GOODNESS-OF-FIT TESTS FOR NONCENTRAL CHI-SQUARE DISTRIBUTION WITH UNKNOWN NONCENTRALITY PARAMETER

Nusrat Jahan & Jane L. Harvill*
Department of Mathematics & Statistics
Mississippi State University

Key Words: Noncentral χ^2, empirical distribution function, goodness-of-fit.

Abstract:

We present Anderson-Darling and Cramér-von Mises goodness-of-fit procedures for testing if a random sample of size n follows an noncentral χ^2 distribution with degrees of freedom ν and noncentrality parameter λ. We consider the case where the null distribution is partially specified; that is, where the degrees of freedom are known, but the value of the noncentrality parameter must be estimated from the sample. The dependence of the size and power of the test upon the method of estimating the noncentrality parameter will be examined and numerically illustrated. Applications of this test extend into testing time series nonlinearity.

1. Introduction

Generally speaking, goodness-of-fit tests are used to determine the likelihood that a random sample was selected from a specified distribution function $F(x, \boldsymbol{\theta})$. The parameter vector $\boldsymbol{\theta}$ may be completely specified or may contain some unknown parameters. Traditional goodness-of-fit tests are based on simple forms of $F(X; \boldsymbol{\theta})$, such as, the uniform, normal, exponential, or gamma distributions. In this paper, we present two goodness-of-fit tests testing $H_0\colon X \sim F(X; \boldsymbol{\theta})$ versus $H_a\colon X \not\sim F(X; \boldsymbol{\theta})$, where F is the cumulative distribution function of a noncentral χ^2 having parameter vector $\boldsymbol{\theta} = (\nu, \lambda)$, where ν represents known degrees of freedom, and λ the unknown noncentrality parameter.

In Section 2., we develop Anderson-Darling and Cramér-von Mises test statistics for the composite null $\chi^2_\nu(\lambda)$. Section 3. discusses how to compute values of the test statistic.

*Dr. Harvill's research was supported in part by NIH grant 1P20RR17661

2. Goodness-of-Fit Tests

Let $X = (X_1, X_2, \ldots, X_n)'$ represent a random sample of size n, and consider testing the hypotheses

$$H_0\colon X \sim F(X; \boldsymbol{\theta}) \text{ versus } H_A\colon X \not\sim F(X; \boldsymbol{\theta}). \quad (1)$$

The parameter vector $\boldsymbol{\theta}$ in the null hypothesis can be either completely specified, in which case H_0 is called a "simple hypothesis," or partially specified, and H_0 is a "composite hypothesis." Very often, the alternative hypothesis is a composite hypothesis and only asserts that the null hypothesis is false. As the name implies, most goodness-of-fit tests attempt to measure the agreement of the data with the null hypothesis.

Let $F_n(x)$ denote the empirical distribution function (EDF), defined by

$$F_n(x) = \begin{cases} 0 & x < x_{(1)} \\ \dfrac{i}{n} & x_{(i)} \le x < x_{(i+1)} \\ 1 & x_{(n)} \le x \end{cases} \quad (2)$$

for $i = 1, \ldots, n$. The estimator $F_n(x)$ in (2) is a consistent estimator of underlying cumulative distribution function $F(x; \boldsymbol{\theta}) = P(X \le x)$. The general EDF test statistic of quadratic class, defined as

$$\begin{aligned} T &= n \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 \psi(x)\, dF(x) \quad (3) \\ &= \|F_n(x) - F(x)\|, \end{aligned}$$

is a measure of agreement between the EDF and the underlying distribution function. Familiar specific forms of the test statistic in (3) include, when $\psi(u) = 1$, the Cramér-von Mises statistic, and the Anderson-Darling statistic, when $\psi(u) = F(u, \boldsymbol{\theta})[1 - F(u; \boldsymbol{\theta})]^{-1}$. For uncomplicated, or well-known null distributions, e.g., the uniform, exponential, gamma, or normal, a number of methods have been devised for testing the simple or composite null hypothesis in (1).

