

Introduction

Frank Wood

July 6, 2010

Introduction

Overview of Topics

Introduction

- ▶ Data mining is the search for patterns in large collections of data
- ▶ Pattern recognition is concerned with automatically finding patterns in data
- ▶ Machine learning is pattern recognition with concern for computational tractability

Example : Handwritten Digit Identification

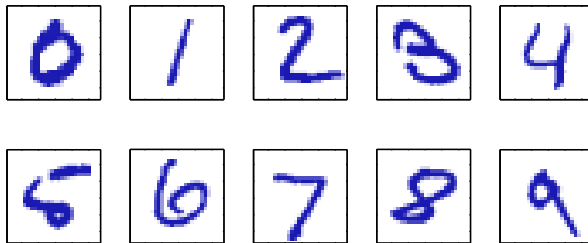


Figure: Hand written digits from the USPS

Approaches to Identifying Digits

Goal

- ▶ Build a machine that can identify digits automatically

Approaches

- ▶ Hand craft a set of rules that separate each digit from the next
- ▶ Set of rules invariably grows large and unwieldy and requires many “exceptions”
- ▶ “Learn” a set of models for each digit automatically from labeled training data.

Formalism

- ▶ Each digit is 28x28 pixel image
- ▶ Vectorized into a 784 entry vector \mathbf{x}

Machine learning approach

- ▶ Obtain a of N digits $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ called the *training set*.
- ▶ Label (by hand) the training set to produce a label or “target” \mathbf{t} for each digit image \mathbf{x}
- ▶ Learn a function $\mathbf{y}(\mathbf{x})$ which takes an image \mathbf{x} as input and returns an output in the same “format” as the target vector.

