The Korean Communications
in Statistics Vol. 10 No. 1, 2003
pp. 211-224

Tests to Detect Changes in Micro-Flora Composition

Donguk Kim[1], Mark C. K. Yang[2]

Abstract

Good's lambda test, a permutation test used to detect the changes of microorganism composition under two pathological conditions, has been quite popular for studying the micro-flora responsible for periodontal disease. A vast number of different micro-flora in the mouth renders the traditional chi-square test inapplicable. The main purpose of this paper is to evaluate the power of this test so that the sample size can be determined at the design stage. The robustness of this test and its comparison to two other intuitive tests are also presented. It is found that a permutation test based on likelihood ratio is more powerful than the lambda test in our simulated cases.

Keywords : closeness, composition, lambda test, likelihood ratio, permutation test, power.

1. Introduction

The λ-test, introduced by Good(1982), seems to be the only available test for testing the change of a multinomial distribution when the number of categories is large. The traditional chi-square test is not applicable because most of the cells have very few observations. Merging cells is also questionable because it can be done only after the data have been gathered. Forming a hypothesis after observing the data makes the true significance level very difficult to compute.

The λ-test was constructed to fill this gap. It is especially useful in determining the change of microorganisms in two pathological conditions when the micro-flora species are numerous, such as the micro-flora in the mouth. This test has been routinely used in periodontal disease research.(Moore et al.(1982, 1991), Socransky et al.(1981))

The λ-statistic is the ratio of the total between group closeness to the total within group closeness but the distribution of the λ-statistic is not known. We use permutation test to find the distribution of λ under the null and under the alternative. When we analyze an experiment with a permutation test, we compare the observed value of the test statistic with the set of other values we obtain by rearranging and relabeling the data. Permutation tests can provide exact significance levels and the large-sample power of permutation tests was

1) Associate Professor, Department of Statistics, Sungkyunkwan University, Seoul 110-745, Korea
E-mail : dkim@skku.ac.kr
2) Professor, Department of Statistics, University of Florida, Gainesville, Florida 32611, U.S.A

studied.(Hoeffding(1952))

We suggest two reasonable tests for comparison, where one is measuring closeness by the square of the Euclidean distance and the other is the likelihood ratio test. The main purpose of this paper is to evaluate the power and the null hypothesis significance level of three tests, so that we can answer the sample size question. Furthermore, we will study the permutational distributions of three tests.

In the next section, we introduce the design, some notation, and the test statistics of interest. Section 3 gives power comparisons of three tests. Section 4 describes in detail the permutational method and permutational distributions of three tests. In section 5, we give concluding remarks.

2. Notation and Tests

To define the basic notation, suppose there are two disease populations, where one may serve as the control. A sample of patients from each population is drawn and for each patient a vector of microorganisms is obtained from the culture of certain tissue or secretion sample from the patient. Let the sample sizes of the two groups be n_1 and n_2 and the micro-flora counts be $\mathbf{X}_i=\{x_{i1}, x_{i2}, \ldots, x_{iK}\}$ for the ith individual in group 1, $i=1,2,\ldots,n_1$ and $\mathbf{Y}_j=\{y_{j1}, y_{j2}, \ldots, y_{jK}\}$ be that of the jth individual in group 2, $j=1,2,\ldots,n_2$. In most practical situations, the number of possible species K is large and unknown, and many of the cells may end up with 0 observation. However, it will be clear that the tests discussed in this paper do not depend on K. Thus, we may assume that K is a large fixed number that covers all possible microorganisms. The goal is to test whether the microorganism compositions are the same in the two populations. Let the proportional composition for $\mathbf{X}_i$ and $\mathbf{Y}_j$ be respectively,

$$\mathbf{P}_i=\{p_{i1}, p_{i2}, \ldots, p_{iK}\}, \text{ and } \quad \mathbf{Q}_j=\{q_{j1}, q_{j2}, \ldots, q_{jK}\},$$

where

$$p_{ik}=x_{ik}/m_{1i}, \quad m_{1i}=\sum_{k=1}^{K} x_{ik}, \text{ and } \quad q_{jk}=y_{jk}/m_{2j}, \quad m_{2j}=\sum_{k=1}^{K} y_{jk}.$$

