Statistica Sinica **16**(2006), 677-682

Editorial

A Dynamic Statistician with Timeless Contributions

This issue features four articles in memory of Professor Ching-Zong Wei who passed away on November 18, 2004, after a long struggle with brain tumors. His close friends Malay Ghosh and Bhramar Mukherjee immediately dedicated their work on sequential design for case control studies to him. The other three articles highlight Wei's contribution to the field.

The article by Ngai Hang Chan contains a comprehensive survey of Wei's contribution to probability and statistics as well as an insightful appreciation of his research style. The second article is based on the joint work of Ching-Kang Ing and Wei himself. It focuses on probability tools developed using convergence systems, and applies these tools to statistical problems such as consistency and model selection. In the third article, Tze Leung Lai and Zhiliang Ying describe Wei's work in recursive estimation and adaptive control. They also extend the ideas and techniques developed by Wei and his coauthors to obtain new results on this topic in semiparametric models.

Together, these three articles map out the various statistical territories that Wei's potent mind explored and left impressive footprints on. While these articles feature his major contributions, I would like to address a few of Wei's lesser known results in order to illustrate his role in creating several statistical methodologies.

One of the driving forces underlying Wei's research was his keen interest and exceptional insight into the dynamics of time in data science. In the early 1990s, Wei studied the following recursion

$$X_{n+1} = X_n + a_n^2 b(X_n) + a_n \sigma(X_n)\varepsilon_{n+1} \ , \tag{1}$$

where $\{a_n\}$ is a sequence of positive constants such that $a_n \to 0$ and $\sum_{n=1}^{\infty} a_n^2 = \infty$. In

Basak, Hu and Wei (1997), Wei and his coauthors showed that under mild smooth and growth conditions on $b(\cdot)$ and $\sigma(\cdot)$, and some moment conditions on the martingale difference sequence $\{\varepsilon_n\}$, the recursion converges to the invariant distribution of the diffusion process

$$dZ(t) = b(Z(t))dt + \sigma(Z(t))dW(t) \ ,$$

where $W(\cdot)$ denotes the Brownian motion process.

Addressing a long-standing issue, Kersting (1978) made a clever observation that the usual normalization procedure in the central limit theorem can be written in the form of (1) with $a_n = (n+1)^{-1/2}$, $b(x) = -x/2$ and $\sigma(x) = 1$. Wei's interpretation of the central limit theorem according to the dynamics of (1) suggests that, at each stage of the recursion, $b(x) = -x/2$ acts as the contraction component pulling probability mass toward the center, whereas the third term on the right hand side of (1) serves as the expansion component pushing probability mass away from the center. As time moves on, these two components will eventually balance out if a_n is of order $n^{-1/2}$ and results in a normal limiting distribution. This interpretation further suggests considering $b(\cdot)$ and $\sigma(\cdot)$ of other forms to obtain a variety of limiting distributions.

In application, the results in Basak, Hu and Wei (1997) can be used to construct simulation algorithms for the invariant distribution of a diffusion process, which, in some cases, is difficult to sample from directly. But it was Wei's keen interest and exceptional insight into the time dynamics in data science that inspired in-depth research into (1).

One of Wei's research projects in the late 1980s was motivated by Bergman and Turnbull (1983). This article held that, in experiments with rodent bioassay and

"What he left us was not only these open problems in probability and statistics, but also the remarkable courage he always showed while fighting brain tumors up until the last minute. He remains a leading expert at the core of statistical estimation and inference, and leaves us all a timeless legacy."

engineering destructive life testing, the information content is closely related to the timing of experiments. The purpose of these experiments is to estimate the onset time of an event (i.e., the presence of tumors in laboratory animals and significant quality deterioration of products in engineering experiments). When an experiment is performed too early, the event will be absent from most subjects. A belated experiment brings about the occurrence of the event on most subjects. In either case, there is little information to estimate the parameter of interest.

If we assume that the onset time follows an exponential distribution, then the likelihood when either $\{x=1\}$ (the presence of the event) or $\{x=0\}$ (the absence of the event) happens equals

$$f(x \mid t,\theta) = (1-e^{-\theta t})^{x}(e^{-\theta t})^{1-x} \ .$$

The Fisher information is given by

$$h(t,\theta) = E_{\theta}\left[-\frac{\partial^2}{\partial\theta^2}\ln f(X \mid t,\theta)\right] = \frac{t^2}{e^{\theta t}-1} \ . \tag{2}$$

As a function of t, $h(\cdot,\theta)$ is unimodal and reaches its maximum at $t^* \approx 1.5936/\theta$.

