The Stata Journal

The *Stata Journal* publishes reviewed papers together with shorter notes or comments, regular columns, book reviews, and other material of interest to Stata users. Examples of the types of papers include 1) expository papers that link the use of Stata commands or programs to associated principles, such as those that will serve as tutorials for users first encountering a new field of statistics or a major new technique; 2) papers that go "beyond the Stata manual" in explaining key features or uses of Stata that are of interest to intermediate or advanced users of Stata; 3) papers that discuss new commands or Stata programs of interest either to a wide spectrum of users (e.g., in data management or graphics) or to some large segment of Stata users (e.g., in survey statistics, survival analysis, panel analysis, or limited dependent variable modeling); 4) papers analyzing the statistical properties of new or existing estimators and tests in Stata; 5) papers that could be of interest or usefulness to researchers, especially in fields that are of practical importance but are not often included in texts or other journals, such as the use of Stata in managing datasets, especially large datasets, with advice from hard-won experience; and 6) papers of interest to those who teach, including Stata with topics such as extended examples of techniques and interpretation of results, simulations of statistical concepts, and overviews of subject areas.

The *Stata Journal* is indexed and abstracted by *CompuMath Citation Index*, *Current Contents/Social and Behavioral Sciences*, *RePEc: Research Papers in Economics*, *Science Citation Index Expanded* (also known as *SciSearch*, *Scopus*, and *Social Sciences Citation Index*.

For more information on the *Stata Journal*, including information for authors, see the webpage

http://www.stata-journal.com

Subscriptions are available from StataCorp, 4905 Lakeway Drive, College Station, Texas 77845, telephone 979-696-4600 or 800-STATA-PC, fax 979-696-4601, or online at

http://www.stata.com/bookstore/sj.html

Subscription rates listed below include both a printed and an electronic copy unless otherwise mentioned.

U.S. and Canada		Elsewhere	
1-year subscription	$ 79	1-year subscription	$115
2-year subscription	$155	2-year subscription	$225
3-year subscription	$225	3-year subscription	$329
3-year subscription (electronic only)	$210	3-year subscription (electronic only)	$210
1-year student subscription	$ 48	1-year student subscription	$ 79
1-year university library subscription	$ 99	1-year university library subscription	$135
2-year university library subscription	$195	2-year university library subscription	$265
3-year university library subscription	$289	3-year university library subscription	$395
1-year institutional subscription	$225	1-year institutional subscription	$259
2-year institutional subscription	$445	2-year institutional subscription	$510
3-year institutional subscription	$650	3-year institutional subscription	$750

Back issues of the *Stata Journal* may be ordered online at

http://www.stata.com/bookstore/sjj.html

Individual articles three or more years old may be accessed online without charge. More recent articles may be ordered online.

http://www.stata-journal.com/archives.html

The *Stata Journal* is published quarterly by the Stata Press, College Station, Texas, USA.

Address changes should be sent to the *Stata Journal*, StataCorp, 4905 Lakeway Drive, College Station, TX 77845, USA, or emailed to sj@stata.com.

The Stata Journal (2012)
12, Number 3, pp. 368–374

The Chen–Shapiro test for normality

Michal Brzezinski
Faculty of Economic Sciences
University of Warsaw
Warsaw, Poland
mbrzezinski@wne.uw.edu.pl

Abstract. The Chen–Shapiro test for normality (Chen and Shapiro, 1995, *Journal of Statistical Computation and Simulation* 53: 269–288) has been shown in simulation studies to be generally slightly more powerful than the commonly used Shapiro–Wilk and Shapiro–Francia tests, implemented in Stata official commands `swilk` and `sfrancia`. I present the `chens` command, which performs the Chen–Shapiro test in Stata.

Keywords: st0264, chens, normality testing, Chen–Shapiro test

1 Introduction

Testing the hypothesis that data are normally distributed plays an important role in many areas of empirical science. Stata has three official commands to test for univariate normality: `swilk` (SW), `sfrancia` (SF), and `sktest` (SK). They implement, respectively, the Shapiro–Wilk test (Shapiro and Wilk 1965), the Shapiro–Francia test (Shapiro and Francia 1972), and the skewness and kurtosis test proposed by D'Agostino, Belanger, and D'Agostino (1990) with an adjustment made by Royston (1991a).

