The Examining of Generalization Quantitative Scientific Findings by Using the Jackknife Method: An Application

Murat KAYRİ, Şener BÜYÜKÖZTÜRK***

Abstract

The outcomes which cannot be generalized are specific for a sample but are unable to be reflected to the rest of the population. The parameters that are reached at the end of the statistics that are scarce in sample arise doubts in the aspect of generalization. In these cases, parameter estimation may not be very stable and outlier values can produce false outcomes in the model. The situation that there is sample size is not adequate enough in numbers, reliability and generalizations that are reached at the end of restricted sets of data necessitate questioning with a dubious approach. Some statistical methods test the confidence intervals of the parameter values of the samples and the generalization of the parameter estimation values of the samples. These methods are known as "resampling" or "domestic-self copying." Jackknife is one of these methods. The sample in this study is composed of 18 students. The status of the "self-respect" of individuals in making decisions has been tried to be determined for the sample by the application of the Melbourne Decision-making Scale I. In this research, the dependent variable was determined to be the total item score related to the self-respect manners of individuals whereas independent variables were determined to be the academic success averages of the students (acdave), the financial income of their families (economy) and the number of siblings (sibling) they have. It has been seen that the academic success independent variable has a considerable effect at a significance level of .05 on the dependent variable of self-respect dependence of decision making ($p < .05$) and Jackknife has confirmed this generalization.

Key Words

Jackknife Parameter Estimator, Resampling, Generalization.

* *Correspondence:* Asist. Prof., Murat Kayri, Yuzuncu Yil University, Education Faculty, Van / Turkey
E-mail: mkayri@yyu.edu.tr
**: Şener Büyüköztürk, Assoc.Prof., Baskent University, Education Faculty, Ankara / Turkey

Kuram ve Uygulamada Eğitim Bilimleri / Educational Sciences: Theory & Practice
9 (4) • Autumn 2009 • 1769-1780

Decision-making process is described as the cognitive responses that individuals develop for different situations that they face. Decision-making is described as a process in which there are more than one options in a given situation or the process which people may face in responding the problems throughout their social lives (Çakır, 2004; Deniz, 2004; Eldeleklioğlu, 1996; Sardoğan, Karahan, & Kaygusuz, 2006). It is already known that decision-making styles differ from person to person. Pieces of writings on the field indicate that people may experience stress and anxiety during the decision-making process (Çakır, 2004; Kuzgun, 1992; Mann at al., 1998; Yı, & Park, 2003). It is declared that "value judgment," "purposes." and "standards" affect decision-making (Yazıcıoğlu & Erdoğan, 2002). Yazıcıoğlu and Erdoğan (2002) exemplified some direct effects on decision-making such as academic achievement, love, respect, trust, freedom, health, wealth, knowledge, ability, and religion.

The principle that the outcomes of the results of the scientific studies is regarded to be extremely important as one of the features of scientific features (Borg & Gall, 1983). The outcomes of the scientific studies that are the results of the experimental efforts shed light on the future studies of the relevant field thus contribute to the development of the field under consideration. Moreover, the situation in which the outcomes of the scientific studies are unbiased and reliable would strengthen the theoretical framework of the field and would preserve the objectivity and reliability of the long term use of the relevant outcomes. The outcomes which cannot be generalized are sample-specific but are unable to be reflected to the rest of the population (Thompson, 1992). It has been observed that some of the samples that are investigated in the scientific studies are scarce in number. The parameters that are reached at the end of the statistics that are scarce in sample arise doubts in the aspect of generalization.

The relationship between dependent and independent variables is determined according to the beta coefficient that is reached at the result of the regression equation. The probability level (p) that belongs to beta coefficient shows whether the relationship is statistically meaningful or not. Determination of independent variables that affect dependent variables is extremely important in especially meta analysis and conceptual studies. The situation which shows that the independent variables that can affect the dependent variables would determine the replication and

the generalization of the studies. As it is known, the relationship between the dependent and the independent variables can be studied and beta coefficients of the independent variable can be determined by the regression method that has been applied to any given data set. However, convincing expressions concerning scientific generalizations that are regarded to be meaningful independent variables in terms of statistics based on the regression balance cannot be reached. The amount of the effect of the independent variables ($\beta_1, \beta_2, \ldots, \beta_n$) and determining coefficients (R^2) of the model can be reached by the simple linear, multiple or multivariate regression analysis. The situation which shows that coefficients that are meaningful in the model can produce approximate results to similar real studies; in other words, its generalization is expected from scientific studies. The choice of the sample size that can represent the universe seems to be extremely important especially in the studies based on observation. However, in studies which have some constraints it can be hard to acquire samples that represent the universe in numbers. In these cases, parameter estimation may not be very stable and outlier values can produce false outcomes in the model. The situation in which the sample is not adequate enough in numbers, reliability and generalizations that are reached at the end of restricted sets of data necessitate questioning with a dubious approach.

