Journal of Mathematical Research & Exposition
Jan., 2009, Vol. 29, No. 1, pp. 113–123
DOI:10.3770/j.issn:1000-341X.2009.01.015
Http://jmre.dlut.edu.cn

A Modified Chi-Squared Goodness-of-Fit Test

DAI Jia Jia[1,2], *YANG Ai Jun*[1]
(1. College of Applied Sciences, Beijing University of Technology, Beijing 100124, China;
2. College of Science, Guizhou University, Guizhou 550025, China)
(E-mail: jjdai@emails.bjut.edu.cn)

Abstract In goodness-of-fit tests, Pearson's chi-squared test is one of most widely used tools of formal statistical analysis. However, Pearson's chi-squared test depends on the partition of the sample space. Different constructions of the partition of the sample space may lead to different conclusions. Based on an equiprobable partition of sample space, a modified chi-squared test is proposed. A method for constructing the modified chi-squared test is proposed. As an application, the proposed test is used to test whether vectorial data come from an uniformity distribution defined on the hypersphere. Some simulation studies show that the modified chi-squared test against different alternative is robust.

Keywords Pearson's chi-squared test; Von Mises-Fisher distribution; Watson distribution; vectorial data.

Document code A
MR(2000) Subject Classification 62G10; 62H11
Chinese Library Classification O212.1

1. Introduction

Specification or stochastic modeling of data is an important step in statistical analysis of data. It was Karl Pearson who first recognized the problem and proposed a criterion to examine whether the observed data support a given specification. It is called chi-squared goodness-of-fit test, which motivates research in testing of hypotheses and estimating of unknown parameters.

Pearson's[1] chi-squared test statistics

$$X_p^2=\sum(\text{observed-expected})^2/\text{expected}$$

is essentially an omnibus test because it is sensitive to a wide variety of different ways in which the data can be different to the hypothesized distribution.

A random sample $X_1,\ldots,X_n$ of size n comes from a population with completely specified cumulative distribution function $F(x)$, against a general alternative not $F(x)$. Let the sample space be broken into m classes (or cells). And let O_j be the number of observation from the sample that falls into the jth class, where $n=\sum_{j=1}^m O_j$. Let E_j be frequency of falling into the

Received date: 2008-03-26; **Accepted date**: 2008-04-18
Foundation item: the Natural Science Foundation of Beijing (No. 1062001);Academic Human Resources Development in Institutions of Higher Learning Under the Jurisdiction of Beijing Municipality(No. 05006011200702).

jth class when $F(x)$ holds. Then the Pearson's chi-squared test statistics is written as

$$X_p^2 = \sum_{j=1}^{m} (O_j - E_j)^2 / E_j. \tag{1}$$

It has asymptotic chi-squared distribution with $m-1$ degrees of freedom.

If a chi-squared test is to be used, then classes must be constructed. For construction of classes, there are several methods[2−9]. For example, Mann and Wald[2] recommended that the cells be chosen to have equal probabilities under the hypothesized distribution F. However, the sequence of chi-squared test will be determined by the number of classes m that is most critical data. Kempthorne[10] pointed out, different conclusions may be reached if different constructions are used. Therefore, whether the data are categorized or not, the statistician who wants to use Pearson's chi-squared has to choose the number and width of classes.

While Pearson's chi-squared test involves choosing the number and the width of classes, and we are led to ask whether there is a modified chi-squared test, which leads to the same conclusion under different constructions of classes. That is, a modified chi-squared test may eliminate the effects of choosing the number and width of classes. In this article, we propose a modified chi-squared test based on an equiprobable partition of sample space. The basic idea is to maximize the Pearson's chi-squared statistics when the sample space is partitioned equiprobably. Once we cannot reject the null hypothesis against the general alternative in the extreme case, we may conclude that the null hypothesis that the sample comes from $F(x)$ is not rejected at the $100a\%$ level of significance.

