Problem 1

1.

$$\varepsilon = \sqrt{\frac{\ln t - \frac{1}{2}\ln\delta}{N_m}}$$

μ_m：母体得到 coin 的机率

$\frac{c_m}{N_m}$：抽样得到 coin 的机率

$$\exp\left(-2 * \frac{(\ln t - \frac{1}{2}\ln\delta)}{\cancel{N_m}} * \cancel{N_m}\right)$$

$$= \exp(-2(\ln t - \tfrac{1}{2}\ln\delta))$$

$$= \exp(-2\ln t + \ln\delta)$$

$$= \delta t^{-2}$$

$\Rightarrow$ $P\left(\mu_m > \frac{c_m}{N_m} + \varepsilon\right) \le \delta t^{-2}$ #

Problem 2

2.
prove $P\left(\mu_m \le \frac{c_m}{N_m} + \sqrt{\frac{\log t + \log M - \frac{1}{2}\log \delta}{N_m}}\right) \ge 1-\delta$

$= 1 - P\left(\mu_m > \frac{c_m}{N_m} + \sqrt{\frac{\log t + \log M - \frac{1}{2}\log \delta}{N_m}}\right) \ge 1-\delta$

$= P\left(\mu_m > \frac{c_m}{N_m} + \sqrt{\frac{\log t + \log M - \frac{1}{2}\log \delta}{N_m}}\right) \le \delta$

by one-sided Hoeffding's inequality,

for all $m=1,\dots,M$, $P\left(\mu_m > \frac{c_m}{N_m} + \sqrt{\frac{\log t + \log M - \frac{1}{2}\log \delta}{N_m}}\right) \le \exp\left(-2\frac{(\log t + \log M - \frac{1}{2}\log\delta)}{\cancel{N_m}}\cancel{N_m}\right)$

$$\le \exp(\log t^{-2} + \log M^{-2} + \log \delta)$$

$$\le \exp(\log((tM)^{-2}\delta))$$

$$\le t^{-2}M^{-2}\delta$$

$$\le \frac{1}{t^2}\frac{1}{M^2}\cdot\delta$$

$$\le 1\cdot 1\cdot \delta \quad \because t > M \ge 2$$

$$\le \delta$$

Problem 3

3 5 张 ticket 的排列可能，AB 與 CD 不能同時出现

种类只有1种 {A}, {B}, {C}, {D} 的排列数 = 4

2种 ~~{AB}~~, {AC}, {AD}

{BC}, {BD}

~~{CD}~~

的排列数 = $4 \times \left(\frac{5!}{4!} + \frac{5!}{2!3!} + \frac{5!}{3!2!} + \frac{5!}{4!}\right) = 120$

3种 ~~{ABC}~~, ~~{ABD}~~, ~~{BCD}~~

4种 ~~{ABCD}~~

$\frac{124}{1024} = \frac{31}{256}$ X

Problem 4

4.

A、C不能出現

5 份 ticket 的排列可能

種類有1種 ~~{A}~~, {B}, ~~{C}~~, {D} 的排列數 = 2

2種 ~~{AB}~~, ~~{AC}~~, ~~{AD}~~

~~{BC}~~, {BD}

~~{CD}~~

的排列數 = $\left(\frac{5!}{4!}+\frac{5!}{2!3!}+\frac{5!}{2!3!}+\frac{5!}{4!}\right)=30$

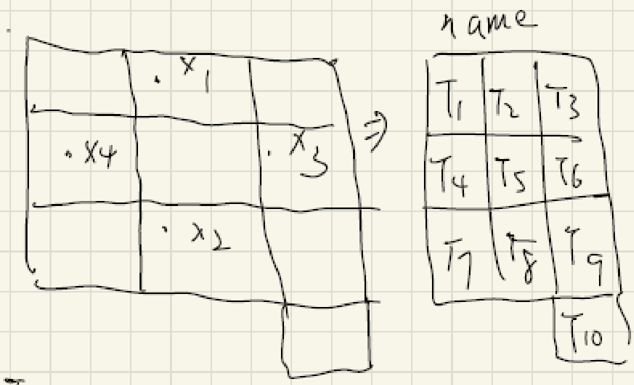
3種 ~~{ABC}~~, ~~{ABD}~~, ~~{BCD}~~

4種 ~~{ABCD}~~

$\frac{32}{1024}=\frac{1}{32}$

Problem 5

5.



