Week 3 Reading Guide – LDA & Naive Bayes

**Why would we consider using methods other than logistic regression for classification problems?**

## Linear Discriminant Analysis for

In the formula

**What does** represent?

**What does** represent?

**How do we estimate** ?

**What classifier has the lowest error rate of all classifiers?**

**Is linear discriminant analysis a parametric or nonparametric method?**

**Suppose there is no information for estimating** . In this instance, what value does LDA use for ?

## Linear Discriminant Analysis for

**If the two predictor variables are uncorrelated, how does the multivariate Gaussian distribution look?**

**What is a “null classifier”?**

**In what scenarios does a null classifier perform well?**

**What is sensitivity?**

**What is specificity?**

**Fill in the table below with the following terms:** *sensitivity, specificity, Type I error, Type II error*

|  |  |  |
| --- | --- | --- |
|  | **Actual** | |
| **Predicted** | No | Yes |
| No |  |  |
| Yes |  |  |

## Quadratic Discriminant Analysis

**Which classifier (LDA, QDA) is more flexible?**

**Which classifier (LDA, QDA) has lower variance?**

**Suppose you have few training observations, which method (LDA, QDA) would be a better choice?**

**Suppose you have a large training dataset, which method (LDA, QDA) would be a better choice?**

## An Empirical Comparison

**In what scenarios does logistic regression perform the best?**

**In what scenarios does LDA perform the best?**

**In what scenarios does QDA perform the best?**

**In what scenarios does Naive Bayes perform the best?**

**In what scenarios does KNN (with CV) perform the best?**