The problem addressed here is an EDF goodness-of-fit test where the composite null hypothesis specifies the sample was selected from a noncentral χ^2

distribution, with known degrees of freedom ν, and unknown noncentrality parameter λ. The cumulative distribution function of a $\chi^2_\nu(\lambda)$ is given by

$$F(x;\nu,\lambda) = \sum_{j=0}^{\infty}\frac{(\lambda/2)^j}{j!2^{(\nu/2)+j}\Gamma\left[(\nu/2)+j\right]} e^{-\lambda/2}\int_0^x y^{(\nu/2)+j-1}e^{-y/2}\,dy. \quad (4)$$

Applying the distribution in (4) to the test statistic in (3) results in a problem under which, finding a null distribution is an intractable. Faced with this, we apply a robust transformation from noncentral χ^2 to standard normal. The primary advantage of this is to have available the known critical values for both Anderson-Darling and Cramér-von Mises test statistic under $N(0,1)$ null. Abdel-Aty (1954) provides an approximation for noncentral χ^2 to normal. In that paper, he equates the skewness and kurtosis of the transformed normal variable to that of the noncentral χ^2 distribution. Specifically,

$$Y = \left(\frac{X}{\nu+\lambda}\right)^h, \quad (5)$$

where $X \sim \chi^2_\nu(\lambda)$. Abdel-Aty sets $h = 1/3$, claiming for this value, the higher-order terms in the third, fourth, and fifth cumulants of Y vanish, and that Y converges in distribution to normal more rapidly than the noncentral χ^2. Sankaran (1959) improved upon the Abdel-Aty transformation by finding an expression for the exponent h in (5) in terms of the degrees of freedom and the noncentrality parameter,

$$h = 1 - \frac{2r(\nu+3\lambda)}{3s^2}. \quad (6)$$

where $r = \nu+\lambda$ and $s = \nu+2\lambda$. With h taking this expression, Sankaran shows that Y has an approximate normal distribution with mean and variance given respectively by

$$\mu = 1 + h(h-1)\frac{s}{r^2} \quad (7)$$

$$-h(h-1)(2-h)(1-3h)\frac{s^2}{2r^4} \quad (8)$$

$$\sigma^2 = h^2\frac{2s}{r^2}\left[1-(1-h)(1-3h)\frac{s}{r^2}\right]. \quad (9)$$

In a 1963 paper, Sankaran developed two additional transformations of a noncentral χ^2 to an approximate normal based on Cornish-Fisher expansion of Y. Johnson and Kotz (1995) provide a numerical comparison of upper and lower 5% points of a good number of normal approximations of the noncentral χ^2, including the Abdel-Aty and the three approximations by Sankaran. Their comparison illustrates that numerically Sankaran's 1959 Biometrika approximation is more accurate in the tails. For the purposes of our study we use Sankaran's 1959 approximation. For $Z = (Y-\mu)/\sigma$, empirical studies have shown that even if parameters are unknown, Z converges in distribution to N(0,1) as long as $\hat{\lambda}$ converges to λ.

3. Computing the Test Statistics

Let $Q_{(i)}$ denote the quantiles computed from the order statistics under the null distribution. Then the Cramér-von Mises statistic may be computed using

$$W^2 = \frac{1}{12n} + \sum_{i=1}^{n}\left[Q_{(i)} - \frac{(2i-1)}{2n}\right]^2, \quad (10)$$

and the Anderson-Darling test statistic can be computed using

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\log Q_{(i)} + \log(R_{(n+1-i)})\right], \quad (11)$$

where $R_{(\cdot)} = 1 - Q_{(\cdot)}$. (See, for example D'Agostino and Stephens (1986).) In either case, an estimate of the noncentrality parameter is needed to compute the quantiles.

We investigated three common estimators of λ (see Johnson and Kotz (1995)). The first was an approximate maximum likelihood estimator given in (12), and two method of moment estimators: one based on the sample median, given in (13), and the other based on the sample mean in (14).

$$\hat{\lambda} = \begin{cases} \frac{1}{2}\left(n^{-1}\sum_{i=1}^{n}\sqrt{x_{(i)}}\right)^2 & \text{if } \bar{x} > \nu \\ 0 & \text{if } \bar{x} \le \nu \end{cases} \quad (12)$$

$$\tilde{\lambda} = \begin{cases} \tilde{x} - \nu & \text{if } \tilde{x} > \nu \\ 0 & \text{if } \tilde{x} \le \nu \end{cases} \quad (13)$$

$$\bar{\lambda} = \begin{cases} \bar{x} - \nu & \text{if } \bar{x} > \nu \\ 0 & \text{if } \bar{x} \le \nu \end{cases} \quad (14)$$

Tables 1 and 2 contain a numerical study of the three estimators based on 1000 replications for samples of size 25 and 100. An examination of these tables reveals that the approximate MLE is, by far, the worst of the three estimators, possessing no desirable properties in an estimator. Specifically, the approximate MLE has a bias and a standard error that increase as degrees of freedom increase. On the other hand, the