Good(1982) defined the closeness between two general proportional vectors $\mathbf{P}=\{p_1, p_2, \ldots, p_K\}$ and $\mathbf{Q}=\{q_1, q_2, \ldots, q_K\}$ as

$$\delta(\mathbf{P}, \mathbf{Q})=\sum_{k=1}^{K} \min(p_k, q_k). \tag{1}$$

This denotes the similarity index and we note that $\delta(\mathbf{P}, \mathbf{Q})=1$ when $\mathbf{P}=\mathbf{Q}$ and $\delta(\mathbf{P}, \mathbf{Q})=0$ when $\mathbf{P}$ and $\mathbf{Q}$ have mutually exclusive sets of positive elements. Also we note that (1) is not affected when K is unknown. Let the total within group closeness W be defined by

$$W=W_1+W_2,$$
$$W_1=\sum_{i>j}\delta(\mathbf{P}_i, \mathbf{P}_j),\quad W_2=\sum_{i>j}\delta(\mathbf{Q}_i, \mathbf{Q}_j),$$

and the total between group closeness be

$$B=\sum_i\sum_j\delta(\mathbf{P}_i, \mathbf{Q}_j).$$

Then the λ-statistic is defined as

$$\lambda=\frac{B/(n_1n_2)}{2W/[n_1(n_1-1)+n_2(n_2-1)]}, \tag{2}$$

and the null hypothesis should be rejected when λ is too small. The λ-statistic is the ratio of the average value of δ between groups to the average within groups. Hence if λ is too small, then it shows some degree of difference in compositions between two groups. Unfortunately, the distribution of λ, even under the null hypothesis that the two populations are identical, seems intractable. Good(1982) suggests that permutation be used to find the distribution of λ under the null hypothesis. More precisely, one will partition the combined n_1 and n_2 data randomly into two groups with sizes n_1 and n_2, compute the λ of this partition, repeat the process over a large number of times, and use the Monte-Carlo sample to estimate the distribution of λ. The p-value is the quantile point of the original λ in the estimated distribution. But at the design stage, a question often asked by the experimenter is what should be the sample size? Since the determination of the micro-flora composition is based on counting the micro-flora colonies in culture plates, the process is usually time consuming and expensive. In this paper we will give answer the sample size question based on power study.

There seems to be no discussion on the optimality of using (1) and (2) to test the change in two multinomial distributions. The reason must be due to the vast amount of possibilities of the alternative hypothesis. Here we suggest two reasonable tests for comparison. The one is to measure closeness by the square of the Euclidean distance

$$d(\mathbf{P}, \mathbf{Q}) = \sum_{k=1}^{K} (p_k - q_k)^2. \tag{3}$$

This distance is again unaffected by the unknown number K and if the δ used for (2) is replaced by d in (3), then we should reject the null hypothesis if λ is too large.

We may also choose to use the likelihood ratio as a rejection criterion. It is obvious that both $\mathbf{X}_i$ and $\mathbf{Y}_j$ have multinomial distributions. The likelihood functions under the null hypothesis is then

$$L(w) = \prod_{i=1}^{n_1} \frac{m_{1i}!}{x_{i1}! \dots x_{iK}!} \prod_{j=1}^{n_2} \frac{m_{2j}!}{y_{j1}! \dots y_{jK}!} p_1^{x_{\cdot 1} + y_{\cdot 1}} \cdots p_K^{x_{\cdot K} + y_{\cdot K}},$$

where $x_{\cdot k} = \sum_{i=1}^{n_1} x_{ik}$, $y_{\cdot k} = \sum_{j=1}^{n_2} y_{jk}$, and the p_k is the true proportion of species k, $k = 1, \dots, K$. Similarly, under the alternative hypothesis the likelihood function becomes

$$L(\Omega) = \prod_{i=1}^{n_1} \frac{m_{1i}!}{x_{i1}! \dots x_{iK}!} p_1^{x_{\cdot 1}} \cdots p_K^{x_{\cdot K}} \prod_{j=1}^{n_2} \frac{m_{2j}!}{y_{j1}! \dots y_{jK}!} q_1^{y_{\cdot 1}} \cdots q_K^{y_{\cdot K}},$$

where the q_k is the true proportion of species k in the second group. We reject the null hypothesis if the statistics Λ is too large, and

$$\begin{aligned} \Lambda = & \sum_{k=1}^{K} x_{\cdot k} \ln x_{\cdot k} + \sum_{k=1}^{K} y_{\cdot k} \ln y_{\cdot k} - M_1 \ln M_1 - M_2 \ln M_2 \\ & - \sum_{k=1}^{K} (x_{\cdot k} + y_{\cdot k}) \ln (x_{\cdot k} + y_{\cdot k}) + (M_1 + M_2) \ln (M_1 + M_2), \end{aligned} \tag{4}$$

where $M_1 = \sum_{k=1}^{K} x_{\cdot k}$ and $M_2 = \sum_{k=1}^{K} y_{\cdot k}$. By the convention $0 \ln 0 = 0$, this test does not depend on K either. Since the number of empty cells is large, we do not expect the asymptotic chi-square distribution for 2Λ to work, and also, we do not expect any analytic solution for the tests (2) and (4). The distribution can be obtained by the simulation based on random partitions as before.

