Süleyman Demirel Üniversitesi Ziraat Fakültesi Dergisi 12 (1): 92-99, 2017
ISSN 1304-9984, *Araştırma Makalesi*

Canonical Correlation Analysis on Egg Production Traits of Quails

Çiğdem TAKMA[1] Yakut GEVREKÇİ[1] Abdullah Nuri ÖZSOY[2] Muzaffer ÇEVİK[1]

[1]Ege University, Faculty of Agriculture, Department of Animal Science, Biometry and Genetics Unit, Bornova/İZMİR
[2]Süleyman Demirel University, Faculty of Agriculture, Department of Animal Science, Biometry and Genetics Unit, ISPARTA
Corresponding author: cigdem.takma@ege.edu.tr

Received date: 08.05.2017, Accepted date: 19.07.2017

Abstract: In this study, canonical correlation analysis was applied to quail data for estimating egg productions. The model is defined that weight at sexual maturity (SMW), hatch time and chick length traits are the first set of variables (**X**); egg number of total 10 weeks that were weekly collected from age of first egg to 10th weeks and egg weight traits of this total 10 weeks are the second set of variables (**Y**). For obtaining the maximum relationship between sets of variables was aimed. Canonical variables were constructed from the linear combinations of the variables in each set. First canonical correlation between the first and the second pair of canonical variates were found 0.41 and statistically significant (P<0.01). Canonical weights and loadings from canonical correlation analysis indicated that SMW had the largest contribution as compared with hatch time and chick length to variation of egg number of total 10 weeks, egg weight of total 10 weeks.

Key words: Canonical correlation; egg number; egg weight, Coturnix coturnix japonica

Bıldırcınların Yumurta Verimi Özelliklerinde Kanonik Korelasyon Analizi

Özet: Bu çalışmada bıldırcın yumurta verimi, yumurta ağırlığı, çıkış zamanı, civciv uzunluğu ve eşeysel olgunluk ağırlığı özellikleri arasındaki ilişkiler kanonik korelasyon analizi ile incelenmiştir. Bu amaçla birinci ve ikinci değişken seti oluşturulmuştur. Çalışmada eşeysel olgunluk ağırlığı (EOA), çıkış zamanı ve civciv uzunluğu birinci değişken kümesinde (X), toplam on haftalık yumurta sayısı ve yumurta ağırlığı ise ikinci değişken kümesinde (Y) tanımlanmıştır. Böylece iki değişken kümesi arasındaki maksimum ilişkinin (korelasyonun) tahminlenmesi amaçlanmıştır. Değişkenlerin doğrusal kombinasyonlarından elde edilen yeni kanonik değişken çiftleriyle maksimum korelasyona ulaşılmıştır. Birinci ve ikinci kanonik varyans çifti arasında tahminlenen ilk kanonik korelasyon 0.41 ve istatistiksel olarak anlamlı bulunmuştur (P<0.01). Kanonik ağırlık ve kanonik yükler incelendiğinde, eşeysel olgunluk ağırlığının çıkış zamanı ve civciv uzunluğuna göre yumurta sayısı ve yumurta ağırlığı üzerine daha fazla katkıda bulunduğu belirlenmiştir.

Anahtar kelimeler: Kanonik korelasyon; yumurta sayısı; yumurta ağırlığı, Japon bıldırcını

Introduction

Quails are widely used animals in selection studies and because of short generation interval, low feed intake, high fertility and disease tolerance (Yang et al., 1998). The egg quality characteristics of quails affect hatchability, chick quality, and therefore future performance of breeding flock. The sexual maturity, egg weight and body weight are also some of the egg characteristics that effect on the egg production. In quail breeding, understanding and estimating the relationships between egg production characteristics is essential to determine important traits in selection for improving the effectiveness of the selection studies. It is especially important to decide which traits should be used in the selection studies. A number of methods are used to determine the relationship between the dependent and independent variables.