Estimator	$\bar{\lambda}$			$\tilde{\lambda}$			$\hat{\lambda}$		
ν	2	10	100	2	10	100	2	10	100
$\lambda = 0.5$	.5138 (.4221)	.6701 (.7808)	1.6381 (2.305)	.1571 (.4569)	.4995 (.7401)	1.7266 (2.563)	.8868 (.5568)	3.4506 (3.874)	30.45 (39.06)
$\lambda = 1$	.9867 (.5305)	1.0378 (.9159)	1.9009 (2.333)	.4377 (.7446)	.8019 (.9411)	1.9556 (2.582)	1.1895 (3.404)	4.4028 (3.997)	32.764 (40.17)
$\lambda = 2$	1.9785 (.6633)	1.9251 (1.108)	2.5162 (2.489)	1.2923 (1.034)	1.5622 (1.288)	2.4946 (2.706)	1.6487 (.4646)	5.5009 (3.707)	38.143 (42.58)
$\lambda = 5$	4.9694 (.9447)	4.83339 (1.374)	4.9523 (2.923)	4.3079 (1.336)	4.3731 (1.721)	4.7163 (3.266)	3.0684 (1.984)	7.0982 (2.203)	50.099 (46.32)
$\lambda = 10$	9.592 (1.283)	9.7487 (1.687)	9.6655 (3.102)	9.3625 (1.696)	9.2994 (2.098)	9.38 (3.673)	5.5267 (4.518)	9.5121 (.9592)	54.53 (44.56)

Table 1: Mean estimate of λ for $n = 25$. (Standard errors in parenthesis.)

Estimator	$\bar{\lambda}$			$\tilde{\lambda}$			$\hat{\lambda}$		
ν	2	10	100	2	10	100	2	10	100
$\lambda = 0.5$	.4777 (.2586)	.5422 (.4518)	.9705 (1.047)	.0245 (.4821)	.1617 (.4462)	.5448 (.8555)	.9537 (.4957)	4.1602 (4.148)	37.475 (43.08)
$\lambda = 1$	.9630 (.3044)	.9740 (.5329)	1.3078 (1.08)	.2008 (.8341)	.4027 (.7472)	.736 (1.021)	1.1837 (.2251)	5.03 (4.162)	42.71 (45.52)
$\lambda = 2$	1.9452 (.3724)	1.93 (.5896)	2.1213 (1.161)	1.0287 (1.051)	1.838 (1.043)	1.286 (1.437)	1.6204 (.4135)	5.6687 (3.684)	48.49 (47.69)
$\lambda = 5$	4.9098 (.5246)	4.8545 (.7003)	4.8806 (1.227)	3.9262 (1.234)	4.0058 (1.282)	3.788 (2.011)	3.0232 (1.992)	7.0906 (2.117)	52.14 (47.15)
$\lambda = 10$	9.87 (.7078)	9.7695 (.8476)	9.6317 (1.314)	8.8581 (1.426)	8.8465 (1.5231)	8.473 (2.279)	5.4753 (4.547)	9.5026 (.6365)	54.50 (44.51)

Table 2: Mean estimate of λ for $n = 100$. (Standard errors in parenthesis.)

method of moments estimator based on the mean of the sample (14) is the best of the three estimators, with a standard error that is unaffected by the value of the degrees of freedom. Moreover, it is not difficult to show that this estimator is consistent, and the numerical studies bear that out. Consequently, it is recommended that the MME based on the mean (14) be the estimate used in computing the test statistic.

To compute the test statistic,

1. Transform the $\chi^2_\nu(\lambda)$ random variable X, into Y, an approximate $N(\mu, \sigma^2)$ random variable using $\bar{\lambda}$ as the estimate of the noncentrality parameter in (5).
2. Standardize the Y values using (7) and (9), and replacing λ in each with $\bar{\lambda}$, yielding for each Y a variable to be denoted Z.
3. Compute the quantiles $Q_{(i)}, i = 1, \ldots, n$ of the ordered Z (using the standard normal cumulative distribution function).
4. Apply these quantiles to the expressions in (10) or (11) to compute the value of the desired test statistic.

Since both mean and variance of the transformed random variable are unknown, if the desired test is the Anderson-Darling test, then the Anderson-Darling test statistic modified for the upper tail should be used; that is, compare

$$A^2_* = A^2\left(1 + \frac{3}{4n} + \frac{9}{4n^2}\right),$$

to the appropriate critical value.