3. Power Estimation

Let the true frequency distribution of the populations 1 and 2 be $\mathbf{F}=\{p_i\}_{i=1}^{K}$ and $\mathbf{G}=\{q_i\}_{i=1}^{K}$, respectively. Thus, under the null hypothesis $\mathbf{F}=\mathbf{G}$, it is difficult to specify a general alternative hypothesis. We can think of two scenarios that may have some general application;

In population 2,
(A1) there is a small shift of prevalence in all the categories, or
(A2) the main difference is in a few species.

We feel the latter seems more applicable in pathology; i.e., several pathogens flourish during the disease. There are many necessary parameters in power computation. The two obvious parameters are the significance level α and the sample sizes n_1 and n_2. We assume $n_1=n_2$ at the design stage. The data $\mathbf{X}_i$ and $\mathbf{Y}_j$ depend on the colony sizes m_{1i} and m_{2j} in the culture. For example, the abundance of micro-flora j of patient i in group 1, $p_{ij}=$ the number of colonies of micro-flora j / total number of colonies in the culture. Thus, if the number of colonies is small, it is not possible to have a large number of nonzero p_{ij}'s. These number of colonies per culture will vary somewhat in the actual culturing, but we assume it is a constant $m=m_{1i}=m_{2j}$ in the design stage. Moreover, we need the configurations of $\mathbf{F}$ and $\mathbf{G}$, which may contain many unknown parameters.

We first approximate the configurations of $\mathbf{F}$ and $\mathbf{G}$ with a gamma density function. Without loss of generality, we may consider $\mathbf{F}$ as the discretized distribution of a continuous density function $f(x)$, i.e.,

$$p_i=\int_{i-1}^{i} f(x)dx,\ i=1,2,\ldots,K,\quad \text{with}\quad \int_{K}^{\infty} f(x)\,dx\approx 0. \tag{5}$$

Similarly, we define $q_i=\int_{i-1}^{i} g(x)dx$. The two parameter gamma density family is rich enough to represent a large variety of frequency distributions.(Fisher et al.(1943)) A simulation program is written for the power for any input of gamma densities $f(x)$ and $g(x)$, K, $n_1=n_2$, the colony size m, and the level of significance α. We can choose the design parameters, but we have found that the power is not very sensitive to the choice of the gamma density functions. Thus, a guideline can be provided even without a good specification of the actual configurations of $\mathbf{F}$ and $\mathbf{G}$. The major finding can be summarized as follows.

We have found that under the alternative hypothesis (A1), tests based on (1), (3) and (4) have little advantage over each other, but (1) and (4) are much more powerful than (3) when the alternative is (A2). Also, between (1) and (4), (4) is better. A representative result for (A1) is given in Table 1, where under the null hypothesis $f(x)=g(x)=$ gamma$(\alpha=2, \beta=0.077)$ according to the usual gamma density function parameterization. For power computation, $f(x)$ is still gamma(2, 0.077), but $g(x)$ has been shifted to gamma(2.5, 0.082), or (3.0, 0.092). These numbers were chosen to satisfy requirement (5) with $K\approx 100$. Figure 1 shows the magnitude of the frequencies and the shift in composition. Every species has a small change in frequency. This table is representative because it covers a large range of powers. Some other tables may go to the extreme so quickly that the powers are either very small or close to 1.0, and in these cases, the powers for all three tests are very similar.

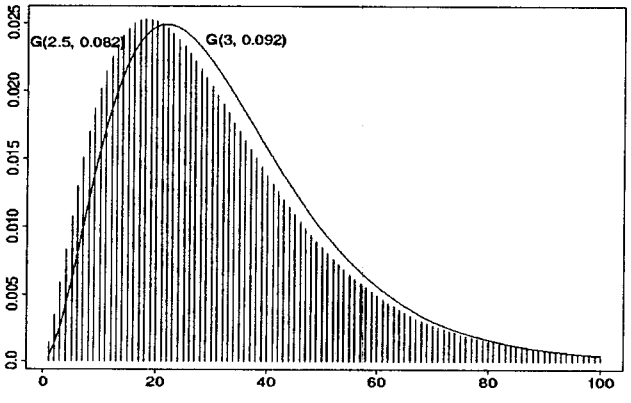


Figure 1. Gamma densities $(\alpha, \beta)=(2.5, 0.082)$(vertical lines) and $(\alpha, \beta)=(3, 0.092)$(curve)

Table 1. Power comparisons of λ, Euclidean distance and likelihood ratio permutation tests when the shift of species frequencies is everywhere. Table I is for the null hypothesis significance level and Tables II and III are for the powers.

