DOI: 10.15415/mjis.2012.11001

Simultaneous Testing For the Goodness of Fit to Two or More Samples

Narinder Kumar
Gobind P. Mehta

Department of Statistics, Panjab University, Chandigarh
Email: nkumar@pu.ac.in

Abstract

In this paper we have considered the problem to test for the simultaneous goodness of fit of an absolutely continuous distribution function to many samples. The proposed test is seen to have many desirable properties.

Keywords: Shift function; distribution free procedure, nonparametric hypothesis.

1. INTRODUCTION

Testing for the goodness of fit of a probability distribution is a very well-known problem. Among others, the tests of Kolmogorov and that of Cramer von-Mises, for testing the goodness of fit of a probability distribution to single sample, are available for such a problem. In this paper, we consider the simultaneous testing for the goodness of fit of an absolutely continuous probability to k $(k \geq 2)$ samples. Let us denote by $F_1, \ldots, F_k$, the k absolutely continuous distribution functions and assume that a sample of size n is available from each of these distributions. Let F_0 be a known absolutely continuous distribution function. The problem is to test the null hypothesis $H_0 : \mathrm{F}_1(x) = \ldots = \mathrm{F}_k(x) = \mathrm{F}_0(x)$ for all x against the alternative $H_A : \mathrm{F}_i(x) \geq \mathrm{F}_0(x)$ for all x and all i and strict inequality for some x and some $i \varepsilon \{1, 2, \ldots, k\}$. We shall call $\mathrm{F}_i(x)$ to be better than $\mathrm{F}_0(x)$ if $\mathrm{F}_i(x) \geq \mathrm{F}_0(x)$ for all x and strict inequality for some x.

Kiefer (1959) considered the problem of testing $H_0 : \mathrm{F}_1(x) = \ldots = \mathrm{F}_k(x) = \mathrm{F}_0(x)$ for all x, but against all possible alternatives, that is, against the alternative $\mathrm{F}_i(x) \neq \mathrm{F}_0(x)$ for some x and some $i \varepsilon \{1, 2, \ldots, k\}$. Although the alternative taken in this paper is a part of the alternative of Kiefer (1959), yet the alternative H_A is nonparametric in nature. The alternative H_A may be the appropriate alternative in a number of real life situations. For example, let us consider the random variable X that denotes the duration of certain illness. Suppose that a practitioner is using for a long time some popular drug to treat that illness. Let F_0 be the distribution of the random variable X when using that popular drug. So F_0 may very well be assumed to be known. If k new drugs for that particular illness come in the market, then the practitioner would like to know if some among these new drugs is better than the one already

Mathematical Journal of Interdisciplinary Sciences
Vol. 1, No. 1,
July 2012
pp. 17–22

being used. This will at least help him to decide whether to shift to some new drug or to stick to the old treatment. Here obviously the practitioner is interested in the alternative H_A and not in merely rejecting H_0 for all possible alternatives. For some other references one may refer to Barlow et al (1972), Miller (1981), Hochberg and Tamhane (1987), Shaked and Shanthikumar (1994) among others.

In Section 2 we formulate the problem and propose the test for testing the null Hypothesis H_0 against the alternative hypothesis H_A.The properties of the proposed test are studied in Section 3. It is seen that the proposed test has many desirable properties.

2. FORMULATION OF THE PROBLEM AND THE PROPOSED TEST

In this section we shall first define, what may be called a shift function. Let F and G be two arbitrary absolutely continuous distribution functions such that, $F(x) \geq G(x)$, for all x.

