



POLSKA AKADEMIA NAUK
Instytut Badań Systemowych

**METHODS OF ESTIMATION
OF RELATIONS OF:
EQUIVALENCE,
TOLERANCE
AND PREFERENCE
IN A FINITE SET**

Leszek Klukowski

Warsaw 2011



**SYSTEMS RESEARCH INSTITUTE
POLISH ACADEMY OF SCIENCES**

**Series: SYSTEMS RESEARCH
Volume 69**

Series Editor:

Prof. dr hab. inż. Jakub Gutenbaum

Warsaw 2011

Editorial Board

Series: SYSTEMS RESEARCH

Prof. Olgierd Hryniewicz - chairman

Prof. Jakub Gutenbaum – series editor

Prof. Janusz Kacprzyk

Prof. Tadeusz Kaczorek

Prof. Roman Kulikowski

Prof. Marek Libura

Prof. Krzysztof Malinowski

Prof. Zbigniew Nahorski

Prof. Marek Niezgódka

Prof. Roman Słowiński

Prof. Jan Studziński

Prof. Stanisław Walukiewicz

Prof. Andrzej Weryński

Prof. Antoni Żochowski



**SYSTEMS RESEARCH INSTITUTE
POLISH ACADEMY OF SCIENCES**

Leszek Klukowski

**METHODS OF ESTIMATION
OF RELATIONS OF:
EQUIVALENCE
TOLERANCE
AND PREFERENCE
IN A FINITE SET**

Warsaw 2011

**Copyright © by Systems Research Institute
Polish Academy of Sciences
Warsaw 2011**

dr Leszek Klukowski
Systems Research Institute
Polish Academy of Sciences
Newelska 6, 01-447 Warsaw, Poland
email: Leszek.Klukowski@ibspan.waw.pl

Papers reviewers:

Prof. dr hab. inż. Ignacy Kaliszewski
Prof. dr hab. Tadeusz Trzaskalik

The work has been supported by the grant No N N111434937
of the Polish Ministry of Science and Higher Education

Printed in Polands
Systems Research Institute
Polish Academy of Sciences
Newelska 6, 01-447 Warsaw, Poland
www.ibspan.waw.pl

ISSN 0208-8029
ISBN 9788389475374

BIBLIOGRAPHY

- Abraham, A., Grosan C., eds., (2006) *Swarm Intelligence in Data Mining*. Springer
- Brunk, H. D. (1960) *Mathematical models for ranking from paired comparisons*. JASA Vol. 55, p. 503 – 520.
- Bradley, R. A. (1976) *Science, statistics and paired comparisons*. Biometrics vol. 32, p. 213 – 232.
- Bradley, R. A. (1984) *Paired comparisons: some basic procedures and examples*. Handbook of Statistics Vol. 4, ed. P. R. Krishnaiah and P. K. Se. Amsterdam: North-Holland, 299 – 326.
- Cormack R., M. (1971) *A review of classification (with discussion)*. JRSS A, p. 321 – 367.
- Cormen T., H., Leiserson C., E., Rivest R., L. Stein C. (2001): *Introduction to Algorithms*. Second ed. The MIT Press.
- Daniel, W. W. (1990) *Applied nonparametric Statistics (2nd ed.)*. Boston: PWS-Kent Publishing Company.
- David H. A. (1970) *Order Statistics*. J. Wiley.
- David H. A. (1988) *The Method of Paired Comparisons, 2nd ed.* Ch. Griffin, London.
- Davidson, R., R. (1976). *A bibliography on the Method of Paired Comparisons*. Biometrics Vol. 32, p. 241 – 251.
- Davoodzadeh, J., Beaver, R. J. (1982) *Models for multivariate paired comparisons with ties*. J. Math. Psychol. Vol. 25, p. 269 – 281.
- DeCani, J. S. (1972) *A branch and bound algorithm for maximum likelihood paired comparison ranking*. Biometrika 59, p. 131 – 136.
- deCastro L., N., (2002) *Artificial Immune Systems: A New Computational Intelligence Approach*. Springer
- Diday, E., Lechevallier, Y., Schader, M., Bertrand, P., Burtschy B. (1994) *New Approaches in Classification and Data Analysis*. Springer.