I. $(\alpha, \beta)=(2, 0.077)$ under the null.

	$n_1=n_2=5$			$n_1=n_2=10$			$n_1=n_2=15$		
m	Lambda	Squared	LRT	Lambda	Squared	LRT	Lambda	Squared	LRT
25	0.056	0.050	0.059	0.067	0.058	0.056	0.049	0.054	0.038
50	0.064	0.055	0.065	0.051	0.055	0.056	0.052	0.056	0.051
75	0.058	0.056	0.049	0.044	0.051	0.049	0.058	0.058	0.048
100	0.051	0.053	0.052	0.062	0.065	0.051	0.053	0.048	0.048
125	0.043	0.054	0.054	0.052	0.057	0.055	0.042	0.046	0.040

II. $(\alpha, \beta)=(2, 0.077)$ and $(2.5, 0.082)$ under the alternative.

	$n_1=n_2=5$			$n_1=n_2=10$			$n_1=n_2=15$		
m	Lambda	Squared	LRT	Lambda	Squared	LRT	Lambda	Squared	LRT
25	0.093	0.092	0.103	0.158	0.146	0.187	0.253	0.249	0.317
50	0.175	0.162	0.195	0.349	0.344	0.395	0.577	0.547	0.629
75	0.254	0.226	0.295	0.605	0.574	0.665	0.827	0.837	0.891
100	0.297	0.316	0.359	0.769	0.740	0.814	0.948	0.941	0.970
125	0.417	0.405	0.466	0.880	0.844	0.923	0.981	0.983	0.996

III. $(\alpha, \beta)=(2, 0.077)$ and $(3, 0.092)$ under the alternative.

	$n_1=n_2=5$			$n_1=n_2=10$			$n_1=n_2=15$		
m	Lambda	Squared	LRT	Lambda	Squared	LRT	Lambda	Squared	LRT
25	0.201	0.202	0.231	0.441	0.418	0.531	0.711	0.698	0.784
50	0.437	0.413	0.499	0.880	0.858	0.926	0.991	0.989	0.995
75	0.658	0.675	0.764	0.989	0.982	0.996	1.0	1.0	1.0
100	0.847	0.820	0.905	1.0	1.0	1.0	1.0	1.0	1.0
125	0.945	0.929	0.977	1.0	1.0	1.0	1.0	1.0	1.0

Three representative results for (A2) is given in Figures 2-4 and Tables 2-4, where we assume that there is a 1% increase in each of the 5 least prevalent species, i.e., the proportion of the last 5 species has each increased from nearly 0 to 0.01. The proportions of the rest species were normalized so that the total proportions became 1. We can see that in this situation, the likelihood ratio test is much better than the Euclidean distance test and considerably better than the λ-test. These tables also show that the power is insensitive to the change in $f(x)$. Thus, we may conclude that both the likelihood ratio test and the λ-test are robust to the species composition.

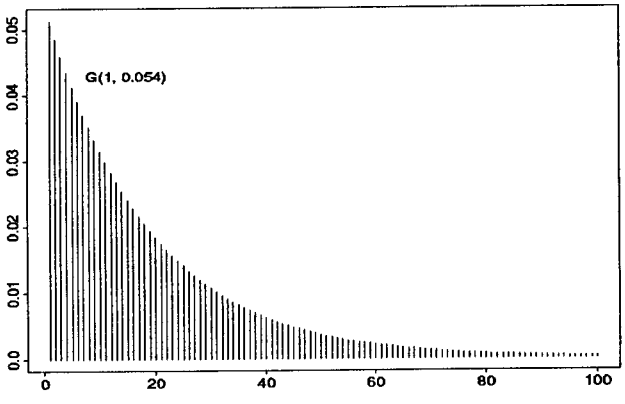


Figure 2. Gamma density $(\alpha, \beta)=(1, 0.054)$.

Table 2. The null hypothesis significance level and power comparisons of the three tests when the alternative is a 1% increase in each of the 5 least prevalent species.

$(\alpha, \beta) = (1, 0.054)$ under the null.

