



Determination of Outliers in Growing Quail's Data with Different Sample Size

Ufuk KARADAVUT^{1*}, Atilla TAŞKIN¹

¹Ahi Evran University, Agricultural Faculty, Animal Science Department, Kirsehir, TURKEY

*Corresponding author: ukaradavut@ahievran.edu.tr

Abstract

The aim of this study was to use as an alternative to MSS estimators M of Robust Regression estimators method is to examine the outlier in Japanese quail body weight data. During 15 weeks in the study, body weight measurements of 150 Japanese quails were recorded weekly. To determine the effect of outliers, quails were randomly divided into three groups and 10, 20 and 30 samplings were performed from each group, respectively. To conclude, it was concluded that the estimator M of outliers on the results of estimation methods can be used with success in this regard. Also, the number of samples increases that marred the outliers was identified and therefore they cannot emerge.

Key Words: M estimator, Outlier, Quail, Regression, Richards function

Bıldırcın Büyüme Verilerinde Farklı Örnek Büyüklüklerinin Aykırı Değerlerinin Belirlenmesi

Öz

Bu çalışmanın amacı, Japon bıldırcını vücut ağırlığı verilerinde aykırı değerleri incelemektir. Robust Regresyon tahmin yöntemi olan M tahmin yönteminin MSS alternatifi olarak kullanılmaktadır. Çalışmada 15 hafta boyunca, 150 Japon bıldırcınının vücut ağırlığı ölçümleri haftalık olarak kaydedilmiştir. Aykırı değer etkisini belirlemek için, bıldırcınlar rastgele üç gruba ayrılmış ve her bir gruptan 10, 20 ve 30 örnekleme yapılmıştır. Sonuç olarak, bu tahmin yöntemlerinin sonuçlarına göre M aykırı değer tahmin edicisinin bu konuda başarı ile kullanılabilceği sonucuna varılmıştır.

Anahtar Kelimeler: M tahmincisi, Aykırı, Bıldırcın, Regresyon, Richards fonksiyonu

Introduction

Among the studies focusing on improving the ability to satisfy the human demand for food are a number of researches into the breeding of quail as an alternative food source (Testik et al., 1993). The main motivation behind these studies is the possibility that the acquired valuable data can be applied to other poultry (Kocak and Ozkan, 2000). The Japanese quail in particular is considered to be an important potential source owing to the short time period between generations, its low feed consumption compared to other poultry and the ability to obtain rapid results in breeding studies (Gurcan et al., 2012).

Animal development is highly responsive to changes in environmental factors (Balcioglu et al., 2005; Narinc et al., 2010) and so there are a number of modelling techniques that are used to predict the effects that changes in environmental factors will have on production (Coelho and Dale, 1980). The main purpose behind the formulation and use of growth and development models is to determine accurately the growth and development conditions of the animals; to obtain results within a shorter time-frame, especially compared to time-consuming breeding studies; and to determine the priorities that need to be considered during the growth,

development and differentiation stages (Minvielle, 2004; Willemsen et al., 2008).

Outliers represent one of the main obstacles to developing and increasing the accuracy of models (Minvielle, 2004), and determining outlier values is very important for the proper evaluation of data, and consequently, sound and appropriate decisions. Most studies tend to disregard normality tests, acting on the assumption that the normality assumption is met (Quackenbush, 2002), which actually represents a very fundamental error (Bek and Efe, 1987; Akdeniz, 1998). In scientific studies, it is not uncommon during the evaluation of obtained numerical values for one or a number of observed values to be quite distant/different from the others, and such values have been given a number of different names, such as extreme values, discordant values, suspicious observed values, surprise values, dirty data, contaminants and outliers (Jain, 2005). These values may stem from natural randomness, human or mechanical error, or other similar reasons (Tserveni-Gousi, 1987).

The existence of a single Outlier (OL) within a sample can interfere with the information provided by the other data, and render all of the statistical results unreliable (Cook, 1977). The disruption of Least Squares (LS) predictions due to the presence of OLs can also lead to significant problems. When

using the LS method during a regression analysis, the method places equal weight on the evaluated data, and consequently, errors caused by OLs reduce the sum of the squares (Quackenbush, 2002). To overcome this, Robust Regression (RR) estimators are used as an alternative to the LS method (Rousseeuw and Yohai, 1984).

