ST 516: Foundations of Data Analytics

Rules of Probability

Fundamentals of Probability

Remember that the probability of an event is the relative frequency of that event's occurrence over an infinite series of trials.

We have defined:

$$\text{relative frequency} = \frac{\text{number of times an event occurs}}{\text{number of trials}}$$

This means that:

- A probability must be a number between 0 and 1.
- Mathematically, if we let E denote the occurrence of event E, then $0 \leq \mathsf{P}(E) \leq 1$, where $\mathsf{P}(E)$ is the probability of E occurring.

Fundamentals of Probability

A consequence of the fact that $0 \leq \mathsf{P}(E) \leq 1$ is that the probability that E does not occur must be:

$$\mathsf{P}(E^c) = 1 - \mathsf{P}(E)$$

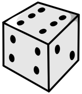
where the symbol, E^c, is read "E complement" or "not E."

Intuitively, if we conduct a trial for which E is a possible outcome, then the trial must result in either E occurring or E not occurring.

Probabilities for Several Events

So far, we've been talking about a single event, but in many data examples, there is more than one event possible.

- If we roll a fair, six-sided die, there are six possible outcomes, the members of the set $\{1, 2, 3, 4, 5, 6\}$.



  - One event resulting from this roll is that a 2 appears
  - Another event is that an even number appears

We need some rules for evaluating probabilities related to multiple events.

Probabilities for Several Events

Let's continue our consideration of the roll of a fair, six-sided die.

- We say that the outcomes (members of the set $\{1, 2, 3, 4, 5, 6\}$) are mutually exclusive (if you roll a 1, you can't have rolled a 4) and exhaustive (the only possible results of the die roll are given by this set).
  Because the outcomes are exhaustive,

$$\mathsf{P}(1) + \mathsf{P}(2) + \mathsf{P}(3) + \mathsf{P}(4) + \mathsf{P}(5) + \mathsf{P}(6) = 1$$

- And since the die is fair,

$$\mathsf{P}(1) = \mathsf{P}(2) = \mathsf{P}(3) = \mathsf{P}(4) = \mathsf{P}(5) = \mathsf{P}(6) = \frac{1}{6}$$

Probabilities for Several Events

Because the outcomes of a roll of a die are mutually exclusive,

$$\mathsf{P}(1 \text{ or } 4) = \mathsf{P}(1) + \mathsf{P}(4) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

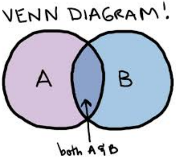
But what about two events that are not mutually exclusive? For example, let A denote the event that the die shows an even number, and let the event B denote the event that the die shows a 4.

- Now, $\mathsf{P}(A) = \frac{1}{2}$ and $\mathsf{P}(B) = \frac{1}{6}$.
- But notice that the event $C = (A \text{ or } B)$ is actually equivalent to the event A.

So how do we calculate $\mathsf{P}(A \text{ or } B)$ when A and B are not mutually exclusive events?

Venn Diagrams

Consider the following Venn Diagram that depicts two events that are not mutually exclusive (because they overlap)



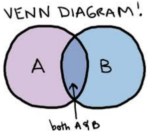
The reason that we can't just add probabilities of non-mutually exclusive events is that we'll end up double counting the probability associated with the overlapping part of the Venn diagram, where both A and B occur.

Joint Probability

The general rule for the probability of one event or another event occurring is:

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B).$$

In our Venn Diagram below, the event A includes the event (A and B), and the event B also includes the event (A and B). Therefore, we must subtract off $P(A \text{ and } B)$



$P(A \text{ and } B)$ is called the **joint probability** of A and B.

Conditional Probability

There is a multiplication rule for calculating joint probabilities, but to understand it you need to learn about *conditional probability*.

Conditional probability is the probability that some event will occur, given that another event has occurred.

- We write $\mathsf{P}(A|B)$ to denote the conditional probability of A given B. B is called the conditioning event.
- From our fair, six-sided die example, suppose that I tell you that a roll of the die has resulted in an even number (a 2, 4 or 6). What is $\mathsf{P}(4)$?

Conditional Probability

There is a multiplication rule for calculating joint probabilities, but to understand it you need to learn about *conditional probability*.

Conditional probability is the probability that some event will occur, given that another event has occurred.

- We write $\mathsf{P}(A|B)$ to denote the conditional probability of A given B. B is called the conditioning event.
- From our fair, six-sided die example, suppose that I tell you that a roll of the die has resulted in an even number (a 2, 4 or 6). What is $\mathsf{P}(4)$?

  Answer: Since the possible outcomes are now just members of the set $\{2, 4, 6\}$, $\mathsf{P}(4) = \frac{1}{3}$.

Joint Probability

The multiplication rule for joint probability is

$$\mathrm{P}(A \text{ and } B) = \mathrm{P}(A|B)\mathrm{P}(B) \tag{1}$$

Since $\mathrm{P}(A \text{ and } B) = \mathrm{P}(B \text{ and } A)$, we can also write

$$\mathrm{P}(A \text{ and } B) = \mathrm{P}(B|A)\mathrm{P}(A). \tag{2}$$

You can use either equation above to calculate the joint probability, it's just that sometimes it's easier to consider B the "conditioning" event as in (1) above, and sometimes it's easier to consider A the conditioning event as in (2). The conditioning event is the event that appears to the right of the | symbol.

Example

Suppose that you want to calculate the joint probability that a student in a statistics class passes the final exam AND passes the class. Let A denote the event that the student passes the final exam, and let B denote the event that the student passes the class.

- We want to calculate $\mathrm{P}(A \text{ and } B)$.
  - It makes sense for A to be the conditioning event here, since A comes before B in time.
  - Using the product rule, we need to calculate:

$$\mathrm{P}(A \text{ and } B) = \mathrm{P}(B|A)\mathrm{P}(A)$$

  - Are there situations in which it's reasonable to think of B as the conditioning event?

Example Continued

Now, suppose that the course instructor knows from previous experience that the probability that a student will pass her class given that the student passed the final exam is 0.98, and that the probability that a student will pass the final exam is 0.93.

- We calculate:

$$\begin{aligned} \mathsf{P}(A \text{ and } B) &= \mathsf{P}(B|A)\mathsf{P}(A) \\ &= 0.98(0.93) \\ &= 0.911 \end{aligned}$$

So, the probability that a student in the class passes both the final and the class is 0.911..