Türkiye Klinikleri Biyoistatistik Dergisi
Turkiye Klinikleri Journal of Biostatistics
Turkiye Klinikleri J Biostat. 2021;13(1):28-38

ORİJİNAL ARAŞTIRMA ORIGINAL RESEARCH
DOI: 10.5336/biostatic.2020-78896

Re-evaluating the Monte Carlo Simulation Results by Using Graphical Techniques

Monte Carlo Simülasyon Sonuçlarının Grafik Teknikler Kullanılarak Yeniden Değerlendirilmesi

Mehmet MENDEŞ[a]

[a]Department of Biometry and Genetics, Çanakkale Onsekiz Mart University Faculty of Agriculture, Çanakkale, TURKEY

ABSTRACT Objective: When results of simulation studies are examined, it is seen that there are noticeable differences among the simulation results due to differences in the experimental conditions. This study aim at re-evaluating the results of 35 simulation studies carried out the same purpose by using graphic techniques. **Material and Methods:** Type I error estimates of 35 simulation studies carried out under different experimental conditions (pairing type, variance ratio, sample size ratio, number of simulation, skewness, kurtosis, total sample size, equality of sample size, and number of group) were used as a material of this study. Two different graphical techniques, namely Automatic Linear Modeling and Regression Tree Analysis were used in evaluating Type I error estimates of these studies. **Results:** Statistical analyses results indicated that the results of simulation studies were affected by different factors and these factors should be considered in order to get more reliable and stable estimates. It was observed that the most important factors affecting Type I error estimates were as pairing type, variance ratio, number of simulation, and sample size ratio. Therefore, these factors can be considered the primary factors that might cause getting different results among the studies. **Conclusion:** Both methods are promising and can be used efficiently to determine the factors that affect the response variable when there is a large and complex data set.

Keywords: Classification and regression trees; data mining; interaction; simulation

ÖZET Amaç: Simülasyon çalışmalarının sonuçları incelendiğinde, deneysel koşullardaki farklılıklardan dolayı simülasyon sonuçları arasında önemli farklılıklar olduğu görülmektedir. Bu çalışma, aynı amaçla yürütülen 35 simülasyon çalışmasının sonuçlarının, grafik teknikler kullanılarak yeniden değerlendirilmesini amaçlamaktadır. **Gereç ve Yöntemler:** Eşleştirme tipi, varyans oranı, örnek genişliği oranı, simülasyon sayısı, eğrilik ve diklik katsayıları, toplam örnek genişliği, örnek genişliğinin eşitliği ve grup sayısı gibi farklı deneysel koşullarda gerçekleştirilen 35 simülasyon çalışmasının Tip I hata tahminleri, bu çalışmanın materyali olarak kullanılmıştır. Veri setinin analizinde, Otomatik Doğrusal Modelleme ve Regresyon Ağacı Analizi olmak üzere 2 farklı grafik tekniği kullanılmıştır. **Bulgular:** Yapılan istatistiksel analiz sonucunda, dikkate alınan simülasyon çalışmaları sonuçlarının farklı faktörlerden etkilendiğini, daha güvenilir ve istikrarlı tahminlerin elde edilebilmesi için tespit edilen faktörlerin dikkate alınması gerektiğini göstermiştir. Tip I hata tahminlerini etkileyen en önemli faktörlerin ise sırasıyla eşleştirme tipi ya da varyans oranları ile örnek genişlikleri arasındaki ilişkiler, varyans oranı, çalışmada dikkate alınan simülasyon sayısı ve örneklem büyüklüğü oranı olduğu görülmüştür. Bu bulgulardan hareketle bu 4 faktörün, çalışmada dikkate alınan 35 simülasyon çalışması sonucunda elde edilen Tip I hata olasılıklarının farklılaşmasında başlıca rol oynayan faktörler oldukları sonucuna varılabilir. **Sonuç:** Her iki yöntem de umut vericidir, büyük ve karmaşık bir veri kümesi olduğunda yanıt değişkenini etkileyen faktörleri belirlemek için verimli bir şekilde kullanılabilir.

Anahtar kelimeler: Sınıflandırma ve regresyon ağaçları; veri madenciliği; etkileşim; simülasyon

Many studies are conducted to investigate the relationships between variables or determine the factors that affect the interested variable significantly. Regression based techniques or approaches (i.e. multiple linear regression and logistic regression) are commonly used for these purposes.[1] Two other techniques which can be able to use for the same purposes are Automatic Linear Modeling (ALM) and Classification and Regression Tree (CART).[2,3] Since they are graphical techniques, they have some important advantages over

Correspondence: Mehmet MENDEŞ
Department of Biometry and Genetics, Çanakkale Onsekiz Mart University Faculty of Agriculture, Çanakkale, TURKEY/TÜRKİYE
E-mail: mmendes@comu.edu.tr

Peer review under responsibility of Turkiye Klinikleri Journal of Biostatistics.

