StatPlayground: Exploring Statistics through Visualizations

Krishna Subramanian
RWTH Aachen University
52062 Aachen, Germany
krishna@cs.rwth-aachen.de

Jan Borchers
RWTH Aachen University
52062 Aachen, Germany
borchers@cs.rwth-aachen.de

CHI'17 Extended Abstracts, May 6–11, 2017, Denver, CO, USA.
ACM ISBN 978-1-4503-4656-6/17/05.
http://dx.doi.org/10.1145/3027063.3052970

Abstract
Statistical analysis is a crucial part of many research fields: it is used by the researcher to validate her hypothesis and to communicate her findings to the community. Null Hypothesis Significance Testing (NHST), a commonly used statistical analysis approach in many research fields, has been criticized over the years due to the prevalence of misconceptions and improper practice. We introduce StatPlayground, an exploratory tool as a viable solution to the root problem of statistical illiteracy. StatPlayground allows users to control data characteristics (e.g., mean, variance of distributions) by directly manipulating visualizations (e.g., box plots) to see the effect on the resulting inferential statistics (e.g., p-value, effect size) and vice versa. We believe that StatPlayground has the potential to help users improve certain statistical literacy skills. We elaborate on the motivation behind this tool and demonstrate its features through use cases.

Author Keywords
Statistical Analysis; Simulation; Exploratory Learning; Data Visualization

ACM Classification Keywords
G.3 [Probability and Statistics]: Statistical software; H.5.2 [User Interfaces]: Graphical user interfaces (GUI).

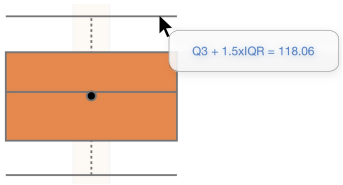

Figure 1: Hovering over the visualization reveals details.

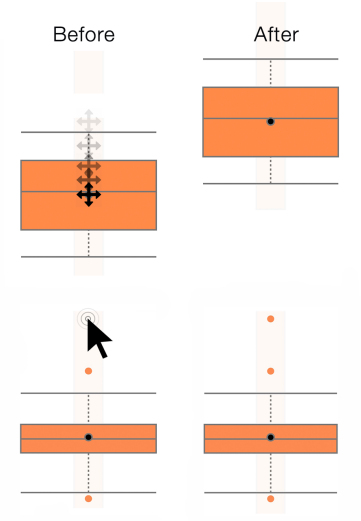

Figure 2: Users can directly manipulate the data: click-and-drag to change the mean (top) and click to create a data point (bottom).

Introduction

NHST is an important part of many research fields such as HCI [1], psychology [11], and medicine [13]. However, the practice is strewn with several misconceptions and issues [1, 4, 18]. An underlying cause of this problem is inadequate statistical literacy [1].

To resolve this problem at a pedagogical level, researchers have identified commonly misunderstood statistical topics [15] and proposed principles for improving teaching [6]. Researchers have also made use of simulations (for particular abstract concepts e.g., confidence intervals) to improve the learning experience [10].

We introduce StatPlayground, an exploratory tool, as a potential solution. StatPlayground performs statistical analysis automatically. It does this by checking the statistical assumptions (e.g., normality, homogeneity of variances) and, based on which assumptions are satisfied, choosing the appropriate test. While automation of the statistical analysis process is not novel [16], StatPlayground also gives the user the ability to explore statistical practice by 1) changing the data properties to see how it impacts other data properties and the computed inferential statistics and 2) changing the computed inferential statistics to see how other data properties change. We believe that such exploration is novel and worth investigating.

StatPlayground is not intended to replace traditional instruction-based learning. We believe it can improve certain statistical literacy skills: the ability to make sense of statistical information [5], identify the relationships between statistical concepts [8], have better data awareness [12], and understand statistical procedures better [7]. StatPlayground uses many concepts shown to improve learning: active discovery learning [9, 3], animated transitions [14, 17], and simulations [2, 10].

Figure 3: Non-linear sliders with intuitive tick marks.

