Pre-Test / Post-Test Report

Dr. Clifton Franklund

Spring 2016

Introduction

This is an R Markdown report that summarizes a pre-test / post-test analysis of student achievement at Ferris State University. This approach to assessment is commonly used to estimate student learning gains over a course. One recent example was a study of Social Statistics course (Delucchi 2014). The actual data for this report were collected using Blackboard tests and analyzed using R. This report template can be used to rapidly create similar reports for additional courses in the future at Ferris or anywhere where Blackboard is used.

Session information

This document was generated using the R statistical language (R Core Team 2013) and the RStudio (RStudio Team 2015) itegrated development environment (IDE). The systems settings used are listed below. You should be able to replicate this report using the .Rmd code and data files with a similarly configured computer.

R version 3.4.0 (2017-04-21) Platform: x86_64-apple-darwin15.6.0 (64-bit) Running under: macOS Sierra 10.12.4

Matrix products: default BLAS: /Library/Frameworks/R.framework/Versions/3.4/Resources/lib/libRblas.0.dylib LAPACK: /Library/Frameworks/R.framework/Versions/3.4/Resources/lib/libRlapack.dylib

attached base packages: [1] stats graphics grDevices utils datasets methods base

other attached packages: [1] xtable_1.8-2 ggplot2_2.2.1

loaded via a namespace (and not attached): [1] Rcpp_0.12.10 digest_0.6.12 rprojroot_1.2 plyr_1.8.4 [5] grid_3.4.0 gtable_0.2.0 backports_1.0.5 magrittr_1.5 [9] evaluate_0.10 scales_0.4.1 stringi_1.1.5 lazyeval_0.2.0 [13] rmarkdown_1.5 tools_3.4.0 stringr_1.2.0 munsell_0.4.3 [17] yaml_2.1.14 compiler_3.4.0 colorspace_1.3-2 htmltools_0.3.6 [21] knitr_1.15.1 tibble_1.3.0

Raw data files

The raw data for this report was aggregated using a Blackboard quiz. A 40-question multiple-choice quiz was given during the first and last weeks of instruction. The questions were randomly drawn from a bank of 123 questions. The students were given five bonus points for attempting the quiz (to encourage participation) and another five bonus points if they scored at or above 75% correct (to encourage intellectual effort). The results of the pre- and post-tests were downloaded from the Blackboard grade center by selecting "download results" from the contextual menu for the appropriate grade columns. The download parameters used were:

1) Tab-delimited
2) By Question and User
3) Only Valid Attempts

The downloaded files were opened in Microsoft Excel (for some reason, R on my Macintosh does not like the UTF-16 files that Blackboard makes). Using Excel, the files were then saved as comma-separated value (.csv) files - Pre-Course.csv and Post-Course.csv. These two files are the raw data necessary to generate this report. The original files are not shared in this project because they contain personally identifiable information about

my students. If you provide your own files, they will be processed in place of my files and the resulting data will be used to create all of the graphs and tables in this document.

Processed (de-identified) data files

An R script, process.R, was written to de-identify and summarize the data from the raw Blackboard files. When executed, this script removes all personally identifiable information and produces two new comma-separated data files - PrePostScores.csv and PrePostQuestions.csv.

PrePostScores.csv is a list containing scores by students for each student that completed both the pre- and post-exams. Rubric scores for the pre- and post-test performances were also generated based upon the following parameters:

Table 1: Conversion percent correct to rubric scores

Rubric	Percent	Performance
0	0 to 49.9%	Unsatisfactory
1	50 to 59.9%	Beginning
2	60 to 69.9%	Progressing
3	70 to 84.9%	Proficient
4	85 to 100%	Advanced

The variables included in the PrePostScores.csv file are:

- Pre — The number of correct responses on the pre-test
- Post — The number of correct responses on the post-test
- preRubric — A rubric score based upon the percent correct on the pre-test
- postRubric — A rubric score based upon the percent correct on the post-test

PrePostQuestions.csv is another list containing information about the class' performance by question. The columns in this file include:

- Question — The prompt text for each question
- Pre — The fraction of students that correctly answered the question on the pre-test
- Pre.N — The number of students that attempted the question on the pre-test
- Post — The fraction of students that correctly answered the question on the post-test
- Post.N — The number of students that attempted the question on the post-test
- Diff — The difference obtained by taking Post minus Pre

Results and Discussion

The overall distribution of scores for the pre- and post-tests are compared in Figure 1. A marked increase in scores for the post-test is observed compared to those of the pre-test. The average score for the pre-test is 19.3, while the post-test average is 27.4 points. This result is statistically significant (t=5.53, df=86.42, p=3.4e-07) as measured by a two-factor non-paired t-test. Pairing was not used so as to calculate a more accurate estimate of d. The effect size for the observed difference is medium (d=0.59). Therefore, it is concluded that the class gained a significant and meaningful amount of knowledge over the course.