It seems that one can view (2) as a regression function and design a sequential

procedure to search for its maximum following Kiefer and Wolfowitz (1952). However, Wei realized that it cannot be formulated as such due to the arrow of time. When one conducts an experiment at time t , it is not possible to carry out another experiment at $t' < t$ because time only moves forward. He envisioned formulating this problem as a multi-armed bandit problem similar to Lai and Robbins (1985) with a twist: there exists a linear order among arms, and one can only do arm-pulling that obeys this linear order. The problem then reduces to that of finding the optimal stopping time to leave the current arm when one of the future arms yields a greater reward (in this case more Fisher information) and to stay with the current arm when it gives the greatest reward. With the aid of optimal sequential testing theory, the idea was implemented in Hu and Wei (1989), which once again demonstrated Wei's profound interest in, and penetrating analysis of, the role of time in data science.

The techniques and ideas developed in Hu and Wei (1989) continue to produce new results; see for example, Fuh and Hu (2000), Hu and Li (2003), and Chan, Fuh and Hu (2006). In particular, the paper by Chan, Fuh and Hu (2006) adopts a framework which allows both the regular and the irreversible multi-armed bandit problems to be addressed simultaneously.

Wei continued researching and thinking about challenging problems even after being diagnosed with brain tumors. His handwritten manuscripts from the last few years contain unfinished work on several topics, one of them being Bayesian control. There he combined the Bayesian and certainty equivalence approach as mentioned in the article by Lai and Ying. Instead of using the least-square estimate in the certainty equivalence approach, Wei employed Bayes estimates (posterior means). He referred to his main idea on the subject as “back to the future”. He was exploring the conditional expectation of the parameters in a stochastic regression model on the ultimate σ - field, which contains observable information from the infinite future. He obtained some interesting results based on this idea. Wei's students and colleagues plan to carry on the unfinished work and make known yet another of his contributions to probability and statistics.

In summary, the three aforementioned articles are not only about the past achievements of Wei but also about possible future developments. As pointed out in the article by Chan, and demonstrated by the new results in the second and third articles, when reviewing Wei's publications, one often finds possible areas of extension. We believe Wei will continue to exert his influence in probability and statistics for years to come.

Looking back over the years, it has been a painful yet memorable experience to document Wei's contributions and to remember our friendship. He had several publications commenting on "lottery wizards", "Bible code puzzles", "statistical education", and "mathematical thinking" that were prepared for a general audience and had a wide impact on younger generations in his homeland. He never stopped teaching, writing and speaking, even after the first major brain operation in 1998. We feel sorry to have lost a wonderful friend and productive colleague. What he left us was not only these open problems in probability and statistics, but also the remarkable courage he always showed while fighting brain tumors up until the last minute. Ching-Zong Wei, a leading expert in the field and my dearest friend, continues to live in my heart and leaves us all a timeless legacy.

References

Basak, G. K., Hu, I. and Wei, C. Z. (1997). Weak convergence of recursion. *Stochastic Process Appl*. **68**, 65-82.

Bergman, S. and Turnbull, B. (1983). Efficient designs for destructive life testing with application to animal serial sacrifice experiments. *Biometrika* **70**, 305-314.

Chan, H. P., Fuh, C. D. and Hu, I. (2006). Optimal strategies for a class of sequential control problems with precedence relations. Technical Report, Department of Statistics and Applied Probability, National University of Singapore.

Fuh, C. D. and Hu, I. (2000). Asymptotically efficient strategies for a stochastic scheduling problem with order constraints. *Ann. of Statist*. **28**, 1670-1695.

Hu, I. and Wei, C. Z. (1989). Irreversible adaptive allocation rules. *Ann. of Statist*. **17**, 801-823.

Hu, I. and Lee, C. W. (2003). Bayesian adaptive stochastic processes termination.

Math. Oper. Res. **28**, 361 - 381.

Kersting, G. D. (1978). A weak convergence theorem with application to the Robbins-Monro process. *Ann. Probab*. **6**, 1015-1025.

Kiefer, J. and Wolfowitz, J. (1952). Stochastic estimation of the maximum of a regression function. *Ann. of Math. Statist*. **23**, 462-466.

Lai, T. L. and Robbins, H. (1985). Asymptotically efficient adaptive allocation rules. *Adv. Appl. Math*. **6**, 4-22.

- Inchi Hu

Inchi Hu is Professor of Information and Systems Management at Hong Kong University of Science and Technology. He received his PhD from Stanford University in 1985. From 1985 to 1986, he was Visiting Assistant Professor at Department of Mathematics, University of Maryland, College Park, where he first met Ching-Zong Wei.