STATISTICAL TECHNIQUES APPLIED IN THREE JOURNALS OF AGRICULTURAL SCIENCES WITH A FOCUS ON ANIMAL SCIENCE

Marcos BUSANELLO[1]
Thiago Sergio de ANDRADE[1]
Carolina Naves AROEIRA[1]
Carlos Tadeu dos Santos DIAS[2]

▪ ABSTRACT: There are few previous studies that investigate the most used statistical techniques in animal science. Due to the large number of tools and methods available for statistical analysis, it is important to identify the most applied ones for this area of research. Therefore, we aimed to identify the use of different statistical techniques (designs, software and analysis) used in two Brazilian journals (Ciência Rural and Revista Brasileira de Zootecnia) and one international journal (Journal of Animal Science). In order to do this, scientific articles published during the years 2011 to 2015 were selected to form a database. Our article discusses the use of designs, software and analyses most commonly used in the journals studied. To study this, we used descriptive statistics and multivariate approaches. Completely randomized and randomized blocks design were the principal designs used in animal science. The SAS® software was the principal software used. Finally, analysis of variance was the principal statistical method, followed by regression analysis. There were no differences between the journals over time regarding the use of statistical analyses. The results highlight the importance of hypothesis testing within animal science.

▪ KEYWORDS: Animal research; experimental design; statistical methods; statistical software.

1 Introduction

Scientific experimentation consists in planning, implementing, collecting sample material, performing statistical analysis, interpreting results, and drawing conclusions. Considering statistical analysis, which allows one to describe and interpret results, besides facilitating the understanding of scientific research. The use of statistical analysis is extremely important because there are non-controlled factors related to biological systems, and its standardization is relevant to the correct interpretation of data (FESTING and ALTMAN, 2002).

Statistical analysis allows verification if there are differences between the factors studied, as well as the interpretation of such significance. However, the researcher must know how to differentiate between statistical significance results and biological

[1] Universidade de São Paulo – USP, Escola Superior de Agricultura "Luiz de Queiroz", Departamento de Zootecnia, CEP: 13418-900, Piracicaba, São Paulo, SP, Brasil. E-mail: *marcosbusanello@hotmail.com; tsazoo@hotmail.com; carolinaaroeira@usp.br*
[2] Universidade de São Paulo – USP, Escola Superior de Agricultura "Luiz de Queiroz", Departamento de Ciências Exatas, CEP: 13418-900, Piracicaba, São Paulo, SP, Brasil. E-mail: *ctsdias@gmail.com*

significance results, as well as considering the power of statistical tests (THOMAS and JUANES, 1996; BAGULEY, 2004). Furthermore, the development of computational tools used in many countries has facilitated the implementation of statistical analyses and expanded the diversity of available methods. Currently, there are numerous tests and software that carry out statistical analysis faster and more practically.

In this way, it is essential to have tools for the correct evaluation and interpretation of results (ROBINSON *et al.*, 2006). From this, the choice of which software to use is very important. According to Cavalieri (2015), researchers with experience in using statistical software, even related to a completely different mathematical problem, are strongly convinced that the tool used in the past is still the right choice, even for a new problem. The same author cites that such a decision is due, mainly, to the trust that the researcher has in his or her knowledge of the software, instead of taking into consideration what the software is really able to do.

In addition to selecting the correct software, it is important to evaluate the possibilities of statistical analysis to be used to obtain the correct results of the hypotheses tested. Thus, assumptions can be made when researchers compare the methods used. Furthermore, Udén *et al.* (2012) stated that the decision which rejects or accepts a study for publication may be based, first, on a clear description of the samples used and also the methods that were used in the statistical analysis.

Thus, there is concern in knowing the statistical methods most used in different areas of knowledge. To this effect, Cantuarias–Avilés and Dias (2008) studied the statistical methods most used in fruit-growing areas in a Brazilian journal and an international journal. They found that in the Brazilian journal, the most used design was completely randomized and in the international journal, most of the scientific papers did not use any design. In addition, the analysis used most, in both journals, was the analysis of variance and mean comparison (ANOVA).

On the other hand, López (2000) studied scientific papers on smoking in four Spanish medical journals and found that the contingency table was the most used statistical method (37% of the papers) followed by descriptive statistics (18%). Moreover, the author reports that it is necessary to evaluate the quality and relevance of the statistical methods used for research on smoking and other medical fields. This may be considered for other areas of knowledge, for example, agricultural sciences.

One of the major areas of research is agricultural sciences; this area consists of multidisciplinary studies involving: agronomy, agroecology, forestry, fishing engineering, veterinary medicine, animal science, agricultural engineering, food science, and aquaculture engineering. It is an area, that aims to seek technical improvement, increased productivity, and improvements in the management and conservation of natural resources. Nowadays, agricultural sciences include some of the most promising fields in technological research, especially genetic engineering and biofuel.

Another important factor is that the growing demands for food and the need to preserve and reuse natural resources placed this field among the most important in the context of current scientific research. Within agricultural sciences, animal science aims to develop and enhance the potential of domestic animals to advance the areas of animal husbandry, food production, nutrition, agribusiness, behaviour, welfare, and biotechnology. Animal science may include the fields of genetics, microbiology, animal behaviour, nutrition, physiology and breeding, soil, agricultural economics, marketing, and legal aspects related to the environment.

Thus, there are few studies about the statistical methods more commonly used in agricultural sciences, for example, Cardellino and Siewerdt (1992), Santos *et al*., (1998), Lúcio *et al.*, (2003), and some others. Consequently, in the sub-areas, such as animal science, there is a lack of scientific papers that demonstrate which are the most used statistical techniques.

Therefore, the aim of this study is to identify the use of different statistical techniques (designs, software, and analyses) used in two Brazilian and one international Agricultural Sciences journals, with a focus on Animal Science. In addition, we seek to find similarity between analyses used within each journal, characterize such analyses and examine the differences between the journals regarding the use of statistical analyses along the years.

2 Material and methods

2.1 Data base

A database was obtained from analyses of original scientific papers from 2011 to 2015 of three important journals in agricultural sciences focused on animal science. Two of these journals are Brazilian journals, Ciência Rural (CR) and Revista Brasileira de Zootecnia (RBZ), and the other was an international journal, Journal of Animal Science (JAS). Both CR and RBZ were chosen because they are the principal targets of studies with animal science in Brazil, and because their scientific papers have open access through Scientific Electronic Library Online (SCIELO). On the other hand, JAS was chosen as an international journal with a specific focus in animal science, but not so specific compared to others, for example, Journal of Dairy Science and Animal Feed Science and Technology, and it is also one of the principal targets of study within animal science in the world.

The CR is a scientific journal started in 1971 and currently publishes an average of 360 scientific papers distributed in 12 volumes per year. Ciência Rural publishes scientific papers, notes and reviews related to agricultural sciences, which is a wide area of knowledge. Revista Brasileira de Zootecnia is a monthly published which has been running by the Brazilian Society of Animal Science since 1972. It publishes original scientific papers in the areas of aquaculture, forage, breeding, genetics and reproduction, ruminants, non-ruminants, animal production systems, and agribusiness. However, RBZ has a more limited scope compared to CR. Journal of Animal Science publishes over 600 scientific papers and scientific techniques notes annually, in addition to being one of the most traditional journals in animal science around the world. Studies published in JAS are focused on animal science in terms of animal production and the fundamental aspects of genetics, nutrition, physiology and animal products.

This study was done using a total of 849 scientific papers, with 396 papers from JAS, 293 from CR, and 160 from RBZ. The selection of scientific papers was done randomly, where one volume was selected for each. After that, two serial numbers within each volume were randomly selected and all the scientific papers within these serial numbers were evaluated. This procedure was adopted after questioning a statistician with prior knowledge in this type of research who advised such an approach. The volumes and number series used from JAS were: v.89, n.3, n.6, 2011; v.90, n.1, n.2, 2012; v.91, n.8, n.12, 2013; v.92, n.7, n.10, 2014 and v.93, n.1, n.2, 2015. The volumes and number series used from CR were:

v.41, n.1, n.2, 2011; v.42, n.4, n.8, 2012; v.43, n.3, n.8, 2013; v.44, n.3, n.11, 2014 and v.45, n.1, n.4, 2015.

In RBZ, there were a small number of scientific papers published in the last two years (2014 and 2015). For this reason, we opted to use three volumes from 2014 and four volumes from 2015, ensuring an accurate representation of the data in this journal. The volumes and number series used from RBZ were: v.40, n.5, n.10, 2011; v.41, n.10, n.11, 2012; v.42, n.7, n.10, 2013; v.43, n.3, n.8, n.12, 2014 and v.44, n.1, n.2, n.3, n.9, 2015.

Data were collected from the materials and methods, in addition to results and discussions, and then compiled into Microsoft Excel® 2010 spreadsheet software. The main information that was sought was statistical analyses, design, and software. Statistical techniques of analyses that were used more than once in the same scientific paper were disregarded, and only software used for statistical analyses were considered (in some studies more than one software was found).

2.2 Statistical analysis

From the database, descriptive statistics (averages and frequencies) of experimental designs, software and statistical analyses were calculated. With the calculated values, cluster analysis, principal components analysis (PCA), multivariate analysis of variance (MANOVA), and profile analysis (PA) were performed. Cluster analysis aimed to find similarity between the analyses within each one of the journals, as well as the similarity between the journals with respect to each analysis specifically. The distance of 0.75 and 0.40 were used to form the groups of analysis within each journal and between journals within each analysis, respectively. On the other hand, the PCA aimed to characterize the analyses in the three journals along the period studied (2011 to 2015).

In addition, MANOVA was performed aiming to find differences between the journals in relation to the use of statistical analyses in general, using the combined information of the variables involved as an advantage. According to Field and Miles (2010), traditional assumptions for MANOVA are tested using analysis of variance (ANOVA) for each variable. However, the same authors reported that it is a mistake do not consider the set of variables as a whole and, for that reason, the PROC DISCRIM of the SAS® should be used, for example, which uses discriminant analysis to test these assumptions. This procedure assumes the relationship between the variables and the biggest advantage of this approach is to reduce and explain the dependent variables in terms of a set of underlying dimensions. Therefore, assumptions were tested using the PROC DISCRIM.

In testing such assumptions, it was observed that some variables were not being classified properly in their journals, because there were observations with the same value for some of the analyses. However, there was a presence of outliers and heterogeneity of variances for some variables. Consequently, a $\log_{10}$ transformation was required to meet all the assumptions. Also, it was necessary to add a constant (constant = 0.5) to all values because there were values equal to zero and such values would not be transformed without the addition of the constant.

At last, profile analysis with a multivariate approach for repeated measures was used to test the hypothesis of parallelism, coincidence and horizontality of average journal profiles, as well as the time effect within each journal and the journal effect within each time period (EYDURAN *et al.*, 2008). Thus, the averages were calculated from the sum of percentage values for all analyses in each journal. This resulted in 15 average values, five for each journal and one for each year, representing the average profile for each journal in the period. All the statistical analyses were performed using the procedures: CLUSTER, PRINCOMP and GLM with MANOVA and REPEATED options, using the statistical software SAS® (version 9.3, 2012). The significance level considered was 5% ($p < 0.05$).

3 Results and discussion

3.1 Experimental design

According to Udén *et al.* (2012), experimental design, statistical analysis and experimental objectives must all be consistent and clearly described. For Robinson *et al.* (2006) description of the statistical model is one of the principal points for the acceptance or rejection of a scientific paper, and the statistical model should be consistent with the experimental design described in materials and methods. Therefore, it is necessary to consider the type of experimental design to perform a correct data analysis (HU and BAO, 2012). The results of designs used in scientific papers published in the journals evaluated are presented in the Table 1.

Table 1 - Percentage of experimental designs used according to the number of scientific papers published for JAS, CR and RBZ

Designs	Journal							
	JAS		CR		RBZ		Total	
	n[1]	%	n	%	n	%	n	%
CRD[2]	70	17.7	83	28.3	85	53.1	238	27.9
RBD[3]	84	21.2	49	16.7	26	16.2	159	18.6
LSD[4]	22	5.6	0	0.0	18	11.2	40	4.7
Others[5]	12	3.0	3	1.0	3	1.9	18	2.1
Without design	211	53.3	158	53.9	29	18.1	398	46.7

[1]Total of scientific papers; [2]Completely randomized design; [3]Randomized blocks design; [4]Latin square design; [5]Crossover (10), Split-plot (5), Lattice design (1), Youden square design (1) and Design of reversal (1). JAS – Journal of Animal Science; CR – Ciência Rural; RBZ – Revista Brasileira de Zootecnia.

Considering the use of designs in each journal, it was observed that the CRD was used in 53% of scientific papers published in RBZ, followed by RBD (16%) (Table 1). This is justified by the presence of a larger number of experiments involving non-ruminant species (swine and poultry, especially) that use these experimental designs. Other designs were used less and 18% of published studies did not use any design.

Ciência Rural, as well as RBZ, used the CRD (28%) and RBD (17%) as the principal experimental designs (Table 1). However, almost 54% of published studies did not use or present any design. This may have occurred because the CR is a more comprehensive journal and accept basic research compared to JAS and RBZ. CR includes more specific areas such as agribusiness, rural extension, veterinary medicine, and others. For example, considering reports studies of clinical cases published in veterinary medicine, the use of statistical designs is less common.

For JAS, the RBD was detected (21%) as the principal experimental design used, followed by CRD (18%) (Table 1). These designs are often used in studies with swine, sheep, and beef cattle. However, as well as CR, almost 53% of the published studies did not show any design, possibly due to the large part of genetic studies published in this journal, which do not often use experimental designs.

In general, less use of designs such as LSD and others was detected. The LSD represented only 5% in JAS and 1% in RBZ and was not observed in the scientific papers of CR. This may be due to the particularities of this design, which is used more frequently in experiments involving milk production, ruminal degradability and digestibility in ruminants. Despite LSD to provide better local control than CRD and RBD, their use has been lower, probably due to the fact that there is a need for the number of repetitions to be equal to the number of treatments or progenies, which limits their use in some studies (RESENDE, 2007). In general, CRD was the experimental design most used (28%), followed by RBD (19%).

3.2 Statistical software

Considering the software used, SAS® was found to have 46% of the scientific papers (Table 2). The percentage of scientific papers published with SAS® in JAS was 65%, in CR was 15%, and in RBZ was 62%. This may be explained due to the recent and relevant versions of SAS®, where PROC MIXED is included. This procedure makes it possible to add random effects in the statistical model, allowing modeling the data covariance structure, which is important for data analysis of repeated measures over time (LITTELL *et al.*, 1998). The term "repeated measures over time" refers to data with multiple observations on the same experimental sample, where, normally, these various observations are made over time. Generally, observations on the same unit are correlated (LITTELL *et al.*, 2000).

The R® and SPSS® software also had significant use in the journals, but were most cited and used in JAS (4% both the software). In the other hand, CR and RBZ used statistical software developed in Brazil, for example, SAEG® (3% and 8%, respectively) and SISVAR® (FERREIRA, 2011) (7% and 5%, respectively), while these software were not used in JAS. In contrast, other software, such as GenStat® (3%) and ASReml® (4%), were important to the studies published in JAS. Those software were not used in scientific papers published in the Brazilian journals. Robinson *et al.* (2006) report that all these software have strengths and weaknesses and anyone can provide the most appropriate statistical models for specific circumstances.

Moreover, in CR 54% of the scientific papers did not present the software used. CR do not requires the citation of software used for statistical analysis, but its focus is on the description of the statistical procedure (the same occurs to JAS). In addition, the study of case reports frequently does not need statistical analysis and are more common in CR. On the other hand, RBZ requires such information on the material and methods and 15% of its scientific papers still do not have citation of software. It is possible that some RBZ's reviewers and authors do not know the instructions for authors well enough.

Table 2 - Percentage of the principal software programs used in scientific papers published for JAS, CR and RBZ

Software	Journal							
	JAS		CR		RBZ		Total	
	n[1]	%	n	%	n	%	n	%
SAS®	263	64.9	45	14.9	85	62.0	393	45.7
SAEG®	0	0.0	9	3.0	11	8.0	20	2.3
GenStat®	11	2.7	0	0.0	0	0.0	11	1.3
SISVAR®	0	0.0	20	6.6	7	5.1	27	3.1
ASReml®	16	3.9	0	0.0	0	0.0	16	1.9
SPSS®	11	4.2	3	1.0	3	2.2	17	2.0
R®	17	4.2	4	1.3	1	0.7	22	2.6
Others[2]	28	17.3	61	19.3	12	6.6	101	11.7
Without Software	70	2.7	162	53.8	21	15.3	253	29.4

[1]total of scientific papers used; [2]Assistat®, Bionumerics®, BioStat®, Canoco®, Cervus®, DMV®, EcoWeight®, Genes®, GenSel®, Graph Pad Prism®, InStat®, Juse-Stat Works®, Microsoft Excel®, Minitab®, MTDFREML®, MULTIV®, PASW Statistics®, PD Quest®, REMLF90®, Sanest®, SASM Agri®, Sigma Stat®, Sigma Plot®, SYSTAT®, SOC-NTIA®, Stata®, StatGraphics®, Centaurion XV®, Statistic®, Statistica®, Stat Soft®, TFPGA®, Unscrambler®, VCE®, WinStat®. JAS – Journal of Animal Science; CR – Ciência Rural; RBZ – Revista Brasileira de Zootecnia.

3.3 Statistical analysis

The statistical analysis most frequently used in all consulted journals was the analysis of variance (ANOVA) and mean comparison (Table 3), which occurred in 77% of scientific papers. Other important statistical analyses were regression (REG) and correlation (CORR), present in 26% and 17% respectively. The other analysis totaled 13%. Such behavior was also observed in the journals individually, except for JAS, where there was a slight superiority of CORR compared to REG. Regarding descriptive statistical analysis (DS), there was a higher incidence in CR with 16% of the scientific papers, followed by JAS and RBZ with 8% and 4% respectively (Table 3). This difference can be caused by the scope of the journals.

The REG analysis had higher incidence in the RBZ (37%) compared to CR (23%) and JAS (24%). Considering that JAS includes scientific papers on genetic studies and CR includes a broader area, it was expected that the regression analysis for RBZ would present greater proportions compared with the other journals. Thus, RBZ published more studies using quantitative levels as treatments than JAS and CR. The use of quantitative levels has been reported by St-Pierre (2007), who noted that in animal science research , especially relating to the animal nutrition area, some changes has been observed in recent years, such as a notable increase in publications with quantitative measures, mainly in regard to diets.

Correlation analysis was most used in JAS (25%) compared with CR (12%) and RBZ (7%) (Table 3). A possible explanation is that CR and RBZ published studies that comprise the area of genetics less frequently, where it is common to use CORR as a statistical technique. Concerning the less used analyses, multivariate analyses represented 1%, 1% and 0.1% in JAS, CR and RBZ, respectively, being equivalent to 2% of all scientific papers (Table 3). Therefore, the researchers chose to use the resources offered by such statistical analyses in a smaller proportions.

Table 3 - Percentage of statistical analysis according to the number of scientific papers published for JAS, CR and RBZ

Statistical Analyses	Journal							
	JAS		CR		RBZ		Total	
	n[1]	%	N	%	n	%	n	%
ANOVA[2]	315	79.5	200	68.3	138	86.3	653	76.9
REG[3]	95	24.0	67	22.9	59	36.9	221	26.0
CORR[4]	97	24.5	34	11.6	11	6.9	142	16.7
DS[5]	32	8.0	46	15.7	6	3.7	84	9.9
Multivariate analyses[6]	4	1.0	5	1.0	8	0.1	17	2.0
Other analyses[7]	8	2.0	0	0.0	0	0.0	8	0.9
Without analysis	25	6.3	33	11.3	5	3.1	63	7.4

[1]total of scientific papers used; [2]Analysis of variance and mean comparison; [3]Regression analysis; [4]Correlation analysis; [5]Descriptive statistics; [6]Cluster analysis (n = 3), MANOVA (n = 1), Canonical correlation analysis (n = 3), Correspondence analysis (n = 2) and Path analysis (n = 2), not informed (n = 6); [7] Genome association analysis (n = 5), Sampling (n = 3). JAS – Journal of Animal Science; CR – Ciência Rural; RBZ – Revista Brasileira de Zootecnia.

Moreover, it was possible to identify studies that used other statistical analyses (e.g., genome association analysis, sampling). Only in JAS was the use of these analyses observed. Some scientific papers did not use statistical analysis, with a higher occurrence in the CR (11%), followed by JAS (6%) and the RBZ (3%), corresponding to a total of 7% from the selected papers (Table 3).

3.4 Similarity between analyses

The use of cluster analysis allowed studying similarity between journals with respect to each analysis. By using cluster analysis with ANOVA (1) (Figure 1) two groups were formed: one by CR and another by JAS and RBZ, confirming the information from table 3 where CR used less ANOVA compared with the other two journals. Moreover, the majority of scientific papers used ANOVA (Table 3) in all journals - JAS (80%), CR (68%) and RBZ (86%). However, CR has less focused a scope than RBZ and JAS, publishing scientific papers in various areas. Thus, the presence of hypothesis testing in studies published by CR is less common than in the other two journals. For this reason, CR has less scientific papers with ANOVA than the others. Moreover, CR has a higher percentage of scientific papers with DS than RBZ and JAS, and this also could confirm the above mentioned comment.

Cluster analysis for REG (2) (Figure 1) resulted in the formation of two groups: one group formed by the RBZ and the other by CR and JAS. It was observed that regression analysis presented greater proportion in RBZ (37%) when compared to CR (23%) and JAS (24%) (Table 3). This demonstrates that RBZ published more studies with quantitative levels of factors, mainly with animal nutrition.

Considering the CORR (3) (Figure 1), it was observed that two groups were formed; the Brazilian journals were grouped together and the JAS remained in another group. This is explained by the fact that the publication of scientific papers with genetic studies are more common in JAS, and also that this type of study used CORR frequently as a statistical technique for the evaluation of the data. The proportion of CORR was higher for JAS (25%) compared with CR (12%) and RBZ (7%) (Table 3).

On the other hand, CR published studies in several areas of agricultural science and, therefore, has higher percentages of studies using DS as the principal analysis (16%), being twice the amount of JAS (8%) and four times the amount of RBZ (4%) (Table 3). So, it was confirmed by cluster analysis, in which DS analysis (4) (Figure 1) formed two groups: one formed by CR and the other by the RBZ and JAS.

In the case of analyses less used, the cluster analysis formed two groups for the multivariate analysis (5) (Figure 1). One group was formed by RBZ and another by CR and JAS, and this analysis represented 1%, 1% and 0.1% in JAS, CR and RBZ, respectively (Table 3). It can be said that the researchers chose to use such statistical analyses less frequently in animal science.

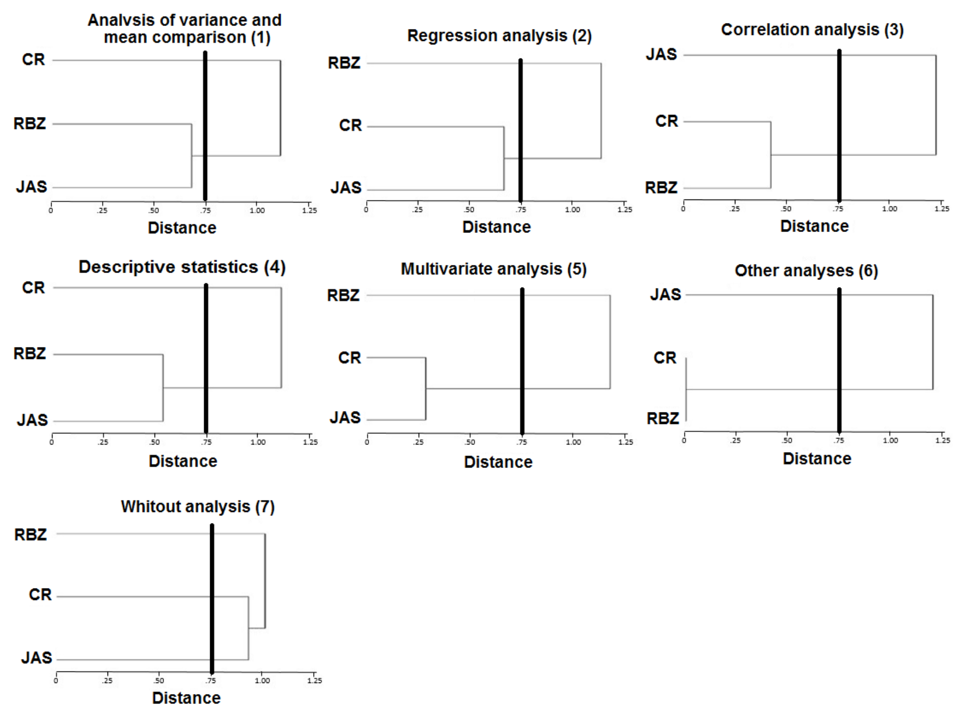

Figure 1 - Cluster analysis among journals for each statistical analysis. JAS – Journal of Animal Science; CR – Ciência Rural; RBZ – Revista Brasileira de Zootecnia.

It was also possible to identify studies that made the choice to use other statistical analyses - genome association and sampling analysis - (6) (Figure 1). Cluster analysis had formed two groups: one group was formed by JAS and the other by RBZ and CR. In this case, only the JAS used another analysis. Moreover, articles that did not used statistical analysis were also recorded (7) having greater occurrence in CR (11%), followed by JAS (6%) and RBZ (3%) (Table 3). Thus, cluster analysis formed a group with CR, another by JAS, and the other by RBZ. Therefore, the similarity between the journals is variable between the different analyses, where different combinations between journals within each statistical analysis were formed.

In addition to comparison of the same analysis between journals, similarities in the use of different analyses within each journal were also examined. Thus, cluster analysis to JAS (1) (Figure 2) formed three groups of analyses: one group formed by ANOVA; the second group formed by analysis REG and CORR; and the third group formed by DS, other analyses, multivariate analysis, and scientific papers without analysis.

