The Future is Clear: Developing Models to Predict Student Performance

Jennifer Ward[1], Erika G. Offerdahl[2,4], Jeffrey Boyer[3,4]
[1]St. Norbert College, [2]Department of Biological Sciences,
[3]College of Science and Mathematics, [4]North Dakota State University

Introduction

- Instructors gather data from multiple sources and hope to be able to use these data to inform instruction
  - E.g., course exams and Introductory Molecular and Cell Biology Assessment (IMCA)[4]
- Instructors can then analyze these data to make credible inferences about student performance and differences in how various groups perform
- Inferences can lead to better support for students during semester and changes in curriculum and instruction

Which student variables contribute to differences in course performance and IMCA performance?

- Data pooled from 2012 and 2014 course – BIOC 460
  - 426 students total-39.67% male and 11.5% non-white
- Multiple linear regression allows for making predictions on the outcome of a single variable (e.g., final grade) based on other variables (e.g., GPA, pre-IMCA, major, gender) that vary in influence

Predicting Final Grade

Predictor	Unstandardized Coefficient	P-value
Intercept	0.37	<0.001
Exam 1	0.32	<0.001
GPA	0.04	<0.001
Pre-IMCA	0.07	<0.001
Chem rigor	0.01	0.08

R^2: 0.71, Adjusted R^2: 0.71
P-value0<.001

Variables:
Exam 1
GPA
Pre-IMCA
Bio rigor
Chem rigor
Final grade
Major
Gender
Ethnicity
Semester

Discussion:

- Based on how student performance is measured, different variables influence that performance
- Regression model predicting final grade explains over 70% of variation
- GPA has been previously shown to be predictive of student performance[1]
  - This model confirms GPA matters
  - GPA based on course grades which is based on other factors besides content, unlike post-IMCA
- Since GPA and pre-IMCA score are predictive of final course grade, instructors may develop special instruction to assist students with low GPAs and IMCA scores
- Course has no biases towards major, gender or ethnicity, so does not cause or strengthen a gap in achievement
- Before any instruction, students' prior knowledge and confidence in chemistry preparation are most predictive of final grade

Predicting Post IMCA

Variables:
Exam 1
GPA
Pre-IMCA
Bio rigor
Chem rigor
Final grade
Major
Gender
Ethnicity
Semester

Predictor	Unstandardized Coefficient	P-value
Intercept	-0.03	0.52
Pre-IMCA	0.46	<0.001
Final Grade	0.46	<0.001
Gender	0.04	<0.001
Ethnicity	0.04	<0.01
Semester	-0.02	0.02

R^2: 0.56, Adjusted R^2: 0.56
P-value<0.001

Discussion:

- Regression model predicting post-IMCA score explains over 50% of the variation
- GPA not being predictive makes sense since GPA takes into account other factors, whereas IMCA only measures content knowledge
- Final grade being predictive suggests the course aligns well with IMCA content
- Being male and being white are statistically significant predictors of post-IMCA score (see right panel)
- 2014 students appear to perform better
  - ***Is there really a statistically significant difference in performance?***
  - *t*-tests
    - Normalized change→ p= 0.38
    - Raw change → p= 0.41

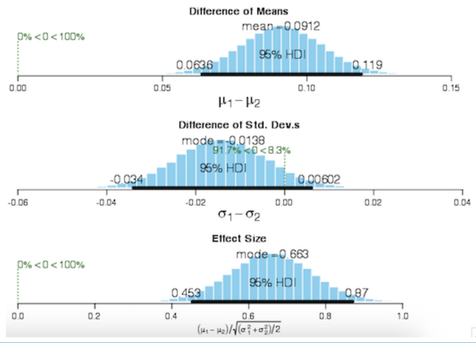
    No Significant Difference

Does a credible difference exist between gender groups and ethnicity groups?

Bayesian Estimation of Group Differences on Post-IMCA

- Reveals "relative credibility of every possible difference of means, difference of standard deviations, and all possible effect sizes" [3]

Male vs. Female Performance



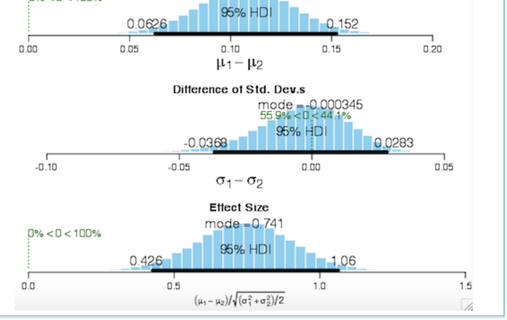
- Y1=male students
- Y2=female students
- Males outperformed females on pre-IMCA

Males perform better than females on post-IMCA

White vs. Non-White Performance

- Y1= white students
- Y2= non-white students
- White students outperformed non-white students on pre-IMCA

White students perform better than non-white students on post-IMCA



Discussion:

- Bayesian estimation reveals a credible difference between male and female performance
  - ➢ Further research should investigate whether IMCA is a gendered assessment
- Small number of non-white students could possibly contribute to what is seen as a difference in white vs. non-white performance

Future Research

- How do A/B students differ in motivation/study skills than others?
- To what degree does perceived rigor of preparation in chemistry and biology influence course performance?
- How does completing a cell biology course prior to BIOC 460 affect student performance?

References

1. Theobald, R., & Freeman, S. (2014). Is It the Intervention or the Students? Using Linear Regression to Control for Student Characteristics in Undergraduate STEM Education Research. *CBE-Life Sciences Education*, *13*(1), 41-48.
2. Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers & Education*, *54*(2), 588-599.
3. Krushke, John K., (2013). Bayesian Estimation Supersedes the *T*-test. *Journal of Experimental Psychology: General, 142*(2), 573-603.
4. Shi, Jia, et al., (2010). A Diagnostic Assessment for Introductory Molecular and Cell Biology. *CBE-Life Sciences Education, 9*(4), 453-461.

Acknowledgments

- Thanks to Jessie Arneson, Shannon Anderson and Andrew Calascione for their help
- Thanks to CiDER REU faculty, participants, and North Dakota State University
- Thanks to the National Science Foundation for the funding of this project – Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation
- Research supported in part by NSF REU, DUE #1156974

NSF
St. Norbert College
NDSU
ICEBERG
Investigating Conceptual Evolution Biochemistry Education Research Group