ISSN 1684-8403
Journal of Statistics
Volume 19, 2012.pp.11-30

Some Powerful Simple and Composite Goodness of Fit Test Based on UIT and related Bahadur Efficiency

Abdolreza Sayyareh[1]

Abstract

In this paper, we have established powerful Goodness of Fit Tests for the basic situation in which the Hypothesized Distribution is known. A new approach of Parameterization is proposed which is a useful approach to construct a Goodness of Fit Test based on Parametric approaches. Various Goodness of Fit Test procedures have been used in literature. We consider the Union-Intersection approach to make some powerful tests to Goodness of Fit Test problem. We simulate the percentage points of introduced statistics. Also, we study the Bahadur Efficiency of the proposed test.

Keywords

Anderson-Darling test, Berk-Jones statistics, Cumulative distribution function, Goodness of fit test, Likelihood ratio, Power of test, Union-Intersection test

1. Introduction

An essential problem in statistics is whether or not a set of measurements is compatible with the assumption that the measurements are an independently identically distributed sample from a known distribution. A difficulty in testing such a statistical Hypothesis is that the Alternatives (Rival Models) are enormously large and could not be described clearly. As the purpose of a Goodness of Fit Test, these tests are intended as tests for distributional form, not as tests of Parametric values. These kind of problems may be called testing Goodness of Fit. They have some strengths and weaknesses. Goodness of Fit Tests are Hypothesis testing problems. But there are some differences.

[1] Department of Statistics, Razi University, Kermanshah,Iran
Email: asayyareh@razi.ac.ir

Hypothesis testing is formulated in terms of Null and Alternative Hypotheses, type one and type two Errors and Power of Tests. In search of the best decision, we turn to the search of a test with acceptable Power. On the other hand, there is no specific Alternative Hypothesis for Goodness of Fit Test, so it is impossible to define the Power of Test simply. Traditionally, Goodness of Fit Tests formulated based on the Cumulative Distribution Functions (c.d.f) of a random variable Y, denoted by $F(y)$, is defined by

$$F(y) = P(Y \leq y), \quad \text{for all } y \in \Psi$$

We consider two directions as Simple and Composite Goodness of Fit Tests.
The Simple Hypotheses are given as

$$H_0: F(y) = F_0(y) \ \forall y \in \Psi \text{ against } H_1: F(y) \neq F_0(y) \text{ for some } y \in \Psi \qquad (1.1)$$

where $F_0(.)$ is a known Distribution Function. We may consider $F_0(.)$ as $F(.,\theta_0)$ where $\theta_0 \in \Theta$ is a specified value of the Vector Parameter θ.
In the Composite situation, we wish to test

$$H_{0c}: F(y) = F(y,\theta) \ \forall y \in \Psi \text{ against } H_{1c}: F(y) \neq F(y,\theta) \text{ for some } y \in \Psi \qquad (1.2)$$

where $\theta \in \Theta$ is an unknown vector of Parameters.

This kind of Hypothesis testing is the problem of whether the underling Distribution Function belongs to a given family of Distribution Functions as $\Phi = \{F(.;\theta), \quad \theta \in \Theta\}$. To test this Composite Hypothesis, we have to estimate θ by $\hat{\theta}_n$ which is a regular estimate of θ. We denote the true value of θ by θ_0. A natural approach to this testing problem is to use the Empirical Distribution Function as an approximation to the true underling Distribution, where the Empirical Distribution Function $F_n(.)$ is defined by

$$F_n(y) = \frac{1}{n}\sum_{i=1}^{n}(Y_i \leq y)$$

Most notable Goodness of Fit Test based on Empirical Distribution Function and $d(F_n(.), F_0(.))$ are the Anderson-Darling A_n^2 Test (1952), Cramer (1928), Kolmogorov (1933), Kuiper V Test (1960), Smirnov (1939, 1941), Von-Mises(1931) and Watson's U^2 Test (1958). The first approach to the problem of testing fit to a fixed distribution is Pearson's (1930) Chi-Squared Test. A way to improve Pearson's statistics consists of employing a functional distance as $d(F_n(.), F_0(.))$. Possibly the best known test statistics based on the Empirical

Distribution Function are the Cramer (1928) and in a more general form Von-Mises(1931) statistics. They proposed

$$\omega_n^2 = n\int_{-\infty}^{\infty}(F_n(y) - F_0(y))^2\zeta(y)dy$$

for some Weight Function ∂ as an adequate measure of discrepancy. The Kolmogorov Test (1933) is the easiest and also most natural Non-Parametric Test. It is based on the Lnorm and computes the distance between an Empirical and the Hypothesized (theoretical) Distribution Function under the Null Hypothesis. Under Alternative Hypothesis, the difference between the Empirical and Theoretical Distribution Functions will be noticeable. This statistic is given by

$$D_n = \sqrt{n}\sup_{y\in \mathrm{P}} | F_n(y) - F_0(y) | .$$

A problem mathematically similar to Kolmogorov's statistic was studied by Smirnov (1939,1941). He has considered D_n^+ and D_n^- where

$$D_n^+ = \sqrt{n}\sup_{y\in \mathrm{P}}(F_n(y) - F_0(y))$$

$$D_n^- = \sqrt{n}\sup_{y\in \mathrm{P}}(F_0(y) - F_n(y)).$$

The statistics D_n, D_n^+ and D_n^- are known as Kolmogorov-Smirnov statistics. They have the advantage of being distribution free. Thus the same p-values can be used to obtain the significance level when testing it to any Continuous Distribution. In search of this property for ω_n^2Cramer-Von-Mises have introduced a simple modification. A modification for Cramer-Von-Mises distance is

$$W_n^2(\psi) = n\int_{-\infty}^{\infty}\psi(F_0(y))\{(F_n(y) - F_0(y))^2\}dF_0(y)$$

which was proposed by Smirnov (1936, 1937).

