확률 및 통계

기말고사 문제풀이

2023.06.13.

기말고사

- **문제풀이**
  - 1) Problem 2.10 (중간고사 3번)
  - 2) Problem 2.35 (과제#4 1번)
  - 3) Problem 3.6 (중간고사 6번)
  - 4) Problem 4.41 (과제#4 2번)
  - 5) Problem 5.27 (과제#4 3번)
  - 6) Example 5.6
  - 7) Problem 7.10 (과제#4 5번)
  - 8) Problem 7.5
  - 9) Problem 6.11 (과제#4 4번)
  - 10) Problem 6.24 (과제#3 5번)

1) Problem 2.10 (Probability Model)

Consider a 12-sided (dodecahedral) unfair die. In rolling this die, the even numbered sides are twice as likely as the odd numbered sides. Define the events: A={odd numbered side} and B={4, 5, 6, 7, 8}.

(a) Find the probability Pr[A]
(b) Find the probability Pr[AB]
(c) Are the events A and B independent?

(a) $Pr[A] + Pr[A^C] = Pr[A] + 2 \cdot Pr[A] = 1 \rightarrow Pr[A] = 1/3$

$Pr[odd] = 1/3 \cdot 1/6, \quad Pr[even] = 2/3 \cdot 1/6$

$Pr[B] = Pr[\{2 \cdot odd + 3 \cdot even\}] = 2 \cdot (1/3 \cdot 1/6) + 3 \cdot (2/3 \cdot 1/6) = 4/9$

(b) $Pr[AB] = Pr[\{5, 7\}] = 2 \cdot 1/3 \cdot 1/6 = 1/9$

(c) $Pr[AB] \neq Pr[A] \cdot Pr[B] \rightarrow$ A and B are not independent.

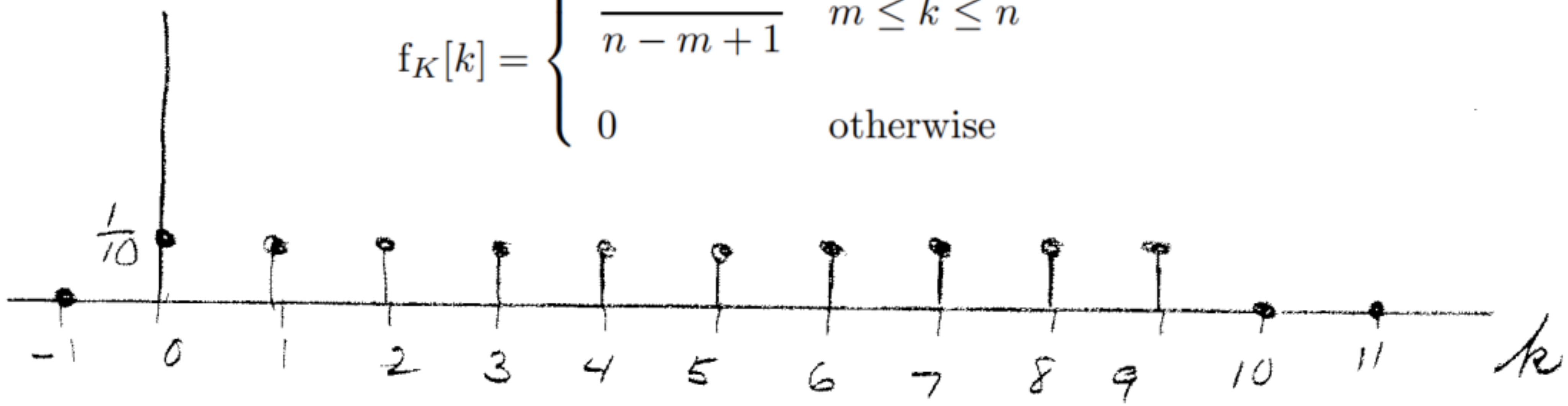
3) Problem 3.6 (PMF)

Sketch the PMF $f_K[k]$ for the following random variables and label the plots.

