DOCUMENT RESUME

ED 427 032 TM 029 409

AUTHOR	Wind, Brian M.; Kim, Jwa K.
TITLE	The Effect of Sample Size and Variance on the Johnson-Neyman Technique.
PUB DATE	1998-11-00
NOTE	29p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (27th, New Orleans, LA, November 4-6, 1998).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Monte Carlo Methods; *Regression (Statistics); *Sample Size; Simulation
IDENTIFIERS	*Johnson Neyman Technique; *Variance (Statistical)

ABSTRACT

The Johnson-Neyman (J-N) technique (P. Johnson and N. Neyman, 1936) is used to determine areas of significant difference in a criterion variable between two or more groups in situations of linear regression. In using this technique, researchers have encountered difficulties with results, possibly related to the J-N technique's sensitivity to violations of certain assumptions and conditions. For this study, Monte Carlo simulations were performed to determine the effect that sample size and variance have on the J-N technique. The simulations examined the hypothesis that unequal ratios of sample size and variance between two groups may create anomalies in the results of the J-N computation. The results do not show anomalies in the output, and further show that the J-N technique produces wider regions of significance as the total sample size increases. The size of variance ratios, as well as the equality of variance and sample size ratios, did not seem to affect the results dramatically. Appendixes contain the Statistical Analysis System program for the simulations and the results of the simulations in table form. (Contains 17 references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 427 032

The Effect of Sample Size and Variance on the Johnson-Neyman Technique

Brian M. Wind

and

Jwa K. Kim

Middle Tennessee State University

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Jwa Kim

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

TM029409

A paper presented at the annual meeting of the Mid-South Educational Research Association, New Orleans, Louisiana, November, 1998.

ERIC

BEST COPY AVAILABLE

Abstract

The Johnson-Neyman (J-N) technique is used to determine areas of significant difference in a criterion variable between two or more groups in situations of linear regression. In utilizing this technique, researchers have encountered difficulties with results, possibly related to the J-N technique's sensitivity to violations of certain assumptions and conditions. For this study, Monte Carlo simulations were performed to determine the effect that sample size and variance have on the J-N technique. The simulations examined the hypothesis that unequal ratios of sample size and variance between two groups may create anomalies in the results of the J-N computation. The results did not show anomalies in the output, and further showed that the J-N technique produces wider regions of significance as the total sample size increases. The size of variance ratios, as well as the equality of variance and sample size ratios did not seem to affect the results dramatically.

ERIC

The Effect of Sample Size and Variance on the Johnson-Neyman Technique

The Johnson-Neyman technique (Johnson & Neyman, 1936) is a procedure that is employed to determine regions of significant difference in a criterion variable, which exist as a function of one or more predictor variables in a linear regression situation. The regions of significance are areas in which criterion variable values differ significantly between groups (on the y-axis), and are defined in terms of the predictor variable (on the x-axis). To better understand the Johnson-Neyman technique and its use, consider the following hypothetical example. Researchers interested in positive reinforcement and its effect on academic performance design a study. They hypothesize that there is a direct correlation between the amount of positive reinforcement given to students and the amount of grade improvement the students will show. In this case the predictor variable would be positive reinforcement and the criterion variable would be academic performance (grades). Thus, academic performance could be plotted as a function of positive reinforcement.

Suppose, however, that the researchers further hypothesized that females will show a significantly higher improvement (due to a high amount of positive reinforcement) in grades than males, and that they will show a significantly smaller amount of improvement than males when little positive reinforcement is given. In this case separate regression lines for males and females could be plotted, again showing academic performance as a function of positive reinforcement. The researchers notice that their graph shows male and female grades are virtually the same when a moderate amount of positive reinforcement is given. Males and females differ, however, in grades outside of this area of moderate reinforcement. In fact, the graph proves their hypothesis that females would show more improvement with increased positive reinforcement and less improvement with a decreased amount of positive reinforcement, with the male and female regression lines forming an "X" pattern (See Figure 1). At what points along the predictor variable axis (x-axis) are the two lines significantly different with respect to the criterion variable (y-axis)? This is a question that is best answered by the Johnson-Neyman (J-N) technique. Using the J-N technique will answer the question of how much positive reinforcement produces significant differences in the amount of improvement in academic performance. This is done by the J-N technique's provision of lower and upper limits of a region of nonsignificance, outside of which region criterion scores are significantly different (See Figure 2).

The J-N technique is often used as a superior alternative to the analysis of covariance (ANCOVA). This is due mainly to the fact that, while the J-N technique shares almost all of ANCOVA'S assumptions, it can be used effectively when ANCOVA's assumption of homogeneity of the slope of regression is not met. With the exception of the homogeneity of slopes assumption, the J-N technique shares all of the assumptions of ANCOVA (Pigache & Graham, 1976). These assumptions are:

1) The regression of Y on X is a linear relationship.
2) Variances for each value of X are equal (homoscedasticity).
3) Within-group variances are equal.
4) Homogenous variance exists across treatment groups.
5) The measurement of the covariate is error-free.
6) Carry-over effects related to measurement are independent.
7) In the regression of Y on X, there is regression toward the mean (normality).
8) Departures from the regression line are independent.