In this article, we just focus on the modified chi-squared test in the interval $[0, 1]$. There is an important reason. It is well-known that the tests of empirical distribution function are not dependent on the population cumulative distribution function. Let $X_1, \ldots, X_n$ be a random sample from cumulative distribution function $F_0(\cdot)$ which is continuous. Consider the simple hypothesis

$$H_0 : F = F_0 \leftrightarrow H_1 : F \neq F_0.$$

Set

$$U_i = F_0(X_i) \quad i = 1, \ldots, n.$$

Then U_i, $i = 1, \ldots, n$ is a uniformity in the interval $[0, 1]$. The problem that to test whether the random samples $X_1, \ldots, X_n$ come from $F_0(\cdot)$ is substituted by the one to test whether $U_1, \ldots, U_n$ come from uniformity in the interval $[0, 1]$. Here we consider a construction of the modified chi-squared test in the case, in which null hypothesis that we are sampling is from the uniform distribution in the interval $[0, 1]$.

Furthermore, the modified chi-squared test is used to test whether the vectorial data come from the uniformity defined on the hypersphere S_{p-1}, where $S_{p-1} = \{\mathbf{X} \in R^p : \mathbf{X}'\mathbf{X} = 1\}$. In this article, for some dimensions of the hypersphere in some cases, we compare the power of the uniformity tests against the hypothesis of a Von Mises-Fisher distribution or a Watson distribution defined on the hypersphere. These tests include the proposed modified chi-squared test, Rayleigh, Ajne, Giné, and Bingham tests. The involved tests and distributions are presented in the appendix. It is true that certain tests tend to perform better than others in certain types of situation. Such as, the Rayleigh and Ajne tests against the hypothesis of a Von Mises-Fisher distribution have the highest empirical power, but they are invalid when alternative is a Watson distribution defined on the hypersphere. However, the simulation results demonstrate that the modified chi-squared test against different alternatives is valid. In contrast to other uniformity tests, the modified chi-squared test is robust.

The article is organized as follows. The construction of the modified chi-squared test statistic is considered in Section 2. In Section 3, the empirical power of these uniformity tests for some dimensions in some cases are compared. Some problems are proposed in Section 4.

2. Construction of statistic

In this section we consider the construction of the modified chi-squared test in interval $[0, 1]$. Note that the test we propose can accommodate the sample space with any equiprobable partition.

Let the interval $[0, 1]$ be partitioned equiprobably into m classes or cells. Our basic idea is to maximize the Pearson's chi-squared statistics. We propose a construction procedure for the modified chi-squared test statistic. Let k be a positive proper factor of the number m.

Step 1. For a fixed k, m classes are combined arbitrarily into new disjoint k groups, where each group includes b old classes, $b \geq 2$. It is obvious that a combination is a new partition of the sample space. For a given combination, calcualte corresponding chi-squared test statistic using (1), where $E_j = nb/m$.

Step 2. For given k, maximize corresponding values of the chi-square test statistic, which are relative to all combinations.

Step 3. For different k which may be all proper factors of the number m, repeat Steps 1 and 2. A fixed k has a corresponding maximum value of chi-square test statistic based on Step 2. Finally, maximize those maximum values.

In fact, the Pearson's chi-squared test statistic is maximized twice. Next, we show the construction of the modified chi-squared test statistic in detail.

Let $X_1, \ldots, X_n$ be a random sample from uniformity in the interval $[0, 1]$, which is partitioned equiprobably into m classes or cells. The interval $[0, 1]$ is the union of mutually disjoint sets $A_1, \ldots, A_m$. We call $A_1, \ldots, A_m$ a first-partition of the sample space. Further, sets $A_1, \ldots, A_m$ are combined into a new disjoint sets $T_1, \ldots, T_k$, where T_1 is the union of b ($b \geq 2$) sets which are chosen from sets $A_1, \ldots, A_m$, and T_j ($j = 2, \ldots, k$) is the union of b sets chosen from the rest sets $\cup_{i=1}^{m} A_i \setminus \cup_{l=1}^{j-1} T_l$, $j = 2, \ldots, k$, where $m = kb$. Then $T_1, \ldots, T_k$ is also a partition of the sample space. We refer to the partition $T_1, \ldots, T_k$ as a second-partition of the sample space. There are $m!/(b!^k k!)$ second-partition. And let $Y_1, \ldots, Y_k$ denote the frequencies with which the sample is, respectively, an element of $T_1, \ldots, T_k$. Then the joint probability density function $Y_1, \ldots, Y_k$ is the multinomial probability density function with parameter $n, 1/k, \ldots, 1/k$.