The post-test scores are also converted into standardized rubric scores as described in the methods. The average rubric score for the post-test was 2.31. This value falls short of the threshold score for proficiency (2.6). However, it is not significantly different from the threshold value (t=-1.72, df=44, p=0.092) as measured by a one-factor t-test. Moreover, the effect size of this difference is small (d=-0.26). We conclude that the average rubric scores approach, but do not yet meet, the threshold score for proficiency.
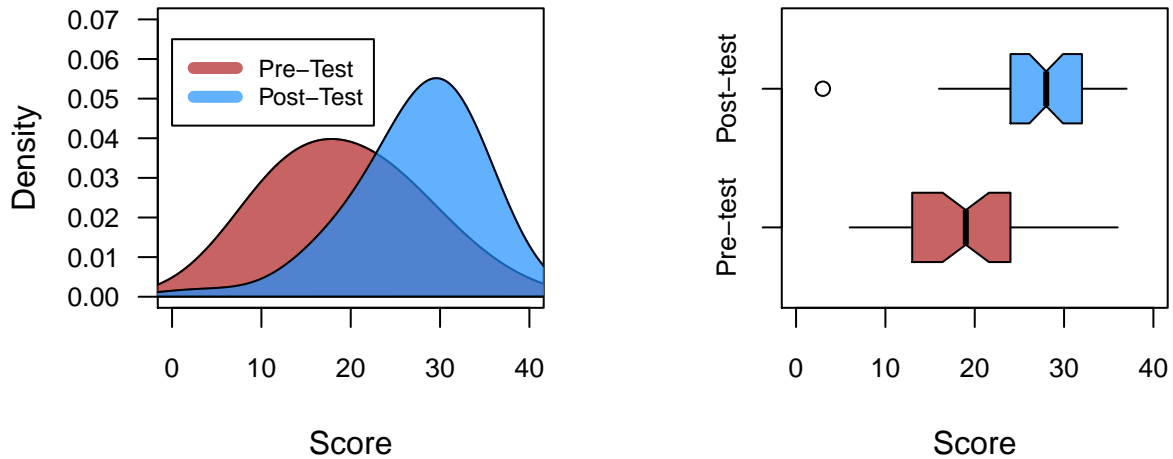
Figure 1: Distribution of pre- and post-test scores. Left panel = smoothed density plot of scores showing marked positive skew for the pre-test and negative skew for the post-test results. Right panel = boxplots of the pre- and post-test scores. The post-test scores were much higher.

The gain of knowledge in this course can also be estimated by comparing the number of proficient test scores (rubric scores of 3 or 4) between the pre- and post-tests. These results are summarized in Table 2.

Table 2: Effect of taking the class on exam proficiency rates.

Test	Proficient	Not Proficient	Total
Pre-test	6	39	45
Post-test	27	18	45

Based upon these results, it is estimated that students are 4.5 times more likely to be proficient in the post-test as compared to the pre-test. Although this is an impressive gain, only 60% of the class scored "proficient" or better on the post-test. This is lower than desired and reflects the low average rubric scores.

