DANIEL PEÑA

by Michael Wiper
mwiper@est-econ.uc3m.es

Professor Peña is one of the best known statisticians in Spain, having written over 100 published papers and a number of textbooks on both Bayesian and classical statistics. Since 1991, Professor Peña has been working in the Statistics and Econometrics Department at the Universidad Carlos III de Madrid. You can find out more information from the Departmental homepage at:

http://halweb.uc3m.es/

We e–mailed Professor Peña a number of questions about his career and the Bayesian world in general. Here are his responses.

1. When and why did you first get interested in statistics, and in particular, Bayesian statistics.

The first time I heard about Bayesian Statistics was in 1972, when I was starting my Ph.D. at the University of Madrid (UPM). I had a degree in industrial engineering (Ingeniero Industrial) and I was interested in a PhD in Operational Research. I had already had a mathematical statistic course and after it I concluded that statistics was a pretty boring subject, hard to understand and not very useful for real problems. During the first year of the PhD I took a course on Decision Theory and the text-book recommended was Decision Analysis by Howard Raiffa. I started reading this book and I was completely fascinated by it. I was very much impressed by the beauty of the concepts and the strong logic of the ideas in the book. After this course, I read Schlaiffer's book and Pratt, Raiffa and Schlaiffer, and I was getting more and more interested in Bayesian Statistics. At this time I did not know anybody in Spain who was interested in this topic, and finally I wrote a thesis, pretty much by myself, on Bayesian Decision Analysis applied to medical diagnosis and treatment. At that time I had a rather left wing point of view, and I did not want to work on anything related to business or economics, so that I chose a medical application.

2. Tell us about some of the people who have influenced your career.

After my PhD I started teaching decision analysis in an Operational Research department. I wanted to integrate dynamics and time series data in the decision process and in 1978 I met Arthur Treadway, an economist from Chicago, who came as a visitor to the University of Madrid. He told me about a new methodology for time series that has just appeared, and we started meeting once a week to study together the Box and Jenkins book. Again, I was fascinated by this book and I was very much attracted by the iterative statistical learning process advocated. In particular, I was very much attracted by the steps of identification and diagnosis of the model, that are very important in time series but also elsewhere. At that time I did not know how to integrate these ideas into the Bayesian framework, but I was convinced that they were very useful and important. I invited George Box to come to Spain to teach a short course on time series and he came with George Tiao. I learned a lot from them and I was very attracted for their approach to time series. After their visit I decided to spend a year in Wisconsin. George Box was always very nice and he helped me to get support to spend the 83-84 academic year in Madison. It was a great year from all points of view. I addition to working with George Box on factor analysis on time series, I visited George Tiao that has just moved to Chicago and we started working together. I also met Irwin Guttman and Dennis Cook, who were visiting Wisconsin this year, and learned a lot about Bayesian Statistics working with Irwin and about influence analysis working with Dennis. Later on I was very much impressed and influenced by the work of Tukey and Efron, among others, but, altogether, I think that the three persons who have had most influence in the way I look at statistics are George Box, George Tiao and Irwin Guttman.

3. You have done a lot of work in time series but mainly using classical statistics. (I don't know if your new book on time series (*A Course in Time Series Analysis*, Daniel Peña, George Tiao and Ruey Tsay eds., Wiley) contains anything on Bayesian methods). Do you think that classical techniques are better suited to this field than Bayesian methods? If so, why?

Yes my book with George Tiao and Ruey Tsay on time series has a chapter on Bayesian Time series written by Ruey, but most of it is from the classical point of view. In many time series applications the sample

information is much larger than the a priori information about the parameters, and thus maximum likelihood estimation is roughly similar to Bayesian estimation. I have not had real experience with short time series in which the prior information can really make a difference. Also, a key part of the time series model building process is the identification and diagnostics of the model, and, for these steps, I believe that the so called classical statistics are better suited than traditional Bayesian Statistics. This point has been stressed by Box (1980) and I fully agree with him that we need Bayes theorem for estimation but to build statistical models we also need many other tools that has been developed in the so called classical statistics. Bayesian time series has sometimes been identified with the structural approach using the Kalman filter, whereas ARIMA models have been considered mostly from the classical statistics viewpoint. I think that both approaches are complementary, and both can be estimated by Bayesian or maximum likelihood techniques. I usually prefer the reduced form, (ARMA type models) because we have better tools available for identification and diagnosis but there are many cases, for instance dynamic factor models, for which I believe that Bayesian recursive estimation using the Kalman filter is more useful. So, I do not think it is so relevant if we use ML (maximum likelihood) or BT (Bayes Theorem) for estimation, because I do not see classical and Bayesian statistics as rival approaches but rather as complementary, and we will be able to solve real problems better if we can use both. The idea of a unique best method to obtain the truth has disappeared in many scientific areas and it is surprising that this dogmatic point of view has such strong roots in our scientific community.