We consider the simple hypothesis (concerning above multinomial probability density function)

$$H_0 : P = P_0 = (\frac{b}{m}, \ldots, \frac{b}{m})'.$$

It is desired to test H_0 against all alternatives.

Now, the modified chi-squared test statistic is defined by

$$X^2_{\max} = \max_k \max_T \sum_{j=1}^{k} \frac{(Y_j - np_j)^2}{np_j}, \tag{2}$$

where k is a positive proper factor of the number m, $p_j = b/m$, Y_j is the frequency of the sample falls into T_j, $j = 1, \ldots, k$, and T is a set of all of second- partitions when k is given.

In (2), for given k, we have to calculate $m!/(b!^k k!)$ Pearson's chi-squared test statistics. To improve the efficiency for counting $X^2_{\max}$, we obtain a simple algorithm. In fact, note that $p_j = b/m = 1/k$, and extend

$$\sum_{j=1}^{k} \frac{(Y_j - np_j)^2}{np_j}.$$

It is immediately seen that, for given k,

$$\max_T \sum_{j=1}^{k} \frac{(Y_j - np_j)^2}{np_j} = C \max_T \sum_{j=1}^{k} Y_j^2 - n,$$

where $C = k/n$. Thus, we just need to maximize $\sum_{j=1}^{k} Y_j^2$. That is, how to choose a second-partition, such that $\sum_{j=1}^{k} Y_j^2$ reached its maximum.

It is easy to obtain the maximum of $\sum_{j=1}^{k} Y_j^2$. According to the frequencies of $A_1, \ldots, A_m$, the $A_1, \ldots, A_m$ are arranged in decreasing order, and are denoted by $A_{1:m}, \ldots, A_{m:m}$. We choose successively b $A_{i:m}$, $i = 1, \ldots, m$ to combine a T_j. Namely, T_1^* is a union of the first b cells $A_{i:m}, i = 1, \ldots, b$. T_2^* is a union of the second b cells $A_{i:m}$, $i = (b+1), \ldots, 2b$, and so on. $T_1^*, \ldots, T_k^*$ is a second-partition of the sample space. Let Y_j^* be the frequency of the sample falls into T_j^*, $j = 1, \ldots, k$. Then

$$\max_T \sum_{j=1}^{k} \frac{(Y_j - np_j)^2}{np_j} = C \sum_{j=1}^{k} Y_j^{*2} - n.$$

Hence, (2) is written as

$$X^2_{\max} = \max_k \big[\frac{k}{n} \sum_{j=1}^{k} Y_j^{*2} - n\big], \tag{3}$$

where Y_j^* is the frequency of the sample falls into T_j^*, $j = 1, \ldots, k$.

In our simulation studies the modified chi-squared test is constructed based on (3).

3. Application to vectorial data

In Section 2, the construction of the modified chi-squared test just requires that the sample space is partitioned equiprobably. It is easy to think that the modified chi-squared test statistic may be used to test whether vectorial data come from a uniformity defined on the hypersphere.

In this section we first compare the power of tests including the modified chi-squared test, Ajne, Bingham, Giné, Rayleigh tests. In some cases, the 95th percentiles of the modified chi-squared test statistic under uniformity are presented. All simulation studies were conducted using R.

First, we compare the power of the modified chi-squared test and Pearson's chi-squared test in some cases. See Figure 1. Here we consider the partial alternative

$$H'_{1n} : p = p_0 + \frac{\gamma \cdot \delta}{\sqrt{n}},$$

where $p_0 = (\frac{b}{m}, \ldots, \frac{b}{m})'$, and $\delta = (\delta_1, \ldots, \delta_k)'$, with $\sum_{j=1}^{k} \delta_j = 0$, $0 \leq \gamma \leq 1$. The inversion method is used to generate random sample from the partial alternative. In Figure 1, the real line represents the empirical power of the modified chi-squared test. And the dashed line is the empirical power of Pearson chi-squared test. From Figure 1, in some cases the modified chi-squared test is more powerful than Pearson's chi-squared test.

