Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112.

3622 citations in Google Scholar as of 4/1/2016

In preparing for this statement, I reread the paper. I must say it reads very well and is quite lucid in its exposition. I believe we have the then-editors of *Psychometrika* to thank for this in that, contrary to editorial strictures currently imposed, they did not demand that we shorten the paper. The pace of the exposition and the examples presented made it possible for any interested reader to apply the methods we were describing. (Samuel W. Greenhouse; February 2, 1982; Citation Classic Commentary)

Paul Horst and Dorothy Adkins were the two co-editors of *Psychometrika* when the Greenhouse-Geisser paper was reviewed and published.

What are Profile Data?

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112.

Profile data as discussed in the Greenhouse and Geisser paper can be characterized by subjects within groups observed over a battery of tests (or one test, say, that is repeatedly measured over a number of occasions).

Three questions are typically of interest:

a) are the group profiles "parallel"? — here, "parallel" refers to the group means being equidistant at each measurement occasion;

b) assuming that profiles are parallel, are they also at the same level (that is, are they "coincident")?

c) assuming coincident profiles, are they also "horizontal"?

Assumptions

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112.

It is assumed that each individual profile is a random vector sampled from a p-variate normal distribution with arbitrary variance-covariance matrix.

It is also assumed implicitly that the p (test) variables are commensurable (that is, the variables have the same metric).

This last assumption allows meaning to be given to the question of whether the profiles have the same "shape" (here, profiles are said to have the same "shape" when they are parallel and the tests are commensurable).

By assuming the particular form for the variance-covariance matrix called compound symmetry (where the tests have equal variances and are equally correlated in pairs), the classical mixed model for g samples is generated.

Analysis-of-Variance Terms and Tests

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112.

3622 citations in Google Scholar as of 4/1/2016

The resulting analysis-of-variance table would include sources for Tests, Groups, Individuals (within groups), Group x Test; and Individual x Test (within groups)

The number of tests is p; the total number of individuals is N; the number of groups is g

a) Parallel (same shape) group profiles (note that an evaluation of the group-test interaction is really an evaluation of whether group profiles have the same shape):

$$\text{MS(GxT)/MS(IxT) (w. groups)} \sim \mathrm{F}_{(p-1)(g-1),(p-1)(N-g)}$$

F-Tests Continued

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112.

b) Coincident group profiles:

MSGroups/MSInd (w. groups) $\sim \mathrm{F}_{(g-1),(N-g)}$

This test is equivalent to first obtaining a single score for each individual by summing over tests and then performing a one-way analysis-of-variance on the resulting scores.

c) Horizontal group profiles:

MSTests/MS(IxT) (w. groups) $\sim \mathrm{F}_{(p-1),(p-1)(N-g)}$

The Problem With Compound Symmetry

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112.

3622 citations in Google Scholar as of 4/1/2016

When the assumption of compound symmetry for the covariance matrix does not hold, the F-tests for parallel group profiles and for horizontal group profiles are no longer valid.

In fact, both tests are too liberal and would reject the corresponding (null) hypotheses too often; or, in other words, the obtained p-values are too small.

The F-test for coincident group profiles is not affected and remains appropriate because it is based on summed scores for individuals.

The Genius of Greenhouse-Geisser

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112.

3622 citations in Google Scholar as of 4/1/2016

The Greenhouse-Geisser paper showed how to correct the degrees-of-freedom for the tests for parallel and horizontal profiles based on an estimated function, ϵ, obtained from the (sample) variances and covariances among the tests (which is no longer assumed to have a compound symmetry form).

For the parallel profiles test, the F-distribution used would be $F_{\epsilon(p-1)(g-1),\ \epsilon(p-1)(N-g)}$; for the horizontal group profiles test, the F-distribution used would be $F_{\epsilon(p-1),\ \epsilon(p-1)(N-g)}$.

Because $1 \geq \epsilon$, the numerator and denominator degrees-of-freedom are both discounted, making the tests more conservative (and, therefore, more appropriate when compound symmetry does not hold);

when compound symmetry does hold, $\epsilon = 1$, and no discounting of the degrees-of-freedom occurs.

A Lower Bound for ϵ

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112.

The Greenhouse-Geisser paper also includes a lower bound for ϵ (that is, $\epsilon > \frac{1}{p-1}$).

Thus, it is possible to discount the degrees-of-freedom maximally, giving the most conservative tests we would need: use $F_{g-1,N-g}$ to test for parallel profiles; use $F_{1,N-g}$ to test for horizontal profiles.

The argument then continues as follows: if one rejects with the maximally discounted degrees-of-freedom, rejection would also occur with any value of ϵ greater than its lower bound.

The Importance of the Greenhouse-Geisser Paper for its Time

The Greenhouse-Geisser paper was important for the analysis of profile data in an era when computation was rather primitive and done without electronic computers.

The mixed model for multiple groups assuming compound symmetry leads to an analysis-of-variance table that was well within the numerical capabilities of the mechanical calculators of the time.

More importantly, the various tests could be modified to mitigate the effects of having to make the compound symmetry assumption.

How Profile Data May be Analyzed Today

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112.

As is now well-known to anyone taking an applied multivariate analysis class (based, for example, on a text like Johnson and Wichern, *Applied Multivariate Statistical Analysis*), profile analysis can be done in an alternative manner with various multivariate analysis-of-variance techniques on difference scores.

This approach requires significant numerical effort involving sample variance-covariance matrices that soon becomes prohibitive without computers and currently available statistical routines (in SYSTAT, SPSS, R, and Matlab, for example).

Greenhouse-Geisser is Still Relevant Today

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112.

There is still one important contemporary use for the Greenhouse-Geisser approach even in the face of all the computational power we now have available.

This is where the number of individuals is less than the number of tests; in these instances, the multivariate procedures are impossible to carry out because the necessary degrees-of-freedom are zero or negative for some of the F-approximations.

Examples of profile analysis data where subjects are fewer than the number of observation occasions, abound in neuroimaging analyses done using fMRI data.

Seymour Geisser Didn't Think Much of Us

This not-so-nice comment by Geisser is taken from an interview published in *Statistical Science* in 2007:

Wes: What kind of problems did you work on at NIH?

Seymour: Well, at NIH, with Sam Greenhouse, I wrote my most infamous paper. "Infamous," I say because it wasn't a very important or very hard paper. It was just a paper that seemed to have caught on with social scientists and some medical people. It was just this profile analysis paper which ended up being a citation classic, which means that it had a lot of citations. It still has more citations than all of my [other] papers altogether. [Laughs all around.]

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112.

3622 citations in Google Scholar as of 4/1/2016

Wes: This was the Greenhouse-Geisser paper?

Seymour: There are two papers. The first paper was in the *Annals*, which actually worked out all of the quadratic forms, their expectations, and the mathematics. And the second was in *Psychometrika* and that was the citation classic. That was just to show the methodology—how to use this. It wasn't a very big deal. I worked much harder on other papers and I think I produced much better work. But de *gustibus non disputandum est* [there's no accounting for taste].