[과제 #3]

20213064_김종민

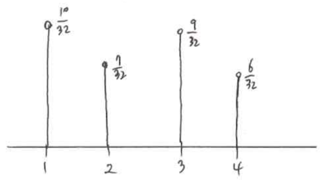
Assignment #3

Problems 5.3 Consider the joint PMF specified below.

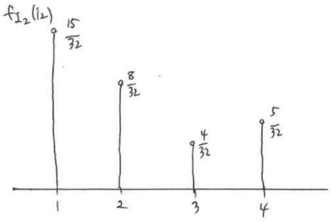
$f_{I_1 I_2}[i_1, i_2]$ ($i_2 \downarrow$, $i_1 \rightarrow$)	1	2	3	4
1	6/32	2/32	4/32	3/32
2	2/32	3/32	2/32	1/32
3	1/32	1/32	2/32	0
4	1/32	1/32	1/32	2/32

(a) Determine the marginal PMFs $f_{I_1}(i_1)$ and $f_{I_2}(i_2)$.

↳ $f_{I_1}(i_1)$

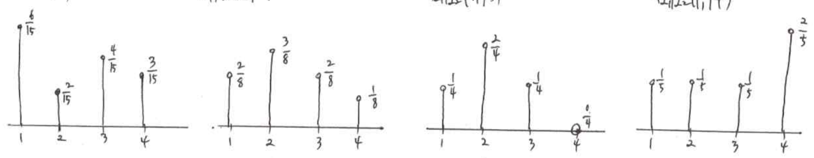


$f_{I_2}(i_2)$



(b) Determine the conditional PMF $f_{I_1|I_2}(i_1|i_2)$.

↳ $f_{I_1|I_2}(i_1|1)$ $f_{I_1|I_2}(i_1|2)$ $f_{I_1|I_2}(i_1|3)$ $f_{I_1|I_2}(i_1|4)$



(c) Find the CDF $F_{I_1, I_2}[i_1, i_2]$

↳

$F_{I_1 I_2}[i_1, i_2]$	1	2	3	4
1	6/32	8/32	12/32	15/32
2	8/32	13/32	19/32	23/32
3	9/32	15/32	23/32	27/32
4	10/32	17/32	26/32	32/32

$\boxed{\text{Problems 5-9}}$ The joint PDF $f_{X_1X_2}(x_1,x_2)$ of two random variables X_1 and X_2 is given by

$$f_{X_1,X_2}(x_1,x_2)=\begin{cases} C(4-x_1x_2) & 0\le x_1\le 4,\ 0\le x_2\le 1\\ 0, & \text{otherwise.}\end{cases}$$

(a) Find C to make this a valid PDF.

↳ $$\int_0^4\int_0^1 C(4-x_1x_2)\,dx_2\,dx_1 = C\int_0^4\int_0^1(4-x_1x_2)\,dx_2\,dx_1 = C\int_0^4\left[4x_2 - x_1\cdot\frac{1}{2}x_2^2\right]_0^1 = C\int_0^4\left(4-\frac{x_1}{2}\right)dx_1$$

$$= C\left[4x_1 - \frac{x_1^2}{4}\right]_0^4 = 12C=1 \qquad \therefore C=\frac{1}{12}$$

(b) Find the marginal density functions of X_1 and X_2. Clearly define the ranges of values they take.

↳

$$f_{X_1}(x_1)=\int_0^1\frac{1}{12}(4-x_1x_2)\,dx_2=\frac{1}{12}\int_0^1(4-x_1x_2)\,dx_2=\frac{1}{12}\left[4x_2-\frac{x_1}{2}x_2^2\right]_0^1=\frac{1}{12}\left(4-\frac{x_1}{2}\right)=\frac{1}{3}-\frac{x_1}{24}$$

$$f_{X_2}(x_2)=\int_0^4\frac{1}{12}(4-x_1x_2)\,dx_1=\frac{1}{12}\int_0^4(4-x_1x_2)\,dx_1=\frac{1}{12}\left[4x_1-\frac{x_2}{2}x_1^2\right]_0^4=\frac{1}{12}(16-8x_2)=\frac{4}{3}-\frac{2}{3}x_2$$

$$\therefore f_{X_1}(x_1)=\frac{1}{3}-\frac{x_1}{24}, \qquad f_{X_2}(x_2)=\frac{4}{3}-\frac{2}{3}x_2$$

(c) Are the random variables independent?

