In hierarchical linear models we often find that group indicator variables at the cluster level are significant predictors for the regression slopes. When this is the case, the average relationship between the outcome and a key independent variable are different from group to group. In these settings, a question such as “what range of the independent variable is the difference in the outcome variable statistically significant among groups?” naturally arises. The Johnson–Neyman (J-N) technique answers this kind of question in the analysis of covariance (ANCOVA) settings. In the hierarchical modeling context, the F test, which is widely used in ANCOVA, cannot be applied because the assumption of homogeneity of variance within cluster units is violated. Instead, the approximate Wald test can be used to determine the region of significance. To illustrate the application of the J-N technique in the context of hierarchical linear modeling, an example from research in education is provided.

Keywords: analysis of covariance, hierarchical linear models, Johnson–Neyman technique, Wald statistic

In hierarchical linear models (HLM), each macrounit has its own regression parameters, and these parameters are allowed to vary randomly across macrounits. This implies that each macrounit can have a different regression slope. At the macrolevel, the variability of the dependent variable among macrounits is explained using the macrolevel independent variables, that is, information about the characteristics that belong to the macrounits. Frequently, this information includes qualitative variables that create groups. For example, suppose one would want to study the relationship between students’ socioeconomic status (SES) and achievement using a school-effectiveness study such as the High School and Beyond Survey (Coleman, Hoffer, & Kilgore, 1982) where the students’ data are nested within schools. If the schools are grouped by such sectors as private and public school, then each sector has its own average relationship. If the intercepts and slopes are different between these two sectors, then it would be of interest to ask in which range of student’s SES do, in fact, the two sectors have a statistically significant difference in terms of student achievement.
Miyazaki and Maier

Whenever multilevel data includes grouping variables, determining the range of independent variables/covariates where the expected outcome has significant differences between groups can be of interest. In the analysis of covariance (ANCOVA) context when the linear model is appropriate, this region can be found by applying the Johnson–Neyman (J-N) procedure. Clearly, the J-N technique cannot be applied directly because multilevel data do not satisfy the required assumptions about the distribution of error terms. We can apply the spirit of this technique, however, by changing the test statistic to be computed and the subsequent reference distribution. By casting the J-N problem into a linear hypothesis in multilevel modeling, we can use the Wald statistic instead of the usual $F$ statistic. Although this test is not as exact as the $F$ test in the ANCOVA-linear model context, the Wald test is a good approximation when the sample size for the cluster units is large enough, which is frequently the case for the typical multilevel data setting.

Johnson–Neyman Procedure in ANCOVA

The Johnson–Neyman technique, as it was originally developed by Johnson and Neyman (1936), solves the problems of identifying regions of significance of the covariates in ANCOVA when the regression lines are not parallel. The idea behind the J-N procedure can be described in the following simple scenario where we have two groups and a single covariate. For any value of the covariate, we can test whether the difference in means for the two groups is statistically significant by calculating a $F$ statistic. This statistic is compared to the corresponding critical value of Fisher’s $F$ distribution, the distribution that this test statistic follows if the normality and homogeneity of variance assumptions for the error terms are met. The decision rule of rejecting the null hypothesis, that is, whether the group means are different at the specified value of the covariate, provides the inequality that the value of the test statistic must satisfy in order to reject the null hypothesis.

Because we don’t know the value of the covariate, the inequality specified by the decision rule is solved for with respect to the unknown value of the covariate, $x$. For the design having two groups and one covariate described above, a simple formula that involves solving a quadratic equation is appealing because it can provide insight into how the terms in the formula influence the solution. A relatively simple formula for the explicit solution is available for several groups, but only when the number of the covariates is one (Huitema, 1980, chap. 13). If the number of covariates is two or more, the complexity of the formula does not avail itself to a tractable solution (Johnson & Fay, 1950). In cases of multiple covariates, however, the solution of the inequality can be addressed by casting the equation within a general linear model framework without any limitation imposed for the number of possible covariates (Hunka, 1995; Hunka & Leighton, 1997). Although this alternative formulation of the
J-N Type Technique in HLM

J-N technique cannot be solved directly, the symbolic processing capabilities of computational software such as Mathematica can be used to provide a solution (Wolfram, 1999).

In the linear model, the dependent variable is regressed on the set of independent variables that include the qualitative grouping variables and the quantitative covariates. Suppose we have $g$ groups and $p$ covariates, and we formulate the linear model with an intercept for each group. Then we have $gp$ regression slopes and $g$ intercepts. The general linear model can be written as

$$Y = X\beta + \epsilon,$$

where $Y$ is the $n \times 1$ vector of observations, and $X$ is a $n \times P$ design matrix where $P = g(p + 1)$; $\beta$ is the $P \times 1$ vector of parameters; $\epsilon$ is the $n \times 1$ vector of errors and $\epsilon \sim N(0, \sigma^2 I)$ where $0$ is the $n \times 1$ vector of zeros, $\sigma^2$ is the error variance, and $I$ is the $n \times n$ identity matrix.

To assess whether the group differences are statistically significant within certain ranges or regions of covariates, we assign the values of the covariates to the values at which we want to test if a significant difference exists and create a contrast matrix $K^T$ of order $df_K \times P$ where $rank(K^T) = df_K$. The null hypothesis to be tested is:

$$H_0: K^T\beta = 0.$$

The degrees of freedom for the contrast matrix $K^T$, $df_K$ is the number of rows in $K^T$, which are constructed to be independent.

To test the hypothesis, we calculate the statistic

$$\frac{SS_k}{df_k} / \frac{SS_e}{df_e},$$

which is distributed as an $F$ distribution with the numerator degrees of freedom of $df_k$ and the denominator degrees of freedom of $df_e$. The quantity $SS_k$ in the numerator is the sum of squares of the contrast,

$$SS_k = (K^T \hat{\beta})^T [K^T (X^T X)^{-1} K]^{-1} K^T \hat{\beta},$$

where $\hat{\beta} = (X^T X)^{-1} X^T Y$, which is simply the least square estimator of $\beta$. The quantity $SS_e$ in the denominator is the residual sum of squares, $SS_e = (Y - \hat{Y})^T (Y - \hat{Y})$, where $\hat{Y} = X\hat{\beta}$ is the predicted value of the outcome. The mean squared error

$$MS_e = \frac{SS_e}{df_e}$$
Miyazaki and Maier

is the estimator of \( \sigma^2 \). The statistic can be compared to the central \( F \) distribution because

\[
\frac{SS_k}{\sigma^2} \sim \chi^2_{df_k},
\]

under

\[
H_0 \text{ and } \frac{SS}{\sigma^2} \sim \chi^2_{df_e},
\]

and \( SS_k \) and \( SS_e \) are independent. To test the null hypothesis, the quantity

\[
\frac{SS_k}{df_k} \text{ / } \frac{SS_e}{df_e}
\]

is compared to the critical value \( F_{\alpha;df_k;df_e} \) of the \( \alpha \)-level test:

\[
\frac{SS_k}{df_k} \text{ / } \frac{SS_e}{df_e} \geq F_{\alpha;df_k;df_e}. \quad (4)
\]

If the null hypothesis implied by \( K^T \) is be rejected at the \( \alpha \) level, the following inequality must hold:

\[
(K^T \hat{\theta})^T [K^T(X^TX)^{-1}K]^{-1}K^T \hat{\theta} - (MS_e)(df_k F_{\alpha;df_k;df_e}) \geq 0. \quad (5)
\]

The region of significance can be obtained by solving Equation 5 with respect to the unknown values of the covariates included in the contrast matrix \( K^T \). Thus, to determine these values, one must compute the root of the equation

\[
(K^T \hat{\theta})^T [K^T(X^TX)^{-1}K]^{-1}K^T \hat{\theta} - (MS_e)(df_k F_{\alpha;df_k;df_e}) = 0. \quad (6)
\]

For the case of two groups and one covariate, the quantity \((K^T \hat{\theta})^T [K^T(X^TX)^{-1}K]^{-1}K^T \hat{\theta}\) can be shown to be the ratio of two quadratic polynomials. For more complicated cases, the computation is much more involved and can be carried out using symbolic algebra.

Two types of a Johnson–Neyman region for comparing \( g \) groups may be constructed by the solution of Equation 5: an individual region of significance and a simultaneous region of significance. The difference between these two regions lies in the degrees of freedom specified for the inequality. This specification changes the constant term \( df_k F_{\alpha;df_k;df_e} \).

