Bayes Theorem

Hamid Dehghani
School of Computer Science
Birmingham
September 2020

UNIVERSITY OF BIRMINGHAM

Bayesian Probability

•**"Probability"**: often used to refer to frequency

… but

•**Bayesian Probability:** a measure of a state of knowledge.

•It <u>quantifies uncertainty</u>. Allows us to reason using uncertain statements.

•A Bayesian model is continually updated as more data is acquired.

How did this come about?

•Billiard Table:

•A white billiard ball is rolled along a line and we look at where it stops.

•We suppose that it has a uniform probability of falling anywhere on the line. It stops at a point p.

•A red billiard ball is then rolled n times under the same uniform assumption.

•How many times does the red ball roll further than the white ball?

Bayes' Theorem

Bayes' Theorem shows the relationship between a conditional probability and its inverse.

i.e. it allows us to make an inference from
the probability of a hypothesis given the evidence to
the probability of that evidence given the hypothesis
and vice versa

Bayes' Theorem

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

P(A) – the PRIOR PROBABILITY – represents your knowledge about A before you have gathered data.

e.g. if 0.01 of a population has schizophrenia then the probability that a person drawn at random would have schizophrenia is 0.01

Bayes' Theorem

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

P(B|A) – the CONDITIONAL PROBABILITY – the probability of B, given A.

e.g. you are trying to roll a total of 8 on two dice. What is the probability that you achieve this, given that the first die rolled a 6?

Bayes' Theorem

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

So the theorem says:

- The probability of A given B is equal to the probability of B given A, times the prior probability of A, divided by the prior probability of B.

A Simple Example

•Mode of transport:	Probability he is late:
•Car	50%
•Bus	20%
•Train	1%

- Suppose that Bob is late one day.
- His boss wishes to estimate the probability that he traveled to work that day by car.

- He does not know which mode of transportation Bob usually uses, so he gives a prior probability of 1 in 3 to each of the three possibilities.

A Simple Example

- $P(A|B) = P(B|A)\ P(A) / P(B)$
- $P(car|late) = P(late|car) \times P(car) / P(late)$

- $P(late|car) = 0.5$ (he will be late half the time he drives)
- $P(car) = 0.33$ (this is the boss' <u>assumption</u>)
- $P(late) = 0.5 \times 0.33 + 0.2 \times 0.33 + 0.01 \times 0.33$
- (all the probabilities that he will be late added together)

- $P(car|late) = 0.5 \times 0.33 / 0.5 \times 0.33 + 0.2 \times 0.33 + 0.01 \times 0.33$

  $= 0.165 / 0.71 \times 0.33$

  $= 0.7042$

So why is Bayesian probability useful?

- It allows us to put probability values on unknowns. We can make logical inferences even regarding uncertain statements.

- This can show counterintuitive results – e.g. that the some test may not be useful.