↳ Two random variables X_1 and X_2 are defined to be independent if their joint PDF is the product of the two marginals: $f_{X_1X_2}(x_1,x_2)=f_{X_1}(x_1)\cdot f_{X_2}(x_2)$

$$f_{X_1X_2}(x_1,x_2)=\frac{1}{12}(4-x_1x_2)$$

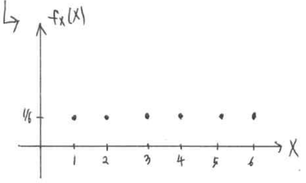
$$f_{X_1}(x_1)\cdot f_{X_2}(x_2)=\left(\frac{1}{3}-\frac{x_1}{24}\right)\left(\frac{4}{3}-\frac{2}{3}x_2\right)$$

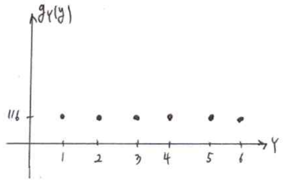
$$f_{X_1X_2}(x_1,x_2)\neq f_{X_1}(x_1)\cdot f_{X_2}(x_2)$$

$\therefore$ random variables are not independent.

problems 5.41 Let X and Y be the number shown on each of two dice and let Z be the sum ($Z = X+Y$). Assume that X and Y are independent and each is uniformly distributed over the integers 1 through 6. Using discrete convolution, show that the PMF for Z has a triangular shape. compare it to Fig 3.2 of Chapter 3.

↳





$$f_X(x) * g_Y(y) = \sum_{n=1}^{6} f_X(n) \cdot g_Y(k-n)$$

($f_X(n)$에서의 n은 X의 확률변수 값, $g_Y(k-n)$에서의 k-n은 Y의 확률변수 값)

(1) $k=2$일때, $\sum_{n=1}^{6} f_X(n) \cdot g_Y(2-n) = f_X(1) \cdot g_Y(1) + f_X(2) \cdot g_Y(0) + f_X(3) \cdot g_Y(-1) + \cdots + f_X(6) \cdot g_Y(-4) = \frac{1}{36}$

$Z = (X+Y) = \{(1+1)\} = 2$

(2) $k=3$일때, $\sum_{n=1}^{6} f_X(n) \cdot g_Y(3-n) = f_X(1) \cdot g_Y(2) + f_X(2) \cdot g_Y(1) + f_X(3) \cdot g_Y(0) + \cdots + f_X(6) \cdot g_Y(-3) = \frac{2}{36}$

$Z = (X+Y) = \{(1+2), (2+1)\} = 3$.

(3) $k=4$일때, $\sum_{n=1}^{6} f_X(n) \cdot g_Y(4-n) = f_X(1) \cdot g_Y(3) + f_X(2) \cdot g_Y(2) + f_X(3) \cdot g_Y(1) + f_X(4) \cdot g_Y(0) + \cdots + f_X(6) \cdot g_Y(-2) = \frac{3}{36}$

$Z = (X+Y) = \{(1,3), (2,2), (3,1)\} = 4$

(4) $k=5$일때, $\sum_{n=1}^{6} f_X(n) \cdot g_Y(5-n) = f_X(1) \cdot g_Y(4) + f_X(2) \cdot g_Y(3) + f_X(3) \cdot g_Y(2) + f_X(4) \cdot g_Y(1) + f_X(5) \cdot g_Y(0) + f_X(6) \cdot g_Y(-1) = \frac{4}{36}$

$Z = (X+Y) = \{(1,4), (2,3), (3,2), (4,1)\} = 5$.