(a) Uniform. m=0, n=9
(b) Bernoulli. p=1/10
(c) Binomial. n=10, p=1/10
(d) Poisson. α=1

(a)

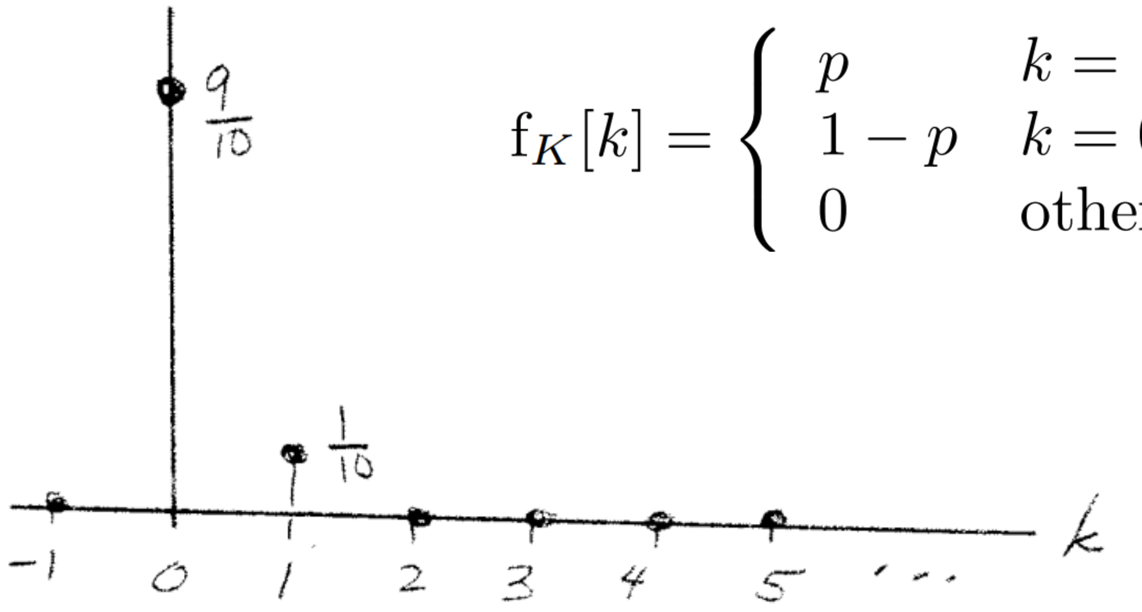
$$f_K[k] = \begin{cases} \dfrac{1}{n-m+1} & m \le k \le n \\ 0 & \text{otherwise} \end{cases}$$



3) Problem 3.6 (PMF)

(b) Bernoulli. p=1/10

(b)

