Web Visualization: Power of a Two-Sample *t*-Test

Instructor Guide

This activity is based on an interactive online resource that helps students get acquainted with the concept of power in hypothesis testing, in the context of a two-sample *t*-test on the null hypothesis of the equality of means.

Power can be a difficult concept for introductory statistics students to understand. In order to aid students in understanding this concept, we use an online interactive resource created by Michael Whitlock at the University of British Columbia. The resource can be accessed at

`https://shiney.zoology.ubc.ca/whitlock/RobustnessOfT/.`

This resource allows the user to modify the means and standard deviations of the two populations, as well as the size of the sample obtained from each. It allows students to visualize how changing these values affects the appearance of the two population distributions and the distribution of the two-sample *t*-statistic, and allows students to anticipate what effect these changes will have on the power of the test.

Additionally, research suggests that how students engage with such resources can impact the learning obtained, as without structure to their interactions with a visualization tool learners may miss important concepts[1]. Engaging with this activity will allow students to get more out of this resource, and will ensure that the most important concepts are being reinforced.

This activity can be used as an introduction to the power of a hypothesis test, or as a means to develop further understanding and intuition after the concept has been presented in the course. However, students should have an introductory knowledge of hypothesis tests in general, including the basic concepts behind the two-sample *t*-test. They should be familiar with the role of the null hypothesis and alternative hypothesis, as well as the interpretation of a type I error rate.

The activity is appropriate for all class sizes, and can be implemented in a wide variety of ways: as a lab activity, pre-class assignment, in-class activity, or homework assignment. Because it is quite long, the instructor may wish to split it into multiple parts. The activity also works well as a group assignment.

The activity consists of the following parts:

- In Part I of the activity, students are introduced to the frequency curves in the top right of the resource. These curves show the distributions of the variable of interest in the two populations. Students are encouraged to use the controls on the left side of the resource to

[1] Lane, D. M., & Peres, S. C. (2006). Interactive simulations in the teaching of statistics: Promise and pitfalls. In *Proceedings of the Seventh International Conference on Teaching Statistics*. International Statistical Institute Voorburg, The Netherlands.

change the values of the parameters for the two populations, and observe how this affects the frequency curves. They are encouraged to anticipate which of the situations will be likely to result in samples that will look more different from each other.

- Part II focuses on understanding the histogram at the bottom right of the resource. This histogram gives the results of 2,000 simulations. For each simulation run, a sample is taken from each of the two populations and the resulting two-sample *t*-statistic is calculated. The histogram shows the empirical distribution of the resulting 2,000 two-sample *t*-statistics. Light blue denotes the *t*-statistics that *would not* lead to rejecting the null hypothesis of equal means, while dark blue denotes the *t*-statistics that *would* lead to rejecting the null hypothesis. Getting acquainted with this histogram will be important for helping students answer more technical questions about the power in Part III.

- In Part III, students learn about the different factors that can affect the power of a two-sample *t*-test: the true difference in means, the values of the two population standard deviations, and the sizes of the samples obtained from the two populations. In each question, students are asked to anticipate what will happen to the power as a result of changing one of the above factors. They are then asked to try this out using the resource, and make conclusions based on what they have observed.

- The activity ends with a problem motivated by a biology research paper. The problem focuses on determining an optimal sample size for obtaining a certain desired value of the power. Students are faced with several realistic problem scenarios involving sample size decisions, and are encouraged to use the resource and what they have learned so far to answer these questions.

It must be noted that the power values displayed in the online resource are obtained through simulation, rather than through theoretical distributions. As a result, the values may differ slightly from student to student, or even for the same student when repeating the task. Although the instructions in the activity already make this clear to the students, it may be worthwhile for the instructor to mention it as well.