Essentials of Data Science With R Software - 1

Probability and Statistical Inference

Probability Theory

:::

Lecture 19
Bayes' Theorem

Shalabh

Department of Mathematics and Statistics

Indian Institute of Technology Kanpur

Law of Total Probability:

Assume that $A_1, A_2, \dots, A_m$ are events such that

- $A_1 \cup A_2 \cup \dots \cup A_m = \Omega$,
- $A_i \cap A_j = \phi$ **(pairwise disjoint) for all** $i \neq j = 1, 2, \dots, m$, **and**
- $P(A_i) > 0$ **for all** i,

then the probability of an event B can be calculated as

$$P(B) = \sum_{i=1}^{m} P(B|A_i)P(A_i)$$

Bayes' Theorem:

Bayes' Theorem gives a connection between $P(A|B)$ and $P(B|A)$.

Consider an example to understand the importance of prior probabilities and Bayes' theorem.

Claim: A blood test for checking the presence/absence of a rare disease is developed with following probabilities:

Events A: Outcome of test is positive

Event D: Person has disease

Bayes' Theorem:

P(Person has disease and test is positive) = $P(A|D)$ = 0.999

P(Person don't has disease and test is negative) = $P(\bar{A} \mid \bar{D})$ = 0.999

Seems to be a good test for a naive person.

Select a person and make a test.

The probability that the person has a disease = $P(D)$

Usually, $P(D)$ is small, say $P(D)$ = 0.0001.

Bayes' Theorem:

We want to know whether the test is good or bad.

$$P(\boldsymbol{D}|\boldsymbol{A}) = \frac{P(\boldsymbol{D})P(\boldsymbol{A}|\boldsymbol{D})}{P(\boldsymbol{D})P(\boldsymbol{A}|\boldsymbol{D}) + P(\overline{\boldsymbol{D}})P(\boldsymbol{A}|\overline{\boldsymbol{D}})}$$

$$= \frac{\mathbf{0.0001}\times\mathbf{0.999}}{\mathbf{0.0001}\times\mathbf{0.999} + (\mathbf{1} - \mathbf{0.0001})\times(\mathbf{1} - \mathbf{0.999})}$$

$$= \mathbf{0.091}\ (\textbf{Not so reliable, too small})$$

So don't rely only on prior probabilities but also look for posterior probabilities.

Bayes' Theorem:

Bayes' Theorem gives a connection between *P*(*A*|*B*) and *P*(*B*|*A*).

For events *A* and *B* with *P*(*A*) > 0 and *P*(*B*) > 0, we get

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

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

Bayes' Theorem:

Assume that $A_1, A_2, \ldots, A_m$ are events such that

- $A_1 \cup A_2 \cup \ldots \cup A_m = \Omega$,
- $A_i \cap A_j = \phi$ **(pairwise disjoint) for all** $i \neq j = 1, 2, \ldots, m$
- $P(A_i) > 0$ **for all** i, **and**
- B **is another event than** A, **then**

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^{m} P(B|A_i)P(A_i)}$$

is known as Bayes' formula (English philosopher Thomas Bayes).

$P(A_i)$ **: Prior probabilities**

$P(A_i|B)$ **: Posterior probabilities**

$P(B|A_i)$ **: Model probabilities**

Bayes' Theorem: Example 1

Suppose someone rents books from two different libraries.

Sometimes it happens that the book is defective due to missing pages.

We consider the following events:

A_i (i = 1, 2): "the book is issued from library i".

Further let B denote the event that the book is available and is not defective.

Assume we know that $P(A_1) = 0.6$ and $P(A_2) = 0.4$ and $P(B|A_1) = 0.95$, $P(B|A_2) = 0.75$ and we are interested in the probability that the rented book from the library is not defective.

Bayes' Theorem: Example 1

We can then apply the law of total probability and get

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) = 0.6 \times 0.95 + 0.4 \times 0.75 = 0.87.$$

We may also be interested in the probability that the book was issued from the library 1 *and* is not defective which is

$$P(B\,A_1) = P(B|A_1)P(A_1) = 0.95 \times 0.6 = 0.57.$$

Now suppose we have a non-defective book issued. What is the probability that it is issued from library 1?

This is obtained as follows:

$$P(\boldsymbol{A_1}|\boldsymbol{B}) = \frac{P(\boldsymbol{A_1B})}{P(\boldsymbol{B})} = \frac{\mathbf{0.57}}{\mathbf{0.87}} = \mathbf{0.6552}.$$

Bayes' Theorem: Example 1

Now assume we have a defective book, i.e. $\overline{B}$ occurs.

The probability that a book is defective given that it is from library 1 is $P(\overline{B}|A_1) = 0.05$.

Similarly, $P(\overline{B}|\ A_2) = 0.25$ for library 2.

We can now calculate the conditional probability that the book is issued from library 1 given that it is defective as follows.

Bayes' Theorem: Example 1

We can now calculate the conditional probability that the book is issued from library 1 given that it is defective:

$$P(\boldsymbol{A_1}|\overline{\boldsymbol{B}}) = \frac{P(\overline{\boldsymbol{B}}|\boldsymbol{A_1})P(\boldsymbol{A_1})}{P(\overline{\boldsymbol{B}}|\boldsymbol{A_1})\mathrm{P}(\boldsymbol{A_1}) + \boldsymbol{P}(\overline{\boldsymbol{B}}|\boldsymbol{A_2})\mathrm{P}(\boldsymbol{A_2})}$$

$$= \frac{\mathbf{0.05 \times 0.6}}{\mathbf{0.05 \times 0.6 + 0.25 \times 0.4}} = \mathbf{0.2308}$$

The result about $P(\overline{B})$ used can also be directly obtained by using

$$P(\overline{B}) = 1 - P(B) = 1 - 0.87 = 0.13.$$

Bayes' Theorem: Example 2

At a certain stage of a criminal investigation, the inspector in charge is 60 % convinced of the guilt of a certain suspect.

