Ciência Rural, Santa Maria, v.43, n.9, p.1553-1560, set, 2013
ISSN 0103-8478

# An alternative method for assessing the equivalence among treatments in survival analysis

## Um método alternativo para avaliar a equivalência entre tratamentos em análise de sobrevivência

**Luiz Alexandre Peternelli[I*] Flávia Sílvia Corrêa Tomaz[II] Diego Paiva Bernardes[I]**
**Gilson Fernandes da Silva[III] Fabricia Gonçalves Lacerda[IV]**

## ABSTRACT

*This work aims to evaluate the use of parametric models instead of a nonparametric procedure to adjust survival curves related to different treatments, and to verify the equivalence among treatments by the multivariate method of cluster analysis. The dataset used to validate the method was obtained from a laboratory experiment with cutting ants. Eight colonies of cutting ants were used, each one receiving different treatments. The exponential, Weibull, log-normal, and logistic models were adjusted for each treatment, along with the usual Kaplan-Meier adjustment. The logistic model used was the best option for evaluating the survival of the ants. Therefore, this model was adjusted for each treatment. The estimates of the parameters of each adjusted model were clustered using Ward's method of multivariate technique of cluster analysis. Finally, heuristic techniques for choosing the number of clusters were applied in order to define the sets of equivalent treatments. For the dataset used, the proposed method was less laborious and as efficient as the logrank for the comparison of many survival curves.*

**Key words**: *logrank test, cluster analysis, curves comparison.*

## RESUMO

*Este trabalho tem como objetivo avaliar o uso de modelos paramétricos, ao invés do usual procedimento não paramétrico, no ajuste de curvas de sobrevivência em relação a diferentes tratamentos e ainda verificar a equivalência entre esses tratamentos através da técnica multivariada de análise de agrupamento. O conjunto de dados usado para validar o método foi obtido através de um experimento de laboratório com formigas cortadeiras. Oito colônias de formigas cortadeiras foram usadas, cada uma recebendo um tratamento diferente. Os modelos exponencial, Weibull, log-normal e logístico foram ajustados para cada tratamento além do usual ajustamento de Kaplan-Meier. O modelo logístico foi o mais aceitável para avaliar a sobrevivência das formigas. Portanto, esse modelo foi ajustado para cada tratamento. As estimativas dos parâmetros de cada modelo ajustado foram agrupadas utilizando o método de Ward da técnica multivariada de análise de agrupamento. Para o conjunto de dados usado, o método proposto foi menos trabalhoso e tão eficiente quanto o teste logrank para a comparação das várias curvas de sobrevivência.*

**Palavras-chave**: *teste logrank, análise de agrupamento, comparação de curvas.*

## INTRODUCTION

Survival analysis is one of the most important areas in statistics. In this kind of analysis the response variable is the time until the occurrence of an event of interest, and its main characteristic is the presence of censure, which is defined as an incomplete or partial observation (KLEINBAUM & KLEIN, 2005). Datasets involving lifetimes or failures are represented by a non-negative random variable, usually continuous, and generally specified by its survival function. A survival function can be achieved by adjusting parametric models, usually nonlinear models, instead of the usual non-parametric Kaplan-Meier function (COLOSIMO & GIOLO,

---

[I] Departamento de Estatística, Universidade Federal de Viçosa (UFV), 36570-000, Viçosa, MG, Brasil. E-mail: peternelli@ufv.br. *Autor para correspondência.

[II] Instituto de Ciências Sociais Aplicadas (ICSA), Universidade Federal de Ouro Preto (UFOP), Mariana, MG, Brasil.

[III] Departamento de Engenharia Florestal, Centro de Ciências Agrárias (CCA), Universidade Federal do Espírito Santo (UFES), Jerônimo Monteiro, ES, Brasil.

[IV] CCA, UFES, Alegre, ES, Brasil.

**Received 09.27.12 Approved 04.08.13 Returned by the author 01.07.13**
**CR-2012-0909.R1**

2006). Since survival data are usually asymmetric, the assumption of normality is not achieved. Therefore, the exponential, the Weibull or the log-normal distributions are generally used for these data (COLLETT & KIMBER, 2011).

