Journal of the Korean
Statistical Society
Vol. 25, No. 3, 1996

A Smooth Goodness-of-fit Test Using Selected Sample Quantiles

Dale Umbach and M. Masoom Ali[1]

Abstract

A new test for goodness-of-fit is presented. It is a modification of a test of LaRiccia (1991). These tests are applicable to continuous location/scale models. The new test statistic is based on a few selected order statistics taken from the sample, while the LaRiccia test is based directly on the full sample. Each test embeds the hypothesized model in a larger linear model and proceeds to test the goodness-of-fit hypothesis by testing the coefficients of this linear model appropriately. The general theory is presented. The tests are compared via computer simulation to a related test of Ali and Umbach (1989) for distributions that could be used as lifetime models. An important aspect of all these tests is that only standard χ^2 tables are used. Selection of the spacings of the order statistics is discussed.

Key Words : Location/Scale Models; Spacings.

[1]Department of Mathematical Sciences Ball State University Muncie, IN 47306, USA.

1. INTRODUCTION

Let F_o be a given completely specified continuous distribution function with density f_o. Define the location/scale family of distributions generated by F_o as

$$\mathcal{F}_o = \{F|F(x) = F_o((x-\alpha)/\beta), \text{ for some } -\infty < \alpha < \infty \text{ and } 0 < \beta < \infty\}.$$

Let $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ be the order statistics corresponding to a random sample of size n from a continuous distribution with distribution function F. We are to use this information to test the goodness-of-fit hypothesis $H_O : F \in \mathcal{F}_o$ vs. $H_A : F \notin \mathcal{F}_o$.

Hartley and Pfaffenberger (1972) introduced the idea of testing a goodness-of-fit hypothesis by the use of quadratic forms based on selected order statistics. Their test, incorporating the probability integral transform, is applicable to a simple null hypothesis. These ideas were soon expanded in Lurie, Hartley, and Stroud (1974), Kirmani and Alam (1974), and Mehrotra (1982).

LaRiccia (1991) presents a quantile function based analog to Neyman's smooth tests for the location/scale free hypothesis testing situation above. His tests are based on the complete collection of order statistics. Our approach is to use relatively few of the order statistics under a similar setup. The rationale for this approach is that one should expect a faster rate of convergence to χ^2 for the test statistics under the null hypothesis when a fixed spacing for a few of the order statistics is used as compared with the full set of order statistics.

The following definitions and results shall be used throughout the paper. For fixed F_o, define the quantile function by $Q_o(u) = \inf\{x|F_o(x) \geq u\}$ and the density quantile function by $q_o(u) = f_o(Q_o(u))$. Note that for each member of $\mathcal{F}_o$, $Q(u) = \alpha + \beta Q_o(u)$.

2. THE LINEAR MODEL

Fix a spacing for r selected order statistics $0 < u_1 < u_2 < \cdots < u_r < 1$. Let $n_i = [nu_i] + 1$. Let $\underset{\sim}{Y} = (X_{n_1:n}, X_{n_2:n}, \ldots, X_{n_r:n})'$. If $F \in \mathcal{F}_o$, then Mosteller (1946) implies that $\underset{\sim}{Y}$ is asymptotically multivariate normal with

$$E(X_{n_i:n}) \approx \alpha + \beta Q_o(u_i),$$

$$Cov(X_{n_i:n}, X_{n_j:n}) \approx \frac{\beta^2}{n} \frac{u_i(1-u_j)}{q_o(u_i)q_o(u_j)} \quad for \quad i \quad j.$$

Define $\mathbf{W} = ((v_{ij}))$ where $v_{ij} = n\text{Cov}(X_{n_i:n}, X_{n_j:n})/\beta^2$. We note that $\mathbf{W}^{-1} = ((w_{ij}))$ where

$$\begin{aligned} w_{ii} &= (q_o(u_i))^2((u_{i+1} - u_i)^{-1} + (u_i - u_{i-1})^{-1}) \\ w_{i,i+1} &= -q_o(u_i)q_o(u_{i+1})(u_{i+1} - u_i)^{-1} \\ w_{i+1,i} &= w_{i,i+1} \\ w_{ij} &= 0 \text{ otherwise.} \end{aligned}$$

These quantities can be expressed in a myriad of forms. They form the basis of many testing and estimation procedures as outlined in Sarhan and Greenberg (1962), David (1981), and Balakrishnan and Cohen (1991).

