

On Statistical Hypotheses Optimal Testing and Identification

Rudolf Ahlswede and Evgueni A. Haroutunian

Universität Bielefeld, Fakultät für Mathematik, Postfach 100131 33501 Bielefeld, Germany,
Institute for Informatics and Automation Problems of NAS of RA,
e-mail alhswede@mathematik.uni-bielefeld.de, evhar@ipia.sci.am

Abstract

A new aspect of the influence of the information-theoretical methods on the statistical theory is considered. The procedures of the probability distributions identification for $K (\geq 1)$ random objects each having one from the known set of $M (\geq 2)$ distributions are studied. N -sequences of discrete independent random variables represent results of N observations for each of K objects. Decisions concerning probability distributions of the objects must be made on the base of such samples. For $N \rightarrow \infty$ the exponential decrease of the test's error probabilities is considered. The reliability matrices of logarithmically asymptotically optimal procedures are explored for some models and formulations of the identification problems. The optimal subsets of reliabilities which values may be given beforehand and conditions guaranteeing positiveness of all the reliabilities are investigated.

References

- [1] Rao C. R., Linear statistical inference and its applications. Wiley, New York, 1965.
- [2] Bechhofer R. E., Kiefer J, and Sobel M., Sequential identification and ranking procedures. The University of Chicago Press, Chicago, 1968.
- [3] Ahlswede R. and Wegener I., Search problems. Wiley, New York, 1987.
- [4] Ahlswede R. and Dueck G., Identification via channels. IEEE Trans. Inform. Theory, vol. 35, no. 1, pp. 15-29, 1989.
- [5] Ahlswede R., General theory of information transfer. Preprint 97-118, SFB., Discrete Structures in der Mathematik, Universität Bielefeld.
- [6] Hoeffding W., Asymptotically optimal tests for multinomial distributions. Annals. of Math. Statist., vol. 36, pp. 369-401, 1965.
- [7] Csiszár I. and Longo G., On the error exponent for source coding and for testing simple statistical hypotheses. Studia Sc. Math. Hungarica, vol. 6, pp. 181-191, 1971.
- [8] Tusnady G., On asymptotically optimal tests. Annals of Statist., vol. 5, no. 2, pp. 385-393, 1977.
- [9] Longo G. and Sgarro A., The error exponent for the testing of simple statistical hypotheses, a combinatorial approach. J. of Combin., Inform. Sys. Sc., vol. 5, No 1, pp. 58-67, 1980.

