20170243 십재윤

$\boxed{\#1}$ [Marginalization]

(X,Y) \ Z	1	2	3
(1, 1)	0.2	0.2	0.1
(1, 2)	0.02	0.1	0.05
(2, 1)	0.03	0.05	0.03
(2, 2)	0.1	0.07	0.05

(a) marginal distributions $p(x,y)$ and $p(z)$

(X,Y) \ Z	1	2	3	$p(X,Y)$
(1, 1)	0.2	0.2	0.1	0.5
(1, 2)	0.02	0.1	0.05	0.17
(2, 1)	0.03	0.05	0.03	0.11
(2, 2)	0.1	0.07	0.05	0.22
$p(Z)$	0.35	0.42	0.23	1

$p(x,y)$ >

$p(1,1)=0.5$, $p(1,2)=0.17$

$p(2,1)=0.11$, $p(2,2)=0.22$

$p(z)$ >

$p(1)=0.35$, $p(2)=0.42$, $p(3)=0.23$

(b) expectation $\mathbb{E}[X+Y^2]$ and $\mathbb{E}[Z]$

$$\mathbb{E}[X] = 1\cdot 0.67 + 2\cdot 0.33 = 0.67 + 0.66 = 1.33$$

$$\mathbb{E}[Y] = 1\cdot 0.61 + 2\cdot 0.39 = 0.61 + 0.78 = 1.39$$

$$\mathbb{E}[Y^2] = 1^2\cdot 0.61 + 2^2\cdot 0.39 = 0.61 + 1.56 = 2.17$$

$$\therefore \mathbb{E}[X+Y^2] = \mathbb{E}[X] + \mathbb{E}[Y^2] = 1.33 + 2.17 = 3.5$$

$$\mathbb{E}[Z] = 1\cdot 0.35 + 2\cdot 0.42 + 3\cdot 0.23 = 0.35 + 0.84 + 0.69 = 1.88$$

(c) conditional distributions $p(x \mid Y=1)$ and $p(z \mid X=1)$

$$p(x \mid Y=1) > \quad p(X=1 \mid Y=1) = \frac{p(X=1, Y=1)}{p(Y=1)} = \frac{0.5}{0.61} = 0.82$$

$$p(X=2 \mid Y=1) = \frac{p(X=2, Y=1)}{p(Y=1)} = \frac{0.11}{0.61} = 0.18$$

$$p(z \mid X=1) > \quad p(Z=1 \mid X=1) = \frac{p(Z=1, X=1)}{p(X=1)} = \frac{0.22}{0.67} = 0.33$$

$$p(Z=2 \mid X=1) = \frac{p(Z=2, X=1)}{p(X=1)} = \frac{0.3}{0.67} = 0.45$$

$$p(Z=3 \mid X=1) = \frac{p(Z=3, X=1)}{p(X=1)} = \frac{0.15}{0.67} = 0.22$$

(d) conditional expectations $\mathbb{E}[X \mid Y=1]$ and $\mathbb{E}[Z \mid X=1]$

$$\mathbb{E}[X \mid Y=1] = \sum_x x \cdot p(x \mid Y=1)$$

$$= 1 \cdot p(X=1 \mid Y=1) + 2 \cdot p(X=2 \mid Y=1)$$

$$= 1 \cdot 0.82 + 2 \cdot 0.18$$

$$= 0.82 + 0.36 = 1.18$$

$$\mathbb{E}[Z \mid X=1] = \sum_z z \cdot p(z \mid X=1)$$

$$= 1 \cdot p(Z=1 \mid X=1) + 2 \cdot p(Z=2 \mid X=1) + 3 \cdot p(Z=3 \mid X=1)$$

$$= 1 \cdot 0.33 + 2 \cdot 0.45 + 3 \cdot 0.22$$

$$= 0.33 + 0.9 + 0.66 = 1.89$$

$\boxed{\#2}$ [Expectation]