The last test, in general, performs worse for nonaggregated or ungrouped data than the first two (Gould and Rogers 1991; Gould 1992; Royston 1991b). As recommended in [R] **sktest**, the Shapiro–Francia test should be used whenever dealing with nonaggregated or ungrouped data. If normality is rejected, SK may be used to determine the reasons for rejection. Moreover, as suggested by D'Agostino, Belanger, and D'Agostino (1990) and Royston (1991b), conclusions about normality should not be drawn solely on the basis of any test statistic—normal quantile plots should be treated as at least an equally meaningful tool for determining normality. These plots are implemented in Stata with the `qnorm` command; see [R] **diagnostic plots** for more information.

Chen and Shapiro (1995) introduced a test for normality that compares the spacings between order statistics with the spacings between their expected values under normality. The test is easy to compute and has been shown in a simulation study by Chen and Shapiro (1995) to be as powerful as or superior to the SW test for most of the alternatives studied. The Chen–Shapiro (CS) test emerged as the best overall test for a wide diversity of alternatives from a comparison of nine normality tests conducted in Seier (2002). She recommended that statistical software should incorporate this test.

Recently, the CS test was included in a comprehensive power comparison study of 33 normality tests by Romão, Delgado, and Costa (2010). Their results indicate that the

CS test is among the best choices for symmetric distributions, while both the CS and the SW tests, among a few others, are recommended for asymmetric distributions, for modified normal distributions (excluding normal distributions with outliers), and when the nature of nonnormality is unknown. In the study, the CS test has on average slightly higher empirical power than the SW test and a bigger advantage over the SF test.

In this article, I present the `chens` command, which implements the CS test for normality in Stata. In section 2, I briefly describe the CS test. In section 3, I present the `chens` command, and I provide an example in section 4. In section 5, I use simulations to compare the empirical power of tests performed by SW, SF, SK, and `chens`. I conclude the article in section 6.

2 CS test

For a sample of size n, $x_1, x_2, \ldots, x_n$, with $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ representing the order of the statistics of that sample, the test for normality introduced by Chen and Shapiro (1995) is based on normalized spacings and is defined as follows

$$QH = \frac{1}{(n-1)s} \sum_{i=1}^{n-1} \frac{x_{(i+1)} - x_{(i)}}{H_{i+1} - H_i}$$

where $H_i = \Phi^{-1}\{(i - 3/8)/(n + 1/4)\}$, Φ^{-1} is the inverse of the standard normal distribution, and s is the sample standard deviation. When the normality hypothesis is true, the distribution of QH will have a mean close to 1. Moreover, Chen and Shapiro (1995) prove that QH is a lower-tailed test. The QH statistic is easily computed by using a function returning the inverse normal cumulative distribution; see `invnormal()` in [D] **functions**. The distribution of QH under the null hypothesis of normality was investigated empirically for $n = 3(1)50, 60, 80, 100, 150, 250, 500, 1000, 2000$ by Chen and Shapiro (1995), with 100,000 samples drawn from the standard normal distribution. To eliminate the need for tables of percentiles, they defined the statistic

$$QH^* = \sqrt{n}(1 - QH)$$

which is an upper-tailed test. Chen and Shapiro (1995) fit regression equations to selected upper quantiles of QH^*, $q = 0.001, 0.005(0.005)0.05(0.01)0.10$, as a function of sample size ($10 \leq n \leq 2000$). Thus the p-value of the QH^* test can be approximated by linear interpolation between two quantiles. In this article, I have reestimated the regression equations proposed by Chen and Shapiro (1995) by using empirical upper quantiles of QH^* computed with 1,000,000 normal samples for sample sizes from 5 to 2,000.

3 The chens command

3.1 Syntax

The syntax for the `chens` command is analogous to that of the SW command:

`chens` *varlist* [*if*] [*in*] [, `noties`]

3.2 Option

`noties` suppresses the use of averaged ranks for tied values when calculating the W test coefficients.