The J-N technique, as noted by Huitema (1980), is also an optimal choice to either a two factor analysis of variance (ANOVA) or simple main effects tests because of the J-N

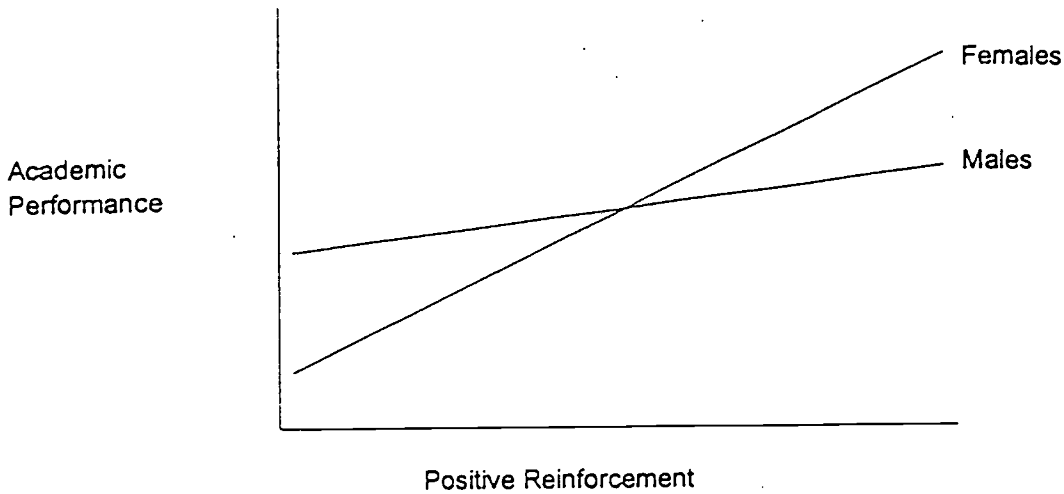

Figure 1. Academic performance plotted as a function of positive reinforcement.

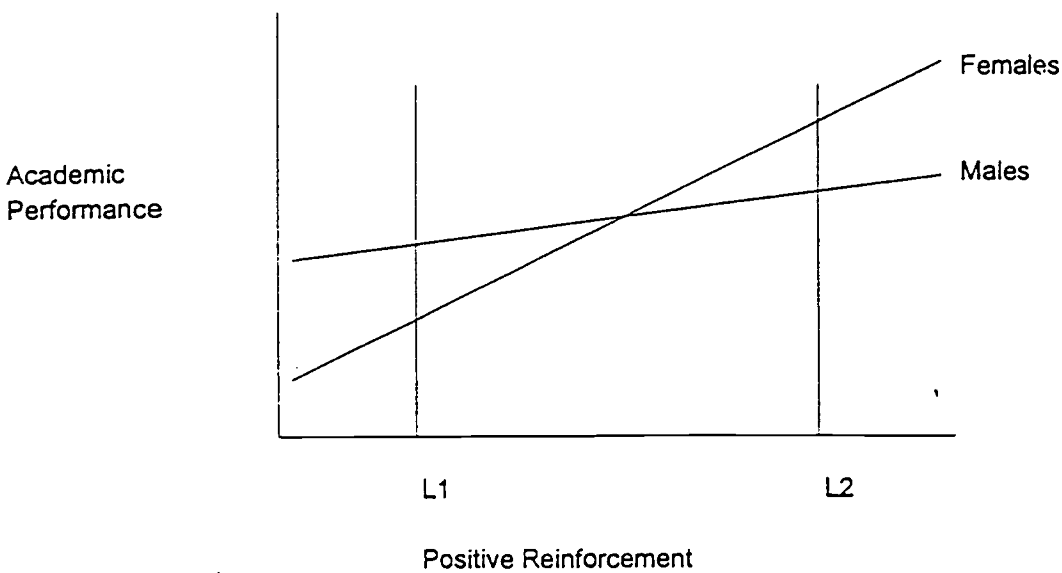

Figure 2. The J-N technique's provision of a region of nonsignificance, where L1 and L2 are lower and upper limits, respectively.

technique's superior power and its allowance of the examination of treatment effects for different covariate levels. Given the J-N technique's value in such cases, it has been used in a variety of applications since its development in the 1930's.

One valuable application of the J-N technique has been in studies where researchers are interested in pre- and post-test differences in the effectiveness of alternative teaching methods (Ceurvorst, 1979). A study of this type by Keim-Abbott and Abbott (1977) examined the interaction of student characteristics (trait) and teaching method (treatment). Upon analysis of the data a significant trait-by-treatment interaction (TTI) was found. Based on this significant TTI, the J-N technique was then utilized to determine the regions of significant difference in the criterion variable as a function of the predictor variable, and results indicated that a teacher-paced method of learning produced significantly higher levels of achievement in students than a self-paced method of learning. Another area in which the J-N technique has been similarly applied is the area of learning disability research. One such study examined the advantages of the J-N technique (with respect to accuracy) over ANCOVA in analyzing regression differences in standardized scores of learning disabled and normal children (McLeskey & Rieth, 1982). Results of this research indicated that the J-N technique was in fact more accurate than ANCOVA at depicting areas of significant difference.

Researchers have also found the J-N technique useful in determining differential prediction for both race and gender in academic and technical programs. In one study done to examine differential prediction of selection for military programs by race, the J-N technique was found to be highly effective in reducing bias that may occur using other methods of regression analysis (Houston & Novick, 1987). Related to gender, Gamache and Novick (1985) showed the J-N technique to produce effective identifications of regions of significance that would otherwise (using other techniques) exclude students from selection to certain programs based on gender. Finally, the J-N technique has been of great assistance to those doing research in clinically related fields. A study by Pigache and Graham (1976) noted the J-N technique's particular usefulness in helping to examine the contribution of baseline levels to treatment effects, as in the baseline level of a drug affecting the level of the criterion variable.