(a) $\mathbb{E}_X[X] = \mathbb{E}_Y[\mathbb{E}_X[X|Y]]$

discrete random variable >

$$\underset{\text{(RHS)}}{\underline{\mathbb{E}_Y[\mathbb{E}_X[X|Y]]}} = \sum_y \mathbb{E}_X[X|Y]\, p(Y=y)$$

$$= \sum_y \left(\sum_x x \cdot p(X=x|Y=y) \right) p(Y=y)$$

$$= \sum_y p(Y=y) \sum_x x \cdot \frac{p(X=x, Y=y)}{p(Y=y)}$$

$$= \sum_y \cancel{p(Y=y)} \cdot \frac{1}{\cancel{p(Y=y)}} \cdot \sum_x x \cdot p(X=x, Y=y)$$

$$= \sum_x \sum_y x \cdot p(X=x, Y=y)$$

$$= \sum_x x \left(\sum_y p(X=x, Y=y) \right)$$

$$= \sum_x x \cdot p(X=x) = \underset{\text{(LHS)}}{\underline{\mathbb{E}_X[X]}}$$

+ continuous random variable >

$$\mathbb{E}_Y[\mathbb{E}_X[X|Y]] = \int_{-\infty}^{\infty} \mathbb{E}_X[X|Y]\, f_Y(y)\, dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot f_{X|Y}(x|y)\, dx \cdot f_Y(y)\, dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot f_{X|Y}(x|y)\, f_Y(y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot f_{X,Y}(x,y)\, dy\, dx$$

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

$$= \int_{-\infty}^{\infty} x f_X(x)\, dx = \mathbb{E}_X[X]$$

(b) $\text{cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$

$$
\begin{aligned}
\text{cov}(X,Y) &= \mathbb{E}\big[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])\big] \\
&= \mathbb{E}\big[XY - X\mathbb{E}[Y] - Y\mathbb{E}[X] + \mathbb{E}[X]\mathbb{E}[Y]\big] \\
&= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] - \cancel{\mathbb{E}[X]\mathbb{E}[Y]} + \cancel{\mathbb{E}[X]\mathbb{E}[Y]} \\
&= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]
\end{aligned}
$$

(c) if $\text{cov}(X,Y)=0$, then X and Y are independent.

disprove > X : random variable $-1, 0, 1$ with each probability $\frac{1}{3}$

Y : random variable $Y = X^2$

$Y \backslash X$	-1	0	1	$P(Y=y)$
0	0	$\frac{1}{3}$	0	$\frac{1}{3}$
1	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{2}{3}$
$P(X=x)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1

$\mathbb{E}(X) = \mu_X = (-1)\cdot\frac{1}{3} + 0\cdot\frac{1}{3} + 1\cdot\frac{1}{3} = 0$

$\mathbb{E}(Y) = \mu_Y = 0\cdot\frac{1}{3} + 1\cdot\frac{2}{3} = \frac{2}{3}$

$\text{cov}(X,Y) = \mathbb{E}[XY] - \mu_X\mu_Y = \left\{(-1)\cdot\frac{1}{3} + 0\cdot\frac{1}{3} + 1\cdot\frac{1}{3}\right\} - 0\cdot\frac{2}{3} = 0$. (condition)

→ And, we want to find an example of $P(X=x, Y=y) \neq P(X=x)P(Y=y)$ because it means that X and Y are not independent.

Pick $X=-1, Y=0 \rightarrow P(X=-1, Y=0) = 0$ & $P(X=-1) = \frac{1}{3}$, $P(Y=0) = \frac{1}{3}$

Then, $\underbrace{P(X=-1, Y=0)}_{=0} \neq \underbrace{P(X=-1)\,P(Y=0)}_{=\frac{1}{3}\times\frac{1}{3}=\frac{1}{9}}$

Therefore, X and Y are not independent.