(5) $k=6$일때, $\sum_{n=1}^{6} f_X(n) \cdot g_Y(6-n) = f_X(1) \cdot g_Y(5) + f_X(2) \cdot g_Y(4) + f_X(3) \cdot g_Y(3) + f_X(4) \cdot g_Y(2) + f_X(5) \cdot g_Y(1) + f_X(6) \cdot g_Y(0) = \frac{5}{36}$

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

(6) $k=7$일때, $\sum_{n=1}^{6} f_X(n) \cdot g_Y(7-n) = f_X(1) \cdot g_Y(6) + f_X(2) \cdot g_Y(5) + f_X(3) \cdot g_Y(4) + f_X(4) \cdot g_Y(3) + f_X(5) \cdot g_Y(2) + f_X(6) \cdot g_Y(1) = \frac{6}{36}$

$Z = (X+Y) = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\} = 7$.

(7) $k=8$일때, $\sum_{n=1}^{6} f_X(n) \cdot g_Y(8-n) = f_X(1) \cdot g_Y(7) + f_X(2) \cdot g_Y(6) + f_X(3) \cdot g_Y(5) + f_X(4) \cdot g_Y(4) + f_X(5) \cdot g_Y(3) + f_X(6) \cdot g_Y(2) = \frac{5}{36}$

$Z = (X+Y) = \{(2,6), (3,5), (4,4), (5,3), (6,2)\} = 8$

(8) $k=9$일때 $\sum_{n=1}^{6} f_X(n) \cdot g_Y(9-n) = f_X(1) \cdot g_Y(8) + f_X(2) \cdot g_Y(7) + f_X(3) \cdot g_Y(6) + f_X(4) \cdot g_Y(5) + f_X(5) \cdot g_Y(4) + f_X(6) \cdot g_Y(3) = \frac{4}{36}$

$Z = (X+Y) = \{(3,6), (4,5), (5,4), (6,3)\} = 9$

(9) k=10일때, $\sum_{n=1}^{6} f_X(n)\cdot g_Y(10-n) = f_X(1)\cdot g_Y(9) + f_X(2)\cdot g_Y(8) + f_X(3)\cdot g_Y(7) + f_X(4)\cdot g_Y(6) + f_X(5)\cdot g_Y(5) + f_X(6)\cdot g_Y(4) = \frac{3}{36}$

$Z = (X+Y) = \{(4,6), (5,5), (6,4)\} = 10.$

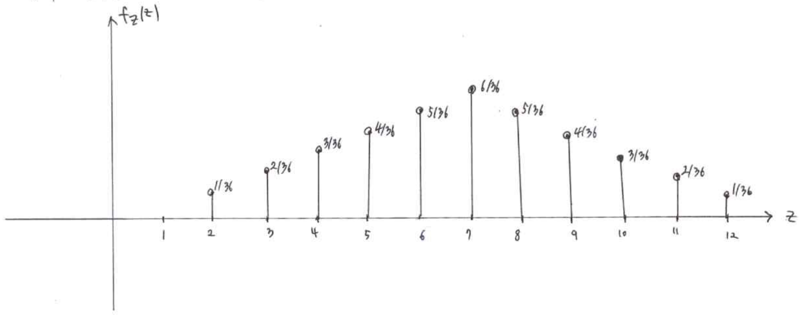
(10) k=11일때, $\sum_{n=1}^{6} f_X(n)\cdot g_Y(11-n) = f_X(1)\cdot g_Y(10) + f_X(2)\cdot g_Y(9) + f_X(3)\cdot g_Y(8) + f_X(4)\cdot g_Y(7) + f_X(5)\cdot g_Y(6) + f_X(6)\cdot g_Y(5) = \frac{2}{36}$

$Z = (X+Y) = \{(5,6), (6,5)\} = 11$

(11) k=12일때, $\sum_{n=1}^{6} f_X(n)\cdot g_Y(12-n) = f_X(1)\cdot g_Y(11) + f_X(2)\cdot g_Y(10) + f_X(3)\cdot g_Y(9) + f_X(4)\cdot g_Y(8) + f_X(5)\cdot g_Y(7) + f_X(6)\cdot g_Y(6) = \frac{1}{36}$

$Z = (X+Y) = (6,6) = 12.$

위의 결과들을 바탕으로 Z의 PMF를 그리면 아래와 같습니다.