For $r \varepsilon [0,1]$,define

$$\Delta(r) = \Delta_{F,G}(r) = FG^{-1}(r) - r.$$

The function $\Delta(.)$ measures the `distance` between two distribution functions F and G. Such a measure of distance between two absolutely continuous distribution functions has been considered by Doksum (1974) in the context of inference problems for nonlinear models in the two sample case. It can easily be seen that:

(i) $\Delta(.)$ is continuous, non-negative and real valued;

(ii) $\Delta_{F,G}(r) \geq \Delta_{F,H}(r)$ for all r if and only if $H(x) \geq G(x)$ for all x;

(iii) $\Delta_{F,G}(r) \geq \Delta_{H,G}(r)$ for all r if and only if $F(x) \geq H(x)$ for all x;

(iv) $\Delta_{F,G}(r) = 0$ for all r if and only if $F(x) = G(x)$ for all x.

Let $\Delta^*(.)$ be such a specified shift function and let $F^*(x)$ be the distribution function such that

$$\Delta_{F^*,F_0}(r) = \Delta^*(r) \text{ for all } r \tag{2.1}$$

This in turn gives $F^*(x) = F_0(x) + \Delta^*(F_0(x)) \geq F_0(x)$ for all x.

Let $X_{i1},\ldots,X_{in}$ be a random sample from distribution , $F_i, i = 1,\ldots,k$, and let the k random samples be independently drawn. Based on the ith random sample, let us define a statistics T_i which depends upon the observations $X_{i1},\ldots,X_{in}$ only through the known distribution function F_0, that is, $T_i = t\left(F_0\left(X_{i1}\right),\ldots,F_0\left(X_{in}\right)\right)$

for i=1,..., k . Let us make the following assumptions about the distribution of the statistics T_i :

Assumption 2.1: The statistics T_i are absolutely continuous random variables.

Assumption 2.2: If $F_i(x) \geq H_i(x)$ for all x, then $G_i(x) \leq K_i(x)$ for all x, where $G_i(x)$ denotes the distribution of the statistics T_i when based on a random sample from $F_i(x)$, and $K_i(x)$ denotes the distribution of the statistics T_i based on a random sample from $H_i(x)$.

For testing the null hypothes is H_0 against the alternative hypothesis H_A we propose the statistics $T = \max_{1 \leq i \leq k} T_i$ and reject H_0 in favour of H_A for large values of T. The statistics T is being proposed as a test statistics for testing H_0 against H_A in view of the following argument. Suppose for some $j \varepsilon \{1,2,\ldots,k\}, F_j(x) \geq F_0(x)$ for all x and strict inequality for some x (this amounts to considering the situation when in fact the alternative H_A is true). In view of assumption 2.2, it would follow that the statistics T_j when based on a sample from F_j would tend to be larger than if T_j were based on a random sample from F_0. Thus if the null hypothesis H_0 is false and the alternative H_A is in fact true, then the statistics T will tend to take larger values. Hence, we reject H_0 in favour of H_A for sufficiently large values of T .

3. PROPERTIES OF THE TEST *T*

We shall first consider the distribution of the test statistics T and then study the properties of the test based on T . Let $\Delta^*(.)$ be specified shift function and let $F^*(.)$ be the distribution function as defined in (2.1). Let us consider the distribution of the statistics T_i when based on a sample from $F^*(.)$ (that is, when for $F_i(x) = F^*(x)$ for all x), and also when based on a sample from F_0 (x) (that is, when $F_i(x) = F_0(x)$ for all x). Obviously, when $F_i(x) = F_0(x)$ for all x, the distribution of $F_0\left(X_{i\alpha}\right)$ for each $\alpha \varepsilon \{1,2,\ldots,n\}$ is uniform and hence the distribution of the statistics Ti would not depend upon known F_0 (x). Also, if $F_i(x) = F^*(x)$ for all x, then from the definition (2.1) of $F^*(.)$, for each $\alpha \varepsilon \{1,2,\ldots,n\}$,

$$p\left\{F_0\left(X_{i\alpha}\right) \leq r\right\} = F^* F_0^{-1}\left(r\right) = r + \Delta^*\left(r\right) \quad \text{for all } r.$$

So in this case also the distribution of $F_0\left(X_{i\alpha}\right)$ for each α, and hence that of T_i, does not depend upon known $F_0(x)$ and depends only upon the shift function $\Delta^*(.)$. This in turn implies that if each F_i is either F^* or F_0, the distribution of the statistics T would not depend upon the knowledge of the distribution F_0 . Hence the test based on the statistics T is a distribution free procedure.

