Filomat 30:3 (2016), 681–688
DOI 10.2298/FIL1603681H

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Optimal Testing of Statistical Hypotheses and Multiple Familywise Error Rates

Farshin Hormozinejad[a]

[a] *Department of Mathematics and Statistics, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran*

Abstract. In this article the author considers the statistical hypotheses testing to make decision among hypotheses concerning many families of probability distributions. The statistician would like to control the overall error rate relative to draw statistically valid conclusions from each test, while being as efficient as possible. The familywise error (FWE) rate metric and the hypothesis test procedure while controlling both the type I and II FWEs are generalized. The proposed procedure shows simultaneous more reliability and less conservative error control relative to fixed sample and other recently proposed sequential procedures. Also, the characteristics of logarithmically asymptotically optimal (LAO) hypotheses testing are studied. The purpose of research is to express the optimal functional relation among the reliabilities of LAO hypotheses testing and to judge with FWE metric.

1. Introduction

In this paper the problems of hypotheses LAO testing for a model consisting of multiple families of probability distributions is studied. Hoeffding [15] and later Csiszár and Longo [7], Tusnady [18] and others studied asymptotically optimal tests. Haroutunian [11] solved the problem of LAO testing of multiple statistical hypotheses. Ahlswede and Haroutunian [1, 12, 13] formulated some problems of multiple hypotheses testing and identification for many objects. Multiple hypotheses LAO testing for many independent objects is also investigated [14] and the multistage tests of multiple hypotheses are studied by Bartroff and Lai [2].

If each experiment is considered as a hypothesis test about the corresponding data stream, then what is needed is a combination of a multiple hypothesis test and a sequential hypothesis test. This situation was addressed by Bartroff and Lai [4] who gave a procedure that sequentially tests S hypotheses while controlling the type I familywise error rate, i.e. the probability of rejecting any true hypotheses, at a prescribed level. Their procedure requires only the existence of basic sequential tests for each data stream and makes no assumptions about the dependence between the different data streams. The preceding situation occurs in a number of real applications including multi-channel change point detection (Tartakovsky et al. [17]) and its applications to biosurveillance (Mei [16]), genetics and genomics (Dudoit and vander Laan [10]).

The procedures of De and Baron [8, 9] simultaneously controlled both the type I and II FWEs. Bartroff and Lai [4] allowed arbitrary acceptances of null hypotheses while controlling the type I FWE, hence the

2010 *Mathematics Subject Classification*. Primary 62F03; Secondary 94A15
Keywords. Familywise error rate, hypotheses testing, information theory, reliability
Received: 26 July 2015; Revised: 19 September 2015; Accepted: 25 September 2015
Communicated by Ljubiša D.R. Kočinac and Ekrem Savaş
Research supported by Ahvaz Branch, Islamic Azad University, Ahvaz, Iran
Email address: hormozi-nejad@iauahvaz.ac.ir (Farshin Hormozinejad)

relationship between these acceptances and the power of the procedure was necessarily only available by analysis on a case-by-case basis.

In this article, the statistical hypotheses testing to make decision among hypotheses concerning many families of probability distributions is considered. Description of the characteristics of LAO hypotheses testing is studied. The purpose of research is to express the optimal functional relation among the reliabilities of LAO hypotheses testing and to judge with FWE metric. The remainder of the paper is organized as follows. Section 2 describes the model, preliminaries and formulations. Section 3 presents the optimal testing for a pair of familywise error rates. Section 4 discusses the optimal testing for multiple familywise error rates and Section 5 concludes.

2. Preliminaries and Formulations

Suppose random variable (RV) X characterizing an object takes values in the finite set $\mathcal{X}$ and $\mathcal{P}(\mathcal{X})$ is the space of all distributions on $\mathcal{X}$ and S hypothetical probability distributions (PDs) of X are given, but are divided in K disjoint families. The first family includes R_1 hypotheses $P_1, P_2, ..., P_{R_1}$, the second family includes R_2 hypotheses $P_{R_1+1}, P_{R_1+2}, ..., P_{R_1+R_2}$ and etc. the K-th family includes R_K hypotheses P_{S-K+1}, P_{S-K+2}, ..., P_S all from $\mathcal{P}(\mathcal{X})$. The considered object is characterized by RV X following to one of this S hypotheses. The statistician is trying to make reliable decision about correct distribution using sample $\mathbf{x} = (x_1, \ldots, x_N)$ of results of N independent observations of the RV X.