The individual region of significance is appropriate when performing pairwise comparisons for all the groups at a specified point of interest. Analogous to mak-
ing pairwise comparisons in a one-way analysis of variance, one might want to take the simultaneity into consideration in this case. A simple way of doing this, usually more powerful than using Scheffé, is to use Bonferroni, substituting \( \alpha/(g(g-1)/2) \) for \( \alpha \). Then both Scheffé for all points and individual tests for specified points can be performed for each pair, keeping the overall level of significance at \( \alpha \). For any specified values of covariates contained within the individual region of significance, the difference in outcome among groups is statistically significant at the \( \alpha \) level. In a setting having \( g \) groups and \( p \) covariates, the individual region of significance implies that the groups are different for any specified individual point in the J-N region. If we want to compare all the groups by pairwise comparison, then the constant \( df_k F_{\alpha, d_k; d_f e} \) in Equation 5 is replaced with \((g-1)F_{\alpha, g-1, d_f e}\).

The simultaneous region of significance defines a region that the groups are different simultaneously for all points contained in the region for all possible pairs of groups with confidence level of \( 1 - \alpha \) by a \( \alpha \)-level test (Potthoff, 1964, 1983). To construct the simultaneous region, the Scheffé approach can be used. To construct a simultaneous region of significance in the case of \( g \) groups and \( p \) covariates, \( df_k \) in Equation 5 is \( df_k = (p + 1)(g - 1) \). Thus, the constant \( df_k F_{\alpha, d_k; d_f e} \) in the above inequality Equation 5 is replaced with \((p + 1)(g - 1)F_{\alpha; (p + 1)(g - 1); d_f e}\). A simultaneous region of significance results in a larger constant \( df_k F_{\alpha, d_k; d_f e} \) as compared to an individual region of significance. For example, when \( g = 2 \) and \( p = 1 \) (this is a typical ANCOVA design where there are control and treatment groups and a single covariate), we use \( F_{\alpha, 1; d_f e} \) to calculate an individual region of significance and \( 2F_{\alpha, 2; d_f e} \) to calculate a simultaneous region of significance.

A chi-square test could also be used for the hypothesis test because the quantity

\[
\frac{SS_k}{MS_e}
\]

is approximately distributed as a chi-square with the degrees of freedom \( df_k \). The resulting decision rule for testing the null hypothesis in Equation 2 is:

\[
\text{Reject } H_0 \text{ if } \frac{SS_k}{MS_e} \geq \chi^2_{df = df_k, \alpha}
\]

(7)

where \( \chi^2_{df = df_k, \alpha} \) is the critical value for the \( \alpha \)-level test for the chi-square distribution with the degrees of freedom \( df_k \). Note that the chi-square test is exact under a model with known variance, but is a good approximation if it can be estimated with high accuracy (i.e., with large degrees of freedom).

**Johnson–Neyman Type Procedure in Hierarchical Linear Model**

When interactions between macrolevel group membership and microlevel covariates exist in multilevel data, we want to determine the region of covariates for which group differences are statistically significant. In the case of a simple linear model, the J-N procedure could be directly applied to construct a region of significance. To construct an analogous region for multilevel data, such as the
school-effectiveness study described earlier, we apply a J-N type procedure. This type of procedure is not the exact J-N procedure as outlined in the ANCOVA example of earlier, but is a procedure analogous to the original formulation.

As in the J-N procedure for the linear model, the J-N type procedure in HLM can be expressed as a linear hypothesis on the fixed effects. Because linear hypothesis testing is concerned about the fixed effects, the number of the levels in the multi-level model is not of consequence—the same procedure can be applied to three level or higher hierarchical models. For the purpose of clarity, we present the application of the J-N type procedure to a two-level HLM model, the simplest case.

A two-level HLM can be written as:

$$Y_j = A_{1j} \gamma + A_{2j} u_j + \epsilon_j, \quad (j = 1, 2, \ldots, J),$$

where $Y_j$ is the $n_j \times 1$ vector of the dependent variable; $A_{1j}$ is the $n_j \times F$ matrix of fixed covariates; $A_{2j}$ is the $n_j \times R$ matrix of random covariates; $\gamma$ is the $F \times 1$ vector of fixed effects parameters; $u_j$ is the $R \times 1$ vector of Level 2 random effects; and $\epsilon_j$ is the $n_j \times 1$ vector of Level 1 random effects. Let $Y_1, \ldots, Y_j, \ldots, Y_J$ be independent. The distribution of $u_j$ and $\epsilon_j$ are assumed to be independent and normally distributed: $u_j \sim N_R(0, \tau)$ where $\tau$ is a $R \times R$ positive definite symmetric matrix; and $\epsilon_j \sim N(0, \sigma^2 \mathbf{I}_{n_j})$ where $\sigma^2$ is the Level 1 error variance and $\mathbf{I}_{n_j}$ is the $n_j \times n_j$ identity matrix. The design matrix $A_{1j}$ for the fixed effects parameter $\gamma$ involves the qualitative grouping variables at Level 2 and the key covariates at Level 1.

While we can directly apply the J-N procedure and use the $F$ test in an ANCOVA setting, we can use the Wald test (Wald, 1941) in a multilevel data setting. The Wald test is appropriate to test the hypothesis on the fixed effects because the errors in a hierarchical linear model are not independent and identically distributed (i.i.d.). Therefore, we cannot obtain the observed $F$ statistic because the unknown error variances do not cancel out. In this case, we can take advantage of the asymptotic property of the Wald statistic that makes this test asymptotically correct as $J$, the number of clusters (schools in the later example), and, thus, the overall sample size

$$N \left( \sum_{j=1}^{J} n_j \right)$$

approaches infinity. The procedure of constructing the Wald statistic and the test for a HLM are described below.

The form of the Wald test can be constructed from a general form of the linear hypothesis. Suppose we want to test $d_k$ linearly independent hypotheses on the fixed effect $\gamma$ in Equation 8. Using a $d_k \times F$ contrast matrix $K^T$, the null hypothesis can be written as

$$H_0: K^T \hat{\gamma} = 0,$$

where $\text{rank}(K^T) = d_k$. The Wald statistic for testing the hypothesis in Equation 9 is

$$H = (K^T \hat{\gamma})^T \hat{V}_K (K^T \hat{\gamma}),$$

238
where $\hat{V}_K^{-1}$ is the estimator of $V_K^{-1}$, and the $V_K$ is the variance–covariance matrix of $K^T\hat{\gamma}$,

$$V_K \equiv \text{Var}(K^T\hat{\gamma}) = K^T\text{Var}(\hat{\gamma})K.$$  \hfill (11)

$\hat{\gamma}$, the MLE of $\gamma$, can be obtained by:

$$\hat{\gamma} = \left( \sum_{j=1}^{J} A_{ij}^T \hat{V}_j^{-1} A_{ij} \right)^{-1} \sum_{j=1}^{J} A_{ij}^T \hat{V}_j^{-1} Y_j,$$  \hfill (12)

where $\hat{V}_j$ is the MLE of $V_j$, and $V_j$ is the variance–covariance matrix of the observed vector $Y_j$, that is, $V_j \equiv \text{Var}(Y_j)$, and from Equation 8 it can be represented as:

$$V_j = A_{2j}^T \tau A_{2j} + \sigma^2 I_{n_j}.$$  \hfill (13)

The variance–covariance matrix of $\hat{\gamma}$ is

$$\text{Var}(\hat{\gamma}) = \left( \sum_{j=1}^{J} A_{ij}^T V_j^{-1} A_{ij} \right)^{-1}.$$  \hfill (14)

The maximum likelihood estimators of $V_j$ and $\text{Var}(\hat{\gamma})$ are obtained by substituting $\tau$ and $\sigma^2$ in $V_j$ (see Equation 13) by their respective maximum likelihood estimators, which are in turn obtained by means of iterative methods such as Fisher scoring or the EM algorithm. We denote the estimate of $\text{Var}(\hat{\gamma})$ as $\text{Var}(\hat{\gamma})$. Then $\hat{V}_K$, the estimate of $V_K$, is obtained by $\hat{V}_K = K^T\text{Var}(\hat{\gamma})K$. Note that $\text{Var}(\hat{\gamma})$ in Equation 14 is identical to the Fisher information obtained from taking the inverse of the negative expectation of the second derivative of the log-likelihood. Therefore $\text{Var}(\hat{\gamma})$ is the Fisher information evaluated at the maximum-likelihood estimates of $\sigma^2$ and $\tau$.