The Parametric version of this statistics when related Parameter is estimated by $\hat{\theta}_n$ is given by

$$\hat{W}_n^2(\psi) = n\int_{-\infty}^{\infty}\psi(F(y;\hat{\theta}_n))\{(F_n(y) - F(y;\hat{\theta}_n))^2\}dF(y;\hat{\theta}_n).$$

The property exhibited by D_n and $W_n^2(\psi)$ of being distribution free does not carry over to the Parametric cases. However, in some cases the Distribution of $F(Y_i;\hat{\theta}_n)$; $i = 1,2,...,n$ does not depend on θ, but only on Φ, the family of underline densities. In those cases, the Distributions of Parametric Goodness of

FitTests are Parameter free. This happens if Φ is a location scale family and $\hat{\theta}_n$ is an equivariant estimator(see, David and Johnson, 1948). All the statistics which can be obtained by varying ψ are usually refereed to as statistics of Cramer-Von-Mises type, two of them are as follows. The Cramer-Von-Mises's statistic obtained by W_n^2 for $\psi(.) = 1$, $W_n^2 = n\int_{-\infty}^{\infty}(F_n(y) - F_0(y))^2 dF_0(y)$

and the Anderson-Darling's statistic (1952) for $\psi(t) = (t(1-t))^{-1}$

$$A_n^2 = n\int_{-\infty}^{\infty} \frac{(F_n(y) - F_0(y))^2}{F_0(y)(1 - F_0(y))} dF_0(y)$$

with Parametric version as

$$\hat{A}_n^2 = n\int_{-\infty}^{\infty} \frac{(F_n(y) - F(y;\hat{\theta}_n))^2}{F(y;\hat{\theta}_n)(1 - F(y;\hat{\theta}_n))} dF(y;\hat{\theta}_n).$$

Consideration of different Weight Functions ψ allows the statistician to put special emphasis on the detection of particular sets of Alternatives. Some people prefer employing Cramer-Von-Mises statistics instead of Kolmogorov-Smirnov statistics. It is because Kolmogorov-Smirnov statistics accounts only for the largest deviation between $F_n(t)$ and $F(t)$, while the other one is a weighted average of all the deviations between $F_n(t)$ and $F(t)$. Anyway we reject H_0 if in each case the value of the statistic is large. The supremum version of the Anderson-Darling statistics is given by

$$B_n^2 = \sup_{-\infty \le y \le \infty} \frac{|(F_n(y) - F_0(y))|}{\sqrt{F_0(y)(1 - F_0(y))}}.$$

Eicker (1979) considered $\psi(t) = (t(1-t))^{-1} = \{F_n(y)(1 - F_n(y))\}^{-1}$, rather than the Hypothesized variance. Berk and Jones (1979) used the Divergence Function which prepares an approach which give us a test statistic using known Likelihood Ratio Test. More precisely the Berk-Jones statistics as the supremum of the Kullback-Leibler (KL) discrepancy between Hypothesized and Empirical Distribution Functions could be defined as a supreme of $K(F_n(y), F_0(y))$ as

$$K(F_n(y),F_0(y)) = \begin{cases} F_n(y)(\log(\frac{F_n(y)}{F_0(y)})) + (1-F_n(y))\log\frac{1-F_n(y)}{1-F_0(y)} & \text{if} 0 \le \mathrm{F_0(y)} < \mathrm{F_n(y)} \le 1 \\ 0 & \text{if} 0 \le \mathrm{F_n(y)} \le \mathrm{F_0(y)} \le 1 \\ \infty & \text{otherwise.} \end{cases}$$

where $K(F_n(y),F_0(y)))$ is the Kullback-Leibler (KL) discrepancy between two Distributions.

It is known that $K(F_n(y),F_0(y))$ behaves as $\frac{1}{2}\frac{(F_n(y)-F_0(y))^2}{F_0(y)(1-F_0(y))}$.

This last term is half of the Pearson statistics for $F_n(y)$. When we consider the Goodness of Fit Test for Multinomial Distribution, the Pearson χ^2 statistic is asymptotically equivalent to the Likelihood Ratio statistic. Berk-Jones proposed that we can fix ‘y’ and construct a test statistic by Likelihood Ratio Test for Goodness of Fit Test problem. Then, we turn to $F_n(y)$. For each fixed sample $\underline{Y}=(Y_1,...,Y_n)$, $F_n(y)$ is a Distribution Function as a function of $y \in \mathrm{P}$. On the other hand, for each fixed value of ‘y’, $F_n(y)$ is a random variable as a function of the sample and also it is known that $F_n(y)$ is a Unbiased Maximum Likelihood Estimator for $F(y)$. The variance of the Empirical Distribution converges to zero as ‘n’ goes to infinity. These indicate that $F_n(.)$ is weakly and strongly consistent for estimation of $F(y)$. As we know, $nF_n(y): \ Bin(n,F(y))$ for a fixed ‘y’, then under H_0; $nF_n(y): \ Bin(n,F_0(y))$. We concluded that the Likelihood Ratio statistic for testing H_0 against H_1 in fixed $y \in \mathrm{P}$ is given by

$$\lambda_n(y) = \frac{\sup_{F(y)} \Lambda_n(F(y))}{\Lambda_n(F_0(y))}.$$

where $\Lambda_n(F(y))$ and $\Lambda_n(F_0(y))$ are Likelihood Functions evaluated at $F(y)$ and $F_0(y)$ respectively. A suitable relation between Berk-Jones statistics and Likelihood Ratio statistic is as follows

$$\sup_{y\in[0,1]} K(F_n(y),F_0(y)) = \sup_{y\in[0,1]} n^{-1}\log\lambda_n(y).$$

Einmahl and McKeague (2003) introduced an integral form of Berk-Jones statistic. They also considered testing for symmetry, a change point, independence and for exponentiality. Wellner and Koltchinskii (2002) have given proofs of the Limiting Null Distribution of the Berk and Jones (1979) statistic.

In section 2, we bring our objective to using Berk and Jones idea and using this idea with Union-Intersection Test (UIT). This section is a sketch of UIT. Section 3 shows how we make our test. The Power comparisons are given in section 4. After constructing test using UIT, we will search some good Weight Functions as means of a powerful test in two directions as Simple and Composite Hypothesis. In fact, we develop an approach for Simple Hypothesis and then we will use the results for simulation study in both situations. In section 5, we study the Bahadur Efficiency of the proposed test.