4 Example

I illustrate the `chens` command by using `sp500.dta`, which comes with Stata. Let us test whether a variable called `open` is distributed normally. Using existing commands, we obtain the following results:

```
. sysuse sp500
(S&P 500)

. sktest open

                   Skewness/Kurtosis tests for Normality
                                                        ------- joint ------
    Variable |    Obs   Pr(Skewness)   Pr(Kurtosis)  adj chi2(2)    Prob>chi2
-------------+---------------------------------------------------------------
        open |    248      0.8984         0.0500          3.89         0.1432

. swilk open

                  Shapiro-Wilk W test for normal data
    Variable |    Obs       W           V         z       Prob>z
-------------+--------------------------------------------------
        open |    248    0.98998      1.805     1.374    0.08475

. sfrancia open

                 Shapiro-Francia W' test for normal data
    Variable |    Obs       W'          V'        z       Prob>z
-------------+--------------------------------------------------
        open |    248    0.99214      1.539     0.903    0.18316
```

All three tests suggest that we cannot reject the hypothesis that `open` is normally distributed at the 5% significance level. Now let us repeat the test by using the `chens` command:

```
. chens open

               Chen-Shapiro QH* test for normal data
    Variable |    Obs       QH         QH*      P-value
-------------+----------------------------------------
        open |    248    0.99986    0.00218    0.04816
```

The result suggests that according to `chens`, the normality of `open` is marginally rejected at the 5% level. However, the differences between the p-values for the SW and `chens` commands are rather small, and it is difficult to decide which test should be preferred.

A normal quantile plot for `open` is shown in figure 1. There are noticeable deviations from normality in the upper tail of the distribution, which may justify why `chens` rejects normality at the 5% significance level. This suggests that at least for the variable considered, `chens` may be a useful alternative for testing normality in Stata.

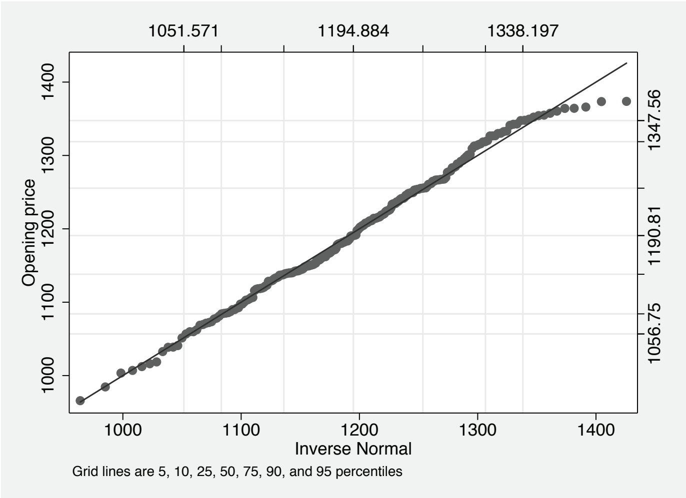

Figure 1. Normal quantile plot for `open` variable

5 Simulation study

In this section, I present the results of a simulation study comparing the empirical power of the CS test and the tests performed by SW, SF, and SK. The setup of the simulation largely follows Gould (1992). The following basic alternative distributions were included: uniform$(0, 1)$; two cases of Student's t distribution with 5 and 20 degrees of freedom, $t(5)$ and $t(20)$; two cases of chi-squared distribution with 5 and 10 degrees of freedom, $\chi^2(5)$ and $\chi^2(10)$; two cases of beta(a, b) distribution, beta$(2, 2)$ and beta$(2, 1)$; and one case of Tukey(λ) distribution, Tukey(0.5).

In addition, three modified normal distributions were used. A location-contaminated standard normal distribution, LoConN, consists of elements drawn with $p = 0.95$ from the standard normal distribution, $N(0, 1)$, and with $p = 0.05$ from $N(5, 1)$. A scale-contaminated standard normal distribution, ScConN, consists of a distribution drawn with $p = 0.95$ from the standard normal distribution, $N(0, 1)$, and with $p = 0.05$ from $N(0, 10)$. Finally, a long-tail normal distribution, LongTailN, consists of standard normal deviates, u, multiplied by 1.25 with probability $p = 0.5$ if the absolute value of

u is greater than 1.5. The standard normal distribution was included as well to confirm the nominal significance levels of tests. The study used 1,000,000 Monte Carlo samples and was carried out for three sample sizes ($n = 20, 50, 100$) and three significance levels ($\alpha = 0.01, 0.05, 0.10$). Table 1, below, gives the empirical power of each test considered for a given alternative distribution.