O = 代表 $y=-1$

X = 代表 $y=1$

	x_1	x_2	x_3	x_4
d_1	X	X	X	X
d_2	O	X	X	X
d_3	X	O	X	X
d_4	X	X	O	X
d_5	X	X	X	O
d_6	O	O	X	X
d_7	O	X	O	X
d_8	O	X	X	O
d_9	X	O	O	X
d_{10}	X	O	X	O
d_{11}	X	X	O	O
d_{12}	O	O	O	X
d_{13}	O	O	X	O
d_{14}	O	X	O	O
d_{15}	X	O	O	O
d_{16}	O	O	O	O

d_1: T_{10}

d_2: T_2, d_3: T_8, d_4: T_6, d_5: T_4

d_6: $T_2 T_5 T_8$, d_7: $T_2 T_3 T_5 T_6$

d_8: $T_1 T_2 T_4 T_5$, d_9: $T_5 T_6 T_8 T_9$

d_{10}: $T_4 T_5 T_7 T_8$, d_{11}: $T_4 T_5 T_6$

d_{12}: $T_2 T_3 T_5 T_6 T_8 T_9$, d_{13}: $T_1 T_2 T_4 T_5 T_7 T_8$

d_{14}: $T_1 T_2 T_3 T_4 T_5 T_6$, d_{15}: $T_4 T_5 T_6 T_7 T_8 T_9$

d_{16}: $T_1 \ldots T_9$

prove 了

Problem 6

VC dimension = 2M

6. 欲 prove $d_{vc} \le 2M$

假設 some $2M+1$ inputs can shattered. $(I_1, I_2, \cdots I_{2M+1})$. 先將其排序成 $(X_1, X_2, \cdots, X_{2M-1}, X_{2M}, X_{2M+1})$，那一定出現某種 dichotomy $(o, x, o, x, \cdots, o, x, o)$，共 $M+1$ 個區間，但最多 M 個區間，因此假設不成立，any $2M+1$ inputs can not be shattered $\Rightarrow d_{vc} \le 2M$

欲 prove $d_{vc} \geq 2M$,

2M input $(I_1, I_2, \cdots I_{2M})$, 先排序成 $(x_1, x_2, \cdots x_{2M})$, 以下 hypothesis 可以將 2M input be shattered

$$a_1 = x_1 + \varepsilon_1, \quad b_1 = x_2 - \varepsilon_2$$

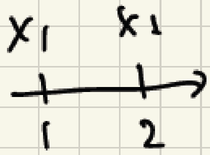
$$a_2 = x_3 + \varepsilon_3, \quad b_2 = x_4 - \varepsilon_4$$

$$\vdots$$

$$a_M = x_{2M-1} + \varepsilon_{2M-1} \quad b_M = x_{2M} - \varepsilon_{2M}$$

s.t. $0 \leq \varepsilon_i < \dfrac{x_{2i} - x_{2i-1}}{2}, \ 1 \leq i \leq 2M$

e.g



2 個 input

1 個 interval

(M=1)

$a_1 = x_1, \ b_1 = x_2$

→ o , o

$a_1 = x_1 + \frac{1}{3}, \ b_1 = x_2$

→ × , o

$a_1 = x_1, \ b_1 = x_2 - \frac{1}{3}$

→ o , ×

$a_1 = x_1 + \frac{1}{3}, \ b_1 = x_2 - \frac{1}{3}$

→ × , ×

Problem 7

7. $m_H(N)$ 為 $2N$

Assume $\text{sign}(0)=+1$

因為 perceptron 一定过原点，所以當一点

$h(x)=+1$, $h(c_1 x)=\begin{cases} +1, & c_1 \geq 0 \\ -1, & c_1 < 0 \end{cases}$ T圖

此可以想成是在一半圓上又因 PLA

會分割兩段，可想成為 positive

and negative ray，最終 $m_H(N)=2N$

Problem 8

Problem 9

9, as $s=+1$, $E_{out}(h_{+1,\theta}) = 0.1 + 0.4|\theta|$

$p(n)$: the probability of noise flip the sign

$p(\alpha|n)$: flip後,預測錯的机率

$p(\beta|n)$: flip後預測對的机率

$$E_{out}(h_{1,\theta}) = \mathop{\mathcal{E}}_{x\sim[-1,1]} [[h_{1,\theta}(x) - \text{sign}(x)]] + \text{noise}$$

positive ray

$$= \frac{|\theta|}{2} + p(n)\left[p(\alpha|n) - p(\beta|n)\right]$$

$$= \frac{|\theta|}{2} + 0.1\left[\frac{2-|\theta|}{2} - \frac{|\theta|}{2}\right]$$

$$= \frac{|\theta|}{2} + 0.1(1-|\theta|)$$

$$= 0.1 + 0.4|\theta|$$

as $s = -1$, $E_{out}(h_{1,\theta}) = 0.9 - 0.4\theta$

$$E_{out}(h_{-1,\theta}) = \underset{x \sim [-1,1]}{\mathcal{E}} [[h_{-1,\theta}(x) - \text{sign}(x)]] + \text{noise}$$

negative value

$$= \frac{2-\theta}{2} + P(n) [P(d(n)) - P(\beta(n))]$$

$$= 1 - 0.5\theta + 0.1 \left(\frac{|\theta|}{2} - \frac{2-|\theta|}{2} \right)$$