System Description

Overview

StatPlayground allows the user to

1. Control data properties such as mean, median, outliers, and shape of the distribution by interactive visualizations and see the effect on the computed statistics such as the statistical test, effect size, and p-value.
2. Change the computed inferential statistics and view the changes in the data that caused this change.

StatPlayground consists of a dataset selection screen where the user can select one of the predefined datasets. These datasets are representative of typical HCI experiments e.g., a keyboard layout comparison study. We prevent the user from using custom datasets in order to avoid misappropriation of the tool (e.g., data tampering).

Use Cases

The following use cases illuminate the benefits of StatPlayground.

1. StatPlayground currently shows box plots of three distributions, all of which are distributed normally and have similar variances. A one-way ANOVA test has been performed and the statistics are displayed (Fig. 4).

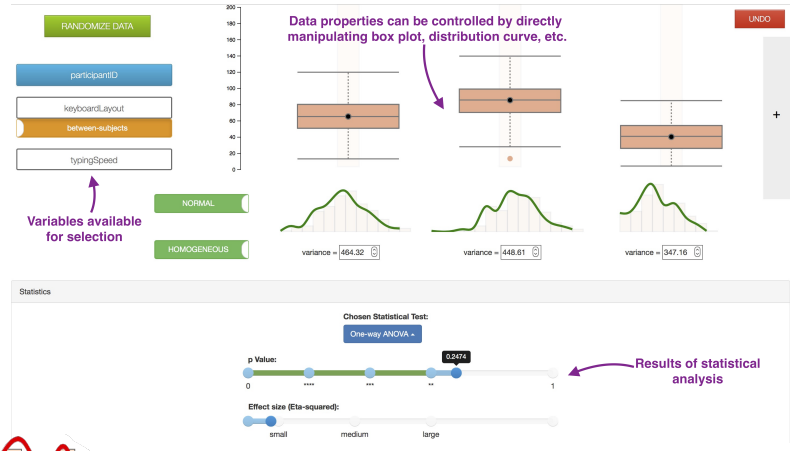

Figure 4: The main screen of StatPlayground. (Annotations are made in purple.)

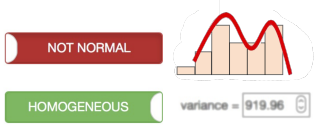

Figure 5: The indicators for whether assumptions have been satisfied (green: satisfied; red: violated) are accompanied by interactive visualizations.

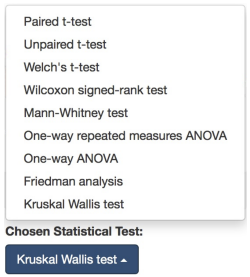

Figure 6: The drop-up menu used to choose a statistical test.

When the user clicks beyond the whiskers of a box plot, she creates an outlier (Fig. 2). The distribution curve changes to indicate that it is no longer distributed normally. StatPlayground now performs a Kruskal Wallis test and updates the reported statistics.

2. StatPlayground displays box plots of two distributions, both distributed normally and having similar variances. The distributions represent two levels of a between-subjects factor. An unpaired t-test has been performed and the statistics are displayed.

   When the user clicks on the toggle button (see Fig. 4) to set the independent variable as a within-subjects factor, StatPlayground performs an unpaired t-test and updates the reported statistics.

3. StatPlayground displays box plots of two distributions which have a Cohen's d effect size value of 0.2 (indicates a small effect size).

   When the user changes the Cohen's d value to 0.8 (indicates a large effect size) using the slider (Fig. 3), the box plots move farther apart to comply with the modified statistic.

4. StatPlayground displays box plots of two distributions, both of which are not distributed normally. A Mann Whitney test has been performed and the statistics are displayed.

   When the user changes the statistical test (using the drop-up menu shown in Fig. 6) to an unpaired t-test, both distributions get transformed into normal distributions.

Limitations & Future Work

StatPlayground currently supports many statistical tests commonly used in HCI (Fig. 6). It has a Model-View-Controller (MVC) architecture and can be easily extended to support multiple independent variables, Bayesian inference, etc. When the user changes the computed statistics to investigate the changes in data properties, StatPlayground chooses one possible dataset (and properties) from an infinite set of datasets. In future, we plan to expand StatPlayground to show many possible distributions. We also plan to allow a more fine-grained control of data properties e.g., the user could 'pin' certain data properties or computed statistics as constants and modify the other factors.