#3) [Mixture of Gaussian] "Bernoulli" → 0, 1

$\alpha \in [0,1]$, $Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \in \mathbb{R}^2$ s.t. $p(y|X=0) = N(y; \mu, \Sigma)$ and

$p(y|X=1) = N(y; \mu', \Sigma')$ where $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \in \mathbb{R}^2$, $\mu' = \begin{bmatrix} \mu_1' \\ \mu_2' \end{bmatrix} \in \mathbb{R}^2$,

$\Sigma = \begin{bmatrix} \Sigma_{1,1} & \Sigma_{1,2} \\ \Sigma_{2,1} & \Sigma_{2,2} \end{bmatrix} \in \mathbb{R}^4$, $\Sigma' = \begin{bmatrix} \Sigma_{1,1}' & \Sigma_{1,2}' \\ \Sigma_{2,1}' & \Sigma_{2,2}' \end{bmatrix} \in \mathbb{R}^4$

(a) distribution Y using $N(y; \mu, \Sigma)$ and $N(y; \mu', \Sigma')$

$$p(y) = p(X=0) \cdot p(y|X=0) + p(X=1) \cdot p(y|X=1)$$
$$= (1-\alpha) \cdot p(y|X=0) + \alpha \cdot p(y|X=1)$$

$$N\left(\begin{bmatrix} (1-\alpha)\mu_1 \\ (1-\alpha)\mu_2 \end{bmatrix}, (1-\alpha)^2 \Sigma\right) \quad N\left(\begin{bmatrix} \alpha\mu_1' \\ \alpha\mu_2' \end{bmatrix}, \alpha^2 \Sigma'\right)$$

$$\therefore N(y; (1-\alpha)\mu + \alpha\mu', (1-\alpha)^2\Sigma + \alpha^2\Sigma')$$

(b) set of α making X and Y independent to each other

We need to have

① $p(X=0)\, p(Y=y) = p(X=0, Y=y) = p(X=0)\, p(Y=y \,|\, X=0)$

② $p(X=1)\, p(Y=y) = p(X=1, Y=y) = p(X=1)\, p(Y=y \,|\, X=1)$

(i) If $\alpha = 0$ or $\alpha = 1$,

① $p(X=0, Y=y) = p(X=0)\, p(Y=y|X=0) = (1-\alpha)\, N(y; \mu, \Sigma)$

$p(X=0)\, p(Y=y) = (1-\alpha)^2 N(y; \mu, \Sigma) + \alpha(1-\alpha)\, N(y; \mu', \Sigma')$

$-\alpha(1-\alpha)\, N(y; \mu, \Sigma) + \alpha(1-\alpha)\, N(y; \mu', \Sigma') = 0 \rightarrow \alpha = 0, 1$

② $P(X=1, Y=y) = p(X=1)\, p(Y=y \mid X=1) = \alpha\, \mathcal{N}(y; \mu', \Sigma')$

$p(X=1)\, p(Y=y) = \alpha(1-\alpha)\, \mathcal{N}(y; \mu, \Sigma) + \alpha^2 \mathcal{N}(y; \mu', \Sigma')$

$\alpha(1-\alpha)\, \mathcal{N}(y; \mu, \Sigma) + \alpha(\alpha-1)\mathcal{N}(y; \mu', \Sigma') = 0 \rightarrow \alpha = 0, 1$

So, X and Y are independent.

