Probability paradoxes

Ivan Murashko

Contents

Introduction

The goal for the article is to demonstrate several paradoxes that are related to probability theory and how can they can be solved.

Chapter 1

Base definitions of probability theory

I am going to provide several definitions. I will give the both formal and informal definitions and show how they are related each other.

1.1 Example and motivation

We will start with the simplest example.

Example 1.1. In the example we have (see fig. 1.1) $N = 5$ balls. There are $N_G = 2$ green balls and N_R red balls. I.e. $N = N_G + N_R$.

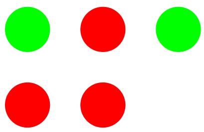

Figure 1.1: Probability example

We can define the probability to get green ball as

$$P_G = \frac{N_G}{N} = \frac{2}{5}$$

and the probability to get red ball as

$$P_R = \frac{N_R}{N} = \frac{3}{5}.$$

We can get only a red or a green ball and

$$P_G + P_R = 1.$$

We will formalize the probability calculations for such cases via the counting principle and define the probability as a ratio of number desired outcomes $|A|$ to the number of all possible outcomes $|\Omega|$:

$$P = \frac{|A|}{|\Omega|}.$$

There is a simple example on counting principle below.

Example 1.2 (Fish in a pond)**.** The example is from a Russian Biological Olympiad. Consider a pond with fishes. 15 of them were marked. After sometime we took 15 fishes and 5 of them were marked. How many fishes are in the pond.

The accepted answer was 45 with the following explanation:

$$\frac{15}{5} = \frac{n}{15}$$

therefore $n = 45$.

Lets try to solve the task with probability theory and convert the question to the following one: How many fishes are in the pond if in every 15 fishes with high probability we get 5 marked ones?

Let n the number of fishes in the pond. Then the size of the sample space $|\Omega|$ is the following

$$|\Omega| = \binom{n}{15}$$

i.e. how many ways to get 15 fishes from n.

The desired outcome has the size $|A|$ combined from 2 ones: getting 5 fishes from 15 and get the rest (10) from non marked fishes: $n - 15$:

$$|A| = \binom{15}{5} \cdot \binom{n-15}{10}.$$

Therefore the result can be calculated as follows [1]

$$P_n = \frac{|A|}{|\Omega|} = \frac{\binom{15}{5} \cdot \binom{n-15}{10}}{\binom{n}{15}}. \tag{1.1}$$

Quick calculations [1] show that $n = 45$ is very close to real answer:

```
$ stack repl
> map fish_in_pond [55, 50, 45, 40, 35, 30]
[0.21391501072376837,0.24492699593153727,0.26162279575176545,
 0.2440273978093778,0.17082265318953427,5.813662441225208e-2]
```

[1] Alternative way to calculate it can be found in Fish in a pond (conditional) (Example 1.12) .

Example 1.3 (Birthday paradox)**.** Consider n people and give a prediction that there is at least one pair of people who have a birthday at the same day. We will call the event as A and there fore are required to find $P(A)$. The straightforward calculations are not easy and we will try to find a probability for another event A^c that is the complement of event A. I.e. event A^c states that there is no such pairs and all birthdays are different. We can calculate the desired probability via

$$P_n = P(A) = 1 - P(A^c).$$

We will use the standard counting here. There are totally 365^n possible options for different birthdays in the group on n people. The number of outcomes that satisfied A^c is

$$365 \cdot 364 \cdot \dots \cdot (365 - n + 1) = \prod_{i=1}^{n} (365 - i + 1).$$

Therefore

$$P(A^c) = \frac{\prod_{i=1}^{n} (365 - i + 1)}{365^n}$$

and

$$P_n = 1 - \frac{\prod_{i=1}^{n} (365 - i + 1)}{365^n}.$$

Calculations gives us $P_{10} = 0.1169, P_{30} = 0.7063, P_{60} = 0.9941$. I.e. we can say with high probability that in the group of 60 people there will be at least 2 persons with the same birthday.

1.2 Definitions

Now we are ready to give several formal definitions.

1.2.1 σ-algebra

Definition 1.4 (Power set)**.** Let Ω is a set than the set of all possible subsets of Ω is called *power set* and denoted as $\mathcal{P}(\omega)$.

