Problem 2.22

References for LFD book page 63, 64.

~~Biase~~ To prove =

$$E_D(E_{out}(g^{(D)})) = \cancel{E_{x,y}}\{\, \sigma^2 + \text{bias} + \text{var}$$

∴ We know, bias $= (\bar{g}(x)^2 - f(x)^2)$ where $\bar{g}(x) = E_D[g^{(D)}(x)]$

Given,

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

where $y(x) = f(x) + \epsilon$

LHC =

$$\therefore E_{x,y}[(g^{(D)}(x) - y(x))^2]$$

$$\therefore \underbrace{E_D[E_{out}(g^{(D)})]}_{LHS} = \underbrace{E_D[E_{x,y}[(g^{(D)}(x) - y(x))^2]]}_{RHS}.$$

∴ LHS = $E_{x,y}[E_D[(g^{(D)}(x) - y(x))^2]]$

... from page 63. of LFD.

$$= E_{x,y}[(E_D[g^{(D)}(x)] - y(x))^2].$$

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

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

put $y(x) = f(x) + \epsilon$

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

~~Add~~ we know, $\bar{g}(x) = E_D\left[g^{(D)}(x)\right]$

$\therefore$ substituting with $\bar{g}(x)$,

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

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

Now, Add & subtract $\bar{g}(x)^2$.

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

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

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

$$= E_{x,y}\left[\text{Var}(x) + \text{bias}(x) - 2\bar{g}(x)(\epsilon) + 2f(x)(\epsilon) + (\epsilon)^2 \right]$$

$$= E_x\left[\text{Var}(x)\right] + E_x(\text{bias}(X)) - 2\epsilon E_x\left(\bar{g}(x) - f(x)\right) + (\epsilon)^2]$$

$$= \text{Var} + \text{bias} + \sigma^2$$

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