In the many years since the J-N technique's development, this procedure has undergone numerous changes. One area of change has been the attempt of researchers to bring the J-N technique into the computer arena. As noted by Ceurvorst (1979), the J-N technique has not been widely taught or used over the years because of its requirement of lengthy computations. The past unavailability of the J-N technique in most statistical software thus left researchers with three options: enduring the computational tasks involved in using the J-N technique, finding alternative methods of analyzing regression slopes, or developing computer programs to put this technique to work in their studies. Some opted to undertake the job of developing a computer program for the J-N technique to use in their research.

The programs written vary along certain lines, such as the type of system on which they may be used, language (e.g. FORTRAN), and the number of variables that may be used in the program. Karpman has designed several J-N programs, one of which is an ANCOVA program that proceeds to the J-N technique when it is determined that there is heterogeneity of slopes. This program is limited to six groups, and output consists of the results of multiple comparisons and a definition of regions of significance

(Karpman, 1980). Karpman later devised a program for use with SPSS-X or SAS (1986). This program requires input of slope, intercept, covariate sums of squares, covariate mean, sample size, critical F-value, and the pooled residual sum of squares for the criterion for each group. The output yields regions of significance and can be plotted using the PLOT procedure on SPSS-X or SAS (Karpman, 1986). Both of these Karpman programs were designed for use in the one covariate case. Earlier, Karpman (1983) wrote a program for use with SPSS or BMDP that could also be used in the two covariate case; however, this program can not produce exact regions of significance. Instead, regions of significance are approximated through the use of confidence tests at each value of the covariate. Such difficulties as this have most often left researchers to design J-N programs for use in the one covariate case.

Scialfa (1987) developed a BASIC program for the J-N technique. This program requires the input of the following parameters for each group: predictor mean, standard deviation, sample size, residual sums of squares, the least estimate for slope, and the least squares estimate for intercept. These data may be calculated using virtually any major statistical software package. In addition, the critical F value is needed, with 1 and N-4 degrees of freedom (df), where N = combined sample size (Scialfa, 1987). The output for this program consists of values for the point of intersection and the two regions of significance. This program is designed to employ the J-N technique in cases where there are two groups and one continuous predictor.

Computer programs for the J-N technique have also been written in FORTRAN. Ceurvorst (1979) devised a FORTRAN V program that requires input in much the same manner as SPSS. Separate "control cards" are entered for variable names, format of input, number of groups, and an alpha level (per-comparison) - required for constructing regions of significance. "Data cards" follow, and an output is produced that gives the following information for each group: number of cases, mean and standard deviation (for both variables), correlation and regression coefficients, and regression line intercept on the dependent variable axis. In addition, the F-ratio for homogeneity of regression (with degrees of freedom and probability) are shown, along with a printout of the point of intersection for the regression lines (if one exists). This program is designed for J-N comparisons between as many as 20 groups, with one independent variable and one dependent variable (Ceurvorst, 1979).

Another FORTRAN program was designed by Pigache and Graham (1976) for the purposes of researching pre- and post-treatment effects of drugs, with consideration of baseline levels. In this program data are entered related to minimum and maximum pre-treatment values, the number of treatments, and the number of participants. A second area of data entry contains pre-treatment values for treatments 1 and 2 and participants through value *n*. These data are immediately followed in the program by the post-treatment numbers for all participants and treatments. From this, two sets of output are obtained. The first output yields the pre- and post-treatment means for each regression line, the correlation coefficient, the t-value, the slope, and the intercept on the criterion variable axis. The second output shows information about the treatments and their comparisons, as well as related significance levels (Pigache & Graham, 1976). It is important to note another aspect of this study - the researchers' modification of the J-N technique. The technique was modified so as not to assume equal variances. This modification consisted of two major stages: estimating the regression line of post-treatment on pre-treatment level separately for each of the treatments together with the slope's standard error, and superimposing a series of intercepts along the pre-treatment axis of values (Pigache & Graham, 1976). For the purposes of the study, this appears to be an effective J-N technique modification.

Although similar modifications of the J-N technique have been done, overall research on the J-N technique has been somewhat limited. Hunka (1995) has done research in the area of the J-N technique, pointing out that this technique has limitations related to many things: the number of groups and covariates, the nature of the regression model, the use of confidence intervals or confidence regions in determining boundaries for the covariate, and whether or not actual regions of significance can be obtained in the case of two or more covariates. As a result, Hunka developed an alternative to the J-N technique by "cast[ing] the analysis into the form of a general linear model," representing the unknown boundary values for the covariate in symbolic form on a contrast matrix, and solving the matrix equation for the sum of squares (to determine significant group effects) using the Mathematica software system (Hunka, 1995). Hunka also provides examples of how this technique can be used effectively in the case of more than one covariate.

