HOME WORK 5

Gawon Kim
305 − 186 − 572

①

ⓐ $P(A \mid C=0) = P(A=1 \mid C=0) = \dfrac{P(A=1 \text{ and } C=0)}{P(C=0)}$

$= \dfrac{0.224 + 0.056}{0.224 + 0.056 + 0.024 + 0.096} = 0.7$

$P(B \mid C=0) = P(B=1 \mid C=0) = \dfrac{P(A=1 \text{ and } C=0)}{P(C=0)}$

$= \dfrac{0.024 + 0.056}{0.224 + 0.056 + 0.024 + 0.096} = 0.2$

$P(A, B \mid C=0) = P(A \mid C=0)\,P(B \mid C=0)$

$= 0.7 \cdot 0.2 = 0.14$

ⓑ $P(A \mid C=1) = P(A=1 \mid C=1) = \dfrac{P(A=1 \text{ and } C=1)}{P(C=1)}$

$= \dfrac{0.27 + 0.03}{0.27 + 0.03 + 0.27 + 0.03} = \dfrac{0.3}{0.6} = 0.5$

$P(B \mid C=1) = P(B=1 \mid C=1)$

$= \dfrac{0.03 + 0.03}{0.27 + 0.03 + 0.27 + 0.03} = \dfrac{0.06}{0.6} = 0.1$

$P(A, B \mid C=1) = 0.05$

ⓒ To satisfy conditional independent, the given information should satisfy $P(A=1 \text{ and } B=1 \mid C=0) = P(A \mid C=0)\,P(B \mid C=0)$

when $C=0$.

$P(A \mid C=0) \cdot P(B \mid C=0)$, its value is $\boxed{0.14}$

and $P(A=1 \text{ and } B=1 \mid C=0)$ is $\dfrac{0.056}{0.4} = \boxed{0.14}$

So they are same each other, so that when $C=0$

A an B are conditionally independent.

when $C=1$.

$P(A, B \mid C=1) = \boxed{0.05}$ from 1 − ⓑ

when $P(A=1 \text{ and } B=1 \mid C=1)$ the value is $\dfrac{0.03}{0.6}$

which is $\boxed{0.05}$

So they are same each other, so that when $C=1$

A an B are conditionally independent.

∴ A and B are conditionally independent

ⓓ $P(A) = 0.224 + 0.056 + 0.27 + 0.03 = \boxed{0.58}$

$P(B) = 0.024 + 0.056 + 0.03 + 0.03 = \boxed{0.14}$

$P(A, B) = 0.056 \times 0.03 = \boxed{0.00168}$

ⓔ $P(A, B) \neq P(A) \cdot P(B)$ → they are not equal so not independe.

②

$$f_{X,Y}(x,y) = \frac{\exp\left\{\frac{-1}{2(1-\rho_{XY}^2)}\left[\left(\frac{x-m_1}{\sigma_1}\right)^2 - 2\rho_{XY}\left(\frac{x-m_1}{\sigma_1}\right)\left(\frac{y-m_2}{\sigma_2}\right) + \left(\frac{y-m_2}{\sigma_2}\right)^2\right]\right\}}{2\pi\sigma_1\sigma_2\sqrt{1-\rho_{XY}^2}}.$$

$X \sim N(m_1, \sigma_1^2)$

$Y \sim N(m_2, \sigma_2^2)$

$z = [X, Y]$ Let $\Sigma^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$[X-\mu_1 \quad Y-\mu_2] \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} X-\mu_1 \\ Y-\mu_2 \end{bmatrix}$

$[a(X-\mu_1) + c(Y-\mu_2) \quad b(X-\mu_1) + d(Y-\mu_2)] \begin{bmatrix} X-\mu_1 \\ Y-\mu_2 \end{bmatrix}$

$a(X-\mu_1)^2 + c(Y-\mu_2)(X-\mu_1) + b(X-\mu_1)(Y-\mu_2) + d(Y-\mu_2)^2$

$= a(X-\mu_1)^2 + (c+b)(X-\mu_1)(Y-\mu_2) + d(Y-\mu_2)^2$

$= \dfrac{1}{(1-\rho_{XY}^2)}\left[a'(X-\mu_1)^2 + (c'+b')(X-\mu_1)(Y-\mu_2) + d'(Y-\mu_2)^2\right]$

