Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal. (2011), Vol. 4, Issue 1, 103 – 105
e-ISSN 2070-5948, DOI 10.1285/i20705948v4n1p103

REVIEW OF THE BOOK TITLED "INFORMATION AND COMPLEXITY IN STATISTICAL MODELING" BY JORMA RISSANEN, PUBLISHED BY SPRINGER, N. Y., 2007

Soubhik Chakraborty*

Department of Applied Mathematics, Birla Institute of Technology, Mesra, Ranchi, India

Received 02 August 2010; Accepted 22 October 2010
Available online 26 April 2011

Abstract*: No statistical model is right or wrong, true or false in a strict sense. We only evaluate and compare their contributions. Based on this theme, Jorma Rissanen has written a short but beautiful book titled "Information and Complexity in Statistical Modeling" (Springer, 2007), where modeling is done primarily by extracting the information from the data that can be learned with suggested classes of probability models. The note reviews this book and on the way rediscovers the chain information-knowledge-wisdom.*

Keywords*: Information, complexity, statistical modeling*

1. Introduction

Statistical modeling has an interesting link with writing a poem. The strength of a poem depends more on the ideas, rather than rhythm, and the optimum use of words to convey these ideas by the poet. Likewise, the strength of a statistical model depends more on the statistician's motivation behind the modeling and a proper selection of design and rightful analysis of the data. Just as rhythm is not mandatory in a poem, so statistical modeling is very much possible even in non-random data, e.g. in computer experiments, where it is used to generate cheap and efficient predictors, and it is this idea that is important. A computer experiment is a series of runs of a code for various inputs. Most computer experiments are deterministic, i.e., re-running the code for identical inputs leads to identical response. But we opt for stochastic modeling because we can then predict the response for untried inputs of huge volume for which it may be computationally cumbersome to run the code. Also, a deterministic model requires knowledge of the entire input. In stochastic modeling, prediction can be made for

E-mail: soubhikc@yahoo.co.in

partial knowledge of the input (e.g. knowing only the input size suffices if the response is complexity of the underlying algorithm).. This explains "cheap" and "efficient prediction" respectively [1]. The errors associated with such prediction can be controlled and to do this is a statistical issue.

Statistical models are not perfect. They are only working models aimed to capture some phenomenon. There are three steps in stochastic modeling: (i) decide which model to fit, (ii) estimate the model parameters and (iii) verify the goodness of fit.

In Jorma Rissanen's short but beautiful book titled "Information and Complexity in Statistical Modeling" (Springer, 2007), we discover another angle of looking at stochastic modeling. It is done primarily by extracting the information from the data that can be learned with suggested classes of probability models. The shortest code length which achieves this is called "stochastic complexity" by the author, inspired by the concept of Kolmogorov's complexity of a sequence, the latter meaning the shortest code length for outputting the given sequence. Recall that a sequence is random if its Kolmogorov's complexity equals its length. The proposed technique ultimately amounts to what the author calls the MDL (Minimum Description Length) Principle where the stochastic

complexity splits into the shortest code length for the optimum model (in the aforesaid suggested classes of models) with the rest comprising noise defined as *the incompressible part in the data without useful information.*

The book has two parts, part one dealing with the coding of random variables and part two dealing with statistical modeling. Some of the topics covered in part one include Shannon-Wiener information, entropy, Markov processes, coding algorithms by Tunstall, Lempel-Ziv and Ma. The major topics covered in the second part include

Kolmogorov's complexity, stochastic complexity, the MDL principle and its applications followed by a useful bibliography.

While reading this book, I re-discovered an amusing philosophical insight worth mentioning. Although information is not knowledge in itself, it can be argued that the mind does not receive knowledge directly. What it receives are pieces of information that are then screened to produce knowledge. Next, knowledge is further screened to produce wisdom. Let me explain this interesting chain with an example.

Consider the following pieces of information:

- Professor P.C. Mahalanobis was the founder director of the Indian Statistical Institute, the first and the only one of its kind in Asia.
- It was he who first brought the computer to India.
- He contributed prominently to independent India's Five Year Plans starting with the second.
- He created the Central Statistical Organization and the National Sample Survey Organization in India.
- He was a fellow of the International Statistical Institute.
- He was also the founding editor of *Sankhya*, the Indian journal of statistics that enjoys an international repute.

- He has made pioneering contributions to statistics especially in sample surveys, biostatistics, and anthropometry; topics such as the famous Mahalanobis distance and fractile graphical analysis are well known terms among the practitioners of the subject.
- Today, rightfully, he is called *the Prophet of Indian Statistics*.

If you now say that "Mahalanobis was a truly great statistician," it is *knowledge* that your mind has screened, summarizing the eight pieces of information given above. Notice that "Mahalanobis was a truly great statistician" is not explicitly written--this is what the mind creates! Next, if you further say that "all statisticians must emulate Mahalanobis," it is *wisdom*, again an invention of the mind, produced by screening the knowledge acquired previously [2]. My point is proved.

2. Conclusion

The moral is that information is still the basic building block of knowledge and knowledge, in turn, the pillar of wisdom. Therefore, is it not rational to expect that statistics, the science of gathering data and extracting relevant information--when seen from the perspective of learning from data and model building to solve complex problems--should also follow this philosophy?

References

[1]. Sacks Welch, J., Mitchell, T.W., Wynn, H. (1989). Design and Analysis of Computer Experiments. *Statistical Science*, vol. 4 (4).

[2]. Bellinger, G., Castro, D., Mills, A.. *Data, Information, Knowledge and Wisdom*. URL: http://www.systems-thinking.org/dikw/dikw.htm