COS 436 / ELE 469
Spring 2013

A4: Quantitative Data Analysis

Due 4/8 at 11:59 PM

Overview

You will analyze data from the in-class keyboard comparison exercise in order to assess the relative performance of the two keyboard layouts. You will also answer a few general questions about hypothesis testing.

You can review the keyboard layout exercise here:
http://www.cs.princeton.edu/courses/archive/spring13/cos436/assignments/A4/Keyboards.pdf

The Assignment

1. Download two data files:

http://www.cs.princeton.edu/courses/archive/spring13/cos436/assignments/A4/Assignment4AllData.xls

This is an Excel spreadsheet with a subset of the data from class, with some preliminary analysis. **Note that this file contains 4 sheets.**

http://www.cs.princeton.edu/courses/archive/spring13/cos436/assignments/A4/WPMValues.txt

This contains only the words-per-minute data from the class exercise, in the format expected by the ANOVA tool you'll be using. In this file, columns from left to right are: L1P1, L1P2, L1P3, L1P4, L1P5, L2P1, L2P2, L2P3, L2P4, L2P5, G, where LnPm is the speed, in words per minute, for that participant to enter phrase m using layout n (where layout 1 = "Method A" / Opti), and G holds the group number (Group 1 did Method A / Opti first).

2. Read Scott MacKenzie's ANOVA discussion at http://www.yorku.ca/mack/Anova2.html. This will help you learn to use his ANOVA software, but it should also provide some helpful information on using ANOVA, in general.

3. Download Scott MacKenzie's free, cross-platform ANOVA tool from
http://www.yorku.ca/mack/anova.zip
You'll be using this tool in the assignment. (Free, cross-platform, multi-purpose ANOVA tools are hard to come by!)

Note on degrees of freedom for F: MacKenzie's tool will output an ANOVA table in nearly the same format as our slides from class (slides 57–59 from Lecture 14). However, his table uses a slightly different set of row labels. In lecture, I told you to look for the rows labeled Error (for between-subjects factors) or Error(*name of source*) to identify the second degrees-of-freedom (df2) parameter for your F-distribution when reporting results. However, instead of labeling a row as "Error" or as "Error(*name-of-source*)", MacKenzie's tool labels these rows as "Participant" or as "*name-of-source*_x_P(Grp)", respectively. So, when you're trying to figure out what to use for reporting df1 and df2 for your F statistics, use the appropriate source degrees-of-freedom for df1, and use the degrees-of-freedom for Participant or *name-of-source*_x_P(Grp) for df2.

A good sanity check: No matter which type of ANOVA you use, make sure that your reported value for F is equal to the value MS1/MS2, where MS1 is the mean square value for the row corresponding to df1, and MS2 is the mean square value for the row corresponding to df2. For example, based on the output below, I would report my test result for the significance of the task type effect as $F(1, 7) = 14.217$, $p < .05$. Note that MS1 = 2745.188, df1 = 1, MS2 = 193.092, df2 = 7, and MS1/MS2 = 14.217 = our value for F, as found in the table.

Source	Sum of Squares / SS	df	Mean Square / MS	*F*	p-value / significance
Task type	2745.188	1	2745.188	14.217	0.007
Error (task type)	1351.646	7	193.092		

4. Check out the free online t-test tool at http://www.graphpad.com/quickcalcs/ttest1.cfm Use this tool if/when you perform a t-test in this assignment.

5. If you don't already know how to do this, read about fitting trendlines to data in Excel at http://office.microsoft.com/en-us/excel-help/add-a-trendline-to-a-chart-HP005198462.aspx

6. Check out the free online tool for finding a 95% confidence interval for Pearson's *r* at http://www.vassarstats.net/rho.html
Use this tool to compute a confidence interval for *r* in the assignment.

7. Download and read "Reporting results of common statistical tests in APA format" from the University of Washington:
http://www.psych.uw.edu/writingcenter/writingguides/pdf/stats.pdf

8. Answer the questions at http://tinyurl.com/cos436A4
You will be able to edit your responses after submitting the form, if you hang onto the URL after submission. You will receive a confirmation email the first time you submit the assignment, but not after you make further edits.

Questions?

Please direct all questions regarding the assignment to Piazza (do not send e-mail). Private Piazza messages are fine.

Acknowledgements

The keyboards tested in this assignment are based on an exercise by Scott MacKenzie.