Canonical correlation is a method to measure the interrelationships between sets of multiple dependent and independent variables. The analysis is a generalized version of multiple regression analysis with more than one set of dependent variables. At the same time, this multivariate statistical technique is designed to assist the researcher in studying the complex interactions of data from two sets of variables. It also conciders two sets of variables related to each other and with how much variance of one set is common with or predictable from the other set (Barbosa et al., 2005). Canonical correlation analysis (CCA) is used to find the linear relationships in the analyzed to reveal the highest correlation between these sets (Yang et al., 2006; Mendeş and Akkartal, 2007).

In the poultry, although increasing the use of statistical package programs that are including CCA, there are still few studies (Gürbüz, 1989; Jaiswal et al., 1995; Yang et al., 1998; Keskin and Özsoy, 2004; Akbaş and Takma, 2005; Rosário et al., 2006; Mendeş and Akkartal, 2007; Çankaya et al., 2008; Ribeiro et al., 2016). However, there is lack of studies using canonical correlation in quails (Hidalgo et al. 2015; Ribeiro et al., 2016).

Hidalgo et al. (2015), applied CCA to estimate the relationship of nine traits: body weight at 28 days, egg weight, age at first egg and egg production at 30, 60, 90, 120, 150 and 180 days after onset of lay. Age at first egg and early stages of egg production were very influent and showed great importance in defining the canonical variates and the estimated canonical correlations. Ribeiro et al. (2016), used CCA to identify the correlation between egg production and reproduction traits of 629 meat type quails. The canonical coefficients showed that egg production and reproductive traits were moderately related (0.3475). Their study indicated that selecting animals with lower age at sexual maturity might lead to an increase on number of eggs, although it might also lead to a decrease on egg weight.

In this study it was aimed to determine the relationship between the weight at sexual maturity (SMW), hatch time, chick length as independent variable set and total 10 week egg number, total 10 week egg weight as dependent variable set by CCA.

Material and Methods

Material

The material of the study consists of 191 quail and their eggs obtained from Süleyman Demirel University, Agriculture Faculty, and Department of Animal Science. The SMW, hatch time and chick lengths were measured. The data set of arguments is defined by these variables. In addition, weekly egg numbers and egg weights were taken for during 10 weeks after the quail started laying eggs. At the end of this period, the total of 10 weeks of egg numbers and weights are defined in the dependent variable set.

Method

The simplest relationship known in statistics is the "simple correlation" between two random variables, denoted X and Y. If the number of variables is p, the correlation coefficient between one of the variables and the remaining p-1 variables is called "multiple correlation coefficients". The purpose of this analysis is to determine the independent variables affects or not on dependent variable set (Dillon and Goldstein, 1984). In calculations, it is tried to maximize the correlation between variable sets, and therefore new variables (canonical variable) pairs are obtained from linear combinations of variables which in both variable sets.

Variables can be determined as in the first set $X_1, X_2,......,X_p$ and in the second set $Y_1, Y_2,.......,Y_q$ Their linear combinations can be written as Equation 1 (Tatsuoka, 1971; Sharma, 1996; Özdamar, 2004; Çankaya, 2005):

$$Z = U_1X_1 + U_2X_2 + \ldots . + U_pX_p \quad W = V_1Y_1 + V_2Y_2 + \cdots + V_qY_q \qquad (1)$$

where, Z and W canonical variables, U_i and V_i coefficient of canonical variables, X_i and Y_i linear combination of set of variables, Relations between similar pairs of canonical variables are called "canonical correlations" (Keskin and Özsoy, 2004). The coefficients in the linear combinations show the amount of change from the canonical variance to the standard deviation in response to a standard deviation increase from the original variance, while another variation defines the standardized canonical coefficients. If we show the coefficient matrices as U=[U_1, U_2,....,U_p] and V=[V_1, V_2,....,V_q] in Equation 1, the highest combination between two set, can be determined as canonic correlation coefficient (r_{zw}) which is function of U and V (Johnson and Wichern, 2002; Özkan et al., 2008; Koşkan et al., 2011).

$$r_{zw} = \frac{u'\sum 12v}{\sqrt{(u'\Sigma 11u)(v'\Sigma 22v)}}$$

The Canonical correlation coefficient is equal to the number of variables in the X and Y sets that have less than the variable. The first canonical variable pair represents the highest correlation, while the second canonical variable pair represents the second highest correlation. In other canonical variable pairs, the correlation is gradually decreasing (Hamarat and Özen, 2015). To determine the significance of the correlation coefficients, the test is started from the first correlation coefficient to the statistically insignificant correlation coefficient (Tatlıdil, 2002). This test is carried out with the "Wilks' Lambda" statistic (Johnson and Wichern, 2002).