Suppose now that a *new* piece of evidence that shows that the criminal has a certain characteristic (such as left-handedness, baldness, brown hair, etc.) is uncovered.

If 20 percent of the population possesses this characteristic, how certain of the guilt of the suspect should the inspector now be if it turns out that the suspect is among this group?

Bayes' Theorem: Example 2

Let

G : Event that the suspect is guilty and

C : Event that he possesses the characteristic of the criminal,

we have

$$P(\boldsymbol{G}|\boldsymbol{C}) = \frac{P(\boldsymbol{GC})}{P(\boldsymbol{C})}$$

Now

$$P(GC) = P(G)P(C|G) = (.6)(1) = 0.6$$

We have supposed that the probability of the suspect having the characteristic if he is, in fact, innocent is equal to 0.2.

Bayes' Theorem: Example 2

To compute the probability that the suspect has the characteristic, we condition on whether or not he is guilty, we find

$$P(C) = P(C|G)P(G) + P(C|\overline{G})P(\overline{G})$$

$$= (1)(0.6) + (0.2)(0.4) = 0.68$$

Hence

$$P(G|C) = \frac{P(GC)}{P(C|G)P(G) + P(C|\overline{G})P(\overline{G})} = \frac{0.60}{0.68} = 0.882$$

and so the inspector should now be 88% certain of the guilt of the suspect.

Bayes' Theorem: Example 3

A plane is missing and it is presumed that it was equally likely to have gone down in any of three possible regions.

The constants p_i are called *overlook probabilities* because they represent the probability of overlooking the plane; they are generally attributable to the geographical and environmental conditions of the regions.

Let $1-p_i$: Probability the plane will be found upon a search of the i^{th} region when the plane is, in fact, in that region, i =1, 2, 3.

What is the conditional probability that the plane is in the i^{th} region, given that a search of region 1 is unsuccessful, i = 1, 2, 3?

Bayes' Theorem: Example 3

Let R_i , i =1, 2, 3, be the event that the plane is in region i; and

let E be the event that a search of region 1 is unsuccessful.

From Bayes' formula, we obtain

$$P(R_1|E) = \frac{P(ER_1)}{P(E)}$$

$$= \frac{P(ER_1)}{P(E|R_1)P(R_1)+P(E|R_2)P(R_2)+P(E|R_3)P(R_3)}$$

$$= \frac{(p_1)(1/3)}{(p_1)(1/3)+(1)(1/3)+(1)(1/3)} = \frac{p_1}{p_1+2}$$

Bayes' Theorem: Example 3

For *j* = 2, 3,

$$P(R_j|E) = \frac{P(ER_j)}{P(E)}$$

$$= \frac{P(ER_j)}{P(E|R_1)P(R_1)+P(E|R_2)P(R_2)+P(E|R_3)P(R_3)}$$

$$= \frac{(1)(1/3)}{(p_1)(1/3)+(1)(1/3)+(1)(1/3)} = \frac{1}{p_1+2}.$$

Thus, for instance, if p_1 = 0.4, then the conditional probability that the plane is in region 1 given that a search of that region did not uncover it is 1/6.

Bayes' Theorem: Example 4

In answering a question on a multiple-choice test, a student either knows the answer or guesses.

Let p be the probability that the student knows the answer and $1 - p$ the probability that the student guesses.

Assume that a student who guesses at the answer will be correct with probability $\frac{1}{m}$, where m is the number of multiple-choice alternatives.

What is the conditional probability that a student knew the answer to a question, given that he or she answered it correctly?

Bayes' Theorem: Example 4

Solution: Let

C **: Events that the student answers the question correctly and**

K **: Event that he or she actually knows the answer.**

Now

$$P(K|C) = \frac{P(KC)}{P(C)} = \frac{P(C|K)P(K)}{P(C|K)P(K)+P(C|\bar{K})P(\bar{K})}$$

$$= \frac{p}{p + (\frac{1}{m})(1-p)} = \frac{mp}{1+(m-1)p}$$

For example, if m = 4, p = 0.5, then the probability that a student knew the answer to a question he or she correctly answered is 4/5.

Bayes’ Theorem: Example 5

A laboratory blood test is 95 percent effective in detecting a certain disease when it is, in fact, present.

However, the test also yields a "false positive" result for 1% of the healthy persons tested. (i.e., if a healthy person is tested, then, with probability .01, the test result will imply he or she has the disease.)

If 0.5% of the population actually has the disease, what is the probability a person has the disease given that the test result is positive?

Bayes' Theorem: Example 5

Solution: Let D : Event that the tested person has the disease and

E : Event that the test result is positive.

The desired probability P(D | E) is obtained as follows:

$$P(D|E) = \frac{P(DE)}{P(E)} = \frac{P(E|D)P(D)}{P(E|D)P(D)+P(E|\bar{D})P(\bar{D})}$$

$$= \frac{(0.95)(0.005)}{(0.95)(0.005)+(0.01)(0.995)} = \frac{95}{294} = 0.323 \text{(Approx.)}$$

Thus only 32% of those persons whose test results are positive actually have the disease.

Surprised!! As we expected this figure to be much higher, since the blood test seems to be a good one.