Received: 10 Sep 2020 ***Received in revised form:*** 18 Nov 2020 ***Accepted:*** 20 Nov 2020 ***Available online:*** 29 Apr 2021

traditional regression analyses like multiple regression and logistic regression. One of the major advantages of these techniques is that they enable researchers to investigate the complex relations among any types of variables. By using these techniques in a correct way, it will also be possible to uncover latent relations between/among the variables and to determine important factors that affect on interested variable(s). One of the other advantages of these techniques is that it is easy to understand and interpret the analysis results.[1,2,4,5] Morgan reported that CART analysis is a simple and powerful analytic tool that helps researchers to determine the most important factors in a dataset, and craft a potent explanatory model.[6] Simulation studies are increasingly being used in many areas such as statistics, engineering, physics, military, and chemistry. Simulations studies are generally used in statistics for the purpose of: a) comparing performance (e.g. Type I error rate and test power) of different statistical methods, b) evaluating consequences of violation of assumptions (e.g. normality and homogeneity of variance for t-test or ANOVA), c) determining optimum sample size of a statistical method during the planning phase of a study, d) obtaining the sampling distribution for a statistic (Learn how a statistic may vary from sample to sample by drawing sufficiently large number of random samples from a specific population), e) obtaining critical table values for a statistical test (i.e. t-test, ANOVA, Spearman-Rank corr, Tetrachoric corr, etc.), especially when the test statistic doesn't have a closed form distribution, f) obtaining parameter estimation, g) finding solutions to mathematical problems that cannot easily be solved, and ı) Obtaining approximate solutions to mathematical problems (probability).[7-11]

When literature related to simulation studies is examined, it is seen that there is an obvious variation among the simulation studies due to differences in experimental conditions in terms of Type I error and test power estimates, even when conducted for the same purpose. A simple literature search shows that the Type I error estimates vary from 0.00% to 40.00%.[8,9,12-14,15-42] Significant differences among the estimates make the reader confused and cause to have trouble determining which test is more appropriate to use. From this point of view**,** the results of some previous simulation studies were re-evaluated by using two different graphical methods. Thus, both it will be possible to determine the variables that affect the performance of the tests the most, and to estimate the amount of changes in the Type I error rate when there is one unit increase in experimental conditions.

MATERIAL AND METHODS

In this study, Type I error estimates of 35 simulation studies carried out under different pairing type (PT), variance ratio ($VR=Var_{min}/Var_{max}$), sample size ratio (SSR) (ratio of the numbers of observations in the groups to be compared), number of simulation (SN), skewness (Sk), kurtosis (Kr), total sample size (TSS), equality of sample size (ESS), and number of group (NG) were re-evaluated by using two graphical techniques namely ALM and Regression Tree Analysis (RTA). Pairing type stands for relation between sample size and variance of the population which samples are taken. If the number of observations in the group with a large variance is higher, direct pairing is in question. On the other hand, if the number of observations in the group with the largest variance is the lowest, then there is inverse pairing is in question. The Type I error estimates under those experimental conditions were taken as dependent or output variable. Other experimental conditions were added to the models as independent variables. After terminal nodes (homogeneous subgroups) were formed, analysis of means (ANOM) Technique was used in order to compare the means of the terminal nodes in terms of Type I error estimates.

AUTOMATIC LINEAR MODELING

ALM, is considered relatively a new method, introduced in SPSS software (version 19 and up), enabling researchers to select the best subset automatically especially when there are a large numbers of variables. In ALM, the predictor variables are automatically transformed in order to provide an improved data fit, and

SPSS uses rescaling of time and other measurement values, outlier trimming, category merging and other methods for the purpose.[3,43,44]

REGRESSION TREE ANALYSIS

RTA is a recursive partitioning method for predicting continuous predictor variables and it has become increasingly popular for all branches of sciences especially in the presence of large and complex data sets. Since it does not require any *priori* assumptions about the nature of the relationships among the dependent and independent variables and allows for the possibility of interactions and non-linearities among variables the RT has clear advantages over classical statistical methods.[1,2,45-48] Yang et al. reported that the tree-based regression models are popular in literature due to their flexibility to model higher order non-linearity and great interpretability.[49] RTA was used for three purposes in general: a) to investigate the complex relations between dependent and independent variables, b) to determine the combinations of factors that provide the protection of the Type I error rate at 0.05 level and, c) to determine the factor that affect the Type I error rates of the simulation studies.

RESULTS

RESULTS OF AUTOMATIC LINEAR MODELING

Results of the ALM are presented in Figure 1, Figure 2, Figure 3, and Figure 4. In determining an appropriate model for fit our data set, many models have been run (not discussed here) and it has been observed that except the model has been used in this study (-4.982) the other models have large information criterion values and above). The accuracy level of the model which is equivalent to the adjusted R-squared value used to fit data and estimate the changes in Type I error rates is found to be 74.6%, means that our model is a good one which can be able to use in fitting and estimating process (Figure 1).

The lower the information criterion is, the better the model is compared to models with a higher information criterion. Since the model used here has been the lowest information criterion value compared to the many other models (not discussed in this document), this model has been preferred in investigating the relations between the Type I error rate and independent variables (experimental conditions).

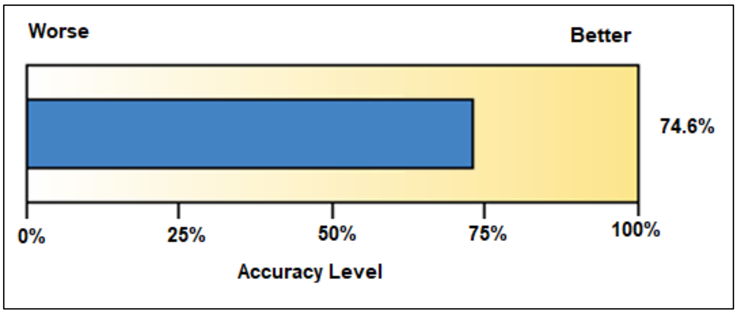

FIGURE 1: Accuracy level of the model.

Importance levels of the predictors have been presented in Figure 2. The importance value is defined as the percent improvement with respect to the most important predictor. Figure 2 shows the predictors in the final model in rank order of importance. For linear models, the importance of a predictor is the residual sum of squares with the predictor removed from the model, normalized so that the importance values sum to 1. When Figure 2 is examined, it is seen that the most important variables or factors that affect alpha estimates are PT, VR, SSR, and SN. Therefore, it can be concluded that the factors related to sample size, relations between sample size and VR, and SN should be taken into consideration in order to get reliable and stable estimates when comparing the performances of tests that can be used for the same purpose.

