Usability of Ridge Regression in the Existence of Multicollinearity for Body Measurements in Saanen Kids

Cem Tirink\textsuperscript{1}\textsuperscript{*} Samet Hasan Abacı\textsuperscript{2} Hasan Önder\textsuperscript{3}

*: cem.tirink@gmail.com
1: Ondokuz Mayıs University, Faculty of Agriculture, Dept. of Animal Science, Samsun, Turkey
2: Ondokuz Mayis University, Faculty of Agriculture, Dept. of Animal Science, Samsun, Turkey
3: Ondokuz Mayis University, Faculty of Agriculture, Dept. of Animal Science, Samsun, Turkey

Abstract

One of the biggest problems in the least squares (LS) multiple regression models, is multicollinearity. multicollinearity problem can be eliminated using Ridge Regression (RR), which is a biased estimation method, and it is possible to obtain models that have more reliable determination coefficient ($R^2$). It is aimed in this study to examine the usability of ridge regression in the existence of multicollinearity on the effect of some biometric measurements (height at withers and rumps, body length, chest width, chest girth and chest depth) on body weights obtained from 40 Saanen kids. The multicollinearity problem was determined between height at withers and height at rump to estimate body weight, due to Variance Inflation Factor (VIF) value was higher than ten. This multicollinearity problem has been eliminated using the RR. It was observed that the standard errors of coefficients obtained from RR were decreased according to LS estimations. $R^2$ was determined as 0.88 from LS method and the $R^2$ was determined 0.875 with $k=0.0136$ from RR method. Results showed that the model obtained from RR model was more reliable than LS.

Keywords: Least Squares Multiple Regression Model, Multicollinearity, Ridge regression