Next, we use the modified chi-squared test to test whether vectorial data come from a uniformity defined on the hypersphere.

In fact, if vectorial data come from a uniformity defined on the sphere S_{p-1}, then vectorial data is also uniformity in each quadrant. We may consider that p quadrants divide the hypersphere into 2^p equiprobable fields. Therefore, the modified chi-squared test may be used to test whether the data come from the uniformity defined on the sphere. We have to deal with how to divide the sphere S_{p-1} into an equiprobable partition. The simplest way is that a quadrant is a cell. For example, if there is a 10-dimension unit sphere, we have to consider the modified chi-squared test in $2^{10} = 1024$ quadrants. To reduce the calculation quantity involved during the modified chi-squared test statistic calculations, we propose a method to construct the modified chi-squared test on the hypersphere.

The method is as follows. Let $X_1, \ldots, X_n$ be a p-dimension random sample of size n, denote $X = (X_1, \ldots, X_n)$. Translate the matrix X into a $(0, 1)$ matrix Y.

Define

$$y_{ij} = \begin{cases} 1, & x_{ij} \geq 0, \\ 0, & x_{ij} < 0. \end{cases}$$

Then $Y = (Y_1, \ldots, Y_n)$, where $Y_i = (y_{1i}, \ldots, y_{pi})'$, $i = 1, \ldots, n$. Considering the sum of nonzero numbers of each axis, denote s_i which is defined by

$$s_i = \sum_{i=1}^{n} y_{ij}, \quad i = 1, \ldots, p.$$

Indeed, arrange $s_i\,(i = 1, \ldots, p)$ in non-increasing sort. Then we may choose the first m axes which have corresponding first m maximum s_i, $i = 1, \ldots, m$, respectively. Then 2^m quadrants is a first-partition. And (3) is can be used.

Based on (3), we determine the 95th percentiles of the modified chi-squared test statistic in some cases. While we do not know the exact distribution of the modified chi-squared test statistics under uniformity on the sphere S_{p-1}, in some cases we determinate by generating 10000 replicates of the statistics under uniformity. See Table 1. Here we use the method proposed by

Sibuya[11] to simulate the uniformity defined on the sphere S_{p-1}.

We know that once several tests have been proposed, they are usually compared on the basis of power in simulation studies. Some authors studied empirical power of uniformity tests on the sphere. Diggle et al.[12] compared the power of uniformity tests proposed by Beran and Giné in some particular cases. For some dimensions of sphere in some cases, the power of Bingham and Giné tests of uniformity defined on S_{p-1} against a Bingham population or a mixture of Bingham population were compared by Figueiredo[13]. And Figueiredo[14] also studied the power of uniformity tests against the hypothesis of a Von Mises-Fisher distribution defined on the hypersphere in some special cases.

$p\backslash n$	10	20	50	80	120	150	200
3	11.60000	12.40000	12.40000	12.70000	12.46667	12.50667	12.64000
4	20.40000	22.40000	23.60000	23.80000	23.73333	23.86667	23.68000
5	44.40000	44.00000	43.44000	43.60000	44.00000	44.34667	44.16000
6	79.60000	79.20000	80.56000	80.80000	81.60000	81.68000	81.28000
7	130.8000	152.8000	152.2400	153.6000	153.0667	152.9333	153.2800
8	271.6000	287.2000	293.0400	294.4000	291.7333	292.0267	292.8000
9	553.2000	517.6000	554.1600	560.0000	562.6667	563.3867	562.8800
10	1014.000	1055.200	1076.400	1097.600	1100.267	1099.280	1095.360

