DigitalCommons@WayneState

Journal of Modern Applied Statistical Methods

Volume 5 | Issue 2 Article 18

11-1-2005

A Comparison of the Spearman-Brown and Flanagan-Rulon Formulas for Split Half Reliability under Various Variance Parameter Conditions

David A. Walker

Northern Illinois University, dawalker@niu.edu

Journal of Modern Applied Statistical Methods
November, 2006, Vol. 5, No. 2, 443-451

1538 – 9472/06/$95.00

A Comparison of the Spearman-Brown and Flanagan-Rulon Formulas for Split Half Reliability under Various Variance Parameter Conditions

David A. Walker
Northern Illinois University

Differences between the Spearman-Brown and Flanagan-Rulon formulas are examined when the variance parameters for two halves of a test had the following ratios: 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0 and also had a correlation between the two halves of a test at 1.00, .95, .90, .80, .70, .60, .50, .40, .30, .20, .10, .05. It was found that use of the Spearman-Brown formula to estimate the population ρ when the ratio between the standard deviations on two halves of a test is disparate, or beyond .9 to 1.1, was not warranted. Applied and theoretical examples are employed, as well as syntax for user application.

Key words: Split half reliability, Spearman-Brown, Flanagan-Rulon

Introduction

Examination of the difference between estimators of the population ρ for split-half reliability has been studied in the past (Charter, 1996, Cronbach, 1951, Feldt & Brennan, 1989, Kelley, 1942, Rulon, 1939, Stanley, 1971). To estimate the score reliability of a test split in half, the Spearman-Brown formula is the typical method used. Charter (1996) showed that:

$$r = 2r_{xy} / (1 + r_{xy}) \qquad (1)$$

where, r_{xy} = the correlation between the two halves of a test.

One major assumption with this formula is that the two halves of a test have equal variance parameters. Rulon (1939, attributed to Flanagan) proposed a split-half formula for use when the variance parameters where unequal. Charter (1996) showed that:

$$r = (4r_{xy}\, SD_x\, SD_y) / (SDx^2 + SDy^2 + 2r_{xy}\, SD_x\, SD_y) \qquad (2)$$

where, SD_x and SD_y are the standard deviations for the two halves of the test.

Cronbach (1951), and more recently Charter (1996), found that when the standard deviations of the two halves of a test are not equal, that is, the previously-noted major assumption guiding the Spearman-Brown formula is violated; its use will lead to an over-estimation of the reliability coefficient.

Purpose

The purpose of this research is to elaborate upon the work of Cronbach (1951) and Charter (1996) and strengthen the evidence in the literature that indicates that in most instances when unequal standard deviations for two halves of a test are present, regardless of the correlation between the two halves, the Flanagan-Rulon formula is the better estimator of ρ in a split-halves reliability situation. Thus, this research will build upon Cronbach's work and show, via various graphs and a detailed table, the differences between the Spearman-Brown formula and the Flanagan-Rulon formula when the variance parameters for two halves of a test have the following ratios (either greater or lesser): 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0 and also have a correlation between the two halves of a test at 1.00, .95, .90, .80, .70,

David Walker is Associate Professor of Educational Research and Assessment at Northern Illinois University. His research interests include structural equation modeling, effect sizes, factor analyses, predictive discriminant analysis, predictive validity, weighting, and bootstrapping.

.60, .50, .40, .30, .20, .10, .05. As well, SPSS (Statistical Package for the Social Sciences) syntax is provided in the Appendix section for users to create Tables 1 and 2 or calculate a Spearman-Brown or Flanagan-Rulon value given certain data.

Results

As was found by both Cronbach (1951) and Charter (1996), a large discrepancy between the variance parameters on the two halves of a test results in a substantial decrease in r. A small difference between standard deviations has the opposite effect, which is to be expected. Cronbach noted that when the ratio between the standard deviations of the two halves of a test are .9 to 1.1, the results from either the Spearman-Brown or the Flanagan-Rulon formulas are virtually the same and thus, the former formula should be used.

For example from empirical data, Gordon (1970) used the Musical Aptitude Profile (i.e., n = 190, under normal distribution, and with non-missing data) and estimated the split half reliability via the Spearman-Brown method. These data yielded a range of ratios between variance parameters of .9 to 1.1 (i.e., .949 to 1.141), which produced Spearman-Brown and Flanagan-Rulon values that were nearly identical. That is, there was no discrepancy between the Spearman-Brown formula results and Flanagan-Rulon estimates (i.e., the ratio of Spearman-Brown to Flanagan-Rulon ranged from 1.000 to 1.005). Thus, in this instance, the Spearman-Brown formula did not over-estimate the population ρ and was the proper choice.

