APPLICATION OF BAYESIAN STATISTICS IN OBJECTIFICATION OF EVIDENCE EVALUATION: IS THE ETHICS OF STATISTICAL RESEARCH MEASURABLE?

DEDIĆ VELIMIR, AJDUKOVIĆ GORDANA

The European Center for Peace and Development, Belgrade, Serbia

***Abstract**: This paper deals with application of Bayesian statistics on hypothesis testing. In contrary to classical approach, historical data and subjective judgment plays a great deal in finding validity hypothesis, but Bayesian statistics is superior in terms of judging validity of a hypothesis having known outcome.*

***Keywords:** Beysian statistics, statistical research.*

1. INTRODUCTION

Moral principles constitute autonomy, beneficence, justice and not doing evil. When it comes to statistical research, the range of terms, activities, processes, principles and ways of understanding covered by this term is wide. In this paper, we concentrate on those statistical studies that show the cause of a phenomenon, which make judgments about the manner of occurrence of a consequence, the operating concepts of causality providing judgments about a process. The component of the ethical considerations of such research is very important, because the consequences of the judgments can be serious, sometimes devastating. Therefore, one should remember Feyerabend and his testimony appealing to reject the pressures and prejudices with assessments of any testimony, particularly testimony backed by force of authority. It is especially important, in light of Kant's categorical imperative, in all phases of the statistical survey, to conduct strictly scientific method based on the principles of critical thinking, because in doing so we avoid, on the one hand, profanation of statistics as a science, and on the other hand, we are protecting the majority secular population of the manipulation of the profession. Further, it is a great responsibility to statistician and statistics when it comes to creating public opinion, directing the public attention, selection of important and unimportant, and in forming habits of consumers. Here, special attention must be paid to the morality of customer survey, integrity of statisticians, public familiarity and clarity of the purpose of research that is performing. In this sense, morality and integrity of the participants, their professional education and conciseness statements represent structural elements of good statistics.

There is no statistical method of application of statistics or industry in which there is no manipulation or that is not suitable for handling. In our opinion, there are two main reasons for this: first, any otherwise fair presentation of results outside the proper context is the

manipulation *per se*, and second, incomplete understanding of the specific statistical method leaves open space for subsequent interpretation and distortion.
Observation of evidence submitted to the court or any other entity which assesses validity evidence, may be regarded as entirely subjective process. Rating evidence validity and actions that individuals take due to consequences of a proven fact that products, can have very serious ethical, economic, social and moral consequences. From this, we can see the importance of the proof reading process, presentation of evidence and presentation of all the limitations that some evidence warrants. However, if we introduce the notion of the context in which the evidence is viewed and perform a small thought experiment the following contents, the evidence D is represented in the context of the present set of facts S, and then let the court, the public or only one independent and disinterested individual come out on the validity of the evidence, a then let the same evidence D to be presented in the context of the T set and let all of the above then to vote on the validity of the same evidence D, and now in light of the new set of facts T. Then you can ask a question, How it to quantify the difference in evaluating the same evidence in the light of different contexts?

2. PROBLEM

Suppose that the cheap test is found (let us call it A test) based on a blood sample to determine whether a person is suffering from an illness. Let us suppose that illness is serious, long and expensive treatment, and in particular, the social environment and the diagnosis of the disease carry a risk of stigmatization. Thus, a person suffering from the disease enters a long period of risk, and all sorts of troubles. Let us ask ourselves; what does it mean, and what is the value of the positive test for the disease? Let the court declare on the validity of such evidence in terms of disease probability - p (D) is the probability that the test is valid as proof of disease (D, the evidence), that supports the claims of imposed illness. For example, if D is a proof of claim that "a person suffering from a disease", then p (D) denotes the probability that it is indeed so. Now, suppose that the test is not perfect (intuitively, this means that the test cannot always give the correct diagnosis) so we introduce following codes:
p (D|A) the probability that a person is sick, if you know that A is a positive test
p (A|D) the probability that the A test is positive, if we know that the person is sick
p (AD) the probability that the A test is positive and that the person is sick
p (A) the probability that the A test is positive
p (D) the probability that a person is sick

Bayes' theorem states [1]:

$$p(A|D) = \frac{p(D|\mathrm{A})\mathrm{p}(\mathrm{A})}{p(D)} \quad (1)$$

or

$$p(A|\mathrm{D}) = \frac{\mathrm{p}(\mathrm{AD})}{\mathrm{p}(\mathrm{D})} \quad (2)$$