Let $N(x|\mathrm{x})$ be the number of repetitions of the element $x \in \mathcal{X}$ in the vector $\mathrm{x} \in \mathcal{X}^N$, then

$$Q_{\mathrm{x}} \triangleq \left\{ Q_{\mathrm{x}}(x) \triangleq \frac{N(x|\mathrm{x})}{N}, \ x \in \mathcal{X} \right\},$$

is the PD, called in statistics *the empirical probability distribution* of the sample x, and in information theory *the type* of x [5, 6].

Let $\mathcal{P}^N(\mathcal{X})$ be the set of all possible types of samples from $\mathcal{X}^N$ and $\mathcal{T}_Q^N$ be the set of all vectors x of the type $Q \in \mathcal{P}^N(\mathcal{X})$. The entropy of RV X with PD Q and the divergence (Kullback-Leibler distance) of PDs P and Q, are defined as follows [5, 6, 11]:

$$\mathrm{H}_Q(X) \triangleq -\sum_{x\in\mathcal{X}} Q(x)\log Q(x),$$

$$\mathrm{D}(Q \,\|\, P) \triangleq \sum_{x\in\mathcal{X}} Q(x)\log \frac{Q(x)}{P(x)}.$$

Let us remind the following useful properties of types [5, 6]:

$$|\mathcal{P}^N(\mathcal{X})| \le (N+1)^{|\mathcal{X}|},$$

$$(N+1)^{-|\mathcal{X}|}.\exp\{N\mathrm{H}_Q(X)\} \le |\mathcal{T}_Q^N| \le \exp\{N\mathrm{H}_Q(X)\},$$

$$P^N(\mathrm{x}) = \exp\{-N(\mathrm{H}_Q(X) + \mathrm{D}(Q\|P))\}, \ \text{for} \ \mathrm{x} \in \mathcal{T}_Q^N.$$

Let Θ be the parametric space and the individual parameter $\theta \in \Theta$ be a r-dimension vector. Based on data coming from a parametric family $P_s(\theta)$, $\theta \in \Theta$, of distributions, we will be concerned with testing a set of hypotheses $H_1, H_2, ..., H_S$. A hypothesis H_s is true if the true θ lies in H_s. If $T(\theta) \subseteq \{1, 2, ..., S\}$ is the set of indices of the true hypotheses, then the familywise error rate (FWE) is defined as the probability

$$\mathrm{FWE}(\theta) = P(\text{some } H_t \text{ is rejected}, \ t \in T)$$

Suppose there is an observable data x and S hypotheses divided into two disjoint families of PDs. It is desired to test about the parameter θ lies in one of the hypotheses H_s, $s = \overline{1, S}$. The sample space $\mathcal{X}^N$ has disjoint subsets of the space A_1 and A_2. The set A_1(or A_2) consists of all vectors x for which H_t, $t \in T$ is adopted(or rejected). The type I and II FWE, are defined as

$$FWE_I(\theta) = P(\text{some } H_t \text{ is rejected for } \ \mathrm{x} \in A_1,\ t \in T),$$

$$FWE_{II}(\theta) = P(\text{some } H_t \text{ is accepted for } \ \mathrm{x} \in A_2,\ t \in T).$$

This definition of FWE_I is the same as the standard one for fixed-sample testing and FWE_{II} is defined analogously. The quantity 1- FWE_{II} has been called familywise power by some authors (Ye et al. [19]).

3. The Optimal Testing for a Pair of Familywise Error Rates

The set of indices of all hypotheses are arranged into two sets: the true hypotheses $T(\theta) = \{\overline{1,R}\}$ and it's compliment $T'(\theta) = \{\overline{R+1,S}\}$. Therefore the pair of disjoint families of PDs $\mathcal{P}_1(\theta)$ and $\mathcal{P}_2(\theta)$ are:

$$\mathcal{P}_1(\theta) = \{P_s(\theta),\ s \in T\}, \quad \mathcal{P}_2(\theta) = \{P_s(\theta),\ s \in T'\}.$$

The decision making consists in using sample x for selection of a family of PDs is denoted by a test $\varphi_1^N(\mathrm{x})$, which can be defined by division of the sample space $\mathcal{X}^N$ on the pair of disjoint subsets

$$\mathcal{A}_i^N(\mathrm{x}) \triangleq \{\mathrm{x} : \varphi_1^N(\mathrm{x}) = i\},\ \ i = 1, 2.$$

The set $\mathcal{A}_i^N(\mathrm{x})$ consists of all vectors x for which i-th family $\mathcal{P}_i(\theta)$ of PDs is adopted.

