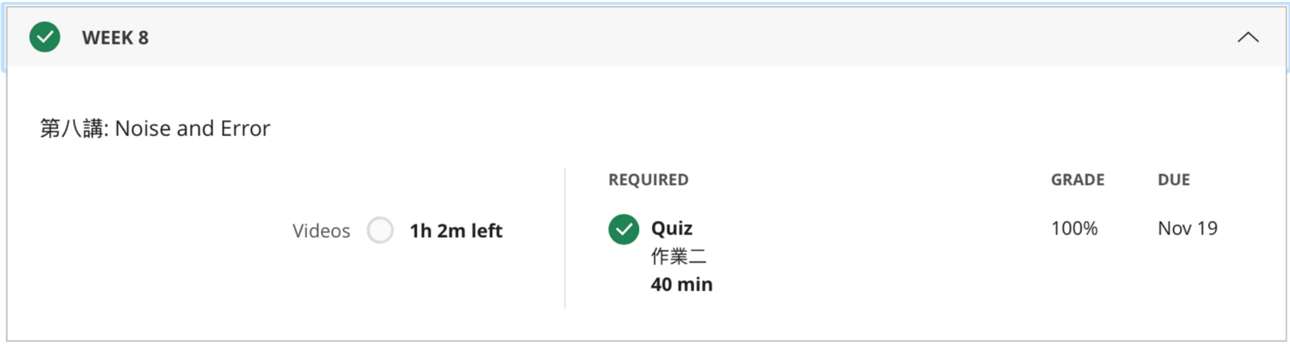
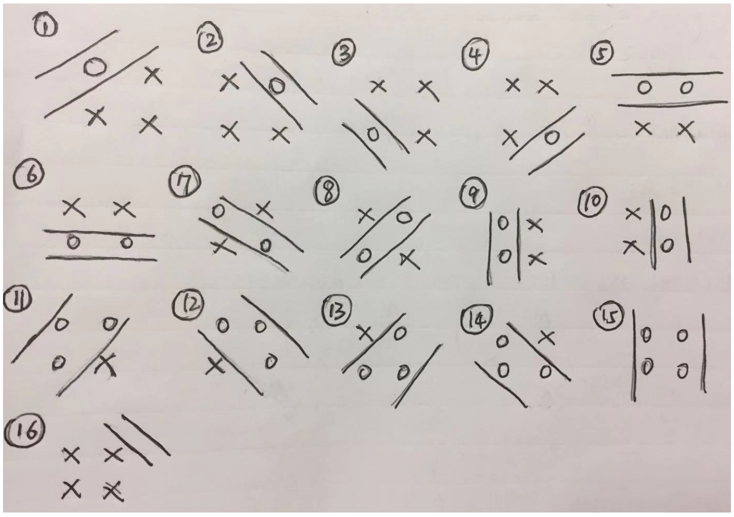
1.




2.

Let “O” = +1, “X” = -1

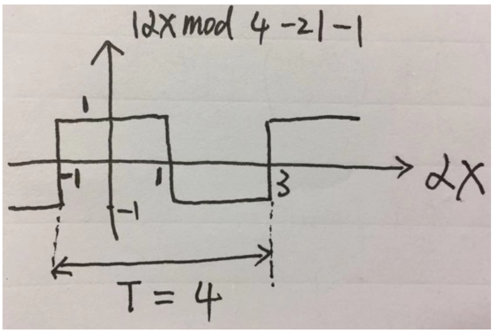


∴ the hypothesis set can shatter 4 inputs, and the VC-Dimension is no less than 4.

3.

