Conditional Probability and Bayes' Rule

Charlie Gibbons
University of California, Berkeley

ARE 210

Fall 2015

Notes

Outline

Notes

Definition

Two events A and B are *independent* if

$$\Pr(A \cap B) = \Pr(A)\Pr(B);$$

the occurrence of one does not change the probability of the other.

Notes

Rolling two dice

Suppose that we roll two dice. Are the outcomes of the two rolls independent?

Yes. One way to see this would be to enumerate the 36 equally-likely probabilities:

1,2	1	2	3	4	5	6
1	1,1	1,2	1,3	1,4	1,5	1,6
2	2,1	2,2	2,3	2,4	2,5	2,6
3	3,1	3,2	3,3	3,4	3,5	3,6
4	4,1	4,2	4,3	4,4	4,5	4,6
5	5,1	5,2	5,3	5,4	5,5	5,6
6	6,1	6,2	6,3	6,4	6,5	6,6

We can confirm independence by showing, for every event A and B, that $\Pr(A, B) = \Pr(A)\Pr(B)$.

Notes

A different case

Imagine instead that we care about the outcome on the first die and the sum of the outcomes on the two dice. Then we have:

1, 2	1	2	3	4	5	6
1	1,2	1,3	1,4	1,5	1,6	1,7
2	2,3	2,4	2,5	2,6	2,7	2,8
3	3,4	3,5	3,6	3,7	3,8	3,9
4	4,5	4,6	4,7	4,8	4,9	4,10
5	5,6	5,7	5,8	5,9	5,10	5,11
6	6,7	6,8	6,9	6,10	6,11	6,12

As we would imagine, the result of the first die influences the value of the sum, so they shouldn't be independent.

Notes

Dependence example

Let's prove it: Are the events first roll $= \{1, 2\}$ and the sum $= \{1, 2, 3, 4, 5\}$ independent?

1,2	1	2	3	4	5	6
1	1,2	1,3	1,4	1,5	1,6	1,7
2	2,3	2,4	2,5	2,6	2,7	2,8
3	3,4	3,5	3,6	3,7	3,8	3,9
4	4,5	4,6	4,7	4,8	4,9	4,10
5	5,6	5,7	5,8	5,9	5,10	5,11
6	6,7	6,8	6,9	6,10	6,11	6,12

$$\Pr(\text{first roll less than 3, sum less than 6}) = \frac{7}{36}$$
$$\neq \frac{5}{54} = \frac{2}{6} \times \frac{10}{36} = \Pr(\text{first roll less than 3}) \times \Pr(\text{sum less than 6})$$

Notes

A quick test

A quick way to demonstrate dependence is to pick clearly impossible cases.

Example: Are the draws of two cards independent?

Notes

Redefining the sample space

Suppose that we have two events, but we want to consider the distribution of one given that the other happens. That is, what is the distribution of B when A occurs?

Note that we are *redefining our sample space*—we only care about the observations such that A occurs. Of this subgroup, what is the frequency of B?

Notes

Conditional frequencies

Example: What is the probability that a person on the street is carrying an umbrella given that it is raining?

$$\frac{\#\{\text{carrying an umbrella and it is raining}\}}{\#\{\text{it is raining}\}}.$$

This is the sample frequency of B given or conditional upon A—the *conditional sample frequency*.

Notes

Conditional probability

The population analogue of conditional frequency, the *conditional probability* of B given A, forms the core of econometrics. The probability of B given A is

$$\Pr(B|A) = \frac{\Pr(B \text{ and } A)}{\Pr(A)}.$$

We divide by the probability of A to account for the fact that we are only considering a subpopulation.

If A and B are independent, then we have

$$\Pr(B \mid A) = \Pr(B).$$

This should be intuitive; if A and B are independent, then knowing A doesn't help you predict B.

Notes

Conditional probability example

What is Pr(second roll is a 3 | first roll is less than 3)?

1,2	1	2	3	4	5	6
1	1,1	1,2	1,3	1,4	1,5	1,6
2	2,1	2,2	2,3	2,4	2,5	2,6

$$\Pr(\text{second roll is a 3} \mid \text{first roll is less than 3}) = \frac{2}{12} = \frac{1}{6}$$

Notes

Monty Hall problem

A famous example is the Monty Hall problem. Suppose that you choose a door at random (*wlog*, door 1). A car is behind one door and donkeys are behind the others.

What's the probability that you chose the car?