To test the goodness-of-fit hypothesis, we consider an alternative model where

$$E(X_{n_i:n}) \approx \alpha + \beta Q_o(u_i) + \sum_{j=1}^{k} \delta_j h_j(u_i)$$

$$Cov(X_{n_i:n}, X_{n_j:n}) \approx \frac{\beta^2}{n}\frac{u_i(1-u_j)}{q_o(u_i)q_o(u_j)} \quad for \quad i \quad j.$$

for some appropriate functions $h_1(\cdot)$, $h_2(\cdot)$, $\ldots$, $h_k(\cdot)$. Note that this model has the variance-covariance matrix for error given by $(\beta^2/n)\mathbf{W}$. Let $\underset{\sim}{\delta} = (\delta_1, \delta_2, \ldots, \delta_k)'$. The goodness-of-fit hypothesis then becomes $\underset{\sim}{\delta} = \underset{\sim}{0}$, for this model.

In practice, one wishes to avoid multicollinearity in the linear model. Thus, the choice of the h functions is crucial. One way to help avoid the problem is to make $h_j(u)$ a function of the hypothesized quantile function $Q_o(u)$. Another good choice for one or more of the h functions is the quantile function of a likely alternative distribution.

3. TEST PROCEDURES

Now for a fixed spacing, we define

$$\begin{aligned} \underset{\sim}{h}_j &= (h_j(u_1),\ h_j(u_2), \ldots,\ h_j(u_r))' \\ \underset{\sim}{Q}_0 &= (Q_o(u_1),\ Q_o(u_2),\ \ldots,\ Q_o(u_r))'. \end{aligned}$$

Using these quantities, the X-matrix for a regression analysis based on the previous linear model can be partitioned as

$$\mathbf{X_1} = (\underset{\sim}{1},\ \underset{\sim}{Q}_0)$$
$$\mathbf{X_2} = (\underset{\sim}{h}_1,\ \underset{\sim}{h}_2,\ \ldots,\ \underset{\sim}{h}_k).$$

Now, for i and $j = 1,\ 2$, let $\mathbf{C_{ij}} = \mathbf{X_i}'\mathbf{W}^{-1}\mathbf{X_j}$, forming

$$\mathbf{C} = \begin{bmatrix} \mathbf{C_{11}} & \mathbf{C_{12}} \\ \mathbf{C_{21}} & \mathbf{C_{22}} \end{bmatrix}.$$

Let $\mathbf{C^{ij}}$ represent the corresponding elements of $\mathbf{C}^{-1}$. Then the regression estimator of the coefficients $(\delta_1,\ \delta_2,\ \ldots,\ \delta_k)'$ can be expressed as

$$\hat{\underset{\sim}{\delta}} = (\mathbf{C^{21}X_1}' + \mathbf{C^{22}X_2}')\mathbf{W}^{-1}\underset{\sim}{Y}.$$

Thus, for this model, we have $\hat{\underset{\sim}{\delta}}$ asymptotically normal with mean $\underset{\sim}{\delta}$ and variance covariance matrix given by $(\beta^2/n)\mathbf{C^{22}}$. So under the null hypothesis, the quadratic form

$$\frac{n\hat{\underset{\sim}{\delta}}'(\mathbf{C^{22}})^{-1}\hat{\underset{\sim}{\delta}}}{\beta^2}$$

is asymptotically χ^2 with k degrees of freedom.