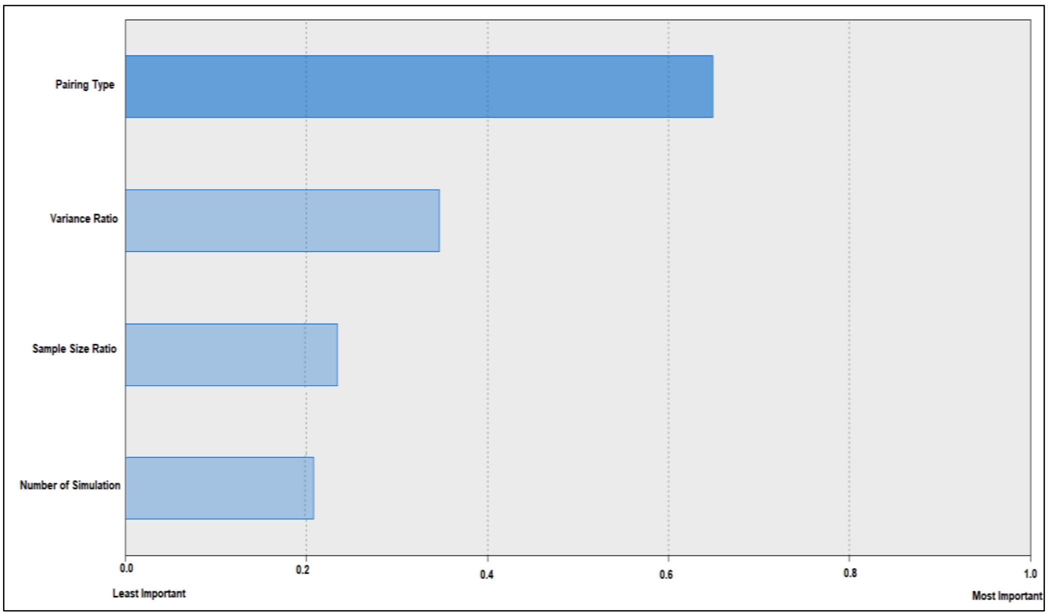

FIGURE 2: Importance levels of the variables or predictors.

After the most important factors or variables are determined, the ALM is re-run using just importance variables namely PTs, VR, SSR, and SN (Figure 2). The final results of ALM are presented in Figure 3, and Figure 4.

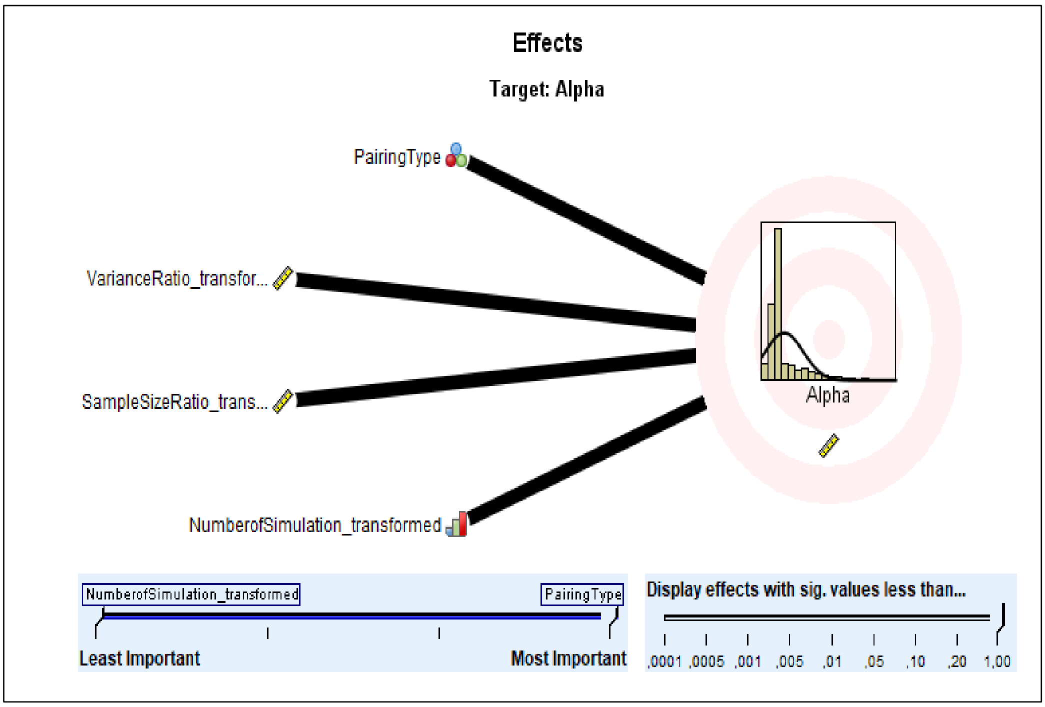

FIGURE 3: Diagram for showing statistical significance of variables or factors.

This diagram is corresponding to the ANOVA table after multiple linear regression analysis. The predictors are ordered from top to bottom in order of importance levels, and the thickness of each line shows the statistical significance (p values) of the relevant effect. As it is seen from the diagram in Figure 3, the PT is the most important predictor or the factor affect the Type I error estimates, and it is followed by VR, SSR, and SN based on importance levels of them. Effect of all four predictors are statistically significant (p=0.001). Based on the results of Figure 3, the regression equation for predicting the alpha is created as alpha=0.131-0.102PT+0.076VR+0.031SSR-0.041NS. The interpretation of the coefficients of this equation is the same as traditional multiple regression coefficients.