證明：$d_{vc}(H) = \infty$

將$\alpha x$看作變量，函數$|\alpha x \ mod\ 4 - 2| - 1$是 T = 4 的方波，如下圖：




Let $x_i = 4^i$ $\quad \alpha = -\frac{1}{2}\sum_{i=1}^{n}(y_i-1)4^{-i}$ $\quad (i \geq 1)$

when $y_j = 1$：

$$\alpha x_j = 4^j * \left(-\frac{1}{2}\right)\sum_{i=1,y_i=-1}^{n}(-2)4^{-i}$$

$= \sum_{i=1, y_i=-1}^{n} 4^{j-i}$

$= \sum_{y_i=-1, i<j}^{n} 4^{j-i} + 0 + \sum_{y_i=-1, i>j}^{n} 4^{j-i}$

$= 4k + 0 + \sum_{y_i=-1, i>j}^{n} 4^{j-i}$ (k 為某個整數)

$\because 0 \le \sum_{y_i=-1, i>j}^{n} 4^{j-i} < \sum_{i=1}^{+\infty} 4^{-i} < \frac{1}{1-4^{-1}} - 1 = \frac{1}{3}$

$\therefore 0 + 4\mathrm{k} \le \alpha x_j < \frac{1}{3} + 4k$

$\because y_i = 1$， when $-1 + 4\mathrm{k} < \alpha x_i < 1 + 4k$

$\therefore h_\alpha(x_j) = 1$, when $y_j = 1$

when $y_j = -1$：

$\alpha x_j = 4^j * \left(-\frac{1}{2}\right) \sum_{i=1, y_i=-1}^{n} (-2) 4^{-i}$

$= \sum_{y_i=-1, i<j}^{n} 4^{j-i} + 4^0 + \sum_{y_i=-1, i>j}^{n} 4^{j-i}$

$= 4k + 1 + \sum_{y_i=-1, i>j}^{n} 4^{j-i}$ (k 為某個整數)

$\therefore 1 + 4\mathrm{k} \le \alpha x_j < \frac{4}{3} + 4k$

$\because y_i = -1$， when $1 + 4\mathrm{k} \le \alpha x_i \le 3 + 4k$

$\therefore h_\alpha(x_j) = -1$, when $y_j = -1$

∴ 任何 y 的+1 和-1 排列組合，都有一個$h_\alpha$可以以滿足

∴ $h_\alpha$可以 shatter 無窮多 inputs，$d_{vc}(H) = \infty$

4.

假設: $S_1$ = {dichotomies of $H_1$}，$S_2$ = {dichotomies of $H_2$} $|S|$為集合內 dichotomy 的數量

$\because |S_1 \cap S_2| \le |S_2|$

$\because m_{H1 \cap H2}(N) = |S_1 \cap S_2|$，$m_{H2}(N) = |S_2|$

$\because \mathrm{N} = d_{vc}(H_2)$:

$m_{H1 \cap H2}(N) \le m_{H2}(N) = 2^N$

∴ $m_{H1 \cap H2}(N)$有可能 shatter $d_{vc}(H_2)$ 個點

$\because \mathrm{N} = d_{vc}(H_2) + 1$:

$m_{H1 \cap H2}(N) \le m_{H2}(N) < 2^N$

∴ $m_{H1 \cap H2}(N)$不可能 shatter $d_{vc}(H_2) + 1$ 個點

$\therefore d_{vc}(H_1 \cap H_2) \le d_{vc}(H_2)$

5.

∵ $m_{H1\cup H2}(N) = 2\mathrm{N} \quad d_{vc}(H_1) = d_{vc}(H_2) = 1$

when $\mathrm{N} = d_{vc}(H_1) = 1,\ m_{H1\cup H2}(N) = 2 = 2^N$

∴ $can\ shatter\ 1\ input!$

∵ when $\mathrm{N} = d_{vc}(H_1) + 1 = 2,\ m_{H1\cup H2}(N) = 4 = 2^N$

∴ $can\ shattere\ 2\ inputs!$

∵ when $\mathrm{N} = d_{vc}(H_1) + 2 = 3,\ m_{H1\cup H2}(N) = 6 < 2^N$

∴ $can\ not\ shatter\ 3\ inputs!$

∴ $d_{vc}(H_1 \cup H_2) = d_{vc}(H_1) + 1 = 2$

6.

