Detection and Contribution of Outliers for Subjective Evaluation of Sound

Samir N. Y. Gerges

Federal University of Santa Catarina (UFSC), Mechanical Engineering, Florianopolis, SC, Brazil 88040-900.
Institute Federal de Santa Catarina (IFSC) / Mechatronics, Florianopolis, SC, Brazil 88020-300.
NR consultancy and Training, Florianopolis, SC, Brazil 88035-200.

Roberto A. Dias

Institute Federal de Santa Catarina (IFSC) / Mechatronics, Florianopolis, SC, Brazil 88020-300.

Rafael N. C. Gerges

NR consultancy and Training, Florianopolis, SC, Brazil 88035-200.

(Received 8 February 2015; accepted 3 May 2016)

The subjective evaluation of noise perception is a very broad topic that has many applications in the field of acoustics. Large variability is usually associated with a subjective evaluation that appears in the standard deviation. This is due to a small amount of subjects (the outliers), who had different responses compared to most of the other subjects. By using the Bootstrap statistical method, this paper shows how to identify the outliers and quantify the contribution to the final results with and without considering the outliers in the calculation.

1. INTRODUCTION

The subjective evaluation of noise perception plays an important role in the decision making of many applications in the field of acoustics, such as the evaluation of noise perception (annoyance) in communities located near airports and studies on traffic noise, product sound quality, environmental soundscapes, sleep disturbance, and hearing protector noise attenuation. The subjective perception of noise by a group of human evaluators usually shows a large variability, as observed through the standard deviation. This is because the subjects differ in terms of their experience, attitudes, expectations, age, personal state of mind, sensitivity to noise, fear of harm connected with the source, personal evaluation of the source, coping capacity with respect to noise, trust in or lack of confidence in the relevant authorities, and a history of noise exposure, among other factors. Some subjects paid greater attention to the assessment and provided more accurate responses while others did not concentrate properly on the task and performed the evaluation simply for the payment. In general, a small amount of the subjects gave responses that differed from most of the other subjects. These few subjects tended to have a considerable influence on the final results and were the main reason behind the high standard deviation. Thus, they will be considered as "outliers." This is a very broad topic that has extensive applications in different fields.[1–3]

2. STATISTICAL DETERMINATION AND DETECTION OF OUTLIERS

There were different methods available for the detection of outliers,[4] including the box plots, miss match models, the Dixon test, the Grubbs test, and Z-scores. In the study reported, a mismatch model and the bootstrap method were used to investigate the statistical distribution and to identify outliers. Additionally, the quantitative effect on the final results when the outliers were removed from the dataset was determined.

The subjects considered as outliers, who generally represented around 3 to 5 out of the total 20 to 30 subjects in the cases presented herein, were not true outliers. However, their subjective evaluation was very different from that of the other subjects. The objective of this paper is to describe a way to detect these outlier subjects and evaluate their effect on the final results by eliminating them. Some real cases were described here to show the application of the bootstrap statistical technique to the identification of outliers and to evaluate their contribution to the results.

2.1. Bootstrap Method

The bootstrap method was introduced by B. Efron in 1979[5] and its use in statistical sciences became widespread within a few decades. This method involves taking the original dataset of N elements and sampling from it by using a computer in order to generate a new sample with size $N/2$. The elements are then exchanged randomly between these two datasets (each of size $N/2$).

This process is repeated many times and for each of these bootstrap samples, the final parameter (e.g. the mean) is computed. The histogram of this final parameter is obtained with the mean value and standard deviation together with the contribution of the N original elements to the final parameter. The statistical distribution is then observed and some outliers will show a non-Gaussian distribution. The outliers can also be identified and their contribution to the final results evaluated. If these outliers are eliminated and the distribution is recalculated, it becomes more Gaussian with a better estimation of the parameters.

In this study, 20 subjective responses were obtained, with 3 to 7 questions given to each subject. The random selection