Table 1 95th percentiles of the modified chi-squared test under uniformity

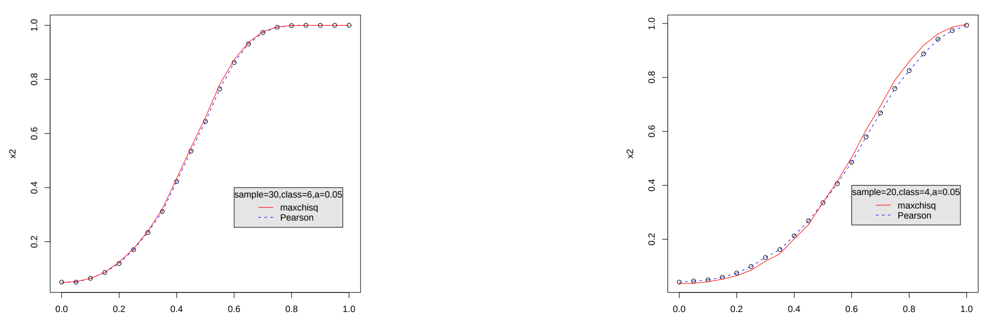

Figure 1 Empirical power of modified chi-squared(real line) and Pearson chi-squared test(dashed line)

Next, we investigate the power of tests under alternative Von Mises-Fisher distribution. Suppose the direction mean $\mu = (0, 0, \ldots, 1)$ and concentration parameter $\xi = 0, 1, 2, \ldots, 10, 15$. For simulation of the Von Mises-Fisher distribution defined on S_{p-1}, the method proposed by Wood[15] is used. We calculate the empirical power of tests including Rayleigh, Ajne, Bingham, Giné and modified chi-squared test. See Table 2, Table 3 and Figure 2. Table 2, Table 3 and Figure 2 depict the simulation results for the following cases: $p = 4(n = 20)$, $p = 6(n = 80)$, $p = 10(n = 40, 120)$. In Figure 2, the real line is the empirical power of the modified chi-squared test statistic. The upper lines are empirical power of the Rayleigh and Ajne tests which have identical power for the cases analyzed here. The lower lines represent the empirical power of Giné and Bingham tests, respectively. Compared with the results in Table 2, Table 3 and Figure 2, we can see that the empirical power of the modified chi-squared test is higher than those of Giné and Bingham tests, and is close to those of Rayleigh and Ajne tests.

ξ	0	1	2	3	4	5	6	7	8
Modified chi-squared	0.050	0.442	0.984	1.000	1.000	1.000	1.000	1.000	1.000
Rayleigh	0.036	0.758	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Ajne	0.036	0.764	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Giné	0.056	0.062	0.364	0.940	1.000	1.000	1.000	1.000	1.000
Bingham	0.058	0.064	0.374	0.948	1.000	1.000	1.000	1.000	1.000

Table 2 Empirical power of the tests against Von Mises-Fisher distribution for $p = 6(n = 80)$

ξ	0	0.2	0.5	1	1.5	2.5	3	4	5
Modified chi-squared	0.052	0.064	0.106	0.336	0.766	0.966	1.000	1.000	1.000
Rayleigh	0.036	0.060	0.152	0.660	0.960	1.000	1.000	1.000	1.000
Ajne	0.036	0.058	0.146	0.650	0.960	1.000	1.000	1.000	1.000
Giné	0.040	0.036	0.068	0.066	0.080	0.364	0.73	0.996	1.000
Bingham	0.058	0.038	0.068	0.068	0.082	0.374	0.73	0.994	1.000

Table 3 Empirical power of the tests against Von Mises-Fisher distribution for $p = 10(n = 120)$

Now, we consider Watson distribution as an alternative. We examine the empirical power of tests statistics presented in this article. See Table 4 and Figure 3. For the simulation of Watson distribution, we have used the acceptance-rejection method. To save space, we present only the simulation results with directional parameter $\mu = (0, \ldots, 0, (1-\cos^2\theta)^{1/2}, \cos\theta)$, where $\theta = \pi/4$. Results for other directional parameters are similar.

