Most Powerful Test for Multiple Hypotheses

Evgueni A. Haroutunian
Institute for Informatics and Automation Problems of NAS RA
Yerevan, Armenia
e-mail: eghishe@sci.am

Parandzem M. Hakobyan
Institute for Informatics and Automation Problems of NAS RA
Yerevan, Armenia
e-mail: par_h@ipia.sci.am

ABSTRACT
The aim of this paper is to present a generalization of the classical Neyman-Pearson Lemma to the case of more than two simple hypotheses.

Keywords
Multiple statistical hypotheses, Most powerful statistical test, Neyman-Pearson Lemma

1. INTRODUCTION
The principle of Neyman-Pearson plays a central role in both the theory and practice of statistics, because the Neyman-Pearson lemma is an important base of the mathematical theory of statistical hypothesis testing.

We call the statistical hypothesis each supposition statement which must be verified concerning the probability distribution of an observable random object. The task of statistician is to construct an algorithm (test) for effective detection of the hypothesis which is realized. The decision must be made on the base of vector (called a sample) of results of N independent identically distributed experiments, denoted by $\mathbf{x} \triangleq (x_1, ..., x_n, ..., x_N)$, the elements of $\mathcal{X}^N$, where $\mathcal{X}$ is the space of possible results of each experiment.

There exists a vast literature where the theory of the hypothesis testing and the Neyman-Pearson lemma are expounded in detail [1]–[9]. The paradigm of Neyman-Pearson is frequently used in different applications [10]–[12]. But the most part of these texts is dedicated to the case of two hypotheses.

Since the testing of multiple hypothesis is actual in applications we present a version of the Lemma for the case of three, or more hypotheses.
The concept of this study was formulated in [13].

It deserves to mention that the idea of consideration of exponential increase of error probabilities for all pairs of hypotheses amongst a known number of them was proposed and published in [15]–[18], and then developed for different models (see [19]–[24]).

2. PROBLEM STATEMENT AND RESULT FORMULATION
Let $\mathcal{P}(\mathcal{X})$ be the space of all probability distributions (PDs) on $\mathcal{X}$. Let X be RV taking values in $\mathcal{X}$ with one of M PDs $G_m \in \mathcal{P}(\mathcal{X})$, $m = \overline{1, M}$. Let the sample $\mathbf{x} = (x_1, ..., x_n, ..., x_N)$, $x_n \in \mathcal{X}$, $n = \overline{1, N}$, be a vector of results of N independent observations of X.

Based on data sample a statistician makes a decision which of the proposed hypotheses $H_m : G = G_m$, $m = \overline{1, M}$, is correct.

The procedure of decision making is a non-randomized test $\varphi_N(\mathbf{x})$, it can be defined by division of the sample space $\mathcal{X}^N$ on M disjoint subsets $\mathcal{A}_m$, $m = \overline{1, M}$. The set $\mathcal{A}_m$, $m = \overline{1, M}$ consists of vectors $\mathbf{x}$ for which the hypothesis H_m is adopted.

We study the probabilities of the erroneous acceptance of hypothesis H_l provided that H_m is true

$$\alpha_{l|m}(\varphi_N) \triangleq G_m(\mathcal{A}_l) = \sum_{\mathbf{x}:\ \mathbf{x} \in \mathcal{A}_l} G_m(\mathbf{x}),$$

$$m, l = \overline{1, M}, \ \ m \neq l.$$

If the hypothesis H_m is true, but it is not accepted then the probability of error is the following:

$$\alpha_{m|m}(\varphi_N) \triangleq \sum_{l: l \neq m} \alpha_{l|m}(\varphi_N) = 1 - G_m(\mathcal{A}_m), \quad m = \overline{1, M}.$$

As it was noted in [8] the case $N = 1$ "contains the general one and there is no need to restrict attention to independent drawings".

For the given preassigned values
$0 < \alpha^*_{1|1}, \alpha^*_{2|2}, ..., \alpha^*_{M-1|M-1} < 1$ we choose the numbers T_1, T_2, ..., T_{M-1} and the sets $\mathcal{A}_m$, $m = \overline{1, M}$, such that

$$\mathcal{A}^*_1 = \left\{ \mathbf{x} : \min \left(\frac{G_1(\mathbf{x})}{G_2(\mathbf{x})}, ..., \frac{G_1(\mathbf{x})}{G_M(\mathbf{x})} \right) > T_1 \right\},$$

$$1 - G_1(\mathcal{A}^*_1) = \alpha^*_{1|1},$$

$$\mathcal{A}^*_2 = \overline{\mathcal{A}^*_1} \cap \left\{ \mathbf{x} : \min \left(\frac{G_2(\mathbf{x})}{G_3(\mathbf{x})}, ..., \frac{G_2(\mathbf{x})}{G_M(\mathbf{x})} \right) > T_2 \right\},$$

$$1 - G_2(\mathcal{A}^*_2) = \alpha^*_{2|2},$$

....

