Probability Refresher

Keith A. Lewis

July 25, 2020

Abstract

This note collects salient facts about Probability Theory.

Probability Theory

In order to understand statistics one must first understand *probability theory*.

Events are assigned a probability between 0 and 1 representing a degree of belief that the event will happen if observed. Random variables are variables in the sense that they are a symbol that can be used in place of a number in equations and inequalities with additional information about the probability of the values it can have.

Probability Model

A *probability model* specifies a *sample space* and a *probability measure* for the possible *outcomes*. A *partition* of the sample space into *events* models partial information.

Sample Space

A *sample space* is the set of what can happen in a model. An *outcome* is an element of a sample space. An *event* is a subset of a sample space.

A sample space for flipping a coin can be modelled by $\{H, T\}$ where H indicates heads and T indicates tails. Of course any two element set could be used for this.

A sample space for flipping a coin twice is $\{HH, HT, TH, TT\}$ where the outcome specifies the individual outcomes of the first and second flip. The event 'the first flip was heads' is the subset $\{HH, HT\}$. The partition $\{\{HH, HT\}, \{TH, TT\}\}$ represents the partial information of knowing the outcome of the first coin flip.

The first event in the partition indicates the first flip was heads. The second event in the partition indicates the first flip was tails.

People seem to be surprised probabilities are modeled using sets. Sets have no structure, they are just a bag of things (*elements*). People also seem to be rather cavalier about specifying sample spaces. The first step in any probablity model is to specify the possible outcomes. The second step is to assign probabilities to the outcomes.

[^monte-hall]

Probability Measure

A *probability measure* assigns a number between 0 and 1 to events. If Ω is a sample space and P is a probability measure then the measure of the union of sets is the sum of the measure of each set minus the measure of the intersection

$$P(E \cup F) = P(E) + P(F) - P(E \cap F)$$

for any events E and F. This is the mathematical way of expressing the requirement that measures do not double count. A probability measure must also satisfy

$$P(\emptyset) = 0 \text{ and } P(\Omega) = 1$$

where $\emptyset$ is the *empty set* that contains no elements. This implies $P(E \cup F) = P(E) + P(F)$ if E and F are *disjoint* events, $E \cap F = \emptyset$.

Exercise. *If Q is a measure with $Q(\emptyset) = a$ and $Q(\Omega) = b$ where $a \neq b$ show $(Q - a)/(b - a)$ is a probability measure.*

If Ω consists of a finite number of elements $\{\omega_1, \ldots, \omega_n\}$ we can define a probabilty measure by $P(\{\omega_i\}) = p_i$ where $0 \leq p_i \leq 1$ and $\sum_i p_i = 1$. Every subset of Ω corresponds to a subset $I \subseteq \{1, \ldots, n\}$. The probability of the event $E = \{\omega_i : i \in I\}$ is $P(E) = \sum_{i \in I} p_i$.

Exercise. *Show this defines a probability measure.*

For the two coin flip model (assuming the coin is fair) we assign equal probability to the oucomes. The probability of the first flip being heads is $P(\{HH, HT\}) = P(\{HH\} \cup \{HT\}) = P(\{HH\} + P(\{HT\}) = 1/4 + 1/4 = 1/2$.

Partition

A *partition* of a set Ω is a collection of subsets (events) that are *pairwise disjoint* with union equal to Ω. A partition $\{E_1, ...\}$ satisfies $E_i \subseteq \Omega$ for all i, $E_i \cap E_j = \emptyset$ if $i \neq j$, and $\cup_i E_i = \Omega$. The elements of the partition are called *atoms*.

Partitions represent partial information about outcomes. Complete information means knowing what outcome $\omega \in \Omega$ occured. Partial information means knowing what atom ω belongs to.