The Wald statistic $H$ will be distributed asymptotically, that is, as $J \rightarrow \infty$, as a chi-square distribution with the degrees of freedom equal to the number of independent rows in $K^T$, denoted as $d_K$. The null hypothesis $H_0: K^T\gamma = 0$ is rejected at the specified $\alpha$-level if

$$(K^T\hat{\gamma})^T \hat{V}_K^{-1}(K^T\hat{\gamma}) \geq \chi^2_{df=d_K; \alpha}.$$  \hfill (15)

To obtain the individual region of significance of the covariates, we solve the inequality

$$(K^T\hat{\gamma})^T \hat{V}_K^{-1}(K^T\hat{\gamma}) - \chi^2_{df=d_K; \alpha} \geq 0$$  \hfill (16)

for the unknown values in $K^T$.

It is possible to construct a simultaneous region of significance for a multilevel setting analogous to the procedure illustrated in the ANCOVA setting. The inequality Equation 16 is modified by replacing the degrees of freedom $df = d_K$ by $df = d$, where $d \leq F$ is the dimension of the space spanned by a set of independent
$F \times 1$ vectors that creates the $K^T$ contrast. For example, suppose we have $(p+1)$ covariates (including an intercept) at Level 1 and $q$ covariates at Level 2, and there are $g$ groups classified by a categorical variable at Level 2. If each of the Level 1 coefficients is regressed on all of the $q$ covariates at Level 2, then $F = g(p+1)$ $(q+1)$. If we want to obtain a simultaneous region of significance for all possible pairs of group differences in any points of $pq$ covariates, then $d = (g-1)(p+1)$, as opposed to $df = d_k(=g-1)$ for specified points of the covariates in the case of an individual region of significance. It should be noted that the Wald test chosen for the J-N type technique in multilevel models corresponds to the alternative J-N technique in linear models, which is represented by the decision rule in Equation 7. This method was also approximate and utilized a statistic that is approximately distributed as the chi-square distribution.

In the following example, the results are presented by employing the value of chi-square for the individual region of significance approach because the focus of this article is to illustrate the logic of the proposed procedure. The simultaneous region of significance can be readily obtained by simply modifying the degrees of freedoms of the chi-square value from $d_k$ to $d$ in the right-hand-side term in Equation 15. In the example, we report the results and the interpretations based on the individual region of significance. However, we also report the results based on the simultaneous inference for the first problem and provide comments on the differences as well as the determination of $d$, the degrees of freedom of the chi-square value for the simultaneous approach.

Illustrative Example

In the following, we illustrate the application of the Johnson–Neyman type technique in HLM with a real data set. Because the main purpose of this illustration is to elucidate the way in which the analyses can be carried out, we use an individual region of significance approach instead of simultaneous region of significance approach to simplify the procedure. However, the interpretation will be made as if the results apply to the simultaneous significance region because it is more natural to do so in the setting that the research question was phrased. Therefore, we note that the conclusions that we draw based on the results using a pointwise analysis would be much weaker if we use a simultaneous analysis.

The data are a subsample from the 1982 HS&B Survey, a nationally representative sample of U.S. public and Roman Catholic high schools. The data for our example consist of 7,185 students within 160 schools. Of the 160 high schools, 70 schools are Catholic, and 90 schools are public high schools. Descriptive statistics for this data set are summarized in Table 1. The model examines the relationship between students’ mathematics achievement and their socioeconomic status while controlling for school membership. This model was used in Raudenbush and Bryk (2002). For illustrative purposes, we focus on applying the J-N type procedure to construct an individual significance region.

The outcome variable is students’ mathematics achievement score. The mean scores for the students in public schools and students in Catholic schools are 11.4
TABLE 1
Descriptive Statistics for High School and Beyond Data

Overall

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>MATHACH</td>
<td>7,185</td>
<td>12.75</td>
<td>6.88</td>
<td>-2.83</td>
<td>24.99</td>
</tr>
<tr>
<td>SES</td>
<td>7,185</td>
<td>0.00</td>
<td>0.78</td>
<td>-3.76</td>
<td>2.69</td>
</tr>
<tr>
<td>RSES</td>
<td>7,185</td>
<td>0.00</td>
<td>0.66</td>
<td>-3.66</td>
<td>2.85</td>
</tr>
</tbody>
</table>

Level 2 (School Level)

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEANSES</td>
<td>160</td>
<td>0.00</td>
<td>0.41</td>
<td>-1.19</td>
<td>0.83</td>
</tr>
<tr>
<td>SECTOR</td>
<td>160</td>
<td>0.44</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Classified by SECTOR:
Catholic

Level 1 (Student Level)

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>SES</td>
<td>3,543</td>
<td>0.15</td>
<td>0.74</td>
<td>-2.84</td>
<td>1.76</td>
</tr>
<tr>
<td>RSES</td>
<td>3,543</td>
<td>0.00</td>
<td>0.62</td>
<td>-3.37</td>
<td>2.52</td>
</tr>
</tbody>
</table>

Level 2 (School Level)

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEANSES</td>
<td>70</td>
<td>0.17</td>
<td>0.37</td>
<td>-0.76</td>
<td>0.83</td>
</tr>
</tbody>
</table>

Public

Level 1 (Student Level)

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>MATHACH</td>
<td>3,642</td>
<td>11.36</td>
<td>7.08</td>
<td>-2.83</td>
<td>24.99</td>
</tr>
<tr>
<td>SES</td>
<td>3,642</td>
<td>-0.15</td>
<td>0.79</td>
<td>-3.76</td>
<td>2.69</td>
</tr>
<tr>
<td>RSES</td>
<td>3,642</td>
<td>0.00</td>
<td>0.70</td>
<td>-3.66</td>
<td>2.86</td>
</tr>
</tbody>
</table>

Level 2 (School Level)

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEANSES</td>
<td>90</td>
<td>-0.13</td>
<td>0.38</td>
<td>-1.19</td>
<td>0.69</td>
</tr>
</tbody>
</table>

Note: SECTOR = 1 if Catholic and = 0 if public. RSES is the relative student's SES, relative to the school mean SES. MEANSES was called as the school SES in the main text. MATHACH is mathematics achievement variable.

and 14.2, respectively, thus students in Catholic schools have a mean score 2.8 higher than students in public schools. This difference is statistically significant at $\alpha = 0.05$; the difference in means is equivalent to 0.41 unit of within-sector standard deviation ($SD$) unit (the pooled $SD = 6.88$).
The individual students’ SES ranges from −3.76 to 2.69 with a mean of 0 and a standard deviation of 0.78. The school mean SES, which is computed by the mean of the students’ SES for each school, ranges from −1.19 to 0.83 with a mean of 0 and a standard deviation of 0.41. In terms of sector difference, the students’ SES ranges from −2.84 to 1.76 with a mean of 0.15 and a standard deviation of 0.74 in Catholic schools, and in public schools, SES ranges from −3.76 to 2.69 with a mean of −0.15 and a standard deviation of 0.79. On average, students’ SES is 0.38 (the pooled SD = 0.78) higher in Catholic schools than in public schools. School SES averages range from −0.76 to 0.83 in Catholic schools with a mean of 0.17 and a standard deviation of 0.40 and from −1.19 to 0.69 in public schools with a mean of −0.13 and standard deviation of 0.38. Thus also for school SES, Catholic schools are higher by 0.73 standard deviation units (0.41 for the pooled SD) than public schools.

These differential effects of the sector can be described as an effect size. The effect size of the sector is 0.41 on math achievement, 0.38 on student’s individual SES, and 0.73 on school SES. According to Cohen’s definition of the magnitude of an effect size, the first two are considered to be small to medium, while the third effect size, the effect size of sector on school SES, is considered to be large in social science research (Cohen, 1988). To summarize the data, students in Catholic schools have higher mathematics achievement scores than those in public schools; their parents are more educated and richer; and their schools have greater resources.

