Research Note
Standard Scores Based on the Median and Inter-quartile Range

James Carifio
University of Massachusetts-Lowell, USA

_____

A LOGICAL, GENERAL DESCRIPTION

Recently when asked to help with statistical analyses of the Medicare Provider Analysis and Review (MEDPar) Data, I encountered a common statistical and analytical problem with this national database, which is also common to many national and health databases.

This problem was the problem of extremely skewed distributions (g1 and/or g2 typically exceeding 20), which creates numerous difficulties in comparing and interpreting values and results both between and within variables, never mind statistically testing them.

Typically, a researcher confronted with this problem will convert the raw scores to standard or "z" scores using the mean and standard deviation of each distribution to remove the relativity and make the scores directly comparable and interpretable. However, when the distributions are extremely skewed (as is typically the case with health and medical data), this practice produces much distorted and misleading standard scores that either over-represent (more favorably than they really are) or under-represent (more negative than they really are) individual cases (say doctors, procedures, or hospitals) and statistical indices upon which they are based. The percentage of cases within the standard scores units (and thus the percentiles) also are both variable and relative to the exact distributions making interpretations very difficult and laborious.

Faced with this problem, I developed a new standard score based on the observed or individual value or case's deviation from **the median** of the distribution divided by the distribution's **interquartile range** so that plus or minus one unit would be the middle 50% of the distribution. One could use the standard deviation, but the interquartile range is better and has more advantages in most contexts and situations.

This new standard score could be called an "ordinal" standard score. It is very "distribution free" and makes comparisons within and between distributions fairer and less distorted, and very easy and direct to interpret. These features of the median-based standard score are or would be particularly important in comparing results between procedures, programs, or hospitals, as the data would be expressed in terms of less biased and less distorted **typical results** and deviations from this **"truer" typicality.** The metrics based on these standard scores would be a fairer and less biased type of standardized effects measures. This last point would be particularly important in conducting meta-analyses where the findings using these median-based standard scores would be less distorted, easier to interpret, and more indicative of typicality of results.

I believe that standard scores based on the median should be used in meta-analysis and healthcare result and evaluation. I want to share the points I have made in this note with health profession researchers and practitioners as quickly as possible so that their work may benefit from the solution I have devised to this common and long standing problem and difficulty with health and medical data.

_____

Corresponding Author: James Carifio, University of Massachusetts-Lowell, USA