The main objective of most of planned experiments is to compare treatments. In survival analysis treatments are usually represented by a function relating lifetimes (or failures) as the X variable and percentage of surviving as the Y variable. Comparison of these treatments is usually conducted through nonparametric procedures, with special attention paid to the *logrank* test (COLOSIMO & GIOLO, 2006). This test allows comparisons of two or more survival curves related to different treatments. In the case of multiple comparisons, i.e. comparisons among several treatments, where each treatment is represented by a different survival curve, it is necessary to pay attention to the level of significance used for the test. Since the overall significance level tends to increase when the number of pairwise tests simultaneously performed increases, the use of corrections for the significance level is suggested to avoid erroneous conclusions (QUINN & KEOUGH, 2006). However, when the number of pairwise tests is too high, the corrections can make the procedures very conservative (QUINN & KEOUGH, 2006). Moreover, a large number of pairwise tests can make the process laborious. Thus, these related problems may discourage some researchers from using this technique.

In cases where many non-linear models are adjusted, with each one representing the response of an individual submitted to a specific treatment, it is possible to check the identity of the models based on parametric procedures (BATES & WATER, 1988; NETER et al., 1996; REGAZZI & SILVA, 2004, 2010). However, these comparisons would also be very laborious and time-consuming if many parameters within models or even full models are to be compared.

Facing these hardships, the use of cluster techniques in order to compare models, as proposed in this paper, offers several benefits. Its main advantage is that the fitted models can be grouped in just one step. Other advantages are simplicity of use and easy interpretation. MATOS JÚNIOR et al. (1999) evaluated the curves of orange maturation by cluster analysis. PETERNELLI et al. (2005) proposed the application of this technique to compare treatments defined by different logistic regression equations applied to survival average percentage data. Following the same reasoning, MAIA et al. (2009) used cluster techniques to compare nonlinear regression models in banana trees. All of these studies indicated that the cluster method is promising and much credit should be given to it.

The two objectives of this work were: (i) evaluating the use of parametric models in place of the Kaplan-Meier nonparametric method to estimate the survival functions of cutting ants submitted to different treatments; and (ii) evaluating the use of Ward's method for clustering similar treatments, based on estimates of the parameters according to the more appropriated parametric model. Pursuing these two objectives will help indicate that the approach proposed herein is a promising method for the comparison of curves based on heuristics.

## MATERIAL AND METHODS

The proposed method consists of three steps: i) from a set of candidate parametric models, finding the most acceptable one to fit the data; ii) using estimates of the selected model parameters on the cluster analysis via Ward's method; and iii) considering the graphical analysis defined in KHATREE & NAIK (2000) to choose the number of groups.

The dataset used to demonstrate and to validate the proposed method was obtained from an experiment that evaluated the effect of the litter of the colony on the mortality rate of ants ***Atta sexdens rubropilosa*** *Forel* that were fed with different substrates, as described in detail by LACERDA et al. (2009). Eight colonies of cutting ants were used in the experiment. Twenty Petri plaques were used for each colony. Ten of them contained litter from the anthill, while the other ten contained no litter. Ten worker ants from the respective colonies were placed in each plaque. The plaques were maintained in an appropriate location, in the laboratory, in a controlled environment. Within a period of 30 days the event of interest, namely the death of the ants, was observed and recorded in a daily basis, characterizing the type I censure for this data, since not all ants were dead by the end of the experiment. The treatments received the following identification: $T_1$ - Acalypha (***Acalypha wilkesiana*** Müller.Arg) with the presence of litter; $T_2$ - Acalypha without the presence of litter; $T_3$ - Ligustrum (***Ligustrum japonicum*** Thunb) with the presence of litter; $T_4$ - Ligustrum without the presence of litter; $T_5$ - Eucalyptus (***Eucalyptus urophylla*** S.T. Blake) with the presence of litter; $T_6$ - Eucalyptus without the presence of litter; $T_7$ - mixed substrate, comprising the three types of plants previously mentioned, with the presence of litter; and $T_8$ - mixed substrate without the presence of litter.