$$\mathcal{A}^*_{M-1} = \overline{\mathcal{A}^*_1} \cap \overline{\mathcal{A}^*_2} \cap ... \cap \overline{\mathcal{A}^*_{M-2}} \cap \left\{ \mathbf{x} : \frac{G_{M-1}(\mathbf{x})}{G_M(\mathbf{x})} > T_{M-1} \right\},$$

$$1 - G_{M-1}(\mathcal{A}^*_{M-1}) = \alpha^*_{M-1|M-1},$$

and

$$\mathcal{A}^*_M = \mathcal{X}^N - (\mathcal{A}^*_1 \cup \mathcal{A}^*_2 \cup ... \cup \mathcal{A}^*_{M-1}) = \overline{\mathcal{A}^*_1} \cap \overline{\mathcal{A}^*_2} \cap ... \cap \overline{\mathcal{A}^*_{M-1}}.$$

The corresponding error probabilities are denoted by

$$\alpha^*_{l|m}(\varphi_N),\ m,l=\overline{1,M-1}.$$

Theorem: *The test determined by the sets $\mathcal{A}^*_1$, $\mathcal{A}^*_2$, ..., $\mathcal{A}^*_M$ is optimal in the sense that, for each other test defined by the set $\mathcal{B}_1$, $\mathcal{B}_2$, ..., $\mathcal{B}_M$ with the corresponding error probabilities $\beta_{l|m}$, $m,l=\overline{1,M}$,*

$$\textit{if}\ \ \beta_{m|m}\leq\alpha^*_{m|m},\ \ \textit{for}\quad \textit{any}\quad m=\overline{1,M-1},$$

then there exists at least one index j, $j\in[m+1,M]$ such that

$$\beta_{m|j}\geq\alpha^*_{m|j}.$$

Proof: Let $\Phi_{\mathcal{A}^*_m}$ and $\Phi_{\mathcal{B}_m}$ be the indicator functions of the decision regions $\mathcal{A}^*_m$ and $\mathcal{B}_m$. For all $\mathbf{x}=(x_1,x_2,...,x_N)\in\mathcal{X}^N$, the following inequality is correct

$$(\Phi_{\mathcal{A}^*_m}(\mathbf{x})-\Phi_{\mathcal{B}_m}(\mathbf{x}))\,(G_m(\mathbf{x})-$$

$$-\max(T_mG_{m+1}(\mathbf{x}),...,T_mG_M(\mathbf{x})))\geq 0.$$

Multiplying and then summing over $\mathcal{X}^N$ we obtain

$$\sum_{\mathbf{x}:\ \mathbf{x}\in\mathcal{X}^N}\big[\Phi_{\mathcal{A}^*_m}(\mathbf{x})G_m(\mathbf{x})-$$

$$-\Phi_{\mathcal{A}^*_m}(\mathbf{x})\max(T_mG_{m+1}(\mathbf{x}),T_mG_M(\mathbf{x}))-\Phi_{\mathcal{B}_m}(\mathbf{x})G_m(\mathbf{x})$$

$$+\ \Phi_{\mathcal{B}_m}(\mathbf{x})\max(TG_{m+1}(\mathbf{x}),...,T_mG_M(\mathbf{x}))]\geq 0,$$

$$\sum_{\mathbf{x}:\ \mathbf{x}\in\mathcal{A}^*_m}[G_m(\mathbf{x})-T_m\max(G_{m+1}(\mathbf{x}),...,G_M(\mathbf{x}))]-$$

$$-\sum_{\mathbf{x}:\ \mathbf{x}\in\mathcal{B}_m}[G_m(\mathbf{x})-T_1\max(G_{m+1}(\mathbf{x}),...,G_M(\mathbf{x}))]\geq 0,$$

According to the definition of error probability we obtain the following:

$$1-\alpha^*_{m|m}-T_m\max(\alpha^*_{m|m+1},...,\alpha^*_{m|M})$$

$$-(1-\beta_{m|m})+T_m\max(\beta_{m|m+1},...,\beta_{m|M})\geq 0,$$

$$-\beta_{m|m}+\alpha^*_{m|m}\leq T_m[-\max(\alpha^*_{m|m+1},...,\alpha^*_{m|M})$$

$$+\max(\beta_{m|m+1},...,\beta_{m|M})].$$

We see now that from $\beta_{m|m}\leq\alpha^*_{m|m}$, it follows that

$$\max(\beta_{m|m+1},...,\beta_{m|M})\geq\max(\alpha^*_{m|m+1},...,\alpha^*_{m|M}).$$

From this it follows that if the maximal is $\beta_{m|j}$, $j\in[m+1,M]$ then $\beta_{m|j}\geq\alpha^*_{m|j}$.

The theorem is proved.

3. CONCLUSION

In this paper we generalized Neyman-Pearson criterion of optimality to the case of many continuous hypotheses.

Bayesian testing was considered for the case of two and more hypotheses in [3], [4], [25], [26]. It is desirable to consider multyhypotheses Bayesian optimal testing for the models consisting of many objects.

