SNS COLLEGE OF ENGINEERING

Kurumbapalayam(Po), Coimbatore – 641 107
Accredited by NAAC-UGC with 'A' Grade
Approved by AICTE, Recognized by UGC & Affiliated to Anna University, Chennai

Department of Information Technology

19IT601 – Data Science and Analytics

III Year / VI Semester

Unit 2 – DESCRIPTIVE ANALYTICS USING STATISTICS

Topic 5: Bayes' Theorem

Conditional Probability

- Conditional probability is a way to measure the relationship between two things happening to each other.
- In mathematical notation, the way we indicate things here is that P(A,B) represents the probability of both A and B occurring independent of each other.
- That is, what's the probability of both of these things happening irrespective of everything else.
- Whereas this notation, P(B|A), is read as the probability of B given A. So, what is the probability of B given that event A has already occurred

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

Conditional Probability

- The probability of B given A is equal to the probability of A and B occurring over the probability of A alone occurring, so this teases out the probability of B being dependent on the probability of A

- A as the probability of passing the first test, and B as the probability of passing the second test. What I'm looking for is the probability of passing the second test given that you passed the first, that is, P (B|A).

$$P(B|A) = \frac{P(A,B)}{P(A)} = \frac{0.6}{0.8} = 0.75$$

Bayes' Theorem

- The Bayes theorem is a mathematical formula for calculating conditional probability in probability and statistics.
- In other words, it's used to figure out how likely an event is based on its proximity to another.
- Simply put, it is a way of calculating conditional probability.
- The probability of A given B is equal to the probability of A times the probability of B given A over the probability of B.
- The key insight is that the probability of something that depends on B depends very much on the base probability of B and A.

Bayes’ Theorem

We can find conditional probability using Bayes’ Theorem with the following formula:

- The components has special names:

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

likelihood

posterior

prior

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

evidence

- ‘A’ is the event of interest.

- P(A) represents our prior belief: probability of event A occurring.

- With new evidence B, the posterior belief or updated probability is represented P(A|B): probability of event A given evidence B has occurred.

- P(B | A) is the conditional probability of event B occurring, given that A is true

Bayes' Theorem

Example 1

John flies frequently and likes to upgrade his seat to first class. He has determined that if he checks in for his flight at least two hours early, the probability that he will get an upgrade is 0.75; otherwise, the probability that he will get an upgrade is 0.35. With his busy schedule, he checks in at least two hours before his flight only 40% of the time. Suppose John did not receive an upgrade on his most recent attempt. What is the probability that he did not arrive two hours early?

Let B = {John arrived at least two hours early}, and A = {John received an upgrade},
then ¬B = {John did not arrive two hours early}, and ¬A = {John did not receive an upgrade}.

John checked in at least two hours early only 40% of the time, or $P(B)=0.4$

Therefore $P(\neg B) = 1 – P(B) = 1- 0.4 = 0.6$.

The probability that John received an upgrade given that he checked in early is 0.75, or $P(A|B)=0.75$.

The probability that John received an upgrade given that he did not arrive two hours early is 0.35, or $P(A|\neg B) = 0.35$

Bayes' Theorem

Therefore

$P(\neg A \mid \neg B) = 0.65$

The probability that John received an upgrade P(A) can be computed as shown

$$
\begin{aligned}
P(A) &= P(A \cap B) + P(A \cap \neg B) \\
&= P(B)*P(A|B) + P(\neg B)*P(A|\neg B) \\
&= 0.4 * 0.75 + 0.6 * 0.35 \\
&= 0.51
\end{aligned}
$$

Thus, the probability that John did not receive an upgrade

$P(\neg A) = 1-0.51 = 0.49$

Bayes' Theorem

Exercise

Assume that a patient named Mary took a lab test for a certain disease and the result came back positive. The test returns a positive result in 95% of the cases in which the disease is actually present, and it returns a positive result in 6% of the cases in which the disease is not present. Furthermore, 1% of the entire population has this disease. What is the probability that Mary actually has the disease, given that the test is positive?

Bayes’ Theorem

Let B = {having the disease} and A = {testing positive}. The goal is to solve the probability of having the disease, given that Mary has a positive test result, P(B|A). From the problem description,

$P(B) = 0.01, P(\neg B) = 0.99$

$P(A|B) = 0.95$ and $P(A|\neg B) = 0.06$

Bayes’ theorem defines

$P(B|A) = P(A|B) * P(B) / P(A)$

The probability of testing positive, that is P(A) , needs to be computed first. That computation is shown below

$$
\begin{aligned}
P(A) &= P(A \cap B) + P(A \cap \neg B) \\
&= P(B)*P(A|B) + P(\neg B)*P(A|\neg B) \\
&= 0.01 * 0.95 + 0.99 * 0.06 \\
&= 0.0689
\end{aligned}
$$

According to Bayes’ theorem, the probability of having the disease, given that Mary has a positive test result, is

$$
\begin{aligned}
P(B|A) &= P(A|B) * P(B) / P(A) \\
&= 0.95 * 0.01 / 0.0689 \\
&= 0.1379
\end{aligned}
$$

THANK YOU