Other research on the J-N technique was done by Chou and Huberty (1992). In this study, the authors set out to test the original J-N technique compared to the modification designed by Potthoff (1964). Potthoff's modification of the J-N technique was done so as to create simultaneous regions of significance, and similar work was later done by Rogosa (1981). Chou and Huberty's research consisted of examining the robustness of the original and modified J-N techniques in relation to controlling simultaneous error rates. Monte Carlo simulations were performed (manipulating conditional distribution shape, conditional variance ratio, and sample size ratio) to evaluate the techniques' relative ability to control the Type I error (identifying a region of significance when in reality none exists) and Type III error (when the identified region of significance contains a covariate value at which the expected criterion score for two groups is equal) (Chou & Huberty, 1992). Results showed that the modified J-N technique effectively controlled the Type I error rate and both techniques controlled for Type III error; however, it was found that the original J-N technique was not effective at controlling for Type I error under most conditions (Chou & Huberty, 1992). The researchers then used ANOVA to determine the effect of the three factors that were manipulated (in the Monte Carlo simulations) on the respective error rates. The following was determined: virtually no effect for conditional distribution shape, and virtually no effect for conditional variance ratio when sample sizes were homogeneous. When sample sizes were different, a large conditional variance ratio effect appeared (Chou & Huberty, 1992).

The current problem is related to previous work done by Lavender and Kim (1996). This study examined differences in predicted (based on ACT-C scores) college grade point average (CGPA) across gender. In order to determine what, if any, significant gender differences existed, the authors performed analyses using the J-N technique. This technique was then extended to examine gender differences in CGPA across various college majors. This was done in an attempt to determine whether or not differential course selection plays a part in gender differences across CGPA. Findings indicated that differences in CGPA did exist across gender for the entire sample, but that these differences were minimized when controlling for college major (Lavender & Kim, 1996).

In performing the analyses for this study using the J-N technique, Lavender and Kim encountered some problems. Regions of significance were found using the J-N technique in the college majors of Finance and Psychology; however, these results contained anomalies. For both majors the nonsignificant regions identified by the J-N technique did not contain the intersection of male and female regression lines. This is then taken to mean that the intersection of the male and female regression lines occurs within a region of significance; however, this can not be true since the male and female regression lines can not differ significantly at their point of intersection (See Figure 3). These anomalies fall into the category of what was defined above as Type III error. It was speculated that Lavender and Kim's (1996) resultant anomalies are related to the J-N technique's condition of equal sample sizes and assumption of equal variances between two groups. When examining how the sample size and variance can affect results in such a way, one must consider the formulas involved in the utilization of the J-N technique:

$$X_{L1} = \frac{-B-\sqrt{B^2-AC}}{A}$$

$$X_{L2} = \frac{-B+\sqrt{B^2-AC}}{A}$$

where X_{L1} and X_{L2} = upper and lower limits of the region of nonsignificance, and

$$A = \frac{-F}{N-4}(ss_{res})\left(\frac{1}{\Sigma x^2_1}+\frac{1}{\Sigma x^2_2}\right)+(b_1-b_2)^2 \quad (1)$$

$$B = \frac{F}{N-4}(ss_{res})\left(\frac{\bar{X}_1}{\Sigma x^2_1}+\frac{\bar{X}_2}{\Sigma x^2_2}\right)+(a_1-a_2)(b_1-b_2) \quad (2)$$

$$C = \frac{-F}{N-4}(ss_{res})\left(\frac{N}{n_1 n_2}+\frac{\bar{X}^2_1}{\Sigma x^2_1}+\frac{\bar{X}^2_2}{\Sigma x^2_2}\right)+(a_1-a_2)^2 \quad (3)$$

In these formulas, F is equal to the tabled F ratio (with 1 and N-4 df). N is equal to the total number of participants (number of participants in both groups combined), while n_1 and n_2 are equal to the number of participants in groups 1 and 2, respectively. The ss_{res} represents the residual sum of squares for the regression analysis with both groups combined using a dummy variable, or pooled ss_{res} for the separate regression line of both groups. $\overline{X}_1$ and $\overline{X}_2$ are the means for groups 1 and 2 (respectively) of the independent (X) variable. The values represented by b_1 and b_2 are the regression line slope coefficients for groups 1 and 2, respectively; a_1 and a_2 are the regression line intercepts for groups 1 and 2, respectively. Finally, $\sum x^2_1$ and $\sum x^2_2$ are equal to the sums of squares for groups 1 and 2, respectively (McLeskey & Rieth, 1982). As can be seen in equations (1)-(3), the variance plays an important role in J-N computation. The population variance is directly related to the sample sum of squares, since the sample variance is an unbiased estimate of the population variance and the sum of squares is the product of the sample variance and degrees of freedom. Thus, the variance's close relationship to the sum of squares may affect the results of computation when using the J-N technique, if the assumption of equal variances is violated.