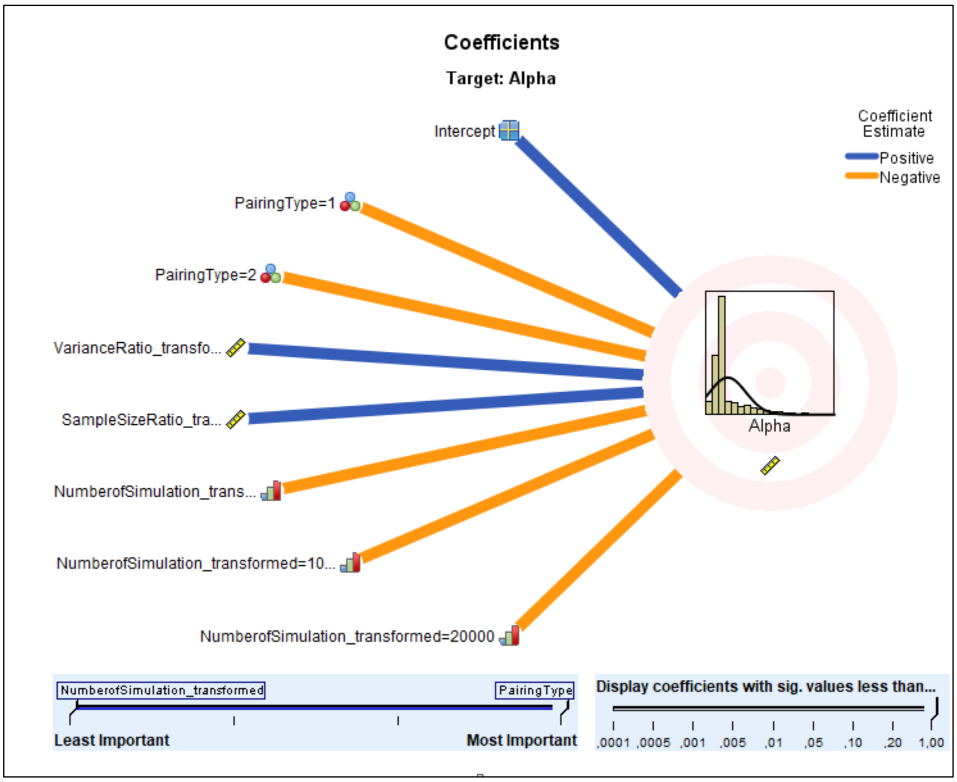

FIGURE 4: Diagram for showing detailed parameter estimates, their signs, and statistical significance of the variables.

This diagram displays the parameter estimate, importance level, and statistical significance of each predictor (Figure 4). Connecting lines in the diagram are colored based on the sign of the coefficient (orange colored shows negative coefficients while blue colored shows positive coefficients) and weighted based on coefficients significance, with greater line with corresponding to more significant coefficients.

RESULTS OF REGRESSION TREE ANALYSIS

Results of Regression Tree and ANOM Techniques are presented in Figure 5, Figure 6, Figure 7, Table 1 and Table 2, respectively. Figure 5 is determined as an optimal tree based on risk value, its standard error, and explained variation in the Type I error rate (Alpha). Explained variation percentage (R^2) of the optimal tree is found to be as 80.6% ($R^2 = 1 - \left(\frac{0.0014}{0.085^2}\right) = 1 - 0.194 = 0.806 = 80.6\%$), means that the optimal tree is a good one to explain the relations between the Type I error rates (dependent variable) and the experimental conditions (independent variables) (Figure 5).