Table 1. Comparison of SB and FR Estimates Using the Musical Aptitude Profile

r_{xy}	SD_1	SD_2	SD Ratio	SB	FR	Ratio of SB to FR
.695	9.57	10.08	.949	.820	.819	1.001
.600	10.01	9.59	1.044	.750	.750	1.000
.818	9.17	8.93	1.027	.900	.900	1.000
.600	7.66	7.29	1.051	.750	.749	1.001
.575	9.33	8.69	1.074	.730	.729	1.002
.786	7.71	6.76	1.141	.880	.876	1.005
.563	9.28	9.03	1.028	.720	.720	1.000
.600	10.20	9.14	1.116	.750	.747	1.004
.538	8.96	8.28	1.082	.700	.698	1.002
.818	7.95	7.22	1.101	.900	.898	1.003
.905	7.06	6.56	1.076	.950	.949	1.001

r_{xy} = correlation between two halves of a test, SD_1 & SD_2 = standard deviations for two halves of a test, SD Ratio = SD_1 / SD_2, SB = Spearman-Brown, FR = Flanagan-Rulon, Ratio of SB to FR = SB / FR

Note. Table adapted from Gordon (1970), where SB was reported only and not $r_{xy.}$ To calculate r_{xy} the inverse of the SB formula was applied, where $r_{xy} = (SB_r/2) / (1-r_{SB})$.

The current research extended Cronbach (1951) and Charter's (1996) work by creating a detailed table and figures which demonstrated that the range of the ratio between the standard deviations for two halves of a test could not be extended beyond .9 to 1.1. Table 2 shows that the deviation between the results yielded by the Spearman-Brown and Flanagan-Rulon formulas when the variance parameter ratio was .9 to 1.1 had a decrease in r of < 1% and a Spearman-Brown to Flanagan-Rulon ratio range difference < 1%. This was not the case, though, when the variance parameter ratio was .8 to 1.2. The Spearman-Brown to Flanagan-Rulon ratio was ≥ 1% starting at r_{xy} = .70, which generated a ratio = 1.010, and ended at r_{xy} = .05 or a ratio = 1.016.

Table 2. Comparison of SB and FR Formulas under Various r_{xy} and Variance Parameter Conditions

Ratio of SDs for Two Halves

r_{xy}	Two Halves	SB	FR	Ratio of SB to FR
1.00	2.000	1.000	.889	1.125
.95	2.000	.974	.864	1.128
.90	2.000	.947	.837	1.132
.80	2.000	.889	.780	1.139
.70	2.000	.824	.718	1.147
.60	2.000	.750	.649	1.156
.50	2.000	.667	.571	1.167
.40	2.000	.571	.485	1.179
.30	2.000	.462	.387	1.192
.20	2.000	.333	.276	1.208
.10	2.000	.182	.148	1.227
.05	2.000	.095	.077	1.238
1.00	1.900	1.000	.904	1.107
.95	1.900	.974	.878	1.109
.90	1.900	.947	.852	1.112
.80	1.900	.889	.795	1.118
.70	1.900	.824	.732	1.125
.60	1.900	.750	.662	1.133
.50	1.900	.667	.584	1.142
.40	1.900	.571	.496	1.152
.30	1.900	.462	.397	1.164
.20	1.900	.333	.283	1.178
.10	1.900	.182	.152	1.194
.05	1.900	.095	.079	1.203

Table 2. Continued

1.00	1.800	1.000	.918	1.089
.95	1.800	.974	.893	1.091
.90	1.800	.947	.866	1.094
.80	1.800	.889	.809	1.099
.70	1.800	.824	.746	1.105
.60	1.800	.750	.675	1.111
.50	1.800	.667	.596	1.119
.40	1.800	.571	.507	1.127
.30	1.800	.462	.406	1.137
.20	1.800	.333	.290	1.148
.10	1.800	.182	.157	1.162
.05	1.800	.095	.081	1.169
1.00	1.700	1.000	.933	1.072
.95	1.700	.974	.907	1.074
.90	1.700	.947	.881	1.076
.80	1.700	.889	.823	1.080
.70	1.700	.824	.759	1.085
.60	1.700	.750	.688	1.090
.50	1.700	.667	.608	1.096
.40	1.700	.571	.518	1.103
.30	1.700	.462	.415	1.111
.20	1.700	.333	.298	1.120
.10	1.700	.182	.161	1.131
.05	1.700	.095	.084	1.137
1.00	1.600	1.000	.947	1.056
.95	1.600	.974	.921	1.058
.90	1.600	.947	.894	1.059
.80	1.600	.889	.837	1.063
.70	1.600	.824	.772	1.066
.60	1.600	.750	.701	1.070
.50	1.600	.667	.620	1.075
.40	1.600	.571	.529	1.080
.30	1.600	.462	.425	1.087
.20	1.600	.333	.305	1.094
.10	1.600	.182	.165	1.102
.05	1.600	.095	.086	1.107