Four parametric models were used to estimate the survival functions: (i) exponential, $S(t)=\exp\left\{-\left(\frac{t}{\alpha}\right)\right\}$; (ii) Weibull, $(t)=\exp\left\{-\left(\frac{t}{\alpha}\right)^{\gamma}\right\}$; (iii) log-normal, $S(t)=\Phi\left(-\frac{\log(t)+\mu}{\sigma}\right)$; and (iv) logistic, $S(t)=\frac{1}{1+\exp\left\{\frac{t-\mu}{\sigma}\right\}}$.

In the above-mentioned models, S(t) is the survival over time (in days); $\alpha$ is a scale parameter associated with the average lifetime in the exponential model; $\gamma$ is a shape parameter associated with the hazard ratio; and $\mu$ and $\sigma$ are, respectively, scale and location parameters (COLOSIMO & GIOLO, 2006; LEE & WANG, 2003).

For each treatment, the survival curve estimated by Kaplan-Meier and the survival curves defined by the parametric models were adjusted using the *survreg* function of the R software (R DEVELOPMENT CORE TEAM, 2011). A graphical technique described by COLOSIMO & GIOLO (2006) was used for selection of the most adequate model. Graphics were constructed for the estimates of survival times achieved by the Kaplan-Meier versus the estimates of survival times achieved through each parametric model. The model whose points were closer to a straight line representing the cumulative distribution function of a random variable with uniform distribution (0, 1) was selected as the best model.

Next, estimates of the parameters of the selected model were clustered using the Ward method, as described by MINGOTI (2005) and KHATREE & NAIK (2000). The clustering was carried out through the *hclust* function of the R software (R DEVELOPMENT CORE TEAM, 2011).

As suggested by KHATREE & NAIK (2000), the following statistics were used to determine the optimum number of groups: sum of squares between groups (RSQ); semi-partial correlation (SPRSQ); and distance between groups (BSS). The calculations of these measures were carried out through proper functions developed on the R software (R Development core team, 2011). Based on these statistics, groups of models considered equivalent were determined.

The clustering result was compared to the result achieved by the usual multiple comparison of the treatments, carried out via the *logrank* test with Bonferroni correction as a control of type I error, as suggested by COLOSIMO & GIOLO (2006). The *logrank* test is described in COLLET & KIMBER (2011) and in COLOSIMO & GIOLO (2006).

We also evaluated the stability of cluster analysis in determining the treatments' equivalency. For this data, the parameters estimated for the selected model were also clustered using other methods: centroid distance, nearest-neighbor, furthest-neighbor, and average distance. The fit between the distance matrix and each dendrogram was estimated using the cophenetic correlation coefficient (SOKAL & ROHLF, 1962).

## RESULTS

Figure 1 presents graphical descriptions of the survival functions estimated by Kaplan-Meier *versus* the survival times estimated by the Weibull, log-normal, exponential and logistic models, for each of the eight treatments. The exponential was not appropriate to describe the dataset of the respective treatment since it presented the longest distance from the straight line y = x for most treatments.

For $T_6$ and $T_8$, only the logistic model adjusted to the data. The graphical representations revealed that the curves for these treatments did not present large distances from the straight line y = x. For the other treatments, the Weibull, log-normal, and logistic models seemed to be most reasonable. Graphics of the linearized Weibull, log-normal, and logistic models were also constructed (data not shown), indicating that the log-normal model could be discarded. Figure 2 presents a comparison between each survival curve estimated by the final parametric models candidates (Weibull and logistic) and the Kaplan-Meier non-parametric model, showing a good fit to the data. For these data, because the Weibull model did not fit for $T_6$ and $T_8$, it was discarded, and the logistic model was finally chosen. Therefore, the use of cluster analysis for the comparison of survival curves was carried out by clustering the estimates of the coefficients of this model. The estimates for the two parameters ($\mu$ and $\sigma$) obtained for each treatment $T_i$ (i = 1 to 8) were, respectively, 10.81 and 3.81 ($T_1$), 14.51 and 3.58 ($T_2$), 8.51 and 2.95 ($T_3$), 17.73 and 4.26 ($T_4$), 5.35 and 1.29 ($T_5$), 15.31 and 3.90 ($T_6$), 9.77 and 2.99 ($T_7$), and 17.43 and 4.33 ($T_8$). These estimates formed the data matrix of dimension 8 × 2 to be used in the clustering of the treatments.

