frac{21}{14} = \frac{3}{2}$

$\mu_Y = \sum_{y=0}^{2} y \cdot h(y)$

$= 0 \cdot \frac{15}{70} + 1 \cdot \frac{40}{70} + 2 \cdot \frac{15}{70}$

$= \frac{70}{70} = 1$

$E(XY) = \sum_{x=0}^{3} \sum_{y=0}^{2} xy\, f(x, y)$

$= 1 \cdot 1 \cdot \frac{18}{70} + 1 \cdot 2 \cdot \frac{9}{70}$

$+ 2 \cdot 1 \cdot \frac{18}{70} + 2 \cdot 2 \cdot \frac{3}{70}$

$+ 3 \cdot 1 \cdot \frac{2}{70}$

OMNIBUS

<「이공학도를 위한 확률 및 통계학」 p.144>

$$= \frac{18+18+36+12+6}{70}$$

$$= \frac{9}{7}$$

따라서, $\sigma_{XY} = \frac{9}{7} - \frac{3}{2} \cdot 1$

$$= \frac{18-21}{14}$$

$$= -\frac{3}{14}$$

$\therefore -\frac{3}{14}$

#4.45

$\sigma_{XY} = E(XY) - \mu_X \mu_Y$

$\mu_X = 1 \cdot 0.10 + 2 \cdot 0.35 + 3 \cdot 0.55$

$= 0.10 + 0.7 + 1.65$

$= 2.45$

$\mu_Y = 1 \cdot 0.2 + 3 \cdot 0.5 + 5 \cdot 0.3$

$= 0.2 + 1.5 + 1.5$

$= 3.2$

$E(XY) = 0.05 + 0.10 + 0.30$

$+ 0.15 + 0.60 + 3.15$

$+ 2 + 1.5$

$= 0.45 + 3.90 + 3.5$

$= 7.85$

따라서, $\sigma_{XY} = 7.85 - 2.45 \times 3.2$

$= 7.85 - 7.84$

$= 0.01$

$\therefore 0.01$

#4.47

$\sigma_{XY} = E[(X-\mu_X)(Y-\mu_Y)]$

$= E(XY) - \mu_X \mu_Y$

$\mu_X = \int_0^1 \int_0^1 x f(x,y)\,dx\,dy$

$= \int_0^1 x \int_0^1 f(x,y)\,dy\,dx$

$\int_0^1 f(x,y)\,dy$

$= \int_0^1 \frac{2}{3}(x+2y)\,dy$

$= \frac{2}{3}\int_0^1 x+2y\,dy$

$= \frac{2}{3}\left[xy + y^2\right]_0^1$

$= \frac{2}{3}(x+1-0)$

$= \frac{2}{3}(x+1)$

$\mu_X = \int_0^1 x \cdot \frac{2}{3}(x+1)\,dx$

$= \frac{2}{3}\int_0^1 x^2+x\,dx$

$= \frac{2}{3}\left[\frac{1}{3}x^3 + \frac{1}{2}x^2\right]_0^1$

$= \frac{2}{3}\left(\frac{1}{3}+\frac{1}{2}\right)$ $\qquad \frac{\cancel{2}}{3} \cdot \frac{5}{\cancel{6}_3}$ $\quad \frac{2+3}{6}$

$= \frac{2}{3} \cdot \frac{5}{6}$

$= \frac{5}{9}$

$\int_0^1 f(x,y)\,dx$

$= \int_0^1 \frac{2}{3}(x+2y)\,dx$

$- \left[\frac{2}{3}\left(\frac{1}{2}x^2 + 2xy\right)\right]_0^1$

$= \frac{2}{3}\left(\frac{1}{2}+2y\right) = \frac{1}{3} + \frac{4}{3}y$

OMNIBUS

$\mu_Y = \int_0^1 y \cdot \left(\frac{1+4y}{3}\right)dy$

$= \int_0^1 \frac{1}{3}y + \frac{4}{3}y^2\,dy$

$= \left[\frac{1}{6}y^2 + \frac{4}{9}y^3\right]_0^1$

$= \frac{3+8}{18} = \frac{11}{18}$

$E(XY) = \int_0^1\int_0^1 xy \cdot \frac{2}{3}(x+2y)\,dx\,dy$

$= \frac{2}{3}\int_0^1\int_0^1 x^2y + 2xy^2\,dx\,dy$

$= \frac{2}{3}\int_0^1 \left[\frac{1}{3}x^3y + x^2y^2\right]_0^1 dy$

$= \frac{2}{3}\int_0^1 \frac{1}{3}y + y^2\,dy$

$= \frac{2}{3}\left[\frac{1}{6}y^2 + \frac{1}{3}y^3\right]_0^1$

$= \frac{2}{3} \cdot \left(\frac{1+2}{6}\right)$

$= \frac{\cancel{2}}{\cancel{3}} \cdot \frac{\cancel{3}}{\cancel{6}_3} = \frac{1}{3}$

$E(XY) - \mu_X \mu_Y$

$= \frac{1}{3} - \frac{5}{9} \times \frac{11}{18}$

$= \frac{54-55}{9 \times 18} = -\frac{1}{162}$

$\therefore -\frac{1}{162}$

#4.48

상관계수 $\rho_{XY}$는 다음과 같이 정의된다.

$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$

$Y = a + bX$ 로부터 Y의 표준편차와 X와 Y의 공분산을 구해보자.

① 공분산

$Cov(X,Y) = Cov(X, a+bX)$

공분산의 선형 특성을 사용하면..

$Cov(X, a+bX)$

$= Cov(X,a) + Cov(X,bX)$

상수 a에 대한 X의 공분산은 0이다.

→ a는 상수이므로 X의 변동과 관련이 없으므로, $Cov(X,a)=0$

$Cov(X,bX) = b \cdot Cov(X,X)$

$Cov(X,X)$는 X의 분산 $Var(X)$와 같으므로, $Cov(X,bX) = b \cdot Var(X)$

$= b \cdot \sigma_X^2$

Y의 표준편차 $\sigma_Y$는 Y의 변동성이 X의 변동성에 b배 만큼 영향을 받는 것을 고려할 때 $b\sigma_X$ 이다.

이를 통해 상관계수를 계산하면,

$\rho_{XY} = \frac{b\sigma_X^2}{\sigma_X |b\sigma_X|} = \frac{b}{|b|}$

따라서 $b<0$ 이면 $\rho_{XY} = -1$ 이고, $b>0$ 이면 $\rho_{XY} = 1$ 이다.