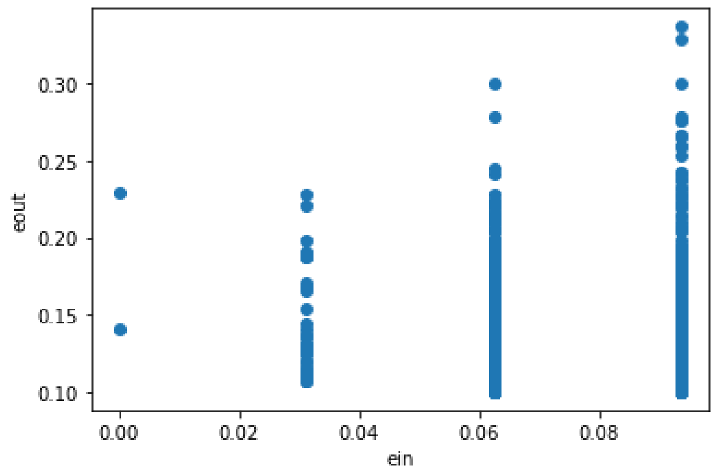
$$= 1 - 0.5\theta + 0.1 \left(\frac{2|\theta| - 2}{2} \right)$$

$$= 1 - 0.4\theta - 0.1 = 0.9 - 0.4\theta$$

so, $E_{out}(h_{s,\theta}) = 0.5 - 0.4s + 0.4s|\theta|$

Problem 10

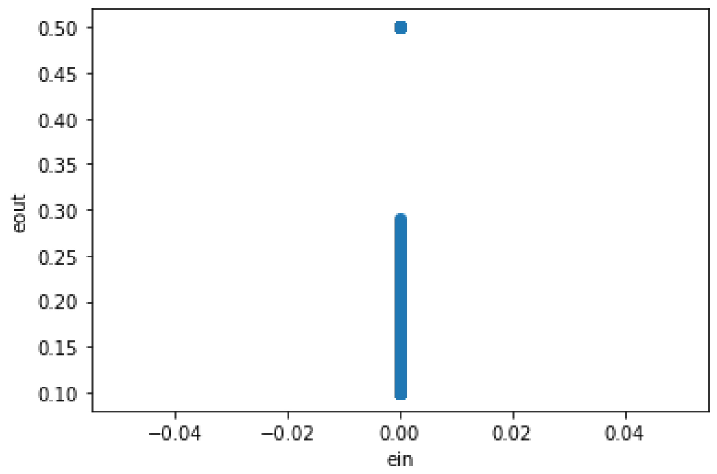
Plot a scatter plot of $(E_{in}(g), E_{out}(g))$



Calculate the median of $E_{out}(g) - E_{in}(g)$ = 0.023338267157456044

Problem 11

Plot a scatter plot of $(E_{in}(g), E_{out}(g))$

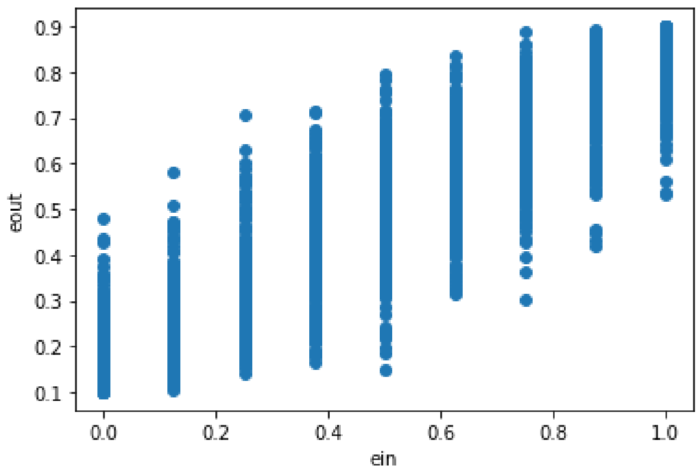


Calculate the median of $E_{out}(g) - E_{in}(g)$ = 0.1322171156279302

Describe your findings: 當資料越少，越容易過擬合到目前的資料，導致 (1) . E_{in} 能越小 (2) $E_{out} - E_{in}$ 的差距越大。

Problem 12

Plot a scatter plot of $(E_{in}(g), E_{out}(g))$



Calculate the median of $E_{out}(g) - E_{in}(g)$ = 0.00637335392449162

Describe your findings: 隨機選 hypothesis 容易選擇到較差的 hypothesis 導致較大 E_{in}