$$P(|E_{in}(g) - E_{out}(g)| > \varepsilon) \leq 2Me^{-2\varepsilon^2 N}$$

M = no of hypothesis
N = no of samples

δ = confidence Interval

$$2Me^{-2\varepsilon^2 N} = \delta$$

$$e^{-2\varepsilon^2 N} = \frac{\delta}{2M}$$

$$-2\varepsilon^2 N = \ln\left(\frac{\delta}{2M}\right)$$

$$2\varepsilon^2 N = \ln\left(\frac{2M}{\delta}\right)$$

$$\varepsilon^2 = \frac{1}{2N}\ln\left(\frac{2M}{\delta}\right)$$

$$\varepsilon = \sqrt{\frac{1}{2N}\ln\left(\frac{2M}{\delta}\right)}$$

$$P(|E_{in} - E_{out}| > \varepsilon) \leq \delta$$

$$P(|E_{in} - E_{out}| \leq \varepsilon) \geq 1-\delta$$

$$|E_{in} - E_{out}| \leq \sqrt{\frac{1}{2N}\ln\left(\frac{2M}{\delta}\right)}$$

$$\Rightarrow \quad E_{out} \leq E_{in} + \sqrt{\frac{1}{2N}\ln\left(\frac{2M}{\delta}\right)}$$

with $p \geq 1-\delta$

$$E_{out} \geq E_{in} - \varepsilon$$

$$E_{out} \leq E_{in} + \varepsilon$$

$$E_{in} - \varepsilon \leq E_{out} \leq E_{in} + \varepsilon$$

$$|E_{out} - E_{in}| \leq \varepsilon$$

For $\boxed{\text{every}}$ hypothesis $h \in H$

$$|E_{in}(h) - E_{out}(h)| \leq \varepsilon$$

if $H = \{h_1, h_2, h_3, h_4\}$

h_1 $E_{in}(0.3)$
h_2 $E_{in}(0.2)$
h_3 $E_{in}(0.10)$
h_4 $E_{in}(0.25)$

our learning algorithm chooses

$$g = h_3$$

$$E_{in}(g) = 0.10$$

$$E_{out}(g) \leq E_{in}(g) + \varepsilon$$

if $\varepsilon = 0.05$

$$E_{out}(g) \leq 0.10 + 0.05$$
$$\leq 0.15$$

Can there exist an h that we did not choose whose true error is $E_{out}(h) = 0.05$

(ie) we chose the wrong hypothesis?

We know $E_{out}(h) \geq E_{in}(h) - \varepsilon$

Suppose h has higher training error than g

So $E_{in}(h) = 0.20$

$$E_{out}(h) \geq 0.20 - 0.05$$
$$\geq 0.15$$

So h cannot have a true error of 0.05

No other hypothesis $h \in H$ has $E_{out}(h)$ significantly better than $E_{out}(g)$

$\Rightarrow$ A hypothesis with worse training error cannot have a better test error

$$E_{in}(g) \leq E_{in}(h) \quad (1)$$

$$E_{out}(g) \leq E_{in}(g) + \varepsilon \quad (2)$$

$$E_{out}(h) \geq E_{in}(h) - \varepsilon \quad (3)$$

From (1) $E_{out}(g) \leq E_{in}(g) + \varepsilon \leq E_{in}(h) + \varepsilon \quad (4)$

From (3) $E_{out}(h) + \varepsilon \geq E_{in}(h)$

$$E_{in}(h) + \varepsilon \leq E_{out}(h) + 2\varepsilon$$

From (4) $E_{out}(g) \leq E_{out}(h) + 2\varepsilon$

$$E_{out}(g) \leq E_{in}(g) + \varepsilon \leq E_{in}(h) + \varepsilon \leq E_{out}(h) + 2\varepsilon$$