Conditional Probability and Bayes' Rule

CMS 380 Simulation and Stochastic Modeling

Joint and Conditional Probabilities

For two events, E and F, the *joint probability*, written $P(EF)$, is the the probability that both events occur.

For example, let E be *a die roll is even* and F be *a die roll is greater than 3*. The following sets describe each event:

$$\begin{aligned} E &= \{2, 4, 6\} \\ F &= \{4, 5, 6\} \\ E \cap F &= \{4, 6\} \end{aligned}$$

The probability that the joint event occurs is the probability that the outcome is in $E \cap F$, $\frac{2}{6}$.

Sometimes, you want to calculate the probability of an event given some *known information about previous events*. The *conditional probability* of event E given event F is

$$P(E|F) = \frac{P(EF)}{P(F)}$$

To interpret this formula, think about the space of events belonging to E and F as partially overlapping boxes (see the "drawing" on the next page). The space that's contained in both boxes (marked with x's in the figure) is the joint event EF. The conditional probability is the fraction of box F that's covered by the joint event.

In other words, if you were required to pick a random point from box F, what fraction of those points would lie within the shared region EF?

```
                  =================
                 |                 |
 -----------------|-------    F    |
|                 |xxxxxxx|        |
|                 |xxxxxxx|        |
|        E         =================
|                        |
 -------------------------
```

Two events are *independent* if $P(E|F) = P(E)$. That is, knowing that F occurred tells us nothing about whether E occurs.

The join probability of independent events is the product of their individual probabilities:

$$\begin{aligned} P(EF) &= P(E|F)P(F) \\ &= P(E)P(F) \end{aligned}$$

The first line is the rearranged definition of conditional probability and the second is applying the definition of independence.

Examples

Coins. *What is the probability that two coins both come up heads given that the first is heads?*

There are two ways to approach this problem. First, we could recognize that the two flips are independent, so the outcome of the first has no effect on the second. The probability depends only on the outcome of the second coin flip, so the answer is $\frac{1}{2}$.

A second, more complete, way of thinking about the problem is to reason about the *reduced sample space*. If the first coin is heads, then the set of possible outcomes has only two elements: $\{HH, HT\}$. Two heads is therefore one-half of the reduced set of possible outcomes.

What is the probability of both coins coming up heads given that ***at least one*** *comes up heads?*

This is slightly trickier, because either the first or second coin or both could come up heads. Reason about the reduced sample space, which has three elements:

$$\{HH, HT, TH\}$$

The probability of two heads in this scenario is therefore $\frac{1}{3}$.

Urns. *I got my probability urn back out and filled it with 8 red balls and 4 white balls. Suppose I draw two balls without replacement. What is the probability that both are red?*

Thinking about sequences of events can be tricky. Here's a plan:

- Define the events in question. Let's have R_1 represent the event *get a red ball on the first draw* and R_2 represent *get a red ball on the second draw*.
- Determine what probability you need to calculate. This usually requires thinking carefully about the difference between joint events and conditional events. Here, the quantity we want is $P(R_1\, R_2)$, the joint probability of getting red balls on both draws. This *is not* the same as $P(R_2|R_1)$, which would consider only the outcome of the second trial.
- Reason about the reduced sample space to determine the probability of the sequence of events.

From the definition of conditional probability,

$$P(R_1\, R_2) = P(R_2|R_1)P(R_1)$$

The probability of getting a red ball on the first trial is easy:

$$P(R_1) = \frac{8}{12}$$

Removing one red ball changes the numbers in the urn, so

$$P(R_2|R_1) = \frac{7}{11}$$

The final result is therefore

$$P(R_1\, R_2) = \frac{8}{12} \cdot \frac{7}{11}$$

Practice question. *What is the probability that the first ball is red and the second is white?*

Practice question. *What is the probability that one ball is red and the other is white? Note that you could have either order.*

Bayes' Theorem

Bayes' Theorem is a variation of the definition of conditional probability, named after the Rev. Thomas Bayes, who first described it in 1769.

Bayes' Theorem allows you to "turn around" a conditional probability $P(A|B)$ and calculate it in terms of $P(B|A)$. The rule is

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

The derivation uses the definition of conditional probability twice:

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(B|A)\,P(A)}{P(B)}$$

The biggest challenge in using Bayes' Theorem is often calculating the unconditional probabilities of the two basic events, $P(A)$ and $P(B)$.

Blood Test

Here's a classic example.

A blood test is 95% effective at detecting the presence of a disease, but it has a 1% false positive rate. This means that if a healthy person is tested, there is a 1% chance that the test result will imply he or she has the disease.

The disease actually occurs in .5% of the population.

What is the probability that a person with a positive test result actually has the disease?

First, let's think about the scenario. An ideal medical test would be one that always detected the target disease, but never produced any false positives.

That isn't possible, of course, so a real test has to accept the fact that some illnesses will go undetected and some healthy people will get false positives. The challenge is in designing the test to balance those two tradeoffs.

In some cases, it's seen as better to have a higher false positive rate than to avoid missing a real illness: anyone with a positive test can get tested again, but anyone with a negative result will be at risk of leaving their disease undiscovered and untreated. In other cases, like prostate cancer screening, a false positive might lead to invasive follow-up test that has unclear long-term clinical benefits.

The goal is to calculate $P(\text{disease} \mid \text{positive test})$. The following facts are available from the problem description:

$$\begin{aligned} P(\text{disease}) &= .005 \\ P(\text{no disease}) &= .995 \\ P(\text{positive test} \mid \text{disease}) &= .95 \\ P(\text{positive test} \mid \text{no disease}) &= .01 \end{aligned}$$

This is a classic set up for Bayes' Theorem. We need to get $P(A|B)$ and the problem statement gives us $P(B|A)$.

Setting everything up:

$$P(\text{disease} \mid \text{positive test}) = \frac{P(\text{positive test} \mid \text{disease})\, P(\text{disease})}{P(\text{positive test})}$$

The numerator is easy to calculate. The denominator, however, is not a basic quantity given in the problem description. The solution is to calculate $P(\text{positive test})$ using total probability:

$$\begin{aligned} P(\text{positive test}) = &P(\text{positive test} \mid \text{disease})\, P(\text{disease}) \\ &+ P(\text{positive test} \mid \text{no disease})\, P(\text{no disease}) \end{aligned}$$

This formulation uses the definition of conditional probability and the fact that there are two possible conditions: disease or no disease. It's possible to extend this kind of calculation to an arbitrary number of conditions.

Crunching the numbers yields $P(\text{disease} \mid \text{positive test}) \approx .3231$—only 32% of people with a positive test actually have the disease. This is surprisingly low!

To understand the counterintuitive result, suppose we have a population of 20000 people.

- 100 (.5%) really have the disease. Of those, 95 are detected by the test.
- 19900 don't have the disease, but 199 of them (1%) get a false positive by random chance.

The total fraction of positive tests due to true illness is therefore

$$\frac{95}{294} \approx .3231$$

Because the disease is so rare, the 1% of the healthy population actually accounts for more positive tests than 95% of those with the disease.