

Review of Definitions for Probability

H. Krieger, Mathematics 156, Harvey Mudd College

Fall, 2008

Motivation: In order to have a basis for probability theory which would make it a tool suitable for the mathematical modelling of random experiments, the Russian mathematician Andrei Kolmogorov formulated (in 1933) a collection of definitions and axioms. On these pages we present his formulation in a manner designed to make all of the commonly used terms have precise meanings.

1. **The Population or Outcome Space or Sample Space** of a random experiment is the **set** of all possible **outcomes** of this experiment. Thus if Ω is the sample space, *each* element of Ω , say ω , represents one possible outcome, or result, of the experiment.
2. **Events:** An event is a **subset** of the sample space. Consequently, an event is a *particular set* of outcomes of the experiment in question. If the sample space is finite or countably infinite, we usually regard *every* subset of the sample space as an event. In particular, this means that the empty set and the sample space itself are events. If the sample space is uncountably infinite, *e.g.* an interval on the real line containing more than one point, then not all subsets are necessarily considered to be events. What we do require is that the collection of events has the following properties:
 - (a) The sample space is an event.
 - (b) If E is an event, then E^c , the complement of E in the sample space, is also an event.
 - (c) If $\{E_n\}$ is a sequence of events, then $\bigcup_{n=1}^{\infty} E_n$ is also an event.

Here are some important consequences of these requirements:

- (a) The empty set must be an event.
- (b) If $\{E_n\}$ is a sequence of events, then $\bigcap_{n=1}^{\infty} E_n$ is also an event.
- (c) The difference between two events E and F , *i.e.* $E \setminus F \equiv E \cap F^c$, is an event.

A collection of subsets of the sample space with the above three properties is referred to as a **σ -algebra** or a **σ -field**.

In the case that the sample space is an interval of the real line we require that this σ -algebra contain all of the sub-intervals and then refer to the smallest such σ -algebra as the **Borel subsets** of that interval (after the French mathematician Émile Borel).

The events which contain precisely one outcome are referred to as **simple events**. Thus these events are indecomposable in the sense that there are *no* two distinct non-empty events with union equal to a simple event. Note that in the case of a finite or countably infinite sample space, where every subset is an event, we see that *every* event is the *union* of a finite or countably infinite number of *simple* events.

The *most important* relation between outcomes and events is the following: If the result of an experiment is the outcome ω and E is an event, then we say that the event E has **occurred** if and only if $\omega \in E$. Otherwise, *i.e.* if $\omega \in E^c$, we say that E has **not occurred**.

3. **Probability Measures:** Once we have a sample space Ω describing the outcomes of our experiment and the collection of subsets of the sample space that we call the events, we want to define what we mean by a probability measure or probability distribution on this collection of events. Kolmogorov's axioms for such a probability, which we denote by P , proceed as follows:

- (a) $P(\Omega) = 1$, *i.e.* we assign probability one to the entire sample space.
- (b) If E is an event, then $P(E) \geq 0$, *i.e.* every event is assigned a non-negative probability.
- (c) If $\{E_n\}$ is a sequence of *mutually exclusive* events, then $P(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} P(E_n)$, *i.e.* the probability assigned to the union of a countable collection of mutually exclusive events is the sum of the probabilities of the individual events. (Recall that two events E and F are mutually exclusive if and only if their intersection is empty.)

Here are some important consequences of these axioms:

- (a) $P(\phi) = 0$, *i.e.* we assign probability zero to the empty set.
- (b) If E is an event, then $P(E) + P(E^c) = 1$, *i.e.* the probability of an event and the probability of its complement sum to one.
- (c) If E is an event, then $P(E) \leq 1$.
- (d) If E and F are events and $E \subset F$, then $P(E) \leq P(F)$, *i.e.* probability is a function which increases as the number of outcomes in an event increases.
- (e) If E and F are events, then $P(E \cup F) + P(E \cap F) = P(E) + P(F)$, *i.e.* the sums of the probabilities of the union and intersection of two events is the same as the sum of the probabilities of the original events.