Raudenbush and Bryk (2002) formulated a two-level model to study the SES-math achievement relationship for a population of schools. The Level 1 model,

\[ Y_{ij} = \beta_{0j} + \beta_{1j}(SES_{ij} - \bar{SES}_j) + \epsilon_{ij}, \epsilon_{ij} \sim i.d. N(0, \sigma^2), \]  

(17)

describes the relationship between the mathematics achievement \( Y_{ij} \) of the student \( i \) in school \( j \) and the student’s SES relative to the school mean SES \( \bar{SES}_j \). At Level 2, the intercept \( \beta_{0j} \) and the SES slope \( \beta_{1j} \) vary from school to school and depend on the school mean SES as well as school sector (public or Catholic):

\[
\beta_{0j} = \gamma_{00} + \gamma_{01}(\bar{SES}_j - \bar{SES}) + \gamma_{02}\text{SECTOR}_j + u_{0j}, \\
\beta_{1j} = \gamma_{10} + \gamma_{11}(\bar{SES}_j - \bar{SES}) + \gamma_{12}\text{SECTOR}_j + u_{1j}. 
\]  

(18)

The random components \( u_{0j} \) and \( u_{1j} \) are assumed normally distributed:

\[
\begin{pmatrix}
  u_{0j} \\
  u_{1j}
\end{pmatrix} \sim \mathcal{N}\left(0,\begin{pmatrix}
  \tau_{00} & \tau_{01} \\
  \tau_{10} & \tau_{11}
\end{pmatrix}\right).
\]

SECTOR\(_j\) = 1, if school \( j \) is Catholic and 0 if it is public. The school mean SES, \( \bar{SES}_j \) is centered around the grand mean (\( \bar{SES} \)). It should be noted that at Level 1,
students’ SES were centered around the school (group) mean and at Level 2, school mean SES was centered around the grand mean. The group-mean centering of the Level 1 model allows us to examine the relationship between mathematics achievement and the students’ SES relative to the school mean SES within a school. This specification decomposes the SES effect on mathematics achievement directly into within-school and between-school portions and, thus, makes the meanings of the parameters clear in terms of contextual effects. Overall, the combined model represents a complex math–SES relationship because the strength of the association between students’ mathematics achievement and his or her individual SES depends not only on the school sector, but also on the school SES. The results of the analysis appear in Table 2.6

Examining the results, we see that Catholic schools do better in promoting high mathematics achievement and provide a more egalitarian education. \( \hat{\gamma}_{00} = 12.10 \) indicates that the average predicted score for the students in public school whose SES is the school mean and the MEANSES is the grand mean is 12.10; \( \hat{\gamma}_{10} = 2.94 \) indicates that on average one unit increase in student’s SES increases the math achievement by 2.94 controlling for SECTOR and MEANSES. In terms of the sector effects, Catholic schools do better by 1.23 (\( \hat{\gamma}_{02} = 1.23 \)) when SES and MEANSES are controlled, but have a lower SES slope \( \hat{\gamma}_{12} = -1.64 \). MEANSES works in a way to increase both the intercept (\( \hat{\gamma}_{01} = 5.33 \)), that is, the mean score given student’s SES and SECTOR, and the SES slope (\( \hat{\gamma}_{11} = 1.03 \)).

Because the mean mathematics achievement gap between Catholic and public schools decreases as the individual SES increases, we expect that for some value of individual SES, the mathematics achievement for students in public schools is

---

**TABLE 2**

*Effect of Student and School SES on Mathematics Achievement*

<table>
<thead>
<tr>
<th>Fixed Effect</th>
<th>Coefficient</th>
<th>SE</th>
<th>T Ratio</th>
<th>Approx. df</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model for school means, ( \beta_{0j} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INTERCEPT, ( \gamma_{00} )</td>
<td>12.10</td>
<td>0.20</td>
<td>60.87</td>
<td>157</td>
<td>0.000</td>
</tr>
<tr>
<td>MEANSES, ( \gamma_{01} )</td>
<td>5.33</td>
<td>0.37</td>
<td>14.45</td>
<td>157</td>
<td>0.000</td>
</tr>
<tr>
<td>SECTOR, ( \gamma_{02} )</td>
<td>1.23</td>
<td>0.31</td>
<td>4.00</td>
<td>157</td>
<td>0.000</td>
</tr>
<tr>
<td>Model for SES-achievement slopes, ( \beta_{1j} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INTERCEPT, ( \gamma_{10} )</td>
<td>2.94</td>
<td>0.16</td>
<td>18.70</td>
<td>157</td>
<td>0.000</td>
</tr>
<tr>
<td>MEANSES, ( \gamma_{11} )</td>
<td>1.03</td>
<td>0.30</td>
<td>3.42</td>
<td>157</td>
<td>0.001</td>
</tr>
<tr>
<td>SECTOR, ( \gamma_{12} )</td>
<td>-1.64</td>
<td>0.24</td>
<td>-6.76</td>
<td>157</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random Effect</th>
<th>SD</th>
<th>Component</th>
<th>df</th>
<th>Chi-square</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>School mean, ( u_{0j} )</td>
<td>1.54</td>
<td>2.38</td>
<td>157</td>
<td>605.30</td>
<td>0.000</td>
</tr>
<tr>
<td>SES-achievement slope, ( u_{1j} )</td>
<td>0.39</td>
<td>0.15</td>
<td>157</td>
<td>162.31</td>
<td>0.369</td>
</tr>
<tr>
<td>Level 1 error, ( \epsilon_{ij} )</td>
<td>6.06</td>
<td>36.70</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The outcome variable is MATHACH.
better than student in Catholic schools. That is, the SES slopes are different for Catholic schools and public schools. In this context, a question such as “In what range of SES are mathematics achievement scores within private and public schools statistically different and which sector does better in that range?” naturally arises. This question is especially important for prospective students and the parents choosing between the public and private school sectors. To answer this question, we need to apply a Johnson–Neyman type procedure.

In the question posed above, we first need to know the expected values of mathematics achievement for Catholic and public schools. These are obtained by taking expectations over the distributions of both $e_{ij}$ and $u_j$;

$$E(Y_{ij}) = \{\gamma_{00} + \gamma_{01}(\overline{SES}_j - \overline{SES}_.) + \gamma_{02}(SECTOR)_j\} + \{\gamma_{10} + \gamma_{11}(\overline{SES}_j - \overline{SES}_.) + \gamma_{12}(SECTOR)_j\} (SES_{ij} - \overline{SES}_j).$$

(19)

Then, the expected mathematics achievement scores for the Catholic and public sectors are as follows:

**Catholic:**

$$E(Y_{ij})_C = \{(\gamma_{00} + \gamma_{02}) + \gamma_{01}(\overline{SES}_j - \overline{SES}_.)\} + \{(\gamma_{10} + \gamma_{12}) + \gamma_{11}(\overline{SES}_j - \overline{SES}_.)\} (SES_{ij} - \overline{SES}_j).$$

(20)

**Public:**

$$E(Y_{ij})_P = \{(\gamma_{00} + \gamma_{01})(\overline{SES}_j - \overline{SES}_.)\} + \{(\gamma_{10} + \gamma_{11})(\overline{SES}_j - \overline{SES}_.)\} (SES_{ij} - \overline{SES}_j).$$

(21)

The fact that overall mean SES is zero simplifies Equations 20 and 21 to the following equations.

**Catholic:**

$$E(Y_{ij})_C = \{(\gamma_{00} + \gamma_{02}) + \gamma_{01}\overline{SES}_j\} + \{(\gamma_{10} + \gamma_{12}) + \gamma_{11}\overline{SES}_j\} (SES_{ij} - \overline{SES}_j).$$

(22)

**Public:**

$$E(Y_{ij})_P = \{(\gamma_{00} + \gamma_{01})\overline{SES}_j\} + \{(\gamma_{10} + \gamma_{11})\overline{SES}_j\} (SES_{ij} - \overline{SES}_j).$$

(23)

Replacing the parameters with estimates provides the predicted values as the function of $SES_{ij}$ and $\overline{SES}_j$.

**Catholic:**

$$\hat{Y}_{ij,C} = \{(\hat{\gamma}_{00} + \hat{\gamma}_{02}) + \hat{\gamma}_{01}\overline{SES}_j\} + \{(\hat{\gamma}_{10} + \hat{\gamma}_{12}) + \hat{\gamma}_{11}\overline{SES}_j\} (SES_{ij} - \overline{SES}_j)$$

$$= (13.32 + 5.33\overline{SES}_j) + (1.30 + 1.03\overline{SES}_j)(SES_{ij} - \overline{SES}_j).$$

(24)
These predicted values create a surface on the plane determined by the student’s SES (SES<sub>ij</sub>) and the school SES (̅<sub>SES</sub>) axes. We can examine sections of the surface by fixing either SES<sub>ij</sub> or ̅<sub>SES</sub> at three different values, that is, one standard deviation above the mean, mean, and one standard deviation below the mean. Here, we fix the school mean SES (mean = 0, SD = 0.41 as in Table 1) at -0.41, 0, and 0.41, and then examine the relationship between mathematics achievement and student SES (see Figure 1).

