Machine Learning - Probability Review

Joshua Saunders

January 2018

1 Introduction

Note: add diagram from notes.

All a_i are *disjoint* and *exhaustive* and b_j is an event.

1.1 Total Probability

The *total probability*, $P(b_j)$ is defined as

$$P(b_j) = \sum_{i}^{n} P(A, B) \tag{1}$$

where there are n events, b_j is an event, and $P(A, B)$ is the probability of the intersection of A and B.

1.2 Conditional Probability

In English, a conditional probability takes on a form similar to

> The probability of A *given* B...

The mathematical equivalent of the above statement is

$$P(A \mid B) = \ldots$$

and the full mathematical formula for conditional probability is

$$P(A \mid B) = \frac{P(A, B)}{P(B)} \tag{2}$$

The intuition behind (2) is quite simple. Bonilla states that it is simply a way to change from the universe of all possibilities (U) to that in which the events A and B intersect [1]. This lets us ask

> What is the probability of A occured *given* that B also occured.

1.3 Bayes' Theory

$$P(A \mid B) = \frac{P(A\,, B)}{P(B)} \tag{3}$$

References

[1] O. Bonilla, "Visualizing bayes theorem." `https://oscarbonilla.com/2009/05/visualizing-bayes-theorem/`, 2009. [Online; accessed 31-January-2018].