AJO-DO

# Analysis of variance to linear regression

**Nikolaos Pandis, Associate Editor of Statistics and Research Design**
*Bern, Switzerland, and Corfu, Greece*

In the article discussing 1-way analysis of variance (ANOVA),[1] we compared the forces generated by 3 orthodontic wires. I used a sample of about 10 wires per group and measured the forces generated on a testing machine. Then I used 1-way ANOVA to assess whether there is a difference on average in the forces generated among any of the 3 wires.

Table I shows the ANOVA output and has been already interpreted in detail in the previous article.[1]

Briefly, the output indicates that there is a significant difference ($P = 0.003$) in the forces generated between at least 2 wires.

Table II uses the same data set and variables to explore the same research question with a simple linear regression model instead of 1-way ANOVA. The findings are the same, but also the regression model includes estimates and confidence intervals that are always of interest.

The interpretation of the regression model in Table II is similar to the previously discussed regression model. To clarify, please note that under the β-coefficient column, there are 2 values (wire_2 and wire_3) and the constant. The *P* values shown in the regression output are from the Wald test and correspond to the difference of each wire level to the baseline of the variable (wire_1). In Table I, the *P* value indicates the overall effect of wire type on the forces generated. The ANOVA *P* value indicates that the type of wire is a significant predictor of the forces generated; however, it does not tell us for which wires the forces generated differ significantly from each other. In the regression output, wire 1 is the reference category, and the coefficient values for wires 2 and 3 indicate the difference in forces generated between those wires and wire 1. Therefore, the force generated by wire 2 is 0.29 units higher than that of wire 1, and the force generated by wire 3 is 0.22 units higher than that of wire 1. The constant indicates the force generated by wire 1 (baseline or reference), and we can use this formula:

$$\text{force} = \text{\_cons} + 0.29{*}\text{wire2} + 0.22{*}\text{wire3}$$

In this equation, wire 2 takes the value of 1 if it is studied; otherwise, the value is 0. The same holds for wire 3. So, to calculate the force generated by wire 2, we write:

$$\text{force} = 0.96 + 0.29{*}1 + 0.22{*}0 = 1.25$$

and the force generated by wire 3 is calculated as follows:

$$\text{force} = 0.96 + 0.29{*}0 + 0.22{*}1 = 1.18$$

In a similar way, we can also use multiple linear regression analysis to conduct 2-way ANOVA, and we will use the same data set that we used previously for the 2-way ANOVA model. As a reminder, there are 2 categorical variables: wire type (3 levels) and bracket type (2 levels). We want to assess simultaneously the effect of both bracket type and wire type on the forces generated.

The *P* values from the 2-way ANOVA analysis are shown in Table III. We can see that overall the model and the wire type are significant predictors, whereas bracket type is not a significant predictor of the forces generated. Again, it does not tell us for which wires the forces generated differ significantly. For bracket type since there are only 2 levels, we know that bracket type_1 does not differ significantly from bracket type_2.

Table IV shows the output from the regression analysis.

ANOVA informs us only about the overall effect of wire type and bracket type on the forces generated, whereas regression shows the effect of wires 2 and 3 compared with wire 1 (reference) after adjusting for bracket type, and the effect of labial compared with lingual brackets (reference) after adjusting for wire type. The *P* values shown in the regression output are from the Wald test and correspond to the difference of each wire or bracket level to the baseline of the variable (wire_1 or bracket_1). For bracket type, there is no difference between ANOVA and regression in the *P* value, since this variable has only 2 levels.

A likelihood ratio test can be used to see the overall effect of wire type or bracket type, or both

Assistant professor, Department of Orthodontics and Dentofacial Orthopedics, School of Dental Medicine/Medical Faculty, University of Bern, Bern, Switzerland; private practice, Corfu, Greece.
Am J Orthod Dentofacial Orthop 2016;149:935-6
0889-5406/$36.00

http://dx.doi.org/10.1016/j.ajodo.2016.03.013

**Table I.** One-way ANOVA output for the effect of wire type on forces generated

| *Source* | *SS* | *df* | *MS* | *F* | P *value* |
|---|---|---|---|---|---|
| Between groups | 0.43 | 2 | 0.22 | 7.40 | 0.003 |
| Within groups | 0.76 | 26 | 0.03 | | |
| Total | 1.19 | | 0.4 | | |

*SS*, Sum of squares; *df*, degrees of freedom; *MS*, mean square.

**Table II.** Linear regression output for the effect of wire type on forces generated

| *Wire type* | β-*coefficient* | *95% CI* | P *value* |
|---|---|---|---|
| Constant | 0.96 | 0.85, 1.07 | |
| Wire_2 | 0.29 | 0.12, 0.46 | 0.001 |
| Wire_3 | 0.22 | 0.07, 0.38 | 0.006 |

simultaneously, and those *P* values will be comparable with those shown in the ANOVA table.

The constant shows the value of the force generated with wire 1 and the lingual bracket: ie, the references or baseline groups for wire type and bracket type, respectively.

In addition, interaction terms can also be added to the regression analysis in a similar way discussed in the 2-way ANOVA article. However, the interpretation of the regression table in the presence of interaction terms is a little more complicated and does not follow the usual approach discussed so far.

In the presence of interaction, we need to calculate the estimates (coefficients) and the associated confidence intervals separately. In other words, if the forces generated by the 3 wires are significantly different at the 2 levels of the other variable (type of bracket), we need to calculate separately the estimates for each type of wire in the lingual and labial brackets.

**Table III.** Two-way ANOVA for the effects of wire type and bracket type

| *Source* | P *value* |
|---|---|
| Model | <0.001 |
| Wire type | <0.001 |
| Bracket type | 0.07 |

Only *P* values are shown, and the rest of the output is omitted.

**Table IV.** Linear regression output for the effects of wire type and bracket type

| | β-*coefficient* | *95% CI* | P *value* |
|---|---|---|---|
| Constant | 1.36 | 1.14, 1.38 | |
| Wire type | | | |
| Wire_2 | −0.08 | −0.22, 0.07 | 0.34 |
| Wire_3 | 0.36 | 0.21, 0.50 | <0.001 |
| Bracket type | | | |
| Bracket_2 | −0.11 | −0.23, 0.01 | 0.07 |

## REFERENCE

1. Pandis N. Analysis of variance. Am J Orthodon Dentofacial Orthop 2015;148:868-9.