Previous studies have evaluated the effect – as well as the approaches for resolving – a single outlier within study data (Davies and Gather, 1993; Hadi and Simonoff, 1993); however, current methods are inadequate for addressing cases in which the data contains multiple OLs. OLs these are fairly close to one another can sometimes mask each other, preventing one or several of them from being identified; while in other cases, OLs can cause reliable data to appear as other OLs due to the sweeping effect. To prevent such undesirable effects, a number of numerical methods and algorithms have been developed for identifying OLs (Satman, 2005).

The aim of this study is to identify outliers within the time-dependent live weight data of quail, and to assess the effectiveness of the M-estimator method of identifying outlier values.

Materials and Methods

Animal materials

The quails used in the study were obtained from 250 hatching eggs, collected from a parent flock of 20 week old Japanese quail (*Coturnix coturnix Japonica*). Prior to incubation, the eggs were stored temporarily at the Animal Physiology Laboratory within an egg chamber maintained at 75-80 % relative humidity and a temperature of 14-16 °C. The eggs were then incubated for 15 days in an incubation machine at 55 % relative humidity and at a temperature of 37.5 °C. The eggs were removed from the incubator on the 15th day and taken to a hatching machine, where they were kept at 75 % relative humidity and a temperature of 37.2 °C.

After hatching, 150 chicks of mixed gender were selected randomly for the study, and divided randomly into three groups of 50, placed in three separate temperature-controlled cages within the separate unit for quails. Each one of the cages measured 50 cm x 90 cm, providing 90 cm² per quail chick. The temperature was set initially at 34 °C, and gradually reduced by 2 °C each week until it was brought to and kept stable at room temperature (22±2 °C and 50-60 % relative humidity). The gender of the chicks was identified after the third week based on their chest feathers, and the

ratio of males and females in each cage was determined accordingly. It was observed that there were no significant differences between the groups with respect to the ratio of males and females.

The chicks were given feed containing 24 % HP and 2,900 kcal ME kg⁻¹ for the first three weeks; 20 % HP and 2,800 kcal ME kg⁻¹ for the following three weeks; and 17 % HP and 2,800 kcal/kg ME after the sixth week until the end of the study. The quails were given food and water *ad libitum* throughout the study period, provided through nipple waterers within the cages. The lighting program ensured light intensity was maintained at 15 lux, illuminating the cages 24 hours/day in the first week, with the lighting period being gradually and sequentially decreased to 16 hours/day starting from the second week. The temperature, humidity and light intensity values were recorded in real time using a data logger (HOBO U12).

Through the 15 week study period, live weight measurements were taken every week for each group using a 0.01 g sensitive digital scale. For the live weight measurements, 10 quails were selected randomly from the first group, 20 from the second group and 30 from the third group. The gender of the animals was taken into consideration when measuring live weights, although evaluations of the study data were

made based on the flock/group totals. This is because the success rate is to identify outliers that may occur depending on the estimated increase in the number of samples.

Statistical analyses

The obtained data was analyzed using the STATISTICA 5.0 V package program, with the quail growth data being employed to determine the LS and M estimators (Huber et al., 1974; Davis, 1991). Following this, the number of outliers within the data was identified. Richards model is a dynamic agent. Therefore, we expect that it shows good performance in growth. The sigmoidal Richard's model used in this study can be expressed with the formula below (Seber, 1984):

$$f(x; \theta) = \frac{\theta_1}{(1 + e^{(\theta_2 - \theta_3 x)})^{1/\theta_4}} \quad (1)$$

in which θ_1 represents the largest possible asymptomatic value for the relevant characteristic; θ_2 represents the value of the relevant characteristic at time t_0 (the baseline); θ_3 represents the net growth rate; θ_4 represents the inflection point of the growth curve for the relevant characteristic (this point is a measure and indication of physiological maturity);

and ε represents the natural logarithm constant ($e = 2.718$).