<「이공학도를 위한 확률 및 통계학」 p.144~145>

# 부록 #8 – 연습문제 (4.49~4.52)

#4.49

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

#4.32 에서 결점수의 기대값은 0.88, $E(X^2) = 1.62$ 였으므로,

$\sigma^2 = 1.62 - 0.88^2$

$= 1.62 - 0.7744$

$= 0.8456$

$\sigma = \sqrt{0.8456} = 0.91956\cdots$

$\therefore \sigma^2 = 0.8456,\ \sigma = 0.9196$

#4.50

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

$\mu = \int_0^1 x \cdot 2(1-x)\,dx$

$= 2\int_0^1 x - x^2\,dx$

$= 2\left[\frac{1}{2}x^2 - \frac{1}{3}x^3\right]_0^1$

$= 2 \cdot \frac{1}{6} = \frac{1}{3}$

$E(X^2) = \int_0^1 x^2 \cdot 2(1-x)\,dx$

$= 2\int_0^1 x^2 - x^3\,dx$

$= 2\left[\frac{1}{3}x^3 - \frac{1}{4}x^4\right]_0^1$

$= 2 \cdot \frac{1}{12}$

$= \frac{1}{6}$

$\sigma^2 = \frac{1}{6} - \frac{1}{9} = \frac{3-2}{18} = \frac{1}{18}$

$\sigma = \sqrt{\frac{1}{18}} = 0.2357\cdots$

$\therefore \sigma^2 = \frac{1}{18},\ \sigma = 0.236$

#4.51

$\sigma_{XY} = E(XY) - \mu_X \mu_Y$

$\mu_X = \sum_{x=0}^{3} x \cdot g(x)$

$= 0 \cdot \frac{5}{70} + 1 \cdot \frac{30}{70} + 2 \cdot \frac{30}{70} + 3 \cdot \frac{5}{70}$

$= \frac{30+60+15}{70} = \frac{105}{70} = \frac{15}{10} = \frac{3}{2}$

$\mu_Y = \sum_{y=0}^{2} y \cdot h(y)$

$= 0 \cdot \frac{15}{70} + 1 \cdot \frac{40}{70} + 2 \cdot \frac{15}{70}$

$= \frac{40+30}{70} = 1$

$E(XY) = \sum_{x=0}^{3}\sum_{y=0}^{2} xy f(x,y)$

$= \cancel{0 \cdot 0 \cdot f(0,0)} + \cancel{0 \cdot 1 \cdot f(0,1)}$

$+ \cancel{0 \cdot 2 \cdot f(0,2)} + \cancel{1 \cdot 0 \cdot f(1,0)}$

$+ 1 \cdot 1 \cdot f(1,1) + 1 \cdot 2 \cdot f(1,2)$

$+ \cancel{2 \cdot 0 \cdot f(2,0)} + 2 \cdot 1 \cdot f(2,1)$

$+ 2 \cdot 2 \cdot f(2,2) + \cancel{3 \cdot 0 \cdot f(3,0)}$

$+ 3 \cdot 1 \cdot f(3,1) + 3 \cdot 2 \cdot f(3,2)$

$= f(1,1) + 2 \cdot f(1,2)$

$+ 2f(2,1) + 4f(2,2)$

$+ 3f(3,1) + 6f(3,2)$

$= \frac{18}{70} + 2 \cdot \frac{9}{70} + 2 \cdot \frac{18}{70}$

$+ 4 \cdot \frac{3}{70} + 3 \cdot \frac{2}{70} + 6 \cdot 0$

$= \frac{18+18+36+12+6}{70}$

$= \frac{36+36+18}{70} = \frac{90}{70} = \frac{9}{7}$

$\sigma_{XY} = \frac{9}{7} - \frac{3}{2} \cdot 1$

$= \frac{18-21}{14} = -\frac{3}{14}$

OMNIBUS

$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$

$E(X^2) = 0^2 \cdot \frac{5}{70} + 1^2 \cdot \frac{30}{70} + 2^2 \cdot \frac{30}{70}$

$+ 3^2 \cdot \frac{5}{70}$

$= \frac{30+120+45}{70}$

$= \frac{195}{70} = \frac{39}{14}$

$\sigma_X^2 = \frac{39}{14} - \left(\frac{3}{2}\right)^2$

$= \frac{39}{14} - \frac{9}{4} = \frac{156-126}{56}$

$= \frac{30}{56} = \frac{15}{28}$

$E(Y^2) = 0^2 \cdot \frac{15}{70} + 1^2 \cdot \frac{40}{70} + 2^2 \cdot \frac{15}{70}$

$= \frac{40+60}{70}$

$= \frac{10}{7}$

$\sigma_Y^2 = \frac{10}{7} - (1)^2 = \frac{3}{7}$

$\sigma_X = \sqrt{\frac{15}{28}},\ \sigma_Y = \sqrt{\frac{3}{7}}$

따라서,

$\rho_{XY} = \frac{-\frac{3}{14}}{\sqrt{\frac{15}{28} \cdot \frac{3}{7}}}$

$= -\frac{3}{14\sqrt{\frac{3^2 \cdot 5}{2^2 \cdot 7^2}}} = -\frac{\cancel{3}}{\cancel{14} \cdot \frac{\cancel{3}\sqrt{5}}{\cancel{14}}}$

$= -\frac{1}{\sqrt{5}}$

$\therefore -\frac{1}{\sqrt{5}}$

#4.52

$\sigma_{XY} = E(XY) - \mu_X \mu_Y$

$\mu_X = \int_0^1 \int_x^1 x \cdot 2\,dy\,dx$

$\int_x^1 2x\,dy = [2xy]_x^1$

$= 2x - 2x^2$

$\int_0^1 2x - 2x^2\,dx$

$= \left[x^2 - \frac{2}{3}x^3\right]_0^1$

$= 1 - \frac{2}{3} = \frac{1}{3}$

$\mu_Y = \int_0^1 \int_0^y 2y\,dx\,dy$

$\int_0^y 2y\,dx = [2xy]_0^y$

$= 2y^2$

$\int_0^1 2y^2\,dy = \left[\frac{2}{3}y^3\right]_0^1 = \frac{2}{3}$

$E(XY) = \int_0^1\int_0^y 2xy\,dx\,dy$

$\int_0^y 2xy\,dx = [x^2 y]_0^y$

$= y^3$

$E(XY) = \int_0^1 y^3\,dy$

$= \left[\frac{1}{4}y^4\right]_0^1$

$= \frac{1}{4}$

$\sigma_{XY} = \frac{1}{4} - \frac{1}{3} \times \frac{2}{3}$

$= \frac{1}{4} - \frac{2}{9}$

$= \frac{9-8}{36} = \frac{1}{36}$

<「이공학도를 위한 확률 및 통계학」 p.145>

# 부록 #9 – 연습문제 (4.53~4.57)

$E(X^2) = \int_0^1 \int_0^y 2x^2\,dx\,dy$

$\int_0^y 2x^2\,dx = \left[\frac{2}{3}x^3\right]_0^y = \frac{2}{3}y^3$

$E(X^2) = \int_0^1 \frac{2}{3}y^3\,dy$

$= \left[\frac{2}{12}y^4\right]_0^1 = \frac{1}{6}$

$\sigma^2_X = \frac{1}{6} - \left(\frac{1}{3}\right)^2$

$= \frac{1}{6} - \frac{1}{9} = \frac{1}{18}$

$\sigma_X = \sqrt{\frac{1}{18}}$

$E(Y^2) = \int_0^1 \int_0^y 2y^2\,dx\,dy$

$\int_0^y 2y^2\,dx = \left[2xy^2\right]_0^y = 2y^3$

$\int_0^1 2y^3\,dy = \left[\frac{1}{2}y^4\right]_0^1 = \frac{1}{2}$

$\sigma^2_Y = \frac{1}{2} - \left(\frac{2}{3}\right)^2$

$= \frac{1}{2} - \frac{4}{9}$

$= \frac{9-8}{18} = \frac{1}{18}$

$\sigma_Y = \sqrt{\frac{1}{18}}$

따라서,

$\rho_{XY} = \frac{\frac{1}{36}}{\sqrt{\frac{1}{18}\frac{1}{18}}} = \frac{\frac{1}{36}}{\frac{1}{18}}$