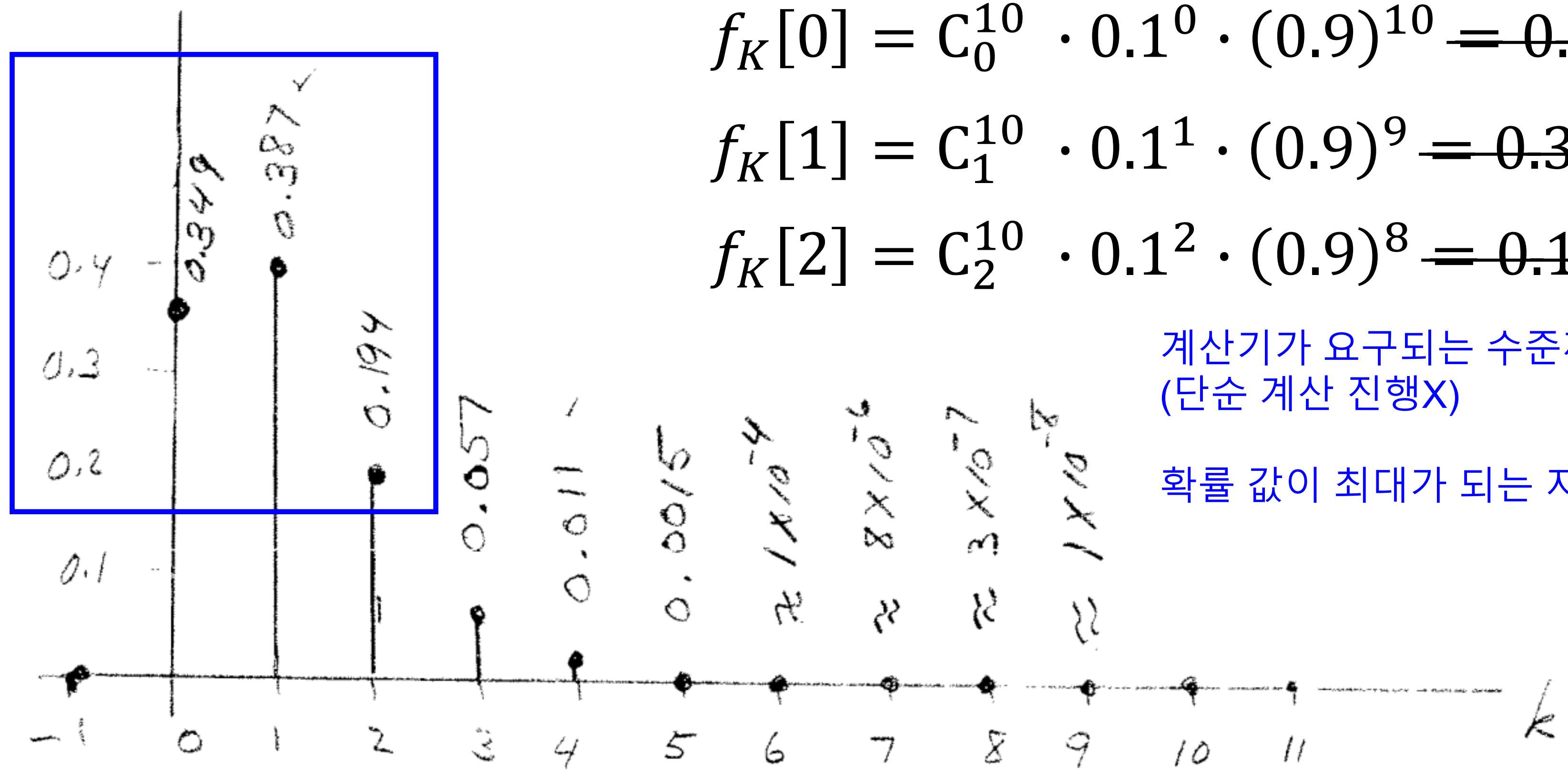


$$f_K[k] = \begin{cases} p & k = 1 \\ 1 - p & k = 0 \\ 0 & \text{otherwise} \end{cases}$$

3) Problem 3.6 (PMF)

