Independence and Conditional Probability

MATH 107: Finite Mathematics

University of Louisville

March 24, 2014

Motivating Questions

Any decent actuary can answer the question:

A simple mortality statistic

What is the probability that a randomly selected person will live to the age of 80?

However, this is not actually a terribly useful answer! In assessing mortality, we often want to take other factors into account and *restrict the sample space*:

More fine-tuned mortality statistics

- What is the probability that a randomly selected person will live to the age of 80, **given that they are male**?
- What is the probability that a randomly selected person will live to the age of 80, **given that they are a smoker**?

The approach to these more refined questions lies in *conditional probability*.

Considering a conditional question

A simple conditional

We pick a random number from 1 to 10, with each number equally likely. Given that we chose an odd number, what is the probability that the number is prime?

We could, under the conditions given, have picked 1, 3, 5, 7, or 9, with equal probabilities of $\frac{1}{5}$.

Of these 5 equally likely selections, 3 of them are prime, so the probability of a prime is $\frac{3}{5}$.

We thus arrive at our answer by *restricting the sample space*.

Symbols and calculations

We explicitly define these conditional probabilities symbolically:

Definition

The probability of an event A given an event B, denoted $P(A|B)$, is the probability that the event A occurs using B, instead of U, as the sample space.

Note that when all elements of the sample space are equally likely, $P(A|B) = \frac{n(A \cap B)}{n(B)}$.
And, in fact, the same basic formula works for arbitrary sample spaces, too:

Formula for conditional probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \text{ if } P(B) \neq 0$$

Note that $P(A|B)$ does not exist when B has zero probability.

Calculating a conditional probability

Let's look at another problem with a more involved sample space:

Rolling dice to get sums

We roll two six-sided dice; the sum of the numbers shown is odd. What is the probability that the sum is also a multiple of 3?

Let's call the event that a sum divisible by 3 is rolled A, and an odd sum B.
We know the probabilities of numbers 2 through 12 are $\frac{1}{36}, \frac{2}{36}, \ldots, \frac{5}{36}, \frac{6}{36}, \frac{5}{36}, \ldots, \frac{1}{36}$.

$$P(B) = \frac{2}{36} + \frac{4}{36} + \frac{6}{36} + \frac{4}{36} + \frac{2}{36} = \frac{18}{36}$$
$$P(A \cap B) = \frac{2}{36} + \frac{4}{36} = \frac{6}{36}$$
$$P(B|A) = \frac{P(A \cap B)}{P(B)} = \frac{6/36}{18/36} = \frac{6}{18} = \frac{1}{3}$$

From conditional probability to intersections

We can twist around the definition of a conditional probability to find the probability of an intersection of events:

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

In other words, the likelihood of A and B both happening is equal to the likelihood of B happening, times the likelihood that, given B, A happens.

From conditional probability to intersections

$$P(A|B)P(B) = P(A \cap B) = P(B|A)P(A)$$

Example

Suppose that 23% of the students at a certain high school are sophomores, and that 64% of the sophomores take a science class. What is the probability that a randomly selected student is a sophomore taking science?

Let A be the event "a randomly selected student takes science" and B the event "a random student is a sophomore". Thus $P(B) = 0.23$ and $P(A|B) = 0.64$, so

