Statistical procedures most used in the analysis of measures repeated in time in the agricultural sector

Sarai Gómez[1], Verena Torres[1], Yoleisy García[1] and J.A. Navarro[2]

[1]*Instituto de Ciencia Animal, Apartado Postal 24, San José de las Lajas, Mayabeque, Cuba*

[2]*Facultad de Medicina Veterinaria y Zootecnia, Universidad Autónoma de Yucatán, km 15.5 Carretera Mérida-Xmatkuil, Apartado Postal 4-116 Itzimná Mérida, Yucatán, México*

Email: sgomez@ica.co.cu

In the agricultural research, situations are presented where it is difficult to use the classical linear models of analysis of variance, because the assumptions of independence, equality of variances and linearity are not fulfilled by making measures repeated in time. This paper had as object to review the statistical procedures used to analyze the designs of measures repeated in time, and determine which analytical strategies are more appropriate for each purpose. In this study, three types of traditionally used analyses are described: univariate variance (ANOVA), multivariate variance (MANOVA), and the recent one, the approach of mixed models. At present, it has been agreed that the latter is the most adequate and versatile, because it provides the possibility of examining data with structures of dependence, unbalance, and lack of normality. Besides, it provides a solution to the limitation of the multivariate analysis of variance in respect to the number of individuals and variables. Also, the model of random effects is described, another member of the wide spectrum of the mixed models that is used in numerous studies in the agricultural field. This approach is strengthened by the use of selection criteria of models, due to the estimation of parameters is based on methods of maximum likelihood or restricted maximum likelihood. The Akaike information criterion (AIC) and the Bayesian information criterion (BIC) are described, permitting the optimum selection of competing mixed models.

Key words: *repeated measures, univariate analysis, multivariate analysis, mixed models, information criteria.*

INTRODUCTION

One of the most used research methods is measuring the response variables of interest in the same experimental units at different times. This design of "repeated measures", with its corresponding analysis, is found among the most used currently in the medical, social, and psychological research. If this type of analysis is applied adequately, it increases the validity of the statistical conclusions, because it has higher accuracy in the estimation of the parameters of the model of analysis, it improves the power of the test and reduces the sample size (Fernández and Vallejo 1996).

There are contexts where it is not possible to use classical linear models for the analysis of variance, because the assumptions of independence, normality, equality of variances and linearity demanded for their utilization are unfulfilled by making measures repeated in time in the same experimental units. This situation limits the application of the classical models. Given the specific characteristics of these experiments of repeated measures, it is determined which analytical strategy is more appropriate. Traditionally, the univariate (ANOVA), the multivariate (MANOVA), and the mixed analyses of variance are applied. This latter has great advantages, and it is the one that has been applied in the SAS software, with the name of PROC MIXED (Balzarini *et al*. 2005). The objective of this paper is to review the statistical procedures used in the design and analysis of measures repeated in time, besides describing the most adequate analytical strategies for studies in the agricultural area.

PARTICULARITIES OF THE DESIGNS OF REPEATED MEASURES

Making repeated measures in the same experimental unit implies that it is not possible to randomize the time factor. This, together with the fact that the measures performed in the same individual are close in time, could provoke the measures be correlated between themselves. Thus, the assumption of independence of errors of the classical models of analysis of variance cannot be supported. Besides, the variances of the repeated measures could change frequently in time. These problems lead to deficiencies in the accuracy and the capacity of prediction of the models fitted to the classical assumptions (Gonzalez and López 2002 and Carrero *et al*. 2008).

There are various statistical methods to analyze data of repeated measures, which go from the most basic up to the most complex. According to Littell *et al*. (1998), these methods include the univariate analysis of variance, the multivariate, and the analysis through mixed models.

UNIVARIATE ANALYSIS OF VARIANCE (ANOVA)

For measures repeated in time in the same experimental unit, the assumption of independence of errors is unfulfilled. This can be determined through the analysis of the matrix of correlation of Pearson, permitting to test if a matrix is identity or not, through the test of sphericity of Bartlett (Balzarini *et al*. 2001).

Torres *et al*. (2003) stated that the univariate analysis of variance is used, according to model of split plots, when there is not significant correlation between the measures in time. Also, it can be stated the assumption of equal correlation between any couple of repeated measures in the same individual (model of compounds symmetry) for the matrices of variance-covariance of the observations in the same experimental unit. In this model, the treatment factor is associated with the experimental units in the main plots, and with the time in the subplots. It is necessary to correct the degrees of freedom of the numerator and the denominator in the tests involving the time factor.

These procedures have been applied in the study of the effect of biological preparations with viable yeasts on the rumen microbial population and fermentative indicators in cows fed roughages. Also, they have been used in an experiment with upgraded Holstein bulls, fitted with rumen cannula and housed in individual pens to study the effect of the concentrate form on the rumen microbial population and on fermentative indicators in diets with sugarcane forage (*Saccharum officinarum*) (Marrero *et al*. 2006 and 2007)

MULTIVARIATE ANALYSIS OF VARIANCE (MANOVA)

The observations obtained in the designs of repeated measures are correlated between themselves, and they are, essentially, of multivariate nature. According to Cole and Grizzle (1966), the multivariate procedure is the adequate method to analyze these designs. These authors also noted that the multivariate approach shares all the assumptions with the univariate, except permitting the variance-covariance matrix to have any structure. The observations from the same individual, besides being correlated, have a matrix of variances-covariances between the repeated measures that has a Toeplitz structure, that is, the closest scores have higher correlation.

In the multivariate model, no particular model is assumed for the matrix, but it is based on the estimation of all the possible covariances between the repeated measures. This model without structure should be used when there are sufficient observations for the estimation of the parameters. For its application, the number of repeated measures should be lower or equal to the number of repetitions of the experiment (Torres *et al*. 2003).

Between the univariate and the multivariate analysis, only the multivariate guarantees that the error of Type I is not above the nominal.

The multivariate analysis of repeated measures is a methodology that has been traditionally used in the study of data of repeated measures, from experiments in the agricultural area. An example of their application is the methane production in the rumen, out of the fermentation of the carbohydrates for a mixed microbial population composed of methanogens, participating also other groups of bacteria and protozoa. Also, it has been used in studies on male buffalo calves (*Bubalus bubalis*) of the Bufalipso breed, fitted with rumen cannula, to evaluate the effect of different levels of supplementation on *in situ* rumen DM intake and degradability of star grass (*Cynodon nlemfuensis*) forage (Galindo *et al*. 2009 and López *et al*. 2009)

The statistical theory describes three types of mathematical models: that of fixed effects, the random, and the mixed.

MODEL OF FIXED EFFECTS

The model of fixed effects of analysis of variance is applied to situations where the experimentator subjects the group or material analyzed to one or various factors. Each is affected only by the mean, and remains the "response variable" (or in an equivalent form the term of error) as the only random variable that has a particular distribution. This model is used when the researcher is interested, only, in the levels of the factor present in the experiment, thus, any variation in the scores could be due to the experimental error (Spiegel *et al*. 2007).

The simplest model of fixed effects is the classical parametric model of analysis of variance with only one factor (or one-way). In this particular model, the "response variable" (or in an equivalent form, the term of error) is assumed as normal, with constant variance. Thus, the levels of the only factor are assumed as fixed, because they are those of interest to be analyzed by the researcher, not extending the inferences to a larger cluster of treatments.

Be y_{ij} the random response observed in the unit j of the treatment i from a population of observations under the treatment i, with normal distribution with mean μ_i and variance σ^2 , the model of one-way analysis of variance of fixed effects for y_{ij} is:

$E(y_{ij}) = \mu_i$ where,

E (.) represents the expectation operator

μ_i is the expected response for an observation under

the treatment *i*.

In this model, called model of means, each μ_i is considered as a constant. These constants (fixed values) represent, in certain form, the magnitudes to be estimated. It can be of interest to estimate μ_i and μ_j or $\mu_i - \mu_j$. The constants to be estimated μ_i´s, with i=1,…, a correspond explicitly to the a treatments tested in the experiment. There are treatments that are of interest, and that, thus, are arbitrarily selected. The effect of the treatment *i* is defined as:

$\tau_i = \mu_i - \mu$ where,

μ is the general mean of the response variable, thus the model can be written as follows:

$E(y_{ij}) = \mu + \tau_i$,

A formal justification of why it is called "of fixed effects" is in the effects of treatment τ_i, that in this model are assumed as constants. If e_{ij} represents the value of the deviation or difference between y_{ij} and its expected value, term called error in y_{ij}, the complete model of fixed effects can be expressed as the sum of its expected value and one error (random) (Balzarini *et al*. 2005):

$y_{ij} = \mu_i + e_{ij}$ or equivalent to $y_{ij} = \mu + \tau_i + e_{ij}$.

This latter expression (or "parameterization") of the model of one-way analysis of variance is called model of effects (Littell *et al*. 2006).

MODEL OF RANDOM EFFECTS

The models of random effects are used to describe situations where there are incomparable differences in the experimental material or group. The simplest example is that of estimating the unknown mean of a population composed of different individuals, where these differences are mixed with the errors of the measurement instrument.

This model is used when the researcher is interested in a population of levels of the factor under study, theoretically infinite, out of which only a random sample is present (t levels) (Spiegel *et al*. 2007). In this instance, each level of the factor treatment is assigned randomly on n experimental units. There are n random observations for each of the a levels of the factor of interest.

The simplest model of random effects is that containing only one random factor. If y_{ij} represents the response observed in the unit *j* of the treatment *i*, the model "of effects" for the data is:

$E(y_{ij} \mid a_i) = \mu + a_i$, where,

μ is the general mean of the response variable.

a_i is the effect of the level i of the factor of interest.

$a_i = \mu_i - \mu$, and $E(y_{ij} \mid a_i)$

$E(y_{ij} \mid a_i)$ represents the conditional expected value of yij, given the random amount a_i (random effect of treatment). Despite the previous expression is quite similar to that corresponding to the model of fixed effects, the underlying assumptions are different. This is due to the levels of the factor treatment represent a random sample of the population. Due to the amounts ai are random variables, it is necessary to characterize their distribution of probabilities. Usually, the amounts ai are considered independent and identically distributed, with expectation zero and variance σ_a^2 for all *i* (Balzarini *et al*. 2005).

The model with only one classical random factor (i.e. with normal error) can be written out of the parameterization by "effects":

$y_{ij} = \mu + a_i + e_{ij}$, where,

- e_{ij} is a component of random error assumed as normal with mean 0 and constant variance, associated with the level *i* of only one factor (random)
- a_i is the effect of the level *i* of the only factor, which is distributed normally with mean 0 and constant variance.

Out of this parameterization, the model contains two "components of variance" for the response variable y_{ij}:

- a component due to the part of random effects of the model a_i.
- another due to the part of the error or residual e_i.

It should be noted that in a recently described model "participates" a fixed effect (μ, the global mean). This explains why the models of fixed and random effects can be unified through the theory of "mixed" models (Littell *et al*. 2006), where effects of the two types (fixed and random) are involved.

MIXED MODELS

The mixed statistical models are used more frequently in the medical sciences and less applied in researches of the agricultural field; thus, their application would be useful due to their advantages. These models permit modeling the response of an experimental or observational study as function of factors or covariables, whose effects are considered fixed constants or random variables (Molinero 2003). The fixed and the random effects are present in the mixed models. They are a combination of both, and to decide when a set of effects is dealt with, as fixed or random, it is important to analyze the data within their context, that is, the environment they come from, the way they are collected and, mainly, the inference space (Verde 2000 and Balzarini *et al*. 2005)

Different types of models are considered within the general frame of the mixed models. It is noteworthy to recall that these models are presented as those permitting to model data sets where the observations are not independent (Balzarini *et al*. 2005).

If it is considered that a model of random effects also contains a fixed effect (the global mean), then, the simplest type of mixed model is that of random effects. Other mixed models are those of random coefficients, which combine random effects and coefficients, and those incorporating covariance patterns of "within" individuals or subjects (being the latter those related to the designs of repeated measures). The selection of the mixed model depends on the characteristics of the experiment and on the objective of the analysis.

The linear and non-linear mixed models emerge from incorporating random effects, different to the associates with the term of error. The greatest advantage of them is in the generality in the inference and in the possibility they provide of modeling the correlation between the observations. The estimation of parameters in these models is performed by likelihood methods (Carrero *et al*. 2008).