With respect to the equal sample sizes property, it can also be seen how violation of this condition could potentially affect results. In the above-listed equation (3), the inclusion of n_1 and n_2 in the computation creates the possibility that unequal sample sizes could cause anomalies, as in the case of Lavender and Kim's study (1996). Unequal values of n_1 and n_2 could also produce flawed results considering n's involvement in the

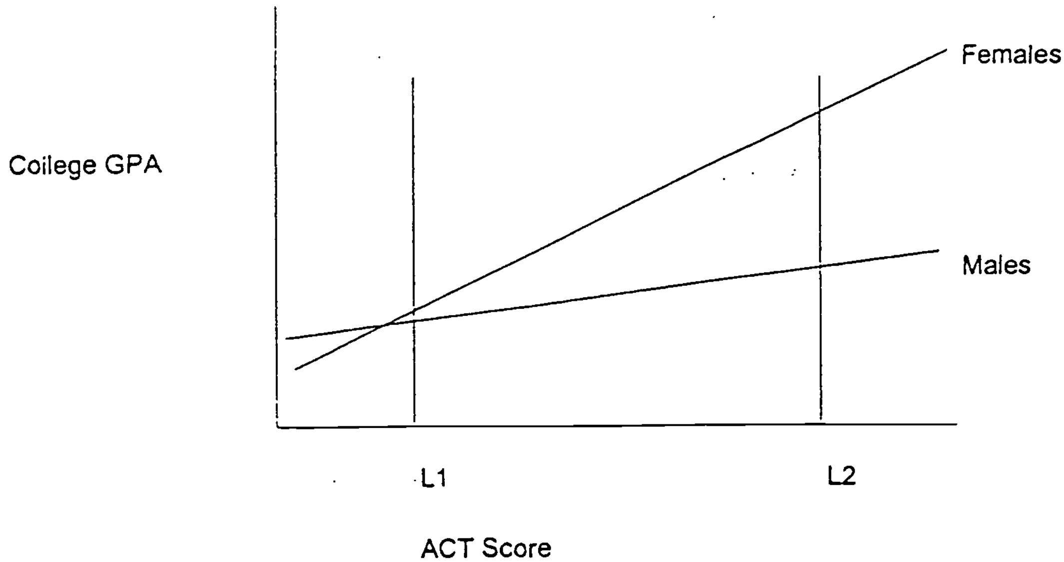

Figure 3. Type III error produced by the J-N technique, as in the case of Lavender and Kim's (1996) research.

formula for the sum of squares (in conjunction with the sample variance), as can be seen in the following:

$$\sum x^2 = s^2(df),$$

where df = n-1.

The purpose of this study was to examine exactly how manipulation of ratios related to sample size and variance affects results when using the J-N technique. The advantage of this approach over other approaches (such as using a hypothetical data set) is that it provides more accurate information with respect to what component(s) of the formulas may affect the outcome of the J-N technique.

Method

Monte Carlo simulations were performed to manipulate various situations of sample size and variance. First, the J-N technique was performed on data to simulate equal sample size ratios, these ratios being paired with both equal and unequal ratios of variance. The J-N technique was then performed again with unequal sample size ratios, which were in turn paired with both equal and unequal variance ratios. Equal ratios for both sample size and variance were represented by the following values: 10:10, 20:20, 30:30, 50:50. Unequal ratios were: 10:20, 10:30, 10:50, 50:10, 30:10, 20:10 (See Figure 4 for a more detailed design of the study). In performing these analyses, areas in the computational formula that do not include the sample sizes and the sums of squares (affected directly by variance) were kept constant, as shown here:

$$A = \frac{-F}{N-4}(ss_{res})\left(\frac{1}{\Sigma x^2_1} + \frac{1}{\Sigma x^2_2}\right) + (b_1 - b_2)^2 \qquad (1)$$

Sample Size Ratio	Variance Ratio	
Equal		
10:10	Equal	10:10, 20:20, 30:30, 50:50
	Unequal	10:20, 10:30, 10:50
20:20	Equal	10:10, 20:20, 30:30, 50:50
	Unequal	10:20, 10:30, 10:50
30:30	Equal	10:10, 20:20, 30:30, 50:50
	Unequal	10:20, 10:30, 10:50
50:50	Equal	10:10, 20:20, 30:30, 50:50
	Unequal	10:20, 10:30, 10:50
Unequal		
10:20	Equal	10:10, 20:20, 30:30, 50:50
	Unequal	10:20, 10:30, 10:50
10:30	Equal	10:10, 20:20, 30:30, 50:50
	Unequal	10:20, 10:30, 10:50
10:50	Equal	10:10, 20:20, 30:30, 50:50
	Unequal	10:20, 10:30, 10:50
50:10	Equal	10:10, 20:20, 30:30, 50:50
	Unequal	10:20, 10:30, 10:50
30:10	Equal	10:10, 20:20, 30:30, 50:50
	Unequal	10:20, 10:30, 10:50
20:10	Equal	10:10, 20:20, 30:30, 50:50
	Unequal	10:20, 10:30, 10:50

Figure 4. Manipulation of sample size and variance ratios.

$$B = \frac{F}{N-4}(ss_{res})(\frac{\overline{X}_1}{\Sigma x^{2_1}} + \frac{\overline{X}_2}{\Sigma x^{2_2}}) + (a_1 - a_2)(b_1 - b_2) \quad (2)$$

$$C = \frac{-F}{N-4}(ss_{res})(\frac{N}{n_1 n_2} + \frac{\overline{X}^{2_1}}{\Sigma x^{2_1}} + \frac{\overline{X}^{2_2}}{\Sigma x^{2_2}}) + (a_1 - a_2)^2 \quad (3)$$

where areas in bold indicate areas in which values were held constant. Regression line equations of y = 1 + x and y = 3 - x were used for groups 1 and 2, respectively; therefore, the intercept produced in each case was one. For each simulation, a modified version of Lavender and Kim's (1996) original SAS program for the J-N technique was used. The program used is located in Appendix A.