2. Motivation

A large family of statistic which embeds χ^2 and Likelihood Ratio Test statistics are obtained by Cressie and Read (1984) family of divergence statistics defined as following

$$2nI_y^{\kappa} = \frac{2n}{\kappa(\kappa+1)}\left\{F_n(y)\left[\frac{F_n(y)}{F_0(y)}\right]^{\kappa} + \{1-F_n(y)\}\left[\frac{1-F_n(y)}{1-F_0(y)}\right]^{\kappa} - 1\right\}$$

for testing the Goodness of Fit of a Multinomial Distribution for the binary sample $X_{1y},\ldots,X_{ny}$ with ‘y’ fixed. Our goal is using Berk-Jones idea in fixing ‘y’ and Likelihood Ratio statistic in fixed ‘y’ in search of a Goodness of Fit Test for two situations, the simple case where $F_0(.)$ is a known Distribution Function and Composite case where as the common approach for Goodness of Fit Test. We have to estimate the unknown Parameter(s) at first and then apply the test. Parametric case will change our situation for model selection from testing for a specified distribution which belongs to the model in more general situation is actually testing for a Family of Distributions (Models). Einmahl and McKeague (2003) have considered the localized Empirical Likelihood Ratio with Likelihood Function as

$$\Lambda(\overline{F}) = \prod_{i=1}^{n}[\Lambda(\overline{F}(Y_i) - \Lambda(\overline{F}(Y_i-)].$$

We would propose the other approach different than Einmahl and McKeague (2003), using some Weight Functions based on classical approach to Hypothesis testing as Union-Intersection approach. This approach has the same advantage as the well known Likelihood Ratio Test approach. The resulting statistics is the same as the Einmahl and McKeague (2003) statistics but our approach makes theUnion-Intersection Test (UIT) suitable for dealing with the Goodness of Fit problem. UIT introduced by Roy (1953) motivated through the multiple comparisons and simultaneous statistical inference. On the other hand, it is easier to illustrate UIT with a Composite Hypothesis testing problem that leads to simultaneous statistical inference. Generally, we use two types of test statistics for testing H_0 against H_1 that can be defined by

$$T = \int_{-\infty}^{\infty} T_z \, dw(z)$$

or

$$T_{max} = \sup_{z \in (-\infty,\infty)} \{T_z w(z)\}.$$

To construct the global test statistic T (or T_{max}) we need some local test statistics as T_z. In this work, we focus on the test statistic as T only, and illustrate our approach to construct the local test statistic.

Consider a random sample as $\underline{Y} = (Y_1, Y_2, ..., Y_n)$ and a Goodness of Fit Test procedure which introduces a Likelihood Ratio Test for each fixed 'z' which could be between any of two Y_i's. Here we must emphases that $F(y)$ is an unknown Distribution Function, whereas $F(z)$ with fixed 'z' is an unknown Parameter. As we saw $nF_n(y) : \; Bin(n, F(y))$, the Likelihood Ratio statistics for testing H_0 against H_1 is given by

$$\lambda_n(z) = \frac{\sup_{F(z)} \Lambda_n(F(z))}{\Lambda_n(F_0(z))} = \frac{\Lambda_n(F_n(z))}{\Lambda_n(F_0(z))} = \left(\frac{F_n(z)}{F_0(z)}\right)^{nF_n(z)} \left(\frac{1-F_n(z)}{1-F_0(z)}\right)^{n(1-F_n(z))}.$$

If we separate the Null Hypothesis $\mathrm{H}_0 : F(y) = F_0(y) \quad \forall y \in \Psi$ (related to a local test) to several Null Hypotheses as $\mathrm{H}_{0z} : F(z) = F_0(z) \quad \forall z \in \mathrm{Z}$, we can construct a Likelihood Ratio for each one of the $\mathrm{H}_{0z}'s$ for each fixed 'z', and then construct a test for our essential Hypothesis testing problem. Fortunately, this concept is known in statistics. The Union-Intersection test (UIT), see Casella and

Berger (2002) andSayyareh (2011), is our proposal to solve this problem. The UIT method is a natural solution to this kind of problem. It is because the overall Hypothesis could be rejected if each local Null Hypotheses could be rejected. As a test statistic, we generalized the logic of the Likelihood Ratio Test. On the other hand, we defined a Weight Function as $w(z)$. This Weight Function permits us to construct different tests. As a choice we consider $w(z) = \cup(F_n(z), F_0(z))$. In the following, after a brief review of UIT, we will construct the Likelihood Ratio Test statistic by UIT, and then we will propose our statistics to model selection. The Likelihood Ratio Test method is a commonly used method of Hypothesis test construction. Another method, which is appropriate when the Null Hypothesis is expressed as an intersection, is the Union-Intersection Test (UIT). In classical statistics we may write

$$\mathrm{H}_0 : \theta \in \bigcap_{\gamma \in \Gamma} \Theta_\gamma$$

where Γ is an arbitrary index set that may be finite or infinite, depending on the problem. By this notation we have

$$\mathrm{H}_1 : \theta \in \bigcup_{\gamma \in \Gamma} \Theta_\gamma^c.$$

Suppose that for each of the testing $\mathrm{H}_{0\gamma} : \theta \in \Theta_\gamma$ against the AlternativeHypothesis $\mathrm{H}_{1\gamma} : \theta \in \Theta_\gamma^c$, we know that the rejection region for the test of $\mathrm{H}_{0\geq}$ is $\{y : T_\gamma(y) \in R_\gamma\}$ where $T_\gamma(.)$ is the test statistic. Thus, if any of the $\mathrm{H}_{0\geq}$ is rejected, then H_0 must also be rejected, it offers a rejection region for UIT as

$$\bigcup_{\gamma \in \Gamma} \{y : T_\gamma(y) \in R_\gamma\}.$$

As a simple example for UIT, we consider a known Hypothesis test in elementary statistics.