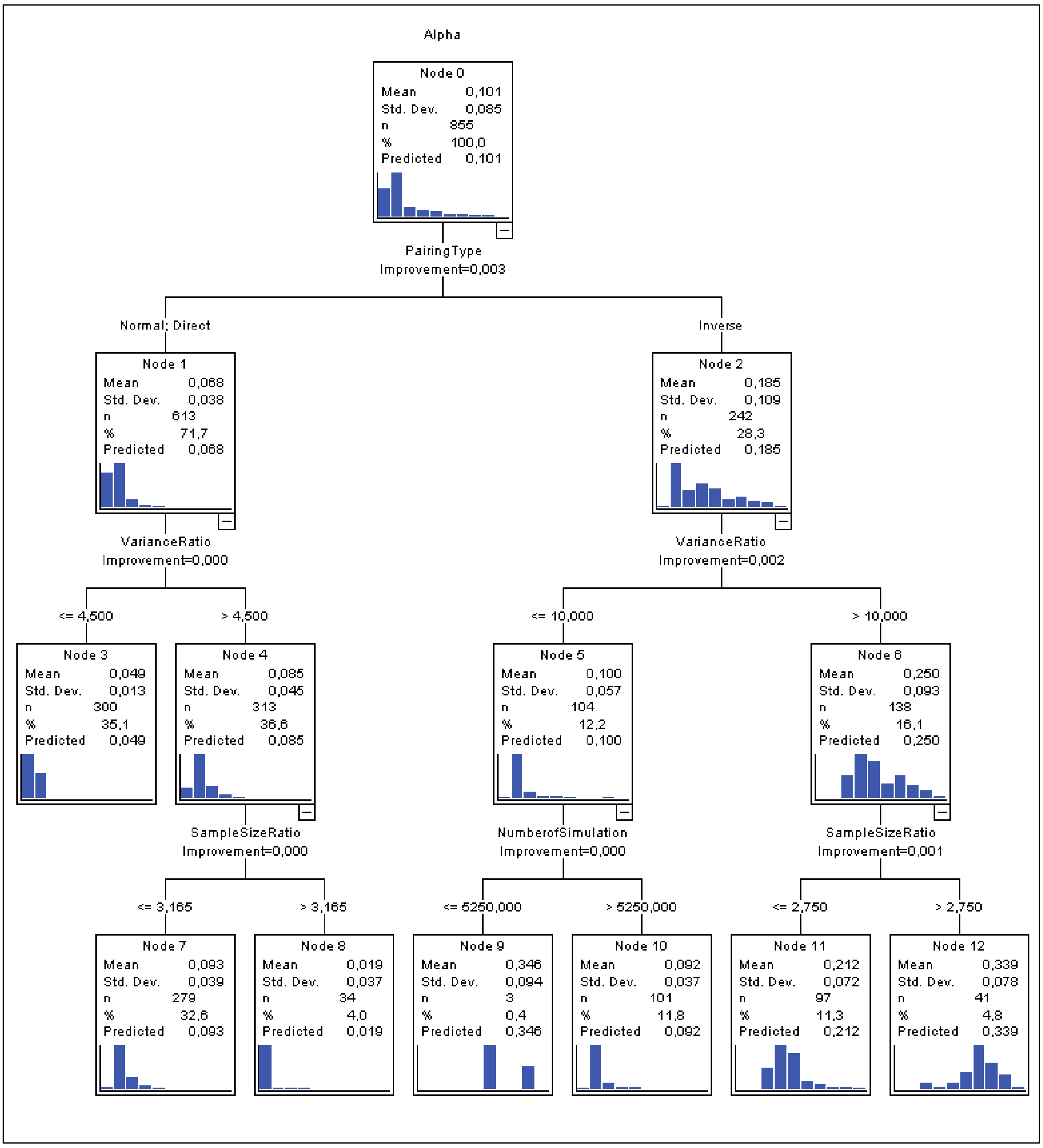

FIGURE 5: Optimal tree for predicting alpha estimates of simulation studies.

When optimal tree is examined, it is seen that this tree has been formed by using four factors namely PT, VR, SN, and SSR. In Figure 5, Node 0 is called the root node and it contains descriptive statistics related to Type I error estimates. Firstly, the effect of each independent variable in predicting the Type I error rate was evaluated separately by computing importance level of each factor. Since distribution shape (Sk and Kr values), TSS, ESS, and NG are not found to be effective in predicting the Type I error, these factors are not included to the optimal tree. As it is seen in Figure 5, firstly, the Type I error estimates in Node 0 or root Node were divided into two nodes as Node 1 (normal pairing and direct pairing) and Node 2 (inverse pairing) according to the PT. Therefore, it is understood that the factor that is the most important or effective in estimating the Type I error rate is **PT**. The mean of the Type I error rate of the simulations studies in Node 1 and Node 2 are predicted as 0.068 and 0.185, respectively. The proportions of the simulation studies in Node

1 and Node 2 in total are 71.7% and 28.3%, respectively. As it can be seen from Node 1 and Node 2, the Type I error average of the studies in Node 2 (0.185) are obviously higher than that of the Node 1 (0.038). It is not sufficient, however, to use only PT in estimating the Type I error rates of the simulation studies. In other words, the simulation studies in Node 1 and Node 2 are not homogeneous enough to get reliable and stable results. That is why, studies in Node 1 are again divided into two new nodes as Node 3 and Node 4, and the studies in Node 2 are divided into two new groups as Node 5 and Node 6 according to **VR's.** Therefore, the **VR** is accepted as the second most important or effective factor in estimating the Type I error rates of the studies. As can be seen from the Figure 5, the splitting of the optimal tree is still gone on according to **sample size** and **SN**. As a result, since the **PT** reflected the highest changes in the Type I error rates, it is determined to be the most important variable or factor, followed by **VR**, **SN, and SSR.** Therefore, among the 9 independent variables or factors, only 4 are selected. Using these 4 factors, 7 terminal nodes are formed and these nodes are accepted as homogenous groups.

TABLE 1: Descriptive statistics for the terminal nodes.

Nodes	n	%	Mean
9	3	0.4	0.346±0.001
12	41	4.8	0.339±0.002
11	97	11.3	0.212±0.006
7	279	32.6	0.093±0.054
10	101	11.8	0.092±0.004
3	300	35.1	0.049±0.007
8	34	4.0	0.019±0.012

TABLE 2: Importance values of independent variables.

Independent variables	Normalized importance
Pairing type	100%
Variance ratio	70.3%
Sample size ratio	67.9%
Number of simulation	34.5%
Total sample size	25.7%
Number of group	16.6%
Skewness	14.0%
Kurtosis	10.2%
Equality of sample size	3.3%

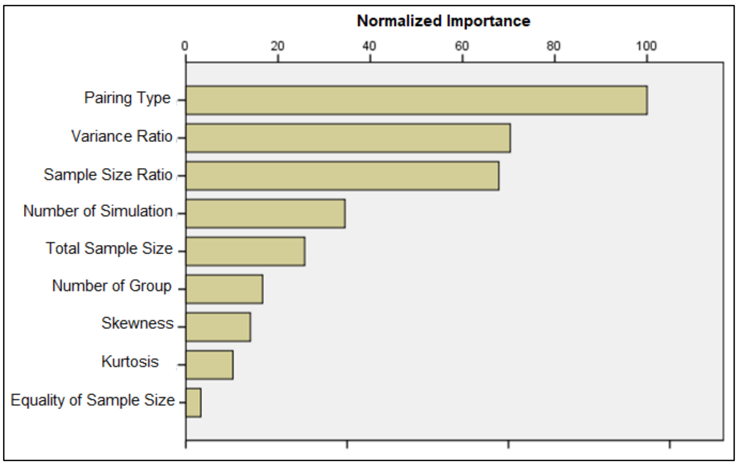

FIGURE 6: Graphical presentation of importance values of independent variables.

