0.1.

1) According to theorem 2.4. if $m_{\mathcal{H}}(N) < 2^N$ then $\exists k$ for which $m_{\mathcal{H}}(N) < 2^k$, so

$$m_{\mathcal{H}}(N) \leq \sum_{i=0}^{k-1} \binom{N}{i}$$

which is a polynomial; therefore, there is no such Hypothesis set.

2) a) error bar $= \sqrt{\frac{1}{2N} \ln \frac{2M}{\delta}}$

$$\epsilon_{in} = \sqrt{\frac{1}{2 \cdot 400} \ln \frac{2 \cdot 1000}{0.05}} = 0.115$$

$$\epsilon_{test} = \sqrt{\frac{1}{2 \cdot 200} \ln \frac{2 \cdot 1}{0.05}} = 0.096$$

$\epsilon_{in} > \epsilon_{test}$ $\Rightarrow$ so the estimate provided by test sample is better.

b) The more examples we use for testing, the fewer examples we use for training. Which means, we may not get a good hypothesis from training set.

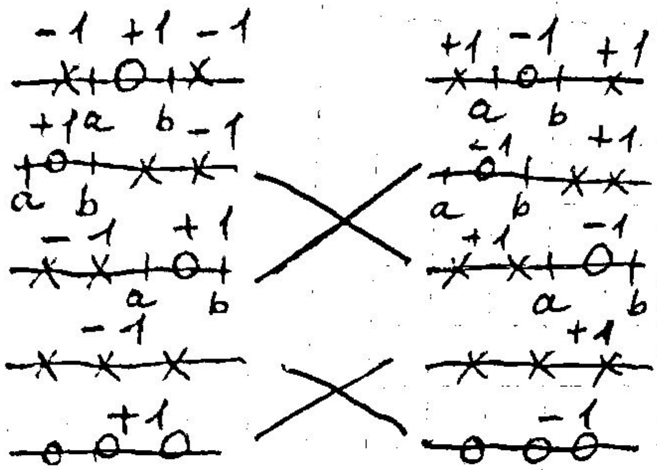
Suppose $N = 1$ for training set & the rest for (K) testing $\Rightarrow$

$$E_{out} \leq E_{in} + \sqrt{\frac{1}{2K} \ln \frac{2}{\delta}}$$

which is bounded tighter.
However, this accounts for overfitting.
Therefore, having less training set & more testing set is not always a good idea.

:)

simpler way :



the same labeled as X

From which we can see that only +1, -1, +1 dichotomy is added , so we have added only 1 dichotomy for $N=3$ points

$$\Rightarrow \frac{N^2}{2} + \frac{N}{2} + 1 + (N-2) = \frac{N^2}{2} + \frac{3N}{2} - 1$$

To find d_{vc} :

$$2^N = \frac{N^2}{2} + \frac{3N}{2} - 1$$

$N_{max} = 3$. Check :

$$2^3 \stackrel{?}{=} \frac{9}{2} + \frac{9}{2} - 1 = \frac{16}{2}$$

$$8 = 8 \quad ✓$$

So, $d_{vc} = 3$

4) a) First we construct a non-singular $(D+1)\times(D+1)$ matrix.

Non-singular matrix, means $\det X \neq 0$ & X is a square matrix. Suppose

$$X = \begin{bmatrix} 1 & x_0^1 & \cdots & x_0^D \\ 1 & x_1^1 & \cdots & x_1^D \\ \vdots & x_2^1 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \\ 1 & x_D^1 & \cdots & x_D^D \end{bmatrix}$$ where all x_k are different.

Consider $y = (y_0, \dots, y_D)^T \in \{-1, +1\}^{D+1}$

& let $c = (c_0, \dots c_D)^T = X^{-1} y \Rightarrow Xc = y$

$\Rightarrow h_c(x) = \text{sign}\left(\sum_{i=0}^{D} c_i x_k^i\right) = y_k \quad k = 0, \dots, D$

$\Rightarrow m_H(D+1) = 2^{D+1}$ & $d_{vc} \geq D+1$.

b) Since X has $D+1$ vectors, any $D+2$ vectors of length $D+1$ have to be linearly dependent.

$$\Rightarrow (x_m^0, x_m^1, \dots, x_m^D) = \sum_{k\neq m}^{D} a_k (x_k^0, x_k^1, \dots, x_k^D)$$

(linear combination) for each $a_k \neq 0$

Suppose $y_m = -1$ & $y_k = \text{sign}(a_k) = \text{sign}\left(\sum_{i=0}^{D} c_i x_k^i\right)$

$\Rightarrow \sum_{i=0}^{D} c_i a_k x_k^i > 0$ (because signs are the same)

let's consider

$$(x_m^0, x_m^1, \dots, x_m^D)\begin{pmatrix} c_0 \\ \vdots \\ c_D \end{pmatrix} = \sum_{k\neq m}^{D} \sum_{i=0}^{D} c_i a_k x_k^i$$

$$\Rightarrow \sum_{k\neq m}^{D} \sum_{i=0}^{D} c_i a_k x_k^i = \sum_{i=0}^{D} c_i x_m^i > 0$$

$\Rightarrow \quad y_m = +1$

$y_m = -1$ & $y_m = +1$ is a contradiction.

so $D+2$ points cannot be shattered by $\mathcal{H}$.