(ii) If $0 < \alpha < 1$, $P(Y=y) = P(Y=y \mid X=0) = P(Y=y \mid X=1)$

$\mathcal{N}\left((1-\alpha)\mu + \alpha\mu', (1-\alpha^2)\Sigma + \alpha^2\Sigma'\right) \quad \mathcal{N}(\mu, \Sigma) \quad \mathcal{N}(\mu', \Sigma')$

According to $\mu, \mu', \Sigma, \Sigma'$, it is different. $\therefore$ $\alpha \in \{0, 1\}$

(c) marginal distributions of Y_1 and Y_2

$Y_1 \sim \mathcal{N}\left((1-\alpha)\mu_1 + \alpha\mu_1', \quad (1-\alpha)^2\Sigma_{1,1} + \alpha^2\Sigma_{1,1}'\right)$

$Y_2 \sim \mathcal{N}\left((1-\alpha)\mu_2 + \alpha\mu_2', \quad (1-\alpha)^2\Sigma_{2,2} + \alpha^2\Sigma_{2,2}'\right)$

(d) means of Y_1 and Y_2

Mean of $Y_1 = (1-\alpha)\mu_1 + \alpha\mu_1'$

mean of $Y_2 = (1-\alpha)\mu_2 + \alpha\mu_2'$

(e) variances of Y_1 and Y_2

variance of $Y_1 = (1-\alpha)^2\Sigma_{1,1} + \alpha^2\Sigma_{1,1}'$

variance of $Y_2 = (1-\alpha)^2\Sigma_{2,2} + \alpha^2\Sigma_{2,2}'$

$\boxed{\#4}$ [Bayes' theorem]

bag

$\boxed{1}$: 4 mango, 2 apple

$\boxed{2}$: 4 mango, 4 apple

biased coin

H = 0.6

T = 0.4

H $\longrightarrow$ $\boxed{1}$

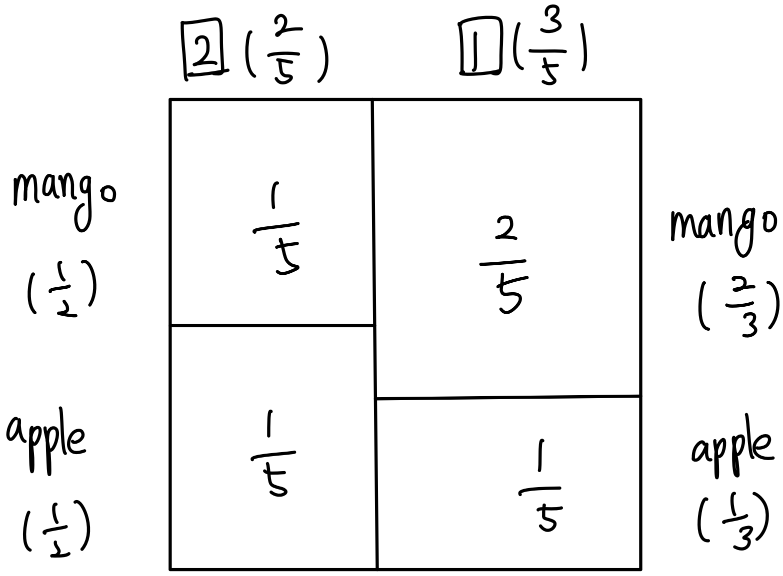
T $\longrightarrow$ $\boxed{2}$

friend flip coin (I cannot see the result)

→ pick a fruit at random from the corresponding bag

→ present me a mango

→ probability that the mango was picked from bag 2?



$$\therefore p(\boxed{2} \mid \text{mango}) = \frac{p(\text{mango} \mid \boxed{2}) \cdot p(\boxed{2})}{p(\text{mango})} = \frac{p(\text{mango} \mid \boxed{2}) \cdot p(\boxed{2})}{p(\text{mango} \mid \boxed{2}) \cdot p(\boxed{2}) + p(\text{mango} \mid \boxed{1}) \cdot p(\boxed{1})}$$

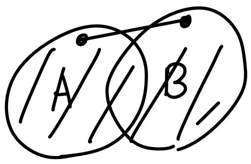
$$= \frac{\frac{1}{5}}{\frac{1}{5} + \frac{2}{5}}$$

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

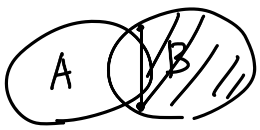
#5 [Basic understanding of convexity]