The questions dataset was also used to compare question performance on the pre- and post-tests. The results of these comparisons are shown in Figure 2. The prediction is that the questions will vary in difficulty, but all show approximately the same gains from pre- to post-test. Therefore, they should form a linear relationship.
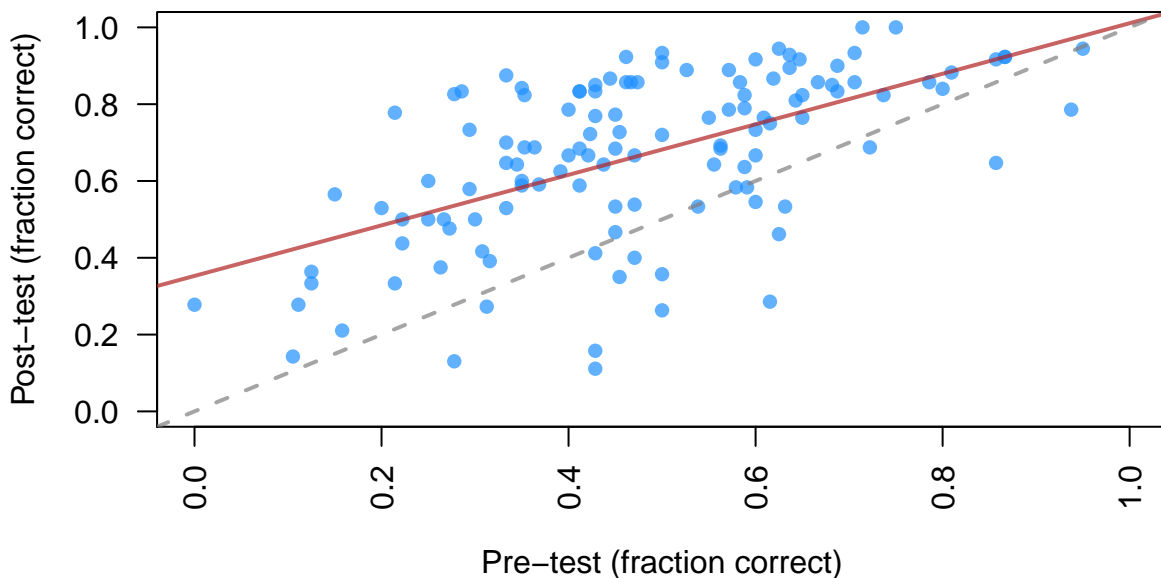

Figure 2: Scatterplot of pre-test vs post-test performance by question. The solid red line is the best linear fit to the data. The broken gray line represents no difference between pre- and post-test performance.

The best linear fit reveals that there was substantial improvement in performance for the many of the more difficult pre-test items. This causes the slope to be shallower than 1 (slope = 0.66). The fit of the linear regression is moderately good as well (r^2=0.34). There were about 15 or 16 questions that did not improve from pre-test to post-test. These need to be identified and examined.

Conclusions

A total of 45 students took both the pre- and post-test assignments. These tests consisted of 40 multiple-choice items randomly drawn from a pool of 123 questions. The average scores on the pre-test and post-test were 19.3 and 27.4, respectively. The difference between these scores is statistically significant (t=5.53, df=86.42, p=3.4e-07) and the effect size for the observed difference is medium (d=0.59). The average converted rubric score for the post-test is 2.31. This is below the criterion of proficiency (2.6), but is not significantly different (t=-1.72, df=44, p=0.092). Based upon the pre-test and post-test rubric scores, students are 4.5 times more likely to test as proficient at the end of the course than at the beginning. Although substantial learning has occured, only 60% of the class scored "proficient" or better on the post-test. Over a dozen exam items did not exhibit improvements in performance between the pre-test and post-test administration.

Proposed actions

- Examine the exam items to determine which questions did not improve over time.
- Possibly delete, edit, or replace defective items
- Determine if pattern exists between lecture module and question difficulty
- Examine scores for outliers and determine their causes and impact on this analysis
- No changes to pedagogy are proposed at this time.

Reuse and distribution of these materials

All of the de-identified data, analysis code, and documentation that constitute this report project may be freely used, modified, and shared. The code files (process.R and report.Rmd) are released under the MIT license. The de-identified dataseta (PrePostScores.csv and PrePostQuestions.csv) are released under the Creative Commons CC0 license. All documentation (including README.md, Codebook.md, and this report) are releaseed under the Creative Commons CC-BY licence. Any questions, comments, or suggestions may be sent to Dr. Franklund.

References

Delucchi, Michael. 2014. "Measuring Student Learning in Social Statistics: A Pretest-Posttest Study of Knowledge Gain." *Teach. Sociol.* 42 (3): 231–39. doi:10.1177/0092055X14527909.

R Core Team. 2013. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

RStudio Team. 2015. *RStudio: Integrated Development Environment for R.* Boston, MA: RStudio, Inc. http://www.rstudio.com/.