To form a test statistic with a $\chi^2(k)$ distribution under the null hypothesis, we can replace β in the denominator of the quadratic form with any estimator that converges stochastically to β under the null hypothesis. A good choice seems to be

$$\tilde{\beta} = [0,\ 1]\mathbf{C_{11}}^{-1}\mathbf{X_1}'\mathbf{W}^{-1}\underset{\sim}{Y},$$

the regression estimator of β under the null hypothesis. Thus, we propose the test statistic

$$T = \frac{n\hat{\underset{\sim}{\delta}}'(\mathbf{C^{22}})^{-1}\hat{\underset{\sim}{\delta}}}{\tilde{\beta}^2}$$

for testing the goodness-of-fit hypothesis.

4. MONTE CARLO STUDY

To get a specific test, one needs to determine which h functions one wishes to use as well as a spacing. In the Monte Carlo study which follows, we have focussed on distributions for F_o whose support is contained in $(0, \infty)$, which could be considered as lifetime distributions. We considered $h_1(u) = Q_o(u)\ln(Q_o(u))$, $h_2(u) = \sqrt{Q_o(u)}$, and $h_3(u) = (Q_o(u))^2$. These particular functions were used by LaRiccia (1991) in his Monte Carlo study. We use them here so that the results are comparable. Clearly, h_1 and h_2 would not be suitable if the support of F_o contained any negative values. Furthermore, if one can target a specific alternative as plausible, then one could use a function of the form $h(u) = Q_1(u)$, where $Q_1(u)$ is the quantile function associated with the standardized form of the targeted alternative distribution.

To investigate the performance of the test statistic, T, a Monte Carlo study was performed. Each of the three possible pairs of h functions above was used, along with either 5, 6, or 9 order statistics. For comparative purposes, LaRiccia's statistic was also calculated as was the X_G^2 statistic of Ali and Umbach (1989). Specifically, the statistic T_1 of LaRiccia (eq. 5 and following) was used. In each case 1000 replications of samples of size 100 from various distributions were used. Additionally, each block of 1000 replications all started with the same seed for consistency.

Table 1 presents some selected results designed to check the statistics as to the accuracy of the χ^2 null distribution. Each of the three statistics has an asymptotically χ^2 distribution under H_0, X_G^2 with degrees of freedom $r - 2$, and the other two statistics with degrees of freedom $k = 2$. In each case 1000 values of the test statistic that were generated using samples of size 100. Specifically with $F_n(x)$ representing the empirical distribution function of these 1000 values, F^* representing the $\chi^2(2)$ or $\chi^2(r-2)$ distribution function as appropriate, and $D = \sup_{x\geq 0}|F_n(x) - F^*(x)|$, the value of $n^{1/2}D = \sqrt{1000}\,D$ was calculated. Thus, the smaller the value of $\sqrt{1000}\,D$ the more supportive the Monte Carlo study is that the test statistic follows a χ^2 distribution. In each case a uniform spacing of the order statistics was used based on either 5 or 9 selected order statistics. The results are reported in the last four columns of Table 1.

Generally, the values seem fairly supportive of χ^2-ness of the new test statistics and X_G^2 for samples of size 100 across a variety of distributions. In fact, the new statistic performed excellently across the board except for the case where h_1 and h_2 were used for the Weibull distribution with shape parameter $= 4$. The same can be said for the X_G^2 statistic except for the Weibull distribution with shape parameter $= 0.5$. The generally larger values reported for the LaRiccia statistic indicate that the true distribution for the test statistic for samples of size 100 is not as close to the $\chi^2(2)$ distribution as the other two statistics. This indicates a generally slower rate of convergence to the $\chi^2(2)$ distribution for the LaRiccia statistic, at least for uniform spacings.

Table 2 presents some selected results of a Monte Carlo study designed to check the power of the various tests. In each case the number of times (out of 100) that the test statistics fell in the critical region of a 10% test was calculated for various combinations of hypothesized distributions and distributions actually used to generate the data. In each case, a uniform spacing was used. These results indicate somewhat poorer power when compared with the LaRiccia test. However, in conjunction with the results above, this is to be expected. As expected, the test based on $r = 9$ order statistics had better power than the test based on $r = 5$ order statistics. Surprisingly, the choice of which two h functions to use seems to be immaterial among the three presented. It would be difficult to generalize here, since the choices were extremely limited.