Example: Let $Y_1, Y_2, ..., Y_n$ be a independently identically distributed (i.i.d.) random sample from $\mathrm{N}(\mu, \sigma^2)$, where μ and σ^2 are unknown Parameters. We want to test that $\mathrm{H}_0 : \mu = \mu_0$ against $\mathrm{H}_1 : \mu \neq \mu_0$, where μ_0 is a specified number. As a UIT, we can write

$$\mathrm{H}_0 : \{\mu : \mu \leq \mu_0\} \cap \{\mu : \mu \geq \mu_0\}$$

This Null Hypothesis could be written as intersection of two new Null Hypotheses as $\mathrm{H}_{0Lower}:\{\mu:\mu\le\mu_0\}$ and $\mathrm{H}_{0Upper}:\{\mu:\mu\ge\mu_0\}$. Now as the classical approach we will test

$\mathrm{H}_{0Lower}:\mu\le\mu_0$ against $\mathrm{H}_{1Lower}:\mu>\mu_0$

with rejection region $\dfrac{1/n\sum_{i=1}^{n}Y_i-\mu_0}{S/\sqrt{n}}\ge t_{Lower}$ and

$\mathrm{H}_{0Upper}:\mu\ge\mu_0$ against $\mathrm{H}_{1Upper}:\mu<\mu_0$

with rejection region $\dfrac{1/n\sum_{i=1}^{n}Y_i-\mu_0}{S/\sqrt{n}}\le t_{Upper}$.

Then the rejection region of the UIT of

$\mathrm{H}_0:\{\mu:\mu\le\mu_0\}\cap\{\mu:\mu\ge\mu_0\}$

against

$\mathrm{H}_1:\{\mu:\mu\ge\mu_0\}\cup\{\mu:\mu\le\mu_0\}$ for $t_{Lower}=-t_{Upper}$ will be express as $\left|\dfrac{1/n\sum_{i=1}^{n}Y_i-\mu_0}{S/\sqrt{n}}\right|\ge t_{Lower}$ which is the two sided test.

3. Proposed Approach to Construct New Tests

Consider $\underline{Y}=(Y_1,Y_2,...,Y_n)$ as anindependently identically distributed (i.i.d.) random sample with unknown Distribution Function $F(.)$. We set $F_0(.)$ as a known Distribution Function. The official Goodness of Fit Test contains testing

$\mathrm{H}_0:F(y)=F_0(y)\quad\forall y\in\Psi$

against

H_1: $F(y)\neq F_0(y)$ for some $y\in\Psi$

A key for proposing a Goodness of Fit Test is that the Distribution Function $F(z)$ for a fixed ‘z’ is an unknown Parameter. It reduces the Goodness of Fit Test to a Likelihood Ratio Test as

$\mathrm{H}_{0z}:F(z)=F_0(z)\quad\forall z\in Z$

against

H_{1z}: $F(z)\neq F_0(z)$ for some $z\in Z$

As it is a case with Composite Hypotheses testing problems, there may not be, in general, an optimal test for testing H_0 against H_1. However, for a general class of testing problems our idea is to rewrite this Hypothesis testing as the UIT, thus we have

$$\mathrm{H}_0 : \bigcap_{z \in Z} \mathrm{H}_{0z}$$

against

$$\mathrm{H}_1 : \bigcup_{z \in Z} \mathrm{H}_{1z}.$$

In this way, there is flexibility in the decomposition of the Hypotheses and choice of appropriate test statistics. To do this, for each 'z' we can define a new random variable (see, Berk and Jones 1979), thus we have

$$Y_{iz} = 1\{Y_i \le z\} = \begin{cases} 1 & \text{if } \mathrm{Y_i} \le \mathrm{z} \\ 0 & \text{if } \mathrm{Y_i} > \mathrm{z} \end{cases}$$

for $i = 1,2,...,n.$

Now, we have a Parametric test with a binary variable with value in $\{0,1\}^n$, i.e.

$$Y_{iz} : \quad Bin(1, F(z))$$

and

$$\sum_{i=1}^{n} Y_{iz} = nF_n(z) : \quad Bin(n, F(z)).$$

The Likelihood Function is

$$\Lambda_n(F(z)) = \Lambda_n(F(z); \underline{Y}_{iz}) = (F(z))^{nF_n(z)}(1 - F(z))^{n(1-F_n(z))}$$

and the Likelihood Ratio Test is given by

$$\approx_n(z) = \frac{\sup_{F(z)} \Lambda_n(F(z))}{\Lambda_n(F_0(z))}$$

$$\approx_n(z) = \Lambda^{F_n(z)/F_0(z)} \;\alpha\; \frac{\Lambda_n(F_n(z))}{\Lambda_n(F_0(z))}$$

for the large value of $\lambda_n(z)$ we reject the Null Hypothesis. The log-Likelihood Function is given by

$$\log \lambda_n(z) = \log \Lambda^{F_n(z)/F_0(z)} = nF_n(z)\log(\frac{F_n(z)}{F_0(z)}) + n(1 - F_n(z))\log(\frac{1 - F_n(z)}{1 - F_0(z)}).$$

The proposed test statistics for testing H_0 against H_1 is

$$U_n = \int_{\mathrm{P}} \log \Lambda^{F_n(z)/F_0(z)} d(w(z)) = \int_{\mathrm{P}} \log \Lambda^{F_n(z)/F_0(z)} d(\psi(F_n(z), F_0(z)))$$

Note that the decision rule is built from the logical equivalence that H_0 is wrong, if and only if, any of its components H_{0z} is wrong or equivalently H_0 is true if all the H_{0z}'s are individually true. Also assume that we can test H_{0z} using a statistic $T_z(y)$ such that for any Hypothesis included in H_{0z}, $p(\{y \in \Psi; T_z(y) \geq c\})$ is known, for all $c \in \mathrm{P}$ and z. Using this idea in the search of the more powerful test we will consider the different $\psi(F_n(z), F_0(z))'$s for U_n in the next section.

4. Investigation of the Power of Some New Tests by Simulation

To generate new tests we have to choose appropriate Weight Function. Then we need to modify $F_n(z)$ at its discontinuity points $Y_{(i)}$ for $i = 1,2,...,n$. It is trivial that for a point $Y_{(i)} = y$ there are $\frac{i-0.5}{2}$ observations among $Y_1,...,Y_n$ which are less than y. This lead us to consider $F_n(Y_{(i)})$ as $\frac{i-0.5}{2}$.