Table 1 shows the results of the cluster analysis and auxiliary statistics on the determination of the number of groups. The graphical representations of these statistics are presented in figure 3. The analysis of table 1 and figure 3 reveals the existence of three groups, following arguments presented in KHATREE & NAIK (2000). The results achieved by cluster

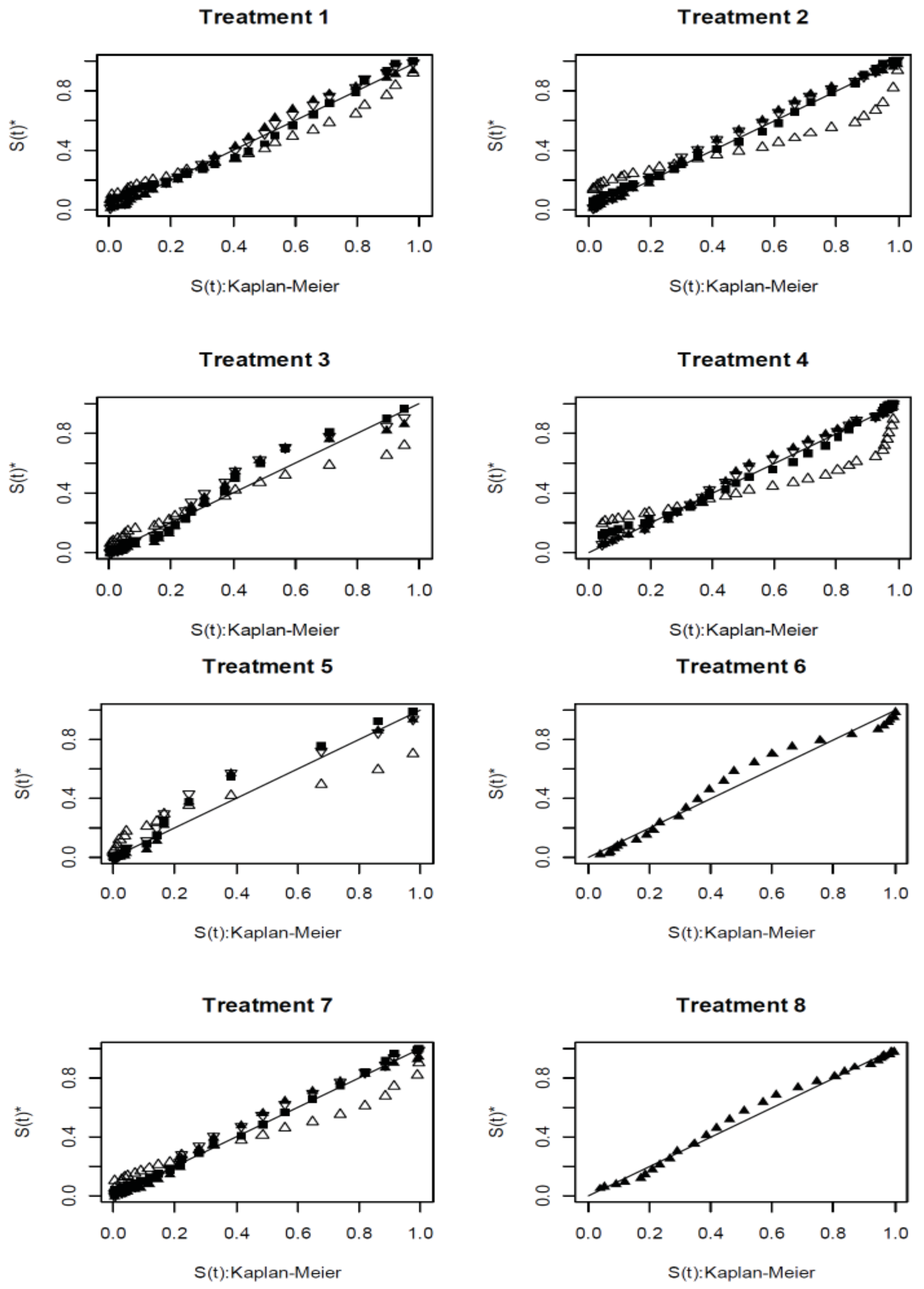

Figure 1 - Survival times estimated over eight treatments by Kaplan-Meier versus survival times, S(t)*, estimated by the following models: exponential (△), Weibull (▽), log-normal (■), and logistic (▲). In treatments 6 and 8, only the logistic model could be adjusted. The continuous line indicates maximum equivalence, or perfect correlation, between pairwise models.