∴ convolution을 사용하여 Z의 PMF를 구한 결과 Z의 PMF는 삼각형이라는 것을 확인했습니다.

$\boxed{\text{problems 5.45}}$ The sum of two independent random variables X_1 and X_2 is given by

$$X = X_1 + X_2$$

where X_1 is an exponential random variable with parameter $\lambda=2$, and X_2 is another exponential random variable with parameter $\lambda=3$.

(a) Find the mean and variance of X.

↳ $f_{X_1}(x_1) = 2e^{-2x_1}$

$E(X_1) = \int_0^{\infty} x_1 \cdot 2e^{-2x_1}\, dx_1$ cf) $\int_a^b uv'\,dx = [uv]_a^b - \int_a^b u'v\,dx$

$u = x_1 \qquad v = -e^{-2x_1}$

$u' = 1 \qquad v' = 2e^{-2x_1}$

$E(X_1) = \int_0^{\infty} x_1 \cdot 2e^{-2x_1}\, dx_1 = [x_1 \cdot (-e^{-2x_1})]_0^{\infty} - \int_0^{\infty} 1\cdot(-e^{-2x_1})\, dx_1 = 0 + \int_0^{\infty} e^{-2x_1}\, dx_1 = \frac{1}{2}$

$E(X_1^2) = \int_0^{\infty} x_1^2 \cdot 2e^{-2x_1}\, dx_1$

$u = x_1^2 \qquad v = -e^{-2x_1}$

$u' = 2x_1 \qquad v' = 2e^{-2x_1}$

$$E(X_1^2) = \int_0^\infty x_1^2 \cdot 2e^{-2x_1}\,dx_1 = \left[x_1^2 \cdot (-e^{-2x_1})\right]_0^\infty - \int_0^\infty 2x_1 \cdot (-e^{-2x_1})\,dx_1 = 0 + \int_0^\infty 2x_1 \cdot e^{-2x_1}\,dx_1 = \frac{1}{2}$$

$$V(X_1) = E(X_1^2) - (E(X_1))^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

$$f_{X_2}(x_2) = 3e^{-3x_2}$$

$$E(X_2) = \int_0^\infty x_2 \cdot 3e^{-3x_2}\,dx_2$$

$u = x_2 \qquad v = -e^{-3x_2}$

$u' = 1 \qquad v' = 3e^{-3x_2}$

$$E(X_2) = \int_0^\infty x_2 \cdot 3e^{-3x_2}\,dx_2 = \left[x_2 \cdot (-e^{-3x_2})\right]_0^\infty - \int_0^\infty 1 \cdot (-e^{-3x_2})\,dx_2 = 0 + \int_0^\infty e^{-3x_2}\,dx_2 = \frac{1}{3}$$

$$E(X_2^2) = \int_0^\infty x_2^2 \cdot 3e^{-3x_2}\,dx_2$$

$u = x_2^2 \qquad v = -e^{-3x_2}$

$u' = 2x_2 \qquad v' = 3e^{-3x_2}$

$$E(X_2^2) = \int_0^\infty x_2^2 \cdot 3e^{-3x_2}\,dx_2 = \left[x_2^2 \cdot (-e^{-3x_2})\right]_0^\infty - \int_0^\infty 2x_2 \cdot (-e^{-3x_2})\,dx_2 = 0 + \frac{2}{3}\int_0^\infty x_2 \cdot 3e^{-3x_2}\,dx_2 = \frac{2}{9}$$

$$V(X_2) = E(X_2^2) - (E(X_2))^2 = \frac{2}{9} - \frac{1}{9} = \frac{1}{9}$$

$$E(X) = E(X_1 + X_2) = E(X_1) + E(X_2) = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}$$

$$V(X) = V(X_1 + X_2) = V(X_1) + V(X_2) = \frac{1}{4} + \frac{1}{9} = \frac{13}{36}$$

$$\therefore E(X) = \frac{5}{6}, \quad V(X) = \frac{13}{36}$$

(b) Determine the PDF of X.

