5.27. Problem 1 & 5. Xinyu Gao 1828036

1. (a) First, $P(X\le x)=\sum_y P(X\le x\mid Y=y)\cdot P(Y=y)$

$= P(X\le x\mid Y=1)\cdot P(Y=1) + P(X\le x\mid Y=2)\,P(Y=2)$

$= \dfrac{x+4}{4} I_{(-4,-2)}(x) + \dfrac{x}{4} I_{(2,4)}(x)$.

Then we can get the pdf of X by $\frac{dP}{dx}$, i.e.

$f_X(x) = \frac{1}{4} I_{(-4,-2)}(x) + \frac{1}{4} I_{(2,4)}(x)$

Hence the marginal distribution of X is uniform.

Then apply the Bayes formula:

$$P(Y=1\mid X=x) = \frac{p(x\mid Y=1)\cdot P(Y=1)}{f_X(x)}$$

$$= \frac{\frac{1}{2} I_{(-4,-2)}(x)\cdot \frac{1}{2}}{\frac{1}{4}} = I_{(-4,-2)}(x) \quad \cdots\cdots ①$$

$$P(Y=2\mid X=x) = \frac{p(x\mid Y=2)\,P(Y=2)}{f_X(x)}$$

$$= \frac{\frac{1}{2} I_{(2,4)}(x)\cdot \frac{1}{2}}{\frac{1}{4}} = I_{(2,4)}(x) \quad \cdots\cdots ②$$

#

(b) Based on the risk $P(Y\ne f_B(x))$,

the loss function here is $L(Y, f_B(x)) = I_{\{Y\ne f_B(x)\}}$

If we want to get the $f_B(x)$, we need to minimize

$$\cdot E[L(Y, f_B(X))] \quad \cdots\cdots \quad (*)$$

and by $E[L(Y, f_B(X))] = E[E[I_{\{Y \neq f_B(X)\}} | X]]$

$$\min \ (*) \iff \min E[I_{\{Y \neq f_B(X)\}} | X]$$

when $f_B(X) = 1$, $E[I_{\{Y \neq 1\}} | X] = P(Y=2 | X) = I_{(2,4)}(X)$

$f_B(X) = 2$, $E[I_{\{Y \neq 2\}} | X] = P(Y=1 | X) = I_{(-4,-2)}(X)$

$\Rightarrow f_B(x) = 2 \cdot I_{(2,4)}(x) + 1 \cdot I_{(-4,-2)}(x)$ and $E[I_{\{Y \neq f_B(X)\}} | X] = 0$.

and $\begin{cases} x \in (-4,-2) \Rightarrow Y = 1, & f_B(x) = 1 \\ x \in (2,4) \Rightarrow Y = 2, & f_B(x) = 2 \end{cases} \Rightarrow P(Y \neq f_B(X)) = 0$.

#

(c)

$\forall X \in (-4,-2)$, we have. $|X - x_i| < |X - x_j|$

for any $x_i \in (-4,-2)$ and $x_j \in (2,4)$, since

$$|X - x_i| = |X - x_j + x_j - x_i| < |X - x_j| + |x_j - x_i| < |X - x_j|$$

$\Rightarrow$ Given a new independent pair (X, Y), if $X \in (-4,-2)$

then $\hat{f}_1(X; S) = 1$.

Similarly if given (X, Y) and $X \in (2,4)$, ~~then~~ $\hat{f}_1(X, S) = 2$.

Therefore, a wrong prediction could be made if.

all $X_{samples} \in (-4,-2)$ while $X_{new} \in (2,4)$

or.

all $X_{samples} \in (2,4)$ while $X_{new} \in (-4,-2)$, i.e.

$$P(Y \neq \hat{f}_1(x;S)) = P(X \in (2,4)) \cdot \prod_{i=1}^{n} P(x_i \in (-4,-2)) + P(X \in (-4,-2)) \cdot \prod_{i=1}^{n} P(x_i \in (2,4))$$

$$= \frac{1}{2} \cdot \frac{1}{2^n} + \frac{1}{2} \cdot \frac{1}{2^n} = \frac{1}{2^n}$$