Results

The results of this study consist of regions of nonsignificance calculated by SAS for each of the 70 cases included in the Monte Carlo simulations (See Appendix B). None of the cases produced anomalies, with each simulated situation yielding an intercept of one, and lower and upper limits below and above one, respectively. It was further noted that the resultant regions of nonsignificance varied in size according to the different manipulations of sample size and variance ratios.

The results show the J-N test to produce the largest amount of variation in accordance with total sample size, since the smallest regions of nonsignificance (and consequently the largest regions of significance) were produced in the series of simulations where the sample size ratio was 50:50. Specific evidence for the importance of total sample size can be seen by first examining the case in which the sample size ratio=50:50 and the variance ratio=10:10, versus the case in which sample size ratio=10:10 and the variance ratio=10:10. The former case resulted in X_{L1}=.88883 and X_{L2}=1.11117, with a region of nonsignificance of .22234 (obtained by subtracting the lower limit from the upper limit); the latter case resulted in X_{L1}=.33548 and X_{L2}=1.66452, with a region of nonsignificance of 1.32904. The smallest region of nonsignificance was produced for the case in which both the sample size and variance ratios were 50:50, resulting in X_{L1}=.88888 and X_{L2}=1.11112, with a region of nonsignificance of .22224. Comparing the results of this case to those of the case in which the variance ratio remained 50:50 but the sample size ratio=10:10, where X_{L1}=.34815 and X_{L2}=1.65185 (region of nonsignificance=1.30370), one can again see the effect of total sample size. Final evidence for the importance of total sample size in using the J-N technique can be seen by comparing the cases in which the sample size ratios were 50:50 and 10:10, respectively, while the variance ratio was held constant at 10:50. In the first case (sample size ratio=50:50, variance ratio=10:50) X_{L1}=.88886 and X_{L2}=1.11114, with a region of nonsignificance of .22228, versus the second case (sample size ratio=10:10, variance ratio=10:50), where X_{L1}=.34190 and X_{L2}=1.65810, with a region of nonsignificance 1.31620. Overall, the smallest regions of nonsignificance were found in the cases in which the sample size ratios were 50:50, while the largest regions of nonsignificance are found in the cases where the sample size ratios were 10:10.

Size and equality of variance ratios seem to have played a minimal role in determining these results, based on cases in which the sample size ratio was held constant at 10:50 while manipulating the variance ratios to 50:50, 10:10, and 10:50. In these cases the resultant regions of nonsignificance were very similar in size. For the first case (sample size ratio=10:50, variance ratio=50:50) X_{L1}=.74532 and X_{L2}=1.25468, with a region of nonsignificance of .50936. For the next two cases in which the sample size ratio remained 10:50 and the variance ratios were 10:10 and 10:50, the regions of nonsignificance were .51080 and .51058, respectively. The results of these three cases (in which both size and equality of variance ratios were manipulated) fall in the .50936 to .51080 range, indicating the minimal importance of variance ratio size and variance ratio equality.

These cases also provide further evidence for the strong effect of total sample size, even when sample size ratios are the most unequal (10:50). This can be seen when comparing the case in which the sample size ratio=10:50 and the variance ratio=50:50 with the case in which the sample size ratio=10:10 and the variance ratio=50:50. As mentioned, the former case yielded X_{L1}=.74532 and X_{L2}=1.25468, and the region of nonsignificance=.50936. The latter case, in which the sample size ratio was both equal and as small as possible (10:10), X_{L1}=.34815 and X_{L2}=1.65185, with a region of nonsignificance of 1.30370. So the case in which the total sample size is larger (10:50) produces a much smaller region of nonsignificance than the case in which the sample size is smallest (10:10). This also provides evidence for the limited importance of equality of sample size ratios in comparison to the strong effect of total sample size. The general trend seen in these results is as follows: regardless of equality of sample size and variance ratios, and regardless of size of variance ratios, regions of nonsignificance are inversely proportional to total sample size.

Discussion

The results of this study indicate that the J-N technique will not produce anomalies in results due to differences in sample size or variance between groups. Each of the cases simulated in this study resulted in valid regions of nonsignificance, with lower and upper limits surrounding the point of intersection. For the purposes of this study, "best case" results were determined by the yielding of the smallest regions of nonsignificance, which in turn produce the largest regions of significance. The best cases were found when the total sample size was largest, while the worst case results occurred when the total sample size was smallest. Thus, it appears that the most important factor in determining the J-N technique's sensitivity is overall sample size. Size of variance did have some effect on the results, since the best overall results were obtained when both the sample size and variance ratios were 50:50; however, even when the variances were smaller and the sample size ratio was 50:50, the results were still very similar.

Another factor that was shown to be of limited importance in increasing the J-N technique's sensitivity was equal ratios. As previously mentioned, the best results were seen when both the sample size and variance ratios were 50:50. In the cases where the sample size ratios remained 50:50 but the variance ratios were made unequal, the results were still very similar (yielding small regions of nonsignificance) to when the variance ratios were equal. It appears, therefore, that the equality of variance ratios did not have a large impact on the results. When the sample size ratios were made to be unequal, results were significantly affected; however, this trend was likely due more to an overall reduction in total sample size than inequality of sample size ratios. So it can be said that, based on these results, the most important factor in determining the J-N technique's sensitivity is total sample size, with higher sample sizes producing the best results.