(a) intersection of any two convex sets → convex T

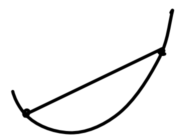
(b) union of any two convex sets → convex F



(c) difference of a convex set A from another convex set B → convex F



convex function : $f(tx+(1-t)y) \preceq tf(x) + (1-t)f(y)$



(d) sum of any two convex functions → convex T

(e) difference of any two convex functions → convex F

$f_1(x) = 0$, $f_2(x) = |x|$

(f) product of any two convex functions → convex F

$f_1(x) = 1-x$, $f_2(x) = 1+x$

(g) maximum of any two convex functions → convex T

#6 [Basic understanding of calculus]

$f(x) = x^3 - 2x^2 + x$

(a) stationary points

$$\frac{\partial f(x)}{\partial x} = 3x^2 - 4x + 1 = 0$$

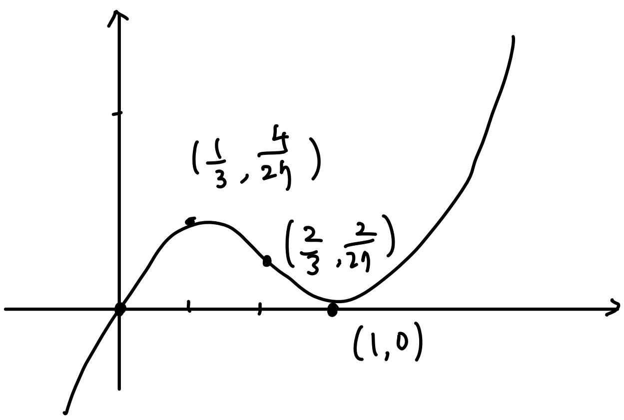
$$\begin{matrix} 3 & -1 \\ 1 & -1 \end{matrix}$$

$(3x-1)(x-1) = 0 \rightarrow x = \frac{1}{3}$ or 1

i) $x = \frac{1}{3}$, $f(\frac{1}{3}) = \frac{1}{27} - 2 \cdot \frac{1}{9} + \frac{1}{3} = \frac{1-6+9}{27} = \frac{4}{27} \rightarrow (\frac{1}{3}, \frac{4}{27})$

ii) $x = 1$, $f(1) = 1 - 2 \cdot 1 + 1 = 0 \rightarrow (1, 0)$

(b) largest interval of x on which $f(x)$ is convex



$f'(x) = 3x^2 - 4x + 1$

$\rightarrow f''(x) = 6x - 4 \geq 0$

$\rightarrow x \geq \frac{2}{3}$ (convex)

$\rightarrow f(\frac{2}{3}) = \frac{8}{27} - 2 \cdot \frac{4}{9} + \frac{2}{3} = \frac{8-24+18}{27} = \frac{2}{27}$

$\therefore [\frac{2}{3}, \infty)$

(c) minimize $f(x)$ on $[-1, 3]$

$f(-1) = (-1)^3 - 2(-1)^2 + (-1) = -1 - 2 - 1 = -4$

$f(3) = 3^3 - 2 \cdot 3^2 + 3 = 27 - 18 + 3 = 12$

$f(\frac{1}{3}) = (\frac{1}{3})^3 - 2 \cdot (\frac{1}{3})^2 + \frac{1}{3} = \frac{1}{27} - \frac{2}{9} + \frac{1}{3} = \frac{1-6+9}{27} = \frac{4}{27}$

$\therefore -4$

(d) minimize $f(x)$ on $[\frac{1}{2}, 3]$

$$f(\tfrac{1}{2}) = (\tfrac{1}{2})^3 - 2\cdot(\tfrac{1}{2})^2 + \tfrac{1}{2} = \tfrac{1}{8} - \tfrac{2}{4} + \tfrac{1}{2} = \tfrac{1}{8}$$

$$f(3) = 12$$

$$f(1) = 0$$

$\therefore$ 0