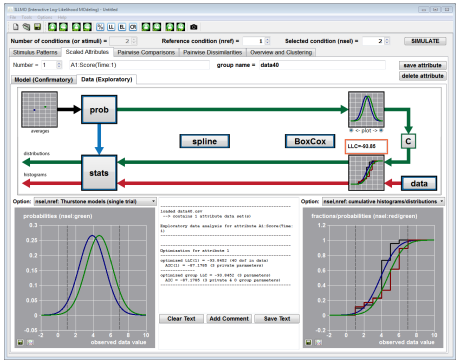

Figure 1: The *graphical user interface* of the ILLMO program, where the image in the lower-right shows two observed (cumulative) histograms (the staircase functions in black and red) together with the (cumulative) Gaussian distributions (the smooth functions in blue and green) that are used to model them. The image in the lower-left shows the corresponding regular Gaussian distributions. The block diagram at the top illustrates the statistical modeling process and contains interaction buttons to open dialog windows with which model characteristics can be changed (e.g., the "prob" dialog allows to change from Gaussian to logistic distributions) or specific statistical analyses can be activated (in "stats").

Insights in Experimental Data through Intuitive and Interactive Statistics

Jean-Bernard Martens
Eindhoven University of Technology
Department of Industrial Design
Eindhoven, the Netherlands
j.b.o.s.martens@tue.nl

ABSTRACT

It is not unusual for empirical scientists, who are often not specialists in statistics, to have only limited trust in the statistical analyses that they apply to their data. The claim of this course is that an improved human-computer interaction with statistical methods can be accomplished by providing a simple mental model of what statistics does, and to support this model through well-chosen visualizations and interactive exploration. In order to support this proposed approach, an entirely new program for performing interactive statistics, called ILLMO, was developed. This course will use examples of frequent statistical tasks such as hypothesis testing, linear regression and clustering to introduce the key concepts underlying intuitive and interactive statistics.

CCS CONCEPTS

• **Mathematics of computing → Probability and statistics**;

CHI'19 Extended Abstracts, May 4–9, 2019, Glasgow, Scotland Uk

ACM ISBN 978-1-4503-5971-9/19/05.
https://doi.org/10.1145/3290607.3298806

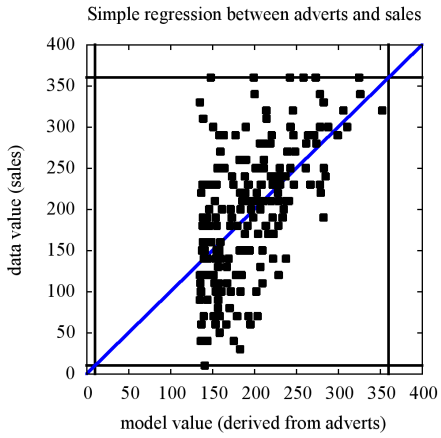

Figure 2: *Linear Regression*: Example of a scatterplot produced by ILLMO, showing the model values, as derived from an independent variable by means of linear regression, against the data values observed for a dependent variable. The figure shows that the implicit assumption in linear regression of an error variance that is independent of the observed value is not valid, as this variance increases with the observed value in this example.

KEYWORDS

Interactive statistics, Likert scales, hypothesis testing, confirmatory statistics (linear regression) and exploratory statistics (clustering).

ACM Reference Format:

Jean-Bernard Martens. 2019. Insights in Experimental Data through Intuitive and Interactive Statistics . In *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI'19 Extended Abstracts), May 4–9, 2019, Glasgow, Scotland Uk.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3290607.3298806

INTRODUCTION AND MENTAL MODEL

Empirical researchers turn to statistics to assist them in drawing conclusions, also called inferences, from their collected data. The data is often experimental, i.e., it consists of (repeated) measurements collected in one or more distinct conditions (e.g., corresponding to alternative interaction devices or software systems used to perform a task). The observed data can be summarized into histograms that specify how often measured values (such as subjective scores on a Likert scale) occur in the distinct conditions.