Let us denote by G_i the distribution function of T_i when based on a random sample from F_i, by G^* if T_i is based on a random sample from F^*, where F^* is defined in (2.1), and by G_0 if T_i were based on a random sample from F_0. As already seen the distributions of G_0 and G^* do not depend upon the known distribution F_0. Let us denote by c, the upper quantile $\alpha-$ of the distribution of T when for all x. So

$$\begin{aligned} 1-\alpha &= p\{T \le c \mid F_i(x) = \text{for all } x, \text{all } i\} \\ &= p\{T_i \le c,\ i=1,\ldots,k \mid F_i(x) = F_0(x) \text{for all } x, \text{all } i\} \\ &= \prod_{i=1}^{k} p\{T_i \le c \mid F_i(x) = F_0(x) \text{ for all } x\} \\ &= [G_0(x)]^k \end{aligned} \tag{3.1}$$

The above equalities follow because under the hypothesis H_0 the statistics $T_i's$ are independent and identically distributed random variables each having distribution G_0.

Relation (3.1) implies that $G_0(c) = (1-\alpha)^{1/k}$, that is, c is the quantile of order $(1-\alpha)^{1/k}$ of the distribution G_0. Thus c does not depend upon the known distribution F_0 and can be determined from the distribution G_0.

Monotonicity of power function: Consider the probability of rejection region for the hypothesis H_0 when in fact either H_0 is true or H_A is true, that is,

$$p\{T \ge c \mid F_i(x) \ge F_0(x) \text{ for all } x, \text{all } i\}$$

We shall now show that this probability behaves monotonically in $F_i's$ in the sense that this decreases (increases) if for some F_α is replaced by H_α, where $F_0(x) \le H_\alpha(x) \le F_\alpha(x)$ for all x ($F_0(x) \le F_\alpha(x) \le H_\alpha(x)$ for all x). we in fact have

$$\begin{aligned} & p\{T \ge c \mid F_i(x) \ge F_0(x) \text{ for all } x, \text{all } i\} \\ &= 1 - p\{T \le c \mid F_i(x) \ge F_0(x) \text{ for all } x, \text{all } i\} \\ &= 1 - p\{T_i \le c \textit{ for all } i \mid F_i(x) \ge F_0(x) \text{ for all } x, \text{all } i\} \\ &= 1 - \prod_{i=1}^{k} p_{F_i}\{T_i \le c\}, \text{ where } F_i(x) \ge F_0(x) \text{ for all } x, \text{all } i \\ &= 1 - \prod_{i=1}^{k} G_i(c), \text{ where } G_i(x) \le G_0(x) \text{ for all } x, \text{all } i \\ &\ge (\le) 1 - \left[\prod_{i=1, i\ne\alpha}^{k} G_i(c)\right] k_\alpha(c), \end{aligned} \tag{3.2}$$

where K_α (c) is the distribution function of T_α based on a random sample from H_α and $F_0(x) \le H_\alpha(x) \le (x)$ for all x ($F_0(x) \le F_\alpha(x) \le H_\alpha(x)$ for all x). The inequality (3.2) follows in view of assumption 2.2. Thus the test of the null hypothesis H_0 against thealternative hypothesis H_A based on the statistics T is monotone. Monotonicity in turn implies that the test is unbiased. We in fact have from (3.2), by replacing each F_i by F_0 on the right hand side and by choice of c,

$$\begin{aligned}\text{Power } &= p\{T \ge c \mid F_i(x) \ge F_0(x) \text{ for all } x, \text{all } i\} \\ &\ge p\{T \ge c \mid F_i(x) = F_0(x) \text{ for all } x, \text{all } i\} \\ &= \alpha.\end{aligned}$$

