Comparing the Type 1 Error Rates of ANOVA When Doing Preliminary Normality Testing

by

Marcelo Munoz
May 2022

A thesis submitted in partial fulfillment of the requirements for a
baccalaurate degree in *Computer Science*
in cursu honorum

Reviewed and approved by

———————————————

Dr. Rebecca Conley

Thesis Supervisor

Submitted to

The Honor's Program, Saint Peter's University

DATE, 2022

Dedications

you can write a dedication

Acknowledgements

you can write acknowledgements

Abstract

Abstract will go here

Contents

Chapter 1 - Introduction and Literature Review

Introduction

Literature Review

Is It Really Robust? Reinvestigating the Robustness of Anova Against Violations of the Normal Distribution Assumption

Overall, the efforts of this paper place attempts at investigating one-way analysis of variance tests against "deviations from the assumption of a normally distributed dependent variable." Several claims are made throughout the paper, they assume that violations of the ANOVA test possibly lead to an increase of the rate of type I and type II errors. Furthermore, the suggestion is made that the use of nonparametric tests is not necessary as they may be less precise, lower power, and inaccurate when compared to the robustness of the ANOVA when facing multiple violations. However, doubt is placed in these previous findings with concers towards the quality of their samples, and the quality of the random number generators they used.

The paper utilizes a "univariate, one-way experimental design with three groups (n =25)" among a normal, rectangular, and exponential distribution. Overall, although the study places emphasis on high-quality samples, there is lack of insight and some questions concerning why the authors chose to only use the 10% best samples when accounting for violations.

Paper 2

To test or not to test: Preliminary assessment of normality when comparing two independent samples

This is how you cite a paper [2] and another citation [1].

Chapter 2 - Methodology

The programming language R is utilized as a primary tool of generating simulations throughout this research for its tendency to focus on statistical capabilities. The initial iterations of these simulations focused on the creation of random samples and creating a "for loop" which will account for the ANOVA tests done on these samples. Then, an "if statement" is used to evaluate the p-value and determine if it is less than or equal to 0.05. In order to demonstrate the consistency of ANOVA tests among normal distributions, the amount of p-values that meet the previously mentioned criteria is printed at the end of the function. Additional functionality was created which provided simulations for mix-normal and uniform distributions. The first step is to create the samples followed by a Shapiro-Wilkes test which introduces the necessary p-values required to conduct the initial analysis. The minimum p-value from the sample set is then assessed to see if it is less than or equal to 0.05 which then results in a Kruskal-Wallis test if the statement is true. Otherwise, the function then undergoes a one-way ANOVA test. The p-value of the test that is conducted throughout the function is then evuluated at printed at the end to give an overview of the error rate over the number of trials. A seperate algorithm was also created with a focus on the uniform distribution which follows the same steps as previously mentioned.

Chapter 3 - Results
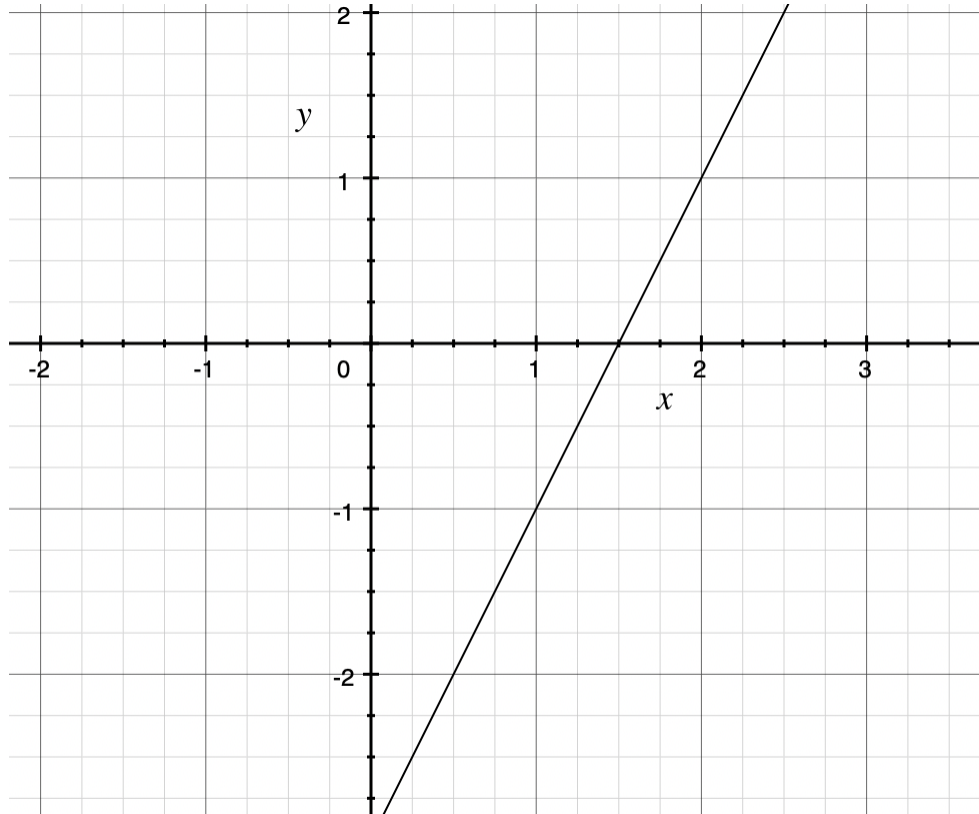
This is how you insert a figure:

Figure 1: Write a caption for the figure

Then you can label the figure and use Insert>Cross-reference to refer to it, see Figure 1.

Chapter 4 - More Results, if you need another chapter

Chapter 6 - Conclusion

Write a conclusion and future work

Work Cited

References

[1] James V Bradley. Robustness? *British Journal of Mathematical and Statistical Psychology*, 31(2):144–152, 1978.

[2] Justine Rochon, Matthias Gondan, and Meinhard Kieser. To test or not to test: Preliminary assessment of normality when comparing two independent samples. *BMC medical research methodology*, 12(1):81, 2012.