	$n_1 = n_2 = 5$			$n_1 = n_2 = 10$			$n_1 = n_2 = 15$		
m	Lambda	Squared	LRT	Lambda	Squared	LRT	Lambda	Squared	LRT
25	0.056	0.056	0.062	0.073	0.069	0.075	0.043	0.040	0.042
50	0.051	0.047	0.055	0.056	0.055	0.055	0.046	0.047	0.052
75	0.053	0.056	0.066	0.059	0.056	0.056	0.040	0.044	0.045
100	0.053	0.054	0.049	0.048	0.052	0.046	0.042	0.046	0.055
125	0.051	0.052	0.050	0.062	0.057	0.052	0.060	0.052	0.042

$(\alpha, \beta) = (1, 0.054)$ and a 1% increase in each of the 5 least prevalent species under the alternative.

	$n_1 = n_2 = 5$			$n_1 = n_2 = 10$			$n_1 = n_2 = 15$		
m	Lambda	Squared	LRT	Lambda	Squared	LRT	Lambda	Squared	LRT
25	0.074	0.083	0.102	0.111	0.079	0.196	0.178	0.109	0.374
50	0.134	0.091	0.196	0.290	0.134	0.545	0.450	0.204	0.816
75	0.197	0.110	0.373	0.473	0.195	0.841	0.763	0.320	0.985
100	0.305	0.138	0.548	0.668	0.278	0.948	0.926	0.494	0.998
125	0.373	0.172	0.665	0.823	0.375	0.990	0.984	0.651	1.0

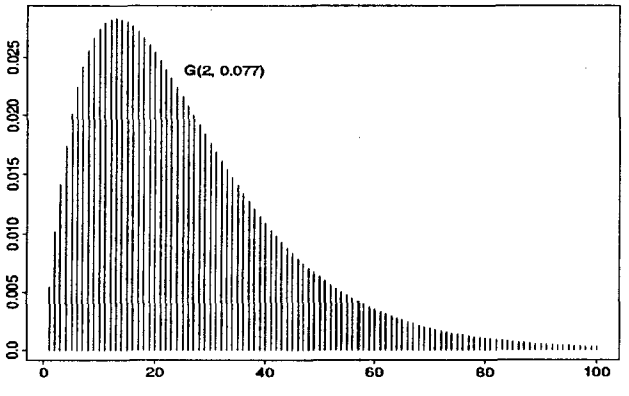


Figure 3. Gamma density $(\alpha, \beta) = (2, 0.077)$.

Table 3. Power comparisons of the three tests when the alternative is a 1% increase in each of the 5 least prevalent species.

$(\alpha, \beta) = (2, 0.077)$ and a 1% increase in each of the 5 least prevalent species under the alternative.

	$n_1 = n_2 = 5$			$n_1 = n_2 = 10$			$n_1 = n_2 = 15$		
m	Lambda	Squared	LRT	Lambda	Squared	LRT	Lambda	Squared	LRT
25	0.103	0.082	0.100	0.113	0.093	0.188	0.163	0.116	0.326
50	0.126	0.091	0.195	0.278	0.162	0.521	0.452	0.231	0.805
75	0.205	0.120	0.351	0.509	0.289	0.826	0.765	0.389	0.978
100	0.260	0.144	0.491	0.680	0.351	0.950	0.930	0.583	0.998
125	0.379	0.190	0.668	0.820	0.455	0.985	0.988	0.745	1.0

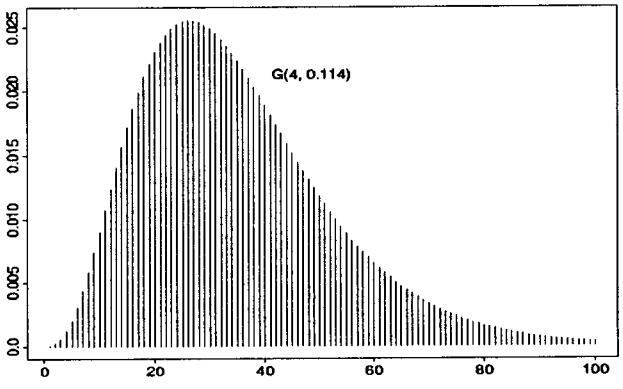


Figure 4. Gamma density $(\alpha, \beta) = (4, 0.114)$.

Table 4. The null hypothesis significance level and power comparisons of the three tests when the alternative is a 1% increase in each of the 5 least prevalent species.

$(\alpha, \beta) = (4, 0.114)$ under the null.