$\Rightarrow \quad d_{vc} \leq D+1$

$\Rightarrow \quad d_{vc} = D+1$

5) suppose $x_1 = 10^1, x_2 = 10^2 \ldots x_N = 10^N$

& let $y = (y_1 \ldots y_N)^T \in \{-1, +1\}^N$

Consider $\alpha = \frac{1}{10^k}$ if $y = -1$ & $\alpha = \frac{2}{10^k}$ if $y = +$

$$\Rightarrow \quad h_\alpha(x_k) = (-1)^{\lfloor \alpha \cdot 10^k \rfloor} = y_k \quad \text{for } k = 1, \ldots, N$$

$$\Rightarrow \mathcal{H}(x_1, \ldots, x_N) = \{-1, +1\}^N \Leftrightarrow m_{\mathcal{H}}(N) = 2^N$$

$\Rightarrow d_{vc} = \infty$

Q.2

a) Let $X = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2^2 & \cdots & 2^d \\ 1 & 3 & 3^2 & \cdots & 3^d \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & d+1 & (d+1)^2 & \cdots & (d+1)^d \end{bmatrix}$

X is invertible matrix, because it's in Vandermonde matrix form & it has $(d+1) \times (d+1)$ dimension.
So, for $y = (y_1, y_2, \ldots y_{d+1})^T$ we have a solution $(x_1, x_2, \ldots, x_{d+1})^T$

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2^2 & \cdots & 2^d \\ 1 & 3 & 3^2 & \cdots & 3^d \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & d+1 & (d+1)^2 & \cdots & (d+1)^d \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{d+1} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{d+1} \end{bmatrix}$$

In other words, $d+1$ points can be shattered, therefore

$$d_{vc} \geq d+1$$

b) We know that any $(d+2)$ vectors in $(d+1)$ dimension are linearly dependent, so we have a linear combination of z_i

$$z_{d+2} = \sum_{i=1}^{d+1} k_i z_i$$

$$w^T z_{d+2} = \sum_{i=1}^{d+1} w^T k_i z_i$$

$$\text{sign}(w^T z_{d+2}) = \text{sign}\left(\sum_{i=1}^{d+1} k_i w^T z_i\right)$$

$$\Rightarrow (\text{sign}(z_1), \text{sign}(z_2) \ldots \text{sign}(z_{d+1}), -\text{sign}(z_{d+2}))$$

This case cannot be shattered, so

$$d_{vc} \leq d+1$$

$$\Rightarrow d_{vc} = d+1.$$

0.3.

1) let $g(X) = (h(x) - f(x))^2$, then

$g(X) = 0$ or $g(X) = 1$, since

$h(x)$	$f(x)$
0	1
1	0
1	1
0	0

$P[h(x) \neq f(x)] = \mathbf{P}(|h(x) - f(x)| = 1)$

then $E(g(X)) = \sum_{n=1}^{N} g(X) P(X = x_n) =$

$= \sum_{n=1}^{N} g(X) \cdot \frac{1}{N} = \frac{1}{N} \sum_{n=1}^{N} (0 \mid 1) = P(h(x) \neq f(x))$

from Ch. 1.

2) $E(g(X)) = \sum_{1}^{N} g(X) P(X = x_n) =$

$= \frac{1}{N} \sum_{1}^{N} (2^2 \mid 0^2) = \frac{1}{N} \sum_{1}^{N} (4 \mid 0)$

$\Rightarrow P[h(x) \neq f(x)] = \frac{1}{4} E(g(X))$

h	f
-1	-1
1	1
-1	1
1	-1

$\max |h(x) - f(x)| = 2$

3) $m_H(N) \leq \sum_{i=0}^{d_{vc}} \binom{N}{i}$

• if $d_{vc} = 1 \Rightarrow \sum_{i=0}^{1} \binom{N}{i} = \binom{N}{0} + \binom{N}{1} = 1 + \frac{N!}{(N-1)!} = 1+N$ possible

• if $d_{vc} = 2 \Rightarrow \sum_{i=0}^{2} \binom{N}{i} = \binom{N}{0} + \binom{N}{1} + \binom{N}{2} =$

$= 1 + \frac{N!}{(N-1)!} + \frac{N!}{(N-2)!\,2!} = 1 + N + \frac{N(N-1)}{2}$ possible

• 2^N is possible

$1 + N + \frac{N(N-1)(N-2)}{6}$ is possible, $d_{vc} = 3$

• if $m_H(N) = 2^{\lfloor \sqrt{N} \rfloor}$, it should be $< 2^k$

min $k = 2$ as $2^1 < 2^2 \Rightarrow d_{vc} = 1$

$\Rightarrow m_H(N) \leq N^{d_{vc}} + 1 = N + 1$

check for $N = 25$ $2^5 \nleq 5+1$ not possible.

• if $m_H(N) = 2^{\lfloor N/2 \rfloor}$ it should be $< 2^k$

min $k = 1$ as $2^{\lfloor 1/2 \rfloor} = 2^0 < 2^1 \Rightarrow d_{vc} = 0$

$\Rightarrow m_H(N) \leq N^0 + 1 = 2$

check for $N = 4$: $2^{\lfloor 4/2 \rfloor} = 2^2 \nleq 2$ not possible