The magnitude of the canonical correlation coefficients between the two sets of variables does not account for the existence of a strong relationship. The high canonical correlation and a high degree of shared variance explained by the dependent variate give a high redundancy index. The redundancy index derived by multiplying the two components (shared variance of the variate multiplied by the squared canonical correlation) to find the amount of shared variance that can be explained by each canonical function and it is calculated as follows (Sharma, 1996):

$$RM_{V_i|W_i} = AV(Y|V_i)xC_i^2$$

$$AV(Y|V_i) = \frac{\Sigma_{j=1}^{q} LY_{ij}^2}{q}$$

In equation *AV(Y|Vi)*, explaining part of the variation in Y variable by i. canonical variable V, LY_{ij}^2, j. variable load between j. variable and i. canonical variable in Y variable set; *q*, number of trait in canonical variables, c_i^2 polled variance between V_i and W_i.

In addition, some assumptions must be done for application of the canonical correlation analysis. These assumptions can be listed as having multiple distributions of data, no multicollinearity between traits and the sample size being as large as possible (Koşkan et al., 2011). In this study, firstly the mentioned assumptions were examined before applying the canonical correlation analysis. Normal distribution test was applied to the data. The chick length and egg number data have non-normal distribution. The logarithmic transformation for chick length data and the arcsin transformation for egg number data were used. SPSS 20 (SPSS, 2011) statistical package program were utilized to analyze all statistics.

Results and Discussion

Descriptive statistics of traits were given in Table 1. The mean, standard deviation and coefficient of variation of five traits which were analyzed by canonical correlation analysis was estimated. Among these traits, coefficient of variation for chick length was found the lowest as 1.57%, and also the highest variation was found for hatch time as 10.54% (Table 1). The phenotypic correlation coefficients between the traits were given in Table 2. There wasn't any multicollinearity between traits. Relationships between all traits was found statistically significant (P<0.01). The highest correlation coefficient among traits was found between SMW and egg weight (r=0.25).

Table 1. Descriptive statistics for traits
Tablo 1. Tanımlayıcı istatistikler

Traits *Özellikler*	Mean *Ortalama*	Standard deviation *Standard sapma*	Coefficient of variation (%) *Varyasyon katsayısı*
SMW EOA	233.09	20.95	8.99
Hatch time *Çıkış zamanı*	8.03	0.85	10.54
Chick length *Civciv uzunluğu*	1.02	0.02	1.57
Egg number *Yumurta sayısı*	7.8	0.59	7.45
Egg weight *Yumurta ağırlığı*	112.84	11.76	10.42

Table 2. Bivariate correlation considering canonical traits
Tablo 2. Kanonik özellikler arası korelasyonlar

	Hatch time *Çıkış zamanı*	Chick length *Civciv uzunluğu*	Egg number *Yumurta sayısı*	Egg weight *Yumurta ağırlığı*
SMW EOA	0.196**	0.220**	0.101	0.249**
Hatch time *Çıkış zamanı*		0.185*	-0.022	0.167*
Chick length *Civciv uzunluğu*			0.004	0.169*
Egg number *Yumurta sayısı*				0.242**

* $P \leq 0.05$ ** $P \leq 0.01$

The canonical correlation coefficients calculated from the CCA were shown in Table 3. Two different canonical variable pairs and two different canonical correlation coefficients between them from were obtained (Table 3). When the Wilk's Lambda values for the significance test of the canonical correlation coefficients were examined, it was shown that the first canonical correlation coefficient (0.41) calculated between only the first canonical variable pair was found statistically significant (P<0.01). On the other hand, the canonical correlation coefficient between the second canonical variable pairs was found non-significant (P>0.05). This result is similar with the findings of Hidalgo et al. (2015) and Ribeiro et al. (2016), who reported the first pair of canonical variates as 0.658 and 0.1208, respectively.