Definition 1.5 (σ algebra)**.** Let Ω is a set then a subset $\mathcal{F}$ of Power set $\mathcal{P}(\Omega)$ ($\mathcal{F} \subseteq \mathcal{P}(\Omega)$) is called σ algebra if the following conditions are satisfied:

1. $\mathcal{F}$ contains Ω: $\Omega \in \mathcal{F}$
2. TBD
3. TBD

In our example 1.1, σ algebra is a collection of any balls.

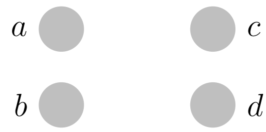

Figure 1.2: Probability space. It consists of elementary events: a, b, c and d, each of them has equal probability $p_{a,b,c,d} = \frac{1}{4}$

1.3 Conditional probability

Definition 1.6 (Conditional probability)**.** The *conditional probability* of event A on event B is defined as follow

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

Example 1.7 (Conditional probability)**.** Lets consider 6 balls each of them can be either two colors (see fig. 1.3).

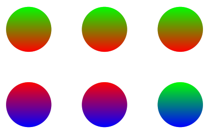

Figure 1.3: Condition probability. Original probability space. $P(A = red) = \frac{5}{6}$, $P(A = blue) = \frac{3}{6}$, $P(A = green) = \frac{4}{6}$

You can see that the probability $P(A)$ to get red ball is $P(A = red) = \frac{5}{6}$, blue one is $P(A = blue) = \frac{3}{6}$, green one is $P(A = green) = \frac{4}{6}$.

Now assume that event A is to get a green ball but event B is to get red ball, how we can define $P(A|B)$ in the case.

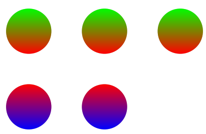

Figure 1.4: Condition probability. $P(A = green|B = red) = \frac{3}{5}$, $P(A = blue|B = red) = \frac{2}{5}$

The situation is displayed on fig. 1.4. We have only 5 possibilities to choose a ball now instead of 6 in the original case. This is because we just got an additional information - "one of the color should be red". Only 3 of the 5 balls are green. Therefore $P(A|B) = P(A = green|B = red) = \frac{3}{5}$.

This result is in correlation with the formal definition of Conditional probability:

$$P(A|B) = \frac{P\left(A \cap B\right)}{P(B)} = \frac{P\left(A = green \cap B = red\right)}{P(B = red)} = \frac{3/6}{5/6} = \frac{3}{5}.$$

The fig. 1.5 gives as the view if event $B = blue$ occurs.

Figure 1.5: Condition probability. $P(A = red|B = blue) = \frac{2}{3}$, $P(A = green|B = blue) = \frac{1}{3}$

In the case we have the following conditional probabilities: $P(A = red|B = blue) = \frac{2}{3}$, $P(A = green|B = blue) = \frac{1}{3}$.

Finally, the fig. 1.6 gives as the view if event $B = green$ occurs.

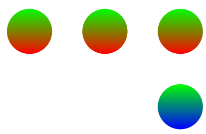

Figure 1.6: Condition probability. $P(A = blue|B = green) = \frac{1}{4}$, $P(A = red|B = green) = \frac{3}{4}$

Example 1.8 (The King's sibling)**.** Suppose that we have a king from a family of 2 children. What's the probability that his sibling is a girl. The important assumption [2] that has to be made is the following: there is no any family planning in the king family and the probability to get a boy P_b and probability to get a girl P_b are equally likely:

$$P_b = P_g = \frac{1}{2}.$$

We have 4 cases: bb, bg, gb, gg and the condition that the king is a boy pick up only 3 options for us: bb, bg, gb. All of them are equally likely and 2 have a girl as sibling. I.e.

$$P(sibling = girl|king) = \frac{2}{3}.$$

[2] For instance if the king family assume to get a new child until the first boy (king) get then we will have the sibling is girl with probability 1.

Proposition 1.9 (Total probability)**.** *The total probability is defined as follows*

$$P(A) = \sum_i P(A|B_i)$$

Example 1.10 (Total probability)**.** Lets assume in the Conditional probability (Example 1.7) that we are interested in the event A that the ball is green. The other color will be either blue or red. I.e. $B_1 = blue$, $B_2 = red$.

$$\begin{aligned} P(A = green) = P(A = green|B = blue)P(B = blue) + \\ +P(A = green|B = red)P(B = red) = \\ = \frac{1}{3} \cdot \frac{1}{2} + \frac{3}{5} \cdot \frac{5}{6} = \frac{4}{6}. \end{aligned}$$

I.e. formula works.