The test $\varphi_1^N(\mathrm{x})$ has two kinds of errors for the pair of hypotheses $\mathcal{H}_i : \mathcal{P}_i(\theta),\ i = 1, 2$. Let $FWE_I(\varphi_1^N)$ be the probability of the erroneous acceptance of the second family $\mathcal{P}_2(\theta)$ provided the first family $\mathcal{P}_1(\theta)$ is true and $FWE_{II}(\varphi_1^N)$ be the probability of the erroneous acceptance of $\mathcal{P}_1(\theta)$ provided the second family $\mathcal{P}_2(\theta)$ is true. The errors are defined as

$$FWE_I(\varphi_1^N) \triangleq \max_{s:s\in T} P_s^N(\mathcal{A}_2^N), \tag{1}$$

$$FWE_{II}(\varphi_1^N) \triangleq \max_{s:s\in T'} P_s^N(\mathcal{A}_1^N). \tag{2}$$

The reliabilities of the sequence of tests φ_1 are defined as:

$$E_i(\varphi_1) \triangleq \liminf_{N\to\infty} \left\{-\frac{1}{N} \log FWE_i(\varphi_1^N)\right\},\ \ i = I, II. \tag{3}$$

The test φ_1 is considered to be LAO if for given value of E_I it provides the largest value to E_{II}. For given E_I^* the test φ_1^{*N} is defined by division of $\mathcal{X}^N$ into two disjoint subsets

$$\mathcal{A}_1^{N*} = \bigcup_{Q_{\mathrm{x}}:\min\limits_{s:s\in T} \mathrm{D}(Q_{\mathrm{x}}\|P_s)\le E_I^*} \mathcal{T}_{Q_{\mathrm{x}}}^N\ , \qquad \mathcal{A}_2^{N*} = \mathcal{X}^N \setminus \mathcal{A}_1^{N*}.$$

Theorem 3.1. *If all distributions $P_s(\theta)$, $s = \overline{1,S}$, are different and E_I^* is positive number such that the following inequality holds*

$$E_I^* < \min_{s:s\in T'} \min_{l:l\in T} \mathrm{D}(P_s\|P_l),$$

then there exists a LAO sequence of tests φ_1^* such that reliability $E_{II}^*(E_I^*)$ is positive and defined as

$$E_{II}^*(E_I^*) = \min_{s:s\in T'} \inf_{Q:\min\limits_{l:l\in T} \mathrm{D}(Q\|P_l)\le E_I^*} \mathrm{D}(Q\|P_s).$$

Proof. The first type of error $FWE_I(\varphi_1^{*N})$ is estimated by applying the properties of types as follows:

$$\begin{aligned} FWE_I(\varphi_1^{*N}) &= \max_{s:s\in T} P_s^N\left(\mathcal{A}_2^{N*}\right) \\ &= \max_{s:s\in T} P_s^N\Big(\bigcup_{Q_{\mathbf{x}}:\min\limits_{l:l\in T} D(Q_{\mathbf{x}}\|P_l)>E_I^*} \mathcal{T}_{Q_{\mathbf{x}}}^N\Big) \\ &< \max_{s:s\in T}(N+1)^{|\mathcal{X}|} \sup_{Q_{\mathbf{x}}:\min\limits_{l:l\in T} D(Q_{\mathbf{x}}\|P_l)>E_I^*} P_s^N\left(\mathcal{T}_{Q_{\mathbf{x}}}^N\right) \\ &\leq \max_{s:s\in T}(N+1)^{|\mathcal{X}|} \sup_{Q_{\mathbf{x}}:\min\limits_{l:l\in T} D(Q_{\mathbf{x}}\|P_l)>E_I^*} \exp\{-ND(Q_{\mathbf{x}}\|P_s)\} \\ &= \exp\Big\{-N\Big[\min_{s:s\in T} \inf_{Q_{\mathbf{x}}:\min\limits_{l:l\in T} D(Q_{\mathbf{x}}\|P_l)>E_I^*} D(Q_{\mathbf{x}}\|P_s) - o_N(1)\Big]\Big\} \\ &\leq \exp\left\{-N\{E_I^* - o_N(1)\}\right\}. \end{aligned}$$