$a' = \dfrac{1}{\sigma_1^2} \quad c'+b' = \dfrac{-2\rho_{XY}}{\sigma_1\sigma_2} \quad d' = \dfrac{1}{\sigma_2^2}$

By comparing the given ⓐ

$$\Sigma^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\sigma_1^2(1-\rho_{XY}^2)} & \dfrac{-1}{\sigma_1\sigma_2(1-\rho_{XY}^2)} \\ \dfrac{-1}{\sigma_1\sigma_2(1-\rho_{XY}^2)} & \dfrac{1}{\sigma_2^2(1-\rho_{XY}^2)} \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho_{XY}\sigma_1\sigma_2 \\ \rho_{XY}\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

$\mu = [\mu_1 \quad \mu_2]^T = [m_1 \quad m_2]^T$

③

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_X^2} + \frac{y^2}{\sigma_Y^2} - \frac{2\rho xy}{\sigma_X\sigma_Y}\right)\right].$$

ⓐ $f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx \quad -\infty < y < \infty$

$\left(\dfrac{y-\mu_y}{\sigma_y}\right)^2 + \left\{\left(\dfrac{y-\mu_y}{\sigma_x}\right)^2 - 2\rho\left(\dfrac{x-\mu_x}{\sigma_x}\right)\left(\dfrac{y-\mu_y}{\sigma_y}\right) + \rho^2\left(\dfrac{x-\mu_x}{\sigma_x}\right)^2\right\} - \rho^2\left(\dfrac{y-\mu_y}{\sigma_y}\right)^2$

$= (1-\rho^2)\left(\dfrac{x-\mu_x}{\sigma_x}\right)^2 + \mu^2$ where $u = \left(\dfrac{x-\mu_x}{\sigma_x}\right) - \rho\left(\dfrac{y-\mu_y}{\sigma_y}\right)$

$\dfrac{\partial u}{\partial x} = \dfrac{1}{\sigma x}$

$f_Y(y) = \int_{-\infty}^{\infty} \dfrac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{-\dfrac{1}{2(1-\rho^2)}\left[(1-\rho^2)\left(\dfrac{y-\mu_y}{\sigma_y}\right)^2 + u^2\right]\right\}\sigma_x\,du$

$= \underbrace{\dfrac{1}{2\pi\sigma_y\sqrt{1-\rho^2}} \exp\left\{-\dfrac{1}{2}\left(\dfrac{y-\mu_y}{\sigma_y}\right)^2\right\} \int_{-\infty}^{\infty} \exp\left\{-\dfrac{1}{2(1-\rho^2)}u^2\right\}du}_{\alpha}$

$\int_{-\infty}^{\infty} \alpha\,du = 1$

So that $f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_y} \exp\left(-\frac{1}{2}\left(\frac{y-\cancel{\mu_y}}{\sigma_y}\right)^2\right)$

In this case the mean (μ_y) is 0

and variance of Y is σ_y^2

$\therefore N(0, \sigma_y^2)$

(b) $f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f(y)}$

$$= \frac{\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}-\frac{2\rho xy}{\sigma_x\sigma_y}\right)\right]}{\frac{1}{\sqrt{2\pi}\sigma_y}\exp\left(-\frac{1}{2}\frac{y^2}{\sigma_y^2}\right)}$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_x\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)\sigma_x^2}\left[x-\left(\mu_x+\rho\frac{\sigma_x}{\sigma_y}(y-\mu_y)\right)\right]^2\right\}$$

ie. $f_{X|Y}(x|y)$ is the PDF of

$$N\left(\mu_x + \rho\frac{\sigma_x}{\sigma_y}(y-\mu_y),\ (1-\rho^2)\sigma_x^2\right)$$

Since μ_x and μ_y are 0

$$N\left(\rho\frac{\sigma_x}{\sigma_y}(y),\ (1-\rho^2)\sigma_x^2\right)$$

(4)

Sample #	HasOutdoorSeating	HasBar	IsClean	HasGoodAtmosphere	IsGoodRestaurant
1	0	0	1	1	1
2	1	0	0	0	0
3	0	1	1	1	1
4	0	0	0	0	0
5	1	1	0	0	0
6	1	0	1	0	1
7	1	0	0	1	1
8	0	0	1	1	1
9	0	1	0	1	?
10	1	1	1	1	?