Table 2. Continued

1.00	1.500	1.000	.960	1.042
.95	1.500	.974	.934	1.043
.90	1.500	.947	.908	1.044
.80	1.500	.889	.850	1.046
.70	1.500	.824	.785	1.049
.60	1.500	.750	.713	1.052
.50	1.500	.667	.632	1.056
.40	1.500	.571	.539	1.060
.30	1.500	.462	.434	1.064
.20	1.500	.333	.312	1.069
.10	1.500	.182	.169	1.076
.05	1.500	.095	.088	1.079
1.00	1.400	1.000	.972	1.029
.95	1.400	.974	.947	1.029
.90	1.400	.947	.920	1.030
.80	1.400	.889	.862	1.032
.70	1.400	.824	.797	1.034
.60	1.400	.750	.724	1.036
.50	1.400	.667	.642	1.038
.40	1.400	.571	.549	1.041
.30	1.400	.462	.442	1.044
.20	1.400	.333	.318	1.048
.10	1.400	.182	.173	1.052
.05	1.400	.095	.090	1.054
1.00	1.300	1.000	.983	1.017
.95	1.300	.974	.957	1.018
.90	1.300	.947	.930	1.018
.80	1.300	.889	.872	1.019
.70	1.300	.824	.807	1.020
.60	1.300	.750	.734	1.022
.50	1.300	.667	.652	1.023
.40	1.300	.571	.558	1.025
.30	1.300	.462	.450	1.027
.20	1.300	.333	.324	1.029
.10	1.300	.182	.176	1.031
.05	1.300	.095	.092	1.033
1.00	1.200	1.000	.992	1.008
.95	1.200	.974	.966	1.009
.90	1.200	.947	.939	1.009
.80	1.200	.889	.881	1.009
.70	1.200	.824	.816	1.010
.60	1.200	.750	.742	1.010
.50	1.200	.667	.659	1.011
.40	1.200	.571	.565	1.012
.30	1.200	.462	.456	1.013
.20	1.200	.333	.329	1.014
.10	1.200	.182	.179	1.015
.05	1.200	.095	.094	1.016

Table 2. Continued

1.00	1.100	1.000	.998	1.002
.95	1.100	.974	.972	1.002
.90	1.100	.947	.945	1.002
.80	1.100	.889	.887	1.003
.70	1.100	.824	.821	1.003
.60	1.100	.750	.748	1.003
.50	1.100	.667	.665	1.003
.40	1.100	.571	.570	1.003
.30	1.100	.462	.460	1.003
.20	1.100	.333	.332	1.004
.10	1.100	.182	.181	1.004
.05	1.100	.095	.095	1.004
1.00	1.000	1.000	1.000	1.000
.95	1.000	.974	.974	1.000
.90	1.000	.947	.947	1.000
.80	1.000	.889	.889	1.000
.70	1.000	.824	.824	1.000
.60	1.000	.750	.750	1.000
.50	1.000	.667	.667	1.000
.40	1.000	.571	.571	1.000
.30	1.000	.462	.462	1.000
.20	1.000	.333	.333	1.000
.10	1.000	.182	.182	1.000
.05	1.000	.095	.095	1.000

Further, Figures 1 to 3 indicate that the Spearman-Brown formula's over-estimation tendencies become worse as the standard deviations for the two halves of a test become dissimilar and the correlation between the two halves moves into the moderate (i.e., $> .30 < .70$) and low ranges (i.e., $\leq .30$). Thus, in these circumstances, the use of the Flanagan-Rulon formula would provide the user with a more accurate estimation of the population ρ.

