Training

training data set

$$\{(y_1, x_1,), \ldots, (y_n, x_n))\}$$

used to find function q that minimizes

Training MSE

$$\hat{f} = \arg\min_q MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - q(x_i))^2$$

Testing

testing data sets (unseen)

$$(y_0, x_0)$$

used to compute **Test MSE**

$$E[y_0 - \hat{f}(x_0)^2]$$

often not so closely related

Training

training data set

$$\{(y_1, x_1,), \ldots, (y_n, x_n))\}$$

used to find function q that minimizes **Training MSE**

$$\hat{f} = \arg\min_q MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - q(x_i))^2$$

Testing

testing data sets (unseen)

$$(y_0, x_0)$$

used to compute **Test MSE**

$$E[y_0 - \hat{f}(x_0)^2]$$

often not so closely related

Training and Testing