- Domański C. (1979) *Statystyczne testy nieparametryczne (Non-parametric statistical tests; in Polish)*. PWE, Warszawa.
- Domański C. (1990) *Testy statystyczne (Statistical tests; in Polish)*. PWE, Warszawa.
- Fisz, M., (1969) *Rachunek prawdopodobieństwa i statystyka matematyczna (Calculus of probability and mathematical statistics; in Polish)*. PWN, Warszawa.
- Falkenauer, E. (1998) *Genetics Algorithms and Grouping Data*. J. Wiley.
- Fraser R. A. (1957) *Nonparametric Models in Statistics*. J. Wiley.
- Gordon, A., D. (1999) *Classification, 2nd Edition*, Chapman&Hall/CRC.
- Hand, D. J. (1986) *Discrimination and Classification*. J. Wiley.
- Hartigan, J. J. (1975) *Clustering Algorithms*. J. Wiley.
- Hansen, P., Jaumard. B., Sanlaville E. (1994) *Partitioning Problems in Cluster Analysis: A Review of Mathematical Programming Approaches*. Studies in Classification, Data Analysis, and Knowledge Organization, Springer- Verlag.
- Hastie, T., Tibshirani, R., Friedman, J. (2002) *The Elements of Statistical Learning. Data Mining, Inference and Prediction*. Springer
- Hoeffding, W. (1963) *Probability inequalities for sums of bounded random variables*. JASA, Vol. 58, pp. 13 – 30.
- Hollander M., Wolfe D. A. (1973) *Nonparametric Statistical Methods*. J. Wiley
- Intelligent Data Analysis (2003). An Introduction*. (second ed.) Berthold, M., Hand, D., J. (editors).
- Kaufman, L., Rousseeuw, P. J. (1990) *Findings Groups In Data: An Introduction to Cluster Analysis*. J. Wiley.
- Klukowski L. (1986) *The method of forecasting of time series including piecewise linear component (in Polish)*. Systems Research Institute Polish Academy of Sciences, Warsaw (unpublished Ph.D. dissertation).
- Klukowski L. (1990a) *Algorithm for classification of samples in the case of unknown number of variables generating them (in Polish)*. Przegląd Statystyczny XXXVII, p. 167 – 177.
- Klukowski L. (1990b) *Ranking of items on the basis of pairwise comparisons in the case of random errors (in Polish)*. In: Kulikowski, R.,

- Kacprzyk, J. (eds.) SYSTEMS RESEARCH, Vol. 3, Omnitech Press, Warsaw, p. 212 - 268.
- Klukowski L. (1991) *Classification of hormonal profiles* (in Polish). The IInd Conference of Operational and Systems Research. Nahorski Z. (ed.), IBS PAN, Warsaw, p. 155 - 163.
- Klukowski L. (1994) *Some probabilistic properties of the nearest adjoining order method and its extensions*. Annals of Operational Research, **51**, 241 – 261.
- Klukowski L. (1997): *The nearest adjoining order method for pairwise comparisons in the form of difference of ranks*. International Conference TRANSITION to Advanced Market Institutions and Economies, Systems Research Institute, Polish Ac. of Sciences, INFORMS, IFORS, IFAC, IIASA, Warsaw, June 18-21, p. 220-225.
- Klukowski L. (2000): *The nearest adjoining order method for pairwise comparisons in the form of difference of ranks*. Annals of Operations Research vol. 97, p. 357-378.
- Klukowski, L. (2002) *Estimation of the tolerance relation on the basis of pairwise comparisons with random errors* (in Polish). In: Bubnicki, Z., Hryniewicz, O., Kulikowski, R. (eds.) Methods and Techniques of Data Analysis and Decision Support. Ac. Publ. House EXIT, Warsaw, p. V-21 – V-35.
- Klukowski L. (2006) *Tests for relation type - equivalence or tolerance - in finite set of elements*. Control and Cybernetics, **35**, 369 - 384.
- Klukowski, L. (2007a) *Estimation of the Preference Relation on the Basis of Medians from Pairwise Comparisons in the Form of Difference of Ranks*. In Kurzynski (et al.) Computer Recognition Systems 2, Advances in Soft Computing 45, Springer.
- Klukowski L. (2007b) *Estimation of tolerance relation on the basis of multiple pairwise comparisons with random errors*. Control and Cybernetics, **36**, 443-466.
- Klukowski L. (2007c) *Estimation of the preference relation on the basis of multiple pairwise comparisons with random errors*. In: Trzaskalik T. (eds.) MODELOWANIE PREFERENCJI A RYZYKO'08, AE Katowice 2008.
- Klukowski L. (2008a) *Determination of tolerance relation – alternative approach to intuitionistic and fuzzy sets*. Advances in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics. Vol. II. K. Atanassov (et al), eds. Ac. Publ. House EXIT, Warsaw.