The results suggest that the ranking of the tests depends on the alternative distribution. For some alternatives—most notably, for Student's t, LongTailN, and $\chi^2(5)$ distributions—the CS test performs worse than at least one of the rival tests. On the other hand, the CS test has in general slightly but noticeably higher power than the rivals for the uniform distribution, especially for smaller sample sizes and significance levels. It exhibits a visibly better performance for beta$(2, 2)$, beta$(2, 1)$, and Tukey(0.5) distributions for the smallest sample size. Moreover, it performs better than SW and SF tests in the cases of beta(2,2) and Tukey(0.5) distributions for all sample sizes and significance levels. These results confirm the conclusions of Romão, Delgado, and Costa (2010) that the CS test is particularly attractive in the case of symmetric alternative distributions. Other simulations, not shown for brevity's sake, revealed that for grouped data, the CS test usually performs significantly worse than the skewness and kurtosis test implemented in SK.

In all, the results in table 1 suggest that the CS test may be a useful complement to the already existing univariate normality tests for nonaggregated data in Stata.

Table 1. Empirical power for 1%, 5%, and 10% tests for selected alternative distributions

Distribution	Significance level											
	0.01				0.05				0.10			
	SK	SW	SF	CS	SK	SW	SF	CS	SK	SW	SF	CS
$\boldsymbol{n = 20}$												
Normal	0.012	0.010	0.010	0.010	0.049	0.051	0.051	0.050	0.098	0.101	0.101	0.100
Uniform	0.017	0.030	0.006	0.040	0.126	0.199	0.080	0.222	0.267	0.360	0.186	0.383
t(5)	0.115	0.094	0.109	0.094	0.217	0.187	0.219	0.181	0.296	0.257	0.299	0.248
$\chi^2(5)$	0.168	0.226	0.210	0.231	0.323	0.439	0.413	0.439	0.436	0.563	0.531	0.562
ScConN	0.480	0.474	0.483	0.474	0.532	0.521	0.530	0.519	0.566	0.555	0.564	0.553
LongTailN	0.032	0.025	0.030	0.026	0.102	0.092	0.109	0.089	0.173	0.157	0.183	0.151
t(20)	0.024	0.018	0.021	0.019	0.077	0.069	0.078	0.068	0.135	0.125	0.136	0.122
$\chi^2(20)$	0.091	0.101	0.098	0.103	0.201	0.244	0.235	0.242	0.294	0.349	0.334	0.346
beta(2, 2)	0.002	0.005	0.002	0.007	0.029	0.053	0.023	0.060	0.084	0.126	0.063	0.135
beta(2, 1)	0.016	0.084	0.049	0.094	0.091	0.304	0.211	0.317	0.201	0.467	0.352	0.478
LoConN	0.412	0.415	0.435	0.415	0.563	0.547	0.571	0.542	0.622	0.602	0.623	0.598
Tukey(0.5)	0.003	0.006	0.002	0.008	0.036	0.063	0.025	0.071	0.100	0.145	0.071	0.157
$\boldsymbol{n = 50}$												
Normal	0.011	0.010	0.010	0.010	0.049	0.051	0.052	0.050	0.099	0.100	0.102	0.100
Uniform	0.495	0.356	0.112	0.443	0.775	0.750	0.474	0.806	0.887	0.880	0.680	0.911
t(5)	0.245	0.230	0.262	0.216	0.385	0.356	0.417	0.333	0.479	0.436	0.506	0.408
$\chi^2(5)$	0.469	0.722	0.659	0.731	0.691	0.889	0.856	0.894	0.826	0.940	0.918	0.943
ScConN	0.818	0.810	0.816	0.808	0.840	0.832	0.840	0.830	0.854	0.847	0.854	0.845
LongTailN	0.043	0.042	0.051	0.037	0.128	0.126	0.163	0.111	0.211	0.200	0.257	0.178
t(20)	0.036	0.028	0.034	0.026	0.099	0.087	0.106	0.081	0.164	0.147	0.174	0.138
$\chi^2(20)$	0.252	0.361	0.322	0.364	0.446	0.590	0.556	0.592	0.585	0.702	0.671	0.705
beta(2, 2)	0.060	0.022	0.004	0.034	0.212	0.151	0.052	0.191	0.358	0.290	0.133	0.341
beta(2, 1)	0.065	0.530	0.326	0.580	0.284	0.841	0.703	0.866	0.607	0.928	0.844	0.941
LoConN	0.851	0.830	0.843	0.823	0.895	0.879	0.892	0.874	0.911	0.898	0.909	0.894
Tukey(0.5)	0.093	0.035	0.006	0.053	0.286	0.207	0.076	0.258	0.449	0.370	0.179	0.429
$\boldsymbol{n = 100}$												
Normal	0.011	0.010	0.010	0.011	0.049	0.049	0.052	0.052	0.102	0.098	0.103	0.103
Uniform	0.971	0.945	0.762	0.976	0.996	0.996	0.968	0.999	0.999	0.999	0.992	1.000
t(5)	0.413	0.418	0.469	0.395	0.582	0.563	0.632	0.530	0.676	0.639	0.710	0.604
$\chi^2(5)$	0.839	0.985	0.973	0.989	0.971	0.998	0.996	0.998	0.994	0.999	0.999	1.000
ScConN	0.966	0.964	0.966	0.963	0.972	0.970	0.972	0.969	0.975	0.973	0.975	0.972
LongTailN	0.058	0.066	0.086	0.058	0.168	0.180	0.235	0.154	0.270	0.272	0.345	0.236
t(20)	0.049	0.040	0.051	0.037	0.124	0.109	0.139	0.100	0.198	0.172	0.214	0.160
$\chi^2(20)$	0.536	0.752	0.702	0.772	0.776	0.901	0.877	0.911	0.895	0.945	0.930	0.951
beta(2, 2)	0.330	0.133	0.033	0.215	0.626	0.452	0.220	0.563	0.776	0.644	0.395	0.737
beta(2, 1)	0.286	0.976	0.915	0.987	0.895	0.998	0.992	0.999	0.986	1.000	0.998	1.000
LoConN	0.982	0.977	0.980	0.976	0.988	0.986	0.988	0.985	0.991	0.989	0.991	0.988
Tukey(0.5)	0.467	0.225	0.066	0.336	0.754	0.599	0.335	0.705	0.871	0.772	0.534	0.846