4. Conversely, in your work on outliers, influence and robustness, you have used both Bayesian and classical techniques.

Yes, and I have found both very useful for different things, but again I think it is not true that we can do everything better from the Bayesian point of view. For instance, many Bayesians do not understand the concept of masking very well. This is an idea that has been developed mostly in the classical robustness literature and there are many so called robust Bayesian procedures published in the last 10 years, that fail completely as soon as we have a small group of high leverage outlier observations. On the other hand I think that we have some classical procedures to deal with outliers in multivariate problems and in regression that are far ahead of the Bayesian alternatives. I find it surprising that some people seem to be more concerned about whether a procedure is truly Bayesian or not than whether or not the procedure is useful to solve the problem it tries to solve. I think that research in Bayesian statistics should concentrate more on solving problems that classical statistics is not well suited for, such as working with small samples, using subjective information in a better way or combining in a robust way different sources of information.

5. You have written fairly extensively on education, and quality improvement in the university sector. What comments do you have on how to improve the teaching of (Bayesian) statistics ?

I hope that Bayesian Statistics will be more used in all scientific areas in the future. I think that we should concentrate our teaching in presenting simple and flexible procedures that people can use in practice to solve the problems they will face in their professions. Sometimes a classical tool could be a convenient and fast approximation and then we should recommend using it. We should teach ALL statistics, that implies how to use subjective information, how to combine information from different sources, and how to incorporate all sources of uncertainty in the problem and all these problems can be better solved using Bayesian Statistics. However, we should also stress exploratory data analysis and model diagnosis, problems in which probability plays a small role and in which the most useful tools have been developed within classical statistics.

6. Also, you have written statistical texts and research papers designed for social scientists, engineers, medics, and have collaborated on research projects with people from many fields. What are the major differences you have found in such diverse areas.

As Tukey has said, the great thing about statistics is that you can play in someone else backyard. I have had a lot of fun

working with people from many different fields in understanding their data. I believe that real problems are the most stimulating source for new statistical developments. I think that we will be better scientists and more useful for society if we concentrate our efforts in solving the difficult problems that are all around in the real world instead of concentrating our efforts in generalizing methods and ideas that are of very limited usefulness in practise. Many of my research interests have come from practical applications. For instance, my interest in outliers came from noticing the important effects they may have, even in large data sets, in our conclusions on public welfare policies in a joint work with the economist Javier Ruiz Castillo. My interest in dynamic factor models came from my collaboration with a historian, Nicolas Sánchez Albornoz, in understanding the effect of political events in wheat prices in Spain in the XIX century. My interest in cluster analysis and data heterogeneity came from building quality indexes for the railroad system in Spain.

7. Continuing along the same lines as questions 3 and 4, what do you think Bayesians can learn from classical statisticians and vice versa.

I think Bayesians can learn methods for exploratory analysis and model diagnostics from Classical statistics. Classical statistics can learn flexible methods for estimation and testing using several sources of information from Bayesian statistics.

8. Also, looking into the future, Bayesian statistics seems to be gaining more converts every year. Do you think that classical statistics will eventually be replaced by Bayesian as the dominant method or do you think other techniques (neural nets, data mining etc.) will start to replace standard statistical analysis?