Another error $FWE_{II}(\varphi_1^{*N})$ is similarly estimated as

$$\begin{aligned} FWE_{II}(\varphi_1^{*N}) &= \max_{s:s\in T'} P_s^N\left(\mathcal{A}_1^{N*}\right) \\ &= \max_{s:s\in T'} P_s^N\Big(\bigcup_{Q_{\mathbf{x}}:\min\limits_{l:l\in T} D(Q_{\mathbf{x}}\|P_l)\leq E_I^*} \mathcal{T}_{Q_{\mathbf{x}}}^N\Big) \\ &\leq \max_{s:s\in T'}(N+1)^{|\mathcal{X}|} \sup_{Q_{\mathbf{x}}:\min\limits_{l:l\in T} D(Q_{\mathbf{x}}\|P_l)\leq E_I^*} P_s^N\left(\mathcal{T}_{Q_{\mathbf{x}}}^N\right) \\ &\leq \max_{s:s\in T'}(N+1)^{|\mathcal{X}|} \sup_{Q_{\mathbf{x}}:\min\limits_{l:l\in T} D(Q_{\mathbf{x}}\|P_l)\leq E_I^*} \exp\{-ND(Q_{\mathbf{x}}\|P_s)\} \\ &= \exp\Big\{-N\Big[\min_{s:s\in T'} \inf_{Q_{\mathbf{x}}:\min\limits_{l:l\in T} D(Q_{\mathbf{x}}\|P_l)\leq E_I^*} D(Q_{\mathbf{x}}\|P_s) - o_N(1)\Big]\Big\}. \end{aligned} \tag{4}$$

And the inverse inequality is obtained as:

$$\begin{aligned} FWE_{II}(\varphi_1^{*N}) &= \max_{s:s\in T'} P_s^N\left(\mathcal{A}_1^{N*}\right) \\ &= \max_{s:s\in T'} P_s^N\Big(\bigcup_{Q_{\mathbf{x}}:\min\limits_{l:l\in T} D(Q_{\mathbf{x}}\|P_l)\leq E_I^*} \mathcal{T}_{Q_{\mathbf{x}}}^N\Big) \\ &\geq \max_{s:s\in T'} \sup_{Q_{\mathbf{x}}:\min\limits_{l:l\in T} D(Q_{\mathbf{x}}\|P_l)\leq E_I^*} P_s^N\left(\mathcal{T}_{Q_{\mathbf{x}}}^N\right) \\ &\geq \max_{s:s\in T'}(N+1)^{-|\mathcal{X}|} \sup_{Q_{\mathbf{x}}:\min\limits_{l:l\in T} D(Q_{\mathbf{x}}\|P_l)\leq E_I^*} \exp\{-ND(Q_{\mathbf{x}}\|P_s)\} \\ &= \exp\Big\{-N\Big[\min_{s:s\in T'} \inf_{Q_{\mathbf{x}}:\min\limits_{l:l\in T} D(Q_{\mathbf{x}}\|P_l)\leq E_I^*} D(Q_{\mathbf{x}}\|P_s) + o_N(1)\Big]\Big\}. \end{aligned} \tag{5}$$

According to the definition of the reliability (3) and from (4) and (5), the proof of the Theorem will be accomplished. If the condition of the Theorem is not held, then by applying the properties of types, $E_{II}^*(E_I^*) = 0$. □

Corollary 3.2. *If $R = 1$, $S = 2$, then there exist only two hypotheses P_1 and P_2. Therefore this case is equivalent to Hoeffding's Theorem [15], where for $E_I^* < D(P_2\|P_1)$,*

$$E_{II}^*(E_I^*) = \inf_{Q:D(Q\|P_1)\leq E_I^*} D(Q\|P_2).$$

4. The Optimal Testing for Multiple Familywise Error Rates

Suppose S possible PDs $P_s,\ \ s=\overline{1,S}$ of X are given and grouped in K disjoint families of PDs. The first family includes R_1 hypotheses, the second family includes R_2 hypotheses and etc., the last family includes R_K hypotheses such that $\sum_{k=1}^{K} R_k = S$. Consider the sets of indices

$$T_1=\{\overline{1,R_1}\},\ \ T_2=\{\overline{R_1+1,R_1+R_2}\},\ ...,\ \ T_K=\{\overline{S-R_K+1,S}\}.$$

Therefore there are K disjoint families of PDs $\mathcal{P}_1(\theta),\mathcal{P}_2(\theta),\ \ldots\ ,\mathcal{P}_K(\theta)$ such that

$$\mathcal{P}_k(\theta)=\{P_s(\theta),\ \ s\in T_k\},\ \ k=\overline{1,K}.$$

The decision making consists in using sample x for selection of one family of PDs is denoted by a test $\varphi_2^N(\mathrm{x})$, which can be defined by division of the sample space $\mathcal{X}^N$ on K disjoint subsets

$$\mathcal{A}_k^N \triangleq \{\mathrm{x}:\varphi_2^N(\mathrm{x})=k\},\ \ k=\overline{1,K}.$$

The set $\mathcal{A}_k^N$ consists of all vectors x for which k-th family of PDs is adopted.