The goal of statistical analysis can be reformulated as characterizing or modeling the change in histograms across conditions. While existing statistical programs (such as SPSS or R) offer a wide range of methods for studying and characterizing such changes, they rely on a familiarity with statistical methods and terminology. They also offer little or no insight into how statistical methods work and into the assumptions that they make. As a rule, expert knowledge is required to change model assumptions, provided it is even possible at all. The lack of insight and exploration of alternatives can lead to erroneous use of statistical methods, as argued more extensively in [2]. A consequence is that many researchers, who are not expert statisticians, often do not feel very confident about (the interpretation of) their statistical outcomes.

In order to validate our alternative approach towards statistics, we developed a new program for interactive statistics, called ILLMO (Interactive Log-Likelihood MOdeling) [3–5]. This software, the interface of which is shown in Figure 1, has been made freely and publicly available, together with supporting material, through the website mentioned at the end of this paper. The goal of this course is to introduce the mental model and operational procedures underlying interactive statistics by means of examples that relate to statistical problems that arise frequently in CHI.

SPECIFIC BENEFITS

This course will focus on three frequent statistical tasks, i.e., hypothesis testing (between two conditions), linear regression (as an example of confirmatory data analysis) and clustering (as an example of exploratory data analysis). Participants will learn how to perform these tasks within the ILLMO

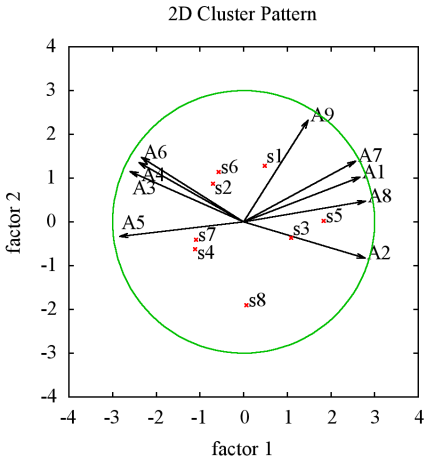

Figure 3: *Cluster Anaysis*: Example of a graph produced by ILLMO, showing 8 different products (i.e., activity bracelets), represented as points labeled s1-s8, being assessed by 5 subjects on a 7-point Likert scale in order to express their agreement with statements: this bracelet can be considered as Jewelry (A1), Feminine (A2), Masculine (A3), Unisex (A4), Sportive (A5), Tough (A6), Luxurious (A7), Graceful (A8) or Easy to Combine (A9). The nine attributes are represented by the vectors labeled A1-A9. The length of a vector is proportional to the squared correlation between the observed attribute data, averaged over subjects, and the 2D model shown in the figure.

program and how to compare the outcomes with those of more traditional methods (such as a T-test or ANOVA). They will also be introduced to a new and powerful approach to statistical modeling called multi-model comparisons [1].

For example, in the case of hypothesis testing, the output of a multi-model comparison can be the likelihoods of two models, one in which the average is the same in both conditions, and one in which the averages differ between conditions. Only in case the ratio between these likelihoods is high (>20) can we conclude that one model is clearly preferred over the other. In the case of (multiple) regression, the output of a multi-model comparison can be the likelihoods of regression models with different (numbers of) independent variables. Such regression models are used to predict the average outcome across conditions of an independent variable. Some examples of visual outputs that can be generated by ILLMO while performing linear regression or clustering are shown in Figures 2 and 3, respectively.

A frequently occurring issue in the CHI community is how to analyze Likert-scale data in a theoretically sound way. The ILLMO program uses an approach proposed by Thurstone in the 1920s, and illustrated in Figure 4. Thurstone modeling allows to analyze discrete data in a way that is very similar to the parametric statistics that most CHI researchers are familiar with for continuous data (e.g., a T-test or ANOVA relies on the modeling of observed histograms by parametric distributions such as Gaussian distributions).

ILLMO implements many traditional statistical methods, as not doing so would increase the threshold for adoption of ILLMO by users who are already familiar with such methods. The integration has been accomplished in such a way that novice users, without prior statistical experience, can ignore such methods. Experienced users can however compare the interactive statistical approach with more traditional statistical methods. On top, ILLMO also offers new statistical methods, such as the recently proposed non-parametric method called empirical likelihood [6], to address cases in which the experimental data cannot be modeled adequately using parametric statistics.

