LING3204 Second Language Acquisition
November 3, 2022

Conditional Probability

1 Basic Concepts and Notations

This section is based on Anderson et al. (2017). If you find it too hard, please skip this section and go to Section 2. If you find it interesting, you are encouraged to read the book. It's a nice introductory textbook to probability theory.

- The sample space Ω is the set of all possible outcomes.
  - For instance, $\Omega = \{\text{open}, \text{waitlist}, \text{closed}\}$ is a sample space consisting of all the possible statuses of a course when you register.
- A subset of the sample space is an event.
  - For instance, $A = \{\text{the course is not open}\} = \{\text{waitlist}, \text{closed}\}$ is an event which means that the course is not open (i.e., either closed or waitlist).
  - $\mathcal{F}$ contains all the subsets of the sample space Ω, i.e. all the possible events.
- $P(A)$ is the probability that event A occurs.
  - For instance, $P(A) = 0.2$ means the probability that the course you are looking at is no longer open is 0.2, i.e. 2 out of 10 times you will see the course being closed or waitlist.
  - For each event A, we have $0 \leq P(A) \leq 1$. $P(A) = 0$ means A is never going to happen, and $P(A) = 1$ means A is always going to happen.
  - $P(\neg A)$ is the probability that event A does not occur. We have $P(\neg A) = 1 - P(A)$.
- $P(AB)$ is the joint probability for events A and B, i.e. the probability that A and B happen at the same time.
  - For instance, A is the event that one got good grades, B is the event that one studied hard, then $P(AB)$ is the probability that you randomly ask a student and found that he or she studied hard and got good grades.
- $P(A \mid B)$ is the conditional probability of A given B, i.e. the probability that A happens given that B happens.
  - For instance, A is the event that one got good grades, B is the event that one studied hard, then $P(A \mid B)$ is the probability that you randomly ask a student who studied hard and found that he or she got good grades.

- We have

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

e.g. A is the event that one got good grades, B is the event that one studied hard. If $\frac{2}{3}$ of the students in the class studied hard ($P(B) = \frac{2}{3}$), and we know that $\frac{1}{2}$ of the students in the class studied hard and got good grades ($P(AB) = \frac{1}{2}$), then we know that if a student studied hard, then he or she would have a chance of $\frac{\frac{1}{2}}{\frac{2}{3}} = \frac{3}{4}$ to get good grades.

2 An Example

Suppose we asked 120 students in a math course whether they studied hard and if they got good grades. The table below shows the number of students in each of these categories.

	Good grades	Bad grades
Studied hard	60	20
Did not study hard	10	30

We calculate the sums of all the rows and columns and get

	Good grades	Bad grades	Sum
Studied hard	60	20	80
Did not study hard	10	30	40
Sum	70	50	120

What is the probability that a student got good grades if he or she studied hard? This is in fact the number of students who studied hard and got good grades, divided by the number of all the students who studied hard. To be specific, it is

$$P(\text{Good grades} \mid \text{Studied hard}) = \frac{60}{60+20} = \frac{3}{4}$$

Another way to do this is to divide everything in the table by 120. We have

	Good grades	Bad grades	Sum
Studied hard	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{2}{3}$
Did not study hard	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{3}$
Sum	$\frac{7}{12}$	$\frac{5}{12}$	1

The conditional probability is

$$P(\text{Good grades} \mid \text{Studied hard}) = \frac{P(\text{Studied hard \& got good grades})}{P(\text{Studied hard})} = \frac{\frac{1}{2}}{\frac{2}{3}} = \frac{3}{4}$$

which gives the same result. In fact, if we divide both the numerator and the denominator by 120 in the first solution, the two are identical.

3 The Contingency Problem

In the Week 7 slides, it is stated that Contingency (ΔP) is the probability of the outcome given the cue $P(O \mid C)$ minus the probability of the outcome in the absence of the cue $P(O \mid \neg C)$. We also have

	Outcome Present (O)	Outcome Absent ($\neg O$)
Cue Present (C)	a	b
Cue Absent ($\neg C$)	c	d

where a, b, c, d represent frequencies.

Wait, isn't it quite similar to the previous example? Let's first calculate the sums of all the rows and columns.

	Outcome Present (O)	Outcome Absent ($\neg O$)	Sum
Cue Present (C)	a	b	$a+b$
Cue Absent ($\neg C$)	c	d	$c+d$
Sum	$a+c$	$b+d$	$a+b+c+d$

Then what is the probability of the outcome being present given the cue being present? We have

$$P(\text{Outcome Present} \mid \text{Cue Present}) = P(O \mid C) = \frac{P(OC)}{P(C)} = \frac{a}{a+b}$$

Similarly, the probability of the outcome being present given the cue being absent is

$$P(\text{Outcome Present} \mid \text{Cue Absent}) = P(O \mid \neg C) = \frac{P(O\neg C)}{P(\neg C)} = \frac{c}{c+d}$$

ΔP is defined as the difference between the two. Thus, we have

$$\Delta P = P(O \mid C) - P(O \mid \neg C) = \frac{a}{a+b} - \frac{c}{c+d}$$

References

David F Anderson, Timo Seppäläinen, and Benedek Valkó. *Introduction to probability*. Cambridge University Press, 2017.