Let $FWE_{m|k}(\varphi_2^N)$ be the familywise error of the acceptance of m-th family of PDs provided k-th family of PDs contains the correct PD:

$$FWE_{m|k}(\varphi_2^N)\triangleq\max_{s:s\in T_k}P_s^N(\mathcal{A}_m^N),\quad m\neq k,\ \ m,k=\overline{1,K}. \tag{6}$$

The familywise error to reject k-th family of PDs, when it is true, is

$$FWE_{k|k}(\varphi_2^N)\triangleq\max_{s:s\in T_k}P_s^N(\overline{\mathcal{A}}_k^N)=\sum_{m\neq k}FWE_{m|k}(\varphi_2^N),\ \ k=\overline{1,K}. \tag{7}$$

The reliabilities of the sequence of tests φ_2 are considered as

$$E_{m|k}(\varphi_2)\triangleq\liminf_{N\to\infty}\Big\{-\frac{1}{N}\log FWE_{m|k}(\varphi_2^N)\Big\},\ \ m,k=\overline{1,K}. \tag{8}$$

It follows from (6), (7) and (8) that

$$E_{k|k}(\varphi_2)=\min_{m\neq k}E_{m|k}(\varphi_2). \tag{9}$$

The test φ_2^* is called LAO if for given by consumer positive values of corresponding $K-1$ diagonal elements of the matrix of reliabilities, the procedure provides maximal values for other elements of it.

Theorem 4.1. *Consider S different PDs take place in K disjoint families of PDs. For given positive numbers $E_{1|1},E_{2|2},...,E_{K-1|K-1}$ let us introduce the regions:*

$$\mathcal{R}_k=\Big\{Q:\min_{s\in T_k}\mathrm{D}\,(Q\parallel P_s)\le E_{k|k}\Big\},\ \ k=\overline{1,K-1},$$

$$\mathcal{R}_K=\Big\{Q:\min_{s\in T_k}\mathrm{D}\,(Q\parallel P_s)> E_{k|k},\ \ k=\overline{1,K-1}\Big\},$$

and the following values for elements of the future matrix of reliabilities $\mathbf{E}(\varphi_2^*)$ of the LAO test sequence φ_2^*:

$$E_{k|k}^*=E_{k|k},\ \ k=\overline{1,K-1}, \tag{10}$$

$$E_{m|k}^*=\min_{s\in T_k}\inf_{Q\in\mathcal{R}_m}\mathrm{D}\,(Q\parallel P_s)\,,\ \ m,k=\overline{1,K},\ m\neq k, \tag{11}$$

$$E_{K|K}^*=\min_{m\neq K}E_{m|K}^*. \tag{12}$$

If the following compatibility conditions take place

$$0 < E_{1|1} < \min_{s\in T_1,\ l\in T_k,\ k=\overline{2,K}} \mathrm{D}(P_l \parallel P_s),$$

$$0 < E_{k|k} < \min[\min_{m=\overline{1,k-1}} E^*_{m|k}\ ,\ \min_{m=\overline{k+1,K},\ l\in T_m,\ s\in T_k} \mathrm{D}(P_l \parallel P_s)],\ \ 2 \le k \le K-1,$$

then there exists a LAO sequence of tests φ^*_2 *with matrix of reliabilities* $\mathbf{E}(\varphi^*_2)$.

Even if one of the compatibility conditions is violated, then the matrix of reliabilities of such test contains at least one element equal to zero.

