Bayes' Rule

Foundations of Data Analysis

January 24, 2019

Brain Teaser: Trick Coin

I have four coins. Three are normal, one side heads, one side tails. One is a trick coin where both sides are heads. I pick one coin at random and flip it. If it shows heads, what is the probability that it is the trick coin?

Bayes' Rule

Let's us "flip" a conditional:

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

Deriving Bayes' Rule

Multiplication rule:

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

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

But these two equations are equal, so:

$$P(B \mid A)P(A) = P(A \mid B)P(B)$$

Dividing both sides by $P(A)$ gives us:

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

Trick Coin Example

A = “heads”, B = “trick coin”

$P(A \mid B) = 1.0$

$P(B) = 0.25$

$$
\begin{aligned}
P(A) &= P(A \mid B)P(B) + P(A \mid B^c)P(B^c) \\
&= 1.0 \times 0.25 + 0.5 \times 0.75 = \frac{5}{8}
\end{aligned}
$$

$$
P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A)} = \frac{1.0 \times 0.25}{5/8} = \frac{2}{5} = 0.4
$$

Random Variables

Definition

A **random variable** is a function defined on a sample space, Ω. Notation: $X : \Omega \to \mathbb{R}$

- A random variable is neither random nor a variable.
- Just think of a random variable as assigning a number to every possible outcome.
- For example, in a coin flip, we might assign “tails” as 0 and “heads” as 1:

$$X(T) = 0, \qquad X(H) = 1$$

Dice Example

Let $(\Omega, \mathcal{F}, P)$ be the probability space for rolling a pair of dice, and let X be the random variable that gives the sum of the numbers on the two dice. So,

$$X[(1,2)] = 3, \quad X[(4,4)] = 8, \quad X[(6,5)] = 11$$

Even Simpler Example

Most of the time the random variable X will just be the identity function. For example, if the sample space is the real line, $\Omega = \mathbb{R}$, the identity function

$$\begin{aligned} X &: \mathbb{R} \to \mathbb{R}, \\ X(s) &= s \end{aligned}$$

is a random variable.

Defining Events via Random Variables

Setting a real-valued random variable to a value or range of values defines an event.

$$[X = x] = \{s \in \Omega : X(s) = x\}$$
$$[X < x] = \{s \in \Omega : X(s) < x\}$$
$$[a < X < b] = \{s \in \Omega : a < X(s) < b\}$$

Joint Probabilities

Two binary random variables:
$C =$ cold / no cold $= (1/0)$
$R =$ runny nose / no runny nose $= (1/0)$

Event $[C = 1]$: "I have a cold"
Event $[R = 1]$: "I have a runny nose"

Joint event
$[C = 1] \cap [R = 1]$: "I have a cold and a runny nose"

Notation for joint probabilities:

$$P(C = 1, R = 1) = P([C = 1] \cap [R = 1])$$

Cold Example: Probability Tables

Two binary random variables:
C = cold / no cold $= (1/0)$
R = runny nose / no runny nose $= (1/0)$

Joint probabilities:

		C = 0	C = 1
R	0	0.40	0.05
	1	0.30	0.25

Cold Example: Marginals

		C = 0	C = 1
R	0	0.50	0.05
	1	0.20	0.25

Marginals:

$$P(R=0) = 0.55, \quad P(R=1) = 0.45$$

$$P(C=0) = 0.70, \quad P(C=1) = 0.30$$

Cold Example: Conditional Probabilities

		$C=0$	$C=1$	
R	0	0.50	0.05	0.55
	1	0.20	0.25	0.45
		0.7	0.3	

Conditional Probabilities:

$$P(C=0 \mid R=0) = \frac{P(C=0, R=0)}{P(R=0)} = \frac{0.50}{0.55} \approx 0.91$$

$$P(C=1 \mid R=1) = \frac{P(C=1, R=1)}{P(R=1)} = \frac{0.25}{0.45} \approx 0.56$$

Probabilistic Logic

Deductive Logic

How about *modus tollens*?

$A \Rightarrow B$	If it's raining, then the sidewalk is wet.
B is false	The sidewalk is not wet.
A is false	It is not raining.

Conditional Probability as Logic

Logic	**Probability**
A, B are propositions	A, B are events
$A \Rightarrow B$	$P(B \mid A) > P(B)$

Weak form of *modus ponens*:

If A is true, B becomes more likely.

A is true.

B is more likely.

Bayesian Logic

Unlike Boolean logic, we can *flip* the implication!

$$P(B \mid A) > P(B) \qquad \text{given}$$

$$\frac{P(A)P(B \mid A)}{P(B)} > P(A) \qquad \text{multiply by } \frac{P(A)}{P(B)}$$

$$P(A \mid B) > P(A) \qquad \text{Bayes' Rule}$$

Flipping the implication: $P(B \mid A) > P(B)$

If A is true, B becomes more likely.

B is true.

A is more likely.

If it's raining, then the sidewalk is more likely to be wet.

The sidewalk is wet.

It's more likely to be raining.

Exercise for You

Given that $P(B \mid A) > P(B)$, show that:

1. If $\bar{B}$ happens, A becomes less likely.
   (weak form of *modus tollens*)
2. If $\bar{A}$ happens, B becomes less likely.

Final Bayesian Logic Rules

Given that $P(B \mid A) > P(B)$, analagous to $A \Rightarrow B$, we have four rules:

1. If A, then B is more likely (weak *modus ponens*)
2. If $\bar{B}$, then A is less likely (weak *modus tollens*)
3. If B, then A is more likely (no logical equivalent)
4. If $\bar{A}$, then B is less likely (no logical equivalent)

Cold Example

		C = 0	C = 1	
R	0	0.50	0.05	0.55
	1	0.20	0.25	0.45
		0.7	0.3	

Remember:
$P(C) = 0.3$
$P(C \mid R) = 0.56$

What if I didn't give you the full table, but just:

$$P(R \mid C) = 0.83 \quad > \quad P(R) = 0.45$$

What can you say about the increase
$P(C \mid R) > P(C)$?

Cold Example

Notice, having a cold *increases* my chance for a runny nose by the factor,

$$\frac{P(R \mid C)}{P(R)} = \frac{0.83}{0.45} = 1.85$$

How does such a ratio increase if I flip the conditional?

$$\begin{aligned}
\frac{P(C \mid R)}{P(C)} &= \frac{P(C \cap R)}{P(R)P(C)} \\
&= \frac{P(R \mid C)}{P(R)} \\
&= 1.85
\end{aligned}$$