

# 16.2) K-Nearest Neighbors

Vitor Kamada

January 2020

Tables, Graphics, and Figures from

James et al. (2017): Chapters: 2.2.3, 3.5

Hastie et al. (2017): Chapters: 2.3

# Training Error and Test Error

$$Y = f(X) + \epsilon$$

$$\hat{Y} = \hat{f}(X)$$

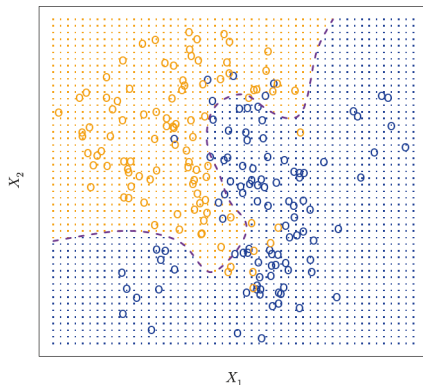
$$\{(x_1, y_1), \dots, (x_n, y_n)\}$$

$$\frac{1}{n} \sum_{i=1}^n I(y_i \neq \hat{y}_i)$$

$$\text{Ave}(I(y_0 \neq \hat{y}_0))$$

# Bayes Classifier and Error Rate

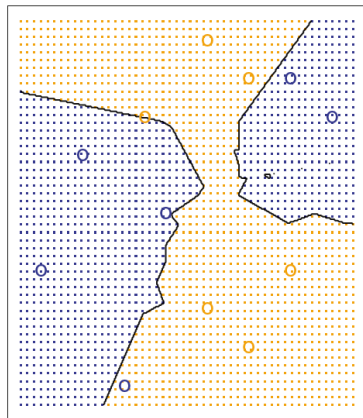
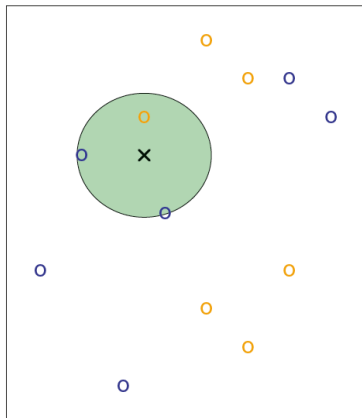
$$Pr(Y = j|X = x_0)$$