The L, R, S and M estimators hold an important place among robust estimators. L estimators are a linear combination of order statistics, and include the sample mean, median and trimmed means as special cases. The R estimators are obtained through rank tests (hence their name). In the event of there being only a single sample, the R estimators are determined only for position problems. S estimators, on the other hand, represent a class of estimators with high break-down points. The aim of these estimators is to minimize the scattering of residues (Huber, 1981). Finally, M estimators generally correspond to Likelihood (L) estimators (Liu and Sirish, 2004), and it is for this reason that M estimators were the preferred method for this study.

Although the M estimator is somewhat protected against outlier values in the dependent variable, they are highly sensitive to, and easily influenced by, outlier values in descriptive variables. Determining the M estimator requires making an initial estimation using the LS method; and based on this estimation, new weights are calculated these are then used in the next series of estimations. Further estimations should be made until the stopping criterions met (Hampel, 1973). M estimators minimize the deviation functions of values obtained

through estimations that are more general than the total of absolute deviations, or the total of the square deviations. M estimators generalize the L estimator for the position parameter in a determined distribution (Huber, 1981).

M estimators are obtained not through the minimization of residue value squares used in the LS method, but through the minimization of residue values with another function. In this regard, they use the ρ function to reduce the disruption sensitivity of the LS estimation, which is a symmetric function of the residues (Huber, 1981). As such, the M estimator is defined as an estimator that uses the ρ function to minimize the residual values defined below:

$$\min_{\hat{\theta}} \sum_{i=1}^n \rho(e_i) \quad (2)$$

The c values are used to determine the highest outlier. In this context, the c value can be taken as $c = 0, 0.5, 0.3, 1/3$ (D'agostino and Stephens, 1986). c value is used as it is the most common factor. As such,

$$\hat{F}(X_{in}; \mu, \sigma) \cong \frac{i-c}{n-2c+1}, i=1,2,\dots,n, 0 \leq c < 1 \quad (3)$$

To determine the outliers, the SMAD values are determined with the aid of the following equations:

$$SMAD = \frac{1}{n-2} \left[\sum_{i=1}^n \rho \left(\hat{U}_c(X_{in}) - \frac{X_{in} - \mu}{\sigma} \right) \right] \quad (4)$$

and $k = 1, 2, \dots, n - \left\lfloor \frac{n}{2} \right\rfloor - 1$ for, (5)

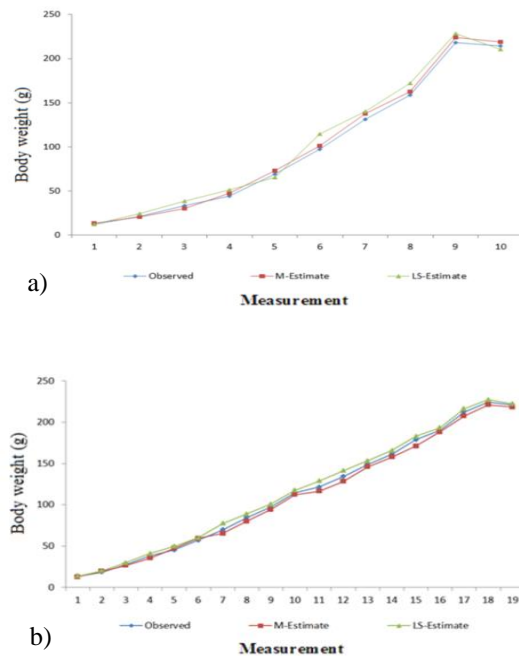
$$SMAD = \frac{1}{n-3} \left[\sum_{i=1}^{n-k} \rho \left(\hat{U}_c(X_{in}) - \frac{X_{in} - \mu}{\sigma} \right) \right] \quad (6)$$

The k value, which has the smallest SMAD value, indicates the largest number of outlier values/observations that can exist within the sample. Thus, by determining the k value with the smallest SMAD value through the use of the relevant c values, it could be calculated the largest possible number of outliers within the sample. In addition, the coefficient of determination and the error sum of squares were used to compare the parameter estimation performance of the M estimation and the LS method.

Results and Discussion

The obtained experimental results were evaluated. The real values for the time-dependent live weight of the Japanese quail, as well as the LS and M estimator values, based on the obtained samples, are shown in Figure 1.

An evaluation of Figure 1 reveals that the LS and M estimator values were fairly close to one another; although it was also noted that the group from which 30 samples were taken provided estimations closer to the real values. The estimation performance of the group from which 10 samples were taken was slightly lower when compared to the other groups. The parameters of the LS and M estimator values, which were calculated based on the collected samples, better illustrate the differences between these estimations (Table 1).



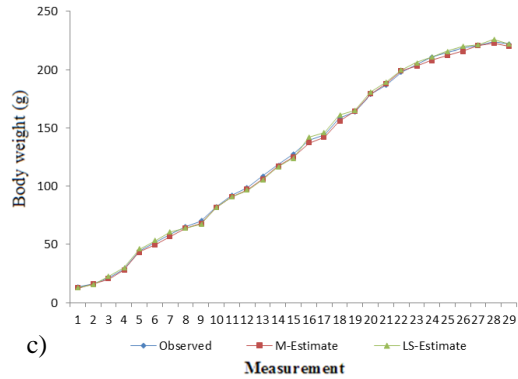


Figure 1. Body weight measured with respect to time with M and LS estimation values in Japanese quail; a)10 samples, b)20 samples, c)30 samples

Şekil 1. M ve LS tahmin değerleri ile zamana göre ölçülen Japon bıldırcını vücut ağırlıkları; a)10 örnek, b)20 örnek, c)30 örnek

An evaluation of Table 1 shows that the θ values were fairly similar for the three groups from which different numbers of sample were collected, with only the group from which 10 samples was taken differing slightly from the others. Looking at the Error Sum of Squares within the same table, it can be seen that for the 10 sample group, the MSS value of the LS method was 746.28, while the value for the M estimator was 731.44. For the 20 sample group, the MSS value of the LS method was 618.18, while the value for the M estimator was 620.11;

and for the 30 sample group, the ESS value of the LS method was 523.47, while the value for the M estimator was 523.47.

According to these results, increasing the number of samples had the effect of increasing the identification performance of both the LS and M methods, resulting in a parallel decrease in the level of error. It is possible to state that the model based on the 10 samples group was weaker when compared to the others, the reason for this being that a low number of samples leads to errors in model forming, to erroneous parameter estimations and to an erroneous analysis of the results (Liu and Sirish, 2004; Liu et al., 2011). Lower numbers of samples are also reported to decrease the coefficient of determination (Hancock and Buehl, 2008), while increasing the number of samples increases also the strength of the test (Cohen, 1992; Marchette and Solka, 2003). An increase in the number of samples is associated with a decrease in the error sum of squares and an increase in the coefficient of determination. As such, increasing the number of samples will serve to increase the effectiveness and strength of a study (Sahinler, 1997).

Based on the obtained values, the k value with the lowest SMAD value provided the largest possible number of outliers within the sample.

The SMAD values obtained when the c values of $c = 0, 0.5, 0.3, 1/3$ were used in the three sample groups for the residues obtained following the first estimated values are shown in Table 2. An evaluation of Table 2 shows that the SMAD value varied according to the number of samples, with the values obtained for the 10 samples and 20 samples groups at $c=0$ indicating that number of samples had the effect of

reducing the number of outliers, although a high number of samples may have actually had the effect of masking some of the outliers.

A study by Satman (2005) suggested that larger sample sizes had the potential to mask outliers, while Wisnowski et al. (2001) claimed that masking is observed more when the number of outliers is high.

Table 1. R^2 and MSS values of LS and M estimators

Çizelge 1. LS ve M tahmin edicilerin R^2 ve MSS değerleri

Parameters Parametreler	10 Samples 10 Örnek		20 Samples 20 Örnek		30 Samples 30 Örnek	
	LS	M	LS	M	LS	M
θ_1	284.11	281.62	278.27	277.12	265.21	265.03
θ_2	-2.156	-2.493	-2.065	-2.380	-1.976	-1.947
θ_3	0.241	0.242	0.241	0.240	0.238	0.238
θ_4	0.028	0.023	0.020	0.0150	0.024	0.018
R^2	92.72	93.06	94.12	95.07	96.13	97.42
MSS	764.28	731.44	618.18	620.11	523.47	492.35

In our study, the number of outliers varied according to the number of samples, and in this context, our results can be interpreted in two ways. The decrease in the number of outliers may have been due to the masking effect; or, alternatively, the increase in the number of samples may have engendered a decrease in the number of outliers.