- [10] Birgé L., Vitesse maximales de décroissance des erreurs et tests optimaux associés, *Z. Wahrsch. verw Gebiete*, vol. 55, pp. 261–273, 1981.
- [11] Haroutunian E. A., Logarithmically asymptotically optimal testing of multiple statistical hypotheses. *Problems of Control and Inform. Theory*. vol. 19, no. 5-6, pp. 413-421, 1990.
- [12] Csiszár I. and Körner J., *Information theory: Coding theorems for discrete memoryless systems*. Academic Press, New York, 1981.
- [13] Csiszár I., Method of types. *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2505-2523, 1998.
- [14] Natarajan S., Large deviations, hypotheses testing, and source coding for finite Markov chains. *IEEE Trans. Inform. Theory*, vol. 31, no. 3, pp. 360-365, 1985.
- [15] Haroutunian E. A., On asymptotically optimal testing of hypotheses concerning Markov chain (in Russian), *Izvestia Acad. Nauk Armenian SSR. Seria Mathem.* vol. 22, no 1, pp. 76-80, 1988.
- [16] Gutman M., Asymptotically optimal classification for multiple test with empirically observed statistics. *IEEE Trans. Inform. Theory*, vol 35, No 2, pp. 401-408, 1989.
- [17] Blahut R. E., *Principles and practice of information theory*. Addison-Wesley, Massachusetts, 1987.
- [18] Blahut R. E., Hypotheses testing and information theory, *IEEE Trans. Inform Theory*, vol 20, no 4, pp. 405-417, 1974.
- [19] Fu F.-W. and Shen S. Y., Hypothesis testing for arbitrarily varying source with exponents type constraint. *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 892-895, 1998.
- [20] Ahlswede R. and Csiszár I., Hypotheses testing with communication constraints. *IEEE Trans. Inform. Theory* vol. 32, no. 4, pp. 533-542, 1986.
- [21] Han T.S. and Kobayashi K., Exponential-type error probabilities for multiterminal hypothesis testing. *IEEE Trans. Inform. Theory*, vol. 35, no. 1, pp. 2-13, 1989.
- [22] Berger T., Decentralized estimation and decision theory, Presented at IEEE Seven Springs Workshop on Information Theory, Mt. Kisco, NY, September 1979.
- [23] Zhang Z. and Berger T., Estimation via compressed information. *IEEE Trans. Inform. Theory*, vol. 34, no. 2, pp. 198-211, 1988.
- [24] Ahlswede R. and Burnashev M., On minimax estimation in the presence of side information about remote data. *Annals of Statist.*, vol. 18, no. 1, pp. 141-171, 1990.
- [25] Han T. S. and Amari S., Statistical inference under multiterminal data compression, *IEEE Trans Inform Theory*, vol. 44, no. 6, pp. 2300-2324, 1998.
- [26] Ahlswede R., Yang E., Zhang Z., Identification via compressed data. *IEEE Trans. Inform. Theory*, vol. 43, no. 1, pp. 48-70, 1997.
- [27] Borovkov A. A., *Mathematical statistics* (in Russian). Nauka, Novosibirsk, 1997.
- [28] Ihara, *Information theory for continuous systems*. World Scientific, Singapore, 1993.
- [29] Chen P.-N., General formulas for the Neyman-Pearson type-II error exponent subject to fixed and exponential type-I error bounds. *IEEE Trans. Inform. Theory*, vol. 42, no. 1, pp. 316-323, 1996.
- [30] Han T. S., *Information-spectrum methods in information theory*. Springer, Berlin, 2003.
- [31] Han T.S. Hypothesis testing with the general source. *IEEE Trans. Inform. Theory*, vol. 46, no. 7, pp. 2415-2427, 2000.
- [32] Maurer U. M., Authentication theory and hypothesis testing. *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1350-1356, 2000.
- [33] Cachin C., An information-theoretic model for steganography. *Proc. 2nd Workshop on Information Hiding* (David Ausmith, ed.), *Lecture Notes in computer Science*, Springer-Verlag, 1998.

Վիճակագրական վարկածների օպտիմալ տեստման և նույնականացման մասին

Ռ.Ալավերդե, Ե. Հարությունյան

Անփոփում

Գիտարկված է ինֆորմացիոն-տեսական մեթոդների վիճակագրական տեսության վրա ազդեցության մի նոր հնարավորություն: Ուսումնասիրված են հավանականությունների բաշխումների նույնականացման ընթացակարգերը $K(\geq 1)$ պատահականացված օբյեկտների համար, որոնցից յուրաքանչյուրը բաշխված է ըստ տրված $M(\geq 2)$ բաշխումներից մեկի: Ընդհատ անկախ պատահական մեծությունների N -հաջորդականությունները ներկայացնում են K օբյեկտների յուրաքանչյուրի N դիտարկումների արդյունքները: Որոշումները պետք է ընդունվեն օբյեկտների հավանականային բաշխումների մասին: Գիտարկված է տեսերի սխալների հավանականությունների ցուցչային նվազումը, երբ $N \rightarrow \infty$: Հետազոտված են լոգարիթմորեն, ասիմպտոտորեն օպտիմալ ընթացակարգերի հուսալիության մատրիցները նույնականացման խնդրի մի քանի մոդելների և ձևակերպումների դեպքում: Նշված են այն հուսալիությունների օպտիմալ ենթաբազմությունները, որոնց արժեքները կարող են տրվել նախօրոք և այն պայմանները, որոնք ապահովում են բոլոր հուսալիությունների դրական լինելը: