Problem 2.22

To prove, $E_D\left[E_{out}(g^{(D)})\right] = \sigma^2 + \text{bias} + \text{var}$

Given, without noise in the data we have,

1) $E_{out}(g^{(D)}) = E_{x,y}\left[(g^{(D)}(x) - f(x))^2\right]$ — ①.

2) we know $\text{bias}(x) = (\bar{g}(x) - f(x))^2$

3) we also know $\text{var}(x) = E_D\left[(g^{(D)}(x) - \bar{g}(x))^2\right]$

where $\bar{g}(x) = E_D\left[g^{(D)}(x)\right]$.

Now proceeding with the solution:-

$$E_D\left[E_{out}(g^{(D)})\right] = E_D\left[E_{x,y}\left[(g^{(D)}(x) - y(x))^2\right]\right]$$

$$= E_D\left[E_{x,y}\left[g^{(D)}(x)^2\right] - 2E_{x,y}\left[g^{(D)}(x)\right] E_{x,y}\left[y(x)\right] + E_{x,y}\left[y(x)^2\right]\right]$$

$$= E_{x,y}\left[E_D\left[g^{(D)}(x)^2\right]\right] - 2E_{x,y}\left[E_D\left[g^{(D)}(x)\right] E_D\left[y(x)\right]\right] + E_{x,y}\left[E_D\left[y(x)^2\right]\right]$$

$$= E_{x,y}\left[\underbrace{\left(E_D\left[g^{(D)}(x)^2\right] - \bar{g}(x)^2\right)}_{\text{part 1}} + \underbrace{\left(\bar{g}(x)^2 - 2\bar{g}(x) E_D\left[y(x)\right] + E_D\left[y(x)^2\right]\right)}_{\text{part 2}}\right]$$

$$= \text{equation } ②$$

$\therefore$ Part 1 $\Rightarrow$ $E_D\left[g^{(D)}(x)^2\right] - \bar{g}(x)^2$

$$= E_D\left[g^{(D)}(x)^2\right] - 2\bar{g}(x)^2 + \bar{g}(x)^2$$

$$= E_D\left[g^{(D)}(x)^2 - 2g^{(D)}(x)\,\bar{g}(x) + \bar{g}(x)^2\right]$$

$$= E_D\left[\left(g^{(D)}(x) - \bar{g}(x)\right)^2\right]$$

Part 2 $\Rightarrow$ $\bar{g}(x)^2 - 2\bar{g}(x)E_D[y(x)] + E_D[y(x)^2]$

$$= \bar{g}(x)^2 - 2\bar{g}(x)E_D[f(x) + \epsilon] + E_D\left[(f(x) + \epsilon)^2\right]$$

$$= \bar{g}(x)^2 - 2\bar{g}(x)E_D[f(x)] - 2\bar{g}(x)E_D[\epsilon] + E_D[f(x)^2] + 2E_D[f(x)\epsilon] + E_D[\epsilon^2]$$

$= \cancel{g(x)} - \cancel{fd.}$

$$= \left(\bar{g}(x)^2 - 2\bar{g}(x)f(x) + f(x)^2\right) - 2\bar{g}(x)E_D[\epsilon] + 2E_D[f(x)\epsilon] + E_D[\epsilon^2]$$

$$= \left(\bar{g}(x) - f(x)\right)^2 - 2\bar{g}(x)E_D[\epsilon] + 2E_D[f(x)\epsilon] + E_D[\epsilon^2]$$

$\therefore$ putting back in equation 2,

$E_D\left[E_{out}(g^{(D)})\right]$

$$= E_{x,y}\Big[E_D\big[(g^{(D)}(x) - \bar{g}(x))^2\big] + (\bar{g}(x) - f(x))^2 - 2\bar{g}(x)\,E_D[\epsilon] + 2E_D[f(x)\epsilon] + E_D[\epsilon^2]\Big]$$

$$= E_{x,y}\Big(\underbrace{E_D\big[(g^{(D)}(x) - \bar{g}(x))^2\big]}_{\text{var.}} + \underbrace{(\bar{g}(x) - f(x))^2}_{\text{bias.}} + E_D[\epsilon^2]\Big)$$

$\therefore$ putting 'var' & 'bias' now,

$= \cancel{\text{var} + \text{bias} + \sigma^2}$ ~~$\Rightarrow$ hence proved~~

$\Rightarrow$ $\cancel{E_x[\text{var}(x)] + E_{x,y}[\text{bias}(x)]} + E_{x,y}[\epsilon^2]$ $\cancel{- 2E_{x,y}[(\bar{g}(x) - f(x))\epsilon]}$

$\therefore \Rightarrow \text{var} + \text{bias} + \text{Var}[\epsilon]$ as $E_\epsilon[\epsilon] = 0$.

$$\boxed{= \text{var} + \text{bias} + \sigma^2 \quad \ldots \text{ hence proved.}}$$