It is said that a previous study using the LS method to identify outliers (Rousseeuw and Yohai, 1984) described that the effect and significance of the masking effect may be quite high in such cases (i.e. with large sample sizes). When there are multiple outliers within the data, the mean calculated from the sample will tend deviate/skew towards the outlier values; consequently, values that are actually outliers may

inadvertently appear as normal (Hawkins, 1984). It is believed that the main reason for the decrease observed in the number of outliers with increasing sample size was associated with the masking of the outliers.

The fact that previous studies obtained smaller MSS values using the LS method supports further the results of our study (Ergunes, 2004; Yildirim, 2010). In addition, based on the standard error values for the LS and M estimators, it is determined that the M estimator had a smaller standard error, which is in agreement with the findings of Hadi and Simonoff (1993), Sahinler (1997) and Karadavut et al. (2005). Rather than minimizing the squares of the residual values used in the LS method, M estimators perform a minimization by employing a different function of the residue values.

Rather than minimizing the squares of the residual values used in the LS method, M estimators perform a minimization by employing a different function of the residue values. In this regard, M estimators are obtained not through the minimization of residue value squares used in the LS method, but through the minimization of residue values with another function, and for this reason, the values obtained from the M

estimators can be considered as the expected results (Huber, 1981).

In this context, it is possible to ask why the LS estimator is used more often than the M estimator if the differences between them are limited. The answer to this is that the LS method is the non-deviating linear estimator with the lowest variance when the assumptions of the classical linear regression model are used. In other words, it is the best of the available estimators. The widespread use of the LS method is also associated with the ease with which the concept can be understood and applied (Karadavut et al., 2005). A general view is that outliers can be identified by evaluating the residues of the LS (Rousseeuw and Leroy, 1987); however, this is not the case for outliers that might be observed in the X direction (Ramsay and Elkum, 2005). In such cases, it becomes impossible to identify the regression curve using the LS approach (Stromberg et al., 2000), in that the LS curve will deviate towards the point in question, which will cause the larger residue values to appear smaller than they are, while the residues of the other points will appear to be larger than they are (Karadavut et al., 2005; Karadavut and Taskin, 2014).

Table 2. SMAD value determined according to the C value

Çizelge 2. C değerine göre belirlenen SMAD değerleri

$c = 0$			
k	10 Samples 10 Örnek	20 Samples 20 Örnek	30 Samples 30 Örnek
0	0.016384	0.016517	0.017426
1	0.016497	0.016640	0.017240
2	0.016622	0.016716	0.017219
3	0.017745	0.016788	0.017036
4	0.017647	0.016819	0.015160*
5	0.01771	0.015112*	
6	0.014335*		
$c = 0.5$			
k	10 Samples 10 Örnek	20 Samples 20 Örnek	30 Samples 30 Örnek
0	0.029162	0.028163	0.286470
1	0.027663	0.028034	0.293341
2	0.02886	0.027141	0.276887
3	0.031574	0.296370	0.261842
4	0.031464	0.294581	0.225591*
5	0.031424	0.246312*	
6	0.025565*		
$c = 0.3$			
k	10 Samples 10 Örnek	20 Samples 20 Örnek	30 Samples 30 Örnek
0	0.022322	0.241622	0.206387
1	0.021812	0.253844	0.196222
2	0.022333	0.248892	0.192584
3	0.024181	0.246312	0.200182
4	0.024063	0.254495	0.168820*
5	0.024091	0.223378*	
6	0.019525*		
$c = \frac{1}{3}$			
k	10 Samples 10 Örnek	20 Samples 20 Örnek	30 Samples 30 Örnek
0	0.023237	0.018244	0.019255
1	0.022611	0.018269	0.018424
2	0.023209	0.017230	0.018216
3	0.025171	0.017062	0.179380
4	0.025054	0.016993	0.145227*
5	0.025073	0.142695*	
6	0.020329*		

