Perceptual and Motor Skills, 1985, 61, 420-422.

SNAPSHOT ANALYSIS OF VARIANCE: COMPARING GROUPS WITH UNEQUAL NUMBERS OF SCORES PER SUBJECT

DAVID J. WEISS

California State University, Los Angeles[1]

Summary.—In ongoing research, at a given time some subjects will have produced more scores than others. It may be desirable to conduct a snapshot analysis of the available data to compare the efficacy of the treatments being administered. For balanced designs, a weighted-means analysis which incorporates nesting is proposed, and an example is given.

In applied settings, it is not uncommon for some subjects to be farther along in the course of treatment than others. If measures on each subject are taken periodically, then at any given time some subjects will have produced more scores than others. If we wish to compare the efficacy of several treatments prior to the conclusion of the data collection, this inequality in the number of scores per subject can complicate the analysis.

For example, suppose we are evaluating behavioral interventions in a clinical setting. As qualified patients come along, they are inducted into the study and randomly assigned to a particular intervention condition. The patients are individually measured on a monthly basis, with each patient scheduled for a year of treatment. Induction might proceed for a year, so it would take two years to collect all of the scores. However, it would often be necessary and usually be prudent to take a snapshot of the data that have been collected after, say, six months or one year. The results might be used to justify further funding, or alternatively, to avoid wasting further efforts on an unpromising line of research.

Let us suppose that the goal is to compare the effects of the treatments at the present moment. A simplistic solution to the statistical problem would be an analysis of variance using the average score for each subject. The defect here is that the average score for a patient who has produced twelve monthly scores should be weighted more heavily than the average score for one who has produced only three monthly scores. Other simplistic schemes, such as using each subject's first score or last score, are even more objectionable because the single scores are inherently less reliable than aggregate scores.

A statistically defensible, yet fairly simple solution is available if the experimental design is balanced. At the time of the snapshot, there must be in each treatment condition the same number of subjects who have produced a given number of scores. This requirement can be assured by using random

---

[1] Department of Psychology, 5151 State University Drive, Los Angeles, CA 90032.

permutations for the assignment of subjects to conditions; at worst, there will be a few extra scores at the time of the analysis. One can either ignore them or wait for balance.

The balance allows the design to have a proportional cell-size structure, which has the statistically desirable property of orthogonality (Kempthorne, 1952). A sensible weighting is achieved, and the standard principle for generating the error term in a mixed design (Dixon, 1970, pp. 503-504) applies.

The technique is illustrated with the artificial data in Table 1. The computations are carried out as though there were a series of one-way designs.

TABLE 1

AN ILLUSTRATIVE EXAMPLE

Time	Treatment 1				Treatment 2				Treatment 3			
	P_1	P_6	P_7	P_{12}	P_2	P_4	P_9	P_{10}	P_3	P_5	P_8	P_{11}
1	7	6	5	2	6	3	4	7	5	2	3	4
2	5	8	9	8	4	3	5	5	8	2	5	2
3	4	4	7		6	3	5		4	3	4	
4	6	7			5	4			5	3		
5	9	3			5	3			6	5		
6	8				3				4			
Σ	39	28	21	10	29	16	14	12	32	15	12	6

Note.—Subjects have been assigned in groups of three. The subscripts used here for patients (P) are intended to convey their actual order of induction into the study.

The nested source, patients, is (temporarily) presumed to have four levels and be crossed with treatments.

$$
\begin{aligned}
SS_{\text{Treatment}} &= (T_1^2 + T_2^2 + T_3^2)/16 - (T^2/N) \\
&= (98^2 + 71^2 + 65^2)/16 - 1140.75 = 38.625\,. \\
SS_{\text{Patients}} &= (P_1^2)/18 + (P_2^2)/15 + (P_3^2)/9 + (P_4^2)/6 - T^2/N \\
&= 100^2/18 + 59^2/15 + 47^2/9 + 28^2/6 - 1140.75 = 22.98 \\
SS_{\text{TxP}} &= (P_1^2)/6 + (\mathrm{P}_6^2)/5 + (P_7^2)/3 + (P_{12}^2)/2 + (P_2^2)/6 + \ldots \\
&\quad + (P_{11}^2)/2 - SS_T - SS_P - T^2/N \\
&= 39^2/6 + 28^2/5 + 21^2/3 + 10^2/2 + 29^2/6 + \ldots + 6^2/2 \\
&\quad - 38.625 - 22.983 - 1140.75 = 15.31 \\
SS_{\text{Residual}} &= \Sigma X^2 - T^2/N - SS_T - SS_P - SS_{T\times P} \\
&= 1310 - 1140.75 - 38.625 - 22.983 - 15.31 = 92.33.
\end{aligned}
$$

To construct the error term against which $MS_{\text{Treatment}}$ will be tested in the analysis of variance, pool SS_{Patients} with $SS_{\text{T}\times\text{P}}$: then divide the sum by the pooled *df* to find MS_{Error}. This is not the true pooling of variances recommended in some situations to provide more *df* for the error term; the present pooling is merely a computational device which can be used for any nested design (Dixon, 1970). For the sample data, $F_{2,9} = 19.31 / 4.25 = 4.54$. This F ratio measures the effectiveness of the treatment variable.

The term labeled "Residual" includes the variation over time periods, along with interactions involving periods. Because this term is not meaningful with differing numbers of periods for the various patients, the *df* (36 in the example) are wasted in this analysis. This statistical inefficiency is the price for the snapshot analysis. When the experiment is completed and all of the scores are available, then time periods will be an additional factor in an ordinary mixed design. It is worth noting that even if a sleepy researcher were to apply the present analysis to a complete design, the correct F ratio for treatments would be obtained.

REFERENCES

DIXON, W. J. (Ed.) *BMD: biomedical computer programs.* Berkeley, CA: Univer. of California Press, 1970.

KEMPTHORNE, O. *Design and analysis of experiments.* New York: Wiley, 1952.

Accepted July 17, 1985.