The Journal of Foot & Ankle Surgery 55 (2016) 1343–1346

Contents lists available at ScienceDirect

The Journal of Foot & Ankle Surgery

journal homepage: www.jfas.org

Investigators' Corner

Random Effects: Variance Is the Spice of Life

Daniel C. Jupiter, PhD

Assistant Professor, Department of Preventive Medicine and Community Health, The University of Texas Medical Branch, Galveston, TX

ARTICLE INFO

Keywords:
ANOVA
random effects
regression
t test
variance

ABSTRACT

Covariates in regression analyses allow us to understand how independent variables of interest impact our dependent outcome variable. Often, we consider fixed effects covariates (e.g., gender or diabetes status) for which we examine subjects at each value of the covariate. We examine both men and women and, within each gender, examine both diabetic and nondiabetic patients. Occasionally, however, we consider random effects covariates for which we do not examine subjects at every value. For example, we examine patients from only a sample of hospitals and, within each hospital, examine both diabetic and nondiabetic patients. The random sampling of hospitals is in contrast to the complete coverage of all genders. In this column I explore the differences in meaning and analysis when thinking about fixed and random effects variables.

In my most recent Investigators' Corner (1), I looked at interaction terms in regression analysis. *Interaction terms* encompass a set of tools that allow us to look more finely at the covariates in our multivariate regression models. We can say, broadly speaking, that interactions allow us to look at "more" levels of our variables. For example, rather than just looking separately at male versus female and diabetes versus nondiabetes we are able, in some respects, to look at all combinations of values of the 2 variables. We concluded the column by mentioning that the variables we had thus far considered in these examples had distinct values that were of particular interest to us (e.g., men vs. women), and that we were examining all such possible values. In contradistinction, I noted that in certain situations we might have variables in our analyses that had actual values not of particular interest. As an example, I mentioned comparing salaries between men and women, from a *random* selection of cities in the United States. In this setting, the difference between Chicago, Illinois, and Galveston, Texas, is not of substantive interest, first because that difference is not the one that we are trying to understand and second because we don't have data on all cities—just the ones in our random sample. Thus, we cannot make any general statements about all possible individual cities.

In this Investigators' Corner I look into including these *random effects* in our models. Why would we even want to include such a variable in our analyses? If we don't think we can or want to say anything about the values of the variable, what could it possibly add of utility to our analysis? Even if we did wish to include such a variable, what information could we glean from it, and how? In order to answer these questions, we need to step back and reconsider the mathematics behind our basic tools, specifically Student's *t* test and the analysis of variance (ANOVA), as well as to rethink why we do

Financial Disclocure: None reported.
Conflict of Interest: None reported.
Address correspondence to: Daniel C. Jupiter, PhD, Assistant Professor, Department of Preventive Medicine and Community Health, The University of Texas Medical Branch, 301 University Boulevard, 1.134G Ewing Hall, Galveston, TX 77555-1150
E-mail address: dajupite@utmb.edu

1067-2516/$ - see front matter © 2016 by the American College of Foot and Ankle Surgeons. All rights reserved.
http://dx.doi.org/10.1053/j.jfas.2016.09.003

multivariate analysis, and what it adds to our understanding of relationships.

The Student's *t* Test and ANOVA

Let us revisit one of the first tests we learn about in biostatistics, the Student's *t* test, but I will try to make its connection to its big brother, ANOVA, a little clearer than normal. The Student's *t* test is designed to compare the means of a continuous variable in 2 groups and tell us if those means are different. Although the test examines *means*, the *variance* of the variable being studied, within each group, sneaks into the picture in a way that is of importance to us. Explicitly, but suppressing some technical details, the formula for the *t* test is

$$\frac{\mu(A) - \mu(B)}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}},$$

where μ(*A*) and μ(*B*) are the mean values in groups A and B, respectively; s_A^2 and s_B^2 are the within group variances in the 2 groups; and n_A and n_B are the numbers of subjects in each of the 2 groups. (This is the most basic form of the *t* test, where we suppress concerns regarding large differences between our 2 groups in terms of within group variances or sample sizes. Although those more complicated forms of the *t* test have accordingly more complicated formulae, the essence of those formulae is the same as that of the formula presented here.)