$= \frac{1}{2}$

$\therefore \frac{1}{2}$

#4.53

$E(Z) = E(3X-2)$

$= 3E(X) - E(2)$

#4.35에서 $E(X) = 4.11$ 이므로,

$E(Z) = 3 \cdot 4.11 - 2$

$= 12.33 - 2$

$= 10.33$

$\sigma^2_Z = \sigma^2_{3X-2} = \sigma^2_{3X} = 9\sigma^2_X$

#4.35에서 $\sigma^2_X = 0.74$ 였으므로

$\sigma^2_Z = 9 \cdot 0.74$

$= 6.66$

$\therefore E(Z) = 10.33,\ \sigma^2_Z = 6.66$

#4.54

$E(Z) = E(5X+3)$

$= 5E(X) + E(3)$

#4.36에서 $E(X) = 1,\ \sigma^2_X = 1$ 이므로

$E(Z) = 5 \cdot 1 + 3$

$= 8$

$\sigma^2_Z = \sigma^2_{5X+3} = \sigma^2_{5X}$

$= 25\sigma^2_X$

$= 25$

$\therefore E(Z) = 8,\ \sigma^2_Z = 25$

#4.55

이익: $-1.2 \times 5 + 1.65 \times x + 0.9(5-x)$

$= -6 + 1.65x + 4.5 - 0.9x$

$= -1.5 + 0.75x$

$g(X) = 0.75X - 1.5$

OMNIBUS

$E[g(X)] = E(0.75X - 1.5)$

$= 0.75E(X) - 1.5$

$E(X) = \frac{2+4+9+16+15}{15}$

$= \frac{6+25+15}{15}$

$= \frac{46}{15}$

$E[g(X)] = \frac{\cancel{75}^{5}}{100} \times \frac{46}{\cancel{15}} - 1.5$

$= 2.3 - 1.5$

$= 0.8$

$\therefore \$0.8$

#4.56

$E(Y) = E(3X-2)$

$= 3E(X) - 2$

$\sigma^2_Y = \sigma^2_{3X-2}$

$= \sigma^2_{3X} = 9\sigma^2_X$

$E(X) = \int_0^\infty x \cdot \frac{1}{4}e^{-\frac{x}{4}}\,dx$

$= \left[-x \cdot e^{-\frac{x}{4}}\right]_0^\infty + \int_0^\infty e^{-\frac{x}{4}}\,dx$

$= \int_0^\infty e^{-\frac{x}{4}}\,dx$

$= \left[-4 \cdot e^{-\frac{x}{4}}\right]_0^\infty$

$= 0 + 4$

$= 4$

$\sigma^2_X = E(X^2) - \mu_X^2$

$E(X^2) = \int_0^\infty \frac{1}{4}x^2 e^{-\frac{x}{4}}\,dx$

$= \left[-x^2 \cdot e^{-\frac{x}{4}}\right]_0^\infty + \int_0^\infty 2x \cdot e^{-\frac{x}{4}}\,dx$

$= \int_0^\infty 2x e^{-\frac{x}{4}}\,dx$

앞서, $\int_0^\infty \frac{1}{4} x e^{-\frac{x}{4}}\,dx = 4$ 였으므로,

이의 8배인 $\int_0^\infty 2x\, e^{-\frac{x}{4}}\,dx = 4 \cdot 8$

$= 32$

$\sigma^2_X = 32 - 4^2 = 16$

따라서, $E(Y) = 3 \cdot 4 - 2 = 10$

$\sigma^2_Y = 9\sigma^2_X = 9 \cdot 16 = 144$

$\therefore \mu_Y = 10,\ \sigma^2_Y = 144$

#4.57

$E(X) = -3 \cdot \frac{1}{6} + 6 \cdot \frac{1}{2} + 9 \cdot \frac{1}{3}$

$= -\frac{1}{2} + 3 + 3$

$= -\frac{1}{2} + 6$

$= 5\frac{1}{2} = \frac{11}{2}$

$E(X^2) = 9 \cdot \frac{1}{6} + 36 \cdot \frac{1}{2} + 81 \cdot \frac{1}{3}$

$= \frac{3}{2} + 18 + 27$

$= 46\frac{1}{2} = \frac{93}{2}$

$E[(2X+1)^2]$

$= E(4X^2 + 4X + 1)$

$= 4E(X^2) + 4E(X) + E(1)$

$= 4 \cdot \frac{93}{2} + 4 \cdot \frac{11}{2} + 1$

$= 186 + 22 + 1$

$= 209$

$\therefore 209$

<「이공학도를 위한 확률 및 통계학」 p.155>

# 부록 #10 – 연습문제 (4.58~4.65)

#4.58

$E(Y) = E(60X^2 + 39X)$

$= 60E(X^2) + 39E(X)$

$E(X^2) = \int_0^1 x^3\,dx + \int_1^2 x^2(2-x)\,dx$

$= \left[\frac{1}{4}x^4\right]_0^1 + \left[\frac{2}{3}x^3 - \frac{1}{4}x^4\right]_1^2$

$= \frac{1}{4} + \frac{2}{3}\cdot 8 - \frac{1}{4}\cdot 16 - \frac{2}{3} + \frac{1}{4}$

$= \frac{1}{4} + \frac{16}{3} - 4 - \frac{2}{3} + \frac{1}{4}$

$= \frac{1}{2} + \frac{14}{3} - 4$

$= \frac{3+28-24}{6}$

$= \frac{7}{6}$

$E(X) = \int_0^1 x^2\,dx + \int_1^2 2x - x^2\,dx$

$= \left[\frac{1}{3}x^3\right]_0^1 + \left[x^2 - \frac{1}{3}x^3\right]_1^2$

$= \frac{1}{3} + 4 - \frac{8}{3} - 1 + \frac{1}{3}$

$= 3 - \frac{6}{3}$

$= 1$

$\therefore E(Y) = 60\cdot\frac{7}{6} + 39\cdot 1$

$= 70 + 39$

$= 109$

$\therefore 109$

#4.59

$E(X^2 - 2X + 1)$

$= E(X^2) - 2E(X) + E(1)$

$= E(X^2) - 2E(X) + 1$

$= 10$

$\Rightarrow E(X^2) - 2E(X) = 9$

$E(X^2 - 4X + 4)$

$= E(X^2) - 4E(X) + E(4)$

$= E(X^2) - 4E(X) + 4$

$= 6$

$\Rightarrow E(X^2) - 4E(X) = 2$

$$\begin{array}{r|l} & E(X^2) - 2E(X) = 9 \\ - & E(X^2) - 4E(X) = 2 \\ \hline & \quad 2E(X) = 7 \end{array}$$