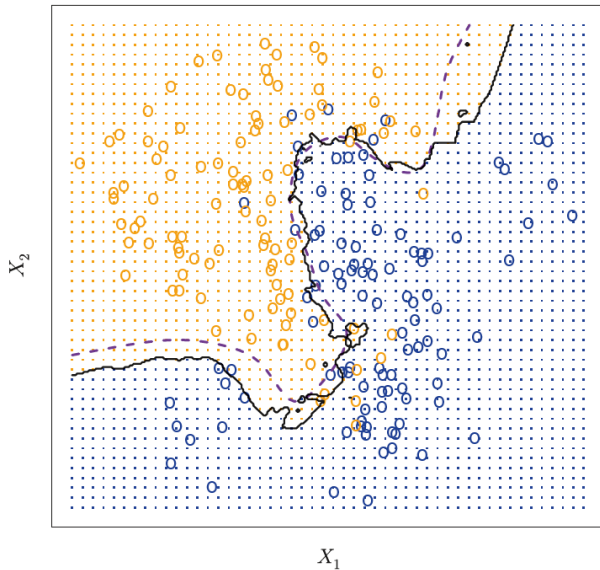
$$1 - E[\max_j Pr(Y = j|X)] = 0.13$$

# K-Nearest Neighbors (K=3)

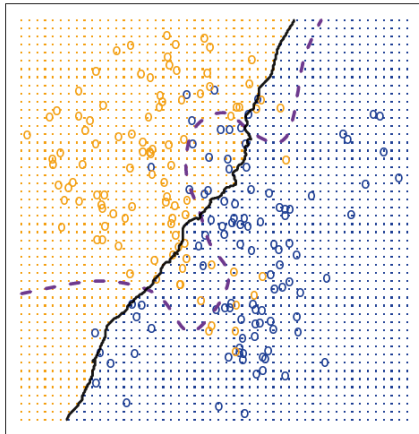
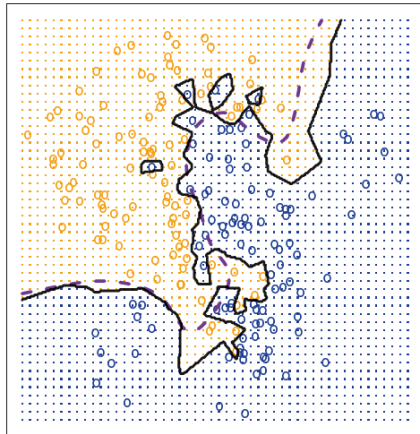
$$Pr(Y = j|X = x_0) = \frac{1}{K} \sum_{i \in N_0} I(y_i = j)$$



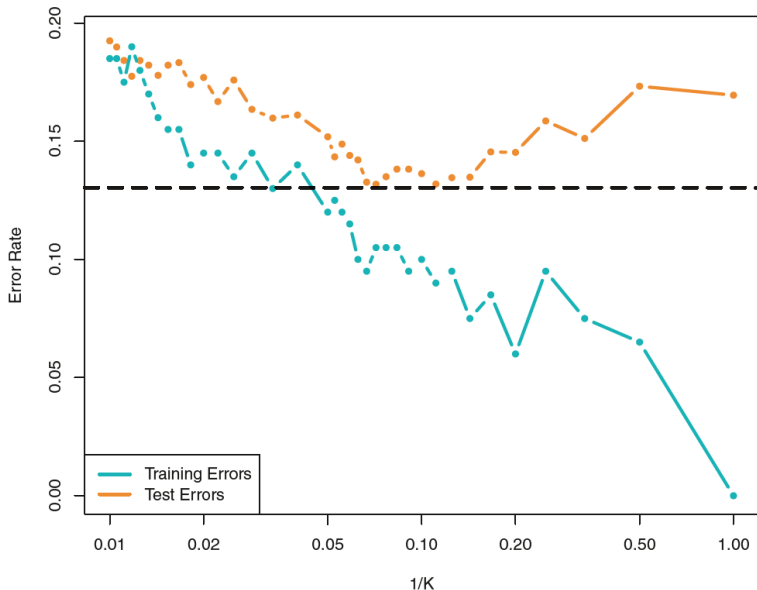
# KNN: K=10



# KNN: $K=1$ and $K=100$



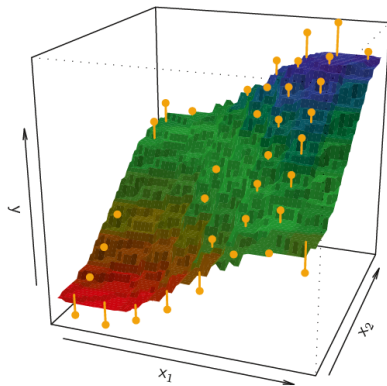
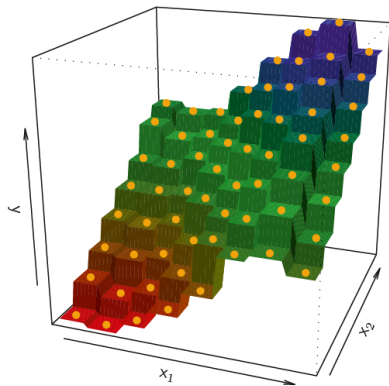
# KNN Training and Test Error Rate



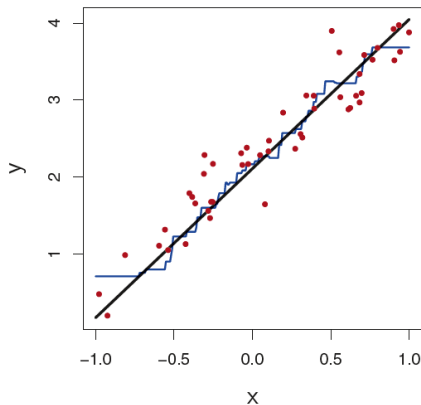
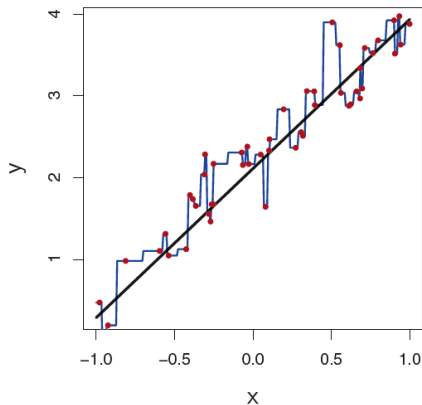