Consider another, not so simple example

Example 1.11 (Total probability paradox)**.** Let we have 6 balls each of them has one color: red or green (see fig. 1.7).

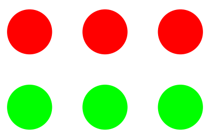

Figure 1.7: Total probability example

Lets event A is an event to get a ball. $P(A) = \frac{1}{6}$. The event B_1 is an event to get green ball: $P(B_1) = \frac{1}{2}$. The same one is for probability to get red ball: $P(B_2) = \frac{1}{2}$. Conditional probabilities can be calculated as follows:

$$P(A|B_1) = P(A|B_2) = \frac{1}{3}. \tag{1.2}$$

As result the total probability is

$$P(A) = \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{2}{6} \neq \frac{1}{6}.$$

The error is in the (1.2). When we consider a concrete ball then it either green or blue and as result one of the conditional probabilities $P(A|B_1)$ or $P(A|B_2)$ is zero. In the case we will get correct answer $P(A) = \frac{1}{6}$.

Lets calculate the required probability from Fish in a pond (Example 1.2) by means of conditional probability tools. This will show us that a same task can be solved in different ways

Example 1.12 (Fish in a pond (conditional))**.** First of all calculate a probability of the following collection

$$A = MMMMMNNNNNNNNNN$$

where M means marked fish and N - non marked. I.e. first 5 fishes were marked and last 10 non marked. The probability to get a marked fish is

$$P_1 = \frac{15}{n}$$

but if we catch a fish then the probability (conditional) to get a new marked fish is

$$P_2 = \frac{14}{n-1}$$

i.e. the probability to catch i-th marked fish is

$$P_i = \frac{15 - i + 1}{n - i + 1}.$$

The probability to catch the first non marked fish is

$$Q_1 = \frac{n - 15}{n - 5},$$

i-th

$$Q_i = \frac{n - 15 - i + 1}{n - 5 - i + 1}.$$

The final probability for event A is

$$P(A) = \frac{\prod_{k=11}^{15} k}{\prod_{i=1}^{15} (n - i + 1)} \prod_{i=1}^{10} (n - 15 - i + 1).$$

The desired collection is any collection with 5 Ms and 10 Ns. All of them have the same probability i.e. the final probability is

$$P = \binom{15}{5} P(A) = \binom{15}{5} \frac{\prod_{k=11}^{15} k}{\prod_{i=1}^{15} (n - i + 1)} \prod_{i=1}^{10} (n - 15 - i + 1) =$$

$$= \binom{15}{5} \frac{\frac{15!}{10!}}{\frac{n!}{(n-15)!}} \frac{(n-15)!}{(n-25)!} =$$

$$= \binom{15}{5} \frac{15!(n-15)!}{n!} \frac{(n-15)!}{10!(n-25)!} = \frac{\binom{15}{5} \cdot \binom{n-15}{10}}{\binom{n}{15}}.$$

That is the same as the equation got in (1.1).

Definition 1.13 (Independence)**.** Two events A and B are independent if

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

Example 1.14 (Non independent events)**.** Consider situation shown in fig. 1.7. Let event A is that ball is green, event B is that ball is red. We have

$$A \cap B = \emptyset,$$

i.e. $P(A \cap B) = 0$. This means that the events cannot be considered as independent accordingly definition 1.13 as soon as $P(A) = P(B) = \frac{1}{2}$.

Really the events are dependent as soon as we can say that A will not occur if B occurs and vice versa.

TBD [2]

Chapter 2

Paradoxes

2.1 Monty Hall problem

TBD

2.2 Waiting time on a bus stop

TBD

Bibliography

[1] Murashko, I. Math experiments in haskell. — https://github.com/ivanmurashko/hsprojects/tree/master/mathexperiments.

[2] А. Н. Колмогоров. Основные понятия теории вероятностей / А. Н. Колмогоров. — Москва: Наука, 1974.