COSC 290 Discrete Structures

Lecture 33: Bayes' rule and Independence

Prof. Michael Hay
Monday, Apr. 23, 2018
Colgate University

Plan for today

1. Conditional Probability (Quick review)
2. Total law of probability
3. Bayes' rule

Logistics

- Exam on Friday
- Pset due Wednesday
- I am away all week (but online through Wednesday night)

Conditional Probability (Quick review)

Recall: Conditional Probability

Definition (Conditional probability)

The conditional probability of event A given event B is

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

where it is required that $Pr(B) \neq 0$.

Intuition: it is the probability of A in light of the information that event B has occurred.

Examples:

- Probability that Cavs win finals *given* that they lost first three games.
- Probability that A is pardoned *given* that Warden says B to be executed.

Exercise: conditional probability

Teams A and B compete in a best of 3 series. The teams are equally matched (probability of A winning is 1/2).

The *conditional probability* of event X given event Y is

$$Pr(X|Y) = \frac{Pr(X \cap Y)}{Pr(Y)}$$

However team A is easily demoralized and so in the game following a loss, the probability that A wins drops to 1/3. (In a game following a win, the teams remain evenly matched.) What is the probability that A wins the series given that they lose *at least* one game? (Hint: Use tree diagram and conditional probability.)

Chain rule

The chain rule is the observation that:

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

(This follows directly from definition of conditional probability, but it's very useful!)

You can apply rule multiple times:

$$\begin{aligned} Pr(A \cap B \cap C \cap D) &= Pr(A|B \cap C \cap D)Pr(B \cap C \cap D) \\ &= Pr(A|B \cap C \cap D)Pr(B|C \cap D)Pr(C \cap D) \\ &= Pr(A|B \cap C \cap D)Pr(B|C \cap D)Pr(C|D)Pr(D) \end{aligned}$$

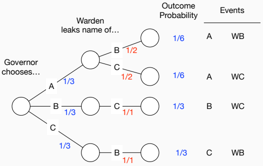
Tree diagrams and conditional probability

In tree diagram, ...

- the edge probability is a *conditional* probability of next event *given* previous events along path to root.
- the outcome probability is the product of these conditional probabilities (using chain rule)

Example: three prisoners problem (next slide)

Three Prisoners Problem



Example: in three prisoners problem, what the Warden says depends on who has been pardoned! (And implicitly on who is asking.)

- $Pr(\text{Warden says C}|\text{A pardoned}) = \frac{1}{2}$
- $Pr(\text{Warden says C}|\text{B pardoned}) = 1$

Total law of probability

Total law of probability

Theorem (Total law of probability)

$$Pr(A) = Pr(A|B)Pr(B) + Pr(A|\bar{B})Pr(\bar{B})$$

Proof:

$$\begin{aligned}
Pr(A) &= Pr((A \cap B) \cup (A \cap \bar{B})) && A = (A \cap B) \cup (A \cap \bar{B}) \\
&= Pr(A \cap B) + Pr(A \cap \bar{B}) && \text{events } A \cap B \text{ and } A \cap \bar{B} \text{ are disjoint} \\
&= Pr(A|B)Pr(B) + Pr(A|\bar{B})Pr(\bar{B}) && \text{chain rule}
\end{aligned}$$

Example application of total law

Consider a deck of cards, well shuffled. What is the probability that the *second* card drawn is an Ace?

Let A be the event the second card is an ace. Let B be that the first card is an ace.

$$\begin{aligned}
P(A) &= Pr(A|B)Pr(B) + Pr(A|\bar{B})Pr(\bar{B}) \\
&= Pr(A|B)\frac{4}{52} + Pr(A|\bar{B})\frac{48}{52} \\
&= \frac{3}{51} \cdot \frac{4}{52} + \frac{4}{51} \cdot \frac{48}{52} \\
&= \frac{(3+48) \cdot 4}{51 \cdot 52} = \frac{51 \cdot 4}{51 \cdot 52} = \frac{4}{52}
\end{aligned}$$

Poll: two urns

Suppose there are two urns. Urn 1 has twelve balls and nine of them are red. Urn 2 has eight balls and two of them are red. Suppose you roll a six-sided die to decide which urn to draw from. If you roll a 1, you pick a ball from urn 1; otherwise you pick a ball from urn 2.

What is the probability that you draw a red ball?

A) $\frac{1}{8}$
B) $\frac{1}{4}$
C) $\frac{1}{3}$
D) $\frac{1}{2}$
E) $\frac{3}{4}$

Bayes' rule

Probabilistic reasoning in the real world

1. Rate of breast cancer among women 40 to 50 years old with no symptoms nor genetic history is 0.8%.
2. Among breast cancer patients, mammogram test reports positive 90% of time.
3. Among cancer-free patients, mammogram test reports positive 7% of time.
4. Imagine woman matching above criteria has a positive mammogram. What is the probability that she has breast cancer?

What does the doc say?

Previous question was posed to doctors as part of a study on risk estimation. Here is the reported reaction of one study participant, a department chief at teaching hospital with more than 30 years experience:

> *"[He] was visibly nervous while trying to figure out what he would tell the woman. After mulling the numbers over, he finally estimated the woman's probability of having breast cancer, given that she has a positive mammogram, to be 90 percent. Nervously, he added, 'Oh, what nonsense. I can't do this. You should test my daughter; she is studying medicine.'"*

http://opinionator.blogs.nytimes.com/2010/04/25/chances-are/

Revisiting example

1. Rate of breast cancer among women 40 to 50 years old with no symptoms nor genetic history is 0.8%.
2. Among breast cancer patients, mammogram test reports positive 90% of time.
3. Among cancer-free patients, mammogram test reports positive 7% of time.
4. Imagine woman matching above criteria has a positive mammogram. What is the probability that she has breast cancer?

Let's formalize the problem:

- Let A be the event the woman has breast cancer.
- Let B be the event that the mammogram test is positive.
- What probability are we being asked to calculate?
- What information are we given?

Bayes' Rule

For any two events A and B,

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

Proof:

$$\begin{aligned} Pr(A|B) &= \frac{Pr(A \cap B)}{Pr(B)} && \text{definition of conditional probability} \\ &= \frac{Pr(B \cap A)}{Pr(B)} && A \cap B = B \cap A \\ &= \frac{Pr(B|A)Pr(A)}{Pr(B)} && \text{chain rule} \end{aligned}$$

We can also "simplify" the denominator using the total law of probability:

$$Pr(B) = Pr(B|A)Pr(A) + Pr(B|\bar{A})Pr(\bar{A})$$

Answer to probability puzzle

Answer worked out on board, around 10%.

Why Bayes' rule is so powerful

Think of A as a cause and B as an observed effect (e.g., symptom).

We often want to know $Pr(A|B)$. (Diagnosis: how likely is that the effects are due to this cause?)

Bayes' rule lets us reason about this unknown using things we know or can *estimate* from data:

- $Pr(B|A)$ – how often this symptom occurs when A is present
- $Pr(B|\bar{A})$ – how often this symptom occurs when A is absent
- $Pr(A)$ – how common is A?

Survey says

In a survey of American doctors, 95 out of 100 estimated the woman’s probability of having breast cancer to be somewhere around 75 percent.

http://opinionator.blogs.nytimes.com/2010/04/25/chances-are/