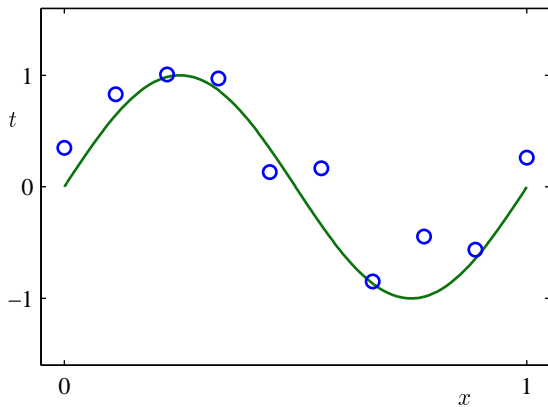


Figure: Chapter 1, Figure 2

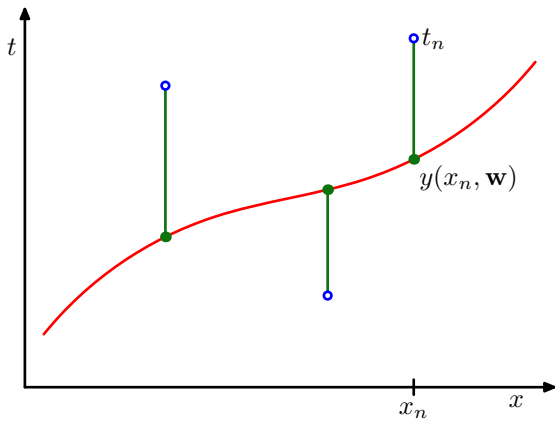


Figure: Chapter 1, Figure 3

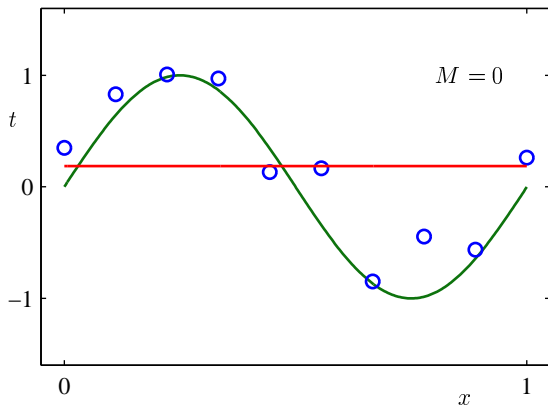


Figure: Chapter 1, Figure 4a

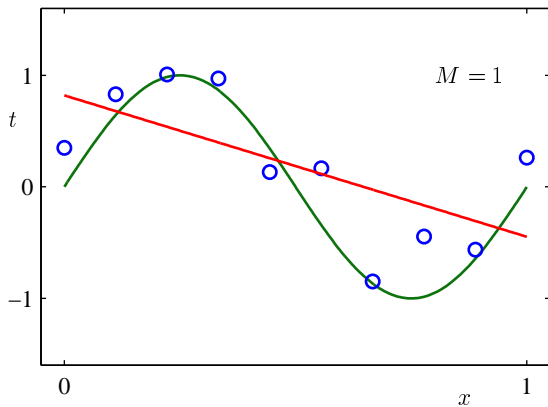


Figure: Chapter 1, Figure 4b

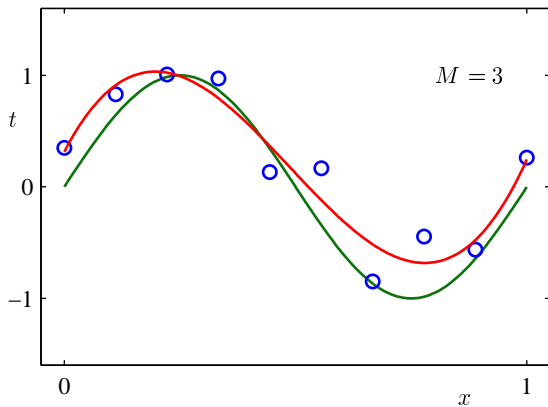


Figure: Chapter 1, Figure 4c

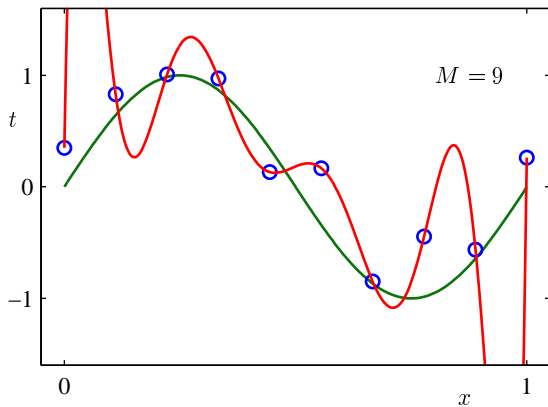


