Least square differences method for quantitative determination of rater bias

Author:

R. Clive Woods, Department of Electrical and Computer Engineering and Microelectronics Research Center, 2128 Coover Hall, Iowa State University, Ames, Iowa 50011, U.S.A.. cwoods@iastate.edu

Abstract *— In order to assess a large number of diverse projects as fairly and as rapidly as possible, a procedure often adopted is to use a panel consisting of a large number of assessors or raters, only a small number of whom assess each project. Since no single rater assesses all the projects, conscious or unconscious bias regarding overall standards by any rater will advantage or disadvantage the projects assessed by that particular rater. 'Rater bias' is the tendency for a project rater to return assessments that deviate from the required and established impartial standards so that the project(s) rated by that assessor are systematically advantaged or disadvantaged. The valuable criterion-referenced rating technique may be used to reduce rater bias, but assessing a project is still usually subjective and criterion-referencing does not entirely eliminate rater bias. Until recently only* ad-hoc *methods of determining rater bias were available. Perhaps the most direct was to distribute to the raters "standard examples" of project reports for assessment, having expected marks known to the overall investigator but not to the raters at the time of assessment; this method is impractical in evaluation procedures typically requiring several hours of work for a single complete assessment. Recently Woods [Woods, R.C., "Iterative processing algorithm to detect biases in assessments", IEEE Trans. Educ., 46(1), 2003, 133–141] proposed a method of determining rater bias quantitatively from the complete set of genuine assessments, based upon a comparison of the 'paired assessments' of each rater with those other raters assessing the same projects. Independently Chan [Chan, K.L., "Statistical analysis of final year project marks in the computer engineering undergraduate program", IEEE Trans. Educ., 44(3), 2001, 258–261] proposed a method of evaluating rater bias using a commercial statistical package. In the present paper, a third method of determining rater bias quantitatively is examined. This method is based upon finding the summed square deviations of each mark from the corresponding paired assessments, and minimizing this summed square deviation by adjustment of the means and standard deviations of each assessor in turn. The method has the advantage of placing the determination of rater bias on a more rigorous mathematical platform than previously, at the expense of more difficult data processing and slower numerical convergence. The algorithm will be presented in detail, together with the results of applying it to the raw data used by both Woods and by Chan in presenting their original methods, so that the differences between the three techniques may be clearly determined.*

Index Terms *— Assessment, grades, iterative algorithm, marks, peer-review, projects, proposals, rater bias.*