One should not expect uniform spacings to be optimal, however. For estimation problems based on selected order statistics, it is typically the case that the optimal spacing is far from uniform under a variety of optimality criteria. However, with such a large class of distributions in the alternative hypothesis, it is difficult to attack the problem of improving power analytically. Another approach is to choose spacings that have proven to be efficient in other settings.

Umbach and Ali (1993) present an approach to choosing spacings which are robust for estimation of location and scale parameters. A parametric hypothesis testing approach which leads to the same results was presented by Saleh and Sen (1985). Umbach and Ali have computed optimal spacings under this approach for a few distributions. In particular, they present the following spacing for the $\chi^2(6)$ distribution.

$$u_1 = 0.00046,\ u_2 = 0.0105,\ u_3 = 0.0849,\ u_4 = 0.8058,\ u_5 = 0.9625$$

Using this special spacing, a Monte Carlo study as described above was carried out. The results are reported in Table 3.

Hassanein (1971) has calculated the spacing that minimizes the generalized variance of the ABLUE (asymptotically best linear unbiased estimator). of (α, β) for the Weibull distribution. In particular, for shape parameter $\delta = 4$, they present the following optimal spacing

$$u_1 = .0011,\ u_2 = .0127,\ u_3 = .0630,\ u_4 = .2063,\ u_5 = .8932,\ u_6 = .9824$$

Using this special spacing, a Monte Carlo study as described above was carried out. The results are reported in Table 4.

Table 1. Results of the simulation of 1000 replications for samples of size 100 when the null hypothesis is true. In each case a uniform spacing was used with $r = 5$ or 9. ($W(\delta)$ refers to the Weibull distribution with shape parameter δ.)

Statistic	r	h functions	Kolmogorov-Smirnov $\sqrt{1000}\ D$			
			$H_0 : \chi^2(2)$	$H_0 : \chi^2(6)$	H_0 :Beta(1,1)	H_0 :Beta(1,2)
X_G^2	5		0.6348	0.6170	0.6365	0.5464
	9		0.9404	0.6088	1.0381	0.7235
LaRiccia		$h_1,\ h_2$	0.7681	1.1560	1.5367	1.7055
		$h_1,\ h_3$	1.7434	1.7632	1.2034	1.2602
		$h_2,\ h_3$	1.3548	1.5907	1.2119	1.4439
New	5	$h_1,\ h_2$	0.8616	0.8035	0.9006	0.6408
		$h_1,\ h_3$	0.8198	0.7043	0.7793	0.6409
		$h_2,\ h_3$	0.8810	0.7442	0.8849	0.5880
	9	$h_1,\ h_2$	0.4809	0.5683	0.8260	0.5068
		$h_1,\ h_3$	0.6741	0.6376	0.8050	0.8169
		$h_2,\ h_3$	0.5165	0.6233	0.7734	0.5462
			$H_0 : W(0.5)$	$H_0 : W(1.5)$	H_0 :W(2)	H_0 :W(4)
X_G^2	5		1.7081	0.6086	0.6288	0.6136
	9		3.1289	0.6568	0.6318	0.8693
LaRiccia		$h_1,\ h_2$	2.7810	0.7580	0.6061	0.8121
		$h_1,\ h_3$	2.5890	1.2000	0.8981	1.0032
		$h_2,\ h_3$	3.0917	1.0488	0.9379	0.9674
New	5	$h_1,\ h_2$	0.6447	0.9058	0.7713	1.4263
		$h_1,\ h_3$	0.5853	0.8645	0.7052	0.6888
		$h_2,\ h_3$	0.4878	0.8735	0.6633	0.8060
	9	$h_1,\ h_2$	0.5307	0.5926	0.7108	1.7287
		$h_1,\ h_3$	0.8115	0.6844	0.6897	0.7186
		$h_2,\ h_3$	0.6404	0.5997	0.7716	0.8306

Table 2. Results of the simulation of 1000 replications for samples of size 100 when the null hypothesis is false. In each case a uniform spacing was used with $r = 5$ or 9. "Number of Rejections" reports the number of times (out of 1000) that the various test statistics fell in the critical region of a 10% test. ($W(2)$ refers to the Weibull distribution with shape parameter 2.)

