Dr. G.W. Snedecor,
Department of Mathematics,
Iowa State College,
Ames, Iowa, U.S.A.

Dear Snedecor,

I was glad to have your letter of May 25th. If you are still satisfied with the statement to which I called your attention, there is no great point in my writing further about it: for I think you have before you the material for seeing why I should consider it misleading.

You say you were disconcerted at not finding in the new portion of Section 49.1 added to the fifth edition, the tests used by Bartlett for the significance of differences among the regressions; I do not quite understand why this should have disconcerted you, seeing that Chapter 5 has, from the first edition, been devoted to showing how such tests of discrepancies among regressions may be treated in a way exactly parallell with the corresponding tests of means. It would have seemed to me out of place to repeat this discussion in Section 49.1, not only because it had been given before, but because the tests considered would not be appropriate for the experimental problem I was discussing.
I need not say that any given set of degrees of freedom may be subdivided into independent components in an infinitude of ways. What perhaps needs stressing is that a uniquely relevant subdivision can only exist in experimental data when the experiment has been properly designed with randomisation in relation to the subdivision proposed. With observational data such as you are principally interested in, the analysis can only be regarded as equally comprehensive on the supposition, which may or may not be true, that the data are, in fact, equivalent to such a set of experimental observations: a supposition, in fact, that nature has supplied effective randomisation, and that certain residual components may properly be regarded as due to chance.

In such cases, additional tests designed to examine the truth of this supposition may be not out of place, though it would be a mistake to think that any limit could be set to these additional tests, or that any single system of them such as that proposed in your letter, could ever be exhaustive; e.g., your first class \( \hat{s} \) may be divided into \( k - 2 \) groups of \( n \) degrees of freedom each, corresponding to the \( n \) which you have separated for regressions \( \hat{s}_2 + \hat{s}_4 \), and with exactly the same justification.

Your statement that the test in my book is not in agreement with Bartlett's about degrees of freedom seems merely to refer to the fact that in my example, \( s_1 \) and \( s_2 \) are both error, and \( s_3 \) and \( s_4 \) are both treatment effects. \( s_4 \) is
merely a particular component of treatment effects, influenced by the independent variate, and in consequence, obtained with somewhat lowered precision, owing to the errors of estimation of the correction. \( s^2_1 \) is in just the same position as the other components of error since it is in fact merely a component of the discrepancies among the comparisons between different treatments in the parallel trials.

You speak as though the theory of the analysis of covariance were in some respect less complete than that of the analysis of variance. That I cannot altogether understand, since the comparisons between regressions and between them and the residual error were among the first examples of the analysis of variance used. What seemed to need attention drawn to it was the simple fact that sums of products could be subdivided in parallel with sums of squares, and that to do so was the first step towards making whatever further work seemed desirable, easy and intelligible.

Yours sincerely,