One-way ANOM technique was used to compare the terminal nodes or homogenous groups in terms of alpha means and to determine different nodes. ANOM is a powerful tool for comparing means, variances, proportions and other location and scale measures. ANOM is commonly used as a graphical alternative to the ANOVA for comparing independent group means.[42,50,51] Mendeş and Yiğit reported that this procedure can also be used efficiently as a multiple comparison test especially when there are large NGs.[51] The results of ANOM showed that there was a statistically significant difference among the terminal nodes (Figure 7). The most reliable results were obtained from the studies where direct or normal pairing was applied and VR's ≤4.50 (Node 3). Therefore, this result was an indicator that the studies where direct or normal pairing was applied and the VR was ≤4.50, the Type I error rate was maintained at 0.05 level (0.049). However, in studies with a VRs >4.5, it was also necessary to consider the SSR as well as the VR. As it will be noticed, the mean Type I error estimates is 0.093 for the studies where sample ratio <3.165, while the mean Type I error estimates is 0.019 for the studies with SSR >3.165. However, the Type I error estimates for both cases could not be maintained at 0.05 level. On the other hand, it is seen that the Type 1 error estimates of the studies with VR <10 where inverse pairing is applied is affected by the SNs and the Type 1 error estimates of the studies whose SNs is smaller than 5,250 shows a significant deviation from 0.05. These findings can be considered as an indication that the SNs has a significant effect on the reliability and stability of the Type I error estimates to be obtained. The Type I error estimates of the studies with a VR >10 is affected by the sample size. The most deviated results, on the other hand, had been observed in Node 9, Node 11, and Node 12. Therefore, it is possible to conclude that the Type I error estimates are negatively affected by inverse pairing application and this negative affect become more obvious especially when VR ≥10.

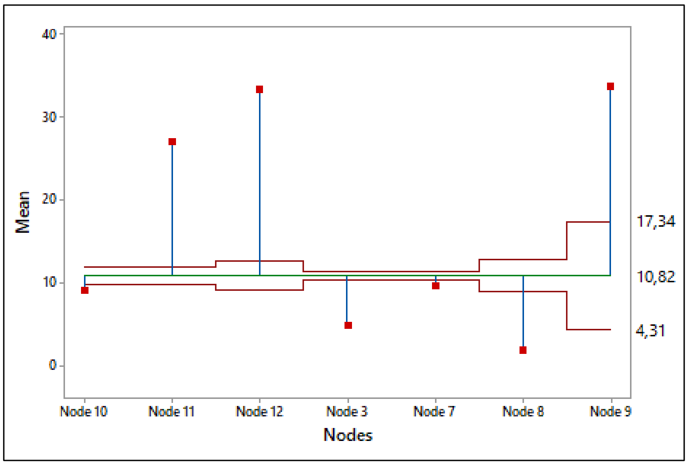

FIGURE 7: Results of one-way ANOM for comparing terminal nodes in terms of alpha estimates.

DISCUSSION

Especially in recent years, regression-based methods which can present the results graphically have become very popular.[4,52] The CART is probably one of the most well-known and commonly used decision tree learning algorithm in the literature.[2] In the popularization of these methods; it can be said that having some advantages over the classical methods (i.e easy understand and interpretation of results, not requiring assumptions like classical methods, providing to investigate complex and latent relationships between variables, providing information about higher order interactions, and making possible to classify individuals and variables) may affective. In this study, two regression-based methods, namely RTA and ALM were used in re-evaluating results of some Monte Carlo simulation studies and it was observed that both methods produced similar results in terms of determining factors affecting Type I rates.

As a result of both methods, the most important factors were determined as PT, VR, SN, and SSR. On the other hand, some differences were also observed between two techniques. For example, RTA produced higher R^2 value (80.6%) than ALM (70.4%) and thus, results of ALM and RTA analyses indicated that at least 70.4% of the variation (70.4% and 80.6%) in the Type I error rate could be explained by the four factors namely PT, VR, SN, and SSR.

Hayes et al. in their simulation study used CARTs and Random Forest to Analyze Attrition.[53] Based on their simulation studies, they reported that although many techniques (i.e., t-tests, logistic regressions) could be successfully used to evaluate the true selection model, pruned CART and random forest analysis appear to perform particularly well. It was also reported that pruned CART might be a strong choice under the various selection models, sample sizes, and amounts of incomplete data.

Gonzales et al. in their study used CART Technique to analyze Monte Carlo simulation studies and they reported that the CART approach would be useful, but it would be beneficial to test the approach in a Monte Carlo simulation that varied even more conditions to use the full potential of CART.[54] They also informed that the CART would be able to overcome the limitations of inferential approaches that use arbitrary effect size cutoffs in significance tests by using variable importance measures and the hierarchy of splits. As a result, they reported that it would be beneficial if the researchers encourage incorporating this tool in their analysis of Monte Carlo data for a better understanding of the complexity of their simulation studies. Similar results have been reached at the end of our study in general.