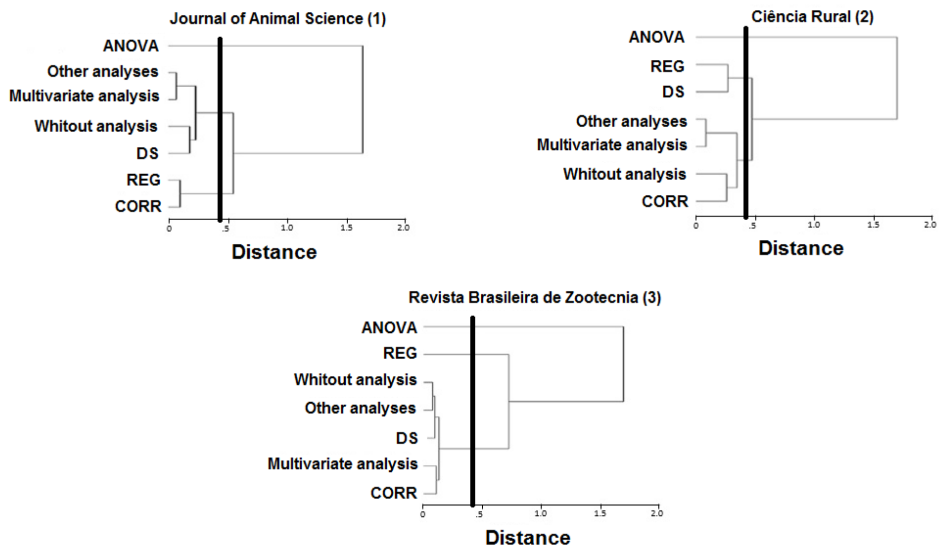

Figure 2 - Cluster analysis between statistical analyses used for each journal. ANOVA - Analysis of variance and mean comparison; REG - Regression analysis; CORR - Correlation analysis; DS - Descriptive and graphic statistics.

On the other hand, cluster analysis to CR (2) (Figure 2) formed three groups with some differences between them when compared to the JAS. One group has formed by ANOVA; the second group by DS and REG; and the third by the analyses of CORR, multivariate analysis, other analyses and scientific papers without analysis. Finally, for RBZ (3) (Figure 2) three groups have also formed, which revealed a group consisting of ANOVA; the second group by REG; and the third by the CORR, DS, multivariate analysis, other analyses and studies without analysis.

Based on these results, ANOVA was the analysis with higher incidence in all three journals analyzed. In most agricultural experiments there is a need to perform multiple comparisons to find differences (or not) between treatments. Such comparisons are based on ANOVA, which needs well-established requirements such as, for example, the homogeneity of the variances and normality of residuals.

The second group was segregated REG analysis for all the journals, but for the JAS, the CORR analysis also made up part of this group, being the analysis regularly used in genetic research. In contrast, DS analysis was the one that remained in this group together with REG for CR, because it contributed substantially with the journal publications, considering that CR includes various areas within Agricultural Sciences. The third group in the three journals was mainly composed of multivariate analysis, other analyses, and studies without analysis. For JAS, DS was also detected within this group; while for CR, CORR was detected; and for RBZ both DS and CORR were detected within this third group.

3.5 Characterization of the statistical analyses

Principal components analysis (PCA) was performed in order to verify the relationship of the statistical analyses (observations) within each journal with the variables (years) characterizing them. Two principal components were considered. The first component had eigenvalue 4.8 and explained 96% of the total variation of the data and the second with 0.1 eigenvalue and explained 2% of the total variance of the data (Table 4). Therefore, the accumulation of the two factors explained 98% of the total variation of the data. However, the results of PCA analysis were discussed based only on the first principal component because the second principal component explains an insignificant part of variation of the data.

Table 4 - Eigenvalues of the principal components

Eigenvalues (CORR)				
Component	Eigenvalues	Difference	Proportion	Cumulative
1*	**4.81**	4.70	0.96	0.96
2	**0.11**	0.06	0.02	**0.98**
3	0.05	0.02	0.01	0.99
4	0.03	0.02	0.01	0.99
5	0.01		0.001	1.000

* Only PCR1 was used because explained 96% of the data variation.

The first principal component explained all five years and eigenvectors of years were highly significant and positive (cutoff value: 0.9) (Table 5). In contrast, the second principal component only explained the years 2011 (eigenvector = 0.23) and 2015 (eigenvector = -0.20), with the value of 0.2 being used as a cutoff. Note that theeigenvector values for the second principal component were much lower compared to the first principal component. The first principal component was named “analyses most used through the years” and the second principal component was named “contrast between the years 2011 and 2015” (Figure 3).

The results in PCA *biplot* (Figure 3) indicated high use of ANOVA for the three journals (CR, JAS and RBZ) in all the years studied (2011 to 2015). Moreover, there was high use of REG and CORR in 2015 by JAS, high use of REG in 2011 and 2012 by RBZ, and high use of DS in 2015 by CR. One more time, ANOVA was evidenced as a statistical method most used in the area, reinforcing the necessity to test hypotheses using such method. Moreover, the use of the other analyses (REG, CORR and DS) in some years could be due of the journals’ scope. Apparently, studies with quantitative levels (REG) are most commons in RBZ and JAS.

Table 5 – Eigenvectors of the variable for the principal components

	Correlation (Structure)				
Variable	PCR1*	PCR2	PCR3	PCR4	PCR5
Year 2011	**0.97**	**0.23**	0.11	0.02	0.02
Year 2012	**0.99**	0.10	-0.08	-0.03	-0.06
Year 2013	**0.99**	-0.08	-0.02	-0.13	0.03
Year 2014	**0.98**	-0.04	-0.12	0.10	0.03
Year 2015	**0.97**	**-0.20**	0.12	0.04	-0.02

*Only PCR1 was used because explained 96% of the data variation.