Table 3. Significance test of canonical correlation coefficients
Tablo 3. Kanonik korelasyon katsayılarının önemlilik testi

Canonical Variable Pairs *Kanonik değişken çiftleri*	Canonical Correlation Coefficient *Kanonik korelasyon katsayısı*	Wilks Lambda *Wilk'in Lamdası*	Chi-Square *Ki-kare*	DF *Serbestlik derecesi*	P *Önemlilik*
U_1V_1	0.41	0.829	31.55	6	0.00
U_2V_2	0.05	0.997	0.47	2	0.79

Standardized canonical coefficients (weights) of canonical variable pairs were given in Table 4. Two different U and V linear components were constructed by using the standardized canonical coefficients of the canonical variable pairs and it was given in Table 4. Because of first canonical variable pair was significant, the linear components of the first canonical variables were written and given as below:

U_1 = -0.766 (SMW)-0.446 (Hatch time)-0.105 (Chick length)

V_1 = -0.339 (Egg number)-0.862 (Egg weight)

Table 4. Standardized canonical coefficients (weights) of canonical variable pairs
Tablo 4. Kanonik değişken çiftlerinin standartlaştırılmış kanonik katsayıları

	X variable set / *X değişken seti*		
	SMW EOA	Hatch time *Çıkış zamanı*	Chick length *Civciv uzunluğu*
U_1	-0.77	-0.45	-0.11
U_2	-0.57	0.23	0.93
	Y variable set / *Y değişken seti*		
	Egg number *Yumurta sayısı*	Egg weight *Yumurta ağırlığı*	
V_1	-0.34	-0.86	
V_2	-0.97	0.56	

Table 4 shows that there were negative contributions of all traits for obtaining U_1 and V_1. Moreover, SMW had the highest contribution to U_1 whereas egg weight had higher contribution to V_1.

In canonical correlation analysis, when the sample size was small and with multicollinearity in the data set, in case of doubt, structural correlations (loadings) were calculated instead of the standardized coefficients. Canonical loadings explain the correlation coefficients between canonical variables and the original variables in that set (Sharma, 1996).

Table 5. Canonical loadings for canonical variable pairs
Tablo 5. Kanonik değişken çiftlerine ait kanonik yükler

	X variable set *X değişken seti*			*Y variable set* *Y değişken seti*	
	U_1	U_2		V_1	V_2
SMW EOA	-0.89	-0.29	Egg number *Yumurta sayısı*	-0.55	-0.84
Hatch time *Çıkış zamanı*	-0.63	0.30	Egg weight *Yumurta ağırlığı*	-0.94	0.33
Chick length *Civciv uzunluğu*	-0.39	0.83			

In Table 5, it was seen that because of the importance of the first canonical variable set, the SMW had the highest loading value (-0.89) when canonical loadings in the independent variable set was examined. In other words, the SMW was determined as the most effective variable in the obtaining the U_1 canonical variable. Thereafter hatch time (-0.63) and chick length (-0.39) were affecting, respectively. In the dependent variable set the highest loading value was determined for egg weight (-0.94) and the loading value of egg number was found as -0.55 constructing the V_1 canonical variable.

Canonical loadings indicate the correlation between the variables and their respective canonical variate, whereas canonical cross loadings indicate the correlation between the variables and the opposite canonical variate. Canonical loadings of original variables in the canonical variable pairs and other variable set were presented in Table 6. All traits that used in CCA, SMW, hatch time, chick length, egg number and egg weight had the negative cross loadings and inverse relationships in the variates U_1 and V_1. When canonical loadings were examined;

the largest contribution to the U_1 canonical variable was the egg weight (-0.39) and it was followed by the egg number (-0.23). By squaring these canonical loadings (0.26 and 0.05), it might be said that 26% of the variance in egg weight and 5% of the variance in egg number are explained by the variate U_1 (Table 6).

For the other set of variables (V_1), The SMW also had the greatest contribution (-0.36) for the V_1 canonical variable, while the hatch time (-0.26) and chick length (-0.16) follow it (Table 6). By squaring these canonical loadings (0.13, 0.07 and 0.03), it can be said that 13% of the variance in SMW, 7% of the variance in hatch time and 3% of the variance in chick length are explained by the variate V_1 (Table 6).