Figure: Chapter 1, Figure 4d

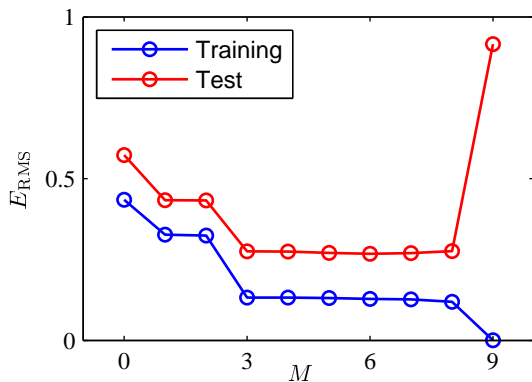


Figure: Chapter 1, Figure 5

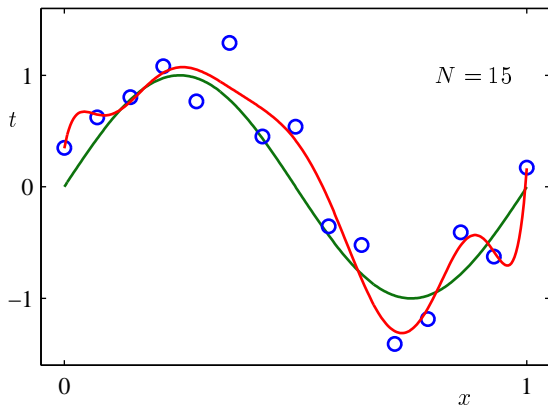


Figure: Chapter 1, Figure 6a

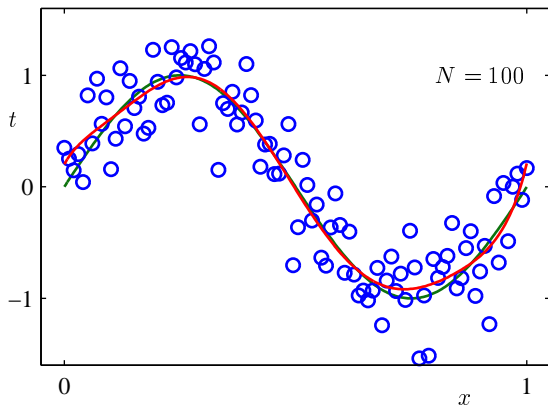


Figure: Chapter 1, Figure 6b

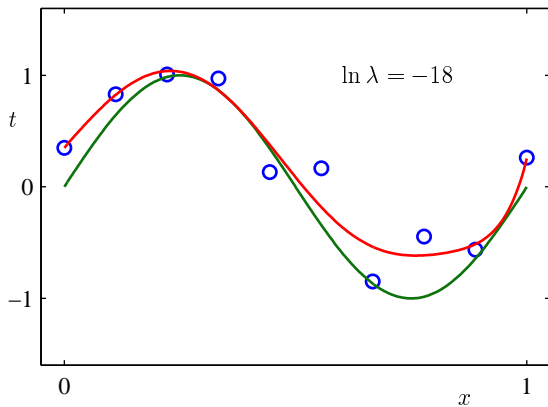


Figure: Chapter 1, Figure 7a

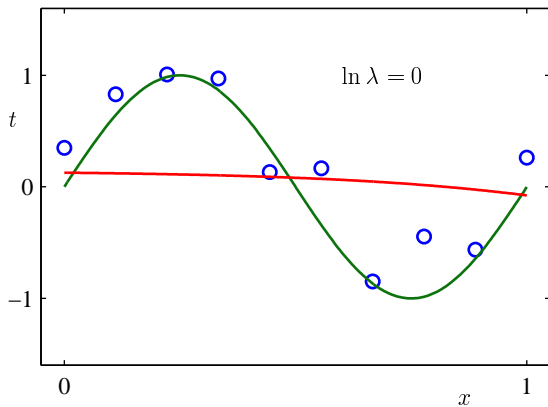


Figure: Chapter 1, Figure 7b

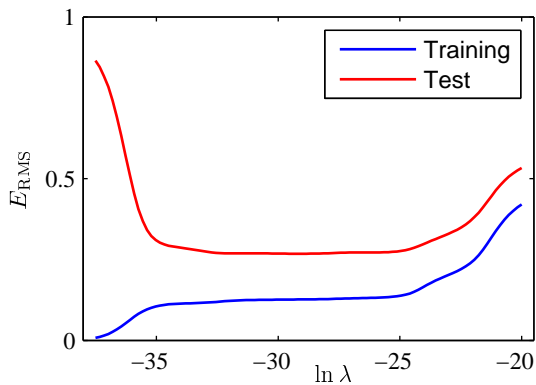