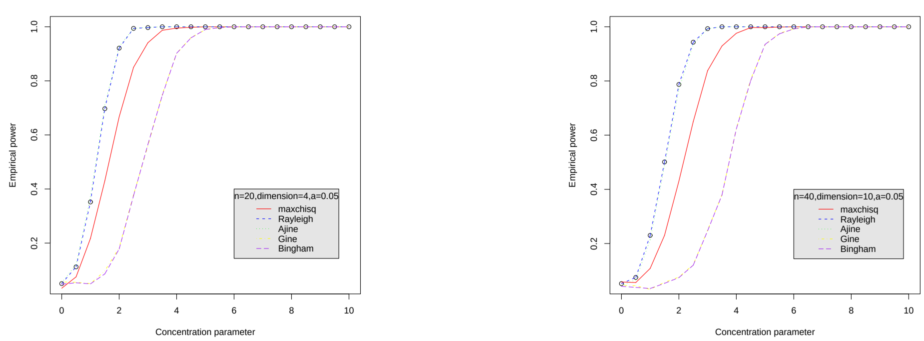

Figure 2 Alternative: Von Mises-Fisher distribution, the real line is the empirical power of the modified chi-squared test, the upper broken line are the empirical power of Rayleigh and Ajne tests, the lower broken line are the empirical power of Giné and Bingham tests.

In Figure 3, the notation is the same as that given in Figure 2. From Figure 3, we can see that the upper lines represent the empirical power of Giné and Bingham tests. On the other hand, Rayleigh and Ajne tests have the lowest empirical power, and even are invalid, contrasting they have the highest empirical power when alternative is Von Mises-Fisher distribution. Therefore, the empirical power of the modified chi-squared test is valid. Although the empirical power of modified chi-squared test is lower than those of Giné and Bingham tests, the difference between them is small in most cases. For example, in Table 4, for $\xi = 1$, the power of modified chi-squared test is 0.058, and the power of Bingham test is 0.100.

ξ	0	1	2	3	4	5	6	8	9
Modified chi-squared	0.044	0.058	0.096	0.256	0.610	0.906	0.994	1.000	1.000
Rayleigh	0.034	0.064	0.052	0.052	0.070	0.090	0.080	0.086	0.118
Ajne	0.040	0.066	0.054	0.054	0.074	0.098	0.088	0.092	0.130
Giné	0.052	0.106	0.406	0.904	1.000	1.000	1.000	1.000	1.000
Bingham	0.052	0.100	0.408	0.908	1.000	1.000	1.000	1.000	1.000

Table 4 Empirical power of the tests against Watson distribution for $p = 6(n = 50)$

In summary, there is no such thing as uniformly "best" test. Whereas, certain tests tend to perform better than others in certain types of simulations. The simulation studies demonstrate that the modified chi-squared test is more robust than other uniformity tests on the hypersphere. And the modified chi-squared test is valid under different alternatives.

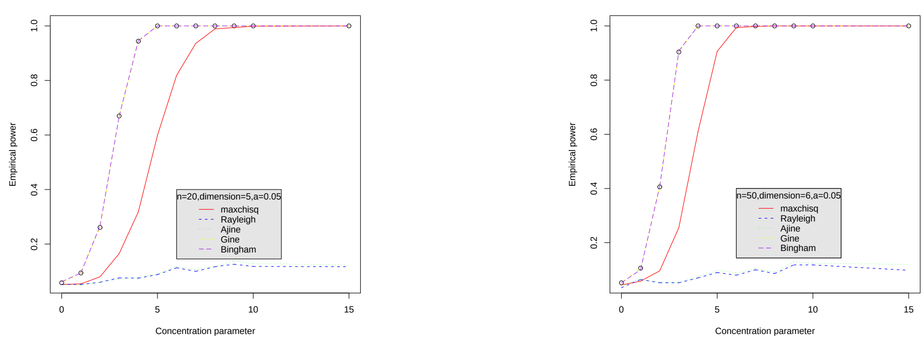

Figure 3 Alternative: Watson distribution, the real line is the empirical power of the modified chi-squared test, the upper broken line are the empirical power of Giné and Bingham tests, the lower broken line are the empirical power of Rayleigh and Ajne tests.