#

(d) We denote $\hat{f}_3(x;S)$ as in 3-NN.

$$P(Y \neq \hat{f}_3(x;S)) = P(X \in (2,4)) \cdot P\left[\sum_{i=1}^{n} I_{(2,4)}(x_i) \leq 1\right] + P(X \in (-4,-2)) \cdot P\left[\sum_{i=1}^{n} I_{(-4,-2)}(x_i) \leq 1\right]$$

$$= \frac{1}{2} \cdot \frac{n+1}{2^n} \times 2 = \frac{n+1}{2^n}$$

#

(e) Since $\frac{1}{2^n} < \frac{n+1}{2^n}$

1-nn has smaller risk than 3-nn. #

5.

1. Since $\lambda = \infty$, $\Rightarrow \int g^2(x)\,dx = 0$.

by $g^2(x) \geq 0 \Rightarrow \hat{g}(x) = 0$

#

2. $\lambda = \infty \Rightarrow \int [g'(x)]^2 dx = 0$

$\Rightarrow$ $g(x)$ is a constant. and assume it to be c.

then $\hat{g} = \underset{c}{\arg\min} \sum_{i=1}^{N} (y_i - c)^2$

where $Q(c) = \sum_{i=1}^{N} (y_i - c)^2$,

$$\frac{dQ}{dc} = 2\sum_{i=1}^{N}(y_i - c) = 0 \Rightarrow c = \frac{1}{N}\sum_{i=1}^{N} y_i \Rightarrow \hat{g} = \frac{1}{N}\sum_{i=1}^{N} y_i$$ #

3. $\lambda = \infty \Rightarrow \int [g''(x)]^2 dx = 0 \Rightarrow g(x) = ax+b, \quad a, b \in \mathbb{R}.$

set $Q(a,b) = \sum_{i=1}^{N} (y_i - ax_i - b)^2$

$$\begin{cases} \frac{\partial Q}{\partial a} = 2\sum_{i=1}^{N} (y_i - ax_i - b)\cdot(-x_i) = 0 \\ \frac{\partial Q}{\partial b} = 2\sum_{i=1}^{N} (y_i - ax_i - b)\cdot(-1) = 0 \end{cases}$$

$$\Rightarrow \begin{cases} \hat{a} = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})y_i}{\sum_{i=1}^{N}(x_i - \bar{x})^2} \\ \hat{b} = \bar{y} - \hat{a}\cdot\bar{x}, \end{cases} \quad \text{and } \hat{g}(x) = \hat{a}x + \hat{b}$$

4. $\lambda = \infty \Rightarrow \int [g^{(3)}(x)]^2 dx = 0 \Rightarrow g(x) = ax^2 + bx + c, \quad a, b, c \in \mathbb{R}.$

set $Q(a,b,c) = \sum_{i=1}^{n} (y_i - (ax_i^2 + bx_i + c))^2$

we assume $P_{1i} = x_i$, $P_{2i} = x_i^2$

$\Rightarrow Q(a,b,c) = (Y - P\beta)^T(Y - P\beta)$

where $P = \begin{pmatrix} 1 & P_{11} & P_{21} \\ 1 & P_{12} & P_{22} \\ \vdots & \vdots & \vdots \\ 1 & P_{1N} & P_{2N} \end{pmatrix}$, $\beta = \begin{pmatrix} c \\ b \\ a \end{pmatrix}$

$\begin{pmatrix} \hat{c} \\ \vdots \end{pmatrix} \quad (P^TP)^{-1}P^TY$

$$\begin{pmatrix} \hat{b} \\ \hat{a} \end{pmatrix} = (P^T P)^{-1} P^T Y.$$

$$\Rightarrow \hat{g}(x) = \hat{a}x^2 + \hat{b}x + \hat{c}. \qquad \#$$

5. $\lambda = 0$. we just need to minimize $\sum_{i=1}^{N} [y_i - g(x_i)]^2$.

we will find some function that satisfy $g(x_i) = y_i$ for $i = 1, 2, \cdots, N$.

we can fit it by cubic spline.