# KNN Regression: K=1 and K=9

$$\hat{f}(x_0) = \frac{1}{K} \sum_{x_i \in N_0} y_i$$



# One-dimension KNN Regression: $K=1$ and $K=9$



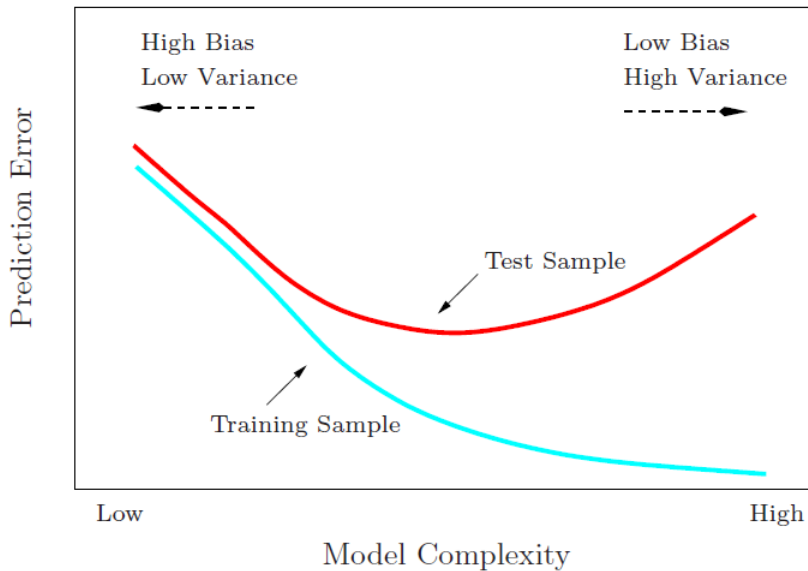
# Mean Squared Error (MSE)

$$\begin{aligned} E(Y - \hat{Y})^2 &= E[f(X) + \epsilon - \hat{f}(X)]^2 \\ &= E[f(X) - \hat{f}(X)]^2 + \text{Var}(\epsilon) \end{aligned}$$

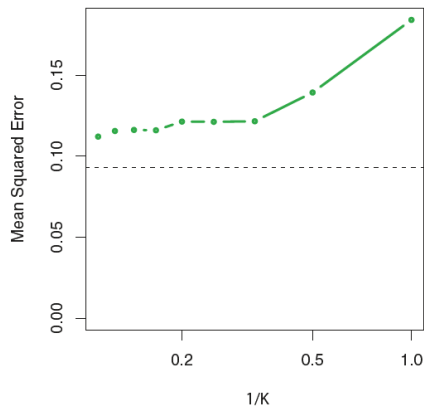
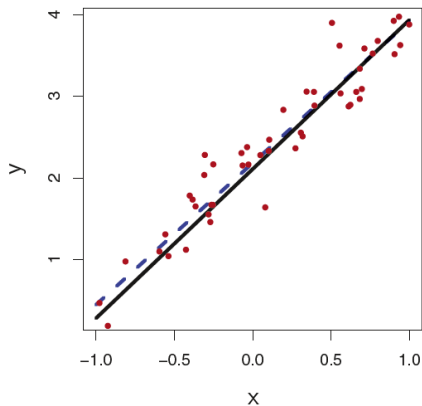
$$MSE(x_0) = E_{\tau}[f(x_0) - \hat{y}_0]^2$$

$$\begin{aligned} E_{\tau}[\hat{y}_0 - E_{\tau}(\hat{y}_0)]^2 &+ [E_{\tau}(\hat{y}_0) - f(x_0)]^2 \\ \text{Var}_{\tau}(\hat{y}_0) &+ [\text{Bias}(\hat{y}_0)]^2 \end{aligned}$$