4. Discussion

In this article we proposed and constructed the modified chi-squared test. We further used it to test whether vectorial data come from a uniformity distribution defined on the hypersphere. And we compared the empirical power of five tests presented in the article for some cases. Although the simulation studies show that the modified chi-squared test statistic against different alternatives is more stable, it may lead to lower empirical power in some case (see Table 4). Thus it is interesting to improve empirical power of the modified chi-squared test.

In order to improve the empirical power of the modified chi-squared test, we may consider a translation of coordinate. To address this issue, let X be a $p\times n$ matrix, where $X=(X_1,\dots,X_n)$, X_i, $i=1,\dots,n$ denote random variables defined on the sphere S_{p-1}. If there exists a $p\times p$ orthogonal matrix, such that $Y=PX$, then we may construct the modified chi-squared test based on Y. The problem how to find out an orthogonal matrix P may be considered.

In addition, in the article we have not discussed the exact distribution or asymptotic properties of the modified chi-squared test. The modified chi-squared test may also be used to test the distribution of multi-dimension data. These research topics are beyond the scope of this article. Further research is needed.

5. Appendix

The Von Mises-Fisher distribution (also known as Langevin distribution) defined on S_{p-1} is usually denoted by $M_p(\mu,\xi)$. Its probability density function with respect to the uniform distribution is given by

$$f(\mathbf{x};\mu,\xi)=(\frac{\xi}{2})^{p/2-1}\frac{1}{\Gamma(p/2)I_{p/2-1}(\xi)}\exp(\xi\mu'\mathbf{x})$$
$$\mathbf{x}\in S_{p-1},\mu\in S_{p-1},\quad \xi\ge 0$$

where I_u denotes the modified Bessel function of the first kind and order ν defined by

$$I_\nu(\eta)=\frac{1}{2\pi}\int_0^{2\pi}\cos(\nu t)e^{\eta\cos t}\mathrm{d}t.$$

Parameters ξ and μ are the concentration and mean direction parameter, respectively. For $\xi=0$, the Von Mises-Fisher distribution reduces to the uniformity on the hypersphere.

Watson distribution is denoted by $W_p(\mu,\xi)$ and has a probability density function given by

$$f(\mathbf{x};\mu,\xi)=\{{}_1F_1(\frac{1}{2},\frac{p}{2},\xi)\}^{-1}\exp(\xi(\mu'\mathbf{x})^2)$$
$$\mathbf{x}\in S_{p-1},\mu\in S_{p-1},\quad \xi\in R$$

where ξ is concentration parameter and μ is directional parameter. The reciprocal of the confluent hypergeometric function ${}_1F_1(\cdot)$ is the normalizing constant and is defined by

$${}_1F_1(\frac{1}{2},\frac{p}{2},\xi)=\frac{\Gamma(\frac{p}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})}\int_0^1\exp(\xi s)s^{-1/2}(1-s)^{(p-3)/2}\mathrm{d}s.$$

For $\xi=0$, Watson distribution is uniformity.

Suppose $X_1,\dots,X_n$ is a p-dimension random sample, where $X_i=(X_{i1},\dots,X_{ip})'$ and $\sum_{j=1}^{p}(X_{ij})^2=1$, $i=1,\dots,n$.

1) Ajne test

The Ajne statistic is defined by

$$A=\frac{n}{4}-\frac{1}{n\pi}\sum_{i<j}\varphi_{ij},$$

where $\varphi_{ij}=\cos^{-1}(\sum_{k=1}^{p}X_{ik}X_{jk})$, $1\le i<j\le n$.

2) Bingham test

The Bingham statistic is defined by

$$B = \frac{p(p+2)}{2n}\sum_{i=1}^{p}(\lambda_i - \frac{1}{p})^{1/2},$$

where λ_i, $i = 1, \ldots, p$ are the eigenvalues of T, $T = \sum_{i=1}^{n}(X_i X_i' - p^{-1}\mathbf{I})$, $\mathbf{I}$ is a $p \times p$ identity matrix.