Klukowski L. (2008b) ***Estimation of the preference relation on the basis of multiple pairwise comparisons in the form of differences of ranks***. Control and Cybernetics, **37**, 711-729.

Klukowski L. (2008c) ***Tests for Relation Type - Equivalence or Tolerance – Based on Averaging of Multiple Binary Comparisons***. In: Development In Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics. Foundations, Vol. I. K. Atanassov (et al), eds. Ac. Publ. House EXIT, Warsaw, p. 175 – 189.

Klukowski L. (2008d) ***Estimation of the preference relation on the basis of medians from pairwise comparisons***. In: Owskiński, J., W., Bruggemann (eds.) Multicriteria Ordering and Ranking: Partial Orders, Ambiguities and Applied Issues, IBS PAN, Warsaw, str. 109 - 120.

Klukowski L. (2010a) ***Estimation of the relations: preference, equivalence and tolerance on the basis of pairwise comparisons***. Eighth Workshop on Intuitionistic Fuzzy Sets and generalized Nets. In: Development In Fuzzy Sets, Intuitionistic Fuzzy Sets, Vol. I: Foundations. IBS PAN, SRI PAS, str. .

Klukowski L. (2010b) ***Tests for relation type – equivalence or tolerance based on medians from multiple pairwise comparisons***. In: Bojar W., Owskiński J., W. (eds.) MATERIAŁY I STUDIA POLSKIEGO STOWARZYSZENIA ZARZĄDZANIA WIEDZĄ. No. 31, 338 - 348.

Klukowski L. (2010c) ***Estimation of the preference relation on the basis of pairwise comparisons indicating direction of preference with random errors*** (in Polish). In: Owskiński J., W. (eds.) MATERIAŁY I STUDIA POLSKIEGO STOWARZYSZENIA ZARZĄDZANIA WIEDZĄ. Nr 31, 57 - 64.

Klukowski, L. (2011a) ***Estimation of Tolerance Relation on the Basis of Pairwise Comparisons***. In: Burduk, R., (et al.) Computer Recognition Systems 4, Advances in Intelligent and Soft Computing 95, Springer.

Klukowski L. (to appear, 2011a) ***Properties of estimators of the preference relation in the form of difference of ranks – simulation survey*** (in Polish) MPaR'2011 Conference, Ustroń, April 3 – 5, 2011. Economic University, Katowice.

Klukowski, L. (to appear, 2011b) ***Properties of estimators of the preference relation based on pairwise comparisons: binary and difference of ranks - simulation survey***. Submitted to Control and Cybernetics.

- Klukowski L., Kuba, E. (2002): *Stochastic optimization of strategy of management of treasury securities* (in Polish). National Bank of Poland, Materiały i Studia, 152, p. 34.
- Kohonen T. (1995) *Self-Organizing Maps*. Springer-Verlag.
- Kolonko J. (1980) *Discriminants Analysis ant its Application in Economy* (in Polish). PWN, Warszawa.
- Lipski W. (1982) *Kombinatoryka dla programistów (Combinatorics for programmers;* in Polish). WNT, Warszawa.
- Koronacki, J., Ćwik J. (2005) *Statystyczne systemy uczące się (Statistical learning systems;* in Polish). WNT, Warszawa.
- Marden J. I. (1995) *Analyzing and Modelling Rank Data*. Monograph on Statistics and Applied Probability 64. Chapman & Hall.
- Maydeu-Olivares, A., (1999) *Thurstonian modeling of ranking data via mean and covariance structure analysis*. Psychometrika, Vol. 64, no. 3, p. 325 – 340.
- McLachlan, G.J. (1992) *Discriminant Analysis and Statistical Pattern Recognition*, J. Wiley.
- Mulawka J., J. (1996) *Systemy ekspertowe (Expert systems;* in Polish). Sec. edition. WNT, Warsaw.
- Pawlak Z. (1982) *Rough Sets*. Int. Journ. Of Information and Computer Sciences., vol. 11, no. 5, p. 341 - 356.
- Philips, J. P. N. (1969) *A further procedure for determining Slater's i and all nearest adjoining orders*. Brit. J. Math. Statist. Psychol. Vol. 20, p. 217 – 225.
- Raghavachari M. (2005) *Measures of Concordance for Assessing Agreement in Ratings and Rank Order Data*. In: Balakrishnan, KannanN., Nagaraja H., N. (editors) *Advances in Ranking and Selection, Multiple Comparisons, and Reliability. Methodology and Applications, STATISTICS FOR INDUSTRY AND TECHNOLOGY*. Birkhäuser, Boston, Basel, Berlin.
- Randles R. H., Wolfe D., A. (1979) *Introduction to The Theory of Nonparametrics Statistics*. J. Wiley.
- Remage, R., Jr. And Thompson, W. A. Jr. (1966) *Maximum-likelihood paired comparison rankings*. Biometrika 53, p. 143 – 149.

