Essentials of Data Science With R Software - 1

Probability and Statistical Inference

Probability Theory

:::

Lecture 17

Conditional Probability

Shalabh

Department of Mathematics and Statistics

Indian Institute of Technology Kanpur

Conditional Probability:

Conditional probability is useful in calculating probabilities when some partial information concerning the result of the experiment is available, or in recalculating them in light of additional information.

In such situations, the desired probabilities are conditional ones.

Sometimes it is often the easiest way to compute the probability of an event is to first “condition” on the occurrence or non-occurrence of a secondary event.

Conditional Probability:

Consider the following example to understand the concept of conditional probability:

Suppose a blood test is developed to diagnose a particular infection. The blood test is conducted over 100 randomly selected persons.

The outcomes of the absolute and relative frequencies are presented in following Tables:

Conditional Probability:

Absolute frequencies of test results and infection status				
		Infection		Total (row)
		Present	Absent	
Test	Positive (+)	30	10	40
	Negative (-)	15	45	60
Total (Columns)		45	55	Total = 100

Conditional Probability:

Relative frequencies of test results and infection status				
		Infection		Total (row)
		Present (IP)	Absent (IA)	
Test	Positive (T+)	0.30	0.10	0.40
	Negative (T-)	0.15	0.45	0.60
Total (Columns)		0.45	0.55	Total = 1

Conditional Probability:

There are the following four possible outcomes:

1. The blood sample has an infection and the test diagnoses it, i.e. the test is correctly diagnosing the infection.

2. The blood sample does not has any infection and the test does not diagnose it, i.e. the test is correctly diagnosing that there is no infection.

Conditional Probability:

There are the following four possible outcomes:

3. The blood sample has an infection and the test does not diagnose it, i.e. the test is incorrect in stating that there is no infection.

4. The blood sample does not has any infection but the test diagnoses it, i.e. the test is incorrect in stating that there is an infection.

Conditional Probability:

If one already knows that the test is positive and wants to determine the probability that the infection is indeed present, then this can be achieved by the respective conditional probability *P*(*IP*|*T*+) which is

$$P(\boldsymbol{IP}|\boldsymbol{T}+) = \frac{P(\boldsymbol{IP}\cap\boldsymbol{T}+)}{P(\boldsymbol{T}+)} = \frac{\mathbf{0.3}}{\mathbf{0.4}} = \mathbf{0.75}$$

Note that *IP* ∩ *T*+ denotes the "relative frequency of blood samples in which the disease is present and the test is positive" which is 0.3.

Conditional Probability:

Let *P*(*A*) > 0. Then the conditional probability of event *B* occurring,

given that event *A* has already occurred, is

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

The roles of *A* and *B* can be interchanged to define *P(A|B)* as follows.

Let *P*(*B*) > 0. The conditional probability of *A* given *B* is

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

.

Conditional Probability:

The definition of conditional probability is consistent with the interpretation of probability as being a long-run relative frequency, i.e., a large number *n* of repetitions of the experiment are performed.

Conditional Probability: Example 1

A coin is tossed twice. If we assume that all four points in the sample space Ω = {(*H*, *H*), (*H*, *T*), (*T*, *H*), (*T*, *T*)} are equally likely, what is the conditional probability that both tosses result in heads, given that the first toss results in head?

Solution:

If *A* = {(*H*, *H*)} denotes the event that both tosses results in heads, and *B* = {(*H*, *H*), (*H*, *T*)} the event that the first toss results in head, then the desired probability is

$$P(A|B) = \frac{P(A \cap B)}{P(A)} = \frac{P(\{(H,H)\})}{P(\{(H,H),(H,T)\})} = \frac{1/4}{2/4} = \frac{1}{2}$$

Conditional Probability: Example 2

An urn contains 10 white, 5 yellow, and 10 black marbles. A marble is chosen at random from the urn, and it is noted that it is not one of the black marbles. What is the probability that it is yellow?

Solution:

Let *Y* denote the event that the marble selected is yellow, and

$\bar{B}$ denote the event that it is not black.

Now, then the desired probability is $Y\bar{B}$

$$P(Y|\bar{B}) = \frac{P(Y \cap \bar{B})}{P(\bar{B})}$$

Conditional Probability: Example 2

However $Y \cap \overline{B} = Y$, since the marble will be both yellow and not black if and only if it is yellow.

Hence, assuming that each of the 25 marbles is equally likely to be chosen, we obtain that

$$P(Y|\overline{B}) = \frac{P(Y \cap \overline{B})}{P(\overline{B})} = \frac{5/25}{15/25} = \frac{1}{3}$$

Conditional Probability: Example 3

A box contains 5 defective, 10 partially defective (that fail after a couple of hours of use), and 25 acceptable (non-defective) transistors. A transistor is chosen at random from the box and put into use. If it does not immediately fail, what is the probability it is acceptable?

Conditional Probability: Example 3

Solution:

Since the transistor did not immediately fail, we know that it is not one of the 5 defectives and so the desired probability is:

$$P(\textbf{acceptable}|\textbf{not defective}) = \frac{P(\textbf{acceptable, not defective})}{P(\textbf{not defective})}$$

Since the transistor will be both acceptable and not defective if it is acceptable.

$$P(\textbf{acceptable}|\textbf{not defective}) = \frac{P(\textbf{acceptable})}{P(\textbf{not defective})} = \frac{\mathbf{25/40}}{\mathbf{35/40}} = \frac{\mathbf{5}}{\mathbf{7}}$$