# Research Class 4

January 6, 2018

The materials covered in this class are:

1. Introduction: Bayesian statistics.

## 1 Bayes Rule

Bayes rule is a very important theorem in the probability theory. It explains how the probability can change given some information. Let's start with the Bayes Formula and an example:

Bayes Formula:

$$P(B \mid A) = \frac{P(A \mid B)\, P(B)}{P(A)}$$

This can be derived from the definition of the conditional probability. We say the probability of event $A$ occurring given the event $B$ occurring is $P(A \mid B) = P(A \cdot B)/P(B)$, where $P(A \cdot B)$ means the probability of both $A$ and $B$ occurring. It can be easily illustrated the following Venn diagram.

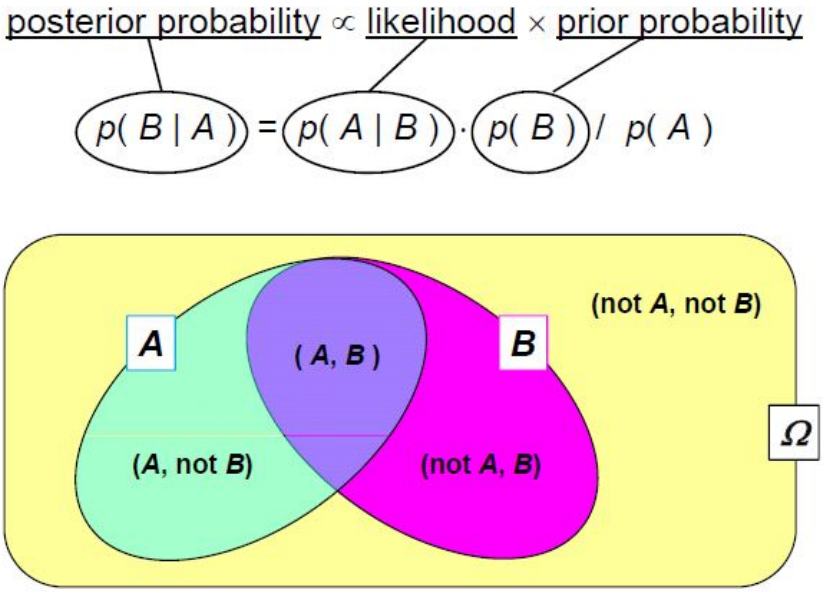


Figure 1: Venn diagram of the Bayes rule

## 2 Parametric Modeling and Likelihood Function

In the past a few classes, we discussed the linear regression model and the logistic regression model. In those models, we basically are modeling the relationship between a response and a set of independent variables. In those models, we were using hand-derivation and python codes to fit parameters and describe such a quantitative relationship.

In this class we will generalize the abstract problem. Let's denote the parameters to fit as $\theta$. For example, in the linear regression model, the functional form between the independent variables and the dependent variable is a linear function:

$$y = X\beta + \epsilon.$$

The inference problem is to calculate the parameters $\theta = (\beta_0, \beta_1, \ldots, \beta_m)$. Let's forget the derivation of $\beta$ for linear regression for a second. Now the $\theta$ is an unknown $m$-dimensional vector. Intuitively, given a set of $X, y$, if we want to fit a quantitative model, we would like to have the probability of this quantitative model correctly explain the relationship as high as possible. To develop this idea further, we need to define a function as this probability. This is the likelihood function, usually denoted by $f(x \mid \theta)$, which means how likely my observation is $X = x$ given I have selected my $\theta$ this way.

# 3 Basic Idea of Bayesian Learning