4. Size and Power of the test statistic

Simulation studies based on 1000 replications were conducted to investigate the size and the power of the goodness-of-fit tests. Motivated by the

	Anderson-Darling			Cramér-von Mises		
	$n = 25$					
$\chi_2^2(2)$	0.102	0.042	0.010	0.096	0.046	0.004
$\chi_2^2(10)$	0.099	0.056	0.006	0.096	0.054	0.008
$\chi_2^2(100)$	0.105	0.059	0.008	0.116	0.051	0.009
	$n = 100$					
$\chi_2^2(2)$	0.109	0.061	0.016	0.113	0.061	0.007
$\chi_2^2(10)$	0.102	0.055	0.009	0.087	0.055	0.010
$\chi_2^2(100)$	0.101	0.057	0.008	0.108	0.069	0.014
α	0.10	0.05	0.01	0.10	0.05	0.01

Table 3: Empirical Rejection Rates

	Anderson-Darling			Cramér-von Mises		
n	Uniform(2,6)					
25	0.418	0.241	0.096	0.358	0.204	0.078
100	0.985	0.954	0.842	0.940	0.874	0.649
250	1.000	1.000	1.000	1.000	1.000	0.997
500	1.000	1.000	1.000	1.000	1.000	1.000
	Normal(4,1)					
25	0.168	0.105	0.030	0.151	0.096	0.024
100	0.418	0.309	0.160	0.373	0.271	0.131
250	0.735	0.648	0.436	0.684	0.567	0.377
500	0.968	0.955	0.952	0.966	0.953	0.943
	F(10,2.67)					
25	0.798	0.738	0.601	0.775	0.712	0.568
100	1.000	0.999	0.994	0.999	0.996	0.991
250	1.000	1.000	1.000	1.000	1.000	1.000
500	1.000	1.000	1.000	1.000	1.000	1.000
	Gamma(8,2)					
25	0.112	0.055	0.013	0.108	0.055	0.014
100	0.136	0.077	0.016	0.132	0.076	0.017
250	0.187	0.099	0.031	0.173	0.094	0.030
500	0.265	0.211	0.083	0.248	0.179	0.058
α	0.10	0.05	0.01	0.10	0.05	0.01

Table 4: Empirical Power Under Different Alternatives.

bispectral-based test of nonlinearity of a univariate time series (Harvill, 1999), the degrees of freedom were set to $\nu = 2$ for all studies. Table 3 contains the results of the study the size of the test for three values of λ at two sample sizes ($n = 25$ and 100). The empirical sizes of both the Anderson-Darling and the Cramér-von Mises tests compare favorably to the nominal rates.

Table 4 contains the study of the power of the test based on 1000 replications of samples of size $n = 25, 100, 250$ and 500. Four alternative hypotheses were considered (1) uniform(2,6), (2) normal(4,1), (3) $F(10, 2.67)$, and (4) gamma(8,2). Note that the mean of each of these four distributions is the same as the mean of a noncentral $\chi_2^2(2)$. Excepting the F alternative, the powers of the tests are moderate to low for small sample size ($n = 25$). However, for moderate to large sample size, the power of the tests increases for all alternatives except gamma(8,2), where the power remains low. This is not surprising since the gamma and noncentral χ^2 distributions are closely related. In general, the Anderson-Darling test outperforms the Cramér-von Mises test. This is also not surprising since it is fairly well-known that the Anderson-Darling test typically has the higher power of the two tests.

5. Conclusions

Numerical studies have illustrated that for estimating the noncentrality parameter of a noncentral χ_ν^2 distribution, a method of moments estimators based on the sample mean is a better estimator than a commonly used approximate maximum likelihood estimator, or an MME based on the sample median. Both method of moment estimators are the superior estimator to the approximate MLE in terms of bias and of standard error. Applying the transformation of Sankaran (1959) transformation yields goodness-of-fit test statistic for testing the composite null distribution involving a noncentral χ^2. The resulting test performs well for distributions not closely related to the noncentral χ_ν^2.

References

Abdel-Aty, S.H. (1954) “Approximate formulae for the percentage points and probability integral of the noncentral χ^2 distribution,” *Biometrika*, **41**, pp. 538-540.

Anderson, T.W. and Darling, D.A. (1952) “Asymptotic Theory of Certain ‘Goodness of Fit’ Criteria Based on Stochastic Processes,” *Annals of Mathematical Statistics*, **23**, pp. 193-212.

Anderson, T.W. and Darling, D.A. (1954) “A Test of Goodness of Fit,” *Journal of American Statistical Association*, **49**, pp. 765-769.

D’Agostino, R.B. and Stephens, M.A. (1986) *Goodness-of-Fit Techniques*, New York, John Wiley & Sons.

Harvill, J.L. (1999) “Testing Time Series Linearity via Goodness of Fit Methods,” *Journal of Statistical Planning and Inference,* **75**, pp. 331-341.

Johnson, N.L., Kotz, S., and Balakrishnan, N. (1995) *Continuous Univariate Distributions*, New York, John Wiley & Sons.

Sankaran, M. (1959) “On the noncentral chi-square distribution,” *Biometrika*, **46**, pp. 235-237.

Sankaran, M. (1963) “On the noncentral chi-square distribution,” *Biometrika*, **46**, pp. 235-237.