11-1-2009

Intermediate r Values for Use in the Fleishman Power Method

Julie M. Smith

Wayne State University, ax7955@wayne.edu

Intermediate r Values for Use in the Fleishman Power Method

Julie M. Smith
Wayne State University

Several intermediate *r* values are calculated at three different correlations for use in the Fleishman Power Method for generating correlated data from normal and non-normal populations.

Key words: Fleishman Power Method, Monte Carlo simulation, correlation.

Introduction

As Headrick and Sawilowsky (1999) observed, "Monte Carlo simulations requiring correlated data from normal and non-normal populations are frequently used to investigate the small sample properties of competing statistics, or the comparison of estimation techniques" (p. 25). Fleishman (1978) introduced the power method for simulating univariate non-normal distributions. This method allows for the systematic control of skew (γ_1) and kurtosis (γ_2) needed in Monte Carlo studies. Fleishman power method models are able to approximate a variety of distributions and require few inputs: a normal random number generator, three constants and an intermediate correlation (Headrick & Sawilowsky, 2000).

A normal random number generator is available as a FORTRAN subroutine, Headrick and Sawilowsky (2000) calculated and provided required constants for various distributions (see Table 1), and the intermediate correlation, *r*, is calculated using the formula

$$r_{xy} = r^2(b^2 + 6bd + 9d^2 + 2a^2r^2 + 6d^2r^4) \text{, (1)}$$

where *a*, *b* and *d* are constants and r_{xy} is the correlation to which all data will be set. The formula, when solved, results in the graph of a

parabola. After graphing, the intermediate *r* value is obtained by determining the point at which the positive horizontal axis intercept is at zero. Establishing the intermediate *r* values may be accomplished via use of a graphing calculator. Values provided in this brief report were obtained using a Texas Instruments (TI) 83-Plus Graphing Calculator by following a five step procedure:

1. Clear all registers and engage the function editor;
2. Enter formula (1) using appropriate constants and desired correlation;
3. Graph the parabola;
4. Use the trace function to position the cursor close to Y = 0 on the positive X-axis;
5. Enlarge the graph using the zoom feature to obtain a precise reading of the positive X value at the point where Y = 0.

Although simple, the process is time-consuming; for this reason several intermediate *r* values have been calculated at three different correlations (See Table 2).

Example

To create correlated data pairs (X, Y) at 0.70 from an exponential distribution (Chi-square, df = 2) with $\gamma_1 = 2$ and $\gamma_2 = 6$, using the constants from Table 1, equation (1) would be as follows:

$$r_{xy} = r^2[(.8263)^2 + (6)(.8263)(.02271) + (9)(.02271)^2 + (2)(-.3137)^2r^2 + (6)(.02271)r^4]$$

$$r_{xy} = r^2[(.68278) + (.11259) + (.004642) + (.19682)r^2 + (.00309)r^4]$$

Julie M. Smith is a Ph.D. Candidate in the College of Education, Department of Educational Evaluation and Research. Email: ax7955@wayne.edu.

$$r_{xy} = r^2[(.8000)+(.19682)r^2+(.00309)r^4]$$

$$0 = .8000\ r^2+.19682r^4+.00309r^6 - 0.70$$

$$X = a + bX_i + (-a)X_i^2 + dX_i^2 \qquad (4)$$

The positive solution using the stated procedure is $r = .859998$. This intermediate r value is placed in the following two equations:

$$Y = a + bY_i + (-a)Y_i^2 + dY_i^2. \qquad (5)$$

$$x_i = rz_1 + \sqrt{1-r^2}z_2 \qquad (2)$$

and

$$y_i = rz_1 + \sqrt{1-r^2}z_3 \qquad (3)$$

The algorithms above produce standardized data centered around $\mu = 0$ and $\sigma = 1$. To realign the values to the χ^2 distribution with df $= 2$, a simple transformation is performed so that $\mu = 2$ and $\sigma = 2$ as follows:

$$\chi_x^2 = (2)(X)+2 \qquad (6)$$

where z_1, z_2 and z_3 are randomly selected standard normal z scores (generated using a random number generator). The data resulting from these equations are not the final correlates, but represent intermediate standard normal variates that will be used to generate the desired correlated data, thus x_i and y_i and the constants appropriate to the distribution are next substituted into the Fleishman equation to produce the final correlates as follows:

and for the Y correlate,

$$\chi_x^2 = (2)(Y)+2 \qquad (7)$$

The last step is optional, because computed values are accurate for the distribution. It is only necessary to perform this step if it is desirable to have values commonly found in the tables for the distribution of interest, such as χ^2 (df $= 2$) in the example.

Table 1: Fleishman Power Constants for Various Distributions*

Distribution	Skew	Kurtosis	Constants		
	γ_1	γ_2	a	b	d
Chi-square (df=1)	$\sqrt{8}$	12	$-.5207$	.6146	.02007
Exponential/Chi-square (df=2)	2	6	$-.3137$	.8263	.02271
Chi-square (df=3)	1.633	4	$-.2595$	.8807	.01621
Chi-square (df=4)	$\sqrt{2}$	3	$-.2269$	.9089	.01256
Chi-square (df=8)	1	1.5	$-.1632$	.9531	.0060
Normal	0	0	0	1	0
Cauchy/t (df=1)	0	25	0	.2553	.2038
t (df=3)	0	17	0	.3938	.1713
t (df=7)	0	2	0	.8357	.05206
Laplace/Double Exponential	0	3	0	.7284	.0679

*From Headrick and Sawilowsky (2000), p. 427.

Table 2: Intermediate *r* Values for Various Distributions at Correlations
0.70, 0.80 and 0.90

Distribution	Intermediate *r* Values at Correlation:		
	0.70	0.80	0.90
Chi-square (df=1)	.88909	.92960	.96633
Exponential/Chi-square (df=2)	.85998	.91319	.95973
Chi-square (df=3)	.79989	.85067	.89771
Chi-square (df=4)	.87870	.93855	.99461
Chi-square (df=8)	.84466	.90058	.95271
Normal	.83666	.89443	.94868
Cauchy/t (df=1)	.88121	.92549	.96472
t (df=3)	.86665	.91814	.96118
t (df=7)	.84006	.89697	.95008
Laplace/Double Exponential	.84248	.89877	.95110

References

Headrick, T. C., & Sawilowsky, S. S. (1999). Simulating correlated multivariate nonnormal distributions: Extending the Fleishman power method. *Psychometrika*, *64*(*1*), 25-35.

Headrick, T. C., & Sawilowsky, S. S. (2000). Weighted simplex procedures for determining boundary points and constants for the univariate and multivariate power methods. *Journal of Educational and Behavioral Statistics*, *25*(*4*), 417-436.

Sawilowsky, S. S. & Fahoome, G. (2003). *Statistics through Monte Carlo simulation with fortran*. Oak Park, MI: JMASM.