Machine Learning and Computational Statistics
4th Homework

Exercise 1:

(a) $E_D[(f(x;D)-E[y|x])^2] = E_D[(f(x;D)-E_D[f(x;D)])^2] + (E_D[f(x;D)]-E[y|x])^2$

Thus, in order for $E_D[(f(x;D)-E[y|x])^2]$ to be equal to zero, both variance ($E_D[(f(x;D)-E_D[f(x;D)])^2]$) and bias ($(E_D[f(x;D)]-E[y|x])^2$) need to be equal to zero.
Bias is zero when the estimated model has the exact same complexity as the model that generates the data. In order for variance to be equal to zero at the same time as the bias, the samples D need to be identical one to another, or in other words to have zero variance. This is achieved for sample size N that tends to infinity, which in realistic conditions cannot occur.

(b) In order to get MSE = 0, we need to know the exact model that generates the data, and have a number of samples that tend to infinity. Both of these are not possible in practice.

Exercise 2:

(a) Large error value on the training set indicates that we have chosen the wrong model to describe the data.

(b) Large error value on the test set may indicate that predicted model overfits. That would mean that model describes perfectly the training data but fails to generalize to other data. It could also occur simultaneously to a large training error, which would indicate that the form of the model chosen is far from reality

(c) Small error value on the training set may indicate that the model describes the mechanism that generates the data sufficiently, but it could also indicate that the model overfits (if test error is large)

(d) Small test error value may indicate that the predicted model describes sufficiently the mechanism that generates the data. Paired with a large train error, it may indicate that train and test sets are not chosen randomly

Exercise 3:

$p(x,y) = \frac{3}{2}$, $x \in (0,1)$, $y \in (x^2, 1)$

(a) In order for $p(x,y)$ to be pdf it needs to stand that:

$$\int_0^1 \int_{x^2}^1 p(x,y)\, dy\, dx = 1$$

$$\int_0^1 \int_{x^2}^1 \frac{3}{2}\, dy\, dx = \frac{3}{2} \int_0^1 \int_{x^2}^1 1\, dy\, dx = \frac{3}{2} \int_0^1 [y]_{x^2}^1\, dx =$$

$$\frac{3}{2} \int_0^1 (1 - x^2)\, dx = \frac{3}{2} \left[x - \frac{x^3}{3}\right]_0^1 = \frac{3}{2}\left(1 - \frac{1}{3}\right) = \frac{3}{2} \cdot \frac{2}{3} = 1$$

Thus $p(x,y)$ is pdf.

(b) $$p_x(x) = \int_{x^2}^1 p(x,y)\, dy = \int_{x^2}^1 \frac{3}{2}\, dy = \int_{x^2}^1 \left(\frac{3}{2}y\right)' dy = \left[\frac{3}{2}y\right]_{x^2}^1 = \frac{3}{2} - \frac{3}{2}x^2 \Rightarrow$$

$$\Rightarrow p_x(x) = \frac{3}{2}(1 - x^2)$$

(c) $$p(y|x) = \frac{p(x,y)}{p_x(x)} = \frac{3/2}{3/2(1-x^2)} = \frac{1}{1-x^2}$$

(d) $$E[y|x] = \int_{x^2}^1 y \cdot p(y|x)\, dy = \int_{x^2}^1 \frac{1}{1-x^2}\, y\, dy = \frac{1}{1-x^2}\left[\frac{y^2}{2}\right]_{x^2}^1 =$$

$$= \frac{1}{1-x^2}\left(\frac{1}{2} - \frac{x^4}{2}\right) = \frac{1}{\cancel{1-x^2}} \cdot \frac{(\cancel{1-x^2})(1+x^2)}{2} = \frac{1}{2} + \frac{1}{2}x^2$$