According to the results of the ALM and RTA, it is possible to conclude that:

a) Results of simulation studies are affected by different factors and these factors should be considered in order to get more reliable and stable estimates.

b) Using the graphical methods instead of classical techniques make possible to investigate the complex and latent relations between dependent and independent variables especially when we have a large and complex data set.

c) Using ALM and RTA enables us to interpret the results easily.

d) Using ALM and RTA enables us to evaluate higher order interactions among the independent variables.

e) The use of these methods allows the identification of similar and different studies by grouping the results of the simulation studies carried out under different trial conditions according to their similarities.

f) Decision trees are sometimes more interpretable than other classifiers such as logistic regression, discriminant analysis, cluster analysis, neural networks and support vector machines because they combine simple questions about the data in an understandable way.

CONCLUSION

As a result, although our results suggest that both ALM and RTA techniques would be useful in determining factors affecting Type I error rates, it shouldn’t be ignored that more reliable results may be obtained in which study that is done with more studies.

Source of Finance

During this study, no financial or spiritual support was received neither from any pharmaceutical company that has a direct connection with the research subject, nor from a company that provides or produces medical instruments and materials which may negatively affect the evaluation process of this study.

Conflict of Interest

No conflicts of interest between the authors and / or family members of the scientific and medical committee members or members of the potential conflicts of interest, counseling, expertise, working conditions, share holding and similar situations in any firm.

Authorship Contributions

This study is entirely author's own work and no other author contribution.

REFERENCES

1. Mendeş M, Akkartal E. Regression tree analysis for predicting slaughter weight in broilers. Italian Journal of Animal Science. 2009;8(4):615-24. [Crossref]
2. Breiman L, Friedman JH, Stone CJ, Olshen RA, Stone CJ. Classification and Regression Trees. 1st ed. Boca Raton, FL: Routledge; 1984.
3. IBM SPSS Inc. IBM SPSS Statistics 21 Algorithms. Chicago, 2012. (Download spss 21 full version 64 bit for free (Windows) (freedownloadmanager.org) [Link]
4. Berk RA. Statistical Learning from a Regression Perspective. 1st ed. New York, NY: Springer Science & Business; 2009. [Crossref] [PubMed]
5. James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning with Applications in R. New York, NY: Springer; First Edition, ISBN: 978-1-4614-7137-0 (Hardcover), 2013.
6. Morgan J. Classification and regression tree analysis. Technical Report No. 1. 2014. [Link]
7. Paxton P, Curran PJ, Bollen KA, Kirby J, Chen F. Monte Carlo Experiments: design and implementation. Structural Equation Modeling: A Multidisciplinary Journal. 2001;8(2):287-312. [Crossref]
8. Mendeş M. Normal dağılım ve varyansların homojenliği ön şartlarının gerçekleşmediği durumlarda varyans analizi tekniği yerine kullanılabilecek bazı parametrik alternatif testlerin I.tip hata ve testin gücü bakımından irdelenmesi. Ankara Üniv. Fen Bil. Enstitüsü, Doktora Tezi, Ankara. 2002. p.278.
9. Mendeş M, Başpınar E, Gürbüz F. Confidence interval for test power in Welch, James-second order and Alexander-Govern tests: a simulation study. YYÜ Fen Bilimleri Enstitüsü Dergisi. 2005;10(1):16-22. [Link]
10. Bandalos DL. The use of Monte Carlo studies in structural equation modeling research. In: Hancock GR, Mueller RO, eds. Structural Equation Modeling: A Second Course. Greenwich, Conn.: IAP; 2006. p.385-426. ISBN 1-59311-015-4 (hardcover) [Link]
11. Lee S. Implementing a simulation study using multiple software packages for structural equation modeling. SAGE Open. 2015;5(3):1-16. [Crossref]
12. Welch BL. On the comparison of several mean values: an alternative approach. Biometrika. 1951;38(3/4):330-6. [Crossref]
13. James GS. The comparison of several groups of observations when the ratios of the population variances are unknow. Biometrika. 1951;38(3/4):324-9. [Crossref]
14. Kohr RL, Games PA. Robustness of the analysis of variance, the Welch procedure and a box procedure to heterogeneous variances. J Exp Educ. 1974;43(1):61-9. [Crossref]
15. Brown MB, Forsythe AB. Robust tests for the equality of variances. J Am Stat Assoc. 1974;69(346):364-7. [Crossref]
16. Brown MB, Forsythe AB. The small sample behavior of some statistics which test the equality of several means. Technometrics. 1974;16(1):129-32. [Crossref]
17. Rogan JC, Keselman HJ. Is the ANOVA F-test robust to variance heterogeneity when sample sizes are equal? An investigation via a coefficient of variation. Am Educ Res J. 1977;14(4):493-8. [Crossref]
18. Bishop TA, Dudewicz EJ. Exact analysis of variance with unequal variances: test procedures and tables. Technometrics. 1978;20(4):419-30. [Crossref]
19. Ramsey PH. Exact Type I error rates for robustness of student's t test with unequal variances. J Educ Behav Stat. 1980;5(4):337-49. [Crossref]
20. Olejnik S. Conditional ANOVA for mean differences when population variances are unknown. J Exp Educ. 1987;55(3):141-8. [Crossref]
21. Rasmussen JL. An evaluation of parametric and non-parametric tests on modified and non-modified data. Br J Math Stat Psychol. 1986;39(2):213-20. [Crossref]
22. Tomarkin AJ, Serlin RC. Comparison of ANOVA alternatives under variance heterogeneity and specific noncentrality structures. Psychologial Bulletin. 1986;99(1):90-9. [Crossref]
23. Krutchkoff RG. One-way fixed effects analysis of variance when the error varianes may be unequal. J Statist Comput Simul. 1988;30(4):259-71. [Crossref]
24. Randolph E, Robey R, Barcikowski R. Type I error of the ANOVA revisited using power analysis criteria. Paper presented at the Annual Meeting of the American Educational Research Association. Boston; 1990. Available from: [Link]
25. Oshima TC, Algina J. Type I error rates James's second-order test and Wilcox's Hm test under heteroscedasticity and non-normality. Br J Math Stat Psychol. 1992;45(2):255-63. [Crossref]
26. Robey RR, Barcikowski RS. Type I error and the number of iterations in Monte Carlo studies of robustness. Br J Math Stat Psychol. 1992;45(2):283-8. [Crossref]
27. Alexander RA, Govern DM. A new and simpler approximation for ANOVA under variance heterogeneity. J Educ Behav Stat. 1994;19(2):91-101. [Crossref]
28. Hsiung TH, Olejnik S. Type I error rates and statistical power for James second-order test and the univariate F test in two-way fixed-effects ANOVA models under heteroscedasticity and/or nonnormality. J Exp Educ. 1996;65(1):57-71. [Crossref]
29. Schneider PJ, Penfield DA. Alexander and Govern's approximation: providing an alternative to ANOVA under variance heterogeneity. J Exp Educ. 1977;65(3):271-86. [Crossref]
30. Murphy KR, Myors B. Statistical Power Analysis: A Simple and General Model for Traditional and Modern Hypothesis Tests. London: Lawrence Erlbaum; First Edition; 1998.
31. Myers L. Comparability of the James' second-order approximation test and the alexander and govern A statistic for non-normal heteroscedastic data. J Stat Comput Simul. 1998;60(3):207-22. [Crossref]
32. Wilcox RR, Charlin VL, Thompson KL. New Monte Carlo results on the robustness of the ANOVA F, W and F* statistics. Communications in Statistics-Simulation and Computation. 1986;15(4):933-43. [Crossref]