Figure: Chapter 1, Figure 8

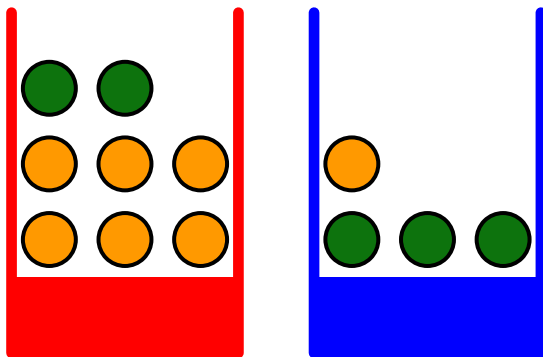


Figure: Chapter 1, Figure 9

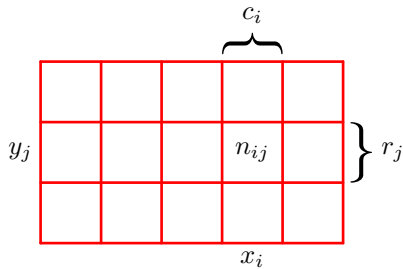


Figure: Chapter 1, Figure 10

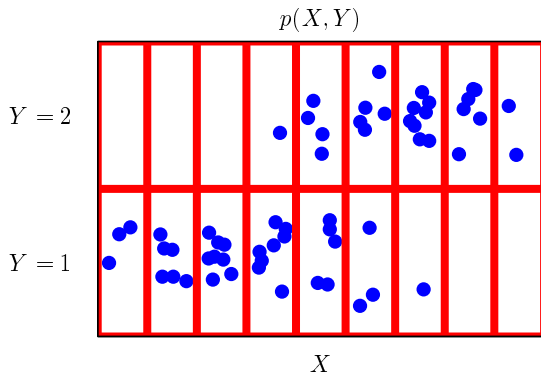


Figure: Chapter 1, Figure 11a

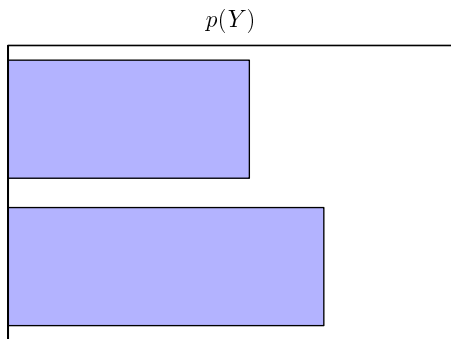


Figure: Chapter 1, Figure 11b

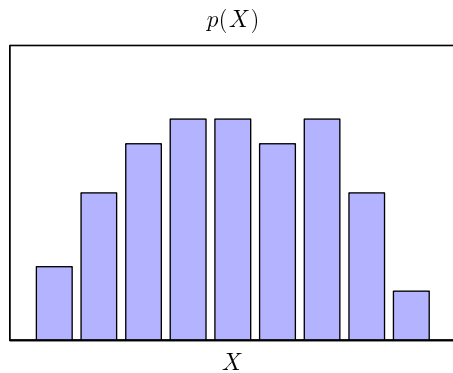


Figure: Chapter 1, Figure 11c

