Probability and Statistics

Ash Bellett

Contents

1 Probability Theory

1.1 Probabilistic Experiments

1.1.1 Experiments

An experiment is a process of observation where the output cannot be predicted with certainty due to random effects.

1.1.2 Trials

A trial is a single performance of an experiment.

1.1.3 Outcomes

An outcome is an observed output of a trial.

1.1.4 Samples

A sample is a set of outcomes. The sample space Ω is the set of all possible outcomes of an experiment.

1.1.5 Events

An event A is a subset of outcomes in a sample space Ω.

$$A \subseteq \Omega$$

The event space $\mathcal{F}$ is the set of all possible events.

$$A \in \mathcal{F}$$

The axioms of set theory apply to events:

- The complement A' of an event A is the set of outcomes in the sample space Ω that are not in A.

$$A' = \{x \mid x \in \Omega \text{ and } x \notin A\}$$

- The union of events $A_1 \cup A_2$ is the set of outcomes in A_1, A_2 or both.

$$A_1 \cup A_2 = \{x \mid x \in A_1 \text{ or } x \in A_2\}$$

- The intersection of events $A_1 \cap A_2$ is the set of outcomes in both A_1 and A_2.

$$A_1 \cap A_2 = \{x \mid x \in A_1 \text{ and } x \in A_2\}$$

- The relative complement of event A_1 in event A_2 is the set of outcomes in A_2 that are not in A_1.

$$A_2 \setminus A_1 = \{x \mid x \in A_2 \text{ and } x \notin A_1\}$$

The event consisting of no outcomes is called the null event $\varnothing$. If events $A_1, A_2, ...$ have no outcomes in common then $A_1, A_2, ...$ are disjoint events (mutually exclusive).

$$A_1 \cap A_2 \cap ... = \varnothing$$

1.1.6 Probability

A probability measure $P : \mathcal{F} \to \mathbb{R}$ is a function that assigns each event in the event space to a real number.

1.2 Axioms

1.2.1 First axiom: non-negative, real

The probability of an event is a non-negative real number.

$$P(A) \geq 0 \ \forall A \in \mathcal{F}$$

This axiom means that the smallest probability of an event is 0 (impossible events). It does not specify an upper bound, however a theorem does.

1.2.2 Second axiom: unitarity

The probability that at least one outcome in the sample space will occur is 1.

$$P(\Omega) = 1$$

This axiom means that it is certain that an outcome will occur from observing an experiment.

1.2.3 Third axiom: countable additivity

If $A_1, A_2, ...$ is an infinite set of disjoint events in a sample space Ω:

$$P(A_1 \cup A_2 \cup ...) = \sum_{i=1}^{\infty} P(A_i)$$

This axiom forms a relationship between a set of disjoint events in a sample space and the individual probabilities of each event.

1.3 Theorems

1.3.1 Probability of the empty set

The probability of the null event is 0.

$$P(\varnothing) = 0$$

Proof: Let an infinite set of events be $\{A_i = \varnothing\}_{i=1}^{\infty}$. Since $\bigcup_{i=1}^{\infty} \varnothing = \varnothing$ and substituting A_i into the third axiom:

$$\begin{aligned} P(\bigcup_{i=1}^{\infty} \varnothing) &= \sum_{i=1}^{\infty} P(\varnothing) \\ P(\varnothing) &= \sum_{i=1}^{\infty} P(\varnothing) \end{aligned} \tag{1}$$

The only solution to this equation is

$$P(\varnothing) = 0$$

.

1.3.2 Additivity

Given a finite set of n disjoint events $A_1, A_2, ...A_n$ in a sample space Ω:

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$$

Proof: Let an infinite set of disjoint events be $E_1 = A_1, E_2 = A_2, ..., E_n = A_n, E_{n+1} = \varnothing, E_{n+2} = \varnothing, ...$. Substituting this into the third axiom:

$$\begin{aligned} P(\bigcup_{i=1}^{\infty} E_i) &= \sum_{i=1}^{\infty} P(E_i) \\ &= \sum_{i=1}^{n} P(E_i) + \sum_{i=n+1}^{\infty} P(E_i) \\ &= \sum_{i=1}^{n} P(E_i) + \sum_{i=n+1}^{\infty} P(\varnothing) \\ &= \sum_{i=1}^{n} P(E_i) + \sum_{i=n+1}^{\infty} 0 \\ &= \sum_{i=1}^{n} P(E_i) \\ P(\bigcup_{i=1}^{n} A_i) &= \sum_{i=1}^{n} P(A_i) \end{aligned} \tag{2}$$

1.3.3 Monotonicity

If an event A_1 is a subset of or equal to another event A_2 which is a subset of or equal to a sample space Ω, then the probability of A_1 occurring is less than or equal to the probability of A_2 occurring.

$$\text{If } A_1 \subseteq A_2 \subseteq \Omega \text{ then } P(A_1) \leq P(A_2)$$

Proof: Let event A_1 be a subset of or equal to event A_2. The event $A_2 \setminus A_1$ is the set difference of A_2 and A_1 and is the set of outcomes in A_2 that are not in A_1. The union of A_1 and $A_2 \setminus A_1$ is equal to A_2.

$$A_2 = A_1 \cup (A_2 \setminus A_1)$$

Taking probabilities of both sides:

$$P(A_2) = P(A_1 \cup (A_2 \setminus A_1))$$

From the probability theorem of additivity where $n = 2$:

$$P(A_2) = P(A_1) + P(A_2 \setminus A_1)$$

From the first axiom, $P(A_2 \setminus A_1) \geq 0$. Therefore, $P(A_1) \leq P(A_2)$.

1.3.4 Complement rule

If A is an event in a sample space Ω, then the probability of the complement of A is given by:

$$P(A') = 1 - P(A)$$

Proof: Let event A be a subset of or equal to a sample space Ω. Then $S = A \cup A'$ and A and A' are disjoint events. From the first axiom:

$$P(\Omega) = P(A \cup A') = 1$$

From the third axiom:

$$P(A \cup A') = P(A) + P(A') = 1$$

Rearranging to make $P(A')$ the subject:

$$P(A') = 1 - P(A)$$

1.3.5 Numeric bounds

If A is an event in a sample space Ω, then the probability of A is bounded between 0 and 1.

$$0 \leq P(A) \leq 1$$

Proof: From the first axiom and the complement rule:

$$1 - P(A) \geq 0$$

Rearranging to make $P(A)$ the subject:

$$P(A) \leq 1$$

From the first axiom:

$$0 \leq P(A) \leq 1$$

1.3.6 Sum rule

If A_1 and A_2 are events in a sample space Ω, the probability that either A_1 or A_2 will occur is the sum of the probabilities that A_1 will occur and that A_2 will occur minus the probability that both A_1 and A_2 will occur.

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

Proof: From the third axiom

1.3.7 Probabilities from elementary events

1.3.8 Equally likely outcomes

2 Independence and Conditional Probability

2.1 Independent Events

2.2 Conditional Probability

2.3 Bayes' Theorem

3 Counting

3.1 Product Rule

If E_1 is an experiment with n_1 outcomes and E_2 is an experiment with n_2 outcomes, then the experiment which consists of performing E_1 and then E_2 has $n_1 n_2$ possible outcomes.