↳ n개의 독립 랜덤변수들의 합에 대한 결합 밀도함수는, 밀도함수들의 n중 합성곱이 됩니다.

$$f_Y = f_{X_1} * f_{X_2} * \cdots * f_{X_n}$$

$$f_X = f_{X_1} * f_{X_2}$$

$$f_X(x) = \int_0^x 2e^{-2t} \cdot 3e^{-3(x-t)}\,dt = 6\int_0^x e^{t-3x}\,dt = 6\left[e^{t-3x}\right]_0^x = 6(e^{-2x} - e^{-3x})$$

$$\therefore \text{PDF of } X \quad f_X(x) = 6(e^{-2x} - e^{-3x})$$

$\boxed{\text{problems } 6.24}$ 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.

↳ • 신뢰구간 설정 : 유의수준 α가 정해지면 z-분포 x축의 임계점 K가 결정됩니다.

신뢰구간 α가 95% 일때 K는 1.96, α가 98% 일때는 2.33입니다.

신뢰구간의 하한 $L = \bar{X} - K\sigma/\sqrt{n}$, 상한은 $U = \bar{X} + K\sigma/\sqrt{n}$ 입니다.

$\bar{X} = M_{500} = 3.25$, $n = 500$. $\sigma = 5$.

(1) 신뢰구간이 95% 일때 : $L = 3.25 - 1.96 \cdot 5/\sqrt{500}$, $U = 3.25 + 1.96 \cdot 5/\sqrt{500}$

$L = 2.812$ $\quad U = 3.688$

(2) 신뢰구간이 98% 일때 : $L = 3.25 - 2.33 \cdot 5/\sqrt{500}$, $U = 3.25 + 2.33 \cdot 5/\sqrt{500}$

$L = 2.729$ $\quad U = 3.771$

∴ 95% confidence intervals (2.812, 3.688), 98% confidence intervals (2.729, 3.771)

$\boxed{\text{problems } 7.5}$ 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).

↳ sample mean: $M_n = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{4}\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\} = \frac{1}{4}\begin{bmatrix} 4 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

sample correlation matrix : $R_n = \frac{1}{n}\sum_{i=1}^{n} X_i X_i^T$

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

$$= \frac{1}{4}\begin{bmatrix} 6 & 3 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} & \frac{3}{4} \\ \frac{3}{4} & \frac{3}{2} \end{bmatrix}$$

sample covariance matrix : $C_n = \frac{n}{n-1}\left(R_n - M_n M_n^T\right)$

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

∴ sample mean $= \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, sample correlation matrix $= \begin{bmatrix} \frac{3}{2} & \frac{3}{4} \\ \frac{3}{4} & \frac{3}{2} \end{bmatrix}$, sample covariance matrix $= \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$

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

$$m_X = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad R_X = \begin{bmatrix} 4 & -1 \\ -1 & 3 \end{bmatrix}$$

(a) compute the covariance matrix C_X.

↳ $C_X = R_X - m_X m_X^T$

$$C_X = \begin{bmatrix} 4 & -1 \\ -1 & 3 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & -1 \\ -1 & 3 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -2 & 2 \end{bmatrix}$$

$$\therefore C_X = \begin{bmatrix} 3 & -2 \\ -2 & 2 \end{bmatrix}$$

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

↳ correlation coefficient $\rho = \dfrac{\mathrm{COV}[X_1, X_2]}{\sigma_1 \sigma_2}$

$$C_X = \begin{bmatrix} \sigma_1^2 & C_{12} \\ C_{21} & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -2 & 2 \end{bmatrix}$$