Some statistical techniques that test the confidence intervals of the parameter values of the samples and the generalization of the parameter estimation values of the samples that are mentioned above are used. In the situations that the sample size is scarce in numbers, the generalization of the outcomes in hand are reached through jackknife, double-cross validation, and bootstrap methods (Avşar, 2006; Bekiroğlu, Konyalıoğlu, & Aydın, 2007). These methods are known as "resampling" or "domestic-self copying" techniques. Resampling methods embody different features from each other. Among these, jackknife and bootstrap methods are used especially in investigations that are small in size (Efron, & Tibshirani, 1998; Miller, 1974). The Jackknife and bootstrap methods that are widely used in the literature shows great similarities while they differ in the calculation of the standard error estimation (Avşar, 2006; Efron, 1982; Jun & Dongsheing, 1995). The Jackknife method takes all of the data into consideration while it eliminates the bias that results from the situation that some incidents apart from the normal find place in the example in a great amount. The facts that the

Jackknife approach is the method that gives the least deviation in estimating the features of the population and that it can make parameter estimation in an extremely conservative cycle are constantly indicated in the literature (Bekiroğlu et al., 2007; Efron, 1985; Wu, 1986).

Method

Sample

The main aim of this study is to identify the Jackknife methods with using the Melbourne Decision Scale. The generalization of scientific findings is so crucial when it is an obligation to study with a small sample size. In this study, relational survey (cause-effect based) research design was used. On the other hand, the study conducted according to correlational search design.

The sample in this study is composed of Department of Computer and Instructional Technology Education students. The total number of students is 18 in the sample. The average age of the students is 20.72 years and the standard deviation of the age distribution is calculated to be 1.41 years. Seven of the students are females and 11 are males in the studied sample. The status of the "self-respect" of individuals in making decisions has been tried to be determined for sample by the administration of the Melbourne Decision Making Scale I.

Research Instruments

The Melbourne Decision-making Scale: The Melbourne Decision Making Scale is composed of two sub-scales which are named as MDMS I and II. In this study, MDMS I sub-scale was used. MDMS I is composed of six items and three of them are graded in reverse. Scoring is as follows: Correct answer to the items recevies 2 points. Sometime correct answer is 1 point, Not Correct answer is 0 points. The maximum score that can be achieved from the scale is 12. Higher grades mean higher self-respect in making-decisions.

Validity and Reliability: *The validity and reliability of research instruments.* Whereas the Cronbach alpha coefficient obtained from the original form of the scale is 0.74, the Cronbach alpha value obtained from this study related to the scale was found to be 0.749, which is very close to the original value.

Procedure

The Melbourne Decision Making Scale I has been administered to 18 students. In this research, the dependent variable was determined to be the total item score related to the self-respect, whereas independent variables were determined to be the academic success averages (acdave), the financial income of families (economy), and the number of siblings (sibling). In a model similar to Y = f (x), the regression equation, Self-respect point = intercept (β0) + β1 * acdave + β2 * economy + β3 * sibling is tried to be estimated. After making parameter estimation in the multiple regression analysis procedure by using "Least Squares Method," the generalization of the estimator variables for which the significant level is found to be expressive at a level of 0.05 was tested. Because the sample size was small, the generalization problematic of the findings of this study to the population will be investigated by the Jackknife method.