33. Wilcox RR. A new alternative to the ANOVA F and new results on James's second-order method. Br J Math Stat Psychol. 1988;41(1):109-17. [Crossref]
34. Wilcox RR. Adjusting for unequal variances when comparing means in one-way and two-way effects ANOVA models. J Educ Behav Stat. 1989;14(3):269-78. [Crossref]
35. Wilcox RR. A one-way random effects model for trimmed means. Psychometrika. 1994;59:289-306. [Crossref]
36. Wilcox RR. ANOVA: A paradigm for low power and misleading measures of effect size? Review of Educational Research. 1995;65(1):51-77. [Crossref]
37. Wilcox RR. Comparing variances of two independent groups. Br J Math Stat Psychol. 2002;55(1):169-75. [Crossref] [PubMed]
38. Moder K. How to keep the type I error rate in ANOVA if variances are heteroscedastic. Austrian Journal of Statistics. 2007;36(3):179-88. [Link]
39. Mendeş M, Akkartal E. Comparison of ANOVA-F and WELCH tests with their respective permutation versions in terms of Type I error rates and test power. Kafkas Üniversitesi Veteriner Fakültesi Dergisi. 2010;16(5):711-6. [Link]
40. An Q. A Monte Carlo Study of Several Alpha-Adjustment Procedures Used in Testing Multiple Hypotheses in Factorial Anova [Thesis]. Ohio: Ohio University; 2010. [Link]
41. Zahayu MY, Suhaida A, Sharipah SSY, Othman AR. A robust alternative to the t-test. Modern Applied Science. 2012;6(5):27-33. [Crossref]
42. Mendeş M, Yiğit S. Comparison of ANOVA-F and ANOM tests with regard to type I error rate and test power. Journal of Statistical Computation and Simulation. 2013;83(11):2093-104. [Crossref]
43. LinkedIn Learning. How to use automatic linear modeling | Lynda.com tutorial. Release date: 22.8.2011. Available from:(Access date: 28.8.2019) [Link]
44. Yang H. The case for being automatic: introducing the automatic linear modeling (LINEAR) procedure in SPSS statistics. Multiple Linear Regression Viewpoints. 2013;39(2):27-37. [Link]
45. Liaw A, Wiener M. Classification and regression by randomForest. R News. 2002;2(3):18-22. [Link]
46. Merkle EC, Shaffer VA. Binary recursive partitioning: background, methods, and application to psychology. Br J Math Stat Psychol. 2011;64(Pt 1):161-81. [Crossref] [PubMed]
47. Gordon L. Using Classification and Regression Trees (CART) in SAS® Enterprise MinerTM for applications in public health. SAS Global Forum 2013. [Link]
48. Therneau T, Atkinson B, Ripley B. rpart: Recursive partitioning and regression trees. R package version 4.1-15. 2019. Retrieved from [Link]
49. Yang L, Liu S, Tsoka S, Papageorgiou LG. A regression tree approach using mathematical programming. Expert System with Application. 2017;78:347-57. [Crossref]
50. Wludyka PS, Nelson PR. An analysis-of-means-type test for variances from normal populations. Technometrics. 1997;39(3):274-85. [Crossref]
51. Mendeş M, Yiğit S. An alternative approach for multiple comparison problems when there are a large number of groups: ANOM technique. Journal of Animal and Plant Sciences. 2018;28(4):1074-9.
52. Mendeş M. İstatistiksel Yöntemler ve Deneme Planlanması. 1. Baskı. İstanbul: Kriter Yayın Evi; 2019.
53. Hayes T, Usami S, Jacobucci R, McArdle JJ. Using Classification and Regression Trees (CART) and random forests to analyze attrition: Results from two simulations. Psychol Aging. 2015;30(4):911-29. [Crossref] [PubMed] [PMC]
54. Gonzales O, O'Rourke HP, Wurpts IC, Grimm KJ. Analyzing Monte Carlo simulation studies with classification and regression trees. Structural Equation Modeling A Multidisciplinary Journal. 2018;25(3):403-13. [Crossref]