Table 6. Canonical correlations between the variables and related canonical variates (canonical cross loadings)
Tablo 6. Orijinal kanonik değişken çiftleri için diğer değişken setindeki kanonik yükler

	V_1	V_2		U_1	U_2
SMW EOA	-0.36	-0.02	Egg number *Yumurta sayısı*	-0.23	-0.04
Hatch time *Çıkış zamanı*	-0.26	0.02	Egg weight *Yumurta ağırlığı*	-0.39	0.02
Chick length *Civciv uzunluğu*	-0.16	0.04			

In this study, the redundancy index for the first canonical correlation is found 45%, while for the second canonical correlation this value was 29%. It shows that 45% of the explanatory variation in the Y variable set can be explained by the variables in the X variable set or by whole X variable set. In the other sense, for the first canonical correlation, 45% of the variation in the Y variable set is due to the variation in the X variable set. When the total redundancy index value is calculated, 74% of the variation in the Y variable set is explained by the X variable set. It was determined that 60.8% of the total variation (0.45/0.74) was defined by the first canonical pair.

Akbaş and Takma (2005), coefficient of correlation between egg yield characteristics and body weight, egg weight and sexual maturity age in laying hens found significant ($P<0.01$) with 0.81 value ($P<0.01$). They found that the egg number up to the 36 weeks of age of sexual maturity are the most important variables for obtaining U_1 and V_1 canonical variables, respectively. Jaiswal et al. (1995) used canonical correlation analysis to simultaneously examine the relationship between several distinctive traits in chickens. The researchers studied egg weights at 32-week old chickens, 280-day egg yields as Y variable group, the first egg-laying age of chickens, 8-week and 20-week body weights as X variable group. The canonical correlation coefficient between the first canonical variable pair was significant, the explanatory index value and 93% of the total variance in the Y variable group can be explained by the X variable group (Çankaya, 2005). Çankaya et al. (2008), found significantly important ($P<0.01$) canonical correlation coefficient between sexual maturity and egg yield traits as 0.43.

Conclusion

Canonical correlation is a method that determining quantitative and qualitative relationships between traits, in a simultaneous study of many traits in biological researches. But, result of the canonical correlation analysis does not mean that the large canonical correlation coefficient value obtained always correlates strongly between the two sets of characteristics. Canonical correlation analysis maximizes the correlation between the linear combinations of the variables in the two groups and does not maximize the variance calculated in the other variable group according to one variable group. As a result, although the long process of the canonical correlation analysis and the computational complexities are considered to be disadvantageous for this analysis, recent developments in computer technology, it is getting easier.

In this study, when the linear components U_1 and V_1 are examined without regard to the signs of the coefficients of the canonical variable pairs, it is seen that the contribution of the SMW variable has the biggest contribution to obtain the U_1 canonical variable. On the other hand, it has been determined that the egg weight variable significantly contributes to the obtaining the V_1 canonical variable. In addition, according to the calculated canonical loadings, the effect of SMW, hatch time and chick length variables have opposite way on the obtaining of U_1 canonical variable. In our study, it is thought that the selection of quail according to lower SMW will be beneficial to increase egg number and egg weight.

References

Akbaş, Y. and Takma, Ç., 2005. Canonical Correlation Analysis for Studying the Relationship Between Egg Production Traits and Body Weight, Egg Weight and Age at Sexual Maturity in Layers. Czech J. Anim. Sci., 50(4): 163–168.

Barbosa, L., Lopes, P.S., Regazzi, A.J., Guimarães, S.E.F. and Torres, R.A., 2005. Estudo da associação entre características de desempenho e de carcaça de suínos por meio de correlação canô- nica. Rev Bras Zootec 34:2218-2224 (Abstract in English)

Çankaya, S., Kayaalp, G.T., and Ünalan, A., 2005. Estimation of Correlation Among Sex, Birth Weight and Some Body Measurements in South Anatolian Red Calves by Using Canonical Correlation Analysis. In 3rd International Conference of the EMR-IBS.