Proof. The LAO test φ^{*N}_2 can be determined by division of $\mathcal{X}^N$ into K disjoint subsets

$$\mathcal{A}^{*N}_k = \bigcup_{Q_{\mathbf{x}}:Q_{\mathbf{x}}\in\mathcal{R}_k} \mathcal{T}^N_{Q_{\mathbf{x}}}\ ,\ \ k=\overline{1,K}.$$

By applying the properties of types for estimating of familywise errors and using the definition of the reliability, $FWE_{k|k}(\varphi^{*N}_2)$, $k=\overline{1,K-1}$, are estimated as:

$$\begin{aligned}
FWE_{k|k}(\varphi^{*N}_2) &= \max_{s\in T_k} P^N_s\left(\overline{\mathcal{A}^{*N}_k}\right)\\
&= \max_{s\in T_k} P^N_s\Big(\bigcup_{Q_{\mathbf{x}}:\min\limits_{l\in T_k}\mathrm{D}(Q_{\mathbf{x}}\|P_l)>E^*_{k|k}} \mathcal{T}^N_{Q_{\mathbf{x}}}\Big)\\
&\le \max_{s\in T_k}(N+1)^{|\mathcal{X}|} \sup_{Q_{\mathbf{x}}:\min\limits_{l\in T_k}\mathrm{D}(Q_{\mathbf{x}}\|P_l)>E^*_{k|k}} P^N_s\left(\mathcal{T}^N_{Q_{\mathbf{x}}}\right)\\
&\le \max_{s\in T_k}(N+1)^{|\mathcal{X}|} \sup_{Q_{\mathbf{x}}:\min\limits_{l\in T_k}\mathrm{D}(Q_{\mathbf{x}}\|P_l)>E^*_{k|k}} \exp\{-N\mathrm{D}(Q_{\mathbf{x}}\|P_s)\}\\
&= \exp\Big\{-N\Big[\min_{s\in T_k} \inf_{Q_{\mathbf{x}}:\min\limits_{l\in T_k}\mathrm{D}(Q_{\mathbf{x}}\|P_l)>E^*_{k|k}} \mathrm{D}(Q_{\mathbf{x}}\|P_s) - o_N(1)\Big]\Big\}\\
&\le \exp\left\{-N\{E^*_{k|k} - o_N(1)\}\right\},
\end{aligned}$$

where $o_N(1)\to 0$ is received by $N\to\infty$ and from here (10) follows.

The familywise errors for $m,k=\overline{1,K}$, $m\ne k$, are estimated as follows:

$$\begin{aligned}
FWE_{m|k}(\varphi^{*N}_2) &= \max_{s\in T_k} P^N_s\left(\mathcal{A}^{*N}_m\right)\\
&= \max_{s\in T_k} P^N_s\Big(\bigcup_{Q_{\mathbf{x}}:\min\limits_{l\in T_m}\mathrm{D}(Q_{\mathbf{x}}\|P_l)\le E^*_{m|m}} \mathcal{T}^N_{Q_{\mathbf{x}}}\Big)\\
&\le \max_{s\in T_k}(N+1)^{|\mathcal{X}|} \sup_{Q_{\mathbf{x}}:\min\limits_{l\in T_m}\mathrm{D}(Q_{\mathbf{x}}\|P_l)\le E^*_{m|m}} P^N_s\left(\mathcal{T}^N_{Q_{\mathbf{x}}}\right)\\
&\le \max_{s\in T_k}(N+1)^{|\mathcal{X}|} \sup_{Q_{\mathbf{x}}:\min\limits_{l\in T_m}\mathrm{D}(Q_{\mathbf{x}}\|P_l)\le E^*_{m|m}} \exp\{-N\mathrm{D}(Q_{\mathbf{x}}\|P_s)\}\\
&= \exp\Big\{-N\Big[\min_{s\in T_k} \inf_{Q_{\mathbf{x}}:\min\limits_{l\in T_m}\mathrm{D}(Q_{\mathbf{x}}\|P_l)\le E^*_{m|m}} \mathrm{D}(Q_{\mathbf{x}}\|P_s) - o_N(1)\Big]\Big\}.
\end{aligned} \tag{13}$$

