

## **Mixed-Model Analysis of Variance**

**Sohad Murrar  
&  
Markus Brauer**

**University of Wisconsin-Madison**

**The SAGE Encyclopedia of Educational Research, Measurement and Evaluation**

**Target Word Count: 3000 - Actual Word Count: 2755**

### **MIXED-MODEL ANALYSIS OF VARIANCE**

The characteristics of the design and the variables in a research study determine the appropriate statistical analysis. A mixed-model Analysis of Variance (or mixed-model ANOVA) is the right data-analytic approach for a study that contains (a) a continuous dependent variable, (b) two or more categorical independent variables, (c) at least one independent variable that varies between-units, and (d) at least one independent variable that varies within-units. By "units" we refer to the unit of analysis, usually subjects. In other words, a mixed-model ANOVA is used for studies in which independent units are "crossed with" at least one of the independent variables and are "nested under" at least one of the independent variables.

Mixed-model ANOVAs are sometimes called split-plot ANOVAs, mixed factorial ANOVAs, and mixed design ANOVAs. They are often used in studies with repeated measures, hierarchical data, or longitudinal data. Below, we will start out by describing simple ANOVAs before moving on to mixed-model ANOVAs. We will focus mostly on the simplest case of a mixed-model ANOVA – one dichotomous between-subjects variable and one dichotomous within-subjects variable. Then, we will briefly present more complex

mixed-model ANOVAs and discuss these analyses in the context of linear mixed-effects models.

### **Simple ANOVAs**

Between-units (e.g., between-subjects) ANOVAs are characterized by units that are "nested under" one or more categorical independent variables. Between-subjects ANOVAs examine the differences between two or more independent groups. For example, a simple one-way between-subjects ANOVA may test whether girls or boys have better grades in school. Here, there is one dichotomous independent variable that varies between-subjects (gender). The goal of the ANOVA is to examine whether the mean scores for each group (boys vs. girls) are reliably different from each other. The statistical model can be described as:

$$Y = b_0 + b_1X + e \quad (1)$$

Where  $Y$  is the dependent variable (scores),  $X$  is the dichotomous independent variable (gender), and  $e$  refers to the residuals (the errors). If the coefficient  $b_1$  is statistically significant, one would conclude that the data provide evidence for the idea that one of the two genders has better grades than the other. Between-subjects ANOVAs are more flexible than independent-samples t-tests because they allow for multiple independent variables with two or more levels each.

Within-units (e.g., within-subjects) ANOVAs are characterized by units that are "crossed with" one or more categorical independent variables. They frequently examine differences between one measurement of a particular variable and another measurement of the same variable for a given subject. In such cases, the observations are not independent of each other in that two data points from the same subject are likely to be

more similar to each other than two data points from two different subjects. Within-subjects ANOVAs examine the differences between two or more dependent groups. Their goal is often to examine changes in an outcome variable over time. For example, a one-way within-subjects ANOVA may test whether students have better grades in English or Math. Here, there is one dichotomous independent variable that varies within-subjects (Discipline: English vs. Math). The statistical model can be described as:

$$(Y_1 - Y_2) = b_0 + e \quad (2)$$

Where  $Y_1$  is subjects' English grade and  $Y_2$  is subjects' Math grade. Like before, the  $e$  refers to the residuals (the errors). If the coefficient  $b_0$  is statistically significant, one would conclude that the data provide evidence for the idea that students' English and Math grades differ from each other. Compared to paired-samples t-tests, within-subjects ANOVAs are more flexible because they allow for multiple independent variables with two or more levels each.

### **A 2 x 2 Mixed-Model ANOVAs**

A mixed-model ANOVA is a combination of a between-unit ANOVA and a within-unit ANOVA. It requires a minimum of two categorical independent variables, sometimes called factors, and at least one of these variables has to vary between-units and at least one of them has to vary within-units. The explanations below focus on the simplest possible mixed-model ANOVA, a so-called 2 x 2 mixed-model ANOVA: one dichotomous between-subjects variable and one dichotomous within-subjects variable. To better understand mixed-model ANOVAs, consider the following research study.