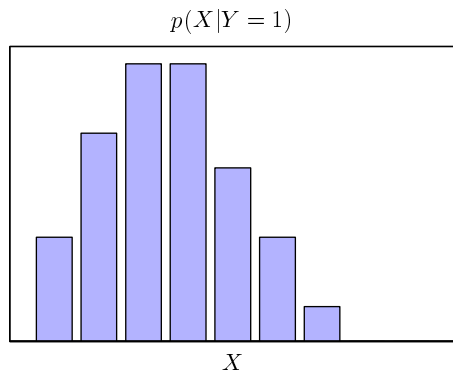


Figure: Chapter 1, Figure 11d

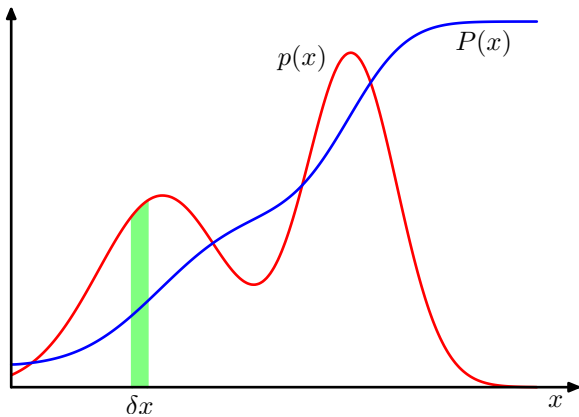


Figure: Chapter 1, Figure 12

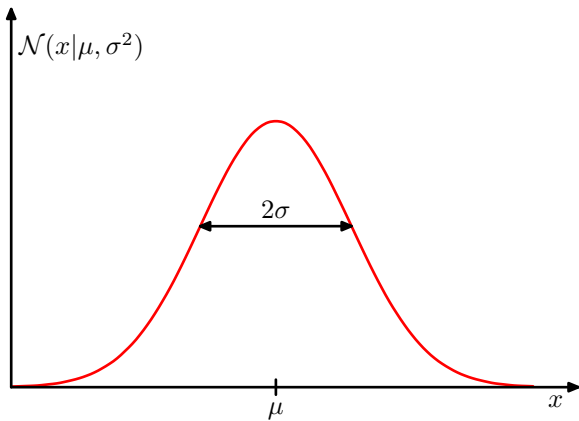


Figure: Chapter 1, Figure 13

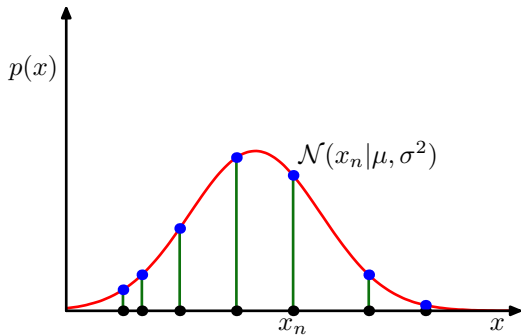


Figure: Chapter 1, Figure 14

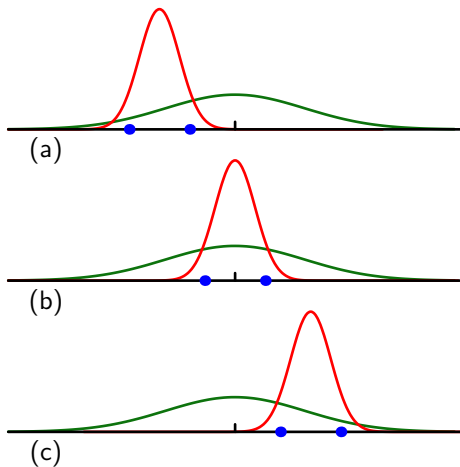


Figure: Chapter 1, Figure 15

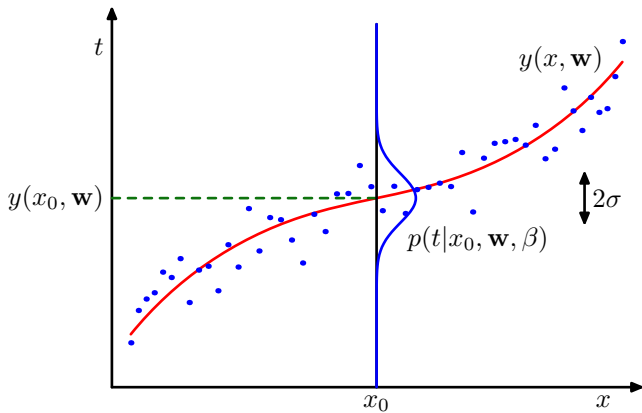


Figure: Chapter 1, Figure 16