Now let us confirm the inverse inequality

$$
\begin{aligned}
FWE_{m|k}(\varphi_2^{*N}) &= \max_{s\in T_k} P_s^N\left(\mathcal{A}_m^{*N}\right) \\
&= \max_{s\in T_k} P_s^N\Big(\bigcup_{Q_{\mathbf{x}}:\min\limits_{l\in T_m} D(Q_{\mathbf{x}}\|P_l)\le E^*_{m|m}} \mathcal{T}^N_{Q_{\mathbf{x}}}\Big) \\
&\ge \max_{s\in T_k} \sup_{Q_{\mathbf{x}}:\min\limits_{l\in T_m} D(Q_{\mathbf{x}}\|P_l)\le E^*_{m|m}} P_s^N\left(\mathcal{T}^N_{Q_{\mathbf{x}}}\right) \\
&\ge \max_{s\in T_k} (N+1)^{-|\mathcal{X}|} \sup_{Q_{\mathbf{x}}:\min\limits_{l\in T_m} D(Q_{\mathbf{x}}\|P_l)\le E^*_{m|m}} \exp\{-ND(Q_{\mathbf{x}}\|P_s)\} \\
&= \exp\Big\{-N\Big[\min_{s\in T_k} \inf_{Q_{\mathbf{x}}:\min\limits_{l\in T_m} D(Q_{\mathbf{x}}\|P_l)\le E^*_{m|m}} D(Q_{\mathbf{x}}\|P_s) + o_N(1)\Big]\Big\}.
\end{aligned} \tag{14}
$$

According to the definition of the reliability and the equations (13) and (14), we are gaining (11) as follows

$$
E^*_{m|k} = \min_{s\in T_k} \inf_{Q:\min\limits_{l\in T_m} D(Q\|P_l)\le E^*_{m|m}} D(Q\|P_s) = \min_{s\in T_k} \inf_{Q\in\mathcal{R}_m} D\left(Q \parallel P_s\right). \tag{15}
$$

Also the equation (12) can be received by (9).

The proof of the first part of the theorem will be accomplished if the sequence of tests φ_2^* is LAO and for agreed $E^*_{k|k}$, $k=\overline{1,K-1}$ and every other sequence of tests φ_2^{**} for all $m,k=\overline{1,K}$, $m\neq k$, $E_{m|k}(\varphi_2^{**})\le E^*_{m|k}$.

Suppose on the contrary, there exists a sequence of tests φ_2^{**} is defined by division of $\mathcal{X}^N$ into K disjoint subsets $\mathcal{B}_1^N,\ldots,\mathcal{B}_K^N$ and

$$
E_{m|k}(\varphi_2^{**}) > E_{m|k}(\varphi_2^{*}), \text{ for some } m,k=\overline{1,K},\ m\neq k. \tag{16}
$$

For large enough N, this condition is equivalent to the inequality

$$
FWE_{m|k}(\varphi_2^{**N}) \le FWE_{m|k}(\varphi_2^{*N}), \text{ for some } m,k=\overline{1,K},\ m\neq k. \tag{17}
$$

Examine the sets $\mathcal{B}_k^N \bigcap \mathcal{A}_k^{*N}$, $k=\overline{1,K}$. This intersection cannot be empty because

$$
\begin{aligned}
FWE_{k|k}(\varphi_2^{**N}) &= \max_{s\in T_k} P_s^N\left(\overline{\mathcal{B}}_k^N\right) \ge \max_{s\in T_k} P_s^N\left(\mathcal{A}_k^{*N}\right) \\
&\ge \max_{s\in T_k} \max_{Q_{\mathbf{x}}:\min\limits_{l\in T_k} D(Q_{\mathbf{x}}\|P_l)\le E^*_{k|k}} P_s^N\left(\mathcal{T}^N_{Q_{\mathbf{x}}}\right) \ge \exp\{-N(E^*_{k|k}+o_N(1))\}.
\end{aligned}
$$

Let's show $\mathcal{A}_k^{*N}\bigcap\mathcal{B}_m^N=\emptyset$, $m,k=\overline{1,K}$, $m\neq k$. Suppose the inverse if there exists $Q_{\mathbf{x}}$: $\min\limits_{s\in T_k} D(Q_{\mathbf{x}}\|P_s)\le E^*_{k|k}$ and $\mathcal{T}^N_{Q_{\mathbf{x}}}\subset\mathcal{B}_m^N$, then

$$
FWE_{m|k}(\varphi_2^{**N}) = \max_{s\in T_k} P_s^N\left(\mathcal{B}_m^N\right) \ge \max_{s\in T_k} P_s^N\left(\mathcal{T}^N_{Q_{\mathbf{x}}}\right) \ge \exp\{-N(E^*_{k|k}+o_N(1))\}.
$$