The Jackknife Method: The Jackknife method is a method that approximates the bias and confidence intervals in guesses for conditions for which they are hard to find by known methods. The Jackknife method was developed to calculate the deviation of the estimator and create strong confidence intervals in order to minimize the sample error (Topuz, 2002). The Jackknife method is considered to be a statistical procedure that aims to put forth the relationship between the variables in the data set (Swingler, 1995) for many areas (mathematics, engineering, health and social sciences etc.) which require parameter estimation. Jackknife is known to be a non-parametric statistical method that does not take into account the assumptions of the distribution that it is studying (Şahinler, & Topuz, 2007). In the parameter estimation procedure, Jackknife, throws out one person at a time from the sample and thus can protect itself from outlier values. Whereas in some studies performed, Jackknife is said to be an unbiased method as a parameter estimator for small samples (Bekiroğlu et al., 2007; Efron, 1982; Quennouilli, 1949, 1956; Tukey, 1958). In some studies, it is stated that it cannot be as strong when compared with other sampling methods (especially Bootstrap) in cases where the sample quantity is low (Fan, & Wang, 1995; Şahinler, & Topuz, 2007). Based on the structure of the variables in the study pattern, it should also be kept in mind that re-sampling methods (such as Bootstrap, Jackknife) may make estimations with deviations. It should not be forgotten that the only way to reach stability in the

estimations is to increase sample size. The fundamental of the Jackknife method is producing *n* different samples which are (n-1) large each by throwing a sample from the sample set one by one. The parameter in question, O', is estimated by a series of estimations belonging to these sub samplings such as o'1, o'2, …, in the following manner;

$$Ji\ (O') = n\ O' - (n-1)o'i, \quad i=1, \ldots, n \qquad (1)$$

The values in Equation 1 are called pseudo values. Here O' is an estimator obtained by *n* observations, which is the whole data set. By keeping out one individual at a time out of the sample, parameter estimations are made for "*n*" sub-samples. After making these calculations for each individual, the final estimation value of the Jackknife estimator is found by taking the average of each estimation obtained for all sub-samples. In this sense, the average pseudo value of the Jackknife estimator can be estimated in this way;

$$J(O') = [\Sigma\ Ji\ (O')] / n, \quad i=1, \ldots, n \qquad (2)$$

As is stated above, Jackknife does not directly act on the parameter estimation value it obtains, but the stability of the estimator is tested by confidence intervals. To calculate the confidence interval for the Jackknife estimator at the level of %100(1-2α), the equation;

$$\theta_{(j)} - t_{n-1,\,\alpha}\ S_J < \theta < \theta_{(j)} + t_{n-1,\,\alpha}\ S_J \qquad (3)$$

is used. Here;

$t_{n-1,\,\alpha}$: The t table value at the specified α level and n-1 degree of freedom level,

$\theta_{(j)}$: Jackknife estimator,

S_J : The standard error estimation for Jackknife.

First, the multiple regression analysis of the study is computed by "Least Squares Method (LSM)" and the β coefficient obtained for each variable is recorded. Also, it is possible to check if the model is significant or not by variance analysis (Analysis of Variance; ANOVA). Similarly, it should also be checked with multiple regression analysis at what ratio (R^2) the independent variables in the model could explain the dependent variable. All of these procedures are performed independent of the Jackknife procedure and create an infrastructure for comparing the Jackknife estimator values in the future. In the regression analysis, Jackknife estimator value will be obtained for variables that are significant in

the model. The related coefficients (β) that will be obtained by the Jackknife estimator for the related variables will be obtained by equations 1 and 2. After that, the confidence interval of the Jackknife estimator is obtained by using equation 3. If the Jackknife estimator calculated previously (Ơ or θ) is within the intervals of the confidence interval calculated, the variable can be said to have "generalization" property. Otherwise, it will be accepted that the variables that seem meaningful after the regression analysis cannot be generalized to the population and are specific only for this sample.

Discussion

Whereas the "self-respect" attitudes of individuals may be effective on their decision-making styles, the self-respect levels of the individuals will also directly affect their ability to make correct decisions. It is known that in the decision-making process of an individual there are lots of direct and indirect factors involved. Their "value judgments," "purpose," and "standards" are described in literature to be directly affecting the process (Yazıcıoğlu & Erdoğan, 2002). Moreover, factors such as success, respect, love, health, and information are thought to shape the decision-making process. Whereas it is known that the process of decision-making sometimes causes stress and distress in individuals (Çakır, 2004; Kuzgun, 1992; Mann et al., 1998; Yı, & Park, 2003), it is thought that the "self-respect" attitude of the individual will have an effect on this stress and distress.