HasOutdoorSeating : O
HasBar : B
IsClean : C
HasGoodAtmosphere : A
IsGoodRestaurant : G.

$P(O=1 \mid G=1) = \frac{2}{5}$ $P(O=0 \mid G=1) = \frac{3}{5}$

$P(O=1 \mid G=0) = \frac{2}{3}$ $P(O=0 \mid G=0) = \frac{1}{3}$

$P(B=1 \mid G=1) = \frac{1}{5}$ $P(B=0 \mid G=1) = \frac{4}{5}$

$P(B=1 \mid G=0) = \frac{1}{3}$ $P(B=0 \mid G=0) = \frac{2}{3}$

$P(C=1 \mid G=1) = \frac{4}{5}$ $P(C=0 \mid G=1) = \frac{1}{5}$

$P(C=1 \mid G=0) = 0$ $P(C=0 \mid G=0) = 1$

$P(A=1 \mid G=1) = \frac{4}{5}$ $P(A=0 \mid G=1) = \frac{1}{5}$

$P(A=1 \mid G=0) = 0$ $P(A=0 \mid G=0) = 1$

(a) Max $P(G)$, Max $P(X \mid G)$ $X \in \{O, B, C, A\}$

$P(G) = \frac{5}{8}$

The max likehood of $P(X \mid G) = \frac{4}{5}$

(b) For sample 9

$P(O=0, B=1, C=0, A=1 \mid G)$

When $P(O=0, B=1, C=0, A=1 \mid G=0)$

$= P(O=0 \mid G=0) \cdot P(B=1 \mid G=0) P(C=0 \mid G=0) \cdot P(A=1 \mid G=0)$

$= \frac{1}{3} \cdot \frac{1}{3} \cdot 1 \cdot 0 = 0$

When $P(O=0, B=1, C=0, A=1 \mid G=1)$

$= P(O=0 \mid G=1) \cdot P(B=1 \mid G=1) P(C=0 \mid G=1) \cdot P(A=1 \mid G=1)$

$= \frac{3}{5} \cdot \frac{1}{5} \cdot \frac{1}{5} \cdot \frac{4}{5} = \frac{12}{5^4} \rightarrow \hat{G} = P(G=1) \cdot \frac{12}{5^4}$

So that for sample 9,

$\boxed{\hat{G} = \frac{5}{8} \cdot \frac{12}{5^4}}$ when $G=1$

For sample 10.

$P(O=1, B=1, C=1, A=1 \mid G)$

When $G=1$

$$P(O=1 \mid G=1)P(B=1 \mid G=1)P(C=1 \mid G=1)P(A=1 \mid G=1)$$
$$= \frac{2}{5}\cdot\frac{1}{5}\cdot\frac{4}{5}\cdot\frac{4}{5} = \frac{32}{5^4} \qquad \hat{G} = \frac{5}{8}\cdot\frac{32}{5^4}$$

When $G=0$

$$P(O=1 \mid G=0)P(B=1 \mid G=0), P(C=1 \mid G=0) P(A=1 \mid G=0)$$
$$= \frac{2}{3}\cdot\frac{1}{3}\cdot 0\cdot 0 = 0 \qquad \cancel{\hat{G}=0}$$