- Ripley, B. D. (1996) ***Pattern Recognition and Neural Networks***. Cambridge University Press.
- Ripley B. D. (1996) ***Pattern Recognition and Neural Networks***. Cambridge University Press, Cambridge.
- Ripley B. D. (2006) ***Stochastic Simulation***. J. Wiley.
- Saaty Th., L. (1980) ***The analytic hierarchy process***. MacGraw-Hill.
- Sachs L. (1978) ***Applied Statistics***, Springer-Verlag, New York.
- Serfling, R., J. (1980): ***Approximation Theorems of Mathematical Statistics***. J. Wiley.
- Sheskin, D. J. (1997) ***Handbook of Parametric and Nonparametric Statistical Procedures***. CRC Press Boca Raton, Boston, London, New York.
- Siegel, S. and Castellan, N. J. Jr. (1988) ***Nonparametric Statistics for the Behavioral Sciences*** (2nd ed.). New York: McGraw-Hill Book Company.
- Slater P. (1961) ***Inconsistencies in a schedule of paired comparisons***. *Biometrika*, **48**, 303 - 312.
- Vapnik, V.N. (1998) ***Statistical Learning Theory***. J. Wiley.
- Vuong, Q., H. (1989) ***Likelihood ratio tests for model selection and non-nested hypothesis***. *Econometrica* vol. 57, p. 307 – 333.
- Webb A. R. (2003) ***Statistical Pattern Recognition***, 2nd Edition. J. Wiley.
- Witten, I. H., Frank E. (2005) ***Data Mining***, sec. ed., Elsevier
- Zuiew U, A. (1986) ***The probabilistic model of the committee of classifiers*** (in Russian). *The Journal of Computational Mathematics and Mathematical Physics*, vol. 26, no. 2, p. 276 -291.

The book presents the estimators of three relations: equivalence, tolerance, and preference in a finite set of data items, based on multiple pairwise comparisons, assumed to be disturbed by random errors. The estimators were developed by the author. They can refer to binary (qualitative), multivalent (quantitative) and combined comparisons. The estimates are obtained on the basis of solutions to the discrete programming problems. The estimators have been developed under weak assumptions on the distributions of comparison errors; in particular, these distributions can have non-zero expected values. The estimators have good statistical properties, including, especially importantly, consistency. Therefore, they produce good results in cases when other methods generate incorrect estimates. The precision of the estimators has been established with the use of simulation methods. The estimates can be validated in a versatile way. The whole estimation process, i.e. comparisons, estimation and validation can be computerized. The approach allows also for inference about the relation type – equivalence or tolerance, on the basis of binary data. Thus, it has features of data mining methods.

The estimators have been applied for ranking and grouping of data from some empirical sets. In particular, estimation of the tolerance relation (overlapping classification) was applied for determination of homogenous shapes of functions expressing profitability of treasury securities and was used for forecasting purposes.

ISSN 0208-8029
ISBN 9788389475374

**SYSTEMS RESEARCH INSTITUTE
POLISH ACADEMY OF SCIENCES**

Phone: (+48) 22 3810246 / 22 3810277 / 22 3810241 / 22 3810273
email: biblioteka@ibspan.waw.pl