#19: Bayes Rule

October 11, 2019

Karle Flanagan and Wade Fagen-Ulmschneider

Tables of Conditional Probability

Example 2: Suppose jurors make the right decisions about guilt and innocence 95% of the time and that 80% of all defendants are truly guilty:

	Acquitted	Convicted	Totals
Innocent			
Guilty			
Totals			**100**

a) What's the probability that a person is convicted given that they are innocent?

b) What's the probability that a person is innocent, given that they are convicted?

Example 3: For years ELISA was the most widely used test for screening donated blood for AIDS. If a blood sample has the AIDS virus, then there is a 99% probability that ELISA will correctly give a positive result. If the blood sample does **not** have the AIDS virus then there is still a 6% probability that ELISA will incorrectly give a positive result. About 1% of the blood samples ELISA tests are truly contaminated.

	Result of Test		
	Positive	**Negative**	**Totals**
Blood has AIDS virus			
Blood does NOT have AIDS virus			
Totals			**10,000**

a) What's the probability someone tested positive given that the person has the AIDS virus?

b) What's the probability someone has the AIDS virus given the person tested positive?

Decomposition into Conditional Probability

Occasionally, we have access to **only** conditional probabilities but need the probability of an event occurring. The probability of an event occurring is _____ **exactly equal to** the sum of two components:

[1]:

[2]:

Mathematically, we express this as the following:

P(**event**) = P(**event** | data)×P(data) + P(**event** | ¬data)×P(¬data)

Bayes Rule

Sometimes, we have the exact opposite conditional data that we need. Bayes Rule provides a way to find a conditional probability:

$$P(h \mid D) = P(D \mid h) \times \frac{P(h)}{P(D)}$$

In English, Bayes Rule tells us we can compute the probability of a hypothesis (**h**) given some data (**D**) given three individual probabilities:

1. The independent probability of the hypothesis being true, P(h),
2. The independent probability of the data being true, P(D), and
3. The conditional probability that the data is true given the hypothesis, P(D | h)

Example 3, using Bayes Rule:

If a blood sample has the AIDS virus, then there is a 99% probability that ELISA will correctly give a positive result. If the blood sample does not have the AIDS virus then there is still a 6% probability that ELISA will incorrectly give a positive result. About 1% of the blood samples ELISA tests are truly contaminated.

Conditionals in Python

Up until today, we have discovered two different control flow techniques in Python:

Control Flow	Python Syntax	Description
	`for i in range( n ):` `  ...`	Repeats indented code n times.
	`def myFunctionName( params ):` `  ...`	Runs indented code when function is called elsewhere in program.

With conditional probability, we need to have a **mechanism to run code conditionally**. In Python, we can do this with an **if-statement**:

```
1 if red == 2:
2   print("The red die rolled a 2.")
```

Four key ideas:

1. [Identical to Pandas]:

2. [Control Flow]:

3. [Syntax]:

4. [Conditional]:

If statements can have an _____ statement for code for cases when the statement is false:

```
1 if red == 2:
2   print("The red die rolled a 2.")
3 else:
4   print("The red die did NOT roll a 2.")
```

Example 4: Write a simulation for the ELISA problem.

If a blood sample has the AIDS virus, then there is a 99% probability that ELISA will correctly give a positive result. If the blood sample does not have the AIDS virus then there is still a 6% probability that ELISA will incorrectly give a positive result. About 1% of the blood samples ELISA tests are truly contaminated.

Practice Problems:

Practice #1: *The following question combines all of the probability rules that we've learned. Suppose you randomly draw from these students.*

	Left-Handed	Ambidextrous	Right-Handed	Totals
Male	**30**	**20**	**266**	**316**
Female	**60**	**27**	**560**	**647**
Totals	**90**	**47**	**826**	**963**

a) What is the chance of getting a female?

b) What is the chance of getting someone who is left-handed?

c) What is the chance you'll get a female if you draw only from the left-handers?

d) What is the chance you'll get a left-hander if you draw only from the females?

e) Draw 3 students *without* replacement. What is the chance that *all* 3 students are left-handed?

f) Draw 3 students *without* replacement. What's the chance that *not all* 3 students are left-handed?

g) Draw 3 students *with* replacement. What is the chance that *at least one* student is right-handed?

Practice #2: What is the probability of getting at least one 5 on six rolls of a die?

Practice #3: What is the probability of rolling a die 3 times and *not* getting all "2"s?

Data Science DISCOVERY – Things To Be Doing:

1. HW 5 is up on PrairieLearn and due Sunday at 11:59pm; HW6 out this weekend
2. lab_birthday due Monday, Oct. 14
3. Open Office Hours: Every M/W/R/F from 4:00pm - 6:30pm in 23 Illini Hall
- Professor office hours Wednesdays, 8:30am-10am in 2215 Siebel Center

4. Extra Credit +1 Notebook every lecture: out after lecture, due 11:30am before next lecture