analysis suggest equivalence among treatments 2, 4, 6, and 8 (treatments without the presence of litter); equivalence among treatments 1, 3, and 7 (treatments with the presence of litter) and differences in treatment 5 compared to the others. These comparisons can be easily seen by depicting the curves related to the logistic model (treatments 1 to 8 from figure 2) in a simple graph. LACERDA et al. (2009) found the same type of equivalence among the treatments using a different methodology for comparison.

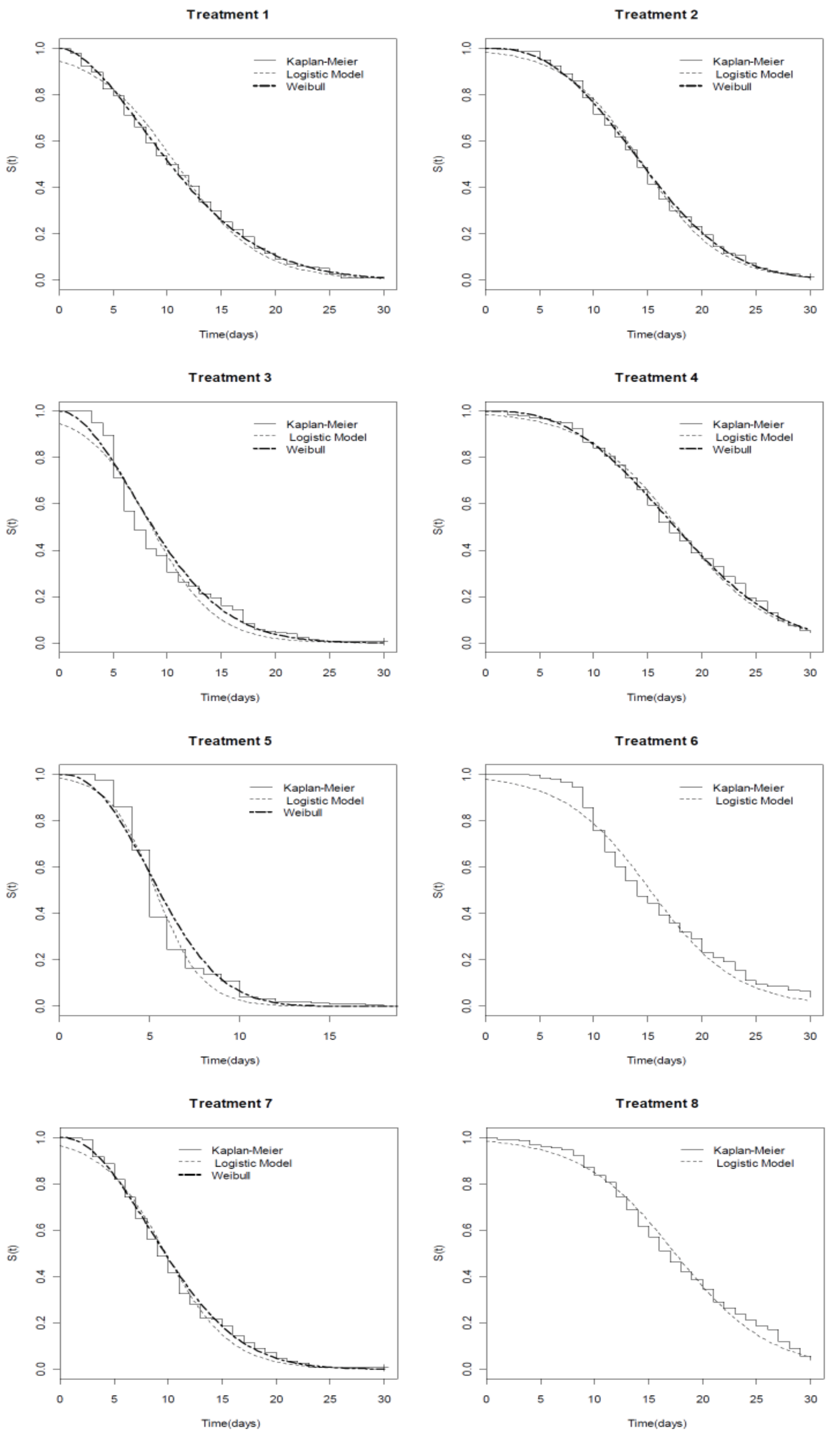

Figure 2 - Comparison between the survival curve estimated by Kaplan-Meier versus curves estimated by the logistic and Weibull models for treatments 1, 2, 3, 4, 5, 6, 7, and 8.

Table 1 - Summary of the auxiliary statistical measurements: distance between groups (BSS); sum of squares between groups (RSQ); and semipartial correlation (SPRSQ) for the clustering of the treatments, according to the parameters' estimates of the adjusted model and the number of groups (NG).

| Step | Groups | NG | BSS | RSQ | SPRSQ |
|---|---|---|---|---|---|
| 1 | {1},{2},{3},{4,8},{5},{6},{7} | 7 | 0.0475 | 1.000 | 0.0003 |
| 2 | {1},{3},{4,8},{5},{7},{2,6} | 6 | 0.3712 | 0.997 | 0.0025 |
| 3 | {1}, {2,6},{3,7},{4,8},{5} | 5 | 0.7946 | 0.992 | 0.0054 |
| 4 | {1,3,7},{2,6},{4,8},{5} | 4 | 2.3297 | 0.976 | 0.0158 |
| 5 | {1,3,7},{2,4,6,8},{5} | 3 | 7.4369 | 0.926 | 0.0503 |
| 6 | {1,3,5,7},{2,4,6,8} | 2 | 17.0513 | 0.810 | 0.1154 |
| 7 | {1,2,3,4,5,6,7,8} | 1 | 119.7491 | 0.000 | 0.8103 |