Selected Weight Function generates a test like a member of the class of the Cramer-Von-Mises. In this section, we propose some Weight Function and simulate their Power against several Alternatives. For each Alternative, the Power result was derived from 10^4 samples of size $n = 50,70,100,120,150,200,250$ depending on choosing κ (related to the Weight Function) and $\psi(F_n(z), F_0(z))$ for Simple and Composite situations.

4.1 Simple Hypothesis: The reasonable choosing of the Weight Function will give us a reasonable test statistic. At the first we consider

$$d\psi(F_n(z), F_0(z)) = 2\{F_n(z)(1 - F_n(z))\}^{-\sqrt{\kappa}} dF_n(z) \quad \text{for} \quad \kappa \geq 0,$$

which is an empiric version of the Weight Function $\psi(F_n(z), F_0(z))$.

By this choice we have

$$T_n = \int_{\mathrm{P}} \log \Lambda^{F_n(z)/F_0(z)} d(w(z)) = \int_{\mathrm{P}} \log \Lambda^{F_n(z)/F_0(z)} 2\{F_n(z)(1-F_n(z))\}^{-\sqrt{\kappa}} d(F_n(z)).$$

Then T_n is given by

$$T_n = 2\frac{1}{n}\sum_{i=1}^{n}\{\log \Lambda^{F_n(X_{(i)})/F_0(X_{(i)})}\{F_n(X_{(i)})(1-F_n(X_{(i)})\}^{-\sqrt{\kappa}}\} =$$

$$2\sum_{i=1}^{n}\{F_n(X_{(i)})\log\frac{F_n(X_{(i)})}{F_0(X_{(i)})}+(1-F_n(X_{(i)})\log\frac{1-F_n(X_{(i)})}{1-F_0(X_{(i)})}\}\{F_n(X_{(i)})(1-F_n(X_{(i)})\}^{-\sqrt{\kappa}}. \qquad (4.1)$$

We generate 10^4 observations from a Beta Distribution, say, $\beta(\eta,\theta)$. Consider the $\beta(1,1)$ as the true (data generate) density. As $F_0(.)$, we consider the Beta Distributions with Parameters as $(\eta,\theta) = (1.5,1.5),(.8,.8),(.6,.6),(1.1,0.8)$ and ≡ =0.1, 0.3, 0.5, 0.7, 1, 1.2, 1.4. For all of tests we set $\alpha = 0.05$ as the level of test. At this given level the critical values of the tests are simulated independently.

Table 1 shows the result of simulations for T_n. For any candidate κ we see that the Power of Test grows when the sample size increases and the Powers converge to 1 very fast. On the other hand, the Power of the new test is always greater than the Power of Anderson-Darling Test. When we set $(\eta,\theta) = (1.5,1.5)$ and $\kappa \le 0.3$ the Power of Anderson-Darling Test is greater than our test.

Table 2 shows the power of the empirical version of the Anderson-Darling test in the same situation as above. In Table 2, the Power of the empirical version of the Anderson-Darling Test in the same situation is stimulated.

Now we consider a simple Weight as $\psi(F_n(z),F_0(z)) = 2dF_n$, which gives T_n as empiric version of the Likelihood Ratio statistic, thus

$$E_n = \int_{\mathrm{P}} [nF_n(z)\log(\frac{F_n(z)}{F_0(z)}) + n(1-F_n(z))\log(\frac{1-F_n(z)}{1-F_0(z)})]d(2F_n(z)) =$$

$$\sum_{i=1}^{n} 2\{F_n(X_{(i)})\log(\frac{F_n(X_{(i)})}{F_0(X_{(i)})}) + (1-F_n(X_{(i)}))\log(\frac{1-F_n(X_{(i)})}{1-F_0(X_{(i)})})\} =$$

$$2\sum_{i=1}^{n}\{F_n(X_{(i)})\log(\frac{F_n(X_{(i)})}{F_0(X_{(i)})}) + (1-F_n(X_{(i)}))\log(\frac{1-F_n(X_{(i)})}{1-F_0(X_{(i)})})\}. \qquad (4.2)$$

This test is more powerful than the Anderson-Darling Test when we set $(\eta,\theta) = (0.6,0.6), \text{or}(1.1,0.8)$. When we choose $(\eta,\theta) = (0.8,0.8)$ our test is the same as the Anderson-Darling Test. But for $(\eta,\theta) = (1.5,1.5)$ our test has a Power which is a little (about 0.1) lower then the Anderson-Darling Test. see Table 3.

The other test of this type may be construct by the Weight Function as

$$d\psi(F_n(z), F_0(z)) = \frac{1}{2}\frac{F_n(z)}{F_0(z)}\frac{1-F_n(z)}{1-F_0(z)}dF_n(z),$$

which introduce a new test, say K_n, where

$$K_n = \int_{\mathrm{P}} \log \Lambda^{F_n(z)/F_0(z)} d(w(z)) = \int_{\mathrm{P}} \log \Lambda^{F_n(z)/F_0(z)} \frac{1}{2}\{\frac{F_n(z)}{F_0(z)}\frac{1-F_n(z)}{1-F_0(z)}\}^{-\sqrt{\kappa}} d(F_n(z)).$$

After simplification K_n is given by

$$K_n = \frac{1}{2}\frac{1}{n}\sum_{i=1}^{n} \log \Lambda^{F_n(z)/F_0(z)}\{\frac{F_n(z)}{F_0(z)}\frac{1-F_n(z)}{1-F_0(z)}\}^{-\sqrt{\kappa}} =$$

$$\frac{1}{2}\sum_{i=1}^{n}\{F_n(X_{(i)})\log\frac{F_n(X_{(i)})}{F_0(X_{(i)})} + (1-F_n(X_{(i)})\log\{\frac{1-F_n(X_{(i)})}{1-F_0(X_{(i)})}\}\}\{\frac{F_n(X_{(i)})(1-F_n(X_{(i)})}{F_0(X_{(i)})(1-F_0(X_{(i)})}\}^{-\sqrt{\kappa}}. \quad (4.3)$$

By this Weight Function, our test has a good Power, but for some $\equiv$ the Power of this test is lower than Anderson-Darling (Table 4). On the other hand, always our test is more powerful than χ^2 test. The Power of χ^2 test for some value of (η,θ) is given in Table 5.