				Number of Rejections			
				True Distribution			
Statistic	r	h functions	H_0	$\chi^2(2)$	$\chi^2(8)$	$W(2)$	F(6,2)
X_G^2	5		$\chi^2(6)$	130	179	995	698
	9			275	117	1000	906
LaRiccia		h_1, h_2		997	225	334	1000
		h_1, h_3		997	221	304	1000
		h_2, h_3		997	232	307	1000
New	5	h_1, h_2		448	110	176	989
		h_1, h_3		444	110	175	988
		h_2, h_3		446	108	177	988
	9	h_1, h_2		758	106	211	1000
		h_1, h_3		755	104	196	1000
		h_2, h_3		754	105	200	1000
X_G^2	5		$\chi^2(4)$	113	427	895	775
	9			196	615	1000	935
LaRiccia		h_1, h_2		955	665	819	1000
		h_1, h_3		952	646	810	1000
		h_2, h_3		955	650	806	1000
New	5	h_1, h_2		292	175	293	975
		h_1, h_3		294	178	289	975
		h_2, h_3		296	178	290	975
	9	h_1, h_2		513	218	401	1000
		h_1, h_3		526	215	389	1000
		h_2, h_3		518	214	388	1000

Table 3. Results of the simulation of 1000 replications for samples of size 100 for the special spacing 0.00046, 0.0105, 0.0849, 0.8058, 0.9625. "Number of Rejections" reports the number of times (out of 1000) that the various test statistics fell in the critical region of a 10% test. ($W(2)$ refers to the Weibull distribution with shape parameter 2.)

			Number of Rejections				
			True Distribution				
Statistic	h functions	H_0	$\chi^2(6)$	$\chi^2(2)$	$\chi^2(8)$	$W(2)$	F(6,2)
X_G^2		$\chi^2(6)$	74	45	125	1000	995
LaRiccia	h_1, h_2		135	997	225	334	1000
	h_1, h_3		138	997	221	304	1000
	h_2, h_3		136	997	232	307	1000
New	h_1, h_2		113	999	63	86	1000
	h_1, h_3		145	1000	82	102	1000
	h_2, h_3		145	1000	76	91	1000

Table 4. Results of the simulation of 1000 replications for samples of size 100 for the uniform spacing with $r = 6$ and the special spacing 0.0011, 0.0127, 0.0830, 0.2063, 0.8932, 0.9824. "Number of Rejections" reports the number of times (out of 1000) that the various test statistics fell in the critical region of a 10% test. ($W(\delta)$ refers to the Weibull distribution with shape parameter δ.)

			Number of Rejections				
			True Distribution				
Statistic	h functions	H_0	$W(4)$	$W(3)$	$W(5)$	$\chi^2(6)$	Beta(2,2)
LaRiccia	h_1, h_2	$W(4)$	131	256	231	999	310
	h_1, h_3		131	280	238	999	357
	h_2, h_3		135	272	230	999	340
Uniform Spacing							
X_G^2			121	272	118	1000	106
New	h_1, h_2		109	113	133	614	128
	h_1, h_3		98	136	124	657	151
	h_2, h_3		97	133	126	652	145
Special Spacing							
X_G^2			77	321	115	1000	35
New	h_1, h_2		78	279	56	995	370
	h_1, h_3		81	317	65	996	508
	h_2, h_3		78	305	72	996	456

5. CONCLUSION

The generally smaller Kolmogrov-Smirnov values reported for the new statistic in Table 1 indicate that the statistic has a null distribution which is closer to its limiting χ^2 than the LaRiccia statistic for uniform spacings. This makes the power comparison in Table 2 more difficult, since the LaRiccia statistic has a propensity toward large values in general. It is interesting to note that there were cases when the X_G^2 statistic proved to be superior to both of the other statistics, e.g. $H_0 : F = \chi^2(6)$ with the actual sample arising from the Weibull distribution with shape parameter = 2, and cases when it was quite inferior, e.g. $H_0 : F = \chi^2(6)$ with the actual sample arising from the $\chi^2(6)$ distribution.