ANALYSES OF DESIGNS OF REPEATED MEASURES BY MEANS OF MIXED MODELS

The univariate analysis of the variance of the repeated measures can be performed through mixed models (Arnau and Balluerka 2004). In the model of repeated measures, the experimental units are considered a random factor, and the time, as fixed. According to Carrero *et al*. (2008), the methodology of mixed models permit analyzing in a correct and efficient form the data from experiments with repeated measures through the modeling of the structure of covariances, which considers the correlations between repeated measures and the presence of heterogeneous variances. Not considering the correlation between subjects with the utilization of models of fixed effects (in SAS they are executed through ANOVA procedures or by general linear model GLM) or mixed models with very simple structures of covariances, could originate the increase in the rate of error type I (rejection of the null hypothesis when it should be accepted) for the test of fixed effects of the model. Nevertheless, a very complex model would affect the power and efficiency of the test for the fixed effects (Pérez *et al*. 2005 and Vallejo *et al*. 2010)

METHODS OF ESTIMATION OF PARAMETERS IN THE MODELS OF REPEATED MEASURES

For estimating the parameters of a mixed model, a wide range of methods can be applied: moments, unbiased quadratic estimation of minimum variance, maximum likelihood, restricted maximum likelihood, and pseudo-likelihood (Littell *et al*. 2006). However, for the case of the analyses of repeated measures, the efficient estimation of the parameters of covariance is attained through the restricted maximum likelihood (REML = Restricted Maximum Likelihood). The basic antecedent to understand intuitionally the idea behind the REML is the method of maximum likelihood.

The method of maximum likelihood (Maximum Likelihood, ML) is a classical method of estimation of parameters associated with functions of density or probability of random variables. The likelihood associated with a sample of random variables is the function of joined density of these variables for the observed values, considered as a function of the parameters that define it. If the density function of a random variable is known, as it is the case of the classical mixed models that assume a normal distribution of the response variable, it is then possible to find out the estimators of maximum likelihood (EML) of the parameters of the mixed model. These EML are the values of the parameters that make maximum the probability (likelihood) of the data occurrence, that is, those being more compatible with the observed data, assuming that it is correct the postulated mathematical model (Molinero 2003).

The procedures of estimation of ML work well, even in instances where there are incomplete records, different to the method of least squares. If the repeated measures are considered as "nested" to the subject factor, the ML estimation permits analyzing an unequal number of observations per experimental unit and different space in time. Also, it admits incorporating parameters of inter-individual variability in growth, whose estimation considers the different precisions derived from different number of measures in each experimental unit (Oliver *et al*. 2000 and Posada and Rosero 2007).

In general, the ML estimators function very well for large samples. From the technical point of view, they possess asymptotic properties, which make them preferable to those obtained with other methods. Besides, they do not require sets of balanced data to maintain these properties (Galán *et al*. 2003 and León 2004). These estimators are characterized by being:

- Consistent
- Invariant to biunivocal transformations, that is, if is the estimator of maximum likelihood of θ_{MV} , and $g(\theta)$ is a biunivocal function of θ, $g(\theta_{MV})$ is the estimator of maximum likelihood of $g(\theta)$.
- If θ is a sufficient estimator of θ, its estimator

of maximum likelihood θ_{MV}, is function of the sample through θ.

- Asymptotically normal.
- Asymptotically efficient, that is, among all the consistent estimators of a parameter θ, those of maximum likelihood are those of minimum variance.

For samples of moderate or small size, ML seems to produce biased estimations of the parameters, that is, the expected value of the estimator of the parameter is not equal to it because it does not consider the degrees of freedom lost when estimating the mean (León 2004). The restricted estimators of maximum likelihood (REML, Restricted Maximum Likelihood) emerged to prevent this problem. They consist in factoring the complete likelihood in two independent parts. One of them does not contain the mean, assuming that by using this part of the likelihood, information is not lost in respect to the complete likelihood. The restricted likelihood is in correspondence with the likelihood associated with a linear combination of the observations, whose mean is null and fulfills the conditions mentioned (being a factor independent from the other with which the complete likelihood is reproduced and not supposing loss of information compared with the use of the original data). The REML method produces unbiased estimations of the variance in mixed models with random or normal residual error. Thus, it surpasses the classical ML method.

INFORMATION CRITERIA FOR THE SELECTION OF THE MODELS IN REPEATED MEASURES

There are several criteria to determine the goodness of fit of the chosen model during the modeling process in the analysis of repeated measures. For comparing models, the most used criterion is that of the deviation or deviance. The deviance is calculated from the logarithm of the function of likelihood or restricted likelihood, depending on choosing a model with identical structure of covariance or of means. Others are also used, such as the Akaike Information Criterion (AIC), the AIC corrected (AICC) and the Bayesian Information Criterion (BIC), as well as various versions emerging from these criteria. Especially, the AIC and the BIC are implemented in most of the programs that fit mixed models. At greater or smaller extent, all penalize the logarithm of the likelihood function by the number of parameters, most of the time out of the marginal formulation of the model, and choose the model that minimizes their value (Vallejo *et al* 2010).

