CS **598**AK Fall 2021

Homework 5

due by email to alexkirlik@gmail.com by 5pm one week after date assigned

1   Description

In this homework assignment, you will demonstrate the statistical techniques discussed in Experimental Human-Computer Interaction, chapters 5.5.2 - 5.5.2.1.

You are provided with an Excel file ("Homework5.xlsx") which contains the data you will use to perform the analysis and worksheets where you should show your solutions to the below questions. Additionally, you will use the PDF files provided in Homework 1 and 3 (t, q, F, U, Wilcoxon, χ^2 tables).

You are to submit 2 files for grading:

- A filled out "Homework5.xlsx", which contains a worksheet labeled "Question 3" to show your solution and work.

- A PDF file containing explanations for the solutions for the questions below.

You are to base your calculations on chapters 5.5.2 - 5.5.2.1 and use the formulas listed in Appendix A1, A2, and A3. You are NOT to use any statistical packages, such as R or SPSS, nor any statistical add-ons or inferential statistical tools found in Microsoft Excel or Google Sheets. You may use descriptive statistic formulas such as SUM, SUMSQ, AVERAGE, and SUMXMY2.

The purpose of this assignment is to perform the analysis and demonstrate the process by showing your work. You may use Excel, Google Sheets, or other spreadsheet software to complete the assignment, but submission of your analysis must be as an Excel file.

2 Instructions

Facebook is examining how different versions of their News Feed is affecting how long users interact with the News Feed. They create 40 different versions and show them to people and measure how long each user interacts. Their participant pool is comprised of 52% female and 48% male. The age demographics are split as follows, 13% are ages 14-17, 52% ages 18 - 32, 25% ages 33 - 55, and 10% ages 56 and older. Use this prompt to answer question 1 below in your PDF file.

1. If the researchers at Facebook decide to perform factor analysis, what would be the main effects they should examine given the information above? What would be the interaction effects that they should examine? When reporting interaction effects, state whether the interaction effect is a two-way, three-way, four-way, etc. interaction.

Adobe is comparing their Photoshop tool to another tool, GIMP. They have 40 users come to their lab and perform 4 tasks for each tool, so every person completes 8 total tasks. After completing each task, they record the error rate. Use this prompt to answer question 2 below in your PDF file.

2. What are the main effects under observation? Based on the information above, what is one additional factor that should be considered and why?

Google is studying how their 2 emailing platforms, Inbox and Gmail, affect the error rate of users sending emails. They bring 30 participants into their lab and have them complete 3 tasks using each email platform, composing & sending a text only email (TEXT), composing & sending an email with images (IMAGE), and composing & sending an email with an attachment (ATTACH). The Excel sheet shows the data the participants. The researchers already found that a two-way ANOVA revealed a significant interaction between task and email platform. Use this prompt and the Excel sheet to answer question 3 below. Show your work in the Excel sheet and explain your finding in the PDF file.

3. Perform the appropriate statistical tests to report the effects of the emailing platform on error rate. You can use Appendix A3 as a reference. You can assume the data is normal and use a threshold of p $= 0.05$. Reflect closely on what tests to run and what to report.

3 Errata

This section contains errors found in the text of the EHCI book that have been discovered so far. Please account for these errors.

- Table A2.6: The approximation of simple "average" n, the equation used ($\frac{k}{\sum_{t=1}^{k} n_k}$) should be inverted, thus it should be $\frac{\sum_{t=1}^{k} n_k}{k}$.

- Page 223 (Appendix A2): The formula for calculating the χ^2 value is incorrect. The denominator $nk(nk+1)$ should be $nk(k+1)$, remove the extra n in your calculations.

- Page 226 (Appendix A2): The formula for standard error for conditions c and d, the sqrt sign should encompass the entire formula, thus it should be $\sqrt{\frac{N(N+1)}{12}(\frac{1}{n_c} + \frac{1}{n_d})}$. Additionally, though this formula is only for when there are no ties in mean ranking (the book does not indicate this), you should use this formula even if ties are present.