Relevance to CHI 2017

Statistical illiteracy is an important issue that concerns the CHI community. Over the years, there has been an increase in statistical awareness through publications, work-

shops[1], and special interest groups[2]. We believe StatPlayground offers a fun, intuitive, and engaging way to explore statistics and hope to gain insights from the knowledgeable CHI audience.

References

[1] Paul Cairns. 2007. HCI... Not As It Should Be: Inferential Statistics in HCI Research. In *Proc. BCS-HCI '07*. British Computer Society, 195–201.

[2] Beth Chance, Dani Ben-Zvi, Joan Garfield, and Elsa Medina. 2007. The Role of Technology in Improving Student Learning of Statistics. *Technology Innovations in Statistics Education* 1, 1 (2007).

[3] Martin A Conway, Gillian Cohen, and Nicola Stanhope. 1991. On the Very Long-Term Retention of Knowledge Acquired Through Formal Education: Twelve Years of Cognitive Psychology. *Journal of Experimental Psychology: General* 120, 4 (1991), 395.

[4] Mark D Dunlop and Mark Baillie. 2009. Paper Rejected (p> 0.05): An Introduction to the Debate on Appropriateness of Null-Hypothesis Testing. *IJMHCI '09* 1, 3 (2009), 86–93.

[5] Iddo Gal. 2002. Adults' Statistical Literacy: Meanings, Components, Responsibilities. *International Statistical Review* 70, 1 (2002), 1–25.

[6] Joan Garfield. 1995. How Students Learn Statistics. *International Statistical Review* 63, 1 (April 1995), 25.

[7] Joan Garfield. 2002. The Challenge of Developing Statistical Reasoning. *Journal of Statistics Education* 10, 3 (2002), 58–69.

[8] Joan Garfield and Dani Ben-Zvi. 2007. How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics. *International Statistical Review* 75, 3 (2007), 372–396.

[9] G Hermann. 1969. Learning by Discovery: A Critical Review of Studies. *The Journal of Experimental Education* 38, 1 (1969), 58–72.

[10] David M Lane and S Camille Peres. 2006. Interactive Simulations in the Teaching of Statistics: Promise and Pitfalls. In *Proc. ICOTS-7*.

[11] Raymond S Nickerson. 2000. Null Hypothesis Significance Testing: A Review of an Old and Continuing Controversy. *Psychological methods* 5, 2 (2000), 241.

[12] Deborah J Rumsey. 2002. Statistical Literacy as a Goal for Introductory Statistics Courses. *Journal of Statistics Education* 10, 3 (2002), 6–13.

[13] Luis Carlos Silva-Ayçaguer, Patricio Suárez-Gil, and Ana Fernández-Somoano. 2010. The Null Hypothesis Significance Test in Health Sciences Research (1995-2006): Statistical Analysis and Interpretation. *BMC medical research methodology* 10, 1 (2010), 1.

[14] Barbara Tversky, Julie Bauer Morrison, and Mireille Betrancourt. 2002. Animation: Can it Facilitate? *International Journal of Human-Computer Studies* 57, 4 (2002), 247–262.

[15] Jessica Utts. 2003. What Educated Citizens Should Know About Statistics and Probability. *The American Statistician* 57, 2 (May 2003), 74–79.

[16] Chat Wacharamanotham, Krishna Subramanian, Sarah Theres Völkel, and Jan Borchers. 2015. Statsplorer: Guiding Novices in Statistical Analysis. In *Proc. CHI '15*. ACM, 2693–2702.

[17] Martin Wattenberg and Jesse Kriss. 2006. Designing for Social Data Analysis. *TVCG '06* 12, 4 (2006), 549–557.

[18] Alain F Zuur, Elena N Ieno, and Chris S Elphick. 2010. A Protocol for Data Exploration to Avoid Common Statistical Problems. *Methods in Ecology and Evolution* 1, 1 (2010), 3–14.

[1] http://transparentstatistics.org/chi2017/

[2] http://transparentstatistics.org/chi2016/