The p-values of the comparison among the curves provided by the *logrank* test are presented in table 2. After a difference between the treatments (*logrank* test, T=810, with 7 d.f., P value<0.001) was observed, the *logrank* test with Bonferroni correction (level of significance = α/c, in which α=0.05, and c = number of comparisons performed) was carried out to compare pairwise curves, as suggested by COLOSIMO & GIOLO (2006). No significant differences were found between $T_1$ and $T_7$; $T_2$ and $T_6$; $T_3$ and $T_7$; $T_4$ and $T_6$; $T_4$ and $T_8$, and $T_6$ and $T_8$. It can also be concluded that $T_5$ differs from all the other treatments. The results presented by the *logrank* test are equivalent to those presented by the cluster analysis based on estimates of the parameters via the logistic model.

The cophenetic correlations found for the clustering methods centroid distance, nearest-neighbor, furthest-neighbor, average distance and Ward's were equal to 0.80, 0.76, 0.80, 0.81, and 0.77, respectively.

## DISCUSSION

The inappropriateness of the exponential model to describe the dataset of the respective treatment was already expected, since the survival curves achieved from the average ant survival percentage per day, for each treatment, presented a sigmoidal behaviour.

The graphical analyses indicated that both the Weibull and the logistic models presented satisfactory adjustment for the dataset, except for $T_6$ and $T_8$, in which only the logistic model adjusted to the data.

Since there was no censure in the data, the eligibility of the logistic or the Weibull model to replace the Kaplan-Meier estimator is ensured. In the present work, however, the logistic model could also be used if there were up to 20% of type I censure, achieved via simulation (data not shown). We finally selected the logistic model because it adjusted well to all treatments in the present dataset.