	$n_1 = n_2 = 5$			$n_1 = n_2 = 10$			$n_1 = n_2 = 15$		
m	Lambda	Squared	LRT	Lambda	Squared	LRT	Lambda	Squared	LRT
25	0.046	0.053	0.043	0.060	0.049	0.052	0.050	0.045	0.044
50	0.039	0.039	0.046	0.052	0.049	0.052	0.042	0.042	0.047
75	0.045	0.053	0.053	0.049	0.044	0.044	0.052	0.050	0.051
100	0.045	0.047	0.049	0.057	0.055	0.065	0.049	0.047	0.041
125	0.054	0.051	0.050	0.041	0.048	0.040	0.040	0.045	0.057

$(\alpha, \beta)=(4, 0.114)$ and a 1% increase in each of the 5 least prevalent species under the alternative.

	$n_1=n_2=5$			$n_1=n_2=10$			$n_1=n_2=15$		
m	Lambda	Squared	LRT	Lambda	Squared	LRT	Lambda	Squared	LRT
25	0.087	0.082	0.104	0.131	0.108	0.184	0.152	0.125	0.327
50	0.123	0.106	0.187	0.300	0.187	0.508	0.420	0.246	0.798
75	0.181	0.111	0.304	0.449	0.243	0.785	0.766	0.434	0.977
100	0.272	0.154	0.477	0.659	0.351	0.938	0.923	0.620	0.997
125	0.375	0.187	0.632	0.822	0.477	0.990	0.982	0.790	1.0

Figure 5 displays power curves for the likelihood ratio tests at $n=5, 10, 15$, where the alternative hypotheses are a) $H_1: (\alpha, \beta)=(2, 0.077), (2.5, 0.082)$ and b) $H_1: (\alpha, \beta)=(2, 0.077)$ and 1% increase in each of the 5 least prevalent species, respectively. Based on these pictures, if the number of species is 100, then the sample size should be at least 10, and colony size should be at least 75 to get a good power.

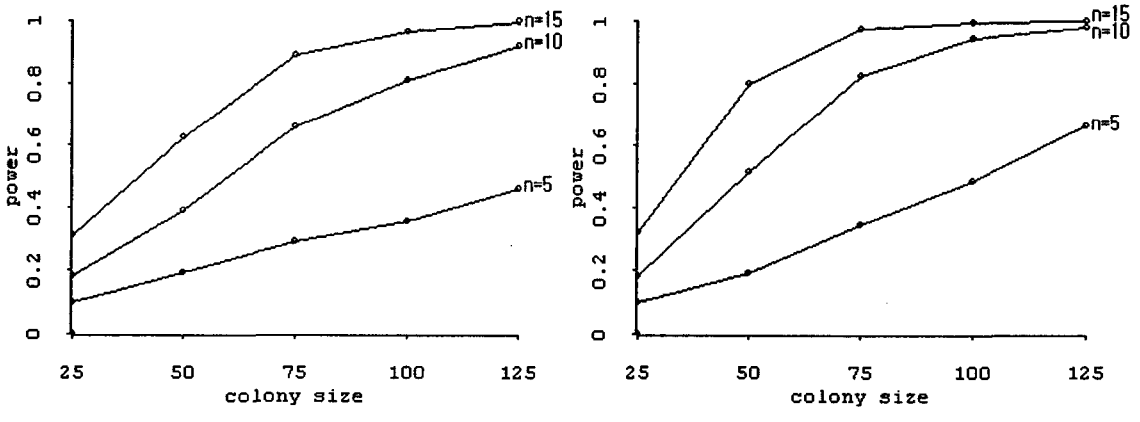


Figure 5. Power curves for the likelihood ratio test: $H_0: (\alpha, \beta)=(2, 0.077)$ for each group.

a) $H_1: (\alpha, \beta)=(2, 0.077), (2.5, 0.082)$

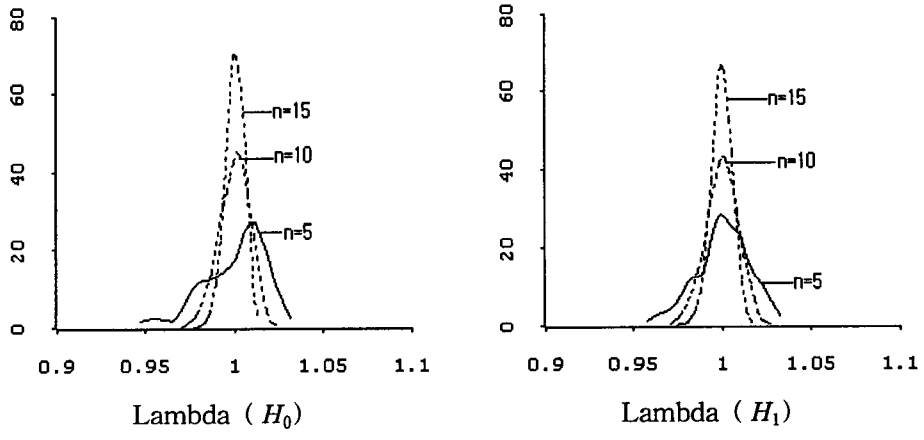
b) $H_1: (\alpha, \beta)=(2, 0.077)$ and 1% increase in each of the 5 least prevalent species

