
# StatNews #37

## Interaction and Association
## January 2000

A common misunderstanding concerns the relation of the concepts of interaction and association, and is manifested by the idea sometimes expressed that two variables cannot interact if they are not associated with each other. The purpose of this article is to review the concepts of interaction and association, and to present a simple example that demonstrates the distinctness of these concepts.

Interaction is the condition where the relation between an explanatory and a response variable is different at different values of another explanatory variable. Association means that information about the value of one variable conveys information about the average value of another variable.

To illustrate these ideas, consider a study in which three variables are observed. The explanatory variables are the presence or absence of a parasite in a host, and the sex of the host. The response variable is the vigor of the host. We might be interested in 1) whether the presence or absence of the parasite is associated with the sex of the host (e.g., do females have more parasites), a question of association, and 2) whether the effect of the presence of the parasite on the vigor of the host is different for the two sexes (e.g., does having the parasite reduce vigor more in males than in females), a question of interaction.

To carry out this study, we collect data on these three variables, and form a two-by-two table of presence versus absence of parasite crossed with female versus male. For each of the four cells formed by this table, the sample size can be labelled as N1, N2, N3, and N4 and the mean vigor as M1, M2, M3, and M4, where the subscripts 1 to 4 refer to, respectively, presence/female, presence/male, absence/female, and asbsence/male.

| | | **Sex of Host** | |
|---|---|---|---|
| | | **FEMALE** | **MALE** |
| **Parasite** | PRESENT | cell 1 | cell 2 |
| | ABSENT | cell 3 | cell 4 |

To answer the question of association, we assess whether N1/N3 is equal to N2/N4, with equality implying no association. That is, if the proportion of females with parasites equals the proportion of males with parasites, then there is no association. This would usually be done by calculating the Pearson's Chisquare statistic from the contingency table of cell frequencies. To answer the question of interaction, we assess whether M1-M3 is equal to M2-M4, with equality implying no interaction. That is, if having the parasite reduces vigor in females the same amount as it reduces it in males, then there is no interaction. This would usually be done by an analysis of variance to compare the means.

The information required to assess association involves only the sample sizes whereas the information required to assess interaction involves only the means. Therefore, it would certainly be possible to observe an interaction (or no interaction) regardless of whether or not there was an association. A question as to whether two explanatory variables are associated is distinct from a question as to whether those two variables interact in their effect on a response variable.

Author: Edward Frongillo

Back to StatNews Table of Contents

(This newsletter was distributed by the Cornell Statistical Consulting Unit. Please forward it to any interested colleagues, students, and research staff. Anyone not receiving this newsletter who would like to be added to the mailing list for future newsletters should contact us at cscu@cornell.edu. Information about the Cornell Statistical Consulting Unit and copies of previous newsletters can be obtained at http://www.cscu.cornell.edu).