Available online at www.sciencedirect.com

ScienceDirect

Journal of Consumer Psychology 25, 4 (2015) 690 – 704

JCP JOURNAL OF CONSUMER PSYCHOLOGY

Research Article

The median split: Robust, refined, and revived☆

Dawn Iacobucci [a,*], Steven S. Posavac [a], Frank R. Kardes [b], Matthew J. Schneider [c], Deidre L. Popovich [d]

[a] *Vanderbilt University, Nashville, TN 37203, USA*
[b] *Lindner College of Business, University of Cincinnati, Cincinnati, OH 45221, USA*
[c] *Medill School of Journalism, Northwestern University, Evanston, IL 60201, USA*
[d] *Rawls College of Business, Texas Tech University, Lubbock, TX 79409, USA*

Received 17 June 2015; accepted 26 June 2015
Available online 3 July 2015

Abstract

In this rebuttal, we discuss the comments of Rucker, McShane, and Preacher (2015) and McClelland, Lynch, Irwin, Spiller, and Fitzsimons (2015). Both commentaries raise interesting points, and although both teams clearly put a lot of work into their papers, the bottom line is this: our research sets the record straight that median splits are perfectly acceptable to use when independent variables are uncorrelated. The commentaries do a good job of furthering the discussion to help readers better develop their own preferences, which was the purpose of our paper. In the final analysis, neither of the commentaries pose any threat to our findings of the statistical robustness and valid use of median splits, and accordingly we can reassure researchers (and reviewers and journal editors) that they can be confident that when independent variables are uncorrelated, it is totally acceptable to conduct median split analyses.

Keywords: Median split; Median-split; Dichotomization; Categorization

Introduction

In Iacobucci, Posavac, Kardes, Schneider, and Popovich (2015), we had documented the enormous popularity of median splits, in consumer research, psychology, and numerous other fields. We had acknowledged the traditional concerns regarding median splits regarding the loss of information and resulting power. More importantly, we sought to investigate the extent to which the more recently expressed concern about median splits held true, that using median splits may give rise to Type I errors. Our approach was more comprehensive than that of the literature to date because we designed full simulation studies rather than relying on an anecdotal data set.

We found that in the presence of multicollinearity, median splits could indeed result in Type I errors, though the effects were often negligible. The results of our studies were clean and unambiguous; in the absence of multicollinearity, median splits do not create misleading results. We made it clear that the findings were not attributable to the use of an ANOVA vs. the regression model, but rather due to the presence or absence of multicollinearity. If a researcher is running an experiment, such as a typical factorial (or other orthogonal design), then letting a median split serve as a factor is completely legitimate.

In our Discussion section, we mentioned that median splits were not likely to have caused problems in published articles and we explained why. We also explained that our statistical results hold for naturally occurring or experimenter-created groups. We demonstrated that our results held even in the presence of extremely non-normal distributions (e.g., quadratic,

☆ The authors are grateful to the Editor, the original submission Area Editor and the Research Dialog Area Editor, and the teams of commentators for their respective roles in this Research Dialog.
* Corresponding author.
E-mail addresses: Dawn.Iacobucci@owen.vanderbilt.edu (D. Iacobucci), Steve.Posavac@owen.vanderbilt.edu (S.S. Posavac), Frank.Kardes@uc.edu (F.R. Kardes), mattschneids@gmail.com (M.J. Schneider), Deidre.Popovich@ttu.edu (D.L. Popovich).

http://dx.doi.org/10.1016/j.jcps.2015.06.014
1057-7408/

natural log, bimodal, and uniform). We also entertained the notion of two median splits in a single study, that while such a practice might not seem advisable, in truth, it may well be less problematic than one might first think.

Finally, in our paper, we stated repeatedly and quite clearly that we were not intending to persuade researchers who like their continuous variables to begin dichotomizing. Rather, our study provides support for researchers who like working with median splits due to the beauty of their parsimony, and the ease with which they may be communicated. The findings of Iacobucci et al. (2015) support those researchers in their preferences for median splits.