The generalization to the population of the findings of empiric studies is a topic of importance in scientific studies (Borg & Gall, 1983). Especially, in cause-effect based studies, it is expected that independent variables that are found to be effective on the dependent variable will also give similar results in further studies. Findings that cannot be generalized are sample-specific and cannot be reflected on the population (Thompson, 1992). In empiric studies, it is very important to select variables that may be suitable for the topic and also that are effective, the selection of a stable and unbiased statistical method and to reach a satisfactory, homogeneous sample number. Among these, it is known that the sampling theory is effective on unbiased and stable findings (Arıcı, 2006; Büyüköztürk et al., 2008; Fraenkal, & Wallen, 2006; Karasar, 2005). However, it can be seen that in scientific studies some of the sample sizes are small. The parameters obtained from a small number of samples arise suspicion about generalization.

One of the statistical methods that test the generalization to the population of the parameters in the regression equation obtained from a small number of samples is the Jackknife parameter estimation method. The Jackknife method can stably test the generalization of the effective parameters for researches having a small number of samples (Avşar, 2006; Bekiroğlu et al., 2007; Efron, 1985; Efron, & Tibshirani, 1998; Jun, & Dongsheing, 1995; Wu, 1986). Jackknife creates new samples by taking out an individual from the sample each time and by estimating parameters for each sample thus created, tries to eliminate the effect of the outlier values in the sample. With this property, the Jackknife method is considered to be an unbiased estimator (Bekiroğlu et al., 2007).

In this study, the regression equation, self respect level = $\beta 0$ + $\beta 1$*academic success + $\beta 2$*economical income + $\beta 3$*number of siblings, is modeled. The variables that might have a considerable effect on the dependent variable have been investigated. In the multiple-regression analysis performed, it has been seen that the academic success independent variable has a considerable effect at a significance level of .05 on the dependent variable of self-respect dependence of decision making ($p < .05$). Similarly, it has been observed that the economical income of the individuals also has an effect but that this effect is not at the same .05 level ($p = .07$). It has been concluded that the independent variable, the number of siblings, has no effect on the dependent variable ($p > .05$). It has also been verified by the Jackknife method that the academic success of the individuals has considerable effect on the dependent variable of "self-respect" in decision making ($p < .05$). Besides, it has been observed that the parameter value obtained by the Jackknife method is within the specified confidence interval (see Table 7). Thus, it has been reached the conclusion that the academic success independent variable is not specific to the sample and that it can be generalized to the population. The independent variable of economical income, which was found to be critical at a p value of ($p = .07$) was found to be not-effective by the Jackknife method at a significance level of 0.05. In this manner, the Jackknife method has shown stability, did not find the effect of economical income considerable and did not generalize this variable to the population. In this study, the self-respect attitude of the individuals is explained by the independent variable of academic success. It has been observed that the number of siblings present in the model has no effect on the "self-respect" attitude level in decision-making.

In conclusion, it has been observed that the Jackknife method which has seen popularity in literature is able to estimate parameters in an unbiased manner for small number of samples. It is especially suggested that researchers investigate the generalization of their findings with the Jackknife parameter estimator for experimental methods where, for due to some limitations, the number of samples is small. It is thought that in this way, the findings will pave the way in a robust and unbiased manner for future studies.

References/Kaynakça

Arıcı, H. (2006). *İstatistik - yöntem ve uygulamalar*. Ankara: Meteksan A.Ş. Baskı Tesisleri.

Avşar, P. E. (2006). *Bootstrap yönteminin regresyon analizinde kullanımına ilişkin olarak Türkiye inşaat sektöründe bir uygulama*. Yayımlanmamış yüksek lisans tezi, Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.

Bekiroğlu, N., Konyalıoğlu, R. ve Aydın, D. (2007, Eylül). *Jackknife tekniği ile bir uygulama*. X. Ulusal Biyoistatistik Kongresi'nde sunulan sözlü bildiri, Cumhuriyet Üniversitesi, Tıp Fakültesi, Sivas.

Borg, W. R., & Gall, D. (1983). *Educational research.* Newyork, NY: Longman.

Büyüköztürk, Ş., Çakmak, E., Akgün, Ö. E., Karadeniz, Ş. ve Demirel, F. (2008). *Bilimsel araştırma yöntemleri*. Ankara: Pegem Akademi Yayıncılık.