假設 $\tilde{\mu}$ 為沒有 noise 情況下的 $E_{out}$

① s=+1, θ ≥ 0時: $\tilde{\mu} = \frac{\frac{N}{2}*\theta}{N} = 0.5 * \theta$

② s=+1, θ < 0時: $\tilde{\mu} = \frac{\frac{N}{2}*|\theta|}{N} = 0.5 * |\theta|$

③ s=-1, θ ≥ 0時: $\tilde{\mu} = \frac{\frac{N}{2}*(2-\theta)}{N} = 1 - 0.5 * \theta$

④ s=-1, θ < 0時: $\tilde{\mu} = \frac{\frac{N}{2}*(2-|\theta|)}{N} = 1 - 0.5 * |\theta|$

∴ $\tilde{\mu}_{(\mathrm{s},\theta)} = 0.5(s+1) * 0.5|\theta| - 0.5(s-1)(1 - 0.5|\theta|)$

$= 0.5 * \mathrm{s} * |\theta| - 0.5s + 0.5$

∵ noise=0.2, 假設$\lambda = 1 - \mathrm{noise} = 0.8$

∴ $E_{out} = \frac{1}{N} * (\lambda\tilde{\mu} * \mathrm{N} + (1-\lambda)(1-\tilde{\mu}) * N)$

$= \lambda\tilde{\mu} + (1-\lambda)(1-\tilde{\mu})$

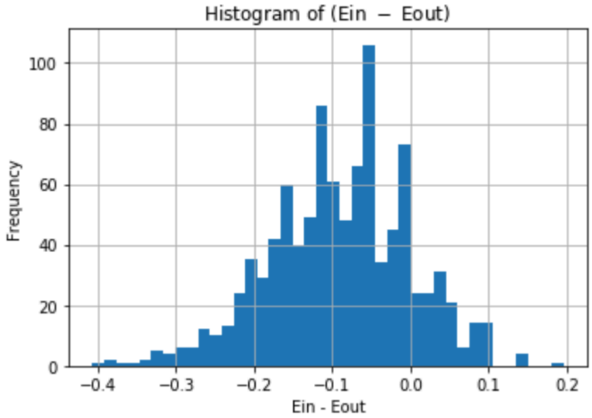
∴ $E_{out}(h_{s,\theta}) = \lambda\tilde{\mu}_{(\mathrm{s},\theta)} + (1-\lambda)\left(1 - \tilde{\mu}_{(\mathrm{s},\theta)}\right)$

$= 1 - \lambda + \tilde{\mu}_{(\mathrm{s},\theta)}(2\lambda - 1)$

$= 0.5 + 0.3 * s(|\theta| - 1)$

7.

```
Average Ein: 0.1645000000000001
Average Eout: 0.2548253120869515
```




($E_{in}$-$E_{out}$)的分佈與高斯分佈相似，mean 在-0.1~0.0 之間，可見即使有 20%的 noise，在 20 筆 training data 下， VC Bound 保證了 $E_{in}$ 和 $E_{out}$ 不會相差太大。($E_{in}$-$E_{out}$)大部分小於零，說明 $E_{in}$ 通常小於 $E_{out}$，演算法往往會選擇 fit noise 的 hypothesis。在含有 20% noise 的情況下，真正正確的 hypothesis(S=1, θ = 0)的平均 $E_{in}$ 和 $E_{out}$ 都約為 0.2，所以 $E_{in}$ 最小的 hypothesis 不一定最好，說不定這個 hypothesis 的 $E_{out}$ 不是最小的。

8.

證明：$B(N,k) \geq \sum_{i=0}^{k-1}\binom{N}{i}$

Goal：construct a special set of $\sum_{i=0}^{k-1}\binom{N}{i}$ dichotomies that dose not shatter any subset of k variable.