Although we suspect that the commentaries as a whole would have added more value if a psychologist or consumer researcher favorable to median splits wrote one of the commentaries, the upside of our having received commentaries written by two teams with a track record of opposition to median splits is that readers can be confident that any possible objection to our results has been generated. Thus, taken together, we are delighted with the commentaries by Rucker, McShane, and Preacher (2015) and McClelland, Lynch, Irwin, Spiller, and Fitzsimons (2015), and this opportunity to clarify and reify the fact that median splits are a perfectly valid, and extremely useful analytical tool for researchers. The commentaries offer a range of opinions, from concepts that are absolutely correct on one hand (e.g., Rucker et al.'s points that, all other things being equal, Type I and Type II errors rise and fall in opposition, and that regressions in and of themselves do not support causal statements, or McClelland et al.'s remarks that a median split is known to reduce power), to erroneous on the other (e.g., McClelland et al.'s claim that our simulations were incomplete or incorrect with technical errors, and their false equivalence logical fallacy when appealing to the ESP literature). The styles of the two commentaries are rather different, with the first being deeper and focused, whereas the second is broader. We address the arguments in each commentary in turn, concurring and clarifying as appropriate.

Commentary by Rucker, McShane, and Preacher

Rucker et al. (2015) offer a number of well thought-out arguments about the treatment of continuous and median split variables, and we feel that when readers compare their perspective with ours, our goal of moving the field toward a more nuanced understanding of median splits is greatly facilitated. Recently, the field had been told to reject median splits because of concerns regarding Type I error based on an overly broad conclusion derived from a highly artificial and constructed data set. The main purpose in our paper was to show that concerns with Type I error are, in fact, bounded within certain methodological contexts. In perhaps the most common experimental scenario in which median splits are used, wherein one factor is an experimental manipulation and the other factor is a median split, our paper shows that Type I error is not increased by median splits. Rucker et al. seem to concede this point, but nevertheless having concerns with median splits, change the focus of the debate from Type I to Type II error. Although we see some of the issues raised in Rucker et al.'s commentary differently, we feel that they make a number of well-reasoned arguments regarding Type II error that help to increase the sophistication of the discussion regarding median splits.

Type I and Type II errors

If this discussion is to revolve around Type I and Type II errors, let us review the basics. Fig. 1 depicts two normal distributions. The distribution at the left, drawn in a solid line, is the distribution around the null hypothesis population mean, $\mu = 0$, the one being tested. The distribution at the right, drawn in a dashed line, is a distribution around a different population mean, $\mu = 1.5$. In the left-hand distribution, the critical regions are drawn at ± 1.96 for a Type I error rate of $\alpha = 0.05$ in a two-tailed test. If the null hypothesis is true and a calculated z exceeds ± 1.96, the researcher would make a Type I error. Type II errors reflect the opposites—the opposite reality and the opposite decision. If the null hypothesis is not true, but z falls short of ± 1.96, the researcher does not reject the incorrect null, committing a Type II error, the likelihood of which is depicted by the shaded area labeled β. Recall the standard label of the probability of committing a Type I error is α, and that for a Type II error is β.

Students of statistics are taught that there is an inverse relationship between Type I errors and Type II errors. It is not a simple relationship, as if α and β sum to some constant value, in part because a Type I error can only occur if the null hypothesis is true, and a Type II error can only occur if the null hypothesis is false, and of course these conditions cannot both hold simultaneously. In Fig. 2, we depict the two distributions with the use of a more conservative $\alpha = 0.01$ In changing critical values from 1.96 (for $\alpha = 0.05$) to 2.58 (for $\alpha = 0.01$), the Type I error probability has decreased. Note that the Type II error, the size of the area under the curve labeled β has increased in Fig. 2 compared with Fig. 1. Figs. 1 and 2 illustrate how the relationship between α and β holds, that as an α-level decreases, the β probability increases. (Conversely, as α increases, say from 0.05 to 0.10, then the likelihood of a Type II error, β, decreases.)

Given that basic frame, let us now add the notion of power to the mix, recall it to be the likelihood of rejecting the

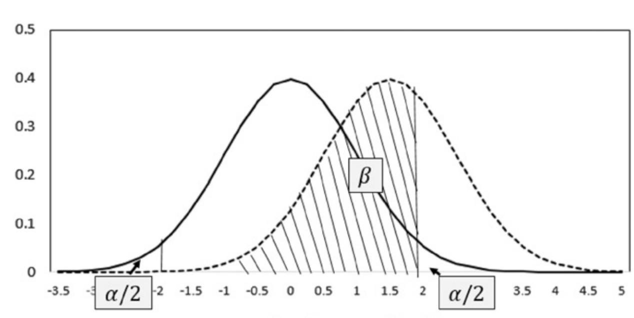

Fig. 1. Standard normal distribution, $\alpha = 0.05$, critical $z = 1.96$.