Çakır, M. A. (2004). B*ir mesleki grup rehberliği programının lise öğrencilerinin mesleki kararsızlık düzeylerine etkisi*. Yayımlanmamış doktora tezi, Ankara Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.

Deniz, M. E. (2004).Üniversite öğrencilerinin karar vermede öz saygı karar verme stilleri ve problem çözme becerileri arasındaki ilişkinin incelenmesi üzerine bir araştırma. *Eğitim Araştırmaları Dergisi*, *4*(15), 23-35.

Efron, B. (1982). *The jacknife the bootstrap and other resampling plans*. Regional Conference Series in Applied Mathematics, Philadelphia, PA.

Efron, B. (1985). Bootstrap confidence interval for a class of parametric problems. *Biometrica*, *72*, 45-58.

Efron, B., & Tibshirani, R. J. (1998). *An introduction to the bootstrap.* New York, NY: Chapman & Hall.

Eldeleklioğlu, J. (1996). *Karar stratejileri ile ana baba tutumları arasındaki ilişki*. Yayımlanmamış doktora tezi, Gazi Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.

Fan, X., & Wang, L. (1995, April). *How comparable are the jackknife and bootstrap results: An investigation for a case of canonical correlation analysis.* The Annual Meeting of the American Educational Research Association, San Francisco, CA.

Fraenkal, J. R., & Wallen, N. E. (2006). *How to design and evaluate research in education* (6th Edt.). New York: McGraw-Hill International Edition.

Jun, S., & Dongsheing, T. (1995). *The jacknife and bootstrap.* Newyork, NY: Spinger-Veriag.

Karasar, N. (2003). *Bilimsel araştırma yöntemi*. Ankara: Nobel Yayınevi.

Kuzgun, Y. (1992, Eylül). *Karar Stratejileri Ölçeği: Geliştirilmesi ve standardizasyonu*. VII. Ulusal Psikoloji Kongresi Bilimsel Çalışmaları'nda sunulan bildiri, Türk Psikologlar Derneği, Ankara.

Mann, L., Radford, M., Burnett P., Ford, S., Bond, M., Leung, K. et al. (1998). Cross-cultural differences in self-reported decision-making style and confidence. *International Journal of Psycholog, 33*(5), 325-335.

Miller, R. G. (1974). The jackknife - a review. *Biometrika*, *61*, 1-15.

Quenouille, M. H. (1949). Approximate tests of correlation in time series. *Journal of the Royal Statistical Society, Series B*, *11*, 18-44.

Quenouille, M. H. (1956). Notes on bias in estimation. *Biometrika*, *61*, 353-360.

Şahinler, S., & Topuz, D. (2007). Bootstrap and jackknife resampling algorithms for estimation of regression parameters. *Journal of Applied Quantitative Methods*, *2*(2), 188-199.

Sardoğan, M. E., Karahan, T. F. ve Kaygusuz, C. (2006). Üniversite öğrencilerinin kullandıkları kararsızlık stratejilerinin problem çözme becerisi, cinsiyet, sınıf düzeyi ve fakülte türüne göre incelenmesi. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, *2*(1), 78-97.

Swingler, D. N. (1995). On the application of the jackknife to the estimation of the parameters of short multisinusoidal signals using the data-matrix formulation. *Ieee Transactions on Signal Processing*, *43*(9), 2135-2143.

Thompson, S. K. (1992). *Sampling.* Newyork, NY: John Willey & Sons, Inc.

Topuz, D. (2002). *Regresyonda yeniden örnekleme yöntemlerinin karşılaştırmalı olarak incelenmesi*. Yayımlanmamış yüksek lisans tezi, Niğde Üniversitesi, Fen Bilimleri Enstitüsü, Niğde.

Tukey, J. W. (1958). Bias and confidence in not-quite large sample. *Ann Math Stat*, *29*, 614.

Wu, C. F. J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis. *The Annals of Statistics*, *14*(4), 1261-1295.

Yı, J. S., & Park, S. (2003). Cross-cultural differences in decision-making styles: A study of college students in five countries. *Social Behavior and Personality*, *31*(1), 35-48.

Yazıcıoğlu, Y. ve Erdoğan, S. (2002). *Aile ekonomisi* (Ed. L. Kılıç). Eskişehir: Anadolu Üniversitesi Yayını.