4.2 Composite Hyopthesis: In Composite case, we are testing the Goodness of Fit for a Family of Distributions. To Goodness of Test in (1.2) our test Function will be

$$T_{nc} = 2\frac{1}{n}\sum_{i=1}^{n}\{\log \Lambda^{F_n(z)/F(z;\hat{\theta}_n)}\{F_n(z)(1-F_n(z)\}^{-\sqrt{\kappa}}$$

$$2\sum_{i=1}^{n}\{F_n(X_{(i)})\log\frac{F_n(X_{(i)})}{F_0(X_{(i)};\hat{\theta}_n)} + (1-F_n(X_{(i)})\log\frac{1-F_n(X_{(i)})}{1-F_0(X_{(i)};\hat{\theta}_n)}\}\{F_n(X_{(i)})(1-F_n(X_{(i)})\}^{-\sqrt{\kappa}} \quad (4.4)$$

And

$$K_{nc}=\frac{1}{2}\frac{1}{n}\sum_{i=1}^{n}\log\Lambda^{F_n(z)/F(z;\hat{\theta}_n)}\{\frac{F_n(z)}{F(z;\hat{\theta}_n)}\frac{1-F_n(z)}{1-F(z;\hat{\theta}_n)}\}^{-\sqrt{\kappa}}=$$

$$\frac{1}{2}\sum_{i=1}^{n}\{F_n(X_{(i)})\log\frac{F_n(X_{(i)})}{F_0(X_{(i)};\hat{\theta}_n)}+(1-F_n(X_{(i)})\log\{\frac{1-F_n(X_{(i)})}{1-F_0(X_{(i)};\hat{\theta}_n)}\}\{\frac{F_n(X_{(i)})(1-F_n(X_{(i)})}{F_0(X_{(i)};\hat{\theta}_n)(1-F_0(X_{(i)};\hat{\theta}_n)}\}^{-\sqrt{\kappa}}\quad(4.5)$$

similar to 4.1 and 4.3.
We assume that Y has a Normal Distribution, say $\mathrm{N}(\mu,\sigma^2)$ with $\theta=(\mu,\sigma^2)$ unknown. We can estimate θ by $\hat{\theta}_n=(\bar{Y}_n,S_n^2)$, the mean and sample variance. Then T_{nc} and K_{nc} will be applied to test the Goodness of Fit Test for normality. For Power study, we verify the Power for

$\mathrm{H}_{0c}:Y:\ \mathrm{N}(\mu,\sigma^2)$

against

$\mathrm{H}_{1c}:Y:\ \mathrm{N}(a+\hat{\mu},b\hat{\sigma}^2)$

where $a\in \mathrm{P}$ and $b\in \mathrm{P}^{+}$. The Power of this test for $\equiv=0.1,0.5,1,1.4$, $a=0.3$ and $b=1$ using T_{nc} and K_{nc} are given in Tables 6 and 7 respectively. As we see these tests have a good Power to choose a Model to describe the data at hand.

5. Bahadur Efficiency of Proposed Test

Considering a case when $\psi(F_n(z),F_0(z))=F_n(z)$ we define a test statistic as

$$S_n=\frac{1}{n}\sum_{z_i\in \mathrm{A}_i}\log\Lambda^{F_n(z_i)/F_0(z_i)}$$

then

$$P_{\mathrm{H}_0}(\frac{1}{n}S_n\geq t)=P_{\cap H_{0z}}\{\frac{1}{n}(\frac{1}{n}\sum_{z}\log\Lambda^{F_n(z)/F_0(z)})\geq t\}$$

$$\leq P_{\mathrm{H}_{0z_{max}}}\{\frac{1}{n}\frac{1}{n}n(\max_{z}\log\Lambda^{F_n(z)/F_0(z)})\geq t\}=$$

$$P_{\mathrm{H}_{0z_{max}}}\{\frac{1}{n}\log\Lambda_{z_{max}}^{F_n(z)/F_0(z)}\geq t\}\leq\sum_{\mathrm{H}_{1z_{max}}}P_{\mathrm{H}_{0z_{max}}}\{\frac{1}{n}\sum_{i=1}^{n}\log\frac{Bin(1,F(z_{max}))}{Bin(1,F_0(z_{max}))}\geq t\}\quad(*)$$

$$\leq|\mathrm{H}_{1z_{max}}|\max_{F(z_{max})}P_{\mathrm{H}_{0z_{max}}}\{\exp\sum_{i=1}^{n}\log\frac{Bin(1,F(z_{max}))}{Bin(1,F_0(z_{max}))}\geq\exp(nt)\}=$$

$$|H_1|\max_{F(z_{max})} P_{H_{0z_{max}}}\{\exp\sum_{i=1}^{n}\log\frac{Bin(1,F(z_{max}))}{Bin(1,F_0(z_{max}))}\geq \exp(nt)\}\leq$$

$$|H_1|\max_{F(z_{max})} E_{H_{0z_{max}}}\{\frac{\exp\sum_{i=1}^{n}\log\frac{Bin(1,F(z_{max}))}{Bin(1,F_0(z_{max}))}}{\exp(nt)}\} = \quad (\text{by Markovin equality})$$

$$|H_1|\max_{F(z_{max})} \exp(-nt)E_{H_{0z_{max}}}\{\prod_{i=1}^{n}\frac{Bin(1,F(z_{max}))}{Bin(1,F_0(z_{max}))}\} =$$

$$|H_1|\max_{F(z_{max})} \exp(-nt)\{E_{H_{0z_{max}}}\frac{Bin(1,F(z_{max}))}{Bin(1,F_0(z_{max}))}\}^n =$$