Little research has been done specifically related to the effect of sample size and variance on the J-N technique. Pigache and Graham (1976) modified a J-N FORTRAN program to eliminate the assumption of equal variances between groups - a modification that improved the results of their study. Hunka (1995) noted that problems exist with the

J-N technique and its production of anomalies in results; however, his research was related to the anomalies being due to variation in the number of groups and covariates, the nature of the regression model, and the use of confidence intervals or regions in determining covariate boundaries. The research done by Chou and Huberty (1992), based on Potthoff's (1964) and Rogosa's (1981) modifications of the J-N technique, was (as in the case of this study) aimed at trying to prevent Type III error rates in using the J-N test. Their research, however, was focused on comparing the effectiveness of the original J-N technique at controlling simultaneous error rates to a modification of this technique (Chou & Huberty, 1992).

The authors of the current study have attempted to prove that differences in sample size or variance ratios are the culprits in creating J-N technique anomalies. The results of this study indicate that differences in these ratios can not be exclusively blamed, no matter how much of an impact sample size and variance seem to have on the calculation of the formulas involved in using the J-N technique. Anomalies may be the result of some other portion of the J-N technique's formulas, and further research should be done to continue ruling out components of these formulas. Another possible difficulty with the current study was that specific cases of sample size and variance ratios may interact poorly with any number of components in the formulas. The case may exist in which an unequal sample size ratio interacts poorly (and produces anomalies) with another component of the formulas only when that component has a value in a certain range. Ideally, further research should be done to manipulate all of the many factors included in the J-N technique's formulas.

This research is of value in that it has revealed what factors increase the sensitivity of the J-N test. Based on the results of this study, researchers can obtain J-N technique results even when sample sizes and variances are different between groups. Better results may be produced, however, by increasing the total sample size and keeping the sample sizes between groups as equal as possible. How equal the groups' sample sizes need to be is a final research question created by this study. Performing simulations again using such sample size ratios as 20:50, 30:50, and 40:50 may help to determine the exact importance of sample size equality in using the J-N technique.

References

Ceurvorst, R. W. (1979). A computer program for the Johnson-Neyman technique. Educational and Psychological Measurement, 39, 205-207.

Chou, T., & Huberty, C. J. (1992, April). The Robustness of the Johnson-Neyman Technique. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Gamache, L. M., & Novick, M. R. (1985). Choice variables and gender differentiated prediction within selected academic programs. Journal of Educational Measurement, 22, 53-70.

Houston, W. M., & Novick, M. R. (1987). Race-based differential prediction in Air Force technical training programs. Journal of Educational Measurement, 24, 309-320.

Huitema, B. E. (1980). Analysis of heterogeneous regression case: Johnson-Neyman technique. The analysis of covariance and alternatives (pp. 271-297). New York: Wiley.

Hunka, S. (1995). Identifying regions of significance in ANCOVA problems having non-homogeneous regressions. British Journal of Mathematical and Statistical Psychology, 48, 161-188.

Johnson, P. O., & Neyman, J. (1936). Tests of certain linear hypotheses and their application to some educational problems. Statistical Research Memoirs, 1, 57-93.

Karpman, M. B. (1980). ANCOVA: A one covariate Johnson-Neyman algorithm. Educational and Psychological Measurement, 40, 791-793.

Karpman, M. B. (1983). The Johnson-Neyman technique using SPSS or BMDP. Educational and Psychological Measurement, 43, 137-147.

Karpman, M. B. (1986). Comparing two non-parallel regression lines with the parametric alternative to analysis of covariance using SPSS-X or SAS: The Johnson-Neyman technique. Educational and Psychological Measurement. 46, 639-644.

Keim-Abbott, S. & Abbott, R. D. (1977). Moderation of achievement prediction in an elementary school metric curriculum by trait by instructional method interactions. Educational and Psychological Measurement. 37, 481-486.

Lavender, T., & Kim, J. K. (1996). Gender differences in the academic performance of college students. Unpublished master's thesis, Middle Tennessee State University, Murfreesboro, TN.

McLeskey, J. & Rieth, H. J. (1982). Controlling IQ differences between reading disabled and normal children: An empirical example. Journal of Learning Disabilities, 15, 481-483.

Pigache, R. M. & Graham, B. R. (1976). A modification of the Johnson-Neyman technique comparing two regressions, applied to treatment effects dependent on baseline levels. Biological Psychology, 4, 213-235.

Potthoff, R. F. (1964). On the Johnson-Neyman technique and some extensions thereof. Psychometrika, 29, 241-256.

Rogosa, D. (1981). On the relationship between the Johnson-Neyman region of significance and statistical tests of parallel within-group regressions. Educational and Psychological Measurement, 41, 73-84.

Scialfa, C. (1987). A BASIC program to determine regions of significance using the Johnson-Neyman technique. Behavior Research Methods, Instruments, and Computers, 19, 349-352.

