Basics of probability

Wenda Zhou

May 24, 2017

Why probability?

- Probability is a mathematical language to talk about uncertainty
- Enables us to describe uncertainty in a rigorous and precise fashion
- Base language of statistics

Axiomatic probability

- In order to make our statements precise, we will define all the concepts
- These concepts may not always correspond exactly to the colloquial notion
- Will make extensive use of mathematical notation, especially from set theory

Sample space

The sample space is the set of possible outcomes.
Usually called Ω (Greek letter capital omega)

Sample space

The sample space is the set of possible outcomes.
Usually called Ω (Greek letter capital omega)

Coin flip

For a coin flip, only two possible outcomes, H and T. May write this mathematically as

$$\Omega = \{H, T\} \tag{1}$$

Sample space

The sample space is the set of possible outcomes.
Usually called Ω (Greek letter capital omega)

Coin flip

For a coin flip, only two possible outcomes, H and T. May write this mathematically as

$$\Omega = \{H, T\} \tag{1}$$

Die roll

For a die roll, six possible outcomes, 1 to 6. Write

$$\Omega = \{1, 2, 3, 4, 5, 6\} \tag{2}$$

Sometimes, sample space will be too large to describe. This is not a problem.

Event

- Events are the main building blocks of probability.
- Objects for which we can ask the probability of.
- Mathematically expressed as a subset of Ω, and usually named A, B, etc.
- Write the probability of the event $\mathrm{P}(A)$

Event

Example - Coin Flip

- The sample space is $\Omega = \{H, T\}$.
- $A = \{H\}$ is an event.
- Also have the event $\varnothing = \{\}$, and the event $\{H, T\} = \Omega$.

The last two events exist in all sample spaces.

Event

Example - Die Roll

- The sample space is $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- $A = \{1\}$, the event that the die lands on 1.
- $B = \{1, 3, 5\}$, the event that that the die lands on a odd number.
- $\varnothing$ and Ω are also events as always.

Probability of events

Assign probability to events.

Fair coin

$$\begin{aligned} \mathrm{P}(\{H\}) &= \mathrm{P}(\{T\}) = 0.5 \\ \mathrm{P}(\Omega) &= 1 \\ \mathrm{P}(\varnothing) &= 0 \end{aligned}$$

Probability of events

Assign probability to events.

Die

$$\begin{aligned} P(\{1\}) &= 1/6 \\ P(\{1, 3, 5\}) &= 1/2 \\ P(\Omega) &= 1 \\ P(\emptyset) &= 0 \end{aligned}$$

Probability of events

Assign probability to events – can have different possible probabilities for same sample space.

Biased coin

$$
\begin{aligned}
\mathrm{P}(\{H\}) &= 0.7 \\
\mathrm{P}(\{T\}) &= 0.3 \\
\mathrm{P}(\Omega) &= 1 \\
\mathrm{P}(\varnothing) &= 0
\end{aligned}
$$

Probability

Probabilities obey the following basic properties:

$$\mathrm{P}(\Omega) = 1$$
$$\mathrm{P}(\varnothing) = 0$$
$$0 \leq \mathrm{P}(A) \leq 1$$

Calculus of probability

Relate the probabilities of related events

Additivity rule

The additivity rule allows us to compute the probability that either A or B (or both) happen.
We write this event $A \cup B$ (read A union B), the event of outcomes that are in A or B.

Additivity rule

The additivity rule allows us to compute the probability that either A or B (or both) happen.
We write this event $A \cup B$ (read A union B), the event of outcomes that are in A or B.
If A and B cannot both happen at the same time, then we have that:

$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) \tag{3}$$

We say A and B are disjoint or mutually exclusive. Mathematically, we write $A \cap B = \varnothing$.

Complement of an event

For every event A, we may define an event A^C, the event that "A does not happen". We have

$$P(A^C) = 1 - P(A) \tag{4}$$

Inclusion-exclusion

Additivity rule requires mutually exclusive condition. What about A and B in general? The Inclusion-exclusion principle gives:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{5}$$

Independence

Independence is the notion that events do not affect each other. Mathematically, we say A and B are independent if

$$\mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}(B) \tag{6}$$

Conditional probability

Conditional probability encodes the idea of partial knowledge of events.
Mathematically, we define the probability of A given B by

$$\mathrm{P}(A \mid B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} \tag{7}$$

Note that we must have $\mathrm{P}(B) > 0$.

Conditional probability

Conditional probability encodes the idea of partial knowledge of events.

Mathematically, we define the probability of A given B by

$$\mathrm{P}(A \mid B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} \tag{7}$$

Note that we must have $\mathrm{P}(B) > 0$.

If A and B are independent, then we have that

$$\mathrm{P}(A \mid B) = \mathrm{P}(A) \tag{8}$$

Conditional probability and additivity

Additivity rule for conditional probability.
If A and B are mutually exclusive, then we have

$$\mathrm{P}(A \cup B \mid C) = \mathrm{P}(A \mid C) + \mathrm{P}(B \mid C) \tag{9}$$

In general, conditional probabilites behave like probabilities

Law of total probability

The law of total probability allows to relate the probability of an event to that of its conditional probabilities.

$$\mathrm{P}(A) = \mathrm{P}(A \mid B)\,\mathrm{P}(B) + \mathrm{P}(A \mid B^C)\,\mathrm{P}(B^C) \tag{10}$$

Bayes Rule

The Bayes rule allows us to relate the two probabilities $\mathrm{P}(A \mid B)$ and $\mathrm{P}(B \mid A)$.

$$\mathrm{P}(A \mid B) = \frac{\mathrm{P}(B \mid A)\,\mathrm{P}(A)}{\mathrm{P}(B)} \tag{11}$$