

IN MEMORIAM JOSEPH B. KRUSKAL 1928–2010

J. DOUGLAS CARROLL AND PHIPPS ARABIE

RUTGERS UNIVERSITY

This text was edited and completed by Willem J. Heiser, after the sudden and untimely death of the authors.

Joseph Bernard Kruskal, Jr. a mathematician who made important contributions to statistical methodology, computer science, and psychometrics, died September 19, 2010 at his home in Maplewood, NJ, of pancreatic cancer. He was 82.

In our community, we know Joe Kruskal best for his work in putting Shepard's original nonmetric multidimensional scaling (MDS) on a much more rigorous and logical basis, locally minimizing an explicit badness-of-fit function called STRESS. In the process, he independently reinvented monotonic regression (after Ayer, Brunk, Ewing, Reis, & Silverman, 1955, to whom he gave full credit for having devised this more general procedure, stating, with his characteristic modesty, that he merely made various steps in the more general algorithm conforming to his MDS innovations). The MDS technique elucidates the main relationships among complex data by representing a matrix of similarities or dissimilarities using interpoint distances which are linear (in the metric case) or monotonic (in the nonmetric case) with the input data of similarities or dissimilarities. Kruskal particularly applied these methods to many disciplines, including the field variously known as lexicostatistics or glottochronology, where he collaborated with Isidore Dyen and Paul Black in mapping the family tree of Indo-European languages. In addition, he wrote several didactic articles trying to remediate many linguists' misunderstandings of statistics as applied to that field.

Joe Kruskal also dealt with a rich class of alternative metrics (or MDS distance functions), which he called the Minkowski L_p metrics; including as a special case the "city-block" or taxicab metric. At Shepard's suggestion, Kruskal then introduced his MFIT monotone regression algorithm into the first procedure fitting a two-way ANOVA model with no interaction terms but with optimal monotonic transformation of the dependent variable (Kruskal, 1965). This prescience provided a practical method of fitting the two-attribute version of Luce and Tukey's conjoint measurement axiomatic model, published as the first article in the 1964 inaugural issue of the *Journal of Mathematical Psychology* in 1964. Professor Paul Green of the Wharton School's Marketing Department also thought this approach would be excellent for fitting what he called "conjoint analysis", which in many ways became a workhorse in marketing science.

Long before most computer scientists foresaw the possibilities of analyzing data sets then unimaginable for computers of the time, Kruskal was devising clever ad hoc methods for MDS with $n = 10,000$ AT&T switching input data values (Kruskal & Hart, 1966), at a time when the conventional upper limit was $n = 100$. Kruskal's co-edited 1983 volume (with David Sankoff) *Time Warps, String Edits, and Macromolecules* is widely recognized as a classic in the field of artificial intelligence. His contributions to and innovations in substantive areas also include such diverse fields as naval electronics (with Aumann in 1958), quantitative cytochemistry (with Giacomelli, Wiener, Pomeranz, & Loud in 1971) and computerized tomography (with Shepp in 1978). The deceased statistician John W. Tukey acknowledged that his developments of projection pursuit were based on ideas Kruskal had published earlier (Kruskal, 1972). At the end of his

career, Joe Kruskal started a thorough theoretical study of the mathematics of three-way models and methods (Kruskal, 1977), which has been very influential in the three-way community, including chemometrics.

Joe Kruskal's earlier research was in combinatorial mathematics. He devised a method now known as "Kruskal's algorithm" for efficiently computing the minimum spanning tree of a network, that is, the overall shortest collection of links connecting all the nodes of a large network. Two other important contributions from that period are Kruskal's Tree Theorem and the Kruskal-Katona Theorem.

Joseph B. Kruskal, Jr. was born in New York City in 1928. His father, Joseph B. Kruskal, Sr., owned a wholesale fur business. His mother, Lillian Kruskal Oppenheimer, was a leading promoter of origami in America from the 1950s onward. Among her gifted students was a young magician named Persi Diaconis, now an eminent mathematician and statistician at Stanford. Joe graduated from the University of Chicago and received his doctorate from Princeton University in 1954. Although his dissertation supervisors officially were Albert W. Tucker and Roger Lyndon, he attributed his main inspiration from conversations he had with Paul Erdős. Dr. Kruskal taught at the University of Wisconsin-Madison and the University of Michigan-Ann Arbor before moving to Bell Laboratories in Murray Hill, NJ, in 1958, where he worked on research until and after his official retirement. While there, his work yielded a patent for a statistical method to assess amplifiers on transatlantic fiber optic telephone cables, enabling the cables to have a useful life of at least fifty years.

Joe Kruskal became a Fellow of the American Statistical Association in 1971 and of the American Association for the Advancement of Science in 1982. He served as second President of the Classification Society of North America (now the Classification Society) in 1972–1973, as well as President of the Psychometric Society in 1974–1975. During the formation of the Classification Society starting in 1969, he had the foresight to realize the need for a bibliographic survey of the far-flung literature relevant to his and colleagues' research. He was highly instrumental in negotiating service from the Institute for Scientific Information and then persuading sociologist Ian C. Ross in the Bell Labs Technical Library to produce the annual bibliography SERVICE, a very useful resource for the classification community, which still exists today on the Website of the society. He received numerous forms of honorific awards from various mathematically and statistically oriented societies.

Joe was both a pure and an applied mathematician *par excellence*. In studying over our folders of Joe's correspondence and publications, we are particularly struck by how encouraging his advice was to junior colleagues like the two of us, and even benign to ideological adversaries in his own peer group. He was a self-effacing person who seemed unaware of his own truly monumental accomplishments. As both a kindly friend and top-internationally recognized major mathematician, his absence is a very sorrowful one.

References

- Aumann, R.J., & Kruskal, J.B. (1958). The coefficients in an allocation problem. *Naval Research Logistics Quarterly*, 5, 111–123.
- Ayer, M., Brunk, H.D., Ewing, G.M., Reid, W.T., & Silverman, E. (1955). An empirical distribution function for sampling with incomplete information. *The Annals of Mathematical Statistics*, 26, 641–647.
- Giacomelli, F., Wiener, J., Kruskal, J.B., Pomeranz, J.V., & Loud, A.V. (1971). Subpopulations of blood lymphocytes demonstrated by quantitative cytochemistry. *The Journal of Histochemistry and Cytochemistry*, 19, 426–433.
- Kruskal, J.B. (1965). Analysis of factorial experiments by estimating monotone transformations of the data. *Journal of the Royal Statistical Society, Series B*, 27, 251–263.
- Kruskal, J.B. (1972). Linear transformation of multivariate data to reveal clustering. In R.N. Shepard, et al. (Eds.), *Multidimensional scaling: theory and applications in the behavioral sciences, vol. I theory* (pp. 181–191). Oxford: Seminar Press.
- Kruskal, J.B. (1977). Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra and Its Applications*, 18, 95–138.

- Kruskal, J.B., & Hart, R.E. (1966). A geometric interpretation of diagnostic data for a digital machine: based on a Morris, Illinois Electronic Central Office. *Bell System Technical Journal*, 45, 1299–1338.
- Sankoff, D., & Kruskal, J.B. (1983). *Time warps, string edits, and macromolecules: the theory and practice of sequence comparison*. Reading: Addison-Wesley.
- Shepp, L.A., & Kruskal, J.B. (1978). Computerized tomography: the new medical x-ray technology. *The American Mathematical Monthly*, 85, 420–439.

Published Online Date: 20 DEC 2011