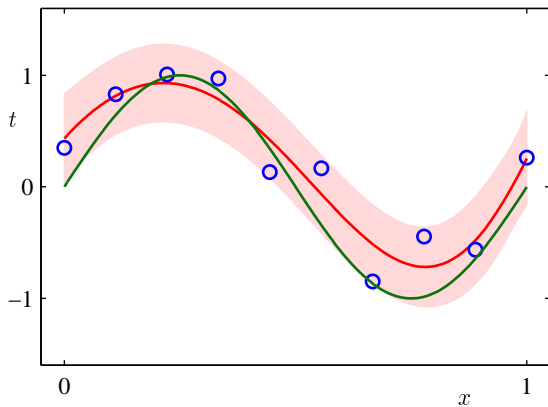


Figure: Chapter 1, Figure 17

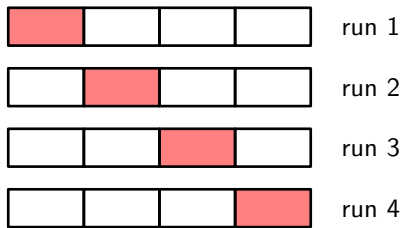


Figure: Chapter 1, Figure 18

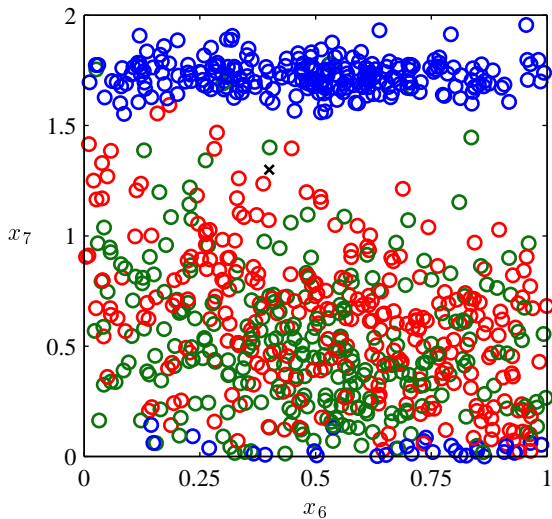


Figure: Chapter 1, Figure 19

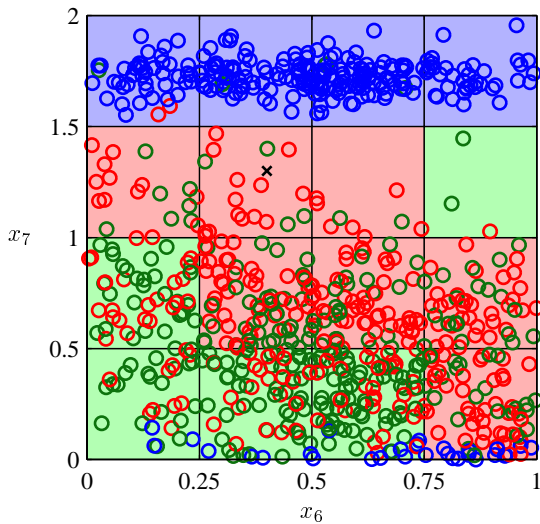


Figure: Chapter 1, Figure 20

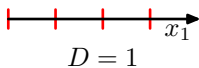
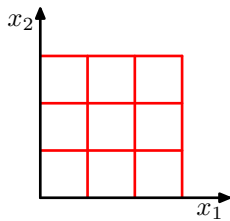
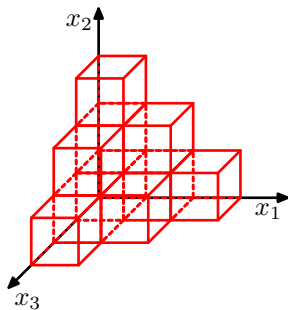