Table 1 brings up an important question regarding the determination of group number. According to KHATREE & NAIK (2000), it is difficult to determine the number of groups, which is based on heuristic criteria. The values of the BSS and SPRSQ statistics must be low, while the RSQ values must be high (SHARMA, 1996). The ideal number of groups is the one corresponding to the point in which

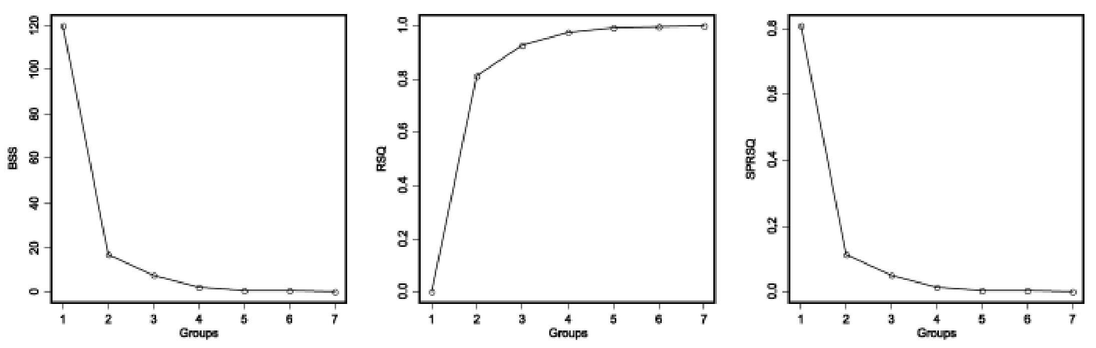

Figure 3 - Graphics of the auxiliary statistics: distance between groups (BSS); sum of square between groups (RSQ); and semipartial correlation (SPRSQ) according to the number of groups.

Table 2 - P-values related to the *logrank* nonparametric test with Bonferroni correction for the comparison of eight treatments ($T_1$ to $T_8$).

| | $T_1$ | $T_2$ | $T_3$ | $T_4$ | $T_5$ | $T_6$ | $T_7$ |
|---|---|---|---|---|---|---|---|
| $T_2$ | $< 0.0018^*$ | | | | | | |
| $T_3$ | 0.0018* | $<0.0018^*$ | | | | | |
| $T_4$ | $< 0.0018^*$ | $< 0.0018^*$ | $< 0.0018^*$ | | | | |
| $T_5$ | $< 0.0018^*$ | $< 0.0018^*$ | $< 0.0018^*$ | $< 0.0018^*$ | | | |
| $T_6$ | $< 0.0018^*$ | $0.0808^{ns}$ | $< 0.0018^*$ | $0.0140^{ns}$ | $< 0.0018^*$ | | |
| $T_7$ | $0.0348^{ns}$ | $< 0.0018^*$ | $0.1070^{ns}$ | $< 0.001^*$ | $< 0.0018^*$ | $< 0.0018^*$ | |
| $T_8$ | $< 0.0018^*$ | $< 0.0018^*$ | $< 0.0018^*$ | $0.9260^{ns}$ | $< 0.0018^*$ | $0.0281^{ns}$ | $< 0.0018^*$ |

$^{ns}$ and $^*$: non-significant and significant at $\alpha/c = 0.0018$ of probability.

the statistical value begins to stabilise. SHARMA (1996) also advises researchers to consider the objective of the study in evaluating the number of groups.

The cophenetic correlation coefficients obtained for the various clustering methods were all greater than 0.7. According to ROHLF (1970) this result indicates that the grouping method used was appropriate to summarize information in the data set. In other words, we could state that the clustering method may be used as an alternative for comparing estimates of fitted models for several treatments. It is important to note that the overall structure of the various methods for clustering was similar, i.e., the groups were formed by the same treatments (results not shown). As the Ward method (WARD, 1963) ensures the maximization of homogeneity within groups, this method has higher statistical appeal than the others, which explains its preference in this proposal.

According to these results, we conclude that the logistic parametric model was efficient and as good as the Kaplan-Meier approach to estimate the survival functions of cutting ants submitted to different treatments, thus representing one more option for describing survival curves. Also, we conclude that the cluster analysis (using Ward's method) based on the estimates of the parameters of the adjusted model was capable of comparing treatment-specific parametric models applied to survival analysis, which highlights the potential of the proposed alternative method for this purpose.

## ACKNOWLEDGMENTS

The authors thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Fundação de Amparo à Pesquisa do estado de Minas Gerais (FAPEMIG) for the financial support with research.

## REFERENCES

BATES, D.M.; WATTS, D.G. **Nonlinear regression analysis and its applications**. New York: John Wiley, 1988. 365p.

COLLET, D.; KIMBER, A. **Modelling survival data in medical research**. 3.ed. London: Chapman & Hall, 2011. 448p.