$E(X) = \frac{7}{2}$, $E(X^2) = 16$

$\sigma^2_X = 16 - \left(\frac{7}{2}\right)^2$

$= 16 - \frac{49}{4}$

$= \frac{64-49}{4}$

$= \frac{15}{4}$

$\therefore \mu = \frac{7}{2}$, $\sigma^2_X = \frac{15}{4}$

#4.60

(a) $E(2X - 3Y)$

$= 2E(X) - 3E(Y)$

$E(X) = 2\cdot 0.4 + 4\cdot 0.6$

$= 0.8 + 2.4 = 3.2$

OMNIBUS

$E(Y) = 1\cdot 0.25 + 3\cdot 0.5 + 5\ 0.25$

$= 0.25 + 1.5 + 1.25$

$= 3$

따라서, $E(2X - 3Y) = 2\cdot 3.2 - 3\cdot 3$

$= 6.4 - 9$

$= -2.6$

(b) X.Y가 서로 독립이므로,

$E(XY) = E(X)E(Y)$

$= 3.2 \times 3$

$= 9.6$

$\therefore$ (a) $-2.6$, (b) $9.6$

#4.62

$\sigma^2_Z = \sigma^2_{-2X+4Y-3}$

$= \sigma^2_{-2X+4Y}$

$= 4\sigma^2_X + 16\sigma^2_Y$

$= 4\cdot 5 + 16\cdot 3$

$= 20 + 48$

$= 68$

$\therefore 68$

#4.63

$\sigma^2_Z = \sigma^2_{-2X+4Y-3}$

$= \sigma^2_{-2X+4Y}$

$= 4\sigma^2_X + 16\sigma^2_Y + 2\cdot(-2)\cdot 4\sigma_{XY}$

$= 4\cdot 5 + 16\ 3 - 16\cdot 1$

$= 20 + 48 - 16$

$= 52$

$\therefore 52$

#4.64

$E(Z) = E(XY)$

$= E(X)E(Y)$

$E(X) = \int_2^\infty x\cdot\frac{8}{x^3}\,dx$

$= \int_2^\infty \frac{8}{x^2}\,dx$

$= \left[-\frac{8}{x}\right]_2^\infty$

$= 4$

$E(Y) = \int_0^1 y\cdot 2y\,dy$

$= \int_0^1 2y^2\,dy$

$= \left[\frac{2}{3}y^3\right]_0^1$

$= \frac{2}{3}$

따라서, $E(Z) = 4\cdot\frac{2}{3}$

$= \frac{8}{3}$

$\therefore \frac{8}{3}$

#4.65

$E(X) = E(Y) = \frac{1+2+3+4+5+6}{6}$

$= \frac{21}{6} = \frac{7}{2}$

(a) $E(X+Y) = E(X) + E(Y)$

$= \frac{7}{2} + \frac{7}{2}$

$= 7$

(b) $E(X-Y) = E(X) - E(Y)$

$= \frac{7}{2} - \frac{7}{2}$

$= 0$

<「이공학도를 위한 확률 및 통계학」 p.156>

# 부록 #11 – 연습문제 (4.66~4.69)

(c) 붉은색 주사위를 던지는 것과 초록색 주사위를 던지는 것은 독립이므로,

$E(XY) = E(X)E(Y)$

$= \frac{7}{2}\cdot\frac{7}{2}$

$= \frac{49}{4}$

$\therefore$ (a) $7$, (b) $0$, (c) $\frac{49}{4}$

#4.66

(a) $\sigma^2_{2X-Y} = 4\cdot\sigma^2_X + 1\cdot\sigma^2_Y$

$E(X^2) = \frac{1+4+9+16+25+36}{6}$

$= \frac{5+25+61}{6}$

$= \frac{91}{6}$

$\sigma^2_X = \frac{91}{6} - \left(\frac{7}{2}\right)^2$

$= \frac{91}{6} - \frac{49}{4}$

$= \frac{182-147}{12}$

$= \frac{35}{12} = \sigma^2_Y$

따라서, $\sigma^2_{2X-Y} = 4\cdot\frac{35}{12} + 1\cdot\frac{35}{12}$

$= \frac{175}{12}$

(b) $\sigma^2_{X+3Y-5} = \sigma^2_X + 9\sigma^2_Y$

$= \cancel{10}^{5}\cdot\frac{35}{\cancel{12}_{6}}$

$= \frac{175}{6}$

$\therefore$ (a) $\frac{175}{12}$, (b) $\frac{175}{6}$

#4.67

$E[g(X,Y)]$

$= E\left(\frac{X}{Y^3} + X^2Y\right)$

$= E\left(\frac{X}{Y^3}\right) + E(X^2Y)$

$= \int_1^2\int_0^1 \frac{2}{7}(x+2y)\cdot\frac{x}{y^3}\,dx\,dy$

$+ \int_1^2\int_0^1 \frac{2}{7}(x+2y)\cdot x^2y\,dx\,dy$

$= \frac{2}{7}\int_1^2\int_0^1 \frac{x^2}{y^3} + \frac{2x}{y^2}\,dx\,dy$

$+ \frac{2}{7}\int_1^2\int_0^1 x^3y + 2x^2y^2\,dx\,dy$

$\int_0^1 \frac{x^2}{y^3} + \frac{2x}{y^2}\,dx$

$= \left[\frac{x^3}{3y^3} + \frac{x^2}{y^2}\right]_0^1$

$= \frac{1}{3y^3} + \frac{1}{y^2}$

$\int_0^1 x^3y + 2x^2y^2\,dx$

$= \left[\frac{1}{4}x^4y + \frac{2}{3}x^3y^2\right]_0^1$

$= \frac{1}{4}y + \frac{2}{3}y^2$

$\frac{2}{7}\int_1^2 \frac{1}{3y^3} + \frac{1}{y^2}\,dy + \frac{2}{7}\int_1^2 \frac{1}{4}y + \frac{2}{3}y^2\,dy$

$= \frac{2}{7}\left[-\frac{1}{6y^2} - \frac{1}{y}\right]_1^2 + \frac{2}{7}\left[\frac{1}{8}y^2 + \frac{2}{9}y^3\right]_1^2$

$= \frac{2}{7}\left(-\frac{1}{24} - \frac{1}{2} + \frac{1}{6} + 1\right) + \frac{2}{7}\left(\frac{4}{8} + \frac{16}{9} - \frac{1}{8} - \frac{2}{9}\right)$

$= \frac{2}{7}\left(\frac{3}{24} + \frac{1}{2}\right) + \frac{2}{7}\left(\frac{3}{8} + \frac{14}{9}\right)$

$= \frac{2}{7}\cdot\frac{15}{24} + \frac{2}{7}\cdot\frac{139}{72}$

$= \frac{\cancel{2}}{7}\cdot\frac{\cancel{184}\,\cancel{92}\,46}{\cancel{72}\,\cancel{36}\,\cancel{18}\,9}$

$= \frac{46}{63}$

$\therefore \frac{46}{63}$

OMNIBUS

#4.68

$E(I) = 15$, $\sigma^2_I = 0.03$, $R = 50$

1) $E(P) = E(I^2R)$

$= 50E(I^2)$

$\sigma^2_I = E(I^2) - \{E(I)\}^2$

$E(I^2) = \sigma^2_I + \{E(I)\}^2$

$= 0.03 + 225$

$= 225.03$

$\rightarrow E(P) = 50 \times 225.03$

$= 11251.5$

#4.69

(a)

$E(X) = \sum_{x=0}^{\infty}\sum_{y=0}^{\infty} x\cdot\frac{1}{4^{(x+y)}}\cdot\frac{9}{16}$

$= \sum_{x=0}^{\infty}\frac{9}{16}x\sum_{y=0}^{\infty}\frac{1}{4^{x+y}}$

$\sum_{y=0}^{\infty}\frac{1}{4^{x+y}} = \sum_{y=0}^{\infty}\frac{1}{4^x}\cdot\frac{1}{4^y}$

$= \frac{1}{4^x}\sum_{y=0}^{\infty}\left(\frac{1}{4}\right)^y$

$= \frac{1}{4^x}\cdot\frac{1}{1-\frac{1}{4}}$

$= \frac{1}{4^x}\cdot\frac{4}{3}$

$\Rightarrow \sum_{x=0}^{\infty}\frac{9}{16}x\cdot\frac{1}{4^x}\cdot\frac{4}{3}$

$= \frac{3}{4}\sum_{x=0}^{\infty} x\cdot\frac{1}{4^x}$

$= \frac{3}{4}\cdot\frac{4}{9} = \frac{1}{3}$

$E(Y) = \sum_{x=0}^{\infty}\sum_{y=0}^{\infty} y\cdot\frac{1}{4^{x+y}}\cdot\frac{9}{16}$