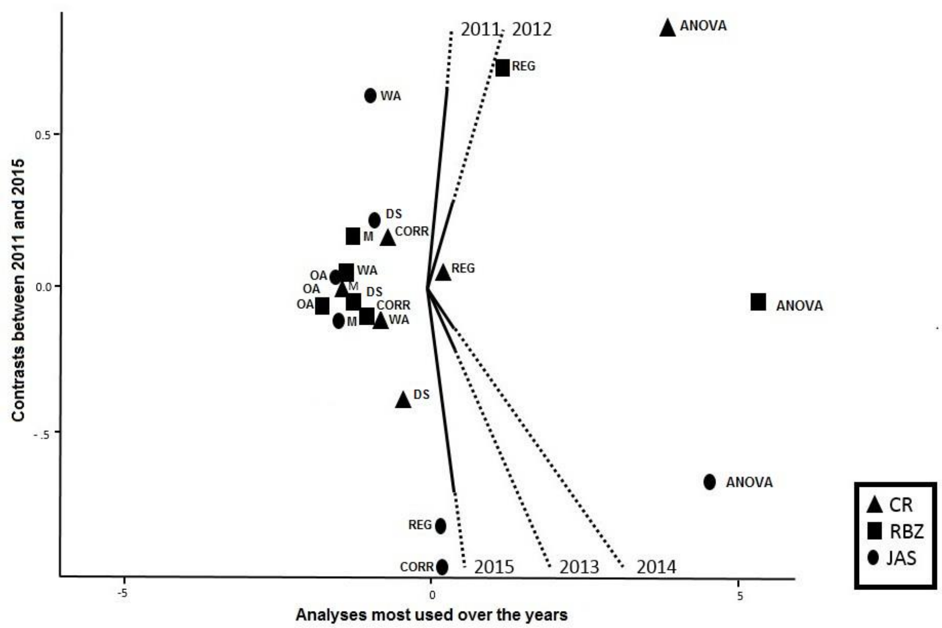

Figure 3 - *Biplot* graphic for principal components analysis on statistical analyses for each journal, according to the years studied. ● = JAS (Journal of Animal Science); ▲= CR (Ciência Rural); ■ = RBZ (Revista Brasileira de Zootecnia); ANOVA - Analysis of variance and mean comparison; REG - Regression analysis; CORR - Correlation analysis; DS - Descriptive statistics; M - Multivariate analysis; OA - Other analyses; WA - without analysis.

3.6 MANOVA and profile analysis

Descriptive statistics was used to study the behavior of different statistical analyses within each journal. There was a slight increase in the use of ANOVA in JAS and RBZ, while decreasing through the years in CR (Figure 4). In JAS, REG analysis increased. By contrast, REG analysis declined from 2013 in RBZ, and remained constant in CR. Furthermore, CORR analysis presented a slight increase in JAS from 2012. Finally, other analyses remained constant along time, taking into consideration that they have not been frequently used.

Based on these descriptive differences between the journals, the aim was then to observe whether there were differences between the journals regarding the use of all the analyses. Thus, multivariate contrasts between journals were studied by MANOVA and, the result obtained was that there was no significant difference between the journals in a general way; that is, all analyses were used with no difference in proportions in the journals (Table 6).

However, time-effect may have occurred with respect to the use of analyses. For this reason, the data were submitted to profile analysis to evaluate the behavior of journals over time, where the hypothesis of profiles parallelism was tested. According the statistical tests there was no statistical difference, indicating that the profiles of the journals are parallel with no difference between the profiles along time (Table 7 and Figure 5).

Table 6 - Multivariate contrasts between journals obtained by multivariate analysis of variance

Statistical test	Contrasts		
	CR vs JAS	CR vs RBZ	JAS vs RBZ
	P > F	P > F	P > F
Wilks' Lambda	0.99	0.92	0.92^{NS}
Pillai's Trace	0.99	0.92	0.92^{NS}
Hotelling–Lawley Trace	0.99	0.92	0.92^{NS}
Roy's Greatest Root	0.99	0.92	0.92^{NS}

^{NS}- non-significant

Table 7 - Results for statistical parallelism profile tests (treatment effect) obtained by profile analysis

Statistical test	Value	Value F	DF	Den DF	P > F
Wilks' Lambda	0.78	1.04	4	15	0.42^{NS}
Pillai's Trace	0.22	1.04	4	15	0.42^{NS}
Hotelling–Lawley Trace	0.28	1.04	4	15	0.42^{NS}
Roy's Greatest Root	0.28	1.04	4	15	0.42^{NS}

^{NS}- non-significant

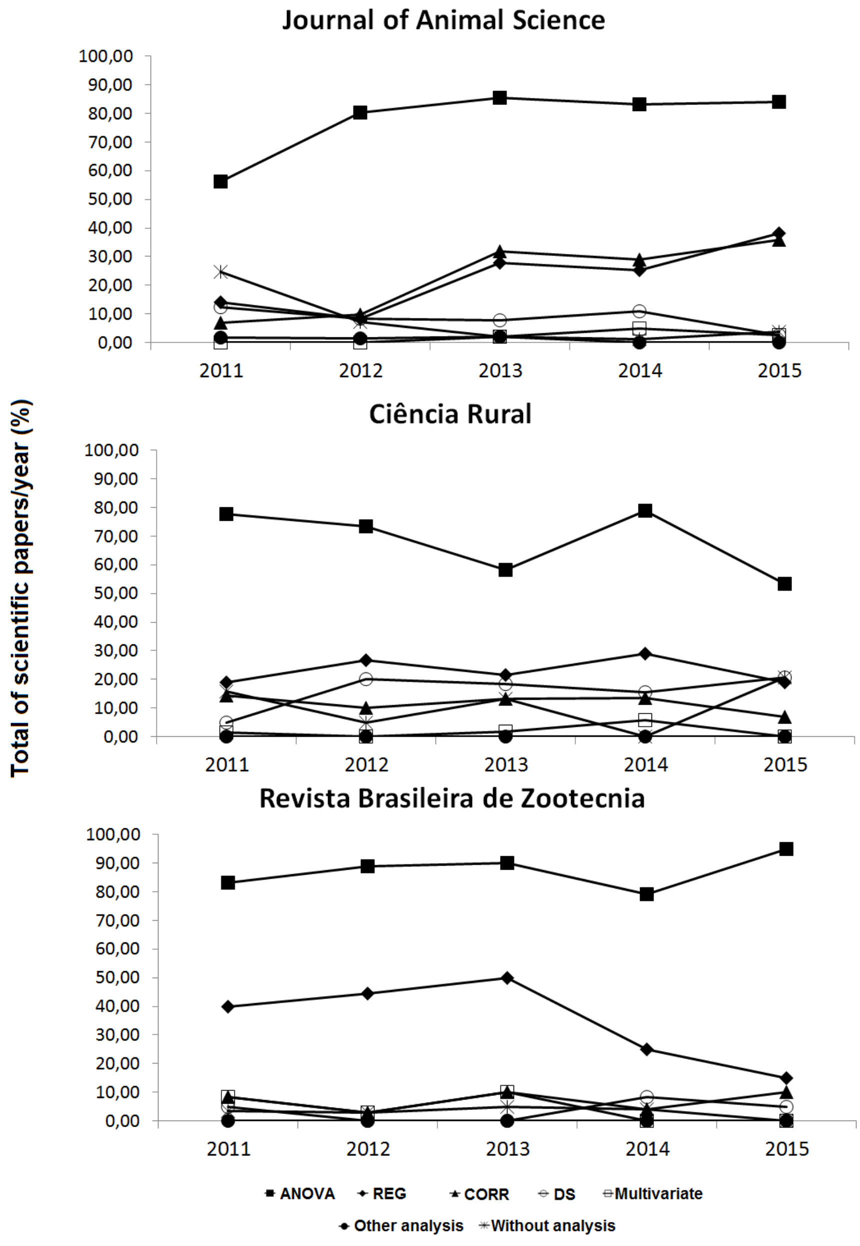

Figure 4 - Statistical analyses used over the years in scientific papers published on JAS, CR and RBZ during the period from 2011 to 2015. ANOVA - Analysis of variance and mean comparison; REG - Regression analysis; CORR - Correlation analysis; DS - Descriptive statistics.