AKAIKE INFORMATION CRITERION (AIC)

Information Criterion (AIC). This method permits determining the efficiency of the models fitted to a database (Posada and Rosero 2007 and Noguera *et al*. 2008). This alternative approach can be applied to nested and non-nested models, and does not rely on values of P (probability) or on the concept of statistical significance:

AIC=-2(ln likelihood - nº parameters)

The selection criterion is to choose models with lower values of the AIC. The model that accounted the best for the data with the minimum number of parameters is the one having the lowest value of AIC (Molinero 2003 and Balzarini *et al*. 2005).

The logic behind this method is not that of the hypothesis tests. Therefore, a null hypothesis should not be stated, nor should be calculated a value of P, and it is not necessary to decide about the trend of the value of P to determine its statistical significance. Besides, the method permits determining which model is the most likeable to the correct and quantifying its resemblance.

The AIC value can be negative or positive, depending on the units in which the data are expressed, and it cannot be interpreted as an individual value. This criterion has great importance when comparing the models; thus, work is conducted with the differences between the AIC (Schermelleh *et al*. 2003)

Define A as the simplest model, and B as the most complex (that is with larger number of parameters). The difference of the AIC between A and B is defined as:

$$\begin{aligned}\Delta AIC &= AIC_B - AIC_A \\ &= N \cdot \left[ln\left(\frac{SC_B}{N}\right) - ln\left(\frac{SC_A}{N}\right)\right] + 2(K_B - K_A) \\ &= N \cdot ln\left(\frac{SC_B}{SC_A}\right) + 2(K_B - K_A)\end{aligned}$$

As in the F test, this analysis establishes a compromise between the goodness of fit characterized by the sum of squares with the change in the number of parameters to be fitted (Posada and Rosero 2007). Because the model A will have almost always the worst fit, the sum of squares of A will be higher than that of B. As the logarithm of a fraction is always negative, the first term of the equation will be negative. As the model B has more parameters, K_B will be higher than K_A, which makes the last term is positive. If the final outcome is negative, it means that the difference in the sum of squares is higher than expected, out of the difference in the number of parameters, hence, the model B is the best. If the difference in the AIC is

positive, thus, the change in the sum of squares is not that large as expected with the change in the number of parameters, thereby showing that the data come from the model A.

The previous equation helps to understand the function of the AIC, out of the balance of the changes in the goodness of fit versus the difference in the number of parameters. In fact, what it is done is to calculate the two individual AIC, and the model with the lowest AIC is chosen, being the most correct (Posada and Rosero 2007).

BAYESIAN INFORMATION CRITERION OF SCHWARZ

The Bayesian statistics emerges, in fact, from the famous theorem of Bayes, which in essence, permits, in case of knowing the probability of an event to happen, to modify its value when having new information (Molinero 2002).

The information criterion of Schwarz is called Bayesian by being based on arguments of the so-called Bayesian statistics. The Bayesian methods are an alternative to the traditional statistics, which is based on the hypothesis contrast. These methods are different in that they incorporate information external to the study. With this information and with the observed data, a distribution of probability is estimated for the magnitude effect under study (Díaz and Batanero 2008).

The formula for the Bayesian information criterion (BIC) is similar to the Akaike criterion, as well as its interpretation:

BIC=G - gl . ln N, where,

G is the quotient of likelihood

gl are the degrees of freedom

N is the size of the sample

The criterion to choose the best model is the same as that of Akaike: that having the lowest value of BIC (Calegario *et al*. 2005 and Carrero *et al*. 2008)

DISCUSSION

The particular characteristics of the researches with repeated measures in the agricultural field make difficult the utilization of the classical models of analysis of variance because the data do not fulfill the basic assumptions. This work describes the methods used to make adequately the analyses of repeated measures. The enumeration of the properties of these procedures has suggested the use of the mixed models to model the experimental data. The advantages of the mixed models are various: permit analyzing data with structures of dependence and unbalances in the data. Besides, it is possible to analyze non-normal data with repeated measures through the so-called generalized mixed models (Littell *et al*. 2006). Therefore, the concern about the fulfillment of the traditional assumptions of normality and equal variances goes into a second order.

