PRML Chapter 2. Probability Distribution by Joohyung Lee

Introduction

Chapter Scope

1) Examples of probability distributions
2) Their properties

Purpose of Introducing Distributions

1) a building blocks for more complex models
2) a recipe to discuss some essential statistical concept, e.g., Bayesian inference
3) to model the probability distribution p(**x**), i.e., density estimation

* Model Selection becomes an issue since density estimation is fundamentally ill-posed problem in that infinitely many distributions can fit the observed data set.

Parametric distribution vs. Non-Parametric distribution

Parametric Distribution:

1) binomial distribution, multinomial distribution, Gaussian distribution (continuous R.V.)
2) For density estimation, the parameters shall be determined with an observed data set.
   A. Frequentist: specific values for parameters (earned by optimizing some criterion, e.g., likelihood function)
   B. Bayesian: estimate posterior distribution with introduced prior distributions over the parameters as well as the observed data
3) Conjugate Priors: To simplify the Bayesian analysis, use conjugate prior which let posterior distribution be in the same form of prior distribution.
4) Exponential family of distributions is presented as it possesses a number of important properties.

Non-Parametric Distribution

1) Form of the distribution is not forced by a user but typically depends on the size of the data set
2) Still contains the parameters but they do not determine the distribution form but the complexity
3) Histogram, nearest-neighbors, kernels

Table. 1 Conjugate prior with posterior distribution in exponential family

Conjugate Prior	**Posterior Distribution**
Dirichlet distribution	Multinomial distribution
Gaussian distribution	Gaussian distribution

Binary Variables