# **Section 1.5.** The Definition of Probability

**Note.** In this section, we assume a probability has been assigned to each event in a sample space. We require that these probabilities satisfy three axioms and then prove some resulting properties of probability.

**Note.** When dealing with the union of disjoint (i.e., mutually exclusive) events, we indicate the disjointness with a special union symbol: $\mathbin{\dot\cup}$.

**Note.** We denote the probability of event $A$ as $\Pr(A)$. As mentioned above, we assume function $\Pr : \mathcal{A}(S) \to \mathbb{R}$ (that is, Pr is defined on a $\sigma$-algebra on the sample space $S$). We impose the following three axioms:

**Axiom 1. Axiom of Non-Negativity.** For any event $A$, $\Pr(A) \geq 0$.

**Axiom 2. Axiom of Total Probability.** For $S$ the sample space, $\Pr(S) = 1$.

**Axiom 3. Axiom of Countable Additivity.** For any infinite sequence of disjoint events $A_1, A_2, \ldots$ we have $\Pr\left(\dot\cup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i)$.

**Note.** In Real Analysis 1 (MATH 5210) you will see that the behavior of the outer measure (and hence of the Lebesgue measure) of a set of real numbers is similar to the required conditions we impose on the probability function. For example, Countable Additivity is addressed in Section 2.1. Introduction and Section 2.3. The $\sigma$-Algebra of Lebesgue Measurable Sets (see Proposition 2.13) of my Real

Analysis 1 notes. In fact, one of the main applications of measure theory is to the theory of probability. For my class notes on this application, see Measure Theory Based Probability.

**Definition.** A function $\Pr : \mathcal{A}(S) \to \mathbb{R}$ (where $\mathcal{A}$ is a $\sigma$-algebra of events) which satisfies Axioms 1, 2, and 3 is a *probability measure*, or simply *probability*, on sample space $S$.

**Note.** We now establish some properties of Pr.

**Theorem 1.5.1.** $\Pr(\varnothing) = 0$.

**Theorem 1.5.2. Finite Additivity.** For any finite sequence of $n$ disjoint events $A_1, A_2, \ldots, A_n$ we have $\Pr\left(\cup_{i=1}^n A_i\right) = \sum_{i=1}^n \Pr(A_i)$.

**Note.** Finite additivity for outer measure (and hence of Lebesgue measure) is addressed in Real Analysis 1 in Proposition 2.6 of Section 2.3. The $\sigma$-Algebra of Lebesgue Measurable Sets.

**Theorem 1.5.3. Probability of the Complement.** For any event $A$, $\Pr(A^c) = 1 - \Pr(A)$.

**Theorem 1.5.4. Monotonicity.** If $A \subset B$ then $\Pr(A) \leq \Pr(B)$.

**Note.** Monotonicity of outer measure (and hence of Lebesgue measure) is addressed in Lemma 2.2.A of my Real Analysis 1 online notes for Section 2.2. Lebesgue Outer Measure.

**Theorem 1.5.5.** For any event $A$, $0 \leq \Pr(A) \leq 1$.

**Theorem 1.5.6.** For any two events $A$ and $B$,

$$\Pr(A \cap B^c) = \Pr(A) - \Pr(A \cap B).$$

**Theorem 1.5.7.** For any two events $A$ and $B$,

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

**Note.** If events $A$ and $B$ are disjoint (or "mutually exclusive") then we have by Theorem 1.5.7 that $\Pr(A \cup B) = \Pr(A) + \Pr(B)$.

**Theorem 1.5.8. Bonferroni Inequality.** For all events $A_1, A_2, \ldots, A_n$.

$$\Pr\left(\cup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \Pr(A_i) \text{ and } \Pr\left(\cap_{i=1}^n A_i\right) \geq 1 - \sum_{i=1}^n \Pr(A_i^c).$$

**Note.** Theorem 1.5.8, the Bonferroni Inequality,is to be proved in Exercise 1.5.13. In measure theory, the first property is called *finite subadditivity*. We can also show that any sequence of events $A_1, A_2, \ldots$ satisfies $\Pr\left(\cup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \Pr(A_i)$. In measure theory, the corresponding property is called *countable subadditivity*; see Proposition 2.3 of my Real Analysis 1 online notes on Section 2.2. Lebesgue Outer Measure. Notice that $\sum_{i=1}^{\infty} \Pr(A_i)$ might diverge to $\infty$ (perhaps this is why DeGroot and Schervish exclude the sequence case as an exercise).

*Revised: 7/16/2019*