(c) Binomial. n=10, p=1/10

(c) $f_K[k] = C_k^n \cdot p^k \cdot (1-p)^{n-k}$



$$f_K[0] = C_0^{10} \cdot 0.1^0 \cdot (0.9)^{10} \cancel{= 0.348678..}$$

$$f_K[1] = C_1^{10} \cdot 0.1^1 \cdot (0.9)^{9} \cancel{= 0.387420..}$$

$$f_K[2] = C_2^{10} \cdot 0.1^2 \cdot (0.9)^{8} \cancel{= 0.193710..}$$

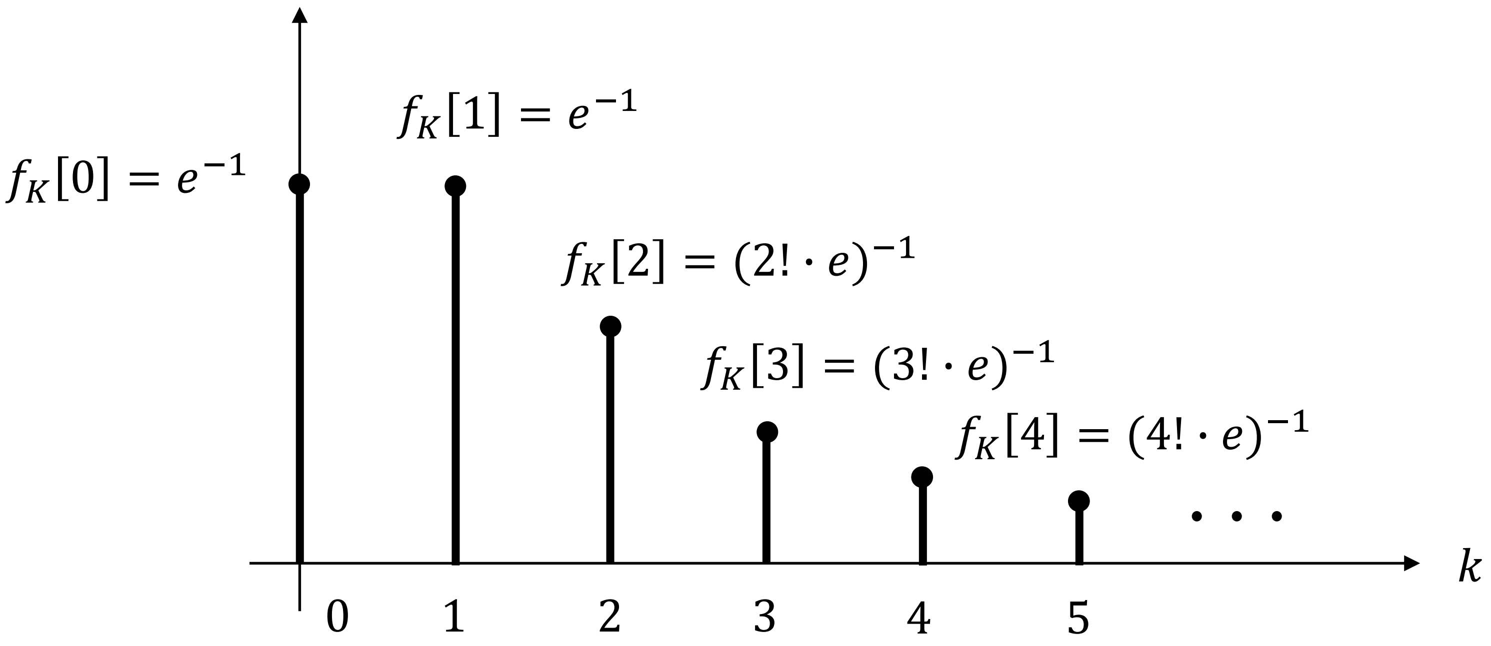
계산기가 요구되는 수준까지만 수식 전개
(단순 계산 진행X)

확률 값이 최대가 되는 지점(=1)은 인지해야 함

3) Problem 3.6 (PMF)

