MIXED MODEL REPEATED MEASURES ANALYSIS OF VARIANCE FOR MILK YIELD DATA

J. V.Ammu[1], V. L.Gleeja[2], K. A. Mercey[3] and A. Prasad[4]
Department of Statistics, College of Veterinary and Animal Sciences, Mannuthy, Thrissur-680651, Kerala

Received: 26.02.2018
Accepted:14.04.2018

Abstract

In most of the animal experiments, observations are taken from the animals at different points of time. Such data is known as repeated measures data and analysis of such data is to be carried out using mixed models. Determination of the most appropriate covariance structure is very importantwhile using mixed models. In present days, some of the journals like Canadian Journal of Animal Science accept only those papers which are reporting mixed procedure to analyse the data sets that include repeated measurements. Hence, a study was conducted about the use of mixed models for repeated measures data. In this paper the method is illustrated with milk yield data.It is recommended that the mixed model be used for analysis of repeated measures designs in animal studies.

Key words: *Mixed model, Repeated measures, Covariance structure*

Repeated measures data is common in animal sciences research (Goonewardene *et al*. 2000, ZoBell*et al.* 2003). They are the observation of same characteristics, which are made in several periods on same experimental unitswhich may be an animal. Thus, the measurements on the same animal or experimental unit are dependent. So, it is essential to define an appropriate covariance structure for such measurements. Conventionally, repeated measures data were either analysed as a univariate or a multivariate analysis of variance (ANOVA), both being handled by the general linear model procedure. Univariate ANOVA basically assumes equal variance and covariance over time and usually this assumption is not true. In multivariate ANOVA, observations are independent, so data with dependent observations will be excluded from analysis. Accommodating covariance structure in univariate and multivariate ANOVA is not an ideal method. Therefore, mixed model has become appealing for analysing repeated data and should be used for analysis. Hence, the present study was taken to discuss the use of mixed models in repeated measures data.

1. PG Scholar
2. Assistant Professor
3. Professor
4. Assistant Professor and Head, Cattle breeding farm, Thumburmuzhy, email: prasad@kvasu.ac.in

Materials and Methods

In an experiment, (Prasad, 2014) six treatments were randomly assigned to 30 cows and milk yields (in litres) were measured over nine weeks. Observations were available on initial milk yield, so they have been used as covariates. The objective was to find out the effect of treatments, weeks and interaction effect of treatment and weeks. In this experiment, observations were obtained from the same animal on different weeks and hence those observations were dependent. Hence, the most suitable method for analysis of this data is repeated measures ANOVA with the mixed model.

Mixed model states that observed data are of two parts, fixed effects and random effects. Fixed effects define the expected value of observations (mean) and random effects define the variance and covariance structure of the observations.The linear mixed model is, where,*y*=(be an n x1 vector of independent observations, *X* is an n x p model matrix, is a p x 1 vector of unknown parameters, = (be an n x 1 vector of errors,Z is a given n x q matrix, is an unobservable random vector of dimensions q x 1. And the random components of model are the vectors and∈.Some of the commonly used covariance structures for the repeated measures model are simple, compound symmetry (CS), autoregressive (AR(1)), ante dependence (ANTE(1)) and unstructured (UN). Simple structure is not realistic for most repeated measures data because it specifies that observations on the same animal are independent. The CS covariance structure is only appropriate when the so-called Huynh-Feldt condition is met, that is an equal correlation between measures on the same subject (Huynh and Feldt 1970, 1976). The AR (1) structure requires equally spaced times, and time must be ordered correctly and the structure needs only two parameter estimates. The ANTE (1) covariance structure allows unequal variances over time and unequal correlations and covariance among different pairs of measurements. With this structure, time periods must be ordered correctly but the equal spacing between times is not necessary. UN covariance structure makes no assumptions about equal variances or correlations over time and this is the most complex structure.

The choice of an appropriately parsimonious covariance structure can improve the efficiency of inferences concerning the mean structure and provide better estimates of standard errors of estimated parameters. The information criteria used in mixed models can be used as a statistical tool to assist a model selection. Different types of information criteria are REML log likelihood (REML logL), Akaike information criteria (AIC), finite sample corrected Akaike information criteria (AICC) and Schwarz's Bayesian criterion (BIC). Among possible within-subject covariance models, the model that minimizes AICC or BIC is preferable. Simpler covariance models are preferred when AICC and BIC values are close. The penalty imposed by BIC is more severe than the one imposed by AIC. The value of information criteria closest to zero indicates a better model fit to the data (SAS Institute, Inc. 1999). That is, covariance structure whose goodness of fit value which is closest to zero is chosen as the best covariance structure. The analysis of data for illustration of the method is done using Statistical analysis system (SAS) version 9.2.

Results and Discussion

Milk yield of 30 cows was observed over nine weeks. As the measurements were taken from the same animals on nine different weeks, the observations from the same animals over days were dependent or correlated.

Table 1. Model fit statistics of milk yield

Criteria	SIMPLE	CS	ANTE(1)	AR(1)	UN
-2 Res Log Likelihood	578.1	450.6	350.3	386.6	291.2
AIC	580.1	454.6	384.3	390.6	381.2
AICC	580.1	454.6	387.3	390.6	404.4
BIC	581.5	457.4	408.2	393.4	444.2

Hence, mixed model repeated measures ANOVA has to be used.A linear mixed model with an appropriate covariate structure put up the features shown in the data. In order to illustrate the mixed model analysis, the same data set were fitted to five different covariance structures and the model fit statistics are given in Table 1.