$E(X)$와 같은 꼴이므로,

$E(X) = E(Y) = \frac{1}{3}$

$E(X^2) = \sum_{x=0}^{\infty}\sum_{y=0}^{\infty} x^2\cdot\frac{1}{4^{(x+y)}}\cdot\frac{9}{16}$

$= \sum_{x=0}^{\infty}\frac{9}{16}x^2\cdot\frac{1}{4^x}\sum_{y=0}^{\infty}\frac{1}{4^y}$

$= \sum_{x=0}^{\infty}\frac{9}{16}x^2\cdot\frac{1}{4^x}\cdot\frac{1}{1-\frac{1}{4}}$

$= \sum_{x=0}^{\infty}\frac{3}{4}\cdot x^2\cdot\frac{1}{4^x}$

$= \frac{\cancel{3}}{\cancel{4}}\cdot\frac{\cancel{20}\,5}{\cancel{27}\,9}$

$= \frac{5}{9}$

따라서, $\sigma^2_X = \frac{5}{9} - \left(\frac{1}{3}\right)^2$

$= \frac{4}{9}$

$\sigma^2_Y$도 $\sigma^2_X$와 동일한 꼴이기 때문에,

$Var(X) = Var(Y) = \frac{4}{9}$

(b)

$E(Z) = E(X+Y)$

$= E(X) + E(Y)$

$= \frac{1}{3} + \frac{1}{3}$

$= \frac{2}{3}$

$Var(Z) = Var(X+Y)$

$= \sigma^2_{X+Y}$

$= \sigma^2_X + \sigma^2_Y$

$= \frac{4}{9} + \frac{4}{9}$

$= \frac{8}{9}$

$\therefore E(X) = E(Y) = \frac{1}{3}$, $Var(X) = Var(Y) = \frac{4}{9}$

$E(Z) = \frac{2}{3}$, $Var(Z) = \frac{8}{9}$

<「이공학도를 위한 확률 및 통계학」 p.156>

# 부록 #12 – 연습문제 (4.70~4.71)

#4.70

(a)

따름정리 4.5에 따르면 확률변수 X와 Y가 서로 독립이면 $\sigma_{XY}=0$임을 이용한다.

$\sigma_{XY} = E(XY) - \mu_X\mu_Y$

$E(XY) = \int_0^1\int_0^1 xy\cdot\frac{3}{2}(x^2+y^2)\,dx\,dy$

$= \frac{3}{2}\int_0^1\int_0^1 x^3y + xy^3\,dx\,dy$

$\int_0^1 x^3y + xy^3\,dx$

$= \left[\frac{1}{4}x^4y + \frac{1}{2}x^2y^3\right]_0^1$

$= \frac{1}{4}y + \frac{1}{2}y^3$

$E(XY) = \frac{3}{2}\int_0^1 \frac{1}{4}y + \frac{1}{2}y^3\,dy$

$= \frac{3}{2}\left[\frac{1}{8}y^2 + \frac{1}{8}y^4\right]_0^1$

$= \frac{3}{2}\cdot\frac{2}{8} = \frac{3}{8}$

$\mu_X = \int_0^1\int_0^1 \frac{3}{2}x(x^2+y^2)\,dx\,dy$

$= \frac{3}{2}\int_0^1\int_0^1 x^3 + xy^2\,dx\,dy$

$\int_0^1 x^3 + xy^2\,dx$

$= \left[\frac{1}{4}x^4 + \frac{1}{2}x^2y^2\right]_0^1$

$= \frac{1}{4} + \frac{1}{2}y^2$

$\mu_X = \frac{3}{2}\int_0^1 \frac{1}{4} + \frac{1}{2}y^2\,dy$

$= \frac{3}{2}\left[\frac{1}{4}y + \frac{1}{6}y^3\right]_0^1$

$= \frac{3}{2}\left(\frac{1}{4} + \frac{1}{6}\right)$

$= \frac{\cancel{3}}{2}\cdot\frac{5}{\cancel{12}\,4}$

$= \frac{5}{8}$

$\mu_Y = \int_0^1\int_0^1 y\cdot\frac{3}{2}(x^2+y^2)\,dx\,dy$

$= \frac{3}{2}\int_0^1\int_0^1 x^2y + y^3\,dx\,dy$

$\int_0^1 x^2y + y^3\,dx$

$= \left[\frac{1}{3}x^3y + xy^3\right]_0^1$

$= \frac{1}{3}y + y^3$

$\Rightarrow \mu_Y = \frac{3}{2}\int_0^1 \frac{1}{3}y + y^3\,dy$

$= \frac{3}{2}\left[\frac{1}{6}y^2 + \frac{1}{4}y^4\right]_0^1$

$= \frac{3}{2}\cdot\frac{2+3}{12}$

$= \frac{5}{8}$

$\therefore \sigma_{XY} = \frac{3}{8} - \frac{5}{8}\times\frac{5}{8} \neq 0$ 이므로

X와 Y는 서로 독립이 아니다.

(b)

(a)에 의해..

$E(X+Y) = E(X) + E(Y)$

$= \frac{5}{8} + \frac{5}{8} = \frac{5}{4}$

$E(XY) = \frac{3}{8}$

(c)

$Var(X) = E(X^2) - \mu_X^2$

$E(X^2) = \int_0^1\int_0^1 \frac{3}{2}x^2(x^2+y^2)\,dx\,dy$

$= \frac{3}{2}\int_0^1\int_0^1 x^4 + x^2y^2\,dx\,dy$

$\int_0^1 x^4 + x^2y^2\,dx$

$= \left[\frac{1}{5}x^5 + \frac{1}{3}x^3y^2\right]_0^1$

$= \frac{1}{5} + \frac{1}{3}y^2$

$\Rightarrow E(X^2) = \frac{3}{2}\int_0^1 \frac{1}{5} + \frac{1}{3}y^2\,dy$

$= \frac{3}{2}\left[\frac{1}{5}y + \frac{1}{9}y^3\right]_0^1$

$= \frac{3}{2}\cdot\frac{9+5}{45} = \frac{\cancel{3}}{\cancel{2}}\cdot\frac{\cancel{14}\,7}{\cancel{45}\,15} = \frac{7}{15}$

$Var(X) = \frac{7}{15} - \left(\frac{5}{8}\right)^2$

$= \frac{7}{15} - \frac{25}{64}$

$= \frac{448-375}{960}$

$= \frac{73}{960}$

Var(Y)도 Var(X)와 같은 꼴이기 때문에 $Var(X) = Var(Y) = \frac{73}{960}$

$Cov(X,Y) = E(XY) - \mu_X\mu_Y$

$= \frac{3}{8} - \frac{5}{8}\times\frac{5}{8}$

$= \frac{24-25}{64} = -\frac{1}{64}$

(d)

$Var(X+Y)$

$= \sigma^2_{X+Y}$

$= \sigma^2_X + \sigma^2_Y + 2\sigma_{XY}$

$= \frac{73}{960} + \frac{73}{960} - \frac{1}{32}$

$= \frac{73+73-30}{960}$ 146 30

$= \frac{116}{960} = \frac{29}{240}$

(a) 독립이 아니다

(b) $E(X+Y) = \frac{5}{4}$, $E(XY) = \frac{3}{8}$

(c) $Var(X) = Var(Y) = \frac{73}{960}$

$Cov(X,Y) = -\frac{1}{64}$

(d) $Var(X+Y) = \frac{29}{240}$

#4.71

(a)