I think that we will move towards a more balanced teaching of statistics. It is clear to me that today a competent statistician cannot ignore Bayesian statistics, as was unfortunately very common in the past. In the same way, I think that the Bayesian statistical community is more mature now and is more interested in solving new problems and finding new procedures than in competing against classical statisticians. But our changing world will force us to develop new tools and new paradigms. For instance, in the last century the standard paradigm for both classical and Bayesian statistics, is that our raw material was a sample from some statistical model. We can simplify this situation saying that our basic assumption was some kind of data homogeneity, may be with some small proportion of outliers. This paradigm is not appropriate today for the analyses of the available large data sets that include hundreds of variables and many thousands of observations. In this situations we do not have a central model, rather we expect that different models will explain the data in different regions of the sample space. This multi-model situation can be called the data heterogeneity situation. The tools to be used in these cases are more complex than the ones considered by cluster analysis or robust methods, and we need new statistical tools to extract the information in these data set. I believe than in order to solve these complicated problems that we have ahead, both Bayesian and classical statistics will be useful. We also need to develop more automatic procedures for data analysis and for this purpose Neural networks and Data mining take advantage of the available computer power. Neural networks are fast, and sometimes not very efficient, ways to build regression or time series models in which the response is a non linear function of linear combinations of the explanatory variables. Thus they are fast procedures for non linear factor models. Data mining includes fast multivariate exploratory methods that can be very appropriate in many situations. These two procedures are useful for gathering information from a given set of data, but if we want to generate knowledge, that is to understand not only the sample data but also similar samples not yet observed, and to be able to generate useful forecasts we need statistical models.

9. Looking back, what are the things you are proudest of in your statistical career?

I am very proud of having had the opportunity of working with such great statisticians as George Box, George Tiao, Irwin Guttman, Dennis Cook and Victor Yohai. I am also very proud of helping to develop the Department of Statistics and Econometrics at the Universidad Carlos III of Madrid, I have excellent

colleagues there and the atmosphere is very stimulating. Also I am very proud of contributing a bit to developing the research potential of my Ph.D. students. I have been very lucky to have excellent Ph.D. students and, to be honest, I feel that I have learned from them more than they have learned from me.

And looking ahead, what about your future plans in statistics?

I have been working for many years now with George Tiao in the problem of data heterogeneity and this is one of my first priorities. Also I am interested in many other problems: diagnostic tests for time series (with J. Rodriguez), new methods for bootstrap in time series (with A. Alonso and J. Romo), outliers in Garch processes (with A. Carnero and E. Ruiz), Bayesian Model Averaging (with I. Guttman and D. Redondas), Robust Bayesian estimation (with R. Zamar), Projection Pursuit methods for multivariate time series (with P. Galeano and R. Tsay), Dynamic factor models (with P. Poncela), random coefficients models for quality (with V. Yohai), Forecasting Multivariate time series (with I. Sánchez) and image analysis (with M. Benito). These problems are going to keep me busy for a while!

10. And finally, what are you looking forward to seeing next year in Valencia 7, or should it be Tenerife 1?

New practical, flexible and iterative methods for getting knowledge from the large and heterogeneous data set that as statisticians we are going more and more to face. These methods should incorporate several dimensions : multivariate, dynamic, robust (in a very broad way) and computationally efficient. I am really looking forward to new advances in this field.

Thanks to Daniel for an interesting interview.

THANKS

TO ALL THOSE WHO SERVED IN THE EDITORIAL BOARD IN 1999 - 2001

Associate Editors: Jim Albert, Sudipto Banerjee, Maria Eugenia Castellanos, Kate Cowles, Maria De Iorio, Duncan Fong, Sujit Ghosh, Gabriel Huerta, Leonhard Knorr-Held, Antonio Lijoi, Javier Morales, Antonio Pievatolo, Raquel Prado, David Rios Insua, Siva Sivaganesan, Dalene Stangl, Isabella Verdinelli, Michael Wiper

Corresponding Editors: Christophe Abraham, Antonia Amaral Turkman, Carmen Armero, Marilena Barbieri, Bill Bolstad, Agata Boratynska, Cinzia Carota, Christopher Carter, Andres Christen, Roger M. Cooke, Petros Dellaportas, Gül Ergun, Josep Ginebra, Paul Gustafson, Eduardo Gutierrez Peña, Katja Ickstadt, Leonhard Knorr-Held, Robert Kohn, Jack C. Lee, Shu-Ing Liu, Hedibert Freitas Lopes, Udi Makov, Marek Meczarski, Renate Meyer, Ioannis Ntzoufras, Yasuhiro Omori, Suleyman Ozekici, Susan Paddock, Carlos Daniel Paulino, Jose Miguel Perez, Raquel Prado, Josemar Rodrigues, Abel Rodriguez, Alfredo Russo, Aldo Sacerdoti, Mike Smith, Daniel Thorburn, Brani Vidakovic, Mattias Villani, Hajime Wago, Cathal D. Walsh, Simon Wilson, Lara Wolfson, Karen Young