Since AICC is less for ANTE (1), it is chosen as the appropriate covariate structure for this particular data even though AIC is lesser for unstructured covariance (UN) structure. It was also because, UN causes over parameterization, which may waste information.

The results of fixed effects for the data on milk yield are shown in Table 2. As p-value for F-test for the milk yield with respect to treatment and week is less than 0.01, it concludes that statistically there exists significant difference in mean milk yield between treatment and over weeks. But there is no significant interaction effect of treatment andweek. Mixed models allow covariate in the model. Also the initial milk yield which is taken as covariate is significantly influencing the milk yield of different treatments over weeks.

For the results of Table 3, fourth treatment is significantly different from second and fifth treatment. Also first and third treatment doesn't show any significant difference from all other treatments. When weeks are considered, milk yield is highest for fifth week. Milk yield for eight and ninth week is very low compared to the first seven weeks. Standard errors of the means with ANTE (1) model are similar within the same time period, and increases over weeks.

A key strength of repeated measurement studies is that it is the only type of design which could obtain information concerning individual patterns of change. However, in practice, repeated measurement studies are described by heterogeneous variances, missing dataand time-dependent covariates. These features make the classical multivariate approaches difficult to apply. The ANOVA method ignores the time-dependent correlations and type I error, which is common in this method. The multivariate ANOVA method cannot directly accommodate the appropriate covariance structure. It assumes unstructured covariance structure, but wastes a large amount of information and makes the tests less powerful. Mainly, these methods cannot handle missing data.

Table 2. F test of fixed effects for milk yield

Effect	DF	F Value	p-value
Treatment	4	7.08**	0.0008
Week	8	13.18**	<.0001
Treatment*Week	32	1.62	0.0803
Initial milk yield	1	258.84**	<.0001

*** significantly different at 1% level*

Table 3. Average milk yield over weeks for five treatments

Treatment	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Overall treatment mean
1	7.99± 0.24	7.93± 0.26	7.80± 0.29	8.04± 0.24	8.72± 0.30	8.32± 0.33	7.92± 0.31	7.67± 0.33	7.52± 0.36	7.99± 0.23^{AB}
2	7.72± 0.24	7.65± 0.26	7.25± 0.30	7.35± 0.24	7.60± 0.30	7.04± 0.33	6.81± 0.31	6.57± 0.33	6.36± 0.36	7.15± 0.23^{B}
3	8.12± 0.24	8.13± 0.26	7.84± 0.30	8.32± 0.25	8.75± 0.30	8.59± 0.33	8.28± 0.30	7.73± 0.33	7.28± 0.36	8.09± 0.23^{AB}
4	8.76± 0.24	8.73± 0.26	8.49 ± 0.30	8.69± 0.25	9.13± 0.30	9.03± 0.33	8.70± 0.31	8.38± 0.33	8.12± 0.36	8.67± 0.23^{A}
5	8.44± 0.24	7.84± 0.26	7.69± 0.30	7.61± 0.24	7.40± 0.30	7.08± 0.33	6.68± 0.31	6.19± 0.33	6.34± 0.36	7.26± 0.23^{B}
Overall week mean	8.21± 0.10^{ab}	8.05± 0.11^{abc}	7.81± 0.14^{bc}	8.01± 0.11^{bc}	8.32± 0.14^{a}	8.0 ± 0.15^{bc}	7.68± 0.14^{c}	7.31 ± 0.16^{d}	7.13± 0.17^{d}	

Means having same superscripts (small letters a-d within a row, capital letters A-B within a column) doesn't differ significantly at 5% level.

So, an alternative procedure based on the linear mixed model in the analysis of animal experiments with repeated measures data is considered. Mixed models have the ability to handle missing data and unequal spacing over time. Importantly it allows covariance analysis in the model. Therefore, mixed model analysis is more precise and accurate, even though assessing an appropriate covariance structure for the data is not an easy task. In a similar study, Wang and Goonewardene (2004) studied the use of mixed models in analysis of animal experiments with repeated measures data. They also recommended that the mixed model should be used for the analysis of repeated measures design in animal experiments.

References

Goonewardene, L. A., Price, M. A., Stookey, J. M., Day, P. A. and Minchau, G. 2000. Handling, electric goad and head restraint: Effects on calves' behaviour. *J. Appl. Anim. Welfare Sci.*3: 5–22.

Huynh, H. and Feldt, L.S. 1970. Conditions under which mean square ratios in repeated measurements designs have exact F-distributions. *J. Am. Stat. Assoc.*65: 1582-1589.

Huynh, H. and Feldt, L.S. 1976. Estimation of the box correction for degrees of freedom from sample data in the randomized block and split plot designs. *J. Edu. Stat.*1: 15–51.

Prasad, A. 2014. Climatic adaption and stress evaluation of cross breed cattle of Kerala. *Ph. D. thesis*, Kerala Veterinary and Animal Sciences University, Pookode, 134p.

SAS Institute, Inc., 1999. *SAS/STAT user's guide*. Version 8. SAS Institute, Inc., Cary, NC.

Wang, Z. and Goonewardene, L.A. 2004. The use of mixed models in the analysis of animal experiments with repeated measures data. *Can. J. Anim. Sci.* **84**: 1-11.

ZoBell, D. R., Goonewardene, L. A., Olson, K. C., Stonecipher, C. A. and Weidmeier, R. D. 2003. Effects of feeding wheat middlingson production, digestibility, ruminal fermentation and carcass characteristics in beef cattle. *Can. J. Anim. Sci.*83: 551–557. ■