Statistica Sinica **23** (2013), 1437-1440
doi:http://dx.doi.org/10.5705/ss.2013.editorial

# Foreword

This is a special issue in honor of Professor David Siegmund's 70th birthday. It consists of 17 papers in the areas of *Statistical genetics*, *Change-point problems*, *Random fields*, and *Sequential analysis*. They represent the major areas of the research of Professor Siegmund, who has made seminal contributions, with far-ranging impacts, to these areas, which will be reviewed below. Some of the papers are expanded versions of talks presented in June 2011 at the Third International Workshop on Sequential Methodologies (IWSM3) and the Conference on Genetics, Probability, and Statistics (GPS), held in honor of his 70th birthday (`http://statistics.stanford.edu/gps`).

## Statistical genetics

The monograph by Siegmund and Yakir (2007) gives an elegant and lucid overview of the statistical methods and problems in identifying, or mapping, the genes associated with certain phenotypes, or traits. The book covers many important topics in experimental and human genetics, together with related models and statistical methods, including significance level and power of genome scans, confidence bounds for the genetic effects of quantitative trait loci (QTL), model selection in mapping QTL, and multipoint linkage analysis, to which Professor Siegmund and his coauthors have made fundamental contributions; see for example Feingold, Brown, and Siegmund (1993), Dupuis, Brown, and Siegmund (1995), Dupuis and Siegmund (1999), Teng and Siegmund (1998), Tang and Siegmund (2002), Siegmund and Yakir (2003), Siegmund (2004), Peng and Siegmund (2004), and Peng, Tang, and Siegmund (2005). Recent major advances by Professor Siegmund's group include mapping quantitative traits for unselected families in Dupuis et al. (2009), and statistical methods for detecting copy number variations; see Siegmund, Yakir, and Zhang (2011a), and Zhang et al. (2010).

The first four papers in this special issue are in the area of statistical genetics. The fifth, by Chen, Fang, and Shao, is on Poisson approximations for dependent trials, which have applications in genomic sequence analysis, as first noted by Arratia, Gordon, and Waterman (1990).

## Change-point problems

Siegmund (2013) has pointed out the close connection between the statistical methodologies for genome scans and change-point problems. In particular, he notes that whereas change-point methodology had its origins in sequential detection of a change-point, as in quality control charts, "applications in biology and, more specifically, genetics" have motivated change-point problems in fixed samples, which "are now completing the circle back to sequential detection, now involving complex signals arising from multiple sources and/or detection by multiple sensors." Professor Siegmund has made fundamental contributions to both sequential and fixed-sample change-point problems; see for example Siegmund (1988), Siegmund and Venkatraman (1995), Siegmund and Worsley (1995), Rabinowitz and Siegmund (1997), Tu and Siegmund (1999), Siegmund and Yakir (2000), Siegmund, Zhang, and Yakir (2011b), and Xie and Siegmund (2013).

In this special issue, Tu's paper introduces a ratchet scan statistic whose scanning window is a grid instead of a rectangular box, to detect cosmic ray pion particles in high energy physics. The paper by Hao, Niu and Zhang develops a novel false discovery rate approach to multiple change-point detection and applies it to the analysis of copy number variations. An alternative approach to multiple change-points using empirical Bayes (EB) filters and smoothers in hidden Markov models (HMM) is used in the paper by Lai and Xing to develop stochastic change-point models with contemporaneous jumps in regression parameters and long-run volatilities for econometric time series. The paper by Dicker, Sun, Zhang, Keenan, and Shepp synthesizes ideas from HMM, filtering, sequential change-point detection algorithms for real-time estimation of blood glucose density from sensor measurements in continuous glucose monitoring systems for managing diabetes.

## Random fields

Scan statistics are closely related to the maxima of random fields. While they represent one direction of Siegmund's work in random fields, another direction was undertaken by him earlier in connection with inference in nonlinear regression models using the Hotelling–Weyl formula for the volumes of tubes around smooth curves and more general manifolds; see Johnstone and Siegmund (1989), Knowles and Siegmund (1989), and Siegmund and Zhang (1993). The papers by Taylor and Worsley, Yakir, and Wang and Woodroofe in this special issue are in the area of random fields.

## Sequential analysis

The last five papers in this special issue belong to the area of sequential analysis and related boundary crossing problems, which was the focus of Professor Siegmund's research from his graduate student days at Columbia University to the time when he finished his influential monograph on the subject in 1985. Another book he published in this area, with Chow and Robbins in 1971, is related to his Ph.D. thesis on optimal stopping. Some of Professor Siegmund's seminal contributions in this area can be found from these two books and in Sections 2.2 and 2.3 of Lai (2009).

On behalf of the Organizing Committees of GPS and IWSM3, I want to thank *Statistica Sinica* for publishing this special issue and to express my appreciation to the Co-Editor, Professor Qiwei Yao, for his excellent and timely editorial work. We would like to use this opportunity to congratulate Professor Siegmund on his achievements and wish him many more happy and productive years.

