XVII Congreso Galego de Estatística e Investigación de Operacións
Ourense, 23 - 25 de outubro de 2025

CONFIDENCE INTERVALS FOR THE AUC WITH MISSING DATA

Susana Rafaela Martins [1], María del Carmen Iglésias-Pérez [2] Jacobo de Uña-Álvarez [3]

[1] Instituto Nacional de Estatística, I.P. (INE), Portugal
[2] Department of Statistics and OR, Universidade de Vigo, Spain
[3] Department of Statistics and OR, Universidade de Vigo, Spain

ABSTRACT

The Area Under the ROC Curve (AUC) plays an important role in various fields, such as biomarker identification or the study of the predictive capacity of regression models. In this work, we consider the construction of confidence intervals for the AUC in the presence of missing data. Several approaches to construct confidence intervals from the AUC are explored, based on empirical AUC: the Newcombe variance-based method and the bootstrap method. To deal with missing data, we use complete case analysis and multiple imputation. The relative performance of the methods is investigated in simulated missing completely at random and missing at random scenarios, considering a setting with a fully observed univariate biomarker and a potentially missing binary outcome. The methods are compared in terms of empirical coverage and interval width, providing practical recommendations for selecting the most appropriate procedure for the construction of confidence intervals for the AUC in the presence of incomplete data.

Keywords: Monte Carlo, Bootstrap, Newcombe, Multiple Imputation, Complete Case Analysis

REFERENCES

Barnard, J. and Rubin, D. B. (1999) Small-sample degrees of freedom with multiple imputation. Biometrika, 86, 948–955.

Cho, S.-J., Kim, S. and Lee, W. (2019) Interval estimation of AUC under missing response data. Statistical Methods in Medical Research, 28, 2412–2426.

Little, R. J. A. and Rubin, D. B. (1987) Statistical analysis with missing data. John Wiley & Sons, New York.

Pepe, M. S. (2003) The statistical evaluation of medical tests for classification and prediction. Oxford University Press, Oxford.

Schomaker, M. and Heumann, C. (2018) Bootstrap inference when using multiple imputation. Statistics in Medicine, 37, 2252–2266.