$E(Y) = \int_0^\infty y\cdot\frac{1}{4}e^{-\frac{y}{4}}\,dy$

$= \left[y\cdot\left(-e^{-\frac{y}{4}}\right)\right]_0^\infty + \int_0^\infty e^{-\frac{y}{4}}\,dy$

$= \left[-4e^{-\frac{y}{4}}\right]_0^\infty$

$= 4$

(b)

$E(Y^2) = \int_0^\infty \frac{1}{4}y^2e^{-\frac{y}{4}}\,dy$

$= \cancel{\left[y^2\cdot\left(-e^{-\frac{y}{4}}\right)\right]_0^\infty} + \int_0^\infty 2y\cdot e^{-\frac{y}{4}}\,dy$

$= 2\int_0^\infty y\cdot e^{-\frac{y}{4}}\,dy$

(a)에서 $\int_0^\infty \frac{1}{4}ye^{-\frac{y}{4}}\,dy = 4$ 이므로

$2\int_0^\infty y\cdot e^{-\frac{y}{4}}\,dy = 4\cdot 8 = 32$

$Var(Y) = E(Y^2) - \{E(Y)\}^2$

$= 32 - 4^2$

$= 16$

(a) $E(Y) = 4$, (b) $E(Y^2) = 32$

$Var(Y) = 16$

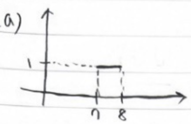
OMNIBUS

<「이공학도를 위한 확률 및 통계학」 p.157>

# 부록 #13 – 연습문제 (4.72~4.77)

#4.72

(a)




(b)

$E(Y) = \int_7^8 y \cdot 1\,dy$

$= \left[\frac{1}{2}y^2\right]_7^8$

$= \frac{64-49}{2}$

$= \frac{15}{2}$

$E(Y^2) = \int_7^8 y^2 \cdot 1\,dy$

$= \left[\frac{1}{3}y^3\right]_7^8$

$= \frac{512-343}{3}$

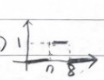
$= \frac{169}{3}$

$Var(Y) = \frac{169}{3} - \left(\frac{15}{2}\right)^2$

$= \frac{169}{3} - \frac{225}{4}$

$= \frac{676-675}{12}$

$= \frac{1}{12}$

(a) 


(b) $E(Y) = \frac{15}{2}$, $E(Y^2) = \frac{169}{3}$, $Var(Y) = \frac{1}{12}$.

#4.73

$E(e^Y) = \int_7^8 e^y f(y)\,dy$

$= \int_7^8 e^y\,dy$

$= [e^y]_7^8$

$= e^8 - e^7$

$= 1884.32$

<테일러 급수 (Taylor series)>

$f(y) = f(a) + f'(a)(y-a) + \frac{f''(a)}{2!}(y-a)^2 + \frac{f'''(a)}{3!}(y-a)^3 + \cdots$

따라서, $e^y = 1 + y + \frac{y^2}{2!} + \frac{y^3}{3!} + \cdots$

$\int_7^8 e^y\,dy$

$= \int_7^8 1 + y + \frac{y^2}{2!} + \frac{y^3}{3!} + \cdots\,dy$

$= \left[y + \frac{1}{2}y^2 + \frac{1}{3!}y^3 + \frac{1}{4!}y^4 + \cdots\right]_7^8$

앞의 20개 항을 사용하여 계산하면 1884.06이 도출된다.

따라서, 정리 4.1을 이용한 계산과 테일러 급수 (Taylor series)를 이용한 계산한 결과가 유사함을 알 수 있다.

#4.74

$E(Z^2) = E(e^{2Y})$

$= \int_7^8 e^{2y}\,dy$

$= \left[\frac{1}{2}e^{2y}\right]_7^8$

$= \frac{e^{16} - e^{14}}{2}$

$\simeq 3,841,753.118$

테일러급수를 사용하면,

$e^{2y} = 1 + 2y + 2y^2 + \frac{4}{3}y^3 + \cdots$

$\int_7^8 e^{2y}\,dy$

$= \int_7^8 1 + 2y + 2y^2 + \frac{4}{3}y^3 + \cdots\,dy$

$= \left[y + y^2 + \frac{2}{3}y^3 + \frac{4}{12}y^4 + \cdots\right]_7^8$

앞의 20개 항을 사용하여 계산한 결과는 3,284,823.839이다.

즉, $f(y)$를 이용하여 $Var(e^Y)$를 구하면,

$3,841,753.118 - (1884.32)^2$

$= 291,091,256$ 이고,

테일러급수를 이용하여 $Var(e^Y)$를 구하면,

$3,284,823.839 - (1884.06)^2$

$= -264.858.245$

서로 다른 방법으로 도출한 분산의 값은 차이가 있음이 파악된다

#4.75

$E(X) = 900,\ \sigma = 50$

㉠ $P(X \le 700)$

$P(X \le 900 - 4 \cdot 50) + P(X \ge 900 + 4 \cdot 50)$

$= 1 - P(900 - 4 \cdot 50 < X < 900 + 4 \cdot 50)$

$P(900 - 4 \cdot 50 < X < 900 + 4 \cdot 50) \ge 1 - \frac{1}{4^2} = \frac{15}{16}$

$\Rightarrow P(X \le 900 - 4 \cdot 50) + P(X \ge 900 + 4 \cdot 50) \le \frac{1}{16}$

$P(X \le 900 - 4 \cdot 50) = P(X \ge 900 + 4 \cdot 50)$ 이므로

$P(X \le 900 - 4 \cdot 50) = P(X \le 700) \le \frac{1}{32}$

$\frac{1}{32} = 0.03125$, 즉 3.125%

$\therefore$ 3.125%

#4.76

$E(X) = 60,\ \sigma = 6$

1000명 지원 → 70명 선발

∴ 상위 7% 선발

$P(60 - 4 \cdot 6 < X < 60 + 4 \cdot 6) \ge 1 - \frac{1}{4^2} = \frac{15}{16}$

$\Rightarrow P(X < 60 - 4 \cdot 6) + P(X > 60 + 4 \cdot 6) \le \frac{1}{16}$

$P(X > 60 + 4 \cdot 6) \le \frac{1}{32}$

84점 초과인 사람들은 3.125% 이므로, 84점을 받은 지원자는 상위 7%에 포함되므로 채용될 수 있다

$\therefore$ 채용 가능

#4.77

$\mu = 10,\ \sigma = 2$

$P(-3 < X - 10 < 3)$

$= P(10 - 3 < X < 10 + 3)$

$= P(10 - 2 \cdot \frac{3}{2} < X < 10 + 2 \cdot \frac{3}{2})$

$\ge 1 - \frac{1}{(\frac{3}{2})^2} = 1 - \frac{4}{9} = \frac{5}{9}$

(a) $P(|X - 10| \ge 3)$

$= 1 - P(|X - 10| < 3)$

$\le 1 - \frac{5}{9} = \frac{4}{9}$

(b) $P(|X - 10| < 3) \ge \frac{5}{9}$

(c) $P(5 < X < 15)$

$= P(10 - 2 \cdot \frac{5}{2} < X < 10 + 2 \cdot \frac{5}{2})$

$\ge 1 - \frac{1}{(\frac{5}{2})^2} = 1 - \frac{4}{25} = \frac{21}{25}$

OMNIBUS

<「이공학도를 위한 확률 및 통계학」 p.157>

# 부록 #14 – 연습문제 (4.78)

(d)