The advantages in the use of the mixed models described in this review to be applied in the analysis of repeated measures can be seen in practice in experimental studies in the agricultural branch. For instance, an experiment was developed with mutant strains of the cellulolytic fungi *Trichoderma viride* M5 and MCX1371 and also it was studied a herd of crossbred animals of the Holstein, Brown Swiss and Brahman breeds, from the commercial farm La Duquesa, located in the "Río Grande" sector, Pan American municipality, Táchira state, in Venezuela. This farm is devoted to dual-purpose cattle rearing, but with a trend to milk yield.

The analysis of the data from the latter was performed initially by the methodology of Torres *et al*. (2003), although it was necessary to prove the fulfillment of the basic hypotheses, specifically the fact that the number of repeated observations was smaller than the number of repetitions of the experiment (Balzarini *et al*. 2005). Sometimes, this hampered the analysis, because it was necessary to reduce the number of schedules and it was not possible to study completely the sequence of the experiment, which affected the results. Thus, it was used the alternative of analysis with mixed models. Even when results similar to the first analysis were obtained, there was greater flexibility in the data processing with the mixed model.

CONCLUSION

For the analysis of the repeated measures in the same experimental unit at different times, adequate statistical procedures are necessary, being different to the traditional, in a way that the validity of the statistical conclusion is accentuated. Thus, the procedures used in the designs of repeated measures in the agricultural field are analyzed: the univariate analysis (ANOVA), the multivariate variance (MANOVA) and the mixed models.

It is concluded that the latter are the most recommendable due to their advantages: 1) solving the inconvenient in the lack of fulfillment of the basic suppositions of the classical methods of analysis of variance; 2) solving the limitations of the univariate analysis of repeated measures in respect to very simple structures of covariance; 3) solving the limitations

of the multivariate analysis of repeated measures in respect to the requirement of large sample sizes, and 4) providing information criteria needed for selecting the best model.

REFERENCES

Arnau, J. & Balluerka, N. 2004. Análisis de datos longitudinales y de curvas de crecimiento. Enfoque clásico y propuestas actuales. Psicothema. 16:156

Balzarini, M., Casanoves, F., Di Rienzo, J.A., González, L.A. & Robledo, C.W. 2001. Infostat. Versión 1. Córdova. Argentina. Pp. 156-159

Balzarini, M., Macchiavelli, R. & Casanoves, F. 2005. Aplicaciones de modelos mixtos en agricultura y forestería. Curso Internacional de Aplicaciones de Modelos Mixtos en Agricultura y Forestería. CATIE. Turrialba, Costa Rica. p.189

Calegario, N., Maestri, R., Leal, C. & Daniels, R. 2005. Estimativa do crescimento de povoamentos de *Eucalyptus baseada* na teoria dos modelos não lineares em multiníveis de efeito misto. Ciência Florestal 15:285

Carrero, O., Jerez, M., Macchiavelli, R., Orlandoni, G. & Stock, J. 2008. Ajuste de curvas de índice de sitio mediante modelos mixtos para plantaciones de *Eucalyptus urophylla* en Venezuela.Interciencia 33:4

Cole, V.W.L. & Grizzle, J.E. 1966. Applications of multivariate analysis of variance to repeated measures experiments. Biometrics 41:505

Díaz, C. & Batanero, C. 2008. ¿Cómo puede el método bayesiano contribuir a la investigación en Psicología y Educación? Ed. Universidad de Granada. Paradigma 27:35

Fernández, P. & Vallejo, G.1996. Diseño de medidas repetidas con dependencia serial en el error bajo la violación de la asunción de homogeneidad. Universidad de Oviedo. Anales de Psicología 12:87

Galán, J., Jiménez, E. & Cervantes, C. 2003. Estimación por máxima verosimilitud restringida de componentes de varianza y covarianza de múltiples características bajo los diseños I y II de Carolina del Norte. Rev. Fitotecnia Mexicana 26:53

Galindo, J., González, N., Delgado, D., Sosa, A., González, R., Torres, V., Aldana, A.I., Moreira, O., Sarduy, L., Noda, A.C. & Cairo, J. 2009. Effect of coconut oil on the population of methanogenic bacteria and its relation to other microbial groups in the rumen under in vitro conditions. Cuban J. Agric. Sci. 43:129