Appendix A

<u>SAS program used for Monte Carlo simulations manipulating sample size and variance ratios; for each case, appropriate variable values were changed in lines 00001 and 00015</u>

```
00001 title "Monte Carlo for J-N n=10/10, var=10/10";
00002 data mcsim;
00003         input f ssres a1 a2 b1 b2 n1 n2 nt xbar1 xbar2 var1 var2;
00004         ss1=var1*(n1-1); ss2=var2*(n2-1);
00005
00006         a=(-f/(nt-4))*ssres*((1/ss1)+(1/ss2))+(b1-b2)**2;
00007         b=(f/(nt-4))*ssres*((xbar1/ss1)+(xbar2/ss2))+(a1-a2)*(b1-b2);
00008         c=(-f/(nt-4))*ssres*((nt/(n1*n2))+(xbar1**2/ss1)+(xbar2**2/ss2))
00009         +(a1-a2)**2;
00010
00011         xl=(-b-sqrt(b**2-(a*c)))/a;c)))/a;
00012
00013         ints=(a2-a1)/(b1-b2);
00014
00015 cards;
00016
00017 4.49 30 1 3 1 -1 10 10 20 1 1 10 10
00018 ;
00019 proc print;
00020 var xl xu ints;
00021 run;
```

Appendix B

Results of Monte Carlo simulations manipulating sample size and variance

Sample Size Ratio	Variance Ratio	Lower Limit	Upper Limit	Intercept
Equal	Equal			
10:10	10:10	.33548	1.66452	1
10:10	20:20	.34348	1.65652	1
10:10	30:30	.34609	1.65391	1
10:10	50:50	.34815	1.65185	1
Equal	Unequal			
10:10	10:20	.33952	1.66048	1
10:10	10:30	.34085	1.65915	1
10:10	10:50	.34190	1.65810	1
Equal	Equal			
20:20	10:10	.70605	1.29395	1
20:20	20:20	.70672	1.29328	1
20:20	30:30	.70694	1.29306	1
20:20	50:50	.70712	1.29288	1
Equal	Unequal			
20:20	10:20	.70639	1.29361	1
20:20	10:30	.70650	1.29350	1
20:20	10:50	.70659	1.29341	1
Equal	Equal			
30:30	10:10	.80996	1.19004	1
30:30	20:20	.81013	1.18987	1
30:30	30:30	.81019	1.18981	1
30:30	50:50	.81024	1.18976	1
Equal	Unequal			
30:30	10:20	.81005	1.18995	1
30:30	10:30	.81007	1.18993	1
30:30	10:50	.81010	1.18990	1
Equal	Equal			
50:50	10:10	.88883	1.11117	1
50:50	20:20	.88886	1.11114	1
50:50	30:30	.88887	1.11113	1
50:50	50:50	.88888	1.11112	1

ERIC Full Text Provided by ERIC

Equal	Unequal			
50:50	10:20	.88885	1.11115	1
50:50	10:30	.88885	1.11115	1
50:50	10:50	.88886	1.11114	1
Unequal	Equal			
10:20	10:10	.56784	1.43216	1
10:20	20:20	.57003	1.42997	1
10:20	30:30	.57075	1.42925	1
10:20	50:50	.57132	1.42868	1
Unequal	Unequal			
10:20	10:20	.56855	1.43145	1
10:20	10:30	.56878	1.43122	1
10:20	10:50	.56897	1.43103	1
Unequal	Equal			
10:30	10:10	.65999	1.34001	1
10:30	20:20	.66106	1.33894	1
10:30	30:30	.66141	1.33859	1
10:30	50:50	.66169	1.33831	1
Unequal	Unequal			
10:30	10:20	.66024	1.33976	1
10:30	10:30	.66033	1.33967	1
10:30	10:50	.66039	1.33961	1
Unequal	Equal			
10:50	10:10	.74460	1.25540	1
10:50	20:20	.74505	1.25495	1
10:50	30:30	.74520	1.25480	1
10:50	50:50	.74532	1.25468	1
Unequal	Unequal			
10:50	10:20	.74467	1.25533	1
10:50	10:30	.74469	1.25531	1
10:50	10:50	.74471	1.25529	1
Unequal	Equal			
50:10	10:10	.74460	1.25540	1
50:10	20:20	.74505	1.25495	1
50:10	30:30	.74520	1.25480	1
50:10	50:50	.74532	1.25468	1

Unequal	Unequal			
50:10	10:20	.74498	1.25502	1
50:10	10:30	.74511	1.25489	1
50:10	10:50	.74521	1.25479	1
Unequal	Equal			
30:10	10:10	.65999	1.34001	1
30:10	20:20	.66106	1.33894	1
30:10	30:30	.66141	1.33859	1
30:10	50:50	.66169	1.33831	1
Unequal	Unequal			
30:10	10:20	.66080	1.33920	1
30:10	10:30	.66108	1.33892	1
30:10	10:50	.66129	1.33871	1
Unequal	Equal			
20:10	10:10	.56784	1.43216	1
20:10	20:20	.57003	1.42997	1
20:10	30:30	.57075	1.42925	1
20:10	50:50	.57132	1.42868	1
Unequal	Unequal			
20:10	10:20	.56933	1.43067	1
20:10	10:30	.56982	1.43018	1
20:10	10:50	.57021	1.42979	1

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

TM029409

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: The effect of sample Size and Variance on the Johnson-Neyman Technique	
Author(s): Brian M. Wind and Jwa. K. Kim	
Corporate Source:	Publication Date:

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1	Level 2A	Level 2B
☑	☐	☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature:	Printed Name/Position/Title: Jwa K. Kim / Associate Prof.	
Organization/Address: Department of Psychology, MTSU Box 522 Murfreesboro, TN 37132	Telephone: (615) 898-2002	FAX: (615) 898-5027
	E-Mail Address: PYKim@frank.mtsu.edu	Date: 11-5-98

ERIC Full Text Provided by ERIC