For example, in Figure 2, when the standard deviation ratio is a moderate 1.5 and the correlation between the two halves of a test is also a moderate .600, use of the Spearman-Brown formula yields an $r = .750$. The Flanagan-Rulon formula produces an $r = .713$, or a discrepancy of nearly 4% showing a ratio of the Spearman-Brown being 1.052 times higher than the Flanagan-Rulon estimate. Thus, the Flanagan-Rulon formula in this case is the more accurate of the two in terms approximating the population ρ. Looking at Figure 3, when the standard deviation ratio is a large 1.9 and the correlation between the two halves of a test is a low .300, use of the Spearman-Brown formula yields an $r = .462$. The Flanagan-Rulon formula produces an $r = .397$, or an even more prominent discrepancy of 6.5% and a ratio of the Spearman-Brown formula yielding results 1.164 times higher than the Flanagan-Rulon estimation.

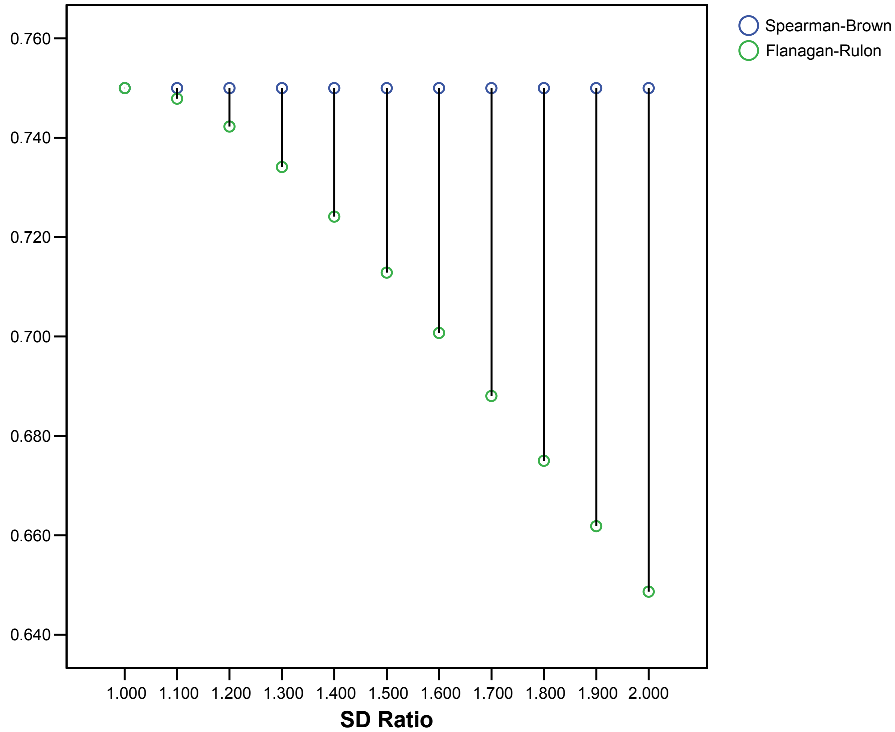

Figure 2. Example when r_{xy} = .600 and SB = .750

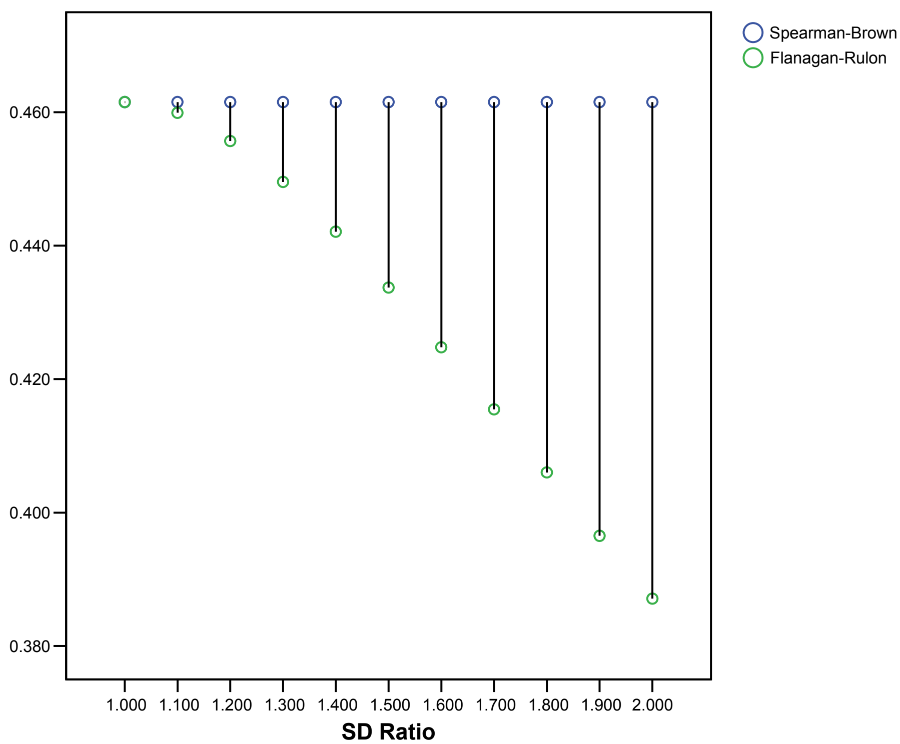

Figure 3. Example when r_{xy} = .300 and SB = .462

Application

The SPSS syntax found in the Appendix serves as a check on the variance parameter ratio range to determine which split half reliability formula to employ. The user types into the Begin Data section of the syntax the two standard deviation values for two halves of a test followed by the correlation between the two halves. The syntax is run and produces the following values: r_{xy}, SD_1, SD_2, the ratio of SD_1 to SD_2, Spearman-Brown, Flanagan-Rulon, and the ratio of Spearman-Brown to Flanagan-Rulon. From these results, the user can determine if the ratio between the standard deviations of the two halves of a test are .9 to 1.1, which would also produce a Spearman-Brown to Flanagan-Rulon ratio $< 1\%$, signifying use of the Spearman-Brown (i.e., no over-estimation of the population ρ). If the ratio range were beyond this threshold, the Spearman-Brown to Flanagan-Rulon ratio would be $\geq 1\%$, which would indicate that the Flanagan-Rulon formula would be the more accurate estimator to use. Future research examining the ratio range of the Spearman-Brown and Flanagan-Rulon formulas should include their performance with empirical data under biased distributional situations, with various sample sizes, and under an assortment of missing data conditions.