Çankaya, S., Yazgan, E., Kayaalp, G.T., Göçmez, Z., and Serbester, U., 2008. Canonical Correlation Analysis for Estimation of Relationship Between Some Body Measurements at The Birth and 6 Month Periods in Holstein Friesian Calves. Journal of Animal and Veterinary Advances, 7(8): 953-958.

Dillon, W.R., Goldstein, M., 1984. Multivariate Analysis Methods and Applications, John Wiley and Sons Inc., New York.

Gürbüz, F., 1989. Değişken Takımları Arasındaki İlişkilerin Kanonik Korelasyon Yöntemi ile Araştırılması. Ankara Üniversitesi. Ziraat Fakültesi Yayınları: 1162, Ankara.

Hamarat, B., Özen, E., 2015. Türkiye'de Tasarruf Tercihlerini Etkileyen Değişkenlerin Kanonik Korelasyon Analizi ile Belirlenmesi, Journal of Life Economics, 3:47-74.

Hidalgo, A.M., Silva, L.P., Mota, R.R., Martins, E.N., 2015. Canonical-Correlation Analysis Applied to Selection-Index Methodology in Quails, Livestock Science, 169:35–41.

Jaiswal, U.C., Poonia, J.S., Kumar, J., 1995. Canonical Correlation Analysis for Studying Relationship Among Several Traits: An Example of Calculation and Interpretation. Indian J. Anim. Sci. 65: 765-769.

Johnson, R.A., Wichern, D.W., 2002. Applied Multivariate Analysis. Prentice Hall Upper Saddle River, N.J., p:77.

Keskin, S. and Özsoy, A.N., 2004. Kanonik Korelasyon Analizi ve Bir Uygulaması, Tarım Bilimleri Dergisi, 10(1): 67-71.

Koşkan, Ö., Önder, E.G., Şen, N., 2011. Değişken Setleri Arası İlişkinin Tahmini İçin Kanonik Korelasyon Analizinin Kullanımı. Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi 1(2): 117-123.

Mendeş, M., and Akkartal, E., 2007. Canonical Correlation Analysis for Studying the Relationships Between Pre and Post Slaughter Traits of Ross 308 Broiler Chickens. Archiv Fur Geflugelkunde, 71(6): 267-271.

Özdamar, K., 2004. Paket Programlar ile İstatistiksel Veri Analizi, 2. Yenilenmiş 5. Baskı. Kaan Kitabevi.

Özkan, M.M., Adak, S.M., Kocabaş, Z., 2008. An Investigation on the Relationship Between Yield and Canopy Components in Wheat (Triticum aestium) Tarım Bilimleri Dergisi, 14: 148-153.

Ribeiro, J.C., Silva, L.P., Caetano, G.C., Crispim, A.C., Pacheco, R.O., Torres, R.A., 2016. Canonical Correlation Analysis Applied to Production and Reproduction Traits of Meat Type Quails, Ciência Rural, Santa Maria, 46(7): 1289-1294.

Rosário, M.F., Silva, M.A.N., Savino, V.J.M., Coelho, A.A.D., and Dias, C.T.S., 2008. Canonical Discriminant Analysis Applied to Broiler Chicken Performance. Animal, 2(3): 419-424.

Sharma, S., 1996. Applied Multivariate Techniques: Canonical Corelation, 391-418. John Willey and Sons Inc., USA.

SPSS, 2011. SPSS for Windows, Version 20, SPSS Inc., Chicago.

Tatlıdil, H., 2002. Uygulamalı Çok Değişkenli İstatistiksel Yöntemler, Cem Web Ofset, Ankara. s:216-252.

Tatsuoka, M.M., 1971. Multivariate Analysis: Techniques for Educational and Psychological Research. Wiley. New York.

Yang, N., Dunnington, E.A., and Siegel, P.B., 1998. Forty Generations of Bidirectional Selection for Mating Frequency in Japanese Quail. Poultry Science, 77: 1469–1477.

Yang, Y., Mekki, D.M., Lv, S.J., Yu, J.H., Wang, L.Y., Wang, J.Y. Xie, K.Z. and Dai, G.J., 2006. Canonical Correlation Analysis of Body Weight, Body Measurements and Carcass Characteristics of Jinghai Yellow Chicken, Journal of Animal and Veterinary Advances, 5(11): 980-984.