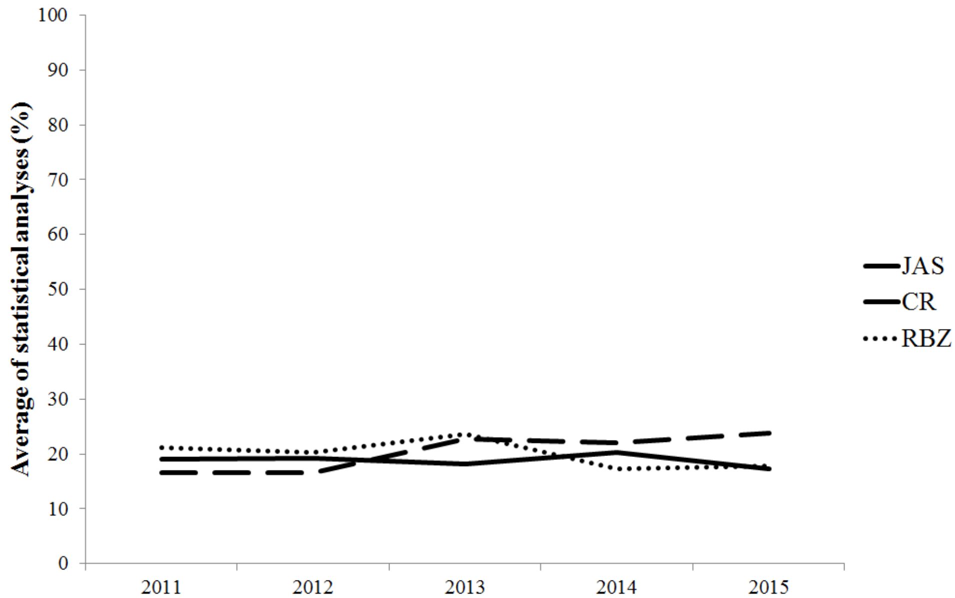

Figure 5 - Graph of average profiles for statistical analyses used as a percentage for each journal in each year of the period studied. JAS – Journal of Animal Science; CR – Ciência Rural; RBZ – Revista Brasileira de Zootecnia.

When parallelism between the profiles is found, we can hypothesize a time-effect within the journals and a journal effect within the time period. As parallelism was not found, it is not necessary to perform these tests (EYDURAN *et al.*, 2008). On the other hand, hypotheses of coincidence and horizontality profiles were performed. The analysis of variance, which checks the coincidence of profiles, found no statistical difference (0.99) (Table 8). This indicates that the profiles were coincident and the average vector of the percentages of analyses for each journal did not differ.

Table 8 - Analysis of variance with regard to the coincidence of profiles obtained by profile analysis

Variation source	DF	Type III SS	Mean Square	Value F	P > F
Journals	2	44.93	22.47	0.01	0.99^{NS}
Residuals	18	68450.02	3802.78		

^{NS}- non-significant

Sequentially, horizontality profile tests demonstrated that there was no statistical difference (Table 9 and Figure 9). This confirms that the profiles for the three journals are horizontal, i.e. the difference between average dependent levels is not significant. Therefore, average profiles of the journals remained constant within periods and there were no differences throughout the years between the profiles of the journals.

Table 9 - Statistical results for horizontality of profiles (time-effect) obtained by profile analysis

Statistical test	Value	Value F	GL	Den DF	P > F
Wilks' Lambda	0.67	0.83	8	30	0.58 ^{NS}
Pillai's Trace	0.36	0.88	8	32	0.54 ^{NS}
Hotelling–Lawley Trace	0.44	0.80	8	19.25	0.61 ^{NS}
Roy's Greatest Root	0.25	1.02	4	16	0.43 ^{NS}

^{NS}- non-significant

Results obtained from PA indicate that the use of statistical analyses did not vary and remained constant within journals along time. Once again, such results emphasize that there was no alteration in the use of statistical analyses by the journals in a general way. This reinforces the results presented previously, where ANOVA was the analysis most frequently used (Table 3). Moreover, also PCA and cluster analysis confirm such information.

3.7 Final considerations

Some results were affected by the CR's thematic amplitude when compared to the JAS and RBZ, because CR covers more areas within agricultural sciences, while JAS and RBZ cover a restricted area involving animal experimentation. Nevertheless, the principal findings in this study indicate that CRD and RBD are the principal experimental designs used in animal science. Both the designs are relatively easier to implement than others (COCHRAN and COX, 1957) and can be used within the greater part of animal studies. On the other hand, experimental designs, such as LSD, are used much more in studies with ruminants. This design could have more importance in a journal focused on studies with ruminants.

With regard to software used to statistical analyses, SAS® was the most used. There was a lot of other software used in the scientific papers. However, currently, SAS® is one of the most well-known and reliable statistical software in the world, despite being commercial software, compared to R® which is free software, for example. In the Brazilian journals some software developed locally are also often used. On the other hand, JAS has software present in its scientific papers that were not present in the papers of the Brazilian journals. All these results highlight the quantity of tools available to perform statistical techniques in animal research.

Analysis of variance is the principal statistical procedure used in animal science. Regression seems be the second most important. It emphasizes the importance of hypothesis testing in animal research. Moreover, through multivariate analyses (PCA, cluster, MANOVA and PA) it was evident that there is no tendency of change in the frequency of use of such analysis, i.e., ANOVA and REG continued being the principal analyses used throughout the years studied. It was an interesting and expected result, but also raises some questions such as, "With the current development of multivariate statistical methods is it possible that such analyses will increase in animal science?", and, "With the development of new statistical methods, could ANOVA decrease its use in animal science?" Such questions may be answered in future studies.

Conclusions

Completely random design and randomized block design are the principal experimental designs used in animal science. The principal software used in animal science was SAS®, however, a lot of other software are used in animal research, but with less frequency. Analysis of variance is the principal statistical analysis used in animal science, followed by regression analysis. Descriptive statistics, correlation analysis, multivariate analysis and other analyses are less frequently used and can depend on the journals' scope. These results highlight the importance of hypothesis testing in animal science.

Acknowledgements

The authors would like to thank the two reviewers and editors for their important comments and suggestions.

