TITLE

Research Report

Marlyne Bosman

November 23, 2017

Abstract

Outliers seriously affect conclusions drawn from analysis of variance (ANOVA). A way of handling the adverse affect of outliers is to use robust measures, that is, measures of central tendency and spread that are relatively unaffected by slight changes in a distribution. The current research aims to investigate the extent of the effect of outliers on the approximate adjusted fractional Bayes factor (AAFBF), developed for the evaluation of informative hypotheses in virtually any statistical model by Gu et al. (2017). Additionally, it is researched to what extent replacing ordinary least squares estimates as input for the AAFBF with robust estimates results in a decreased effect of outliers.

1 Introduction

Analysis of variance (ANOVA) is a statistical approach for comparing means that is used by many researchers. Just as any statistical approach, ANOVA is based on certain assumptions. When the data meet these assumptions, ANOVA can be validly applied. In reality, however, data often violates assumptions. For example, sampling can result in an outlier, i.e. an observation that “deviates so much from other observations as to arouse suspicious that it was generated by a different mechanism” (Hawkins, 1980, p. 1).

Unfortunately, even a small proportion of outliers can seriously affect an ANOVA. Firstly, outliers result in an increased error variance, thereby leading to a reduced power of the statistical test (Wilcox, 2017). Secondly, outliers cause biased parameter estimates (Ruckstuhl, 2014; Wilcox, 2017). Therefore, when an ANOVA is applied to a dataset that contains outliers inferences can be highly inaccurate.

One obvious way of handling the adverse effects of outliers is to remove them from the dataset prior to analysis. However, whether the removal of outliers is advised depends on the source of the outliers. Outliers are often divided in those that are the results of errors in the data and those that come from natural variability (Anscombe, 1960). Outliers that are the result from natural variability can be considered legitimate cases (Osborne & Overbay, 2004). Additionally, sometimes the source of outliers can not be traced back. Ideally, one wants to keep the legitimate cases and outliers for which the source is unknown in the data while at the same time minimizing their influence on estimation and hypothesis testing. One way to achieve that objective is to use robust measures.

Robust measures are measures of central tendency and spread that are relatively unaffected by slight changes in a distribution. (Wilcox, 2017, p.25). Usually, the central tendency and spread of a distribution are measured by the population mean, μ, and standard deviation, σ, respectively. However, while μ and σ result in inaccurate results in the presence of outliers, robust measures give relatively accurate results (Ruckstuhl, 2014; Wilcox, 2017). A simple example of a robust measure is the median. Unlike the mean, the value of the median gives an accurate estimate of the central tendency of the data unaffected by outliers.

Robust measures are mostly discussed in the context of estimation and null hypothesis significance testing (NHST). Another approach for evaluating hypotheses is a Bayesian model selection approach (Klugkist et al., 2005). This approach uses a Bayes factor (BF) to directly evaluate specific scientific expectations, stated as informative hypotheses. In the context of an ANOVA, an informative hypothesis can be used to state an expected ordering of means, for example,

$$H_1 : \mu_1 < \mu_2 < \mu_3, \tag{1}$$

where μ_g represent the mean of group $g = 1, 2, 3$. With the Bayes factor, the relative support in the data can be calculated for an informative hypothesis, H_i, compared with it's complement, H_c (van Rossum, van de Schoot, & Hoijtink, 2013), an unconstrained hypothesis, H_u, or another informative hypothesis, H_i'. For example, H_1, as stated in Equation 1, can be compared with another informative hypothesis,

$$H_2 : H_1 : \mu_1 < \mu_2 = \mu_3. \tag{2}$$

Finding a BF_{12} of 5 then indicates that the support in the data for hypothesis H_1 is five times larger than the support for hypothesis H_2.

Recently, Gu, Mulder, & Hoijtink (2017) developed the approximate adjusted fractional Bayes factor (AAFBF). With the AAFBF, informative hypotheses can be evaluated for virtually any statistical model. Additionally, the AAFBF is implemented in an easy-to-use software package called Bain. For the calculation, only the parameter estimates of the statistical model at hand and their covariance matrix are needed.

In the ANOVA context, the parameter estimates are the group means. In a regular ANOVA, these are estimated by means of the Ordinary Least

Squares (OLS) estimator. However, as previously stated, parameters that are estimated by regular estimation methods can be seriously affected by outliers in the data. Hence, the expectation is that the AAFBF resulting from these estimates is also negatively affected by outliers. However, to our knowledge, this has never been formally investigated.

This paper aims to investigate to what extent the AAFBF based on the regular OLS estimates is affected by outliers. Additionally, it aims to investigate to what extent replacing the OLS estimates as input for the AAFBF with robust estimates results in a decreased effect of outliers. For this purpose, a simulation study is used to show and compare the effect of outliers on the OLS estimator and an estimator that is relatively unaffected by outliers, i.e. a robust estimator. Subsequently, a simulation study will evaluate and compare the effect of outliers on the AAFBF based on the OLS estimator and on the AAFBF based on the robust estimator.

This paper is organized as follows. Section 2 introduces a robust estimator suitable for the ANOVA context. Furthermore, it qualities are discussed and it is proposed to use the robust estimator for estimation of the parameters needed for Bain. Next, in Section 3 the methods of the simulation study are explained. Subsequently, Section 4 shows the results of the simulation study. Finally, in Section 5 the results, implications and limitations of the research are discussed.

2 Robust estimators

The interest of the present paper is in a robust estimator of the population mean and it's covariance matrix, the parameter estimates of an ANOVA that need to be supplied to Bain for the evaluation of informative hypotheses. Ruckstuhl (2014) and Wilcox (2017) describe various robust measures of central tendency and spread with varying qualities and limitations. From the discussion of robust estimators in Wilcox (2017, Chapter 3) it can be concluded that a robust estimator called the 20% trimmed mean appears to be most suitable for the ANOVA context in terms of small-sample efficiency and accurate coverage probability.