Example 3.1: Let us make the statistics $T_i^{'}$s to be

$$T_i = -2\sum_{\alpha=1}^{n} \log_e F_0(X_{i\alpha}), i = 1,2\ldots,k$$

and the known shift function $\Delta^*(.)$ to be

$$\Delta^*(r) = r^{1-\delta^*} - r \text{ for all } r, 0 \le r \le 1,$$

where $\delta^*(0 \le \delta^* \le 1)$ is a pre-specified number. This choice of $\Delta^*(.)$ gives $F^*(x) = [F_0(x)]^{1-\delta^*}$ for all x So = 0 if and only if $F^*(x) = F_0(x)$ for all x and larger the value of δ* more is the distance between F^* and F_0. In this case it can be seen that

$$G_0(x) = \int_0^x \frac{2^{-n} e^{-y/2} y^{n-1}}{(n-1)!} dy,$$

a gamma distribution with scale parameter 1/2 and shape parameter n (or chi-square distribution with 2n degrees of freedom), and

$$G^*(x) = \int_0^x \frac{2^{-n}(1-\delta^*)^n e^{-y(1-\delta^*)/2} y^{n-1}}{(n-1)!} dy,$$

a gamma distribution with scale parameter $(1-\delta^*)/2$ and shape parameter n.

So the critical point in this case can be obtained from gamma/chi-square tables. In case $F_i(x) = F^*(x)$ for all x and for t values of i, and $F_i(x) = F_0(x)$ for all x and for remaining (k-t) values of i, then we can write from (3.2),

$$\text{Power} = 1 - [G^*(c)]^t [G_0(c)]^{k-t}$$

which is increasing in t, in view of assumption 2.2. We see that for $t = 0$ (that is, under H_0), Power $= \alpha$ and for $t = k$ (that is, when all distributions are better than F_0 and are equidistance δ^* from F_0), Power $= 1 - \left[G^*(c)\right]^k$. Thus in this case the power can be obtained by making use of gamma tables. It may be seen that as $\delta^* \to 1, G^*(c) \to 0$, which in turn implies that Power $\to 1\left(\text{as } \delta^* \to 1\right)$.

Remark 3.1: One may consider the test T using $T_i = D_{in}^+ = \sup_x\left(F_i(x) - F_0(x)\right), i = 1, 2\ldots, k,$ the Kolmogorov's statistics. But the distribution of $T_i's$ even when all $F_i's$ are better than F_0 and are equidistant δ^* from F_0, cannot be written in a closed form. So, even for this situation, one cannot obtain an explicit expression for the power of test T.

REFERENCES

[1] Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D. (1972): Statistical Inference under Order Restrictions. John Wiley & Sons, New York.

[2] Doksum, K. (1974): Empirical probability plots and statistical inference for nonlinear models in the two sample case. *Annals of Statistics*, 2, 267–277. http://dx.doi.org/10.1214/aos/1176342662

[3] Hochberg, Y. and Tamhane, A. C. (1987): Multiple Comparison Procedures. John Wiley & Sons, NewYork.

[4] Kiefer, J. (1959): K-sample analogue of the Kolmogorov-Smirnov and Cramer von-Mises tests. *Annals of Mathematical Statistics*, 30, 420–427. http://dx.doi.org/10.1214/aoms/1177706261

[5] Miller, R. G., Jr. (1981): Simultaneous Statistical Inference, 2nd ed., Springer-Verlag, NewYork.

[6] Shaked, M. and Shanthikumar, J.G. (1994): Stochastic Orders and Their Applications, Academic Press, San Diego.

Narinder Kumar, Professor in Department of Statistics, Panjab University, Chandigarh.

Gobind P. Mehta, Professor in Department of Statistics, Panjab University, Chandigarh.