As can be seen in Equations 22 and 23, once the school mean SES, ̅<sub>SES</sub>, is fixed, the relationship between mathematics achievement and SES is assumed linear. The graph shows that for low to middle/high student SES, going to Catholic school produces higher mathematics achievement, but this gap decreases as student SES increases. At higher student SES levels, we would expect that going to public schools will be a better choice. The value of student SES at which public schools catch up to Catholic schools shifts to the right as school mean SES increases, that is, for ̅<sub>SES</sub> = -0.41, SES<sub>ij</sub> = 0.34, for ̅<sub>SES</sub> = 0, SES<sub>ij</sub> = 0.75, and for ̅<sub>SES</sub> = 0.41, SES<sub>ij</sub> = 1.16. This means that the higher the SES of the school, the more students receive the benefit of achieving higher mathematics scores by attending Catholic schools.

FIGURE 1. Relationship between math achievement and student SES by sector (for high, middle, and low school SES).
Although the preceding description shows a general pattern of how mathematics achievement differs in Catholic and public schools as a function of student’s SES given a certain school SES, we are not sure whether these differences are, in fact, statistically significant. To determine the sets of the pairs of student SES and school SES that produce the statistically significant and nonsignificant differences between Catholic and public schools, we need to apply the Johnson–Neyman type procedure described above.

To perform the test, we must first compute the expected difference between Catholic and public schools. For the same student SES and school SES scores, the expected difference is obtained from taking the difference between Equations 22 and 23:

\[ E(Y_{ij})_C - E(Y_{ij})_P = \gamma_0 + \gamma_1(SES_{ij} - \overline{SES}_j). \]  

Thus, the expected difference is the linear function of the relative student’s SES score, relative to the school SES score. The predicted difference is

\[ \hat{Y}_{ij}C - \hat{Y}_{ij}P = \hat{\gamma}_0 + \hat{\gamma}_1(SES_{ij} - \overline{SES}_j). \]  

Replacing the parameter values by their estimates in Table 2 provides the numerical relationship:

\[ \hat{Y}_{ij}C - \hat{Y}_{ij}P = 1.23 - 1.64(SES_{ij} - \overline{SES}_j). \]  

Thus, the difference decreases as the relative SES increases.

Let a student’s relative SES be \( x \), that is, \( x = p - q \), where \( p \) is the student’s SES, and \( q \) is the school SES. Then, the hypothesis that \( E(Y_{ij})_C - E(Y_{ij})_P = 0 \) can be cast into the form of a linear hypothesis:

\[ H_0: K^T \gamma = 0 \]  

where \( K^T = (0, 0, 1, 0, 0, x) \), and \( \gamma^T = (\gamma_{00}, \gamma_{01}, \gamma_{02}, \gamma_{10}, \gamma_{11}, \gamma_{12}) \).

Because \( K^T \) has one row, the degrees of freedom for the reference chi-square distribution is one. Thus, the individual region of significance for SES is given by solving the inequality:

\[ (K^T \hat{\gamma})^T \hat{V}_K^{-1} (K^T \hat{\gamma}) - \chi_{df=1, \alpha}^2 \geq 0, \]  

with respect to \( x \), where \( \hat{V}_K \) is the estimate of \( V_K \) in Equation 11. If we choose \( \alpha = 0.05 \),

\[ (K^T \hat{\gamma})^T \hat{V}_K^{-1} (K^T \hat{\gamma}) - 3.84 \geq 0, \]  

where \( K^T \hat{\gamma} = 1.23 - 1.64x \). Because \( \hat{V}_\gamma \) is a \( 6 \times 6 \) symmetric matrix, and if
\[
\hat{\gamma} = \begin{pmatrix}
\text{var}(\hat{\var_0}) & \text{cov}(\hat{\var_0}, \hat{\var_0}) & \text{cov}(\hat{\var_0}, \hat{\var_{20}}) & \text{cov}(\hat{\var_0}, \hat{\var_{22}}) & \text{cov}(\hat{\var_{20}}, \hat{\var_0}) \\
\text{cov}(\hat{\var_0}, \hat{\var_0}) & \text{var}(\hat{\var_0}) & \text{cov}(\hat{\var_0}, \hat{\var_{20}}) & \text{cov}(\hat{\var_0}, \hat{\var_{22}}) & \text{cov}(\hat{\var_{20}}, \hat{\var_0}) \\
\text{cov}(\hat{\var_0}, \hat{\var_{20}}) & \text{cov}(\hat{\var_0}, \hat{\var_{20}}) & \text{var}(\hat{\var_0}) & \text{cov}(\hat{\var_{20}}, \hat{\var_0}) & \text{cov}(\hat{\var_{20}}, \hat{\var_{22}}) \\
\text{cov}(\hat{\var_{20}}, \hat{\var_0}) & \text{cov}(\hat{\var_{20}}, \hat{\var_{20}}) & \text{cov}(\hat{\var_{20}}, \hat{\var_{20}}) & \text{var}(\hat{\var_0}) & \text{cov}(\hat{\var_{20}}, \hat{\var_{22}}) \\
\text{cov}(\hat{\var_{22}}, \hat{\var_0}) & \text{cov}(\hat{\var_{22}}, \hat{\var_{22}}) & \text{cov}(\hat{\var_{22}}, \hat{\var_{22}}) & \text{cov}(\hat{\var_{22}}, \hat{\var_{22}}) & \text{var}(\hat{\var_0})
\end{pmatrix}
\]

\[
(32)
\]

then
\[
\hat{\gamma} = K^T \hat{\gamma} K = \text{Var}(\hat{\var_{20}}) + 2x \text{cov}(\hat{\var_{20}}, \hat{\var_{22}}) + x^2 \text{Var}(\hat{\var_{22}}).
\]

For this example,
\[
\hat{\gamma} = \begin{pmatrix}
3.95 \times 10^{-2} & 1.80 \times 10^{-2} & -4.24 \times 10^{-2} & 2.30 \times 10^{-3} & 1.05 \times 10^{-3} & -2.47 \times 10^{-3} \\
1.80 \times 10^{-2} & 1.36 \times 10^{-1} & -4.02 \times 10^{-2} & 1.06 \times 10^{-3} & 8.15 \times 10^{-3} & -2.41 \times 10^{-3} \\
-4.24 \times 10^{-2} & -4.02 \times 10^{-2} & 9.38 \times 10^{-2} & -2.47 \times 10^{-3} & -2.40 \times 10^{-3} & 5.63 \times 10^{-3} \\
2.30 \times 10^{-3} & 1.06 \times 10^{-3} & -2.47 \times 10^{-3} & 2.47 \times 10^{-2} & 1.33 \times 10^{-2} & -2.65 \times 10^{-2} \\
1.05 \times 10^{-3} & 8.15 \times 10^{-3} & -2.40 \times 10^{-3} & 1.33 \times 10^{-2} & 9.15 \times 10^{-3} & -2.58 \times 10^{-2} \\
-2.47 \times 10^{-3} & -2.41 \times 10^{-3} & 5.63 \times 10^{-3} & -2.65 \times 10^{-2} & 2.58 \times 10^{-2} & 5.90 \times 10^{-2}
\end{pmatrix}
\]

\[
(33)
\]

Solution of the unknown value \( x \) in the contrast matrix \( K^T \) yields the individual region of significance. Steps for carrying out the computations using Mathematica software for this example appear in the technical appendix. The resulting individual region of significance is shown in Figure 2. The axes labeled as \( H - 3.84 \) represents the value of \( \hat{\gamma}^T K \hat{\gamma} - 3.84 \) for various values of \( x \), the relative student’s SES.

The results indicate that a statistically significant difference in achievement is attained when the groups are equated to a value of the student’s relative SES set to 0.36 or smaller and 1.29 or greater, that is, the point at which the function crosses the horizontal axis. Significant differences in achievement at \( \alpha < 0.05 \) will be found for all values of \( x \) whose value of \( H - 3.84 \) axis is above the horizontal axis. With 95% confidence, we conclude that when a student’s relative SES is less than 0.36, we will expect the student to have a higher mathematics achievement score in a Catholic school than in a public school. On the other hand, if the student’s relative SES is greater than 1.29, the student will be expected to have higher mathematics achievement in public schools than in Catholic schools. If the student’s relative SES is between 0.36 and 1.29, there is no statistically significant difference in achievement whether the student attends a Catholic high school or a public high school.