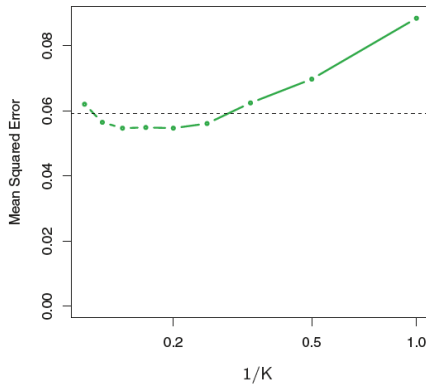
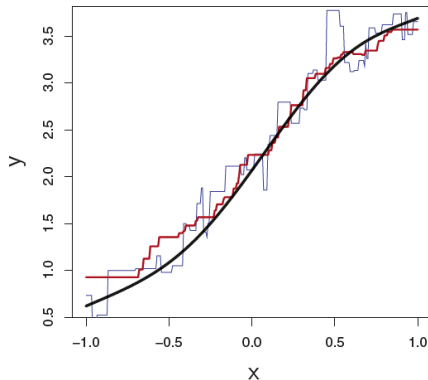
# Bias-Variance Tradeoff



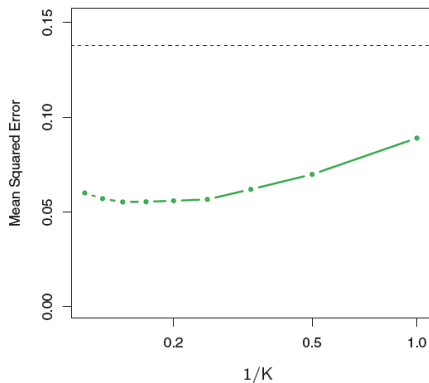
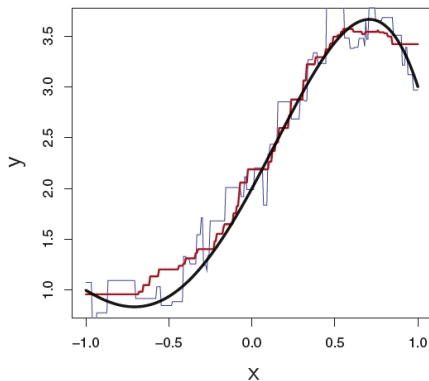
# MSE: OLS vs KNN



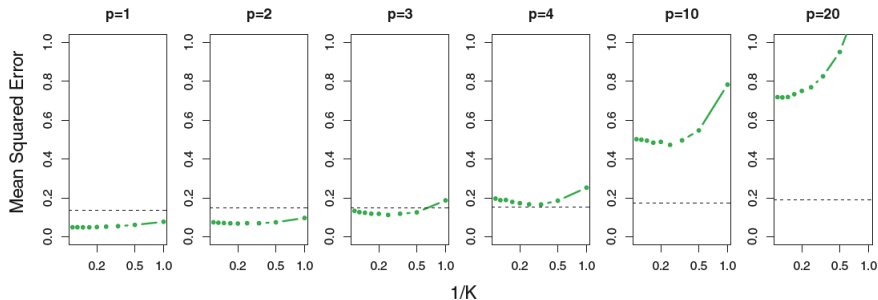
# Slightly Non-Linear Relationship



# Strongly Non-Linear Relationship

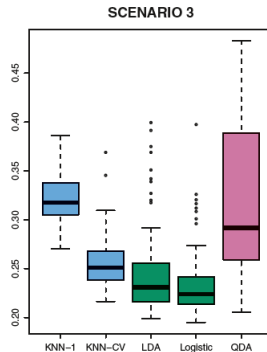
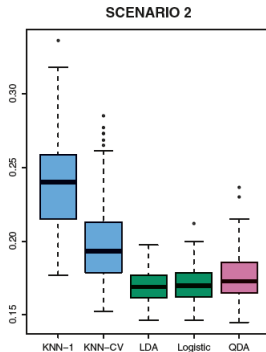
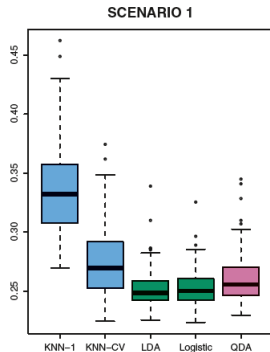


# Additional Noise Variables





# Test Error Rates: Linear Scenarios



# Test Error Rates: Non-Linear Scenarios