4. The Permutational Distribution

Permutation tests have a wide range of applications, and it requires relatively weak assumptions, e.g., that the underlying distributions are symmetric and the alternatives are shifts in value(Good(1994)). It is a computational intensive resampling method. With nowadays computing power, it is widely applicable to many moderate size data sets. The detail of the design and extent of the simulation are given in the Appendix.

The computations of the null hypothesis significance level and powers are based on 1000 Mote Carlo runs. One of these 1000 Monte Carlo runs was illustrated in Figure 6 as an example. All the other runs were similar. For this run, the permutational distributions of the λ-statistics using (1), (3), and likelihood ratio test statistics at $n_1 = n_2 = 5, 10, 15$ were calculated by performing 1000 Monte Carlo runs under the null and under the alternative, separately. The null hypothesis is $(\alpha, \beta) = (2, 0.077)$ in group 1 and 2, and the alternative hypothesis is $(\alpha, \beta) = (2, 0.077)$ in group 1 and a 1% increase in each of the 5 least prevalent species in group 2. The pictures in Figure 6 was made at $m = 100$ using kernel density function.

As sample size increases, the permutation test distributions of three tests become symmetric. For λ-statistics using (1) and (3) the permutation test distributions under the null are similar to those under the alternative. But, for the likelihood ratio test statistic, the distributions under the alternative are shifted to the right compared to the distributions under the null.



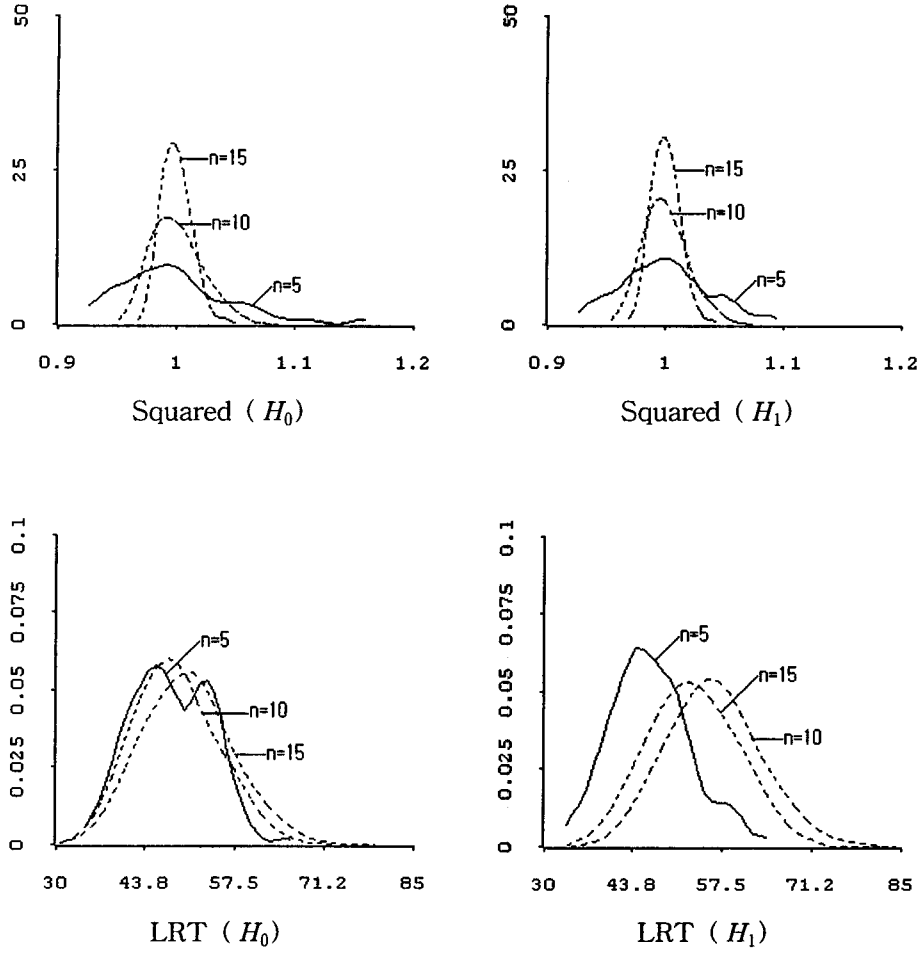


Figure 6. Examples of the permutation test distribution of three tests under the null and under the alternative.