$\sigma_{X_1}^2 = 3$, $\sigma_{X_1} = \sqrt{3}$

$\sigma_{X_2}^2 = 2$, $\sigma_{X_2} = \sqrt{2}$

$\rho_{X_1X_2}\sigma_{X_1}\sigma_{X_2} = -2$

$\rho_{X_1X_2}\sqrt{3}\cdot\sqrt{2} = -2$

$\rho_{X_1X_2} = \dfrac{-2}{\sqrt{6}} = \dfrac{-\sqrt{6}}{3}$

$$\therefore \rho_{X_1X_2} = -\frac{\sqrt{6}}{3}$$

(c) Invert the covariance matrix and write an explicit expression for the Gaussian density function for X.

↳ $$C_X^{-1} = \frac{1}{1-\rho^2}\begin{bmatrix} \frac{1}{\sigma_1^2} & -\frac{\rho}{\sigma_1\sigma_2} \\ -\frac{\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} \end{bmatrix} = \frac{1}{1-\left(\frac{\sqrt{6}}{3}\right)^2}\begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{2} \end{bmatrix} = 3\begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & \frac{3}{2} \end{bmatrix}$$

이를 바탕으로 X에 대한 가우스 밀도 함수로 나타내면 아래와 같습니다.

$$f_X(x) = \frac{1}{(2\pi)^{\frac{n}{2}}|C_X|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-m_X)^T C_X^{-1}(x-m_X)}$$

$|C_X| = \sigma_1^2\sigma_2^2(1-\rho^2)$

$|C_X| = 3\cdot 2\left(1-\frac{6}{9}\right) = 2$

$$f_X(x) = \frac{1}{(2\pi)^{\frac{2}{2}} \, |2|^{\frac{1}{2}}} e^{-\frac{1}{2}\left(x - \begin{bmatrix}1\\1\end{bmatrix}\right)^T \begin{bmatrix}1 & 1\\1 & \frac{3}{2}\end{bmatrix}\left(x - \begin{bmatrix}1\\1\end{bmatrix}\right)}$$

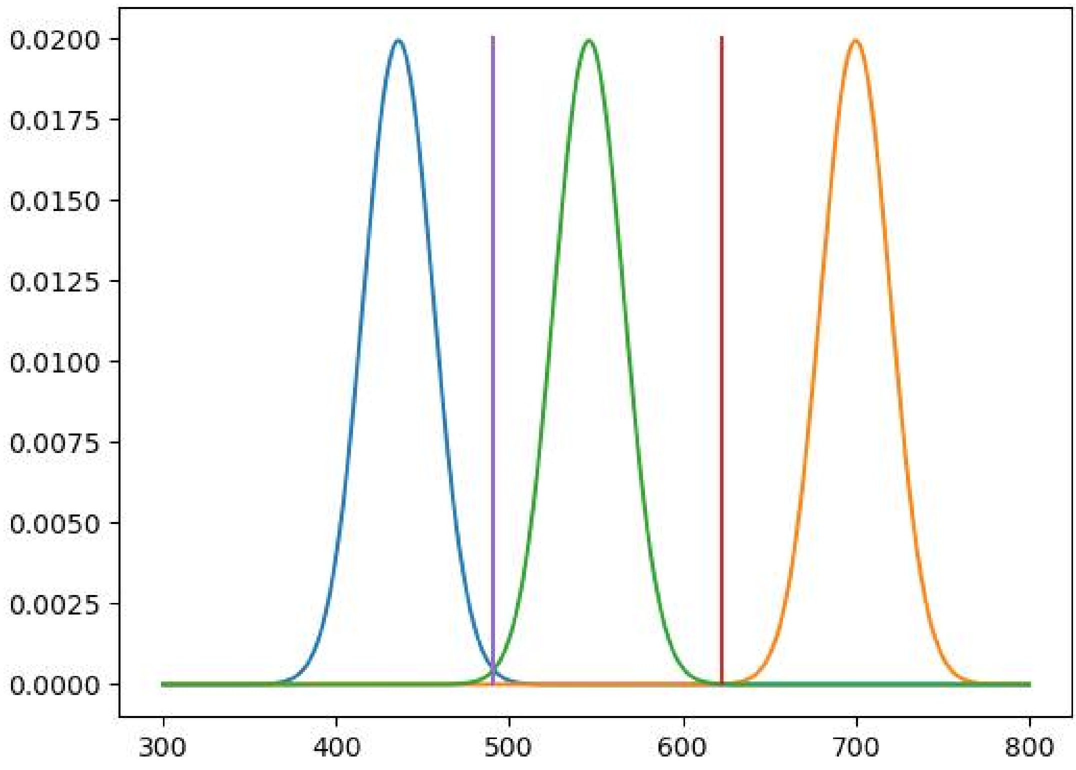
$$f_X(x) = \frac{1}{2\sqrt{2}\,\pi} e^{-\frac{1}{2}\left(x - \begin{bmatrix}1\\1\end{bmatrix}\right)^T \begin{bmatrix}1 & 1\\1 & \frac{3}{2}\end{bmatrix}\left(x - \begin{bmatrix}1\\1\end{bmatrix}\right)}$$