Figure: Chapter 1, Figure 21a



$$D = 2$$

Figure: Chapter 1, Figure 21b



$$D = 3$$

Figure: Chapter 1, Figure 21c

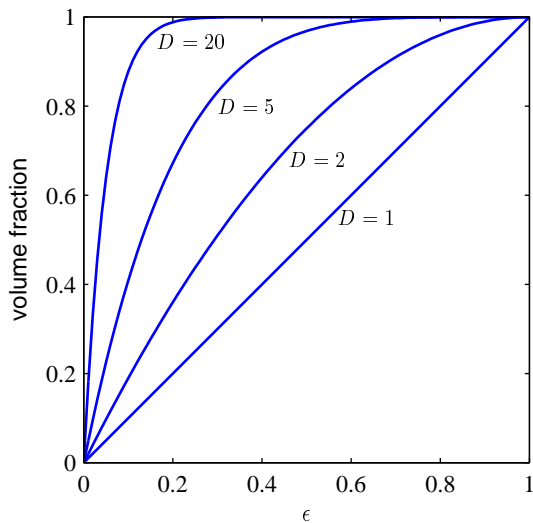


Figure: Chapter 1, Figure 22

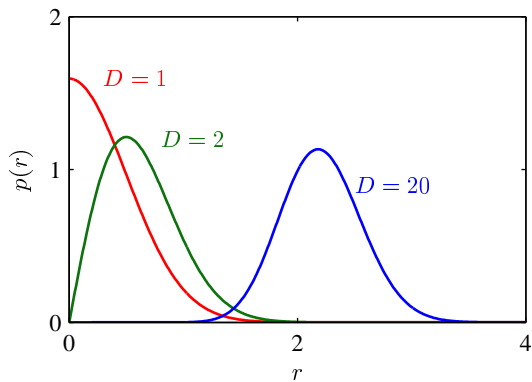


Figure: Chapter 1, Figure 23

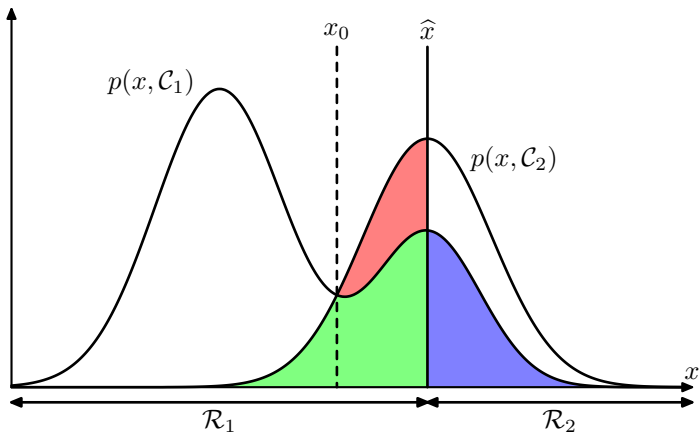


Figure: Chapter 1, Figure 24

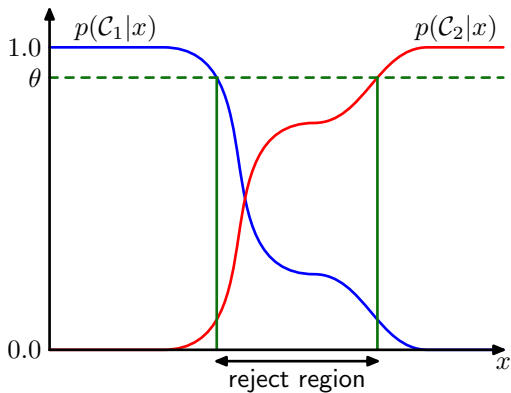


Figure: Chapter 1, Figure 26

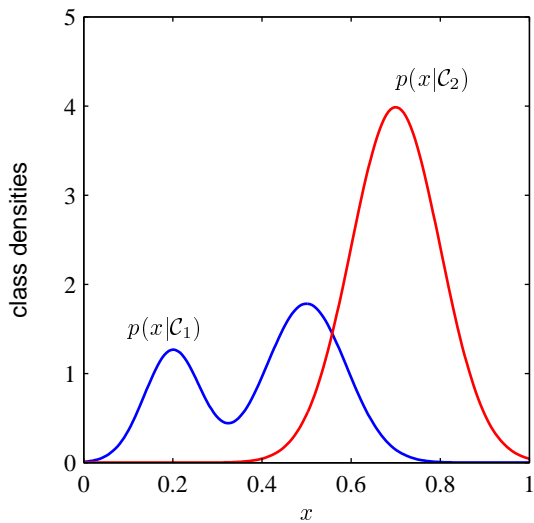


Figure: Chapter 1, Figure 27a

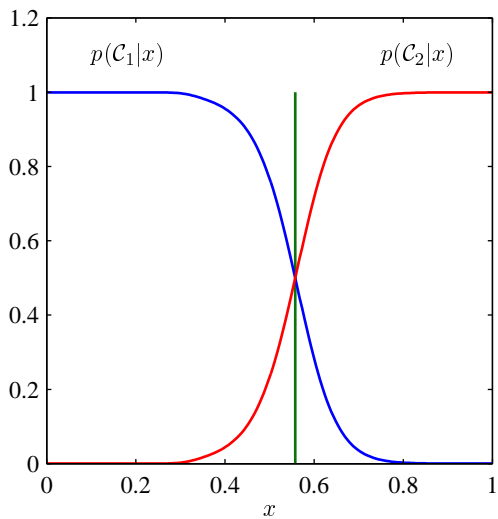


Figure: Chapter 1, Figure 27b