3) Giné test

The Giné statistic is defined by

$$G = \frac{n}{2} - \frac{p-1}{2n}[\frac{\Gamma((p-1)/2)}{\Gamma(p/2)}]^2 \sum_{i<j} \sin\varphi_{ij},$$

where $\varphi_{ij} = \cos^{-1}(\sum_{k=1}^{p} X_{ik}X_{jk})$, $1 \le i < j \le n$, is the smaller one of the two angles between X_i and X_j.

4) Rayleigh test

Let R be the length of the resultant vector defined by

$$R = \{(\sum_{i=1}^{n} X_{i1})^2 + (\sum_{i=1}^{n} X_{i2})^2 + \cdots + (\sum_{i=1}^{n} X_{ip})^2\}^{1/2}$$

and $\bar{R}$ the mean resultant length defined by

$$\bar{R} = \frac{R}{n}.$$

Then Rayleigh statistic is represented as

$$np\bar{R}^2.$$

Acknowledgements The authors cordially thank the Associate Editor and Reviewers for their constructive comments which lead to improvement of the manuscript. They are also very grateful to Prof. Adelaide Figueiredo for his help.

References

[1] PEARSON K. *On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling* [J]. Philosophical Magazine, 1900, **50**: 157–172.

[2] MANN H B, WALD A. *On the choice of the number of class intervals in the application of the Chi-square test*[J]. Ann. Math. Statistics, 1942, **13**: 306–317.

[3] ROSCOE J T, BYARS J A. *An investigation of the restraints with respect to sample size commonly imposed on the use of the Chi-square statistic*[J]. J. Amer. Statist. Assoc., 1971, **66**: 755–759.

[4] Larntz K. *Small-sample comparisons of exact levels for chi-squared goodness-of-fit statistics* [J]. J. Amer. Statist. Assoc., 1978, **73**: 253–263.

[5] Koehler K J, Larntz K. *An empirical investigation of goodness-of-fit statistics for sparse multinomials*[J]. J. Amer. Statist. Assoc., 1980, **75**: 336–344.

[6] LAWAL H B. *Tables of percentage points of Pearson's goodness-of-fit statistic for use with small expectations* [J]. J. Roy. Statist. Soc. Ser. C, 1980, **29**(3): 292–298.

[7] KALLENBERG W C M. *On moderate and large deviations in multinomial distribution*[J]. Ann. Statist., 1985, **13**(4): 1554–1580.

[8] RAYNER J C W, BEST D J. *The choice of class probabilities and number of classes for the simple X^2 goodness of fit test*[J]. Sankhya Ser. B, 1982, **44**: 28–38.

[9] RAYNER J C W, BEST D J, Dodds K G.*The construction of the simple X^2 and Neyman smooth goodness of fit tests*[J]. Statistica Neerlandica, 1985, **39**: 35–50.

[10] KEMPTHORNE O. *The Classical Problem of Inference Goodness of Fit* [M]. California Press, Berkeley, Calif, 1967.

[11] SIBUYA M. *A method for generating uniformly distributed points on n-dimensional spheres* [J]. Ann. Inst. Statist. Math., 1962, **14**: 81–85.

[12] DIGGLE P J, FISHER N I, LEE A J. *A comparison of tests of uniformity for spherical data* [J]. Austral. J. Statist., 1985, **27**(1): 53–59.

[13] FIGUEIREDO A, Gomes P. *Power of tests of uniformity defined on the hypersphere* [J]. Comm. Statist. Simulation Comput., 2003, **32**(1): 87–94.

[14] FIGUEIREDO A. *Comparison of tests of uniformity defined on the hypersphere* [J]. Statist. Probab. Lett., 2007, **77**(3): 329–334.

[15] WOOD A. *Simulation of the Von Mises-Fisher distribution* [J]. Comm. Statist. Simulation Comput, 1994, **23**(1): 157–164.

[16] YANG Zhenhai. *Goodness of Fit Tests* [M]. Hefei: Anhui Education Publishing, 1994. (in Chinese)

[17] COX D R. *Karl Pearson and the Chi-Squared Test* [R]. Goodness-of-Fit Tests and Model Validity. Berlin: Boston, Basel, 2002, 3–8.