CSCI 2540: PROBABILISTIC METHODS IN COMPUTER SCIENCE (NOTES)

ALEXANDER LEE

1. Events and Probability

1.2. Axioms of Probability.

Definition. *A probability space has three components:*

1. *a sample space Ω, which is the set of all possible outcomes of the random process modeled by the probability space;*
2. *a family of sets $\mathcal{F}$ representing the allowable events, where each set in $\mathcal{F}$ is a subset of the sample space Ω; and*
3. *a probability function $\Pr : \mathcal{F} \to \mathbb{R}$ satisfying the following definition.*

Definition. *A probability function is any function $\Pr : \mathcal{F} \to \mathbb{R}$ that satisfies the following conditions:*

1. *for any event E, $0 \le \Pr(E) \le 1$;*
2. *$\Pr(\Omega) = 1$; and*
3. *for any finite or countably infinite sequence of pairwise mutually disjoint events $E_1, E_2, E_3, \ldots$,*

$$\Pr\left(\bigcup_{i \ge 1} E_i\right) = \sum_{i \ge 1} \Pr(E_i).$$

Lemma. *For any two events E_1 and E_2,*

$$\Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2) - \Pr(E_1 \cap E_2).$$

Lemma (Union Bound)**.** *For any finite or countably infinite sequence of events $E_1, E_2, \ldots$,*

$$\Pr\left(\bigcup_{i \ge 1} E_i\right) \le \sum_{i \ge 1} \Pr(E_i).$$

Lemma (1.3)**.** *Let $E_1, \ldots, E_n$ be any n events. Then*

$$\Pr\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} \Pr(E_i) - \sum_{i<j} \Pr(E_i \cap E_j) + \sum_{i<j<k} \Pr(E_i \cap E_j \cap E_k)$$
$$- \cdots + (-1)^{l+1} \sum_{i_1 < i_2 < \cdots < i_l} \Pr\left(\bigcap_{r=1}^{l} E_{i_r}\right) + \cdots .$$

Definition. *Two events E and F are independent if and only if*

$$\Pr(E \cap F) = \Pr(E) \cdot \Pr(F).$$

More generally, events $E_1, E_2, \ldots, E_k$ *are mutually independent if and only if, for any subset* $I \subseteq [1, k]$,

$$\Pr\left(\bigcap_{i \in I} E_i\right) = \prod_{i \in I} \Pr(E_i).$$

Definition. *The conditional probability that event* E *occurs given that event* F *occurs is*

$$\Pr(E \mid F) = \frac{\Pr(E \cap F)}{\Pr(F)}.$$

The conditional probability is well-defined only if $\Pr(F) > 0$.

Theorem (Law of Total Probability)**.** *Let* $E_1, E_2, \ldots, E_n$ *be mutually disjoint events in the sample space* Ω*, and let* $\bigcup_{i=1}^{n} E_i = \Omega$*. Then*

$$\Pr(B) = \sum_{i=1}^{n} \Pr(B \cap E_i) = \sum_{i=1}^{n} \Pr(B \mid E_i) \Pr(E_i).$$

Theorem (Bayes' Law)**.** *Assume that* $E_1, E_2, \ldots, E_n$ *are mutually disjoint events in the sample space* Ω *such that* $\bigcup_{i=1}^{n} E_i = \Omega$*. Then*

$$\Pr(E_j \mid B) = \frac{\Pr(E_j \cap B)}{\Pr(B)} = \frac{\Pr(B \mid E_j) \Pr(E_j)}{\sum_{i=1}^{n} \Pr(B \mid E_i) \Pr(E_i)}.$$