$P(|X-10| \ge c)$

$= P(X-10 \le -c) + P(X-10 \ge c)$

$= P(X \le 10-c) + P(X \ge 10+c)$

$= 1 - P(10-c < X < 10+c)$

$P(10-c < X < 10+c)$

$= P(10 - 2\cdot\frac{c}{2} < X < 10 + 2\cdot\frac{c}{2})$

$\ge 1 - \frac{1}{(\frac{c}{2})^2} = 1 - \frac{4}{c^2}$

$\Rightarrow P(|X-10| \ge c) \le 1 - (1 - \frac{4}{c^2})$

$= \frac{4}{c^2}$

$\frac{4}{c^2} = 0.04$

$\frac{4}{c^2} = \frac{4}{100}$ 이므로, $c = 10$

$\therefore$ (a) $\frac{4}{9}$ 이하, (b) $\frac{5}{9}$ 이상

(c) $\frac{21}{25}$ 이상, (d) $c = 10$

#4.78

$\mu = \int_0^1 x \cdot 6x(1-x)\,dx$

$= 6\int_0^1 x^2 - x^3\,dx$

$= 6\left[\frac{1}{3}x^3 - \frac{1}{4}x^4\right]_0^1$

$= 6 \cdot \frac{1}{12} = \frac{1}{2}$

$E(X^2) = \int_0^1 x^2 \cdot 6x(1-x)\,dx$

$= 6\int_0^1 x^3 - x^4\,dx$

$= 6\left[\frac{1}{4}x^4 - \frac{1}{5}x^5\right]_0^1$

$= 6 \cdot \frac{1}{20} = \frac{3}{10}$

$\sigma^2 = \frac{3}{10} - (\frac{1}{2})^2$

$= \frac{3}{10} - \frac{1}{4} = \frac{6-5}{20} = \frac{1}{20}$

$\sigma = \sqrt{\frac{1}{20}} = \frac{1}{2\sqrt{5}} = \frac{\sqrt{5}}{10}$

㉮ $P(\mu - 2\sigma < X < \mu + 2\sigma)$

$= P(\frac{1}{2} - \frac{\sqrt{5}}{5} < X < \frac{1}{2} + \frac{\sqrt{5}}{5})$

$= \int_{\frac{1}{2}-\frac{\sqrt{5}}{5}}^{\frac{1}{2}+\frac{\sqrt{5}}{5}} f(x)\,dx$

$= \int_{\frac{1}{2}-\frac{\sqrt{5}}{5}}^{\frac{1}{2}+\frac{\sqrt{5}}{5}} 6x(1-x)\,dx$

$= \left[3x^2 - 2x^3\right]_{\frac{1}{2}-\frac{\sqrt{5}}{5}}^{\frac{1}{2}+\frac{\sqrt{5}}{5}}$

$\simeq 0.984$

<체비셰프 정리 활용>

$P(\mu - 2\sigma < X < \mu + 2\sigma)$

$\ge 1 - \frac{1}{2^2} = \frac{3}{4} = 0.75$

앞서구한 0.984의 값는 체비셰프 정리를 활용하여 도출한 확률은 0.75 이상에 포함되므로 옳은 결과임을 파악할 수 있다.

<「이공학도를 위한 확률 및 통계학」 p.158>

# 부록 #15 – MATLAB 코드 (1/17)

- 그림 4.1 (1/3)

```
f = @(x) 2*x.^2 - 5*x + 2;
x_range = linspace(-2, 4, 400);
y_values = f(x_range);

figure;
plot(x_range, y_values, 'b-', 'LineWidth', 2);
hold on;

a = -1.23;
b = 3.45;

c = (a + b) / 2;

fc = f(c);
```

# 부록 #16 – MATLAB 코드 (2/17)

• 그림 4.1 (2/3)

```
plot([a, a], [0, f(a)], 'r--', 'LineWidth', 1);
plot(a, 0, 'ro', 'MarkerSize', 7, 'MarkerFaceColor', 'r');
text(a, 0, '  a', 'VerticalAlignment', 'bottom', 'HorizontalAlignment', 'right', 'FontSize', 20);

plot([b, b], [0, f(b)], 'r--', 'LineWidth', 1);
plot(b, 0, 'ro', 'MarkerSize', 7, 'MarkerFaceColor', 'r');
text(b, 0, '  b', 'VerticalAlignment', 'bottom', 'HorizontalAlignment', 'left', 'FontSize', 20);

plot([c, c], [0, fc], 'k--', 'LineWidth', 1);
plot(c, 0, 'ro', 'MarkerSize', 7, 'MarkerFaceColor', 'r');
text(c, 0, '  c', 'VerticalAlignment', 'bottom', 'HorizontalAlignment', 'left', 'FontSize', 20);

plot([0, c], [fc, fc], 'k--', 'LineWidth', 1);
plot(0, fc, 'ro', 'MarkerSize', 7, 'MarkerFaceColor', 'r');
text(0, fc, '  f(c)', 'VerticalAlignment', 'bottom', 'HorizontalAlignment', 'right', 'FontSize', 20);

xlabel('x');
ylabel('f(x)');
grid on;
```

# 부록 #17 – MATLAB 코드 (3/17)

- 그림 4.1 (3/3)

```
ax = gca;
ax.XAxisLocation = 'origin';
ax.YAxisLocation = 'origin';
ax.XTick = [];
ax.YTick = [];
xlim([-2 4]);
ylim([min(y_values) max(y_values)]);

xl = xlim;
yl = ylim;
line([xl(1) xl(2)], [0 0], 'Color', 'k', 'LineWidth', 1);
line([0 0], [yl(1) yl(2)], 'Color', 'k', 'LineWidth', 1);
plot(xl(2), 0, 'k>', 'MarkerSize', 7, 'MarkerFaceColor', 'k');
plot(0, yl(2), 'k^', 'MarkerSize', 7, 'MarkerFaceColor', 'k');

hold off;
```

# 부록 #18 – MATLAB 코드 (4/17)

- 그림 4.2

```
mu = 2;
sigma1 = 1;

x = mu - 4*sigma1:0.01:mu + 4*sigma1;

y2 = (1 / (sigma1 * sqrt(2 * pi))) * exp(-0.5 * ((x - mu) / sigma1) .^ 2);

figure;
plot(x, y2, 'Color', 'Blue', 'LineWidth', 2);
hold on;
line([mu mu], ylim, 'Color', 'Black', 'LineStyle', '--');
text(mu, min(ylim), '\mu = 2', 'VerticalAlignment', 'bottom', 'HorizontalAlignment', 'center');
xlabel('\it{X}');
hold off;
grid on;
```

# 부록 #19 – MATLAB 코드 (5/17)

- 그림 4.3

```
mu = 2;
sigma2 = 3;

x = mu - 4*sigma2:0.01:mu + 4*sigma2;

y2 = (1 / (sigma2 * sqrt(2 * pi))) * exp(-0.5 * ((x - mu) / sigma2) .^ 2);

figure;
plot(x, y2, 'Color', 'Blue', 'LineWidth', 2);
hold on;
line([mu mu], ylim, 'Color', 'Black', 'LineStyle', '--');
text(mu, 0.01, '\mu = 2', 'VerticalAlignment', 'bottom', 'HorizontalAlignment', 'center'); % Adjust
ed text position
xlabel('\it{X}');
ylim([0 0.4]); % Setting y-axis limit to 0.4
hold off;
grid on;
```