Tze Leung Lai
Stanford University

## References

Arratia, R., Gordon, L. and Waterman, M. S. (1990). The Erdös-Rényi law in distribution, for coin tossing and sequence matching. *Ann. Statist.* **18**, 539-570.

Chow, Y. S., Robbins, H. and Siegmund, D. (1971). *Great Expectations: The Theory of Optimal Stopping.* Houghton Mifflin Co, Boston, Mass..

Dupuis, J., Brown, P. and Siegmund, D. (1995). Statistical methods for linkage analysis of complex traits from high-resolution maps of identity by descent. *Genetics* **140**, 843-856.

Dupuis, J., Shi, J., Manning, A. K., Benjamin, E. J., Meigs, J. B., Cupples, L. A. and Siegmund, D. (2009). Mapping quantitative traits in unselected families: Algorithms and examples. *Genet. Epidemiol.* **33**, 617-627.

Dupuis, J. and Siegmund, D. (1999). Statistical methods for mapping quantitative trait loci from a dense set of markers. *Genetics* **151**, 373-386.

Feingold, E., Brown, P. and Siegmund, D. (1993). Gaussian models for genetic linkage analysis using complete high-resolution maps of identity by descent. *Amer. J. Hum. Genet.* **53**, 234-251.

Johnstone, I. and Siegmund, D. (1989). On Hotelling's formula for the volume of tubes and Naiman's inequality. *Ann. Statist.* **17**, 184-194.

Knowles, M. and Siegmund, D. (1989). On Hotelling's approach to testing for a nonlinear paramter in regression. *Internat. Statist. Rev.* **57**, 205-220.

Lai, T. L. (2009). Martingales in sequential analysis and time series, 1945-1985. *J. Électron. Hist. Probab. Stat.* 5: 31.

Peng, J. and Siegmund, D. (2004). Mapping quantitative traits with random and with ascertained sibships. *Proc. Nat. Acad. Sci. USA* **101**, 7845-7850.

Peng, J., Tang, H. and Siegmund, D. (2005). Genome scans with gene-covariate interaction. *Genet. Epidemiol.* **29**, 173-184.

Rabinowitz, D. and Siegmund, D. (1997). The approximate distribution of the maximum of a smoothed Poisson random field. *Statist. Sinica* **7**, 167-180.

Siegmund, D. (1985). *Sequential Analysis: Tests and Confidence Intervals.* Springer-Verlag, New York.

Siegmund, D. (1988). Approximate tail probabilities for the maxima of some random fields. *Ann. Probab.* **16**, 487-501.

Siegmund, D. (2004). Model selection in irregular problems: Applications to mapping quantitative trait loci. *Biometrika* **91**, 785-800.

Siegmund, D. (2013). Change-points: From sequential detection to biology and back. *Sequential Anal.* **32**, 2-14.

Siegmund, D. and Venkatraman, E. S. (1995). Using the generalized likelihood ratio statistic for sequential detection of a change-point. *Ann. Statist.* **23**, 255-271.

Siegmund, D. and Worsley, K. J. (1995). Testing for a signal with unknown location and scale in a stationary Gaussian random field. *Ann. Statist.* **23**, 608-639.

Siegmund, D. and Yakir, B. (2000). Tail probabilities for the null distribution of scanning statistics. *Bernoulli* **6**, 191-213.

Siegmund, D. and Yakir, B. (2003). Statistical analysis of direct identity-by-descent mapping. *Ann. Hum. Genet.* **67**, 464-470.

Siegmund, D. and Yakir, B. (2007). *The Statistics of Gene Mapping.* Springer, New York.

Siegmund, D., Yakir, B. and Zhang, N. R. (2011a). Detecting simultaneous variant intervals in aligned sequences. *Ann. Appl. Stat.* **5**, 645-668.

Siegmund, D. and Zhang, H. (1993). The expected number of local maxima of a random field and the volume of tubes. *Ann. Statist.* **21**, 1948-1966.

Siegmund, D., Zhang, N. R. and Yakir, B. (2011b). False discovery rate for scanning statistics. *Biometrika* **98**, 979-985.

Tang, H. and Siegmund, D. (2002). Mapping multiple genes for quantitative or complex traits. *Genet. Epidemiol.* **22**, 313-327.

Teng, J. and Siegmund, D. (1998). Multipoint linkage analysis using affected relative pairs and partially informative markers. *Biometrics* **54**, 1247-1279.

Tu, I.-P. and Siegmund, D. (1999). The maximum of a function of a Markov chain and application to linkage analysis. *Adv. Appl. Probab.* **31**, 510-531.

Xie, Y. and Siegmund, D. (2013). Sequential multi-sensor change-point detection. *Ann. Statist.* **41**, 670-692.

Zhang, N. R., Siegmund, D., Ji, H. and Li, J. Z. (2010). Detecting simultaneous changepoints in multiple sequences. *Biometrika* **97**, 631-645.