Another way of examining the student’s relative SES range that produces either significant or nonsignificant differences is to study the expected difference between Catholic and public schools directly, represented by Equation 26 or \( K^T \gamma \) where \( K^T \)
FIGURE 2. Individual region of significance as a function of relative student SES. (Note: R. SES is the student SES relative to the school mean to which he or she belongs; $H$ is the value of the Wald statistic in Equation 30, and 3.84 is the critical value for the chi-square distribution with the degrees of freedom = 1.)

$= (0, 0, 1, 0, 0, x)$ for $x$ will be the student’s relative SES and $\gamma^T = (\gamma_{00}, \gamma_{01}, \gamma_{02}, \gamma_{10}, \gamma_{11}, \gamma_{12})$. Because a contrast $K^T$ in this case has one row, a 95% individual confidence interval on $K^T\hat{\gamma}$ is:

$$K^T\hat{\gamma} \pm \sqrt{3.84\hat{V}_k}$$

(34)

where 3.84 is the critical value at the upper tail cutoff of a chi-square variate with one degree of freedom, $K^T\hat{\gamma}$ is the predicted difference as in Equations 27 or 28, and $\hat{V}_k$ is the estimate of the $\text{Var}(K^T\hat{\gamma})$, which is represented in Equation 11.

The predicted difference and its associated confidence interval can be best described visually, as in Figure 3. The bold straight line in Figure 3 represents the predicted difference, as expressed by Equation 28, or symbolically, by $K^T\hat{\gamma}$. The upper curve represents the upper bound of the 95% confidence interval and the lower curve represents the corresponding lower bound. From the figure, we find that the expected mathematics achievement of students in Catholic schools is higher in the region of $(-\infty, 0.75)$ of students’ relative SES than that in public schools, where 0.75 is the value of the student’s relative SES at which the bold line crosses the horizontal axis. This value was the same value that was observed in Figure 1, where the mean achievement of public schools gets higher in the region of $(0.75, \infty)$ than that of Catholic schools. To see if the difference is statistically significant or not, we focus on the values at which the lower bound line and the upper bound line cross the abscissa. They are 0.36 for the line of lower bound and
1.29 for the upper bound line. Thus, we conclude that the expected mathematics achievement of students in Catholic schools is statistically significantly higher in the region of \((-\infty, 0.36)\) in students’ relative SES than that in public schools. On the other hand, the expected mathematics achievement of students in public schools is statistically significantly higher in the region of \((1.29, +\infty)\) in students’ relative SES than that in Catholic schools.

Out of all 7,185 students in our sample, 4,993 students (69.5\%) have their relative SES lower than 0.36, 2,064 students (28.7\%) have their SES between 0.36 and 1.29, and 128 students (1.8\%) has the relative SES higher than 1.29. Thus, it can be said that, on average, about 70\% of the students will experience a gain in mathematics achievement scores if they choose to go to a Catholic school rather than a public school when the two sectors have the same school SES.

If we classify those three groups of students by sector (Table 3), we find that 2,469 students in Catholic schools and 2,524 students in public schools out of a total of 4,993 students have a relative SES lower than 0.36. Thus, it can be said that, after taking into consideration the difference between his or her own SES and the school SES, 69.6\% (2,469/3,543) of the students in Catholic school made the right choice, and 69.3\% (2,524/3,642) of the students in public schools could have obtained a better mathematics achievement score if they chose to attend a Catholic school instead. For the students whose relative SES is high, 93 students (2.6\% of the all public school students) in public schools had their relative SES (RSES) greater than or equal to 1.29, and thus they benefited by having gone to public schools; 35
TABLE 3

Classification of the Students by Sector and Level of Their SES

<table>
<thead>
<tr>
<th>Sector</th>
<th>RSES &lt; 0.36</th>
<th>0.36 ≤ RSES &lt; 1.29</th>
<th>1.29 ≤ RSES</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catholic</td>
<td>2,469</td>
<td>1,039</td>
<td>35</td>
<td>3,543</td>
</tr>
<tr>
<td></td>
<td>(69.6%)</td>
<td>(29.3%)</td>
<td>(1.0%)</td>
<td>(100%)</td>
</tr>
<tr>
<td>Public</td>
<td>2,524</td>
<td>1,025</td>
<td>93</td>
<td>3,642</td>
</tr>
<tr>
<td></td>
<td>(69.3%)</td>
<td>(28.1%)</td>
<td>(2.6%)</td>
<td>(100%)</td>
</tr>
<tr>
<td>Total</td>
<td>4,993</td>
<td>2,064</td>
<td>128</td>
<td>7,185</td>
</tr>
</tbody>
</table>

students (1.0% of the all Catholic school students) in Catholic schools were disadvantaged by having gone to Catholic schools. Overall, 2,562 (35.7%) students made a right choice in terms of sector, and 2,559 (35.6%) students could have experienced gains in their mathematics achievement scores if they have chosen to attend the other sector. For the rest of the 2,064 (28.7%) students, the sector did not create a statistically significant difference in mathematics scores.

Applying the J-N type procedure to the previous example allows us to explore additional interesting inferences that could not be examined otherwise. The data clearly show that choosing a school with the highest school SES will be beneficial for students’ mathematics achievement scores. If schools in both the Catholic and public sectors have the same school SES, then the individual SES plays a sector-specific role. Students whose individual SES is lower than school SES or whose individual SES is higher up to 0.36 standard deviations of the school SES would benefit more from choosing a Catholic school. Students whose family SES is substantially higher than the school SES (specifically 1.29 standard deviations and above) would benefit from choosing a public high school. For those remaining students whose family SES relative to the school SES is between 0.36 and 1.29, the choice of sector does not have an impact on students’ mathematics scores. For our sample, more than two thirds of the students (69.5%) would have realized an advantage if they had chosen a Catholic school.

Although it is possible to interpret the results directly by the students’ relative SES as was done above, translating students’ relative SES back into the original scales, that is, student’s SES and school mean SES, provides extra insight. Because \( x = p - q \) and the nonsignificance region is represented by the inequality, \( 0.36 \leq x \leq 1.29 \), the information can be mapped into a two-dimensional space, as shown in Figure 4.

The large rectangle delineates the data points in the sample and the two parallel lines, \( q = p - 0.36 \) and \( q = p - 1.29 \) divide the plane into three regions. The upper left area in dark shadow represents the region where students in Catholic schools do significantly better than those in public schools. The area that lies between the parallel lines indicates the region where there is no statistically significant difference between two sectors. The lower right area is the region where
public schools will produce better results than Catholic schools. We can use this figure to answer some interesting questions. For example, let us consider a student with a particular individual SES who is trying to make a decision about which sector of high school to attend in order to achieve better in mathematics. Suppose the student’s SES is 0.5, which is a little lower than one standard deviation above the mean (see the vertical line at SES = 0.5 in Figure 4). The answer depends on the level of school SES. The student will perform better in public school if he or she goes to a school whose school SES is lower than -0.79; there would be no significant difference if the school SES is between -0.79 and 0.14; and the student will perform better in a Catholic school if he or she goes to a school whose school SES is higher than 0.14. Of course, to maximize the students’ mathematics achievement, the higher the school SES, the better the student’s mathematics score will be.\(^{11}\)

We can use Figure 4 to address another scenario. Suppose that an administrator of a school district ponders what kind of students can benefit by going to either a Catholic or a public school if levels of school SES are the same. For simplicity assume the school SES is fixed at the grand mean of zero. Then for students having a SES of up to 0.36, those students will do better if they are sent to Catholic schools. There will be no statistical difference if the student SES is between 0.36 and 1.29. And if the students’ SES is higher than 1.29, they will perform better if they attend a public school.