Conclusions

It possible to state that erroneous data entries will not lead to significant problems if the study sample is sufficiently large; however, in studies with smaller samples, outliers have the potential to cause serious problems. In studies with multiple samples, a high number of extreme values and outliers will prevent researchers from utilizing the obtained data effectively (Wang and Chow, 2003). Our study indicates that increasing the number of samples has the effect of changing the number of outliers, and our study also demonstrates that performing a normality control by itself is not sufficient, and that controlling outliers in addition to the normality would be a better approach. In cases where outliers are observed during studies, the generally adopted method involves the prompt removal of the outlier value/measurement, regardless of the total number of samples in the study. This is a highly inappropriate approach, since without proper knowledge of the actual relevance or importance of the removed value, the researchers may actually reach incorrect results and conclusions. In conclusion, our study has demonstrated that the M estimator, which is one of several different methods for estimating outliers, can be used to good effect in the identification of outliers within study data.

References

- Akdeniz, F., 1998. *Olasılık ve istatistik*, Baki Kitapevi, Adana, Turkey.
- Balcioglu, M. S., Karabag, K., Yolcu, H. I., Sahin, E., 2005. Japon bildircinlarında canlı ağırlığa göre iki yönlü seleksiyonun eşeyssel olgunluk yaşı ve bazı verim özellikleri üzerine etkisi. GAP. IV. Tarım Kongresi, 21-23 Eylül, Sanliurfa, Turkey.
- Bek, Y., Efe, E., 1987. *Araştırma deneme metotları* 1, C.U. Ziraat Fakültesi Ofset ve Teksir Atölyesi, Adana.
- Coelho, D. T., Dale, R. F., 1980. An energy-crop growth variable and temperature function for predicting corn growth and development: Planting to silking. *Agronomy Journal*, 72: 503-510.
- Cohen, J., 1992. Statistical power analysis. *Current Directions in Psychological Science*, 1: 98-101.
- Cook, R. D., 1997. Detection of influential observations in linear regression. *Technometrics*, 19: 15-18.
- D'agostino, R. B, Stephens, M. A., 1986. Goodness of fit techniques. Marcel Dekker Inc., New York.
- Davies, P. L., Gather, U., 1993. The identification of multiple outliers (with discussion), *Journal of Statistical Planning and Inference*, 122: 65-78.
- Davis, R. A., Knight, K., Liu, J., 1991. M-estimation for autoregressions with infinite variance. *Stochastic Processes and their Applications*, 40: 145-180.
- Ergunes, E., 2004. En küçük kareler yöntemi ile ridge regresyon yönteminin karşılaştırılması olarak incelenmesi. C.U. Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Adana.