4. ACKNOWLEDGEMENT

This work was supported in part by SCS of MES of RA under Thematic Program No SCS 13–1A295.

REFERENCES

[1] J. Neyman and E. S. Pearson, “On the problem of the most efficient tests of statistical hypotheses”, Phil. Trans. Roy. Soc. London, Ser. A, 231, pp. 289-337, 1933.

[2] J. Neyman, *First Course in Probability and Statistics*, Holt, Rinehart and Winston, New York, 1950.

[3] A. A. Borovkov, *Mathematical Statistics*, in Russian, Nauka, Moscow, 1997.

[4] H. L. Van Trees, *Detection, Estimation and Modulation Theory*, Part 1. New York: Wiley, 1968.

[5] M. H. DeGroot, *Probability and Statistics*, 2nd ed., Reading, MA, Addison-Wesley, 1986.

[6] M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics*, 2, Inference and relationship, Third edition. Hafner publishing company, London, 1961.

[7] A. K. Bera, “Hypothesis testing in the 20-th century with a special reference to testing with misspecified models”, In: “Statistics for the 21-st century. Methodologies for applications of the Future”, Marcel Dekker, Inc., New York, Basel, pp. 33-92, 2000.

[8] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press, New York, 1981, (Russian translation, Mir, Muscow, 1985).

[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Second Edition. Wiley, New York, 2006.

[10] E. Levitan and N. Merhav, “A competitive Neyman-Pearson approach to universal hypothesis testing with applications”, *IEEE Transactions on Information Theory*, vol. 48, no. 8, pp. 2215–2229, 2002.

[11] P. Moulin, “A Neyman-Pearson approach to universal erasure and list decoding”, *ISIT*, Toronto, Canada, July 6-11, 2008.

[12] P.-N. Chen, “General formulas for the Neyman-Pearson type-II error exponent subject to fixed and exponential type-I error bounds”, *IEEE Transactions on Information Theory*, vol. 42, no. 1, pp. 316–323, 1996.

[13] E. A. Haroutunian, “Neyman-Pearson principle for more than two hypotheses”, *Abstracts of Armenian Mathematical Union Annual Session Dedicated to 90 Anniversary of Rafael Alexandrian*, Yerevan, pp.49–50, 2013.

[14] E. Haroutunian, M. Haroutunian and A. Harutyunyan, Reliability Criteria in Information Theory and in Statistical Hypothesis Testing, *Foundations and Trends in Communications and Information Theory.* vol. 4, no. 2-3, 171 P., 2008.

[15] R. L. Dobrushin, Personal Communication, 1987.

[16] R. L. Dobrushin, M. S. Pinsker and A. N. Shiryaev, “Application of the notion of entropy in the problems of detecting a signal in noise,” (In Russian), *Lithuanian Mathematical Transactions*, vol. 3, no. 1, pp. 107–122, 1963.

[17] E. A. Haroutunian, “On asymptotically optimal tests for Markov chains”, *The First World Congress of Bernoulli Society*, vol. 2, no.3, pp. 153-156, 1989.

[18] E. A. Haroutunian, "Logarithmically asymptotically optimal testing of multiple statistical hypotheses," *Problems of Control and Information Theory*, vol. 19, no. 5–6, pp. 413-421, 1990.

[19] R. F. Ahlswede and E. A. Haroutunian, "On logarithmically asymptotically optimal testing of hypotheses and identification", *Lecture Notes in Computer Science, volume 4123, "General Theory of Information Transfer and Combinatorics", Springer*, pp. 462–478, 2006.

[20] E. Haroutunian and P. Hakobyan, "Multiple hypotheses LAO testing for many independent object", *Scholarly Research Exchange*, Article ID 921574, 2009.

[21] E. Haroutunian and A. Yessayan, "On reliability approach to multiple hypotheses testing and to identification of probability distributions of two stochastically related objects", *Proceedings of IEEE International Symposium on Information Theory* , pp. 2607-2011, 2011.

[22] E. Haroutunian, P. Hakobyan and A. Yessayan, "On Multiple Hypotheses LAO Testing With Rejection of Decision for Many Independent Objects", *Proceedings of International Conference CSIT 2009*. pp. 141-144, 2011.

[23] E. A. Haroutunian, " Reliability in multiple hypotheses testing and identification problems", *in Data Fusion for Situation Monitoring, Incident Detection, Alert and Response Management, IOS Press*, pp. 189–201, 2005.

[24] E. Haroutunian and P. Hakobyan, "Multiple objects: error exponents in hypotheses testing and identification", *Lecture Notes in Computer Science, Ahlswede Festschrift LNCS*, vol. 7777, Springer Verlag, pp. 313–345, 2013.

[25] J. Berger, "The case for objective bayesian analysis", *Bayesian Analysis*, vol. 1, pp. 385–402, 2006.

[26] C. C. Leang and D. H. Johnson, "On the asymptotics of M-hypothesis Bayesian detection", *IEEE Transactions on Information Theory*, vol. 43, no. 1, pp. 280–282, 1997.