

Stat 136 (Bayesian Statistics)

Lesson 1.2: The Bayes Theorem

Conditional Probability and Law of Total Probability

Recall that by definition, the conditional probability of A given B , is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \text{ provided } P(B) > 0. \quad (1)$$

Similarly, we have

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \text{ provided } P(A) > 0. \quad (2)$$

By multiplication, we obtain from (2),

$$P(A \cap B) = P(B|A) \times P(A) \quad (3)$$

Recall further that the total probability, $P(B)$, can be computed as,

$$P(B) = P(B|A) \times P(A) + P(B|\bar{A}) \times P(\bar{A}) \quad (4)$$

where: $P(\bar{A}) = 1 - P(A)$.

The Bayes Theorem: The Key to Bayesian Statistics

Applying (3) and (4) to (1) gives rise to the Bayes formula

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B|A) \times P(A) + P(B|\bar{A}) \times P(\bar{A})} \quad (5)$$

As regards terminology, we call $P(A)$ the **prior** probability of A (meaning the probability of A before B is known to have occurred), and we call $P(A|B)$ the **posterior** probability of A given B (meaning the probability of A after B is known to have occurred).

We may also say that $P(A)$ represents our *a priori* beliefs regarding A , and $P(A|B)$ represents our *a posteriori* beliefs regarding A .

The above result can be easily extended to more than 2 partitions of set B . Let A_1, A_2, \dots, A_k forms a partition of set B . Then for any $i = 1, 2, \dots, k$

$$\begin{aligned} P(A_i|B) &= \frac{P(A_i \cap B)}{P(B)} \\ &= \frac{P(B|A_i) \times P(A_i)}{\sum_{j=1}^k P(B|A_j) \times P(A_j)} \end{aligned} \quad (6)$$

Example 1

The incidence of a disease in the population is 1%. A medical test for the disease is 90% accurate in the sense that it produces a false reading 10% of the time, both: (a) when the test is applied to a person with the disease; and (b) when the test is applied to a person without the disease. A person is randomly selected from population and given the test. The test result is positive (i.e. it indicates that the person has the disease). What is the probability that the person actually has the disease?

SOLUTION:

Let A be the event that the person has the disease, and let B be the event that they test positive for the disease. Then:

- $P(A) = 0.01$ (the prior probability of the person having the disease)
- $P(B|A) = 0.9$ (the *true positive* rate, also called the *sensitivity* of the test)
- $P(\bar{B}|\bar{A}) = 0.9$ (the *true negative* rate, also called the *specificity* of the test)

So,

$$P(A \cap B) = P(A) \times P(B|A) = 0.01 \times 0.9 = 0.009 \text{ and}$$

$$P(\bar{A} \cap B) = P(\bar{A}) \times P(B|\bar{A}) = 0.99 \times 0.1 = 0.099.$$

Thus, the unconditional (or *prior*) probability of the person testing positive is

$$P(B) = P(A) \times P(B|A) + P(\bar{A}) \times P(B|\bar{A}) = 0.009 + 0.099 = 0.108$$

So the required posterior probability of the person having the disease is

$$\begin{aligned}P(A|B) &= \frac{P(A) \times P(B|A)}{P(A) \times P(B|A) + P(\bar{A}) \times P(B|\bar{A})} \\&= \frac{0.009}{0.108} \\&\approx 0.083\end{aligned}$$

Example 2

In a particular population:

- 10% of persons have Type 1 blood, and of these, 2% have a particular disease;
- 30% of persons have Type 2 blood, and of these, 4% have the disease;
- 60% of persons have Type 3 blood, and of these, 3% have the disease.

A person is randomly selected from the population and found to have the disease. What is the probability that this person has Type 3 blood?

SOLUTION:

Let

- A = ‘The person has Type 1 blood’
- B = ‘The person has Type 2 blood’
- C = ‘The person has Type 3 blood’
- D = ‘The person has the disease’

Then,

- $P(A) = 0.1, P(B) = 0.3, P(C) = 0.6$
- $P(D|A) = 0.02, P(D|B) = 0.04, P(D|C) = 0.03$

By the law of total probability we have

$$\begin{aligned}P(D) &= P(D|A) \times P(A) + P(D|B) \times P(B) + P(D|C) \times P(C) \\&= 0.02 \times 0.1 + 0.04 \times 0.3 + 0.03 \times 0.6 \\&= 0.032\end{aligned}$$

Therefore,

$$\begin{aligned}
P(C|D) &= \frac{P(D|C) \times P(C)}{P(D)} \\
&= \frac{0.03 \times 0.6}{0.032} \\
&= 0.5625
\end{aligned}$$

Conclusion

- In the right side of (5), $P(B|A_i)$ represents the *likelihood* and $P(A_i)$ is the *prior*; while, the denominator serves as the *normalizing* constant to make sure that *posterior* probabilities sum up to 1
- Thus, we can view the posterior $P(A_i|B)$ as proportional to the product of the *likelihood* and the *prior*, that is,

$$P(A_i|B) \propto P(B|A_i) \times P(A_i) = \text{likelihood} \times \text{prior}$$