BUSANELLO, M., ANDRADE, T. S., AROEIRA, C. N., DIAS, C. T. S. Técnicas estatísticas aplicadas em três periódicos da área de ciências agrárias com enfoque em zootecnia. *Rev. Bras. Biom*. Lavras, v.36, n.2, p.454-472, 2018.

- *RESUMO: Existem poucos estudos prévios que investigam quais as técnicas estatísticas mais utilizadas na área da zootecnia. Devido à grande quantidade de ferramentas e métodos disponíveis para realizar análises estatísticas, torna-se importante identificar quais os mais aplicados para esta área de pesquisa. Portanto, objetivou-se identificar a utilização de diferentes técnicas estatísticas (delineamentos, softwares e análises) usadas em dois periódicos brasileiros (Ciência Rural e Revista Brasileira de Zootecnia) e um internacional (Journal of Animal Science). Desta forma, foram selecionados artigos científicos publicados durante os anos de 2011 a 2015 para formar uma base de dados. Este artigo aborda o uso de delineamentos, softwares e análises mais comumente utilizadas nos periódicos estudados. Para estudar isso, utilizaram-se estatísticas descritivas e abordagens multivariadas. O delineamento inteiramente aleatorizado e de blocos ao acaso foram os principais utilizados na área de zootecnia. O software SAS foi o principal software utilizado nas pesquisas. Por fim, análise de variância foi o principal método estatístico seguido da análise de regressão. Não houve diferenças entre os periódicos e ao longo do tempo com relação à utilização de análises estatísticas. Os resultados reforçam a importância dos testes de hipóteses dentro da área de zootecnia.*

- *PALAVRAS-CHAVE: Experimentação animal, delineamento experimental, métodos estatísticos, softwares estatísticos.*

References

BAGULEY, T. Understanding statistical power in the context of applied research. *Applied Ergonomics*, v.35, n.2, p.73-80, 2004.

CANTUARIAS-AVILÉS, T.; DIAS, C. T. S. Utilization of statistical techniques in two journals of fruticulture. *Ciência Rural*, v.38, n.8, p.2366-2370, 2008.

CARDELLINO, R. A.; SIEWERDT, F. Utilização adequada e inadequada dos testes de comparação de médias. *Revista Brasileira de Zootecnia*, v.21, n.6, p.985-995, 1992.

CAVALIERE, R. How to choose the right statistical software? - A method increasing the post-purchase satisfaction. *Journal of Thoracic Disease*, v.7, n.12, p.585-598, 2015.

CIÊNCIA RURAL, Santa Maria, v.41, n.1, n.2, 2011; v.42, n.4, n.8, 2012; v.43, n.3, n.8, 2013; v.44, n.3, n.11, 2014 e v.45, n.1, n.4, 2015.

COCHRAN, W. G.; COX, G. M. *Experimental designs*, 2.ed. New York: John Wiley and Sons, 1957.

EYDURAN, E.; YAZGAN, K.; OZDEMIR, T. Utilization of profile analysis in Animal Science. *Journal of Animal and Veterinary Advances*, v.7, n.7, p.796-798, 2008.

FERREIRA, D. F. Sisvar: a computer statistical analysis system. *Ciência e Agrotecnologia*, Lavras, v.35, p.1039-1042, 2011.

FESTING, M. F. W.; ALTMAN, D. G. Guidelines for the design and statistical analysis of experiments using laboratory animals. *Institute for Laboratory Animal Research Journal*, v.43, n.4, p.244-258, 2002.

FIELD, A.; MILES, J. *Discovering statistics using SAS statistics*. Thousand Oaks: Sage Publications Inc., 2010.

HU, L. P.; BAO, X. L. How to choose an appropriate experimental design type (Part 1). *Chinese Journal of Integrative Medicine*, v.10, n.6, p.615-618, 2012.

JOURNAL OF ANIMAL SCIENCE, Champaign, v.89, n.3, n.6, 2011; v.90, n.1, n.2, 2012; v.91, n.8, n.12, 2013; v.92, n.7, n.10, 2014 e v.93, n.1, n.2, 2015.

LITTELL, R. C.; HENRY, P. R.; AMMERMAN, C. B. Statistical analysis of repeated measures data using SAS procedures. *Journal of Animal Science*, v.76, n.4, p.1216-1231, 1998.

LITTELL, R. C.; PENDERGAST, J.; NATARAJAN, R. Modelling covariance structure in the analysis of repeated measures data. *Statistics in Medicine*, v19, n.13, p.1793-819, 2000.

LÓPEZ, J. A. G. Descriptive study of statistical methods in the original articles published on the cigarette smoking habit in four Spanish medical journals (1985-1996). *Revista Española de Salud Pública*, v.74, n.1, p.33-43, 2000.

LÚCIO, A. D.; LOPES, S. J.; STORCK, L.; CARPES, R. H.; LIEBERKNECHT, D.; NICOLA, M. C. Características experimentais das publicações da Ciência Rural de 1971 a 2000. *Ciência Rural*, v.33, n.1, p.161-164, 2003.

RESENDE, M. D. V. *Matemática e estatística na análise de experimentos e no melhoramento genético*. Empresa Brasileira de Pesquisa Agropecuária - Florestas (EMBRAPA), Colombo, PR, 2007.

REVISTA BRASILEIRA DE ZOOTECNIA, Viçosa - MG, v.40, n.5, n.10, 2011; v.41, n.10, n.11, 2012; v.42, n.7, n.10, 2013; v.43, n.3, n.8, n.12, 2014 e v.44, n.1, n.2, n.3, n.9, 2015.

ROBINSON, P. H.; WISEMAN, J.; UDÉN, P.; MATEOS, G. Some experimental design and statistical criteria for analysis of studies in manuscripts submitted for consideration for publication. *Animal Feed Science and Technology*, v.129, p.1–11, n.1-2, 2006.

SANTOS, J. W.; MOREIRA, J. A. N.; BELTRÃO, N. E. M. Avaliação do emprego dos testes de comparação de médias na revista Pesquisa Agropecuária Brasileira (PAB) de 1980 a 1994. *Pesquisa Agropecuária Brasileira*, v.33, n.3, p.225-230, 1998.

ST-PIERRE, N. R. Meta-analyses of experimental data in the animal sciences. *Revista Brasileira de Zootecnia*. v.36, suplemento especial, p.343-358, 2007.

THOMAS, L.; JUANES, F. The importance of statistical power analysis: an example from animal behavior. *Animal Behavior*, v.52, n.4, p.856-859, 1996.

UDÉN, P.; ROBINSON, P. H.; MATEOS, G. G.; BLANK, R. Use of replicates in statistical analyses in papers submitted for publication in Animal Feed Science and Technology. *Animal Feed Science and Technology*, v.171, n.1, p.1-5, 2012.

Received on 21.10.2016

Approved after revised on 28.08.2017