Ödön Vancsó, Budapest

Inverse probabilities in everyday situation (Bayesian-type problems)

Abstract

In this paper the well-known medical diagnosis problem will be discussed from three different points of view. At first, the different types of probability will be analysed which are used as different error rates of the test and the ratio of the infected people relating to the whole population, and the requested chance for the infection in the case of a positive result of the diagnosing test will be discussed. (Concerning the different kinds of probability, see for example [1] and [2].)

Secondly, we deal with questions of the different context of the given information, and the different visualizations of the situation. (See for example [4], [5]). The effects of the visual representations will be analysed in some experiments. These experiments emerge from cooperation with researchers at the Max Planck Institute in Berlin. Some experimental results (without didactical analysis) gained from investigations among different students in Budapest may be traced in [5].

Thirdly, the main point covers the main goal of our project: an interpretation of the result with respect to the tested person and the Board of Health. The discussion of this interpretation leads again into questions of the various interpretations of probability.

Finally, some ideas of a chapter of a new Hungarian textbook [6] dealing with such problems will be summarised. This chapter demonstrates a teaching concept for Bayesian-type problem solving by elementary methods, preparing the Bayes–theorem and related modelling and thinking at an intuitive level.

The theoretical background and the structure of this chapter draw heavily from ideas in [3].

School experiments concerning this textbook have been starting in early 2004; it is too early to report and comment fully on the results.

0. Introduction

In recent years a new reform has started in the Hungarian Mathematics Education. Following some international trends (after the TIMSS and PISA studies), educational politics puts more influence on public education to force a change in school education, above all favouring the practical part of knowledge against the too "academic" side.

Consequently, stochastic gets more weight in the curriculum of school mathematics. In this paper, I suggest a promising way for the teaching of such real problems, which can be handled by modelling with conditional probabilities. Our first example is a modified version of the well-known "Mammography" problem.

The starting problem is:

There is a blood-test to diagnose a specific infection (for example HIV-test) with the following properties:

98% of the infected people will have a positive result and 5% of the non-infected persons will also have a (false) positive result. Let 0.01 (one hundredth)% be the estimated ratio of the infected persons related to the whole population. If a person gets a positive result in this test, what is his chance to be really infected?

1. The analysis of the problem focusing on the resources of the given data

The basis for this chapter is [1] and [2]. There are two types of data in the given problem. The error rates of the test are based normally on medical experiments under surely infected and not infected patients. The ratios come from the *relative frequency* of the different false results at experiences with the test. These are *not given probabilities* but only estimated parameters, characteristics of the situation. Here the frequency aspect of the notion of probability is used, and the reliability of these percents depends on the sample size. The estimated ratio of the rates of infected people related to the whole population is a *subjective probability*, depending on available information and data, and can vary widely among different researchers. The interpretation of an answer of the question depends on the different points of view. In case of the Ministry of Health there is a *frequency-based* interpretation, from the point of view of a tested person a *subjective* interpretation.

In a Hungarian school a short experiment was performed with this problem. There were 32 pupils aged 16 in the class. Our goal was to introduce them to the world of conditional probabilities in real life situations, to practise the model building, working with the model and finally to establish different interpretations of the results. Some remarks were used in our textbook (see section 4. below). They enjoyed the three lessons, were very active and gave many interesting remarks. They suggested for example a more sensitive model for the tested patient, for such cases when we can obtain additional information about the subject with respect to his personal risks. In this case we have to work with another “a priori” probability instead of the estimated ratio of the infected persons related to the whole population. This idea can be the starting-point to introduce elementary notions of Bayes-statistics. In the last few years we are working on an approach towards inferential statistics for the secondary and university level covering classical and Bayesian methods in parallel (see [7], [8], [9]).

We observed that it is very important for the students to be able to differentiate the various interpretations of the notion of probability in real situations. We think it became clear for the students, that the classical frequency-based probability can be used only in the case of the Ministry of Health’s point of view. For a tested person the “classical” result tells nothing. His probability of an infection depends on his lifestyle or his sexual behaviour and therefore his degree of the chance of being infected can be very different. This notion is not part of the classical theory of statistics, probability as degree of belief belonging only to Bayesian-statistics (see for example [10] or [11]).

