

■ Commentary

Comments on Statistical Issues in September 2015

Kyung Do Han, Yong Gyu Park

Department of Biostatistics, The Catholic University of Korea College of Medicine, Seoul, Korea

In this section, we explain the terms of MANOVA and eta-squared which used in the article titled, "Non-coronary patients with severe chest pain show more irrational beliefs compared to patients with mild pain," published in July 2015 by Bahreman et al.¹⁾

MANOVA

Multivariate analysis of variance (MANOVA) is the generalization of the analysis of variance (ANOVA) for testing the equality of mean vectors of several populations. In this expression, 'mean vectors' and 'several populations' indicate that there are several dependent (outcome) variables and at least two groups to be compared, respectively. It is well known that ANOVA is used to compare the means of one dependent variable among several groups. The basic assumptions of MANOVA are the distribution of several dependent variables in each group is a multivariate normal and all groups have the equal covariance matrix, which are the generalization of assumptions of a normal distribution and the equal variance in ANOVA.

In Table 4 by Bahreman et al.,¹⁾ there were four dependent variables (hopeless, demand, problem, and emotional), and two groups (severe vs. mild chest pain). The zero hypotheses for MANOVA were 'all four means of dependent variables are not different between severe and mild chest pain groups.' Therefore, the rejection of these zero hypotheses means that one or more dependent variables are significantly different between groups. We can see that the P-value of MANOVA is 0.001 and P-values of ANOVA for three dependent variables are less than 0.05 (hopeless=0.001, problem=0.002, and emotional=0.001).

Pillai's trace, Wilks' lambda, Hotelling-Lawley's trace,

and Roy's greatest root are frequently used test statistics of MANOVA,²⁾ and if two assumptions (normal distribution and equal variances) mentioned above are satisfied, then the distributions of all four test statistics are F and have the similar P-values. All four statistics have the same P-values (see Table 4 by Bahreman et al.¹⁾) and we think that presenting just one of four results might be enough. Results of F tests in the upper four rows are those of ANOVA applied to each dependent variable and are also the results of Student t-tests for each variable because there are only two groups.

ETA-SQUARED

Eta-squared is a measure of effect size used in ANOVA and analogous to R-squared (coefficient of determination) in regression analysis. Eta-squared is defined as the proportion of variance of dependent variable explained by independent variable (eta-squared = between sum of squared/total sum of squared; exactly the same meaning to R-squared). The reason for using the expression 'analogous' is that eta-squared could be used to express the non-linear relationship while R-squared is used only for linear one. For example, consider the following data.

Group 1: 1, 2, 3 (mean 2); group 2: 7, 8, 9 (mean 8); group 3: 4, 5, 6 (mean 5).

If we code the values of independent variable (group) as 1 (group 1), 2 (group 2), and 3 (group 3), values for eta-squared and R-squared are 0.90 and 0.225, respectively. However, if we code the values as 1 (group 1), 3 (group 2), and 2 (group 3), then both eta-squared and R-squared are the same values, 0.90.

CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

REFERENCES

1. Bahreman M, Saeidi M, Komasi S. Non-coronary patients with severe chest pain show more irrational beliefs compared to patients with mild pain. *Korean J Fam Med* 2015;36:180-5.
2. Johnson RA, Wichern DW. *Applied multivariate statistical analysis*. Englewood Cliffs (NJ): Prentice-Hall Inc.; 1982.