The 20% trimmed mean is a robust measure of central tendency that deals with reducing the effect of outliers by removing 20% of a sample's distribution at both tails. The 20% trimmed mean can be calculated as follows,

$$\mu_t = \frac{1}{1-2\gamma}\int_{x_\gamma}^{x_{1-\gamma}} x dF(x), \tag{3}$$

where $\gamma = 0.2$ is the amount of trimming and x_γ is the

Besides evaluating the performance of the 20% trimmed mean in terms of efficiency and coverage probability, as done in Wilcox (2017, Chapter 3), it's degree of resistance to outliers, i.e. robustness, can be evaluated. Robustness of estimators is evaluated with two measures: the influence function and the breakdown point. As explained by Ruckstuhl (2014), the influence function of an estimator can be seen as a measure of local reliability of an estimator: it measures the influence of the size of a single y-value. Conversely, the breakdown-point can be seen as a measure of global reliability: it measures the minimal proportion of outliers for which an estimator gives completely implausible values (Ruckstuhl, 2014). The two concepts of the influence function and breakdown point will be discussed in more detail below.

Influence function

Imagine a dataset consisting of collected scores on some dependent variable for one group with sample size, $n = 65$, and normally distributed errors for which $\mu = 0.0$ and $\sigma = 1.0$. In Figure 1, the empirical influence functions of the mean and the 20% trimmed mean are shown for the described sample. The empirical influence function shows the effect on the value of the mean estimate of adding an y-value to the dataset with an arbitrary large value. As can be seen in Figure 1, the value of the mean increases without bounds if an increasingly outlying y-value is added to the data. Meanwhile, the 20% trimmed mean has a bounded influence function: additional y-values that are too extreme are trimmed and cannot influence the estimated parameter.

The illustration in Figure 1 demonstrates the concept of infinitesimal robustness. An estimator is said to have infinitesimal robustness if it's influence function is bounded. From Figure 1 it can be concluded that the mean does not have infinitesimal robustness, but the 20% trimmed mean does.

Breakdown point

The 20% trimmed mean as input for Bain

3 Methods

4 Results

5 Discussion

Conclusion

References

Anscombe, F. J. (1960). Rejection of outliers. *Technometrics*, *2*(2), 123–146. doi: 10.1080/00401706.1960.10489888

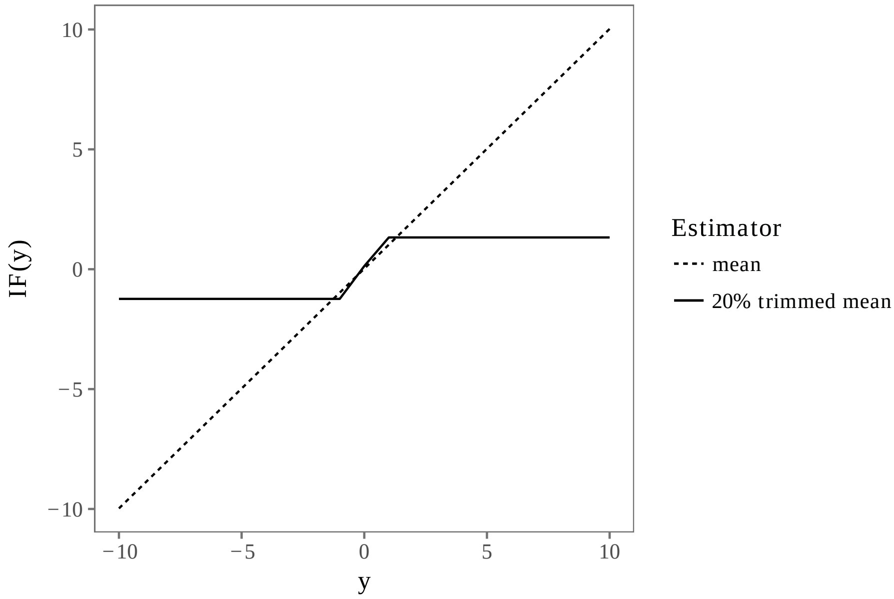

Figure 1: The empirical influence function of the mean and the 20% trimmed mean for a random sample ($n = 65$) drawn from the standard normal distribution. Illustrated is the effect adding an outlier varying between $-10 < y < 10$ on the value of the parameter estimate.

Gu, X., Mulder, J., & Hoijtink, H. (2017). Approximated adjusted fractional Bayes factors : A general method for testing informative hypotheses . (24), 1–43.

Hawkins, D. M. (1980). *Identification of outliers.* Springer Netherlands. doi: 10.1007/978-94-015-3994-4

Klugkist, I., Laudy, O., & Hoijtink, H. (2005). Inequality constrained analysis of variance: A bayesian approach. *Psychological Methods*, *10*(4), 477–493. doi: 10.1037/1082-989x.10.4.477

Osborne, J. W., & Overbay, A. (2004). The power of outliers (and why researchers should always check for them). *Practical assessment, research & evaluation*, *9*(6), 1–12. Retrieved from `http://pareonline.net/htm/v9n6.htm`

Ruckstuhl, A. (2014). Robust fitting of parametric models based on m-estimation. *Lecture notes.[Available at https://stat. ethz. ch/wbl/wbl4/WBL4_robstat14E. pdf.]*.

van Rossum, M., van de Schoot, R., & Hoijtink, H. (2013). “is the hypothesis correct” or “is it not”. *Methodology*, *9*(1), 13–22. doi: 10.1027/1614-2241/a000050

Wilcox, R. R. (2017). *Introduction to robust estimation and hypothesis testing* (4th ed.). New York: Academic press.