The authenticity of the classical (and also the Bayesian) numerical result depends on the reliability of the data. For example, labelling the error of a false negative test with α and the error of a false positive test result with β, we can analyse the sensitivity of the result considering the changes of the parameters α and β (see below). Of course it is possible to vary the infection rate also, and to interpret this number as a frequency based probability. If we denote this rate with p, then the required result depends on these three parameters α, β and p. The numerical analysis of this question is part of section 3.below.

2. Different visualisations of the structure of the problem

Three different models will be shortly discussed: the doubletree diagram, Venn-type diagram and data structured in tabular form. In the latter case we used the ideas of [12]. How can we follow the way of thinking inherent to these models? This can be analysed from psychological aspects.

In the case of the doubletree diagram we use only natural numbers calculating from the given percentages and finally the classical Laplace-probability shows the answer. That is a simple procedure, which can be used in many other cases. An advantage of this method is that the procedure uses only integers first of all powers of ten (in this case the counting with percentages is easier), which are sure ground for most students.

The Venn-diagram is a visual image of the situation, which can help to show the reason behind the paradox result. In constructing the diagram we have to pay attention to the ratios, the area of the drawn sets should be proportional to the real data.

The tabular form is a more abstract algorithm than the doubletree diagram. A didactical analysis of this way of problem-solving can be read in [12].

A Hungarian experimental result shows that the most frequently used model by children is the Venn-type diagram (nearly 60%), the doubletree diagram being used by about 24% and the tabular form used only by 15% among the good solutions. This depends on the style of teaching, the Venn-type diagram being the most well-known in Hungarian schools. Some researchers claim (see f. e. [4]) that the frequency format is more advantageous at the case of similar problems.

The main scope of an investigation among Hungarian students in 2002 was to use these different representations of the given data to reveal the inherent ways of thinking, to study the effects of understanding and problem perception (resp. reformulation). The paper [5] contains some interesting results. According to my view, the doubletree diagram has an algorithmic, more dynamical aspect of the representation in comparison to the Venn diagram, which is more a static "iconic" picture but helps the visual understanding of the paradox result. The tabular form is a more symbolical representation, not enactive or iconic. But it is only a step forward to the more abstract symbolical representation in the direction of the Bayes-Theorem. Our experimental result (till now unpublished) shows that the impact of teaching is better, if the students know all representations of the problem and can choose freely that variation which seems best for a situation. These results are in accordance to Bruner's theory about different representations. He stated that learning is more efficient when different representations are used simultaneously.

We are surprised that even among university students, who have studied the Bayes-Theorem before the experiment, the rate of good solutions (without any help of the representation of the data) was very low (ca. 15%). They know the theory in abstract form, but they are unable to apply this knowledge to a real situation.

3. Interpreting the (sometimes very unexpected) results

Interpreting the results, the cost of a diagnosing test plays a crucial role. If the test is more reliable, then it is more expensive. The organisations of public health have to strive for testing all people in order to stop the expansion of a dangerous disease (caused by e.g. a virus).

If we assume that only 0.01% (e.g.) of the population have such a virus, the decision in favour of the cheep test causes to save a huge amount of money for the Ministry of Health, because the expensive test or other medical examination must be done only in a small percentage of the whole population (namely in the cases of the positive test result). We have to pay “the price” for this decision, as there will be too many false “positives” in the population. The false positive result by the diagnosing test causes troubles and a very bad feeling for the affected person. They are healthy, but “victims” of the procedure. If they (and the doctors also) know the correct meaning of the positive test, we could minimise the disadvantages of this procedure. It is important that students recognise the occurrence of such cases in many real world situations.

Discussing the paradox result and the different interpretations, the pupils learned about the interactive relations between intuition and the analysed, mathematically described model. On one hand they learned an important view, namely the classical analyses of such problems that can always be interpreted in the case of mass phenomena. This is the point of view for example of the state or of a bigger population but this view is of no relevance for the individual. Without subjective probabilities this situation can not be handled from the individual point of view. This remark is very important if we try to use the theory of probability in real life situations. On other hand we have seen that Venn-diagrams provide the best understanding of the paradox result. It can help to see why the desired probability is so low.

One student described the reason of the paradox result in the following very impressive manner: “the big part of a very little set can be much less, than a little part of the complement big set.” This relation also shows how the result can be refined.