The first way of writing formulas is very important because it allows the following interpretation: if a priori probability of event is known A, p (A), then the introduction of known sizes: p (D|A) - the probability that a person is sick, if we know that the A test is positive, as p(D) - the probability that a person is sick, can be viewed as a correction factor

of a priori probability p(A) or as a quantitative measure of the value of evidence contained in the findings that quantifies using the probabilities p(D|A) and p(D). In addition, p(D|A) is called the conditional probability of event D, with the well known fact that event A has already occurred.
Furthermore, we can write using the formula of total probability [2]:

$$p(A|D) = \frac{p(D|\mathrm{A})\mathrm{p}(\mathrm{A})}{p(D)} = \frac{p(D|\mathrm{A})\mathrm{p}(\mathrm{A})}{p(D|A)p(A) + p(D|\bar{A})p(\bar{A})} \quad (3)$$

Or the other way:

$$p(A|D) = \frac{1}{1 + \frac{p(D|\bar{A})p(\bar{A})}{p(D|A)p(A)}} \quad (4)$$

This way of writing of the Bayes' formula is known as Bayesian test, where A is a hypothesis, and D is one of the outcomes. This seemingly terminology change is in fact the essential, because it gives us a way to estimate the probability that a cause (hypothesis) is responsible for any outcome. With label $\bar{A}$ is marked a complementary event of staged hypothesis. In our case, the complementary event is that "the test is negative," because these are the only two possible outcomes of testing (if we discard the possibility there are defective test kits, which are destroyed prior to the application due to the imperfections of the manufacturing process). Known methodological device, the process of assessing the validity of the hypothesis, it is now applied in the context of the evaluation of objective evidence, which facilitates the evaluation of the quality of the proceedings.
Because often we know the probability of events with the known fact that the hypothesis is fulfilled(or not met), to this way of writing the Bayes formula is actually a Bayesian method of testing hypotheses. It is important to emphasize that the above estimated probabilities are based on known data and our experience, which make the Bayesian approach subjective, in some way.
Let us return for a moment on the evidence of Bayes' theorem, according to the formula 3 [3]:

$$p(A|\mathrm{D}) = \frac{\mathrm{p}(\mathrm{AD})}{\mathrm{p}(\mathrm{D})}$$

Also, we notice that when it comes to assessing the quality of the test, the probability that A test is positive and that the person is really sick, and the probability of the disease, can be estimated by epidemiological analysis, analysis of records of the test application on a large sample, estimate of the incidence before the invention of the test, and other methods used by public health services.

If we apply the formula 4 for several potential scenarios, i.e. the more potential hypotheses, evaluating the validity of the assumption that the specific hypothesis led to the known outcome, we can calculate using the Bayes' formula. If the outcome D can occur in several ways, or if several hypotheses can justify the occurrence of event D, and if such hypothesis is marked as H_i, then:

$$p(D) = \sum p(D|H_i)p(H_i) \quad (5)$$

Now, according to Bayes' formula [4]:

$$p(H_i|\mathrm{D}) = \frac{p(D|H_i)p(H_i)}{\sum p(D|H_i)p(H_i)} \quad (6)$$

Thus, the value of some evidence, according to other evidences, we can quantify, if we know all the ways in which the outcome D may occur. In our thought experiment, this situation applies to such a test for a hypothetical illness that could have multiple outcomes, and each outcome would indicate the specific cause of disease. Such a scenario cites the importance of the Bayes' method.

3. RESUME

Bayesian statistics, although non-intuitive, introduces the important ideas in the methodological problem of assessing hypotheses. In addition, it is a good tool for demonstrating the quantitative approach to studying the ethical problems of statistical research. Often with a priori knowledge of the probability of occurrence of the events without conditions, as well as the existence of certain conditions, it is very difficult without the knowledge of Bayesian statistics for estimating the probability that some very specific conditions led to the observed events. By changing the terminology, Bayesian statistic is very useful in assessing the validity of any evidence in support of cited claim. We should not forget the weakness of this approach, which is an inherent need for a priori probability of occurrence of the events, as well as the conditional probability of occurrence of events under known conditions, if calculated or estimated based on available historical data, which sometimes leads to criticism of Bayesian access as subjective.
Finally, this analysis shows that science does not relieve from the inclusion of values and attitudes of the term values. Rather, the value is the foundation of science and one of the most important features of human activity.

REFERENCES

[1] M Bishop, „Computer Security: Art and Science“, Addison Wesley, 2003.
[2] D Pleskonjić, N Maček, B Đorđević, M Carić, „Sigurnost računarskih sistema i mreža“; Mikro knjiga, Beograd, 2007.
[3] T Duval et al "XMeta, A Bayesian approach for computer forensics", French ministry of defence proceedings, 2004.
[4] B Rowe, "Multivariate Bayesian Statistics" CRC Press 2003.
[5] X Wang, H Yu,"How to Break MD5 and Other Hash Functions" In: Ronald Cramer (editor), "Advances in Cryptology - EUROCRYPT 2005", volume 3494 of Lecture Notes in Computer Science, pages 19-35,Springer Verlag, Berlin, 2005.