H_0: $(\alpha, \beta) = (2, 0.077)$

H_1: $(\alpha, \beta) = (2, 0.077)$ in group 1 and a 1% increase in each of the 5 least prevalent species in group 2

5. Concluding Remarks

This paper shows that the permutation test using the likelihood ratio may be more powerful than λ-test in testing the change of micro-flora composition under a likely pathological condition. Tables are provided for the sample size determination under this condition.

Confirming the change in micro-flora is the first step towards the identification of the

bacteria that cause the disease. When the null hypothesis is rejected, an obvious choice for the responsible bacteria is to select bacteria that have the biggest differences in two groups. However, a large number of species makes the threshold for this type of multiple comparisons difficult. It seems that these bacteria can be confirmed only by forming a more specific hypothesis in a subsequent study.

The simulation program is written in FORTRAN. This program, which uses double precision, runs interactively. We can provide input parameters, which are the parameters of the gamma distribution (α, β), the sample size $(n_1 = n_2)$, the number of possible species K, the number of colonies m, the level of significance α, and the number of iterations. The gamma random numbers are generated from the uniform random number generator described in L'Ecuyer(1988). A FORTRAN program implementing these tests is available from the authors.

Appendix : The Simulation Study

We used repeated simulated samples to get the significance level and power. For instance, to compute the power, we generated data sets for the two groups, $\mathbf{X}_i = \{x_{i1}, x_{i2}, \ldots, x_{iK}\}$, $i = 1, \ldots, n_1$, and $\mathbf{Y}_j = \{y_{j1}, y_{j2}, \ldots, y_{jK}\}$, $j = 1, \ldots, n_2$, according to the multinomial probabilities $\{p_j\}$, and $\{q_j\}$ defined by (5). From this data set, we computed the λ-statistics using distance (1) and (3), and also computed the likelihood ratio test statistic (4). Let the values of these statistics be $\lambda_{(1)}$, $\lambda_{(3)}$, and Λ, respectively. The next step was to find the distributions of these statistics using random permutation. More specifically, we partitioned the combined n_1 and n_2 data randomly into two groups of n_1 and n_2 and computed three statistics for this partition. This process was repeated a large number of times (1,000 was used). Let $\lambda^*_{(1)}$ be the $\alpha 100\%$ quantile and $\lambda^*_{(3)}$, Λ^* be the $(1-\alpha)100\%$ quantile from the sample distributions. Our decision on which hypothesis to accept was based on comparing the data values with the quantiles of $\lambda^*_{(1)}$, $\lambda^*_{(3)}$, and Λ^*. We repeated this whole process 1,000 times, and kept a record of the number of null hypothesis rejections. The powers were the average number of rejections, and the null hypothesis significance levels were obtained under the null distribution $\mathbf{F} = \mathbf{G}$.

References

[1] Fisher, R. A., Corbet, A. S. and Williams, C. B.(1943). The relation between the number of species and the number of individuals in a random sample of an animal population, *Journal of Animal Ecology*, vol. 12, 42-58.

[2] Good, I. J.(1982). An index of separateness of clusters and a permutation test for its statistical significance, *Journal of Statistical Computation and Simulation*, 15, 81-84.

[3] Good, P.(1994). *Permutation tests*, Springer-Verlag, New York.

[4] Hoeffding, W.(1952). The large-sample power of tests based on permutations of observations, *Annals of Mathematical Statistics*, 23, 169-192.

[5] Johnson, N. L. and Kotz, S.(1970). *Continuous univariate distribution-1*, Wiley, New York.

[6] L'Ecuyer, P.(1988). Efficient and portable combined random number generators, *Communications of the ACM,* vol. 31, 742-749.

[7] Moore, W. E. C., Holdeman, L. V., Smibert, R. M., Good, I. J., Burmeister, J. A., Palcanis, K. G. and Ranney, R. R.(1982). Bacteriology of experimental gingivitis in young adult humans, *Infection and Immunity*, 38, 651-667.

[8] Moore, W. E. C., Moore, L. H., Ranney, R. R., Smibert, R. M., Burmeister, J. A. and Schenkein, H. A.(1991). The microflora of periodontal sites showing active destructive progression, *Journal of Clinical Periodontology,* 18, 729-739.

[9] Socransky, S. S., Tanner, A. C. R. and Goodson, J. M.(1981). Patterns of subgingival microbial colonization, *J. Dental Research*, 60, 486.

[Received February 2003, Accepted March 2003]