Now, let us consider a slightly different scenario. Suppose we want to know about the range of the student SES that produces a significant difference in mathematics achievement between the Catholic and public sectors if two students who have the same individual SES are assigned to a typical Catholic school and a typical public school. The “typical” indicates that the school SES is the mean SES of all the schools in the same sector. Let the individual student SES be denoted \(x\), the
Miyazaki and Maier

mean school SES of all the Catholic schools be \( s_C \), and the mean school SES of all the public schools be \( s_P \). In our data, these values are \( s_C = 0.166 \) and \( s_P = -0.130 \). The expected values for those students in each typical school are obtained from Equations 22 and 23 by setting \( \text{SES}_j = s_C \) for Roman Catholic schools and \( \text{SES}_j = s_P \) for public schools:

\[
E(Y_{ij})_{\text{Typ. Cath.}} = \{\gamma_{00} + \gamma_{01} s_C + \gamma_{10} + \gamma_{11} s_C\}(x - s_C) + (\gamma_{10} + \gamma_{11} s_C + \gamma_{12})x,
\]

\[
E(Y_{ij})_{\text{Typ. Pub.}} = \{\gamma_{00} + \gamma_{01} s_P + \gamma_{10} + \gamma_{11} s_P\}(x - s_P) + (\gamma_{10} + \gamma_{11} s_P)x.
\]

The predicted values for each sector are obtained by replacing the parameter values by their estimates and \( s_C = 0.1661 \) and \( s_P = -0.1295 \):

\[
\hat{Y}_{ij, \text{Typ. Cath.}} = \{\hat{\gamma}_{00} + (\hat{\gamma}_{01} - \hat{\gamma}_{10}) s_C + \hat{\gamma}_{02} - \hat{\gamma}_{11} s_C^2 - \hat{\gamma}_{12} s_C\} + (\hat{\gamma}_{10} + \hat{\gamma}_{11} s_C + \hat{\gamma}_{12})x
= 13.963 + 1.649x,
\]

and

\[
\hat{Y}_{ij, \text{Typ. Pub.}} = \{\hat{\gamma}_{00} + (\hat{\gamma}_{01} - \hat{\gamma}_{10}) s_P + \hat{\gamma}_{02} - \hat{\gamma}_{11} s_P^2 - \hat{\gamma}_{12} s_C\} + (\hat{\gamma}_{10} + \hat{\gamma}_{11} s_P)x
= 11.768 + 2.804x.
\]

The relationship between the outcome (math achievement score) and the individual SES for each sector are depicted in Figure 5.

Examining Figure 5, we see that the two average lines intersect each other at 1.900, obtained by equating Equations 37 and 38. Upon comparison of this graph with that generated in Figure 1 for the entire sample, we see that the intersection point is shifted to the right, compared to all of the three intersection points (student SES = 0.34, 0.75, and 1.16) in Figure 1. For students attending a “typical” school in the region where SES < 1.9, students in the Catholic schools will have higher math achievement on average and in the region where SES > 1.9, students in public schools will have higher math achievement than the persons in Catholic school given the same individual SES.

To determine in what region of SES is the difference between two group means statistically significant, we can again use the Johnson–Neyman type technique. As before, we first obtain the expected difference from Equations 35 and 36, which is

\[
E(Y_{ij})_{\text{Typ. Cath.}} - E(Y_{ij})_{\text{Typ. Pub.}} = \{(\gamma_{01} - \gamma_{10})(s_C - s_P) + \gamma_{02} - \gamma_{11}(s_C^2 - s_P^2) - \gamma_{12} s_C\} + (\gamma_{11}(s_C - s_P) + \gamma_{12})x
\]

252
Thus, the hypothesis that $E(Y_{ij})_{Typ.\ Cath.} - E(Y_{ij})_{Typ.\ Pub.} = 0$ can be expressed as (9), that is, a linear combination of the fixed effects parameters as:

$$H_0: K^T \gamma = 0,$$

where $K^T = [0, s_C - s_P, 1, -(s_C - s_P), (s_C - s_P)\left(x - s_C - s_P\right), x - s_C]$ and $\gamma^T = (\gamma_{00}, \gamma_{01}, \gamma_{02}, \gamma_{10}, \gamma_{11}, \gamma_{12})$. Because $K^T$ has one row, the degrees of freedom for the reference Chi-square distribution is one. Thus, the individual SES region of significance is given by solving:

$$(K^T \hat{\gamma})^T \hat{V}_K^{-1}(K^T \hat{\gamma}) - \chi^2_{df=1, \alpha} \geq 0$$

(40)

with respect to $x$. If we choose $\alpha = 0.05$, the critical value is 3.84, and in our data, because $s_C = 0.1661$ and $s_P = -0.1295$, we have $(K^T \hat{\gamma})^T \hat{V}_K^{-1}(K^T \hat{\gamma}) - 3.84 \geq 0$ where

$$K^T = [0, 0.2956, 1, -0.2956, 0.2956(x - 0.0366), x - 0.1661].$$

(41)
The contrast matrix $K^T$ involves the unknown value $x$; following the same steps as in the previous computation, we reach the solution for the inequality, Equation 40:

$$x \leq 1.077 \text{ or } x \geq 2.639.$$  

Comparing this solution to the case when two sectors have the same school SES, we see that the significance region shifts more to the right, that is, higher student SES. This shift happens because the school SES is higher in Catholic schools than public schools ($s_c = 0.166$ and $s_p = -0.130$), and the estimated effects of school SES ($\gamma_{01}$ and $\gamma_{11}$) are positive both for the intercept ($\hat{\beta}_0$) and the student’s SES slope ($\hat{\beta}_{11}$). The results suggest that going to a Catholic school becomes advantageous to more students in this scenario than the previous scenario where the comparison was made for schools having the same school SES.

**Summary and Discussion**

The Johnson–Neyman technique was developed to determine the region of significance for a covariate when the parallel slopes assumption does not hold in an ANCOVA context. In the ANCOVA context, the model can be formulated by a linear model, and we can define the Johnson–Neyman problem by an equation that the $F$ statistic must satisfy; thus, the roots of the equation provide the solution. In multilevel modeling, when cross-level interactions exist, the use of J-N region of significance is compelling but not as straightforward as in the ANCOVA context. In this case use of the usual $F$ test (exact even for small samples for the linear model) is inappropriate because the within-group errors are not identical. Instead, we can use the approximate Wald test, which is asymptotically distributed as a chi-square distribution and is valid when the sample size, especially the number of clusters, is large enough.

In the example, we illustrated how the Johnson–Neyman type technique was useful in organizational research where the key covariate is continuous. The example is a straightforward extension of the Johnson–Neyman technique to multilevel modeling settings. Using this technique, we can identify the specific regions of the pairs of individual SES and school SES that produce statistically significant differences. The identification provides more precise information on which sector of schools is advantageous to achieve higher student mathematics achievement scores in what conditions. It should be noted that we interpreted the results from the point significances as if they apply to the simultaneous significance for illustrative purposes. Thus, the conclusions would be much weaker if more appropriate simultaneous significance levels were used.

In this article, we applied the J-N type technique to the two-level hierarchical linear model only. The approach developed in this article can be directly extended to three or higher level models and multilevel models whose outcome variables are not continuous. Alternatively, we can apply a J-N type procedure to longitudinal growth models, as illustrated in Curran, Bauer, and Willoughby (2003) and Miyazaki and Maier (2003). Application of the J-N type procedure to these more complex
models is extremely straightforward because the hypothesis tests are for the fixed effects parameters only.

**Technical Appendix**

The purpose of this technical appendix is to describe the detailed steps to perform the Johnson–Neyman type techniques in the context described on page 246 using Mathematica software. The steps are provided as follows.

**Step 1:** Obtain the estimate of the variance–covariance matrix for $\hat{\gamma}$ when you run the HLM program by using the keyword “Print variance–covariance”. This command produces a text file ‘gamvc.dat’ that contains the variance–covariance matrix of the fixed effects. The matrix (Equation 33) is the one we obtained.

**Step 2:** Define the estimate of the variance–covariance matrix for $\hat{\gamma}$, $\hat{\text{Var}}(\hat{\gamma})$: In Mathematica, scientific notation is represented as $10^\_$. For example, 0.039488440 is $3.9488440 \times 10^{-2}$.

$$V_g = \begin{pmatrix}
3.9488440 \times 10^{-2}, & 1.7959450 \times 10^{-2}, & -4.2425393 \times 10^{-3}, & 2.2987740 \times 10^{-3}, & 1.0537420 \times 10^{-2}, & -2.4745921 \times 10^{-3} \\
1.7959450 \times 10^{-2}, & 1.3628002 \times 10^{-1}, & -4.0245690 \times 10^{-2}, & 1.0585086 \times 10^{-3}, & 8.1498270 \times 10^{-3}, & -2.1743120 \times 10^{-2} \\
-4.2425393 \times 10^{-2}, & -4.0245690 \times 10^{-2}, & 9.3802256 \times 10^{-2}, & -2.1743120 \times 10^{-2}, & -2.4032180 \times 10^{-3}, & 5.6273247 \times 10^{-3} \\
2.2987740 \times 10^{-2}, & 1.0537420 \times 10^{-2}, & -2.4743120 \times 10^{-2}, & 2.4686438 \times 10^{-2}, & 1.3342234 \times 10^{-2}, & -2.6503914 \times 10^{-2} \\
1.0537420 \times 10^{-2}, & 8.1498270 \times 10^{-3}, & -2.4032180 \times 10^{-3}, & 1.3342234 \times 10^{-2}, & 9.1546401 \times 10^{-2}, & -2.5814267 \times 10^{-2} \\
-2.4745921 \times 10^{-2}, & -2.4135827 \times 10^{-2}, & 5.6273247 \times 10^{-3}, & -2.6503914 \times 10^{-2}, & -2.5814267 \times 10^{-2}, & 5.9003012 \times 10^{-2}
\end{pmatrix}$$

**Step 3:** Define the vector $\hat{\gamma}$, the estimate of the fixed effects.

$$g = \begin{pmatrix}
12.0950064, & 5.3330565, & 1.2263840, & 2.9377875, & 1.0344270, & -1.6409540
\end{pmatrix}$$

Note that in order to define the column vector, we enclose each element by a French brace { }.