6 Conclusion

In this article, I introduced the `chens` command, which implements the CS test for the normality of nonaggregated or ungrouped data. Simulation results showed that for some alternative distributions, especially symmetric ones, the test has visibly higher empirical power than the official Stata normality tests.

7 References

Chen, L., and S. S. Shapiro. 1995. An alternative test for normality based on normalized spacings. *Journal of Statistical Computation and Simulation* 53: 269–288.

D'Agostino, R. B., A. Belanger, and R. B. D'Agostino, Jr. 1990. A suggestion for using powerful and informative tests of normality. *American Statistician* 44: 316–321.

Gould, W. 1992. sg3.7: Final summary of tests of normality. *Stata Technical Bulletin* 1: 114–115. Reprinted in *Stata Technical Bulletin Reprints*, vol. 5, pp. 10–11. College Station, TX: Stata Press.

Gould, W., and W. Rogers. 1991. sg3.4: Summary of tests of normality. *Stata Technical Bulletin* 1: 106–110. Reprinted in *Stata Technical Bulletin Reprints*, vol. 3, pp. 20–23. College Station, TX: Stata Press.

Romão, X., R. Delgado, and A. Costa. 2010. An empirical power comparison of univariate goodness-of-fit tests for normality. *Journal of Statistical Computation and Simulation* 80: 545–591.

Royston, P. 1991a. sg3.5: Comment on sg3.4 and an improved D'Agostino test. *Stata Technical Bulletin* 1: 110–112. Reprinted in *Stata Technical Bulletin Reprints*, vol. 3, pp. 23–24. College Station, TX: Stata Press.

———. 1991b. sg3.6: Response to sg3.3: Comment on tests of normality. *Stata Technical Bulletin* 1: 112–114. Reprinted in *Stata Technical Bulletin Reprints*, vol. 4, pp. 8–9. College Station, TX: Stata Press.

Seier, E. 2002. Comparison of tests for univariate normality. *Interstat* 17: 1–17.

Shapiro, S. S., and R. S. Francia. 1972. An approximate analysis of variance test for normality. *Journal of the American Statistical Association* 67: 215–216.

Shapiro, S. S., and M. B. Wilk. 1965. An analysis of variance test for normality (complete samples). *Biometrika* 52: 591–611.

About the author

Michal Brzezinski is an assistant professor in the Faculty of Economic Sciences at the University of Warsaw in Poland. His current research interests are power-law modeling of socioeconomic phenomena and measuring inequality and poverty in Central and Eastern European countries.