Journal of Psychophysiology 3 (1989) 193–194

Synopsis of repeated measurement analysis

Chr. Jennen-Steinmetz
Zentralinstitut für Seelische Gesundheit, Postfach 5970, 6800 Mannheim

In psychophysiology a usual design is the repeated measures design. It arises if a variable is repeatedly observed for the same subjects (or experimental units). For example, repeated observations can correspond to different experimental conditions or different times. Often these repeated measures arise naturally, e.g. in EEG data several observations for each subject arise through the use of multiple electrodes. Further, such designs are frequently preferred to factorial designs with independent groups since sample size may be reduced and since the effect of interindividual variation is decreased. Factors referring to repeated observations are called within-subject factors. Experimental factors dividing the subjects into independent groups are called between-subject factors. The statistical analysis of repeated measures data by analyses of variance procedures requires precautions as is well known. In the psychological literature many papers have appeared which study the methodological issues of repeated measurement analysis of variance. A good exposition of many aspects is given by Rogan, Keselman and Mendoza (1979). There also, further references can be found. Here the main issues will be reviewed briefly. A more elaborate discussion of the problems attached to repeated measures analysis of variance is given by Vasey and Thayer (1987) and the comments on that paper.

Since repeated observations on the same subject are not statistically independent, these designs may not be analysed by the usual analysis of variance for factorial designs. Appropriate statistical procedures are multivariate analysis of variance (MANOVA) or univariate mixed model (split-plot) analysis of variance. Both methods assume that the observations across subjects are statistically independent of each other, and that they are multivariate normally distributed with the same covariance matrix in each group.

Multivariate analysis of variance. If the covariance matrix for the repeated observations is the same for all groups and if the number of subjects is large enough (number of subjects − number of groups > number of repeated measurements), the MANOVA method may be applied.

Univariate analysis of variance. Traditionally, univariate analysis of variance has been applied to analyse repeated measures data. This procedure is based on a mixed-model (split-plot) approach. The traditional univariate F-tests for effects related to between-subject factors (main effects or interactions) are only valid if the sphericity condition (sometimes also called circularity) is satisfied. This is a rather technical condition on the covariance matrix of the observations (see Rouanet and Lépine (1970), Huynh and Mandeville (1979) or Rogan, Keselman and Mendoza (1979)). If this assumption is violated, the F-tests are liberal, i.e. they exceed the stated alpha levels. To keep the nominal level of significance, several adjustments have been proposed typically leading to a reduction of the degrees of freedom of the F-tests. The best known procedures are the ε-corrections due to Greenhouse and Geisser (1959) or Huynh and Feldt (1976). These adjusted tests keep approximately the correct level of significance. It should be noted that for designs with more than one within-subject factors, there are several sphericity conditions and several adjustment indices ε (Mendoza, Toothaker and Crain (1976).

Various tests on sphericity have been proposed. This may suggest to first test on sphericity, and dependent on the results to apply an adjusted or unadjusted F-test. But this procedure cannot be recommended. The tests on sphericity may have poor power (Rouanet and Lepine (1979)), and as a consequence the overall error rate of the procedure can get out of control. Furthermore, in case of sphericity the correction factor ε is generally close to one such that there is actually no big difference between adjusted and unadjusted tests.

Tests on group effects. There is no trouble with tests on effects of between-subject factors, since no assumptions on the structure of the covariance matrix are required. Furthermore, the univariate and multivariate tests coincide. The latter applies as well to tests on main effects of within-subject factors, having only two levels.

Strategies for data analysis. For a given repeated measures data set one has to decide whether to apply adjusted or unadjusted univariate tests or multivariate tests. Since all procedures are now provided by several statistical packages, there is no reason to stick to the traditional univariate F-tests. As the validity of these unadjusted F-tests rests on sphericity, they are only appropriate if there are strong arguments that this assumption holds. For most data with more than two levels of the within-subject factors this condition is violated. Furthermore, since in case of sphericity adjusted and unadjusted tests do not differ much, the recommendation is against the use of unadjusted univariate tests.

It remains to choose between the multivariate procedure and adjusted univariate tests. Various studies have compared the power of the two procedures (Huynh (1978), Rogan, Keselman and Mendoza (1979)), but they do not allow a general recommendation in favour of one method. One might be led to apply one method, and if it does not reach significance to try the other one. It should be stressed that such a procedure leads to an inflated type I error. The procedure is only correct if a Bonferroni correction (α-adjustment) is applied, i.e. the significance levels for the individual tests are halved.

Here, the emphasis was on the sphericity assumption. But it should be remembered that all procedures above assume multivariate normality with homogeneous covariance matrices across groups. The tests are fairly robust against violations of multivariate normality, but heterogeneous covariance matrices can affect the validity of the tests, if sample size differs much between groups.

References

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

Huynh, H. (1978) Some approximate tests for repeated measurement designs. *Psychometrika, 43*, 161–175.

Huynh, H., & Feldt, L.S. (1976) Estimation of the Box correction for degrees of freedom from sample data in the randomized block and split-plot designs. *Journal of Educational Statistics, 1*, 69–82.

Huynh, H., & Mandeville, G.K. (1979) Validity conditions in repeated measures designs. *Psychological Bulletin, 86*, 964–973.

Mendoza, J.L., Toothaker, L.E., & Crain, B.R. (1976) Necessary and sufficient conditions for F ratios in the L × J × K factorial design with two repeated factors. *Journal of the American Statistical Association, 71*, 992–993.

Rogan, J.C., Keselman, H.J., & Mendoza, J.L. (1979) Analysis of repeated measurements. *British Journal of Mathematical and Statistical Psychology, 32*, 269–286.

Rouanet, H., & Lépine, D. (1970) Comparison between treatments in a repeated-measurement design: ANOVA and multivariate methods. *The British Journal of Mathematical and Statistical Psychology, 23*, 147–163.

Vasey, M.W., Thayer, J.F. (1987) The continuing problem of false positives in repeated measures ANOVA in psychophysiology: A multivariate solution. *Psychophysiology, 24*, 479–486.