References

Charter, R. A. (1996). Note on the underrepresentation of the split-half reliability formula for unequal standard deviations. *Perceptual and Motor Skills, 82*, 401-402.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334.

Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R.H. Linn (Ed.), *Educational Measurement* (3rd ed.). New York: Macmillan.

Gordon, E. (1970). Taking into account musical aptitude differences among beginning instrumental students. *American Educational Research Journal, 7*, 41-53.

Kelley, T. L. (1942). The reliability coefficient. *Psychometrika, 7*, 75-83.

Rulon, P. J. (1939). A simplified procedure for determining the reliability of a test by split-halves. *Harvard Educational Review, 9*, 99-103.

Stanley, J. C. (1971). Reliability. In R.L. Thorndike (Ed.), *Educational Measurement* (2nd ed.). Washington, DC: American Council on Education.

Appendix. Syntax for Calculating Tables 1 and 2 or any SB or FR Value

```
**********************************************************************
Copyright David A. Walker, 2005
Contact dawalker@niu.edu
Northern Illinois University, 101J Gabel, DeKalb, IL 60115
                    **APA 5th Edition Citation**
Walker, D. A. (2005). Comparison between the Spearman-Brown and Flanagan-Rulon
          split half reliability formulas [Computer program]. DeKalb, IL: Author.
**********************************************************************.

DATA LIST LIST /sd1 sd2 (2f9.2) r (f9.3).

**NOTE: Below, insert the SD1 and SD2 values and the r value.**


BEGIN DATA
15 10 .95
14 10 .90
13 10 .80
12 10 .70
11 10 .60
10 10 .50
END DATA.
COMPUTE FR1 = (4*r)*(sd1)*(sd2).
COMPUTE FR2 = (sd1**2)+(sd2**2)+(2*r)*(sd1)*(sd2).
COMPUTE FR = FR1/FR2.
COMPUTE SBPF2 = (2*r)/(1+r).
COMPUTE SDRATIO = sd1/sd2.
COMPUTE RATIO = SBPF2/FR.
EXECUTE.
FORMAT FR TO RATIO  (f9.3).
VARIABLE LABELS sdratio 'SD Ratio'/ratio 'Ratio of SB to FR'/sbpf2 'Spearman-Brown'/r
'Correlation Between the Two Halves of a Test (rxy)'/sd1 'Standard Deviation for Test 1'/sd2 'Standard
Deviation for Test 2'/fr 'Flanagan-Rulon'/.
REPORT FORMAT=LIST AUTOMATIC ALIGN (LEFT)
MARGINS (*,120)
  /VARIABLES= r sd1 sd2 sdratio sbpf2 fr ratio
  /TITLE "Comparison Between Spearman-Brown and Flanagan-Rulon".
```