What does this formula mean? We see the difference between means, which we would expect. And we divide that difference, or variation, by some sort of estimate of the variance within the groups; in fact, it looks almost like a weighted average of the variances in the 2 groups, where we weight by the size of the groups. We have then the variation (or variance!) *between* groups, divided by the variance *within* groups. If this ratio, the *t* value, is large, the corresponding *p* value for the *t* test is small, and we have observed a *significant* difference in means.

A heuristic picture will aid in understanding what is happening here. In Fig. 1*A* we see 2 groups in which the difference (variance) between the mean values (illustrated by the red vertical line) is large and, in particular, is large relative to the within group variance (illustrated by the blue vertical line). Our intuition, translated into mathematics in the formula shown earlier, tells us that we should not hesitate to call this difference between means significant. By contrast, in Fig. 1*B* the variance between groups is small relative to the variance within groups, and we would not call the difference in means significant. The variance within groups is so large that we have trouble deciding whether the difference between groups is real or is just secondary to the large variation within groups. Less within group variability allows us to better understand between group variability.

Now that we have done the heavy lifting of understanding the technicalities of the *t* test, the logic of the ANOVA reveals itself to us easily. Indeed, the ANOVA is really just an extension of the *t* test: rather than comparing the between group variance to the within group variance of 2 groups, we do the same exercise with 3 groups. Given that there are 3 or more groups, rather than 2, we cannot simply look at a difference of means in the numerator. The ANOVA handles this by making explicit what we had implied in talking about the *t* test: the numerator is now a *variance* of (or between) group means, and the denominator remains an estimate of *within* group variances. Essentially, this method is another comparison of means, using variances.

But the story does not end here. Indeed, simple linear regression is just another application of the same idea! Fig. 2 shows a scatterplot. Is the trend we observe statistically significant? We again look at variances: is the variance over the whole data set (the variance *between*

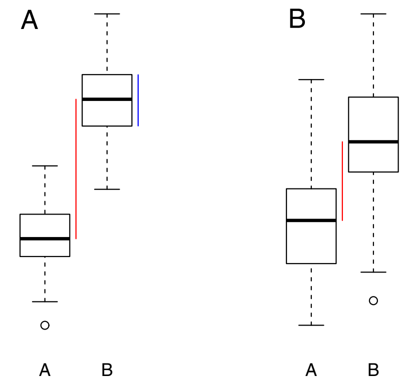

Fig. 1. (*A*) An example of 2 groups where the difference in means is significant. (*B*) An example of 2 groups where the difference in means is not significant. In both figures the red vertical line represents between group variance, and the blue vertical line represents within group variance.

the predicted means of the y values at the various x values) large compared to the variance of y values *within* each x value? As before, these 2 variances are represented with red and blue lines, respectively, and again we compute their ratio. A large ratio yields a significant test, a small ratio a non-significant result. In fact, the statistical test (*F* test) employed in testing significance here is the same as that used in testing significance of our ANOVA ratios.

Multivariate Regression

When I first discussed multivariate regression (2), I introduced it as a tool with which we could ascertain the independent impacts of each variable:

- Is age associated with time to bone union independently of body mass index (BMI)?
- Is BMI associated with time to bone union independently of age?

In other words, the multivariate regression tells us how BMI impacts time to union at any given fixed age, rather than telling us how BMI impacts time to union, in the population overall. These are, at least intuitively, 2 different notions of average:

- The former is an average of the effects of BMI at each age.
- The latter is an average of the effect of BMI in the whole population.

It is in this sense that we view multivariate regression as a way to "control" for the impact of one variable while examining the impact

Fig. 2. A scatterplot, with red vertical line showing deviation of predicted y-values from the mean y-values, and blue vertical line showing deviation at each x-value of the y-values from the predicted y-value.