González, L.M. & López, L. A. 2002. Medidas repetidas con datos faltantes: estimación de parámetros vía análisis de covarianza. Rev. Colombiana de Estadística 25:127

León, E. 2004. Métodos de estimación de componentes de varianza en poblaciones. Una reseña histórica. Instituto de Investigaciones Porcinas. Rev. Computarizada de Producción Porcina 11:1

Littell, R.C., Henry, P.R. & Ammerman, C.B. 1998. Statistical Analysis of Repeated Measures Data Using SAS Procedures. Dep. Statistics Animal Sci. University of Florida, Gainesville. Pp. 42-46

Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D., & Schabenberger, O. 2006. SAS for Mixed Models. Second Edition. Institute Inc., Cary, EE.UU. p. 814

López, J.R., Elías, A., Delgado, D., González, R., & Sarduy, L. 2009. Effect of the supplementation with concentrates on the *in situ* rumen degradability of forage from star grass (*Cynodon nlemfuensis*) in buffalo calves (*Bubalus bubalis*). Cuban J. Agric. Sci. 43:151

Marrero, Y., Galindo, J., Elías, A., Moreira, O. & Cueto, M. 2006. Effect of biological preparations with viable yeasts on the microbial population in rumen and fermentative indictaors in cows fed fibrous diets. Cuban J. Agric. Sci. 40:321

Marrero, Y., Rodríguez, D., Galindo, J., Aldama, AI., Moreira, O. & Noda, A. 2007. Rumen microbial population and fermentative indicators in cattle fed sugarcane (*Saccharum officinarum* L.) and concentrate. Cuban J. Agric. Sci. 41:141

Molinero, L. 2002. El método bayesiano en la investigación médica. Ed. Liga española para la lucha contra la hipertensión arterial. p. 3-10

Molinero, L.M. 2003. ¿Qué es el método de estimación de máxima verosimilitud y cómo se interpreta? Liga Española para la lucha contra la Hipertensión Arterial. p. 1 Available: <www.seh_lelha.org/statl.htm> [Consulted: Enero 2011]

Noguera, R.R., Pereira L.R. & Solarte, C.E. 2008. Comparación de modelos no lineales para describir curvas de crecimiento en cuyes (*Cavia porcellus*) desde el nacimiento hasta la edad de sacrificio. Universidad de Antioquia-Universidad de Nariño. Liv. Res. Rural Dev. 20:3

Oliver, J.C., Rosel, J. & Murray, L. 2000. Análisis de medidas repetidas mediante métodos de máxima verosimilitud. Universidad Jaume I – Universidad del Estado de Nuevo México. Psicothema 12:403

Pérez, E.R., Gutiérrez, E., Herrera-Camacho, J. & Segura, J.C. 2005. Elección del mejor modelo para el análisis de experimentos con medidas repetidas en tiempo. Hormonas en cerdas durante la lactancia. Instituto de Investigaciones Agropecuarias y Forestales. Universidad Michoacana de San Nicolás de Hidalgo - Universidad Autónoma de Yucatán. p. 3

Posada, S.L. & Rosero, R. 2007. Comparación de modelos matemáticos: una aplicación en la evaluación de alimentos para animales. Rev. Col. Cienc. Pec. 20:141

Schermelleh, K., Moosbrugger, H. & Müller, H. 2003. Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. Methods of Psychological Res. 8:23

Spiegel, M.R., Schiller, J. & Srinivasan, R.A. 2007. Análisis de la varianza. Probabilidad y estadística. Ed. McGraw-Hill. Segunda Edición. México, D.F. p. 335-371

Torres, V., Navarro, J.R. & Pérez, T. 2003. Statistical models for processing experiments with repeated measurements in the same experimental unit. Cuban J. Agric. Sci. 37:225

Vallejo, G., Arnau, J., Bono, R., Fernández, P. & Tuero, E. 2010. Selección de modelos anidados para datos longitudinales usando criterios de información y la estrategia de ajuste condicional. Psicothema 22:323.

Verde, O. 2000. Modelos mixtos, máxima verosimilitud y modelo animal. Ed. Facultad de Ciencias Veterinarias. Universidad Central de Venezuela. Rev. Zootec. Trop. 18:3

Received: December 10, 2011