The paradox feature of the solution decreases if the error rates are smaller or the rate of infected persons is higher. It is not too complicated to analyse the paradox mathematically. If we try to do this, for example, using the steps of the solution by a doubletree diagram, then the following formal result can be obtained. The numerical result has a form of a fraction, where the numerator is the number of the positive *and*(at the same time) infected people; the denominator is the number of all people having positive test result.

How can we write probability (classical or subjective) into this formula? This is an important step into the way of mathematical abstraction. The obtained result is a special case of the so called Bayes-Theorem. The analysis of this theorem in the classical probability theory can be read in [13] pp. 70-73. If we use the symbol *P* for probability independently of the interpretation, then we get the following fraction:

$$P(Inf|+)=\frac{P(+|Inf)\cdot P(Inf)}{P(+|Inf)\cdot P(Inf)+P(+|Non\,Inf)\cdot P(Non\,Inf)}=\frac{(1-\alpha)\cdot p}{(1-\alpha)p+\beta(1-p)}.$$

That is equal to

$$\frac{1}{1+\dfrac{\beta(1-p)}{(1-\alpha)p}}.$$

There are three parameters depending on the data α, β and p.

Decreasing both error rates the chance for the infection in case of the medical diagnosing result being positive, is increasing. Provided, that all three parameters are small (in real situations normally α, β are less than 0,05 or maybe 0,01, and the rate of infection taken in view of a rare disease is also less then 0,01), we approximate the factors (1-p) and (1-β) by 1 in these cases. It can be seen that in this case, the value of the fraction depends on the ratio β/p. If β is much smaller than p then the numerical result is near to 1, there will be no paradox. But if the value of p is much smaller than that of β, then the numerical result will be close to 0, which causes the surprising paradox. It is interesting to see that the role of the two types of test-errors is not symmetrical. If we fix p, the numerical value of the result is much more sensitive to changes in β than in α. This analysis highlights the background of the paradox.

We can also analyse the interval into which the numerical result falls, if we use estimated intervals for the parameter values instead of exact parameter values. We can make only a probabilistic statement for the probability of infection. The statement $P(Inf|+)$ lies in the interval $[p_1; p_2]$ with probability 0,95 or 0,99. It is the next step towards refined mathematical modelling of the situation. We do not know the exact values of α, β and p therefore we can not give an exact numerical answer. To estimate intervals for α and β - using the experimental data from the sample for the test, the Law of Large Numbers has to be used. We denote the estimated (f. e.) 0,95-intervals for the parameters α, β with $[\alpha_1; \alpha_2]$ and $[\beta_1; \beta_2]$. In this case, $P(Inf|+)$ will fall in the interval

$$\left[\frac{1}{1+\frac{\beta_2(1-p)}{(1-\alpha_2)p}}; \frac{1}{1+\frac{\beta_1(1-p)}{(1-\alpha_1)p}}\right]$$

with probability 0,95.

The length of the interval for $P(Inf|+)$ can be shorter or longer than that for the intervals for α and β.

Let us examine two numerical examples:

1. $\alpha_1 = 0{,}005$; $\alpha_2 = 0{,}015$; $\beta_1 = 0{,}045$; $\beta_2 = 0{,}055$ finally $p = 0{,}001$. The length of the intervals for α, β are only 0,01 and for $P(Inf|+)$ we get [0,0176; 0,0217]; the length is less than 0,004.
2. $\alpha_1 = 0{,}025$; $\alpha_2 = 0{,}035$; $\beta_1 = 0{,}045$; $\beta_2 = 0{,}055$ finally $p = 0{,}005$.
   $P(Inf|+) \in [0{,}081;\ 0{,}0982]$. In this case, the interval for $P(Inf|+)$ is longer than intervals for α, β.

It is very useful, if we give additional examples, which are more or less paradox. One of them is the so-called "witch-hunt" effect, which often lies behind the bias against a minority.

We judge this bias (prejudice) to be a very important field of the application of Bayesian-type problems. If the rate of a certain attribute is higher in a small part of a population (a minority) than in the greater part (the majority), we tend to decide the person belongs to the minority once we know about him to have this attribute. Many times there is a positive feed back in this process, and the rate of this attribution increases in the case of a person belonging to the minority by false decision. It leads slowly to the very strong prejudice, that everybody of the minority has this attribute. This effect is known in certain cases as the so-called witch-hunt. It was very surprising for the pupils that mathematics can help to understand such processes in society. See [6] for more details.