$$\therefore C_X^{-1} = \begin{bmatrix}1 & 1\\1 & \frac{3}{2}\end{bmatrix}, \quad f_X(x) = \frac{1}{2\sqrt{2}\,\pi} e^{-\frac{1}{2}\left(x - \begin{bmatrix}1\\1\end{bmatrix}\right)^T \begin{bmatrix}1 & 1\\1 & \frac{3}{2}\end{bmatrix}\left(x - \begin{bmatrix}1\\1\end{bmatrix}\right)}$$

실습 7. 분류 에러 분석 실습

In [8]:
```
# 분류 에러 분석 실습
import math
import numpy as np
import matplotlib.pyplot as plt
import scipy.stats as st
H1_mu, H1_sig = 700, math.sqrt(400)
H2_mu, H2_sig = 546, math.sqrt(400)
H3_mu, H3_sig = 436, math.sqrt(400)
th12 = (H1_mu + H2_mu) / 2
th23 = (H2_mu + H3_mu) / 2
x = np.linspace (300, 800, 1000)
h1 = st.norm.pdf(x, H1_mu, H1_sig)
h2 = st.norm.pdf(x, H2_mu, H2_sig)
h3 = st.norm.pdf(x, H3_mu, H3_sig)
plt.plot(x, h3)
plt.plot(x, h1)
plt.plot(x, h2)
plt.plot([th12, th12], [0, 0.02])
plt.plot([th23, th23], [0, 0.02])
plt.show()

e3 = st.norm.sf((th23-H3_mu) / H3_sig, 0, 1)
e1 = st.norm.sf(-(th12 - H1_mu) / H1_sig, 0, 1)
e2Left = st.norm.sf(-(th23-H2_mu) / H2_sig, 0, 1)
e2Right = st.norm.sf((th12 - H2_mu) / H2_sig, 0, 1)
e2 = e2Left + e2Right
print(th12, th23)
print(e3, e1, e2)
```




```
623.0 491.0
0.0029797632350545555 5.905891241892237 4e-05 0.0030388221474734775
```

실습 8. 피어슨의 적합도 검정

In [9]:
```
# 피어슨의 적합도 검정
import pandas as pd
import scipy.stats as stats
f_ob1 = [99, 101, 102, 97, 101, 100]
f_ob2 = [117, 119, 120, 115, 119, 10]
f_exp = [100, 100, 100, 100, 100, 100]
print("1: ", stats.chisquare(f_ob1, f_exp=f_exp))
print("2: ", stats.chisquare(f_ob2, f_exp=f_exp))
```

```
1:  Power_divergenceResult(statistic=0.16, pvalue=0.9994854883416188)
2:  Power_divergenceResult(statistic=97.36, pvalue=1.9021569776120922e-19)
```