假設有一組 N 維度 dichotomy 集合 S，根據每個 dichotomy 中「+1」的數量，S 由以下 K 類 dichotomies 組成：

| | 每個 dichotomy 中「+1」的數量 | dichotomy 數量 |
|---|---|---|
| 第 1 類 | 0 | $\binom{N}{0}$ |
| 第 2 類 | 1 | $\binom{N}{1}$ |
| … | … | … |
| 第 K 類 | k-1 | $\binom{N}{k-1}$ |

∵ S 的 dichotomy 最多只包含 K-1 個「+1」

∴ 從 S 中任取 K 個維度組成新的 dichotomy 集合 $S_2$，$S_2$ 中沒有全為「+1」的 dichotomy

∴ 因為缺少 K 個「+1」的 dichotomy，S 無法 shatter 任意 k 個維度， 所以 S 滿足$B(N,k)$

∴ $B(N,k) \geq |S| = \binom{N}{0} + \binom{N}{1} + \cdots + \binom{N}{k-1}$

∴ $B(N,k) \geq \sum_{i=0}^{k-1}\binom{N}{i}$

證明：$B(N,k) \le \sum_{i=0}^{k-1}\binom{N}{i}$

假設一組 dichotomy 集合 S 維度為 N，$|S|$為集合 S 中元素的個數，$|S| = B(N,k)$。根據最後一個維度，S 的結構可以用下圖表示：

<table>
<tr><td rowspan="5">S</td><td>$x_0$~ $x_{N-2}$</td><td>$x_{N-1}$</td></tr>
<tr><td>$S_1$</td><td>+1</td></tr>
<tr><td>$S_2$</td><td>-1</td></tr>
<tr><td>$S_3$</td><td>+1</td></tr>
<tr><td>$S_3$</td><td>-1</td></tr>
</table>

| $x_0$~ $x_{N-2}$ | $x_{N-1}$ |
|---|---|
| $S_1$ | +1 |
| $S_2$ | -1 |
| $S_3$ | +1 |

∵ S 中任意 K 個點不能被 shattered

∴ $x_0$~ $x_{N-2}$ 中任意 K 點也不能被 shattered

∴ $|S_1 + S_2 + S_3| \le B(N-1,k)$

| $x_0$~ $x_{N-2}$ | $x_{N-1}$ |
|---|---|
| $S_3$ | +1 |
| $S_3$ | -1 |

∵ 若 S3 中任意 K-1 個點可以被 shattered，則搭配所對應的 $x_{N-1}$ 同時有+1 和-1，S 就可以被 K 個點 shattered

∴ S3 不能被 k-1 個點 shattered

∴ $|S_3| \le B(N-1,k-1)$

∴ $|S| = B(N,k) \le B(N-1,k) + B(N-1,k-1)$

∵ N = 1 時

$$\mathrm{B}(1,1) = 1 \quad \sum_{i=0}^{k-1}\binom{N}{i} = \binom{1}{0} = 1$$

$$\mathrm{B}(1,2) = 2 \quad \sum_{i=0}^{k-1}\binom{N}{i} = \binom{1}{0} + \binom{1}{1} = 2$$

∴ $\mathrm{B(N,k)} \le \sum_{i=0}^{k-1}\binom{N}{i}$ 在 N=1 時成立

∴ 假設 $\mathrm{N} \le N_0$ 時，$\mathrm{B(N,k)} \le \sum_{i=0}^{k-1}\binom{N}{i}$成立

$\therefore$ $\mathrm{N} = N_0 + 1$時

$$B(N_0+1,k) \le B(N_0,k) + B(N_0,k-1)$$

$$\le \sum_{i=0}^{k-1}\binom{N_0}{i} + \sum_{i=0}^{k-2}\binom{N_0}{i}$$

$$= \binom{N_0}{0} + \binom{N_0}{1} + \cdots + \binom{N_0}{k-1} + \binom{N_0}{0} + \binom{N_0}{1} + \cdots + \binom{N_0}{k-2}$$

$$= \binom{N_0}{0} + \binom{N_0+1}{1} + \cdots + \binom{N_0+1}{k-1}$$

$$= \sum_{i=0}^{k-1}\binom{N_0+1}{i}$$

$\therefore$ $\mathrm{B(N,k)} \le \sum_{i=0}^{k-1}\binom{N}{i}$成立

$\because$ $B(N,k) \ge \sum_{i=0}^{k-1}\binom{N}{i}$

$\therefore$ $B(N,k) = \sum_{i=0}^{k-1}\binom{N}{i}$