COLOSIMO, E.A.; GIOLO, S.R. **Análise de sobrevivência aplicada**. São Paulo: Edgar Blücher, 2006. 392p.

KHATREE, R.; NAIK, D.N. **Multivariate data reduction and discrimination with SAS software**. New York: John Wiley and Sons, 2000. 584p.

KLEINBAUM, D.G.; KLEIN, M. **Survival analysis**: a self-learning text. 2.ed. New York: Springer, 2005. 606p.

LACERDAetal.Mortalityof***Attasexdensrubropilosa***(Hymenoptera: ***Formicidae***) workers in contact with colony waste from different plant sources. **Bulletin of Entomological Research**, v.100, p.00-103, 2010. Available from: <http://journals.cambridge.org/action/displayFulltext?type=1&pdftype=1&fid=6900140&jid=BER&volumeId=100&issueId=01&aid=6900132>. Accessed: Dec. 20, 2012. doi:10.1017/S000748530900683X.

LEE, E.T.; WANG, J.W. **Statistical methods for survival data analysis**. 3.ed. New York: John Wiley and Sons, 2003. 534p.

MAIA, E. et al. Method of comparison of non-linear regression models in bananas trees. **Ciência Rural**, v.39, n.5, p.1380-1386, 2009. Available from: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-84782009000500012&lng=pt&nrm=iso>. Accessed: Dec. 20, 2011.

MATOS JÚNIOR, D. et al. Evaluation of different orange maturation curves by means of cluster analysis. **Pesquisa Agropecuária Brasileira**, v.34, n.12, p.2203-2209, 1999. Available from: <http://webnotes.sct.embrapa.br/pab/pab.nsf/4b9327fca7faccde032564ce004f7a6a/4bb9658edf31c4238325685c0045401f?OpenDocument>. Acessed: Dec. 20, 2011.

MINGOTI, S.A. **Análise de dados através de métodos de estatística multivariada**: uma abordagem aplicada. Belo Horizonte: UFMG, 2005. 295p.

NETER, J. et al. **Applied linear statistical models**. 4 ed. New York: McGraw-Hill, 1996. 1408p.

PETERNELLI, L.A. et al. Heurística do uso de análise de agrupamento para comparação de modelos de regressão. In: REUNIÃO ANUAL DA REGIÃO BRASILEIRA DA SOCIEDADE INTERNACIONAL DE BIOMETRIA, 50; SIMPÓSIO DE ESTATÍSTICA APLICADA À EXPERIMENTAÇÃO AGRONÔMICA, 2005, Londrina, PR, Brazil. **Anais...** Londrina, PR: RBRAS, 50. e SEAGRO, 11., 2005. CD. Para uso em PC.

QUINN, G.P.; KEOUGH, M.J. **Experimental designs and data analysis for biologists**. Cambridge: Cambridge University, 2002. 556p.

R DEVELOPMENT CORE TEAM. **R**: A language and environment for statistical computing. Vienna, Austria, 2011. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.

REGAZZI, A.J.; SILVA, C.H.O. Teste para verificar a igualdade de parâmetros e a identidade de modelos de regressão não-linear. I. Dados no delineamento inteiramente casualizado. **Revista de Matemática e Estatística**, v.22, n.3, p.33-45, 2004.

REGAZZI, A.J.; SILVA, C.H.O. Testes para verificar a igualdade de parâmetros e a identidade de modelos de regressão não-linear em dados de experimento com delineamento em blocos casualizados. **Revista Ceres**, v.57, p.315-320, 2010. Available from: <http://www.ceres.ufv.br/ceres/revistas/V57N003P19908.pdf>. Accessed: Nov. 20, 2011.

ROHLF, F.J. Adaptative hierarquical clustering schemes. **Systematic Zoology**, v.19, p.58-82, 1970.

SHARMA, S. **Applied multivariate techniques**. New York: John Wiley & Sons, 1996. 512p.

SOKAL, R.R.; ROHLF, F.J. The comparison of dendrograms by objective methods. **Taxonomy**, v.11, n.1, p.30-40, 1962.

WARD, J.H. Hierarquical grouping to optimize an objective function. **Journal of the American Statistical Association**, v.58, p.236-244, 1963.