$$|H_1|\max_{F(z_{max})} \exp(-nt)\{\sum Bin(1,F_0(z_{max}))\}^n \leq |H_1|\exp(-nt)$$

so

$$-\frac{2}{n}\log P_{H_0}(\frac{1}{n}S_n\geq t)\geq 2t-\frac{2\log|H_1|}{n}$$

we know that

$$1/n\log\Lambda^{F_n(z)/F_0(z)} = \frac{1}{n}(\log\Lambda^{F_n(z)}-\log\Lambda^{F_0(z)}) =$$

$$\frac{1}{n}(\log\Lambda^{F(z)}-\log\Lambda^{F_0(z)})+\frac{1}{n}(\log\Lambda^{F_n(z)}-\log\Lambda^{F(z)}) \qquad (\text{iff}) \quad F(z)\neq F_0(z) \quad (\text{under} \quad H_{1z})$$

$$=\frac{1}{n}\sum_{i=1}^{n}\log\frac{Bin(1,F(z))}{Bin(1,F_0(z))}+\frac{1}{n}(\log\Lambda^{F_n(z)}-(\log\Lambda^{F(z)})\xrightarrow{\Pi}E_{H_{1z}}\log\frac{Bin(1,F(z))}{Bin(1,F_0(z))}+(0)\chi^2 =$$

$$KL\{Bin(1,F(z)),Bin(1,F_0(z))\} \quad \text{a.s} \quad \text{under} \quad H_{1z}.$$

Thus

$$\frac{1}{n}\sum_{z}\log\Lambda^{F_n(z)/F_0(z)}\xrightarrow{\Pi}E_{H1}KL(Bin(1,F(Y)),Bin(1,F_0(Y)))$$

$$-\frac{2}{n}\log P_{H_0}(\frac{1}{n}S_n\geq t\geq \inf_{H_0}E_{H1}KL(Bin(1,F(Y)),Bin(1,F_0(Y)))$$

Bahadur (1967) showed that the other part of inequality for all of tests is right, then

$$-\frac{2}{n}\log P_{H_0}(\frac{1}{n}S_n\geq t) = 2\inf_{H_0}E_{H1}KL(Bin(1,F(Y)),Bin(1,F_0(Y))).$$

The inequality which namely $(*)$ is correct because

$$\frac{1}{n}\log\Lambda_{z_{max}}^{F_n(z)/F_0(z)} \le \sup_{F\in H_1}\frac{1}{n}\sum_{i=1}^{n}\log\frac{Bin(1,F(z_{max}))}{Bin(1,F_0(z_{max}))}.$$

So the Bahadur Efficiency is achieved.

6. Conclusion

The Goodness of Fit Tests are used for verifying whether or not the experimental data come from the postulated model. In this direction,one must decide if theoretical and experimental Distributions are the same. Then,Goodness of Fit is a Hypothesis testing problem and the problem is concerned with the choice of one of the Alternative Hypothesis. This problem contained the Parameters or not. In this work, we consider a simple situation where the Distribution Function is completely known, and also the Composite case. We have introduced an approach which is known to all statisticians, the Likelihood Ratio approach to Hypothesis testing problem. For simple situation, the family which we consider to simulation study is a simple family, but sensitive to choice of Parameter. This family is U-shaped if both of its Parameters (η,θ) are less than one, is J-shaped if $(\eta-1)(\theta-1)<0$, and is otherwise Unimodal. In the case (η=1, θ=1), this distribution is Uniform Distribution on $(0,1)$. This sensibility to Parameters lets us verify our test to various situations. For Composite situation, we consider Location-Scale Family as Normal Family. Development of this approach to a Weight Function which could be morepowerful than Anderson-Darling Test for any κ is our idea. On the other hand, we showed that a member of this kind of tests is efficient in Bahadur sense.

Table 1: Power computations of H_0: F(.) = β(1,1) against H_1: F(.) = β(η,θ) at level α= 0.05 for (η,θ) = (1.5,1.5) using test function T_n

κ	n=50	n=70	n=100	n=120	n=150	n=200	n=250
0.1	0.137	0.251	0.485	0.557	0.702	0.890	0.957
0.3	0.257	0.446	0.597	0.706	0.825	0.947	0.984
0.5	0.322	0.475	0.664	0.776	0.888	0.972	0.993
0.7	0.333	0.568	0.684	0.780	0.911	0.973	0.994
1.0	0.398	0.610	0.729	0.827	0.915	0.970	0.995
1.2	0.413	0.570	0.742	0.824	0.912	0.975	0.991
1.4	0.438	0.606	0.785	0.829	0.910	0.955	0.987

Table 2: Power computation of H_0: F(.) = β(1,1) against H_1: F(.) = β(η,θ) at level α= 0.05 for (η,θ) = (1.5,1.5), (0.8,0.8), (0.6,0.6), (1.1,0.8) based onAnderson-Darling Test

(η,θ)	n=50	n=70	n=100	n=120	n=150	n=200	n=250
(1.5,1.5)	0.302	0.395	0.622	0.640	0.829	0.933	0.975
(0.8,0.8)	0.041	0.082	0.104	0.109	0.129	0.212	0.297
(0.6,0.6)	0.261	0.393	0.570	0.690	0.889	0.955	0.992
(1.1,0.8)	0.495	0.638	0.782	0.849	0.906	0.964	0.980

Table 3: Power computation of H_0: F(.) = β(1,1) against H_1: F(.) = β(η,θ) at level α= 0.05 for (η,θ) = (1.5,1.5), (0.8,0.8), (0.6,0.6), (1.1,0.8) using test function E_n

(η,θ)	n=50	n=70	n=100	n=120	n=150	n=200	n=250
(1.5,1.5)	0.083	0.166	0.254	0.420	0.566	0.747	0.894
(0.8,0.8)	0.146	0.176	0.210	0.241	0.256	0.339	0.447
(0.6,0.6)	0.580	0.725	0.876	0.922	0.973	0.995	0.999
(1.1,0.8)	0.506	0.646	0.704	0.844	0.923	0.978	0.993

Table 4: Power computation of H_0: F(.) = β(1,1) against H_1: F(.) = β(η,θ) at level α= 0.05 for (η,θ) = (1.5,1.5) using test function K_n

