

Comparative Evaluation of Residual Diagnostic Measures for Outlier Detection in Linear Regression: An Empirical Study Using Heart Rate Data: Reg No: 410

Shashank Kirti

Integral Institute of Medical Sciences & Research

CORRESPONDING AUTHOR

Dr Shashank Kirti, Integral Institute of Medical Sciences & Research

Email: drshashankkirti@gmail.com

CITATION

Kirti Shashank, . Comparative Evaluation of Residual Diagnostic Measures for Outlier Detection in Linear Regression: An Empirical Study Using Heart Rate Data: Reg No: 410. Journal of the Epidemiology Foundation of India. 2025;3(2Suppl):84.

DOI: <https://doi.org/10.56450/JEFI.2025.v3i2Suppl.084>

ARTICLE CYCLE

Received: 10/11/2025; Accepted: 20/12/2026; Published: 31/12/2025

This work is licensed under a Creative Commons Attribution 4.0 International License.

©The Author(s). 2026 Open Access

ABSTRACT

Outliers in medical datasets can substantially distort regression models, leading to biased estimates and unreliable clinical inferences. This study evaluated residual-based diagnostic methods-including standardized residuals, Cook's Distance, DFFITS, DFBETAS, Covariance Ratio, and standardized scores-for detecting influential observations in a simple linear regression model. Using age as the predictor and heart rate as the outcome, data from 30 individuals were analyzed. The initial model ($R^2 = 0.22$) revealed one highly influential outlier (heart rate = 290 at age 42), consistently identified across multiple diagnostics. Removal of this case improved model adequacy, yielding a more stable regression ($R^2 = 0.10$) with no further significant outliers detected. Findings highlight that even a single outlier can markedly distort regression-based conclusions, underscoring the importance of employing complementary diagnostic tools to ensure valid statistical inference in medical research.

KEYWORDS

Outlier detection linear regression residual diagnostics