- Gurcan, E. K., Cobanoglu, O., Genc, S., 2012. Determination of body weight-age relationship by non-linear models in Japanese quail. *Journal of Animal and Veterinary Advances*, 11: 314-317.
- Hadi, A. S., Simonoff, J. S., 1993. Procedures for the identification of multiple outliers in linear models. *Journal of the American Statistical Association*, 88: 1264-1272.
- Hampel, F. R., 1973. Robust estimation: A condensed partial survey. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 27: 87-104.
- Hancock, G. R., Buehl, M. M., 2008. Second-order latent growth models with shifting indicators. *Journal of Modern Applied Statistical Methods*, 7: 39-55.
- Hawkins, D. M., Bradu, D., Kass, G. V., 1984. Location of several outliers in multiple-regression data using elemental sets. *Technometrics*, 26: 197-208.
- Huber, P. J., 2005. *Robust Statistics*. New York, John Wiley and Sons, 1981.
- Huber, P. J., Dutter, R., 1974. Numerical solutions of robust regression problems, in: G. Brickmann, ed., *COMPSTAT 1974 (Physika Verlag, Wein)*: 165-172.
- Karadavut, U., Genc, A., Tozluca, A., Kinaci, I., Aksoyak, S., Palta, C., Pekgor, A., 2005. Nohut (*Cicerarietinum L.*) bitkisinde verime etki eden bazı karakterlerin alternative regresyon yöntemleriyle karşılaştırılması. *Tarım Bilimleri Dergisi*, 11: 328-333.
- Karadavut, U., Taskin, A., 2014. Estimation of heritability of weight gain of Japanese quail by using analysis of variance, maximum and restricted likelihood tests. *Türk Tarım ve Doğa Bilimleri Dergisi*, 1: 59-63.
- Kocak, C., Ozkan, S., 2000. *Bıldırcın, sülün ve keklik yetiştiriciliği*. Ege Üniversitesi Ziraat Fakültesi Yayınları. No: 538, İzmir.
- Liu, H., Sirish, J. W., 2004. On-line outlier detection and data cleaning. *Computers & Chemical Engineering*, 28: 1635-1647.
- Liu, M., Hancock, G. R., Harring, J. R., 2011. Using finite mixture modelling to deal with systematic measurement error: A case study. *Journal of Modern Applied Statistical Methods*, 10: 249-261.
- Marchette, D. J., Solka, J. L., 2003. Using data images for outlier detection. *Comput. Statist.*, 43: 541-552.
- Minvielle, F., 2004. The future of Japanese quail for research and production. *World's Poultry Science Journal*, 60: 500-507.
- Narinc, D., Aksoy, T., Karaman, E., Curek, D. I., 2010. Analysis of fitting growth models in medium growing chicken raised indoor system. *Trends in Animal and Veterinary Sciences*, 1: 12-18.
- Quackenbush, J., 2002. Microarray data normalization and transformation. *Nature Genetics*, 32: 496-501.
- Ramsay, T., Elkum, N. A., 2005. Comparison of four different methods for outlier detection in bioequivalence studies. *Journal of Biopharmaceutical Statistics*, 15: 43-52.
- Rousseeuw, P. J., Leroy, A. M., 1987. *Robust regression and outlier detection*. New York, John Wiley.
- Rousseeuw, P. J., Yohai, V., 1984. Robust regression by means of S-estimators. *Lecture Notes in Statistics*, 26: 256-272.
- Sahinler, S., 1997. Regresyon analizinde etkili gözlemlerin (Influential Observations) belirlenmesinde kullanılan istatistiklerin karşılaştırmalı olarak incelenmesi, Ç.Ü. Fen

- Bilimleri Enstitüsü, Zootekni Anabilim Dalı, Doktora Tezi, Adana.
- Satman, M. H., 2005. Doğrusal regresyonda aykırı gözlemlerin teşhis yöntemleri. İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, Ekonometri Anabilim Dalı, Yüksek Lisans Tezi, İstanbul.
- Seber, G. A. F., 1984. Multivariate observations. New York, John Wiley and Sons.
- Stromberg, A. J., Hossjer, O., Hawkins, D. M., 2000. The least trimmed differences regression estimator and alternatives. *Journal of the American Statistical Association*, 95: 853-864.
- Testik, A., Uluocak, N., Sarica, M., 1993. Değişik genotiplerdeki Japon bıldırcınlarının (*Coturnix coturnix japonica*) bazı verim özellikleri. *Turkish Journal of Veterinary and Animal Sciences*, 17: 167-173.
- Tserveni-Gousi, A. S., 1987. Relationship between parental age, egg weight and hatching weight of Japanese quail. *British Poultry Science*, 28: 749-752.
- Wang, W., Chow, S. C., 2003. Examining outlying subjects and outlying records in bioequivalence trials. *Journal of Biopharmaceutical Statistics*, 13: 43-56.
- Willemsen, H., Everaert, N., Witters, A., Smith, L., Debonne, M., Verschuere, F., Garain, P., Berckmans, D., Decuypere, E., Bruggeman, V., 2008. Critical assessment of chick quality measurements as indicator of post hatch performance. *Poultry Science*, 87: 2358-2366.
- Wisnowski, J. W. Montgomery, D. C., Simpson, J. R. A., 2001. Comparative analysis of multiple outlier detection procedures in the linear regression model. *Computational Statistics*, 36: 351-382.
- Yildirim, N., 2010. Determination the effects of outliers at the least squares, ridge regression and robust Regression analysis results. Ç.Ü. Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Adana.