**Step 4:** Define the contrast matrix:

$$K' = \begin{pmatrix}
0, & 0, & 1, & 0, & 0, & x
\end{pmatrix}$$

**Step 5:** Compute the variance–covariance matrix of $K'\hat{\gamma}$.

$$V_K = K' V_g \text{Transpose}[K']$$
The output provided by Mathematica for this operation is given below:

\[
\{\{0.0938023 + 0.00562732x + (0.00562732 + 0.059003x)x\}\}.
\]

Mathematica can simplify this result by using the command `Simplify`.

\[\text{Simplify}\left[V_\mathbf{K}\right] \]

\[
\{\{0.0938023 + 0.0112546x + 0.059003x^2\}\}\]

**Step 6: Compute the statistics \( H \).**

\[
H = \text{Transpose}[g].\text{Transpose}[\mathbf{K}']\text{.Inverse}[V_\mathbf{K}].\mathbf{K}' .g
\]

\[
\left\{\begin{array}{c}
\frac{1.22638(1.22638 - 1.64095x)}{0.0938023 + 0.00562732x + (0.00562732 + 0.059003x)x} \\
- \frac{1.64095(1.22638 - 1.64095x)x}{0.0938023 + 0.00562732x + (0.00562732 + 0.059003x)x}
\end{array}\right\}.
\]

Again, a more concise representation of this equation can be obtained by asking for a simplification:

\[\text{Simplify}[H] \]

\[
\left\{\begin{array}{c}
[2.69273(-0.74736 + x)(-0.74736 + x)] \\
0.0938023 + 0.112546x + 0.059003x^2
\end{array}\right\}
\]

As can be seen from the above expression, Mathematica represents \( H = (K'^T \hat{\gamma})^{-1} V_e (K'^T \hat{\gamma}) \) as the ratio of a quadratic polynomial both in the numerator and in the denominator.

**Step 7: Solve the equation \( H - 3.84 = 0 \) for the unknown \( x \).**

\[\text{Solve}[H - 3.84 = 0, x] \]

and the results provided by Mathematica are as follows:

\[
\{\{x \to 0.359527\}, \{x \to 1.29004\}\}.
\]

The two real roots 0.359527 and 1.29004 represent the solution of interest.

**Step 8: Plot the function \((H - 3.84)\) over the range of \(-4 \leq x \leq 4\):**

\[\text{Plot}[H - 3.84, \{x, -4, 3\}, \text{AxesLabel} \to \{\text{"R. SES"}, \text{"H - 3.84"}\}]\]

where “R. SES” represents the student’s SES relative to the school mean.
Notes

1 The Bonferroni approach is possible if we are testing a finite number of points. If the covariates that we are interested in determining the regions of significance are continuous variables, the Scheffé approach can be used for determining the simultaneous regions of significance.

2 Of course, the interactions that occur at the same level can be handled in the same way. But the cross-level interaction is mentioned here because it is a common feature of multilevel modeling and the school effectiveness example contains such a cross-level interaction.

3 The log-likelihood for Equation 8 is

\[
l = \log[f(Y \mid \sigma^2, \tau, \gamma)] = \frac{1}{2} \sum_{j=1}^{J} n_j \log(2\pi) - \frac{1}{2} \sum_{j=1}^{J} \log |V_j| - \frac{1}{2} \sum_{j=1}^{J} (Y_j - A_j \gamma)^T V_j^{-1} (Y_j - A_j \gamma)
\]

and, thus, the information with respect to \( \gamma \) is

\[
I(\gamma) = -E[\partial^2 l / \partial \gamma^2] = \sum_{j=1}^{J} A_j^T V_j^{-1} A_j.
\]

Then, as the sample size increases, the MLE of \( \gamma \) is consistent and asymptotically normally distributed with a mean \( \gamma \) and a variance \([I(\gamma)]^{-1}\), which is represented by \( \hat{\gamma} \sim \mathcal{N}(\gamma, [I(\gamma)]^{-1}) \).

4 The reason behind this can be shown by the similar argument for the simultaneous confidence interval for the linear model. The major change from the case of the linear model is that the quantity represented by a Wald statistic is approximately distributed as a chi-square with \( d \) degrees of freedom. The rest of the proof follows in a similar manner. The key for the proof is to apply the inequality (A4.11) of Seber (1977, p. 388) to the Wald statistic. See Seber (1977, chap. 7) or Miller (1980, chap. 2) for more detail. The subtle part of the simultaneous J–N region of significance is to determine the dimension \( d \).

5 This variable is frequently interpreted as a proxy measure for such school resources as the number of teachers available, facilities, funding, and so forth.

6 The model was fitted via the Restricted Maximum Likelihood Method (REML) using HLM software (Raudenbush, Bryk, Cheong, & Congdon, 2000).

7 Readers should be reminded that statements that made from here on are based on the premise that certain assumptions are satisfied. Those assumptions include: the model is correctly specified, errors for each level are normally distributed, error variances are equal within each level, etc. Please note that the linear model used here was a rather oversimplified one. Therefore, there could be a model specification error such as omission of some important independent variables, which may alter the results and the substantive implications.
Of course, we can take a look at Equations 24 and 25 the other way around, that is, mathematics achievement as a function of school SES by fixing student’s SES. As we can see, once the student’s SES is fixed, Equations 24 and 25 become a quadratic function of school SES. We, however, restrict our attention only on the above case, that is, mathematics achievement as a function of student’s SES given a certain school SES, for the purpose of simplifying the presentation of the idea.

If we want to have the simultaneous region of significance, we should use \( df = 2 \) instead of \( df = 1 \) because the dimension spanned by the intercept, 1, and a single independent variable \( SES_j \) in Equation 26 is two. Thus we should use 5.99 for \( \chi^2_{df} = 2, \alpha = 0.05 \) instead of 3.84 for \( \chi^2_{df} = 1, \alpha = 0.05 \). The roots then become 0.27 and 1.47, corresponding to 0.36 and 1.29, respectively. Clearly, the simultaneous inference results in a wider nonsignificance region than for individual inference. We will omit to report the simultaneous results for the rest of the presentation because a minor modification on \( df \) such as the one shown here will provide the results for simultaneous statement immediately, and the focus of this article is on methodological development, not on substantive implications.

Note that, strictly speaking, this confidence interval is not a simultaneous one. If we want to state the confidence interval at any points of the \( X \), that is, the student relative SES, simultaneously at .05 level, then we need to change the coefficient for \( \hat{V}_k \) in Equation 34, from 3.84 to 5.99 because the degrees of freedom for the chi-square distribution is now two. Regarding this point, see the short notes in the previous section.

Because predicted math achievement is a quadratic function with respect to school SES (see Equations 24 and 25 with the coefficient for the quadratic term be negative), there is a school SES that maximizes the student’s math achievement. For the student whose SES is 0.5, it is 2.20. But there is no school whose school SES is 2.20 in our data set.

References


J–N Type Technique in HLM


Authors

YASUO MIYAZAKI is Assistant Professor, Department of Educational Leadership and Policy Studies, School of Education, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061; yasuom@vt.edu. His areas of specialization are hierarchical models, models for human growth and development, hierarchical measurement models, sample sizes and statistical power analysis, and schooling effects.

KIMBERLY S. MAIER is Assistant Professor, College of Education, Michigan State University, East Lansing, MI 48824-1034; kmaier@msu.edu. She specializes in hierarchical models, missing data methods, Bayesian data analysis, and gender differences in science and math.

Manuscript received October 1, 2003
Revision received March 1, 2004
Accepted April 9, 2004