4. Textbook

The chapter on conditional probabilities in [6] will be described shortly with the main focus on real problem situations and the ideas in the background. The main goal for this textbook is to combine the different ways of the representation of such Bayesian-type problems using the experimental results for the different kinds of thinking of pupils. There are many tasks to practise converting from one representation to another in the same problem. Below there are two illustrative problems. Both problems will be solved in the book in different ways (Venn-, doubletree-diagram and tabular form) and the pupils are requested to produce the other solutions.

Task 1: There are two taxi companies in a town one of them uses green, the other blue cars. One night a taxi caused an accident and drove away (hit and run accident). The court of justice takes action against unknown committer. Two eyewitnesses claim the taxi was green. The judge prescribed an experiment for the witnesses controlling their ability to identify the colour of a taxi in the darkness. The result is given in following table:

according to witnesses → ↓Colour of taxi	Green	Blue	Total
Green	12	3	15
Blue	17	68	85
Total	29	71	100

According to the information about the taxi companies the ratio of the green and blue taxis at night about: 15:85 and this was reproduced in the experiment (see the table above). How probable is that the taxi caused the accident was really green, if the witnesses saw it green?

It is proposed in the textbook to solve this problem by a Venn-diagram.

Task 2: A Hungarian-minority in a neighbouring country of Hungary, comprises 6% of the whole population. Among this minority, there are people speaking at least three languages (Hungarian, majority and a world language) 36%. Among the majority, this rate is only 4% (they usually do not speak Hungarian). We meet a walking man on the street, who speaks three languages.

Is it likely that this man belongs to the minority? What is the chance of this ?

This problem is solved by a doubletree diagram in the textbook.

After solving some similar problems, the numerical way of thinking will be summarised by a formula. That is the mathematical model of such tasks. The interpretation of the results is part of the answer.

We intend to increase the efficiency using this textbook by this problem solving approach.

We will report on some concrete experimental results in the next paper.

A selection of references:

[1] A. Hawkins - R. Kapadia: Children's conceptions of probability – A psychological and pedagogical review. Educational Studies in Mathematics 15 (1984) 349-377.

[2] M. Borovcnik – R. Peard: Chapter 7: Probability. In A.J. Bishop (ed.), International Handbook of Mathematics Education, 1996 Kluwer Academic Publisher

[3] M. Borovcnik: Wechselspiel zwischen Intuition und Mathematik. 1992 Bibliographisches Institut Mannheim

[4] L. Martignon, C. Wassner, S. Krauss: Frequency format versus probability. Manuscript

[5] Ödön Vancsó: Wie verstehen Leute im Alltag bedingte Wahrscheinlichkeiten. GDM Tagungsband 2003 Franzbecker Hildesheim

[6] Ödön Vancsó (ed.): Matematika 10, 11 Textbook in Hungarian Muszaki Kiado 2003-2004

[7] Ödön Vancsó: Die Stichprobe als ein Beispiel dafür, wie im Unterricht die klassische und bayesianische Auffassung gleichzeitig dargestellt werden kann (manuscript presented for publication in Teaching Mathematics and Computer Sciences (3), University Debrecen (Hungary) 2004)

[8] Ödön Vancsó: Klassische und bayesianische Schätzung der Wahrscheinlichkeit dessen, dass beim Würfelwerfen die Sechs herauskommt. In: Disputationes scientificae of the Catolic University of Ruzomberok (3) 2003, p. 91-99

[9] Ödön Vancsó: Die Schätzung der größten Lottozahl aus dem Maximum der gezogenen Zahlen (publishing in GDM Tagungsband 2004)

[10] B. de Finetti: Wahrscheinlichkeitsrechnung Birkhäuser Verlag 1953

[11] D. Lindley: Probability and statistics from the Bayesian point of view Oxford University Press 1962

[12] M. Pfannkuch, G. A. F. Seber, C.J. Wild: Probability With Less Pain. Teaching Statistics Spring 2002 Vol.24. Number 1, pp. 24-30

[13] D. Wickmann: Zur Begriffsbildung im Stochastikunterricht. Journal für Mathematik-Didaktik 19 (1998) Heft 1, 46-80. B. G. Teubner, Stuttgart

The author's address:
Ödön Vancsó
H-1024 Budapest
Margit krt. 7. II. 5.
e-mail: vancso@ludens.elte.hu