The results for the special spacings are mixed. The first column of values in Table 3 indicates that the special spacing for the $\chi^2(6)$ distribution does not yield a test statistic that is closer to a χ^2 than LaRiccia's. It's power is also generally poorer.

However, Table 4 indicates that for the Weibull distribution with shape parameter = 4, the special spacing yields an improvement in both power and significance level over the uniform spacing. The power also compares favorably with that of LaRiccia's test in all cases considered, except for the Weibull(5) case.

Tables 3 and 4 indicate that if a special spacing is to be used instead of the uniform spacing, care must be taken to insure that it is appropriate for the distribution in question. These special spacings have not been computed for many distributions. This makes generalizations difficult as to when it is appropriate to use them. This study indicates that the new test criterion is competitive with the LaRiccia test. It seems to be operating very closely to its advertised α-level over a variety of spacings, especially the uniform spacing, as is indicated by Tables 1, 3, and 4. Further study may indicate particularly effective spacings and h functions to be used. Information concerning h functions would most probably be valid for LaRiccia's test also.

ACKNOWLEDGEMENT

The authors are grateful to the referees for their helpful comments which improved the paper significantly.

REFERENCES

(1) Ali, M. M. and Umbach, D. (1989). A Shapiro-Wilk type goodness-of-fit test using a few order statistics. *Journal of Statistical Planning and Inference*, **22**, 251-261.

(2) Balakrishnan N. and Cohen A. C. (1991). *Order Statistics and Inference: Estimation Methods*, Academic Press, San Diego.

(3) David, H. A. (1981). *Order Statistics*, 2nd ed., John Wiley & Sons, New York.

(4) Hartley, H. O. and Pfaffenberger, R. C. (1972). Quadratic forms in order statistics used as goodness-of-fit criteria. *Biometrika*, **59**, 605-611.

(5) Hassanein, K. M. (1971). Percentile estimators for the parameters of the Weibull distribution. *Biometrika*, **58**, 673-676.

(6) Kirmani, S. N. U. A. and Alam, S. N. (1974). On goodness-of-fit tests based on spacings. *Sankhyā* A, **36**, 197-203.

(7) LaRiccia, V. N. (1991). Smooth goodness-of-fit tests: A quantile function approach. *Journal of the American Statistical Association*, **86**, 427-431.

(8) Lurie, D., Hartley, H. O. and Stroud, M. R. (1974). A goodness-of-fit test for censored data. *Communications in Statistics*, **3**, 745-753.

(9) Mehrotra, K. G. (1982). On goodness-of-fit tests based on spacings for type II censored samples. *Communications in Statistics–Theory and Methods*, **11**(8), 869-878.

(10) Mosteller, F. (1946). On some useful "inefficient" statistics. *Annals of Mathematical Statistics*, **17**, 377-408.

(11) Saleh, A. K. M. E. and Sen, P. K. (1985). Asymptotic relative efficiency of some joint tests for location and scale parameters based on selected order statistics. *Communications in Statistics–Theory and Methods*, **14**(3), 621-633.

(12) Sarhan, A. E. and Greenberg, B. G. (*eds.*) (1962). *Contributions to Order Statistics*, John Wiley & Sons: New York.

(13) Umbach, D. and Ali, M. M. (1993). Conservative spacings for the estimation of functions of location and scale. *Journal of Information and Optimization Sciences*, **14**, 309-319.