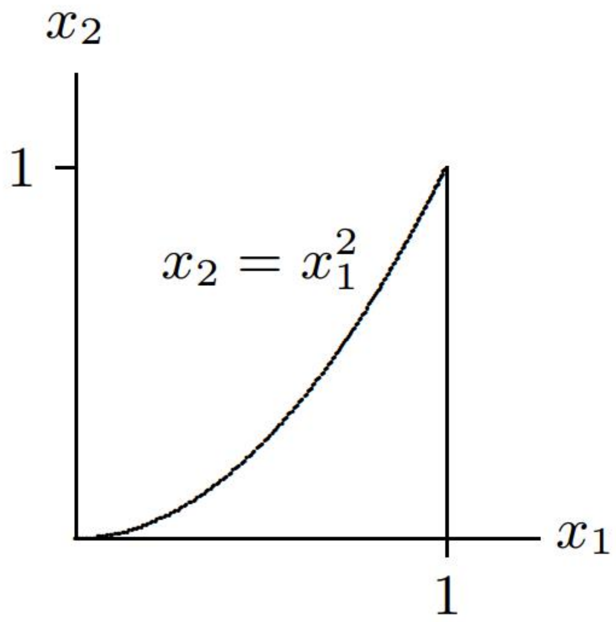
(d) Poisson. α=1

$$\rightarrow f_K[k] = \frac{\alpha^k}{k!} e^{-\alpha}$$

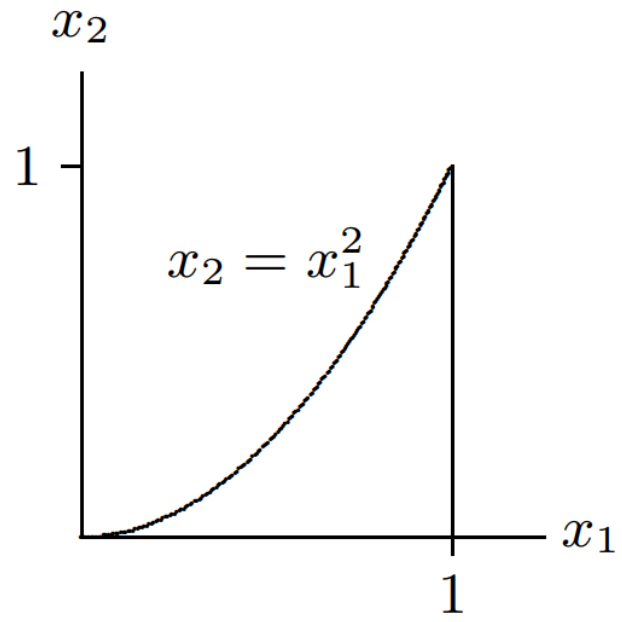


6) Example 5.6 (Two Random Variables)

Example 5.6: Two random variables are described by the joint density function shown below.



$$f_{X_1X_2}(x_1, x_2) = \begin{cases} 10x_2 & 0 \le x_2 \le x_1^2, \quad 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$



$$f_{X_1X_2}(x_1,x_2) = \begin{cases} 10x_2 & 0 \le x_2 \le x_1^2, \quad 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

The mean of random variable X_1 is given by

$$m_1 = \mathcal{E}\{X_1\} = \int_0^1 \int_0^{x_1^2} x_1 \cdot 10x_2\, dx_2\, dx_1 = 10\int_0^1 x_1 \int_0^{x_1^2} x_2\, dx_2\, dx_1 = \frac{5}{6}$$

Likewise the mean of X_2 is given by

$$m_2 = \int_0^1 \int_0^{x_1^2} x_2\, 10x_2\, dx_2\, dx_1 = 10\int_0^1 \int_0^{x_1^2} x_2^2\, dx_2\, dx_1 = \frac{10}{21}$$

6) Example 5.6 (Two Random Variables)

The variances of the two random variables are computed as

$$\mathcal{E}\left\{X_1^2\right\} = \int_0^1 \int_0^{x_1^2} x_1^2\, 10 x_2\, dx_2\, dx_1 \;= \frac{5}{7}$$

$$\sigma_1^2 = \mathcal{E}\left\{X_1^2\right\} - m_1^2 = \frac{5}{7} - \left(\frac{5}{6}\right)^2 = \frac{5}{252}$$

and

$$\mathcal{E}\left\{X_2^2\right\} = \int_0^1 \int_0^{x_1^2} x_2^2\, 10 x_2\, dx_2\, dx_1 \;= \frac{5}{18}$$

$$\sigma_2^2 = \mathcal{E}\left\{X_2^2\right\} - m_2^2 = \frac{5}{18} - \left(\frac{10}{21}\right)^2 = \frac{5}{98}$$

6) Example 5.6 (Two Random Variables)

The correlation r is given by

$$r = \mathcal{E}\{X_1 X_2\} = \int_0^1 \int_0^{x_1^2} x_1 x_2 \, 10 x_2 \, dx_2 \, dx_1 = \frac{5}{12}$$

The covariance is then computed using (5.28) as

$$c = \mathcal{E}\{X_1 X_2\} - m_x m_y = \frac{5}{12} - \left(\frac{5}{6}\right)\left(\frac{10}{21}\right) = \frac{5}{252}$$

Finally, the correlation coefficient ρ is computed as the normalized covariance.

$$\rho = \frac{c}{\sigma_1 \sigma_2} = \frac{5/252}{\sqrt{5/252}\sqrt{5/98}} = \sqrt{\frac{7}{18}}$$

6) Example 7.2 (Random Vector)