# 부록 #20 – MATLAB 코드 (6/17)

- 그림 4.4 (1/3)

```
nPoints = 100;

meanX = 175;
meanY = 70;

stdX = 8;
stdY = 5;

rho = 0.95;
covXY = rho * stdX * stdY;
covarianceMatrix = [stdX^2, covXY; covXY, stdY^2];
[V, D] = eig(covarianceMatrix);
Z = randn(nPoints, 2);

data = Z * sqrt(D) * V';
data = bsxfun(@plus, data, [meanX, meanY]);

X = data(:, 1);
Y = data(:, 2);
```

- 그림 4.4 (2/3)

```
figure;
scatter(X, Y, 'filled');
hold on;

xLimits = [min(X) - 10, max(X) + 10];
yLimits = [min(Y) - 10, max(Y) + 10];
xlim(xLimits);
ylim(yLimits);

line([meanX, meanX], yLimits, 'Color', 'k', 'LineStyle', '--', 'LineWidth', 1);
line(xLimits, [meanY, meanY], 'Color', 'k', 'LineStyle', '--', 'LineWidth', 1);
```

# 부록 #22 – MATLAB 코드 (8/17)

- 그림 4.4 (3/3)

```
text(meanX, min(yLimits) - 1, '\mu_X', 'HorizontalAlignment', 'center', 'VerticalAlignment', 'to
p', 'FontSize', 12);
text(min(xLimits) - 1, meanY, '\mu_Y', 'HorizontalAlignment', 'right', 'VerticalAlignment', 'mid
dle', 'FontSize', 12);

xlabel('X', 'FontSize', 12);
ylabel('Y', 'FontSize', 12);
grid on;
axis equal;

set(gca, 'FontSize', 12);

hold off;
```

# 부록 #23 – MATLAB 코드 (9/17)

• 그림 4.5 (1/3)

```
nPoints = 100;

meanX = 175;
meanY = 70;

stdX = 8;
stdY = 5;

rho = -0.95;
covXY = rho * stdX * stdY;
covarianceMatrix = [stdX^2, covXY; covXY, stdY^2];
[V, D] = eig(covarianceMatrix);
Z = randn(nPoints, 2);

data = Z * sqrt(D) * V';
data = bsxfun(@plus, data, [meanX, meanY]);

X = data(:, 1);
Y = data(:, 2);
```

# 부록 #24 – MATLAB 코드 (10/17)

- 그림 4.5 (2/3)

```
figure;
scatter(X, Y, 'filled');
hold on;

xLimits = [min(X) - 10, max(X) + 10];
yLimits = [min(Y) - 10, max(Y) + 10];
xlim(xLimits);
ylim(yLimits);

line([meanX, meanX], yLimits, 'Color', 'k', 'LineStyle', '--', 'LineWidth', 1);
line(xLimits, [meanY, meanY], 'Color', 'k', 'LineStyle', '--', 'LineWidth', 1);
```

- 그림 4.5 (3/3)

```
text(meanX, min(yLimits) - 1, '\mu_X', 'HorizontalAlignment', 'center', 'VerticalAlignment', 'to
p', 'FontSize', 12);
text(min(xLimits) - 1, meanY, '\mu_Y', 'HorizontalAlignment', 'right', 'VerticalAlignment', 'mid
dle', 'FontSize', 12);

xlabel('X', 'FontSize', 12);
ylabel('Y', 'FontSize', 12);
grid on;
axis equal;

set(gca, 'FontSize', 12);

hold off;
```

# 부록 #26 – MATLAB 코드 (12/17)

- 그림 4.6 (1/3)

```
nPoints = 100;

meanX = 175;
meanY = 70;

stdX = 8;
stdY = 5;

rho = 0;  % Set correlation to 0 for independence
covXY = rho * stdX * stdY;
covarianceMatrix = [stdX^2, covXY; covXY, stdY^2];
[V, D] = eig(covarianceMatrix);
Z = randn(nPoints, 2);

data = Z * sqrt(D) * V';
data = bsxfun(@plus, data, [meanX, meanY]);

X = data(:, 1);
Y = data(:, 2);
```

# 부록 #27 – MATLAB 코드 (13/17)

- 그림 4.6 (2/3)

```
figure;
scatter(X, Y, 'filled');
hold on;

xLimits = [min(X) - 10, max(X) + 10];
yLimits = [min(Y) - 10, max(Y) + 10];
xlim(xLimits);
ylim(yLimits);

line([meanX, meanX], yLimits, 'Color', 'k', 'LineStyle', '--', 'LineWidth', 1);
line(xLimits, [meanY, meanY], 'Color', 'k', 'LineStyle', '--', 'LineWidth', 1);
```

# 부록 #28 – MATLAB 코드 (14/17)

- 그림 4.6 (3/3)

```
text(meanX, min(yLimits) - 1, '\mu_X', 'HorizontalAlignment', 'center', 'VerticalAlignment', 'top', 
'FontSize', 12);
text(min(xLimits) - 1, meanY, '\mu_Y', 'HorizontalAlignment', 'right', 'VerticalAlignment', 'middl
e', 'FontSize', 12);

xlabel('X', 'FontSize', 12);
ylabel('Y', 'FontSize', 12);
grid on;
axis equal;

set(gca, 'FontSize', 12);

hold off;
```

# 부록 #29 – MATLAB 코드 (15/17)

- 그림 4.7

```
temperatures_celsius = [18.49, 21.73, 20.69, 22.22, 21.29, 22.18, 20.85, 19.53, 21.32, 20.18];
ac_sales = [469.72, 523.10, 511.31, 532.91, 520.24, 530.88, 514.95, 492.89, 520.96, 507.16];
temperatures_fahrenheit = (temperatures_celsius * 9/5) + 32;

figure;
subplot(1, 2, 1);
scatter(temperatures_celsius, ac_sales, 'filled', 'blue');
xlabel('Temperature (°C)');
ylabel('AC Sales');
title('Celsius vs AC Sales');
grid on;

subplot(1, 2, 2);
scatter(temperatures_fahrenheit, ac_sales, 'filled', 'red');
xlabel('Temperature (°F)');
ylabel('AC Sales');
title('Fahrenheit vs AC Sales');
grid on;
```

- 그림 4.8 (1/2)

```
mu = 0;
sigma1 = 1;
sigma2 = 2;
k = 2;

x = linspace(mu - 4*max(sigma1, sigma2), mu + 4*max(sigma1, sigma2), 1000);

pdf_normal1 = (1/(sigma1 * sqrt(2 * pi))) * exp(-0.5 * ((x - mu)/sigma1).^2);

pdf_normal2 = (1/(sigma2 * sqrt(2 * pi))) * exp(-0.5 * ((x - mu)/sigma2).^2);

figure;
plot(x, pdf_normal1, 'Color', [0.678, 0.847, 0.902], 'LineWidth', 2);
hold on;
plot(x, pdf_normal2, 'Color', [0.698, 0.933, 0.509], 'LineWidth', 2);
```

- 그림 4.8 (2/2)

```
plot([mu mu], [0 0.4], 'k--', 'LineWidth', 1);

legend('σ = 1', 'σ = 2');

hold off;

xlabel('X');
ylabel('Probability Density');

ylim([0, 0.4]);

grid on;
```