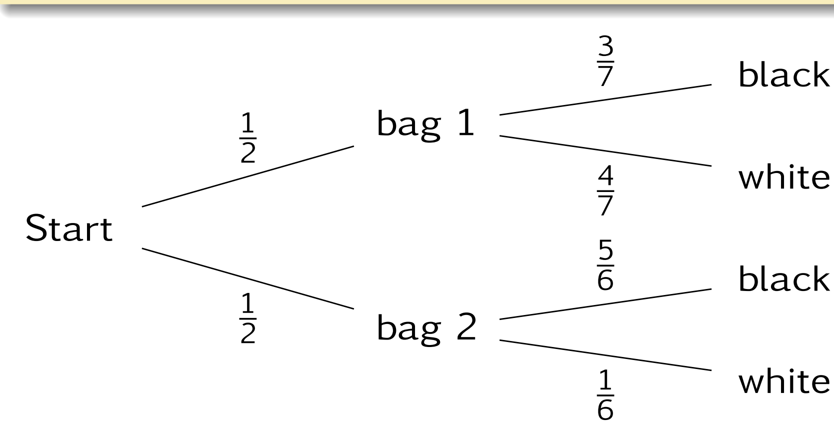
$$P(A \cap B) = P(B)P(A|B) = 0.1472$$

Trees

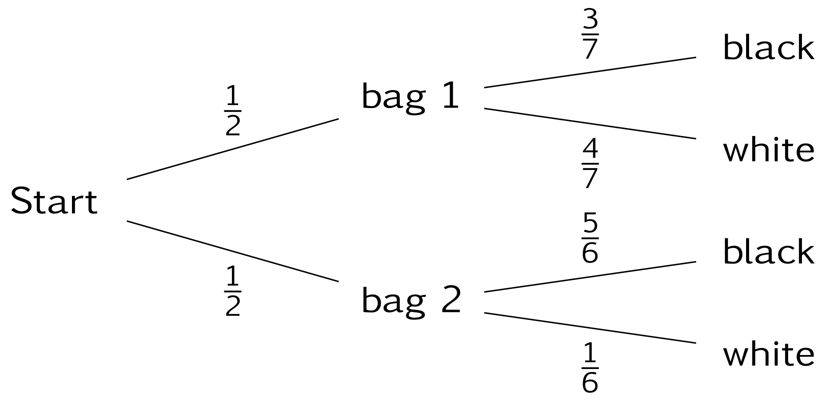
We can use trees to get a better visualization of probability.

A two-stage problem

We have two bags; one contains 3 black balls and 4 white balls, and the other has 5 black balls and 1 white ball. We pick a bag at random and then pick a ball from the bag. What is the probability we draw a black ball?



Trees, cont'd



If we describe the event of picking bag 1 as A and drawing a black ball as B, then we know:

$$P(A) = P(A') = \frac{1}{2}; P(B|A) = \frac{3}{7}; P(B'|A) = \frac{4}{7}; P(B|A') = \frac{5}{6}; P(B'|A') = \frac{1}{6}$$

and we want to find $P(B)$:

$$P(B) = P(B \cap (A \cup A')) = P(B \cap A) + P(B \cap A') = \frac{1}{2} \cdot \frac{3}{7} + \frac{1}{2} \cdot \frac{5}{6} = \frac{53}{84} \approx 63\%$$

Two notable situations can arise with regard to conditional probability. Below, assume that B has nonzero probability.

Common conditional effects

If $P(A|B) = P(A)$, then A and B are said to be *independent*.
If $P(A|B) = 0$, then A and B are said to be *mutually exclusive*.

For example, let's suppose we roll a white and a black die. Then:

- The events "white 3" and "white 5" are mutually exclusive.
- The events "white 3" and "black 5" are independent.
- The events "white 3" and "sum 7" are (surprisingly) independent.
- The events "white 3" and "sum 10" are mutually exclusive.
- The events "white 3" and "sum 5" are neither independent nor exclusive ("sum 5" has probability $\frac{1}{9}$, conditional probability $\frac{1}{6}$).

Independence reconceptualized

We can also say two events are independent if *one has no influence over the other*.

For example, if I flip a coin 3 times and roll a die 4 times, each result is *independent*.

Independent events are very easy to intersect:

Intersection law for independent events

If A and B are independent, then $P(A \cap B) = P(A|B)P(B) = P(A)P(B)$. If $E_1, E_2, \ldots, E_n$ are independent, then

$$P(E_1 \cap E_2 \cap \cdots \cap E_n) = P(E_1)P(E_2)\cdots P(E_n)$$

We used independence last week on the birthday paradox and Chevalier de Méré's problem!

Using independence to solve problems

Interseciton of several events

I roll a die 4 times. What is the probability that my first and second rolls are odd, my third is a 6, and my fourth is a multiple of 3?

Since the four events in question (odd first roll, odd second roll, 6 on third roll, and multiple-of-3 on fourth roll) are independent, I can multiply their probabilities!

$$\left(\frac{1}{2}\right)^2\left(\frac{1}{6}\right)\left(\frac{1}{3}\right) = \frac{1}{72}$$

Independent vs. dependent events

Replacement scenario

I have an urn with 3 black balls and 4 white balls. I draw a ball and put it back. Then I draw another ball. Are the events that the two balls are black independent?

Yes! In each case, regardless of how the other draw went, each draw has a probability of $\frac{3}{7}$.

Scenario without replacement

I have an urn with 3 black balls and 4 white balls. I draw a ball and don't put it back. Then I draw another ball. Are the events that the two balls are black independent?

No! If A and B are the events on our two draws, $P(B|A) = \frac{2}{6} = \frac{1}{3}$, while $P(B|A') = \frac{3}{6} = \frac{1}{2}$.

Why independence matters

Scientific inquiry is often interested in the question of whether two events, enacted on a random selection from the population, are independent.

An experimental inquiry

A random member x of the population is selected. Are the events "x smokes cigarettes" and "x has lung cancer" independent?

Statistical methods are used in the sciences to answer such questions!

When these methods show that two things are *not* independent, we learn about relationships among variables — in this case, how lung cancer and smoking are related.