κ	n=50	n=70	n=100	n=120	n=150	n=200	n=250
0.1	0.196	0.300	0.423	0.496	0.712	0.887	0.945
0.3	0.247	0.355	0.526	0.640	0.727	0.896	0.976
0.5	0.275	0.428	0.590	0.686	0.806	0.910	0.962
0.7	0.342	0.441	0.633	0.696	0.827	0.938	0.982
1.0	0.352	0.502	0.674	0.755	0.870	0.945	0.983

κ	n=50	n=70	n=100	n=120	n=150	n=200	n=250
1.2	0.373	0.529	0.676	0.765	0.871	0.947	0.985
1.4	0.407	0.552	0.704	0.793	0.900	0.959	0.986

Table 5: Power computation of H_0: F(.) = β(1,1) against H_1: F(.) = β(η,θ) at level α= 0.05 for (η,θ) = (1.5,1.5), (0.8,0.8), (0.6,0.6), (1.1,0.8) based on χ^2 test

(η,θ)	n=50	n=70	n=100	n=120	n=150	n=200	n=250
(1.5,1.5)	0.210	0.400	0.500	0.591	0.630	0.785	0.810
(0.8,0.8)	0.102	0.170	0.175	0.185	0.200	0.280	0.320
(0.6,0.6)	0.382	0.575	0.711	0.823	0.900	0.984	0.999
(1.1,0.8)	0.085	0.145	0.155	0.165	0.183	0.212	0.264

Table 6: Power computation of H_{0c}: Y: $N(\mu,\sigma^2)$ against H_{1c}: Y: $N(0.3+\hat{\mu},\hat{\sigma}^2)$ at level α= 0.05 using test function T_{nc}

κ	n=50	n=70	n=100	n=120	n=150	n=200	n=250
0.1	0.654	0.739	0.880	0.918	0.989	1.0	1.0
0.5	0.718	0.765	0.871	0.880	0.900	0.999	1.0
1.0	0.509	0.674	0.792	0.912	0.907	0.951	0.994
1.4	0.460	0.637	0.659	0.718	0.850	0.908	0.991

Table 7: Power computation of H_{0c}: Y: $N(\mu,\sigma^2)$ against H_{1c}: Y: $N(0.3+\hat{\mu},\hat{\sigma}^2)$ at level α= 0.05 using test function K_{nc}

κ	n=50	n=70	n=100	n=120	n=150	n=200	n=250
0.1	0.660	0.829	0.939	0.949	0.963	0.980	0.999
0.5	0.453	0.657	0.790	0.949	0.941	0.920	0.990
1.0	0.436	0.676	0.957	0.960	0.972	0.987	0.992
1.4	0.320	0.505	0.638	0.870	0.880	0.947	0.980

References

1. Anderson, T. W. and Darling, D. A. (1952). Asymptotic theory of certain goodness of fit criteria based on stochastic processes.*The Annals of Mathematical Statistics*, **23(2)**, 193-212.
2. Bahadur, R. R. (1967). Rates of convergence of estimates and test statistics. *The Annals of Mathematical Statistics*, **38**, 303–324.
3. Berk, R. H. and Jones, D. H. (1979). Goodness of fit test that dominate the Kolmogorov statistics. *Zeitschrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **47**, 47-59.
4. Cassella, G. and Berger, R. L. (2002). *Statistical Inference,* 2nd edition, Pacific Grove CA: Duxbury Press.
5. Cramer, H. (1928). On the composition of elementary errors. Second paper: Statistical application. *Skand. Atuar tidskr*, **11**, 141-180.
6. Cressie, N. and Read, T. R. C. (1984). Multivariate goodness-of-fit tests. *Journal of the Royal Statistical Society, Series B*, **46(3)**, 440-464.
7. David, F.N. and Johnson, N.L. (1948). The probability integral transformation of when parameters are estimated from the sample. *Biometrika*, **35(1-2)**, 182-190
8. Eicker, F. (1979). The asymptotic distribution of the suprema of the standardized empirical process. *The Annals of Statistics*, **7(1)**, 116-138.
9. Einmahl, J. H. J. and McKeague, I. (2003). Empirical likelihood based hypothesis testing. *Bernoulli*, **9(2)**, 267-290.
10. Kolmogorov, A. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giorn. Ist. Ital. Attuari*, **4**, 83-91.
11. Kuiper, N. H. (1960). Tests concerning random points on a circle. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen, Series A*, **63**, 38–47.
12. Pearson, E. S. (1930). A further development of test for normality. *Biometrika*, **22**, 239-249.
13. Roy, S. N. (1953). On a heuristic method of test construction and its use in multivariate analysis. *The Annals of Mathematical Statistics*, **24**, 220-238.
14. Sayyareh, A. (2011). Limiting distribution and simulated power of some proposed Cramer-Von-Mises type statistic. *International Mathematical Forum*, **6(49)**, 2427-2446.
15. Smirnov, N. V. (1936). Sur la distribution de W_n^2 . *C. R. Acad. Sci. Paris*, **202**, 449-452.

16. Smirnov, N. V. (1937). On the distribution of Mises's W_n^2 criterion. *Mat. Sb.*, Nov., Series 2, 973-993.
17. Smirnov, N. V. (1939). Sur les ecarts de la courbe de distribution empirique. *Matematiceskii Sbornik N.S.*, **6**, 3-26.
18. Smirnov, N. V. (1941). Approximation laws of distribution of random variables from empirical data.*Uspekhi Matematicheskikh Nauk*, **10**, 179-206.
19. Von- Mises, R. (1931). *Wahrscheinlichkeitsrechnung*. Wein, Leipzig.
20. Watson, G.S. (1958). On chi-square goodness-of-fit tests for continuous distributions. *Journal of the Royal Statistical Society, series B*, **20**, 44-61.
21. Wellner, J. A. and Koltchinskii, V. (2002). A note on the asymptotic distribution of Berk-Jones type statistics under the null hypothesis. (http://www.stat.washington.edu/jaw/JAW-papers/NR/HDP2002-rev%20copy.pdf)