3.1 PROBABILITY II

1. Cards

We discuss some examples with $m = 5$ cards drawn from a standard deck of 52 cards. For cards we may calculate ordered or unordered, as long as both the event space and the sample space is calculated in the same way. With ordered selection the sample space has

$$|S| = P(52, 5)$$

elements, and otherwise we get

$$|S| = C(52, 5).$$

What is the probability of getting one pair? We will do the calculation with both types of sample space, first with ordered selection. A pair is described by its rank and its suit. Therfore $13 \cdot 4 \cdot 3$ possibilities for choosing the pair. The pair can be placed among the five cards in $C(5, 2)$ ways. The remaining cards can be chosen by choosing three different ranks, excluding the rank of the pair, and then determining their suits. This gives $12 \cdot 11 \cdot 10 \cdot 4^3$. The probability is therefore

$$P = \frac{13 \cdot 4 \cdot 3 \cdot 12 \cdot 11 \cdot 10 \cdot 4^3}{P(52, 5)} \approx 42\%.$$

With unordered sample space we have the following calculation. There are $13 \cdot C(4, 2)$ possibilities for choosing the pair. The remaining cards can be chosen in $C(12, 3) \cdot 4^3$ ways. This gives the probability

$$P = \frac{13 \cdot C(4, 2) \cdot C(12, 3) \cdot 4^3}{C(52, 5)} \approx 42\%.$$

2. Independence and conditional probabilities

Two events A and B are called **independent** if

$$P(A \cap B) = P(A) \cdot P(B).$$

For instance, if A is the event that the throw of a dice shows 3, and B is the event that the throw of another dice shows 4, then event A has no influence on the probability of event B, and so these two events are independent.

An example of two events which are not independent is as follows. Let A be the event of drawing a diamond from a deck of cards, and let B be the event of drawing a spade on the second card. These two events are not independent. The probability of drawing a spade on the second card depends on which card was drawn on the first. If the first card is a spade, then

$$P(B) = \frac{12}{51}$$

and if the first card is not a spade, then

$$P(B) = \frac{13}{51}$$

The probability $P(B|A)$ is the probability of an event B happening, given that we know that an event A occurs or has occured. We define

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

when $P(A) \neq 0$. A useful point of view is that $P(B|A)$ is the probability of B when we change the sample space of all outcomes from S to A. The formula

$$P(B|A) \cdot P(A) = P(A \cap B),$$

which follows directly from the definition of $P(B|A)$ is also useful.

If A and B are independent, then $P(B|A) = P(B)$, in other words, knowing that A has happened has no effect on the probability of B happening. Similarly, we have $P(A|B) = P(A)$ if A and B are independent.

In general $P(A|B) \neq P(B|A)$. The correct relationship is Bayes' theorem which says

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

Since we are dividing by $P(A)$, we have to assume that $P(A) \neq 0$ for this to make sense. Example (from Wikipedia):

Example 1. *You are looking at the the effectiveness of a new test for a medical condiion. You know that* 0.5 *per cent of the population has this condition. The test gives a positive result in* 99 *per cent of the cases which have the condition and a negative result in* 99 *per cent of the cases which do no have the condition. Assume that the test is positive on a random person. What is the probabilty that the person has the condition?*

Solution:

We first introduce som notation. Let C be the event that a person has the condition, let $+$ be the event that the test is positive and let $-$ be *the event that the test is negative. The question asks for the probability* $P(C|+)$ *which we will compute using Bayes' theorem.*

We have

$$P(C|+) = \frac{P(+|C) \cdot P(C)}{P(+)}.$$

From the question we have the information $P(+|C) = 0.99$ *and* $P(C) = 0.5$*. The only quantity missing is* $P(+)$*. We have*

$$P(+) = P(+ \cap C) + P(+ \cap \overline{C})$$

Using the definition of conditional probability we have

$$P(+ \cap C) = P(+|C) \cdot P(C) = 0.99 \cdot 0.005$$

and

$$P(- \cap C) = P(+|\overline{C}) \cdot P(\overline{C})) = 0.01 \cdot 0.995$$

Putting all this together gives us $P(+|C) = 0.332$ *or* 33.2 *per cent.*

Exercises

Exercise. 1. *You are playing poker and are dealt* 5 *cards. What is the probabilities that you are dealt*

a) *High card*
b) *Pair*
c) *Two pairs*
d) *Flush*
e) *Straight*

Exercise. 2. *Think of all possible outcomes of a two dice throw as a 6 by 6 matrix. Calculate the probability that the first dice show* 4 *given that*

a) *the sum is* 8.
b) *the sum is at most* 8.
c) *the sum is more than* 8.
d) *the product is* 8.

Exercise. 3. *You have developed a program which tests a document for plagarism. You assume that* 1 *per cent of documents are plagarised. The program gives a positive result for* 90 *per cent of plagarised documents and a negative result on* 95 *per cent of documents which are not plagarised. Given an arbitrary document with a positive test result, what is the probability that the document is plagarised?*

Solutions:

Exercise 1.

Google "Poker probability Wikipedia" and look at the relevant page.

Exercise 2.

a) 1/5

b) $4/26 = 2/13$

c) $2/10 = 1/5$

d) 1/2

Exercise 3.

We first fix the notation. Let *plag* be the event of a plagarised document, let $+$ be the event that the test results in a positive result, and $-$ the event that the test gives a negative result.

Then $P(plag) = 0.01$, $P(+|plag) = 0.90$ and $P(-|\overline{plag}) = 0.95$. We calculate $P(+) = P(+|plag) \cdot P(plag) + P(+|\overline{plag}) \cdot P(\overline{plag}) = 0.90 \cdot 0.01 + 0.05 \cdot 0.99 = 0.0585$. The answer is $P(plag|+) = \frac{P(+|plag) \cdot P(plag)}{P(+)} = \frac{0.90 \cdot 0.01}{0.0585} = 0.1538$ or 15 per cent.