Axioms of Probability

Probability Theory: Foundation for Data Science

with **Anne Dougherty**

Data Science
UNIVERSITY OF COLORADO **BOULDER**

Learning Goals

In this module, we'll learn about the difference between a population and a sample and why probability is the foundation for statistics and data science. We'll also begin our study of the foundations of probability.

- Explain why probability theory is relevant to statistics and data science.

- Describe what it means to predict the outcome of an experiment and organize the outcomes into sample spaces.

- **Calculate probabilities of events using the Axioms of Probability.**

- Calculate probabilities when each simple event is equally likely. Understand permutations and combinations.

What is a Probability?

The goal of **probability** is to assign some number, $P(A)$, called the probability of event A, which will give a precise measure to the chance that A will occur. In **statistics**, we draw a sample from a population, and give an estimate. So, you will be able to understand statistics more thoroughly and deeply if you first understand probabilities.

- Start with an experiment that generates outcomes
- Organize all of the outcomes into a sample space, S
- Let A be some event contained in S. That is, A is some collection of outcomes from the experiment.

What do we expect to be true of $P(A)$?

Axioms of Probability

Axiom 1 : For any event A, $0 \leq P(A) \leq 1$

Axiom 2 : $P(S) = 1$

Axiom 3 : If $A_1, A_2, \ldots, A_n$ are a collection of n mutually exclusive events (i.e. the intersection of any two is the empty set), then

$$P(\cup_{k=1}^{n} A_k) = \sum_{k=1}^{n} P(A_k)$$

Axioms of Probability - continued

Axiom 3' : More generally, if $A_1, A_2, \ldots$ is an infinite collection of mutually exclusive events, then

$$P\left(\cup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$$

These three properties are called the Axioms of Probability and we can derive many results from them.

Example 1

Experiment: Flip a fair coin until the first tail appears. Let 0 represent a head and 1 a tail.

$$S = \{1, 01, 001, 0001, \ldots\}$$

Let A_n represent the event of obtaining a tail on the n^{th} flip, $A_n = \{00 \cdots 01\}$ Find $P(A_1)$, $P(A_2)$, $P(A_5)$ and $P(A_n)$, where n is a positive integer.
$P(A_1) = 1/2$
$P(A_2) = P(\{01\}) = 1/4$
$P(A_5) = P(\{00001\}) = 1/2^5$
$P(A_n) = 1/2^n$

Example 1 - continued

Note: $P(S) = P(\cup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k) = \sum_{k=1}^{\infty} \frac{1}{2^k} = 1$

If B is the event that it takes at least 3 flips to obtain a tail, find $P(B)$.

B^c, the complement of B, is the event that you obtain a tail on the first or second flip.

$P(B^c) = P(\{1, 01\}) = 1/2 + 1/4 = 3/4$

We also note:

$P(S) = P(B \cup B^c) = P(B) + P(B^c) = 1$. So,

$P(B) = 1 - P(B^c) = 1 - 3/4 = 1/4$

Consequences of the Axioms

If A and B are two events contained in the same sample space S,

- $A \cap A^c = \emptyset$ and $A \cup A^c = S$ so, $1 = P(S) = P(A \cup A^c) = P(A) + P(A^c)$ which implies $P(A^c) = 1 - P(A)$
- If $A \cap B = \emptyset$, then $P(A \cap B) = 0$.
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

These three consequences will help us calculate many probabilities.

Example 2

Return to our car example: Recall a randomly selected car is inspected for three defects. The sample space is $S = \{000, 100, 010, 001, 110, 101, 011, 111\}$. Consider the three events:

- A is the event defect 1 is present, $A = \{100, 110, 101, 111\}$
- B is the event defect 2 is present, $B = \{010, 110, 011, 111\}$
- C is the event defect 3 is present, $C = \{001, 011, 101, 111\}$

Example 2 continued

Suppose over many days, data is collected and it is found that 20% of the cars have defect 1, 25% have defect 2, and 30% have defect 3. Further, 5% have defects 1 and 2, 7.5% have defects 2 and 3, 6% have defects 1 and 3, and 1.5% have all three.

Example 2 continued

Calculate the probability of each of the following events for the randomly selected car:

- defect 1 did not occur
- at least one defect occurs
- no defect occurs
- defect 1 and 3 occur but 2 does not.