BOOK REVIEW

Naked Statistics

Komsan Suriya

This is the review of the book titled "Naked Statistics: Stripping the Dread from the Data" written by Charles Wheelan. The book was published by W.W. Norton and Company in 2013.

I remember some words in German when I studied there for my Ph.D. They write three words; Forschen, Glauben, Leben. In English, it means do research, believe and live. In developing countries including Thailand, this is not a common practice of the people. Thai people do less research, do not construct the beliefs from scientific-based knowledge, and live without good guidelines that should lead them to a better level of well-beings. However, to believe what is right and not to believe what is wrong, statistics play a crucial role. Charles Wheelan writes this book "Naked Statistics" with interesting points, critics, and suggestions that are very valuable to practitioners in statistics and econometrics.

Professor Hung T. Nguyen showed me this book when I visited his house in Chiang Mai in May 2013. He was so excited that the book pointed many stories related to his favorite statistical technique, Copula. I remembered that an example that he raised was the failure of jet engine. The probability of failure of two jet engines at the same time was underestimated when a quantitative analyst assumed independence between them. The excitement of Professor Hung brought me to find the book. I found it at Kinokuniya book store in Bangkok. I read it through and realized the feeling of Professor Hung while he read the book.

Charles Wheelan begins his book with an attractive introduction. He quotes the words of Hal Varian who is the chief economist at Google that "Statistician will be the sexy

job over the next decade". Besides he compares a statistician to a detective who uses statistics as a powerful tool to answer many important questions. The positive attitude makes readers comfortable with statistics even though the subject may not be their favorite classes in college.

In many chapters, the author gradually provides basic statistical concepts to readers. This is important to those who are not familiar with statistics but want to know what statistics is and for what statistics is. Graduates with statistical background can still enjoy the way he describe these basic concepts by various examples in daily life.

Wheelan writes in chapter 1 that there are limits on the data and the way that the data are collected. He quotes Donald Rumsfeld's words that a general must go to the war with the army that he has. This is what I am impressed by his book. Once, I have to present statistical results in telecommunications research in front of experts in the field. Many results are not so convincing to the audience. Yet some results welcomes critical comments from the floor. However, I insist that my studies are based on the best data and information that I have at that time. This is the same as Wheelan and Rumsfeld say. Statisticians do their best under the available data, methodologies and resources.

In chapter 4, the author encourages the usage of statistics in finding association between sets of data. He asks an interesting question why a computer program knows his taste. Then he tries to explain the mechanism behind it. He also gives an example of algorithm competition that well attracts youngsters around the world to work on complex statistical algorithm.

A stunning point that Wheelan raises in Chapter 6, problems with probability, is the Value at Risk model, or VaR. Researchers in economics uses VaR heavily to calculate risk. Some examples that were published by EEQEL or other books of CMSE Press include Bunnag et al (2010), Chaithep et al (2012), and Mzoughi and Mansouri (2013). He said that VaR can be potentially catastrophic. Wheelan shows that the drawbacks of VaR can appear on at least two points. The first point is at the assumption that the future necessarily follows the past. The second point is that he one percent that is thought to be dangerous may kill us. I remember the lecture of Prof. Paul Embrechts given in the Fifth International Conference of the Thailand Econometric Society (TES2012) in Chiang Mai that we will not be killed by normal situation but an abnormal one, the one percent that is ignored, instead. The lecture of Professor Embrechts matches so perfectly to the comments on VaR of Charles Wheelan.

Raising Mark Twain's words, Wheelan presents three kinds of lies: lies, damned lies, and statistics. On the one hand, this is just a joke that Wheelan plays around the words of Twain. On the other hand, he states seriously that a researcher may be able to interpret statistical results in many ways. He suggests that, even analyzed by complex

techniques, good statistical results should not be against common sense. He also compares doing wrong statistics to committing crime.

Recalling the story told by Professor Hung about jet engines. Wheelan includes this story in chapter 6, problems with probability. He presents some common probability-related errors, misunderstandings, and ethical dilemmas. The story of jet engines is in the first error which is the assumption of independence between variables. There are much more story of the errors that are very interesting and important to statisticians and econometricians such as the assumption of inter-dependence while the variables are in fact independent.

In Chapter 7, the author states the importance of the data by raising a famous proverb in statistics and econometrics "garbage in, garbage out". It is a big mistake that a researcher believes the results that are against common sense. The problem may come from bad data. Some spurious relationship between variables may mislead a researcher to a wrong interpretation. In the chapter, Wheelan presents several kinds of bias that prevents a researcher to look at the goodness of the data carefully.

Other chapters in his book are also interesting. He deals with the central limit theorem, inference, polling, regression analysis, common regression mistakes and program evaluation. He is good in telling stories that bring readers deeper into the points. One of the best thing in his book is at his writing style, writing a short summary at the end of each chapter to allow readers with and without statistical background to capture the main idea of the chapter.

All in all, Professor Hung gives a summary to this Wheelan's book. He says that it is fine that we use statistics as soon as we do not kill people. He returns to his story of jet engines again. He insists that the miscalculated probability with careless assumption may kill people on a jet plan over Atlantic. After reading Wheelan's Naked Statistics, I believe so. I also remind myself that using statistics is not only beneficial but also dangerous since the data can be interpreted in many ways and can lead to different sets of truth. Once a researcher uses statistics or econometrics, he or she should defend the results with common sense before believing in the results. This is a strong and strict suggestion from Wheelan too.

REFERENCES

Bunnag, Ratanan et al. 2010. Value at risk of international tourist arrivals to Thailand. Chapter 26 in Pisit Leeahtam et al (eds.). **Econometric Leadership of the Knowledge-based Service Economy**. Chiang Mai: CMSE Press.

Chaithep, Kittiya. 2012. Value at risk analysis of gold price returns using extreme value theory.. **The Empirical Econometrics and Quantitative Economics Letters 1**, 4 (December): pp. 151 – 168.

Mzoughi, Hela and Faysal Mansouri. 2013. Computing risk measures for non-normal asset returns using Copula theory. **The Empirical Econometrics and Quantitative Economics Letters 2**, 1 (March): pp. 59 - 70.

Wheelan, Charles. 2013. **Naked Statistics: Stripping the Dread from the Data**. New York: W.W. Norton and Company.