So that when $G=1$, it is Max value

So for sample 10, $\frac{5}{8}\cdot\frac{32}{5^4}$ when G is 1

⑤ training set : $\{(x^{(i)}, y^{(i)});\ i=\{1,\cdots,m\}\}$

$\theta_0 = P(y^{(i)}=0) \qquad x^{(i)} \in \{1,2,3\cdots,s\}^n$

$1-\theta_0 = P(y^{(i)}=1) \qquad y^{(i)} \in \{0,1\}$

ⓐ

$P(x^{(1)},\cdots,x^{(m)}, y^{(1)}\cdots y^{(m)})$

$P(x^{(1)}, y^{(1)}) \cdots P(x^{(m)}, y^{(m)})$

$= P(y^{(1)})P(x^{(1)} \mid y^{(1)}) \cdots P(y^{(m)})P(x^{(m)} \mid y^{(m)})$

* Note. $P(y^i):\ \theta_0^{1[y^i=1]}(1-\theta_0)^{1[y^i=0]}$

$P(X^i \mid y^i):\ \theta_{i,k|y=0}\left(1-\sum_{k=1}^{s-1}\theta_{i,k|y=0}\right)\theta_{i,k|y=1}\left(1-\sum_{k=1}^{s-1}\theta_{j,k|y=1}\right)$

$$= \prod_{i=1}^{m}\theta_0^{1[y^i=0]}(1-\theta_0)^{1[y^i=1]}\prod_{j=1}^{n}\theta_{j,k|y=0}^{1[x_j^i=k,\ y^i=0]}\left(1-\sum_{k=1}^{s-1}\theta_{j,k|y=0}\right)^{1[x_j^i=s,\ y^i=0]}$$
$$\prod_{j=1}^{n}\theta_{j,k|y=1}^{1[x_j^i=k,\ y^i=1]}\left(1-\sum_{k=1}^{s-1}\theta_{j,k|y=1}\right)^{1[x_j^i=s,\ y^i=1]}$$

ⓑ

When taking the log to the equation for the maximize the joint probability

$$L = \sum_{i=1}^{m} 1[y^{(i)}=0]\log\theta_0 + 1[y^{(i)}=1]\log(1-\theta_0) + \sum_{j=1}^{n} 1[x_j=k,\cdots]\log\theta_{j,k|y=0} + 1[x_j=s, y_i=0]\log\left(1-\sum_{k=1}^{s-1}\theta_{j,k|y=0}\right)$$
$$+\sum_{j=1}^{n} 1[x_j=k, y_i=1]\log\theta_{j,k|y=0} + 1[x_j=s, y_i=1]\log\left(1-\sum_{k=1}^{s-1}\theta_{j,k|y=1}\right)$$

when doing derivative of the equation by θ_0

$$\frac{\partial L}{\partial\theta_0} = \frac{\sum_{i=1}^{m}1[y^{(i)}=0]}{\theta_0} - \frac{\sum_{i=1}^{m}1[y^{(i)}=1]}{1-\theta_0} = 0$$
$$(1-\theta_0)\sum_{i=1}^{m}1[y^{(i)}=0] = \theta_0\sum_{i=1}^{m}1[y^{(i)}=1]$$
$$\left(\sum_{i=1}^{m}1[y^{(i)}=1] + \sum_{i=1}^{m}1[y^{(i)}=0]\right)\theta_0 = \sum_{i=1}^{m}1[y^{(i)}=0]$$

$$\boxed{\hat{\theta}_0 = \frac{\sum_{i=1}^{m}1[y^{(i)}=0]}{\left(\sum_{i=1}^{m}1[y^{(i)}=1] + \sum_{i=1}^{m}1[y^{(i)}=0]\right)}}$$

$$\frac{\partial L}{\partial\theta_{j,k|y=0}} = 0 \rightarrow \frac{\sum_{j=1}^{n}1[x_j=k, y_i=0]}{\theta_{j,k|y=0}} - \frac{\sum_{j=1}^{n}1[x_j=s, y_i=0]}{1-\sum_{k=1}^{s-1}\theta_{j,k|y=0}} = 0$$

$$\hat{\theta}_{j,k|y=0} = \frac{\sum_{j=1}^{n}1[x_j=k, y_i=0]}{\sum_{j=1}^{n}1[x_j=k, y_i=0] + \sum_{j=1}^{n}1[x_j=s, y_i=0]} = \boxed{\frac{\sum_{j=1}^{n}1[x_j=k, y_i=0]}{\sum_{j=1}^{n}1[y_i=0]}}$$

Based on the left calculation.

$$\hat{\theta}_{j,k|y=1} = \boxed{\frac{\sum_{j=1}^{n}1[x_j=k, y_i=1]}{\sum_{j=1}^{n}1[y_i=1]}}$$