Notes

The path not chosen

Now, Monty opens a door that you didn't choose, but *that he knows does not have the car.*

Possibilities:

1	2	3	Open?
D	D	C	2
D	C	D	3
C	D	D	2 or 3

$$\Pr(\text{door 2 opened}) = \frac{1}{3} + \frac{1}{3}\frac{1}{2} = \frac{1}{2}$$

$$\Pr(\text{car} \mid \text{door 2 opened}) = \frac{\Pr(\text{car and door 2 opened})}{\Pr(\text{door 2 opened})} = \frac{1/6}{1/2} = \frac{1}{3}$$

Notes

Stay or go?

Opening door 2 gives us no information about what's behind our door. (The same would be true about opening door 3 instead.) But does it give us information about what's behind door 3?

$$\Pr(\text{car behind 3} \mid \text{door 2 opened})$$
$$= \frac{\Pr(\text{car behind 3 and door 2 opened})}{\Pr(\text{door 2 opened})}$$
$$= \frac{1/3}{1/2} = \frac{2}{3}$$

Now the car is much more likely to be behind door 3! Given the option, you should switch to door 3.

Notes

Dictatorships and growth

Example from Bill Easterly's "Benevolent Autocrats" (2011).

Growth Commission Report, World Bank

Growth at such a quick pace, over such a long period, requires strong political leadership.

Thomas Friedman, *NY Times*

One-party autocracy certainly has its drawbacks. But when it is led by a reasonably enlightened group of people, as China is today, it can also have great advantages. That one party can just impose the politically difficult but critically important policies needed to move a society forward in the 21st century.

Notes

Wrong question, wrong interpretation

	Autocracy	Democracy
Growth Success	9	1

$$\widehat{\Pr}(\text{Autocracy} \mid \text{Success}) = \frac{9}{9+1} = 90\%$$
$$\widehat{\Pr}(\text{Democracy} \mid \text{Success}) = \frac{1}{9+1} = 10\%$$

Notes

Typical question

Econometricians generally ask for the

$$\Pr(\text{outcome} \mid \text{treatment and other predictors}).$$

Notes

The right question

	Autocracy	Democracy
Growth Success	9	1
Growth Failure	10	0
Neither	70	12

$$\widehat{\Pr}(\text{Success} \mid \text{Autocracy}) = \frac{9}{9+10+70} = 10\%$$
$$\widehat{\Pr}(\text{Success} \mid \text{Democracy}) = \frac{1}{1+0+12} = 8\%$$
$$\widehat{\Pr}(\text{Failure} \mid \text{Autocracy}) = \frac{10}{9+10+70} = 11\%$$
$$\widehat{\Pr}(\text{Failure} \mid \text{Democracy}) = \frac{0}{1+0+12} = 0\%$$

Notes

Law of total probability

Let $B_1, \ldots, B_k$ be a *partition* of the sample space Ω, such that all events are pairwise disjoint and $\cup_{i=1}^{k} B_i = \Omega$. Additionally, let $\Pr(B_i) > 0 \, \forall \, i$. Then, for every event A in Ω,

$$\Pr(A) = \sum_{i=1}^{k} \Pr(B_i) \Pr(A \mid B_i).$$

This is the *law of total probability*.

Notes

Returning to our definition

Note that we could ask the opposing question of our definition of conditional probability: supposing that B occurred, what's the probability that A occurs?

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}.$$

Using these two versions of the definition brings us to *Bayes' rule*:

$$\Pr(B \mid A) = \Pr(A \mid B) \frac{\Pr(B)}{\Pr(A)}$$

or

$$\Pr(B \mid A) = \Pr(A \mid B) \frac{\Pr(B)}{\sum_{i=1}^{k} \Pr(B_i) \Pr(A \mid B_i)}.$$

Notes

Defective products

Suppose that 20% of a product is produced by machine 1, 30% by machine 2, and 50% by machine 3. Let B_i be the event that a randomly selected item was produced by machine i. Let A be the event that the item is defective. Suppose:

$$\Pr(A \mid B_1) = 0.01$$
$$\Pr(A \mid B_2) = 0.02$$
$$\Pr(A \mid B_3) = 0.03$$

A single item is chosen and it is found to be defective. What is the probability that it is from machine 2?

$$\Pr(B_2 \mid A) = \frac{0.3 \times 0.02}{0.2 \times 0.01 + 0.3 \times 0.02 + 0.5 \times 0.03} = 0.26$$

Notes

Prior and posterior probabilities

$\Pr(B_2) = 0.3$ is the *prior probability* of the event that the item was from machine 2; it is your guess before you receive any additional information.

$\Pr(B_2 \mid A) = 0.26$ is the *posterior probability* of the event; you have *updated* your prior probability to incorporate new information.

Notes

Notes

Notes