Example 7.2: The two jointly-distributed random variables described in Chapter 5, Example 5.6 are taken to be components of a random vector

$$\boldsymbol{X} = [X_1, X_2]^T$$

The mean vector and correlation matrix for this random vector are given by

$$\mathbf{m}_{\boldsymbol{X}} = \begin{bmatrix} \mathcal{E}\{X_1\} \\ \mathcal{E}\{X_2\} \end{bmatrix} = \begin{bmatrix} 5/6 \\ 10/21 \end{bmatrix}$$

and

$$\mathbf{R}_{\boldsymbol{X}} = \begin{bmatrix} \mathcal{E}\{X_1^2\} & \mathcal{E}\{X_1X_2\} \\ \mathcal{E}\{X_2X_1\} & \mathcal{E}\{X_2^2\} \end{bmatrix} = \begin{bmatrix} 5/7 & 5/12 \\ 5/12 & 5/18 \end{bmatrix}$$

where the numerical values of the moments are taken from Example 5.6. The covariance matrix can then be computed from (7.18) as

$$\mathbf{C}_{\boldsymbol{X}} = \begin{bmatrix} 5/7 & 5/12 \\ 5/12 & 5/18 \end{bmatrix} - \begin{bmatrix} 5/6 \\ 10/21 \end{bmatrix} \begin{bmatrix} 5/6 & 10/21 \end{bmatrix} = \begin{bmatrix} 5/252 & 5/252 \\ 5/252 & 5/98 \end{bmatrix}$$

The elements σ_1^2, c, and σ_2^2 correspond to the numerical values computed in the earlier Example 5.6.

8) Problem 7.5 (Random Vector)

7.5 Find the sample mean, sample correlation matrix, and sample covariance matrix (7.21) for the following data:

$$X_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad X_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad X_3 = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \quad X_4 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Check your results using (7.22).

$$M_4 = \frac{1}{4}\left[\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right] = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$R_4 = \frac{1}{4}\left[\begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix}\begin{bmatrix} 1 & 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \end{bmatrix}\begin{bmatrix} 2 & 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \end{bmatrix}\right] = \begin{bmatrix} 3/2 & 3/4 \\ 3/4 & 3/2 \end{bmatrix}$$

8) Problem 7.5 (Random Vector)

7.5 Find the sample mean, sample correlation matrix, and sample covariance matrix (7.21) for the following data:

$$X_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad X_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad X_3 = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \quad X_4 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$C_4 = \frac{1}{3}\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix}\begin{bmatrix} 1 & -1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix}\begin{bmatrix} -1 & 0 \end{bmatrix}\right]$$

$$= \begin{bmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{bmatrix}$$

8) Problem 7.5 (Random Vector)

7.5 Find the sample mean, sample correlation matrix, and sample covariance matrix (7.21) for the following data:

$$X_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad X_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad X_3 = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \quad X_4 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Check your results using (7.22).

$$C_4 = \frac{4}{3}\left(R_4 - M_4 M_4^T\right)$$

$$= \frac{4}{3}\left[\begin{bmatrix} 3/2 & 3/4 \\ 3/4 & 3/2 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \end{bmatrix}\right] = \begin{bmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{bmatrix}$$

8) Problem 7.7 (k-dim. Gaussian)

- **Multivariate Gaussian Density Function**
  - For a K-dimensional Gaussian random vector $\boldsymbol{x} = [x_1, x_2, \ldots, x_K]^T$

$$f_{X_1, X_2, X_2, \cdots, X_K}(\boldsymbol{x}) = f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{K}{2}}|\mathbf{C}_{\boldsymbol{X}}|^{\frac{1}{2}}} \exp -\frac{1}{2}(\boldsymbol{x}-\boldsymbol{m}_{\boldsymbol{X}})^T \mathbf{C}_{\boldsymbol{X}}^{-1}(\boldsymbol{x}-\boldsymbol{m}_{\boldsymbol{X}})$$

- **Bivariate Gaussian Density Function (K=2)**
  - For a 2-dimensional Gaussian random vector $\boldsymbol{x} = [x_1, x_2]^T$

