Stressed-out Intermediate Algebra Students

Michael Lloyd
Professor of Mathematics

Abstract

Anxiety, preparation, and performance are analyzed by measuring the