Useful Probability Prerequisites

Notes on Bayesian Statistics

Devere Anthony Weaver

---

1 Conditional Probability and Independence of Events

The **conditional probability** of an event is the probability (relative frequency interpretation) of the event given the fact that one or more events has already occurred.

Definition 1.1 (Conditional Probability). *The conditional probability of an event A, given that an event B has occurred, is equal to*

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \;\; P(B) > 0.$$

Suppose that the probability of the occurrence of A is unaffected by the occurrence or nonoccurence of event B. When this happens, we would be inclined to say that events A and B are **independent**.

Definition 1.2 (Independent). *Two events A and B are said to be independent if any one of the following holds:*

1. $P(A|B) = P(A)$,
2. $P(B|A) = P(B)$,
3. $P(A \cap B) = P(A)P(B)$.

Otherwise, the events are said to be dependent.

N.B.

- Utilize one of three conditions in the definition to test for independence of events.

---

2 Two Laws of Probability

The following two theorems can be used to obtain the probability of the intersection and union of two events.

Theorem 2.1 (The Multiplicative Law of Probability). *The probability of the intersection of two events A and B is*

$$P(A \cap B) = P(A)P(B|A)$$

$$= P(B)P(A|B).$$

If the two events are independent then,

$$P(A \cap B) = P(A)P(B).$$

Theorem 2.2 (The Additive Law of Probability). *The probability of the union of two events A and B is*

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

If the two events are mutually exclusive then,

$$P(A \cup B) = P(A)P(B).$$

3 The Law of Total Probability and Bayes' Rule

Theorem 3.1 (Law of Total Probability)**.** *Assume that* $\{B_1, B_2, \ldots, B_k\}$ *is a partition of sample space* S *such that* $P(B_i) > 0, i = 1...k$. *Then for any event* A,

$$P(A) = \sum_{i=1}^{k} P(A|B_i)P(B_i).$$

In other words, the probability of an event A is equal to the sum of each intersection of A and B_i. To visualize this, think of a sample space S that is made up of each piece B_i, then A is some arbitrary event that happens to be in the sample space S. A is overlapping in each of the B_i but not in their entirety so their intersection is the portion of A that lies within the subspace B_i of S.

This theorem can be useful in determining $P(A)$ when it is easier to compute $P(A|B_i)$ than to compute $P(A)$ directly.

Theorem 3.2 (Bayes' Rule)**.** *Assume that* $\{B_1, B_2, \ldots, B_k\}$ *is a partition of* S *such that* $P(B_i) > 0, i = 1, 2, \ldots, k$. *Then*

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{k} P(A|B_i)P(B_i)}.$$

This seems confusing at first, but what it appears to be saying is that we want to know what the conditional probability of B_j is given that event A has occurred. Recall that B_j is some arbitrary subspace of the sample space S. Thus if we know the elements in the numerator and can compute the total probability of A in the denominator, then we can "reverse" the probability to find the probability of this subspace B_j.