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{2\pi|\mathbf{C}_{\boldsymbol{X}}|^{\frac{1}{2}}} \exp -\frac{1}{2}(\boldsymbol{x}-\boldsymbol{m}_{\boldsymbol{X}})^T \mathbf{C}_{\boldsymbol{X}}^{-1}(\boldsymbol{x}-\boldsymbol{m}_{\boldsymbol{X}})$$

$$\boldsymbol{m}_{\boldsymbol{X}} = [m_1 \quad m_2]^T$$

$$\rho = \frac{c_{12}}{\sigma_1\sigma_2}$$

$$\mathbf{C}_{\mathrm{X}} = \begin{bmatrix} {\sigma_1}^2 & c_{12} \\ c_{21} & {\sigma_2}^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

$$\mathbf{C}_{\boldsymbol{X}}^{-1} = \frac{1}{1-\rho^2}\begin{bmatrix} \frac{1}{\sigma_1^2} & -\frac{\rho}{\sigma_1\sigma_2} \\ -\frac{\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} \end{bmatrix}$$

8) Problem 7.7 (k-dim. Gaussian)

7.7 The mean vector and covariance matrix for a Gaussian random vector $\boldsymbol{X}$ are given by

$$\mathbf{m}_{\boldsymbol{X}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad C_{\boldsymbol{X}} = \begin{bmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{bmatrix}$$

(b) What is the correlation coefficient $\rho_{X_1 X_2}$?

$$\rho_{X_1 X_2} = \frac{E\left[(X_1 - m_{X_1})(X_2 - m_{X_2})\right]}{\sigma_{X_1} \sigma_{X_2}} = \frac{-\frac{1}{3}}{\sqrt{\frac{2}{3}} \cdot \sqrt{\frac{2}{3}}} = -\frac{1}{2}$$

8) Problem 7.7 (k-dim. Gaussian)

7.7 The mean vector and covariance matrix for a Gaussian random vector $\boldsymbol{X}$ are given by

$$\mathbf{m}_{\boldsymbol{X}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad C_{\boldsymbol{X}} = \begin{bmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{bmatrix}$$

(c) Invert the covariance matrix and write an explicit expression for the Gaussian density function for $\boldsymbol{X}$.

$$|C_{\boldsymbol{X}}| = \frac{2}{3} \cdot \frac{2}{3} - \left(-\frac{1}{3}\right) \cdot \left(-\frac{1}{3}\right) = \frac{1}{3} \qquad C_{\boldsymbol{X}}^{-1} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$f_{\underline{X}}(\underline{x}) = \frac{1}{2\pi |C_X|^{1/2}} e^{-\frac{1}{2}(\underline{x} - \underline{m}_x)^T \underline{C}_x^{-1} (\underline{x} - \underline{m}_x)}$$

→ 역행렬 계산 필요

10) Problem 6.24 (Parameter Estimation)

Confidence Intervals

6.24 Five hundred observations of a random variable X with variance $\sigma_X^2 = 25$ are taken. The sample mean based on 500 samples is computed to be $M_{500} = 3.25$. Find 95% and 98% confidence intervals for this estimate.

The variance of the sample mean is

$$\sigma_{500}^2 = \frac{25}{500} = 0.05 \longrightarrow \sigma_{500} = 0.2236$$

For the 95% CI we can use the value

$z = 1.96$ from the table in Fig. 6.6

The upper and lower limits are $3.25 \pm (1.96)(0.2236)$

The CI is therefore $(2.81, 3.69)$

10) Problem 6.24 (Parameter Estimation)

Confidence Intervals

6.24 Five hundred observations of a random variable X with variance $\sigma_X^2 = 25$ are taken. The sample mean based on 500 samples is computed to be $M_{500} = 3.25$. Find 95% and 98% confidence intervals for this estimate.

For the 98% CI, we use the table of Q functions to find $Q(2.325) = 0.01$, which is half the probability of the critical region. With $z = 2.325$ the computation is $3.25 \pm (2.325)(0.2236)$

or $(2.73, 3.77)$