INTENDED AUDIENCE AND PREREQUISITES

The course is relevant for CHI researchers interested in quantitative data collection, either for confirmatory or exploratory analysis. The course does not require any advanced mathematical or statistical knowledge, only an understanding of the distinction between an observed histogram and a theoretical probability distribution. The former summarizes the observed fractions, e.g., the number of times heads and tails are observed in an actual coin-flipping experiment. The latter defines the model probabilities, such as the a priori probabilities of observing head or tail.

PRACTICAL WORK

The participants can practice with the ILLMO software on their laptops and use the instruction videos and sample data sets provided via the website to try out some of the analyses by themselves. In case

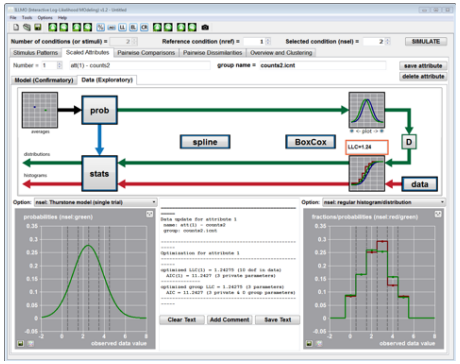

Figure 4: *Thurstone modeling* illustrated in the graphical user interface of the ILLMO program. The image in the lower left shows the parametric Gaussian distribution that is used to model the observed regular histogram in an experimental condition (shown in red in the lower-right image). The areas underneath the green curve in between the dotted lines (at 0.5, 1.5, . . . , 4.5) are the model predictions (shown in green in the lower-right image) for the probabilities of the observed values (of 0, 1, . . . , 5) on a Likert scale.

they have data available in CSV format, they can also try out the analyses on their own experimental data.

INSTRUCTOR BACKGROUND

Jean-Bernard Martens is a full-time professor in the department of Industrial Design at the Eindhoven University of Technology. He has been active in the field of HCI since 2000, and has contributed to topics such as: quantifying interactions in virtual and augmented environments, user experience over time and across subjects, storytelling in design, etc. He is currently focusing on data-driven design. From 1980 to 2000 he worked on image processing and subjective image quality. He graduated in 1979 in electronic engineering, specializing in information and communication theory. A common factor throughout most of these diverse activities has been the statistical modeling of quantitative data.

SOFTWARE AND SUPPORTING MATERIAL

Run-time versions of the ILLMO program for both Mac OS and Microsoft Windows can be downloaded, together with some supporting material, through the project website http://illmoproject.wordpress.com. Different versions of the ILLMO software have been used in teaching statistics to students in Industrial Design at the Eindhoven University of Technology. These experiments *in the wild* have provided valuable feedback for identifying useful system features and for improving the visual appearance of the interface.

REFERENCES

[1] K. Burnham and D. Anderson. 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach.* Springer, New York.

[2] Maurits Kaptein and Judy Robertson. 2012. Rethinking Statistical Analysis Methods for CHI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12).* ACM, New York, NY, USA, 1105–1114. https://doi.org/10.1145/2207676.2208557

[3] Jean-Bernard Martens. 2012. Statistics from an HCI Perspective: Illmo - Interactive Log Likelihood Modeling. In *Proceedings of the International Working Conference on Advanced Visual Interfaces (AVI '12).* ACM, New York, NY, USA, 382–385. https://doi.org/10.1145/2254556.2254629

[4] Jean-Bernard Martens. 2014. Interactive Statistics with Illmo. *ACM Trans. Interact. Intell. Syst.* 4, 1, Article 4 (April 2014), 28 pages. https://doi.org/10.1145/2509108

[5] Jean-Bernard Martens. 2017. Insights into Experimental Data: Interactive Statistics with the ILLMO Program. https://itunes.apple.com/us/book/insights-in-experimental-data/id1210325588?mt=13

[6] A. Owen. 2001. *Empirical Likelihood.* Chapman and Hall, Boca Raton.