A group of researchers is interested in comparing boys' and girls' grades in English and Math. Let's assume they are predicting a gender difference (girls have better grades than boys) and they expect this gender

difference to be greater in English than in Math. In this example, there are two independent variables. The first is gender (boy vs. girl), a dichotomous between-subjects variable. The second is discipline (Math vs. English), a dichotomous within-subjects variable. For ease of interpretation, let's assume that the data confirm the researchers' hypotheses.

The study just described has a classic 2 x 2 design, and its data can be analyzed with a two-way mixed-model ANOVA. This data-analytic approach allows researchers to test whether there are main effects for both gender and discipline. A main effect is the effect of a particular independent variable, averaging across all levels of the other independent variable(s). The data-analytic approach also allows researchers to test whether there is an interaction between the two independent variables. An interaction is present when the effect of one independent variable is stronger at one level of the other independent variable than at the second level of that same independent variable. A mixed-model ANOVA tests whether each of the three effects – the two main effects and the interaction effect – is statistically significant.

These three effects can be obtained with the following statistical models:

$$(Y_1 + Y_2)/2 = b_0 + b_1X + e \quad (3)$$

$$(Y_1 - Y_2) = b_0 + b_1X + e \quad (4)$$

Where  $Y_1$  is subjects' grades in English,  $Y_2$  is subjects' grades Math,  $X$  is the dichotomous between-subjects variable (gender), and  $e$  refers to the residuals (the error) in the model. The term on the left side of the equations is simply the average (equation (3)) or the difference (equation (4)) of the grades.

The coefficient  $b_1$  in equation (3) represents the main effect of gender, the between-subjects independent variable. If  $b_1$  in equation (3) is statistically significant, one would

conclude that girls on average have reliably higher or reliably lower grades than boys, regardless of discipline. The coefficient  $b_0$  in equation (4) estimates the effect of discipline, the within-subjects independent variable, for students with a score of zero on X (gender). If X is coded 1 and 2, then this coefficient is rather meaningless. However, if X is "centered" (i.e., coded -.5 and +.5, or -1 and +1), then  $b_0$  in equation (4) represents the main effect of discipline. If it is statistically significant, one would conclude that the students, regardless of their gender, performed better in one of two disciplines. The coefficient  $b_1$  in equation (4) represents the interaction effect between gender and discipline. If this coefficient is statistically significant, one would conclude that the gender difference is greater for one of the two disciplines. The coefficient  $b_0$  in equation (3) is the grand mean (the average of all scores) and is usually not interpreted.

Note that many menu-based data analysis programs (like SPSS) will automatically center the dichotomous between-subjects variable (X) for the user when the appropriate module is chosen. When using other, more code-based programs, researchers may have to recode the between-subjects variable "by hand" to make sure it is centered prior to estimating the model described in equation (4) if they want to interpret the main effect of the within-subject variable.

### **Advanced Mixed-Model ANOVAs**

Mixed-model ANOVAs are not limited to dichotomous independent variables. For example, they can contain within-subjects independent variables with more than two levels. Imagine a group of researchers interested in comparing boys' and girls' grades in English, Math, and Biology. They are predicting a gender difference (girls have better grades than boys) and they expect this gender difference to be greater in English than in

the other two disciplines (Math and Biology). Now, the within-subjects independent variable (discipline) has three levels (English, Math, and Biology).

In the case of such data, the study has a 2 x 3 factorial design that can also be analyzed with a mixed-model ANOVA. The data analytic approach is the same as before examining two main effects and an interaction effect, but the within-subjects independent variable will most likely be examined with a specific contrast. Given that the researchers predict the gender difference to be greater in English than in the other two disciplines, the appropriate contrast would be 1, -.5, -.5 (which produces the same F- and p-values as the contrasts 2, -1, -1, and .67, -.33, -.33).