It follows $E_{m|k}(\varphi_2^{**})\le E^*_{k|k}$ and from equation (9) it also follows that $E^*_{k|k}\le E_{m|k}(\varphi_2^*)$ and at result $E_{m|k}(\varphi_2^{**})\le E_{m|k}(\varphi_2^*)$ which contradicts to (16). Hence it concludes $\mathcal{B}_k^N\bigcap\mathcal{A}_k^{*N}=\mathcal{A}_k^{*N}=\mathcal{B}_k^{*N}$ which means $\varphi_2^{**}\equiv\varphi_2^*$ and φ_2^* is the LAO test. For the proof of the second part of Theorem 4.1 it is sufficient to remark if one of the compatibility conditions is violated, then from equations (10)–(12), at least one of the elements $E^*_{m|k}$ is equal to zero. □

5. Conclusion

The sequential Holm procedure is a general method for combining individual sequential tests into a sequential multiple hypothesis testing procedure which controls both the type I and II FWEs at prescribed levels (Bartroff and Lai [3]). The sequential Holm procedure exhibits much more efficiency in terms of smaller average total sample size than existing sequential procedures, as well as Holm's fixed-sample test. This paper includes the optimality theory for multiple testing procedures and for calculating of their operating characteristics such as achieved FWEs and reliabilities. Description of the characteristics of LAO hypotheses testing is shown and the optimal functional relations among the reliabilities of LAO hypotheses testing with FWE metric and terms of optimality are expressed.

References

[1] R. Ahlswede, E.A. Haroutunian, On statistical hypotheses optimal testing and identification, Lecture Notes in Computer Science, General Theory of Information Transfer and Combinatorics, Springer, 4123 (2006) 462–478.
[2] J. Bartroff, T.L. Lai, Multistage tests of multiple hypotheses. Commun. Stat. Theory Methods 39 (2010) 1597–1607.
[3] J. Bartroff, J. Song, Sequential tests of multiple hypotheses controlling type I and II familywise error rates, J. Stat. Planning Inference 153 (2014) 100–114.
[4] J. Bartroff, T.L. Lai, M. Shih, Sequential Experimentation in Clinical Trials: Designand Analysis, Springer, New York, 2013.
[5] T.M. Cover, J.A. Tomas, Elements of Information Theory, Second edition, Wiley, New York, 2006.
[6] I. Csiszár, J. Körner, Information Theory: Coding Theorems for Discrete Memoryless Systems, Academic Press, New York, 1981.
[7] I. Csiszár, G. Longo, On the error exponent for source coding and for testing simple statistical hypotheses, Studia Sci. Math. Hungarica 6 (1971) 181–191.
[8] S. De, M. Baron, Sequential Bonferroni methods for multiple hypothesis testing with strong control of family-wise error rates I and II, Sequential Anal. 31 (2012) 238–262.
[9] S. De, M. Baron, Step-up and step-down methods for testing multiple hypotheses in sequential experiments. J. Stat. Planning Inference 142 (2012) 2059–2070.
[10] S. Dudoit, M. vanderLaan, Multiple Testing Procedures with Applications to Genomics, Springer-Verlag, New York, 2008.
[11] E.A. Haroutunian, Logarithmically asymptotically optimal testing of multiple statistical hypotheses, Problems of Control and Information Theory 19:5-6 (1990) 413–421.
[12] E.A. Haroutunian, Reliability in multiple hypotheses testing and identification, Proceedings of the NATO-ASI Conference, NATO Science Series III: Computer and Systems Sciences, Yerevan, Armenia, 198 (2008) 189–201.
[13] E.A. Haroutunian, M.E. Haroutunian, A.N. Haroutunian, Reliability criteria in information theory and in statistical hypothesis testing, Foundations and Trends in Communications and Information Theory 4:2-3, 2008.
[14] E.A. Haroutunian, P.M. Hakobyan, Multiple hypotheses LAO testing for many independent objects, Scholarly Research Exchange (2009) 1–6.
[15] W. Hoeffding, Asymptotically optimal tests for multinomial distributions, Annals Math. Stat. 36 (1965) 369–401.
[16] Y. Mei, Efficient scalable schemes for monitoring a large number of data streams, Biometrika 97 (2010) 419–433.
[17] A. Tartakovsky, X. Li, G. Yaralov, Sequential detection of targets in multichannel systems, IEEE Trans. Inf. Theory 49 (2003) 425–445.
[18] G. Tusnady, On asymptotically optimal tests, Annals Stat. 5 (1977) 385–393.
[19] Y. Ye, A. Li, L. Liu, B. Yao, A group sequential Holm procedure with multiple primary endpoints, Stat. Med. 32 (2013) 1112–1124.