# M339G: September 9th, 2024.

## Bias.

**Def'n.** Let $\hat{\theta}$ be a point estimator for the parameter $\theta$. The **bias** of the estimator $\hat{\theta}$ is defined as

$$\text{bias}(\hat{\theta}) := \mathbb{E}_\theta[\hat{\theta}] - \theta$$

**Def'n.** The **mean squared error** of $\hat{\theta}$ is defined as

$$\text{MSE}[\hat{\theta}] := \mathbb{E}_\theta[(\hat{\theta}-\theta)^2]$$

## Assessing Model Accuracy.

Say, we have the "usual" model: $Y = f(X) + \varepsilon$

Say, we fit our model to some training data:

$$Tr = \{(x_i, y_i) : i = 1, \dots, N\}$$

Let $\hat{f}$ be the fit of the model to our $Tr$

$$\text{MSE}_{Tr} := \underset{i \in \{1,\dots,N\}}{\text{Ave}} (y_i - \hat{f}(x_i))^2 = \frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{f}(x_i))^2$$

We propose to look @ other data

$$Te = \{(x_i, y_i) : i = 1, \dots, M\}$$

These are our testing data.

We calculate:

$$\text{MSE}_{Te} := \underset{i \in \{1,\dots,M\}}{\text{Ave}} (y_i - \hat{f}(x_i))^2 = \frac{1}{M}\sum_{i=1}^{M} (y_i - \hat{f}(x_i))^2$$