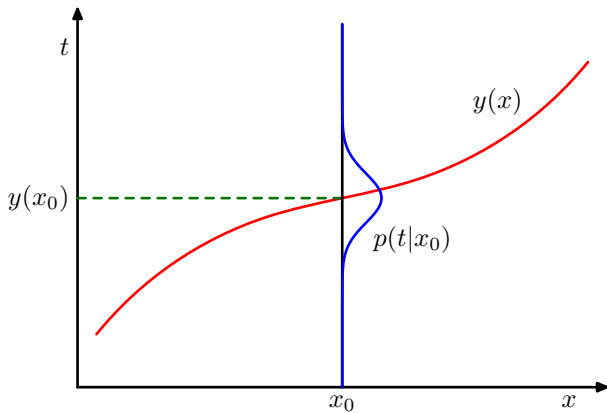


Figure: Chapter 1, Figure 28

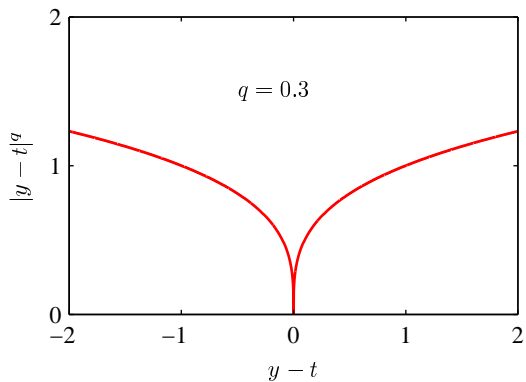


Figure: Chapter 1, Figure 29a

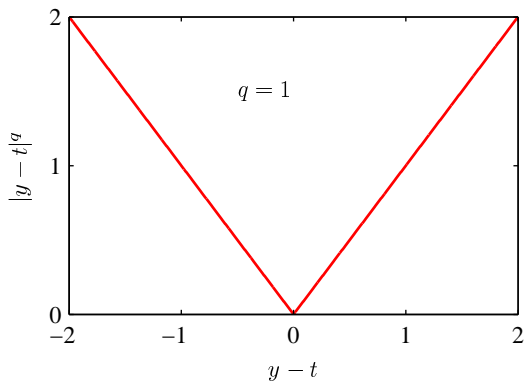


Figure: Chapter 1, Figure 29b

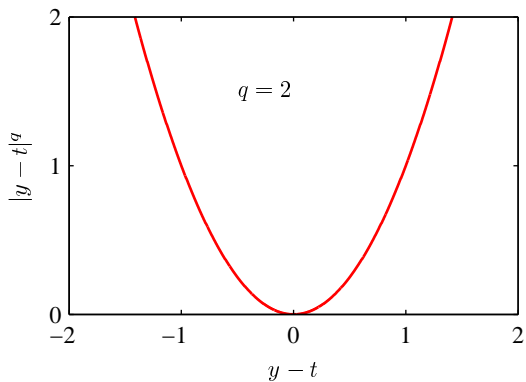


Figure: Chapter 1, Figure 29c

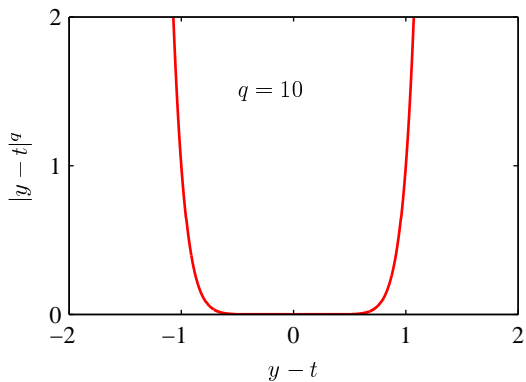


Figure: Chapter 1, Figure 29d

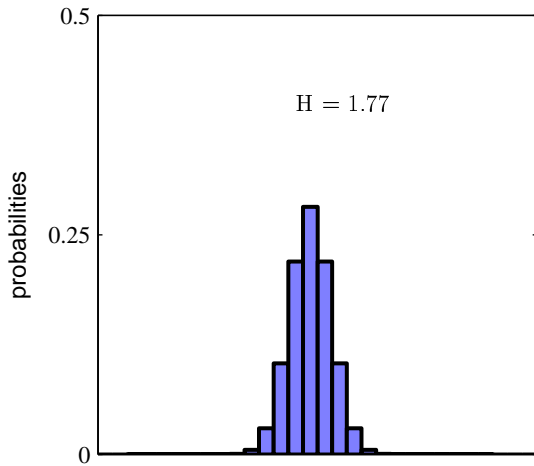


Figure: Chapter 1, Figure 30a

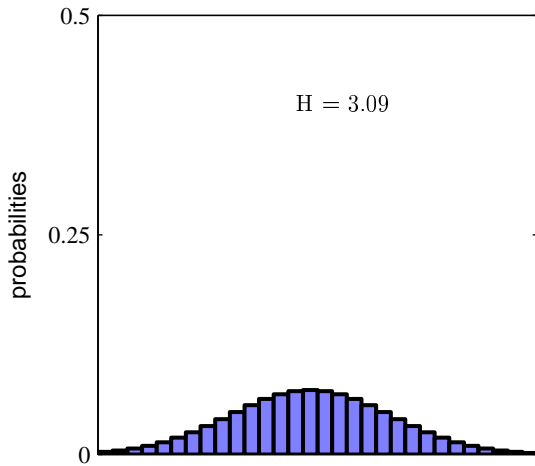


Figure: Chapter 1, Figure 30b

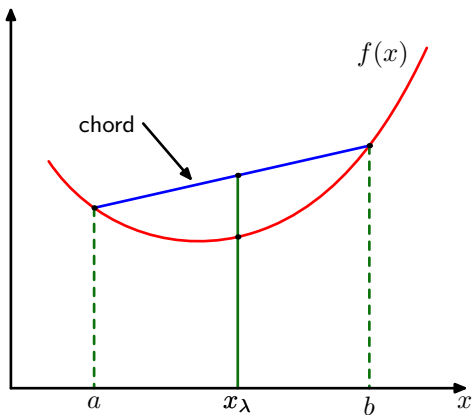


Figure: Chapter 1, Figure 31