These main and interaction effects can be obtained with the following models:

$$(Y_1 + Y_2 + Y_3)/3 = b_0 + b_1X + e \quad (5)$$

$$(Y_1 - ((Y_2 + Y_3)/2)) = b_0 + b_1X + e \quad (6)$$

Where  $Y_1$ ,  $Y_2$ ,  $X$ , and  $e$  have the same meaning as in equations (3) and (4).  $Y_3$  is students' grades in Biology. The term on the left side of the equation is the average (equation (5)) or the weighted difference (equation (6)) of the grades.

The coefficient  $b_1$  in equation (5) represents the main effect of gender. It tests whether girls have on average reliably better or reliably worse grades than boys, regardless of discipline. The coefficient  $b_0$  in equation (6) corresponds to the effect of the within-subjects contrast. If this contrast is statistically significant, one would conclude that the students, regardless of their gender, have higher grades in English than in Math and Biology. The coefficient  $b_1$  in equation (6) describes the interaction between the within-subjects contrast and gender. If it is statistically significant, one would conclude that the

gender difference is greater in English than in the other two disciplines. Like before, the coefficient  $b_0$  in equation (5), the grand mean, is usually not interpreted.

Researchers may also decide to include covariates – sometimes called confounding variables or concomitants – in their analyses. These are variables that are not of primary interest to the study, but may affect the outcome variable. For example, students' parental income can provide them with resources that may influence their grades. Thus, the researchers decide to measure parental income and to account for the effects of this variable in the statistical analysis. Here, a mixed-model ANOVA with a covariate – called a mixed model analysis of covariance (or mixed model ANCOVA) – can be used to analyze the data. This approach allows researchers to examine the main effects of discipline and gender on grades, as well as the interaction between them, while statistically controlling for parental income.

The relevant effects can be obtained with the following statistical models:

$$(Y_1 + Y_2)/2 = b_0 + b_1X + b_2Z + e \quad (7)$$

$$(Y_1 - Y_2) = b_0 + b_1X + b_2Z + e \quad (8)$$

Where  $Y_1$ ,  $Y_2$ ,  $X$ , and  $e$  have the same meaning as in equations (3) and (4) and  $Z$  is parental income.

The coefficient  $b_1$  in equation (7) represents the main effect of gender, over and above the effect of parental income. The coefficient  $b_0$  in equation (8) represents the effect of discipline for students who have a score of 0 on both  $X$  and  $Z$ . Like before, this coefficient is in most cases rather meaningless if  $X$  and  $Z$  are uncentered. In order to be able to interpret  $b_0$ , it is usually necessary to center both the dichotomous  $X$  (by recoding it into  $-0.5$  and  $+0.5$ , or into  $-1$  and  $+1$ ) and the continuous  $Z$  (by "mean-centering" it, i.e., by computing

the average value of all scores and then subtracting this value from every participant's score). When both  $X$  and  $Z$  are centered, the coefficient  $b_0$  in equation (8) represents the main effect of test discipline for the participant with an average score on the covariate. If this coefficient is statistically significant, one would conclude that the students with an average parental income, regardless of their gender, have better grades in one of the two disciplines. The coefficient  $b_1$  in equation (8) is the interaction effect between discipline and gender when controlling for the effect of parental income. A significant  $b_1$  in equation (8) suggests that the gender difference is greater for one of the two disciplines.

Note that none of the data analysis programs, not even the menu-based ones, will automatically mean-center the covariate for the user. It is thus important to manually mean-center the covariate before including it in the analysis. A failure to do so leads to an incorrect interpretation of the main effect of the within-subjects variable in a mixed-model ANCOVA (unless a score of 0 on the covariate is a theoretically meaningful value).

The mixed-model ANOVA is a powerful analytic approach for examining data from complex research designs. It is useful to note that the one of the most common  $2 \times 2$  mixed-model ANOVAs contains one manipulated within-subjects variable, and the between-subjects variable is the order in which the levels of the within-subjects variable were administered (subjects in one order condition first did level 1 and then level 2 of the within-subjects variable, whereas participants in the other order condition started out with level 2 and then did level 1). For more information on this model, refer to the entry on Repeated Measures Designs in this encyclopedia.

Another important issue to note is that researchers sometimes conduct a  $2 \times 2$  mixed-model ANOVA with pre-test and post-test as the within-subjects variable. One



should know that this is not always the best test. Statistical power can be increased by including the pre-test as a covariate, i.e., by regressing the post-test on both the between-subjects variable and the pre-test. However, this data-analytic approach can be chosen only if certain conditions are satisfied (see the entry on [Repeated Measures Designs](#) in this encyclopedia for more details).

### **Model Assumptions for Mixed-Model ANOVAs**

Mixed-model ANOVAs must meet certain assumptions in order to generate unbiased estimates of the main and interaction effects. As with usual ANOVAs, one assumption is that the residuals in both the between-subjects model (equation (3)) and the within-subjects model (equation (4)) must be normally distributed. The second assumption is homogeneity of variances, or homoscedasticity. This assumption holds that the two groups defined by the between-subjects variable have approximately the same error variance. Applying transformations to the data may correct violations of these assumptions.

Mixed-model ANOVAs also have several assumptions that are specific to them. One of these is "homogeneity of the variance-covariance matrices". This assumption is satisfied when the pattern of intercorrelations among the various levels of the within-subjects independent variable(s) is consistent across groups of subjects defined by the levels of the between-subjects independent variable(s). The homogeneity of the variance-covariance matrices assumption is tested using Box's M statistic. If Box's M returns a  $p$ -value that is less than 0.001, then the variance-covariance assumption is violated. Violations of this assumption can be corrected for with data transformations.

The final model assumption of mixed-model ANOVAs is "sphericity" and applies only to models including within-subjects variables with three or more levels. This

assumption is satisfied if the variance of the difference scores for any two levels of the within-subjects independent variable is similar to the variance of the difference scores for any other two levels. Mauchly's test of sphericity can be used to evaluate this assumption, and if it is significant at  $p < 0.05$ , the  $F$ - and  $p$ -values of the coefficients in the mixed-model ANOVA should be adjusted using the Greenhouse-Geisser or the Huynh-Feldt corrections.

### **Extending Mixed-Model ANOVAs to Linear Mixed-Effects Models**

An increasing number of researchers are analyzing data from studies with both within- and between-subjects independent variables as linear mixed-effects models. In this approach, the unit of analysis is the observation rather than the subject. As a result, the data have to be in "long format" (one line per observation) rather than in "wide format" (one line per subject). Data files in the traditional wide format have to be restructured into long format before they can be submitted to a linear mixed-effects model analysis.

When specified correctly and with complete data, the linear mixed-effects model yields the same results (i.e., the same coefficients, the same  $F$ - and  $p$ -values) as the mixed-model ANOVA. And yet, linear mixed-effects models have numerous advantages. They are more flexible in that they allow researchers to analyze the effects of continuous within- and between-subjects variables. They have the ability to incorporate missing data directly (i.e., there is no need to delete incomplete cases or impute for missing values). They can account for multiple sources of nonindependence (e.g., when subjects react to the same set of items). Finally, they allow researchers to relax the above mentioned sphericity assumption under certain circumstances.

*Sohad Murrar and Markus Brauer*

**See also** Analysis of Variance, Analysis of Covariance, Multivariate Analysis of Variance, Repeated Measures Analysis of Variance, Repeated Measures Designs, Two-Way Analysis of Variance

## **FURTHER READINGS**

Judd, C. M., McClelland, G. H., & Ryan, S. (2009). Repeated-Measures ANOVA: Models with Nonindependent Errors. In C. M. Judd, G. H. McClelland, & S. Ryan (Eds.). *Data analysis: A model comparison approach*. (pp. 247-276). New York: Routledge.

Keppel, G. & Zedeck, S. (1989). The mixed two factor design. In R. Atkinson, G. Lindzey, & R. Thompson (Eds.), *Data analysis for research designs*. (pp. 293-314). New York: W. H. Freeman and Company.

Laerd Statistics (2013). Mixed ANOVA using SPSS. Retrieved from <https://statistics.laerd.com/spss-tutorials/mixed-anova-using-spss-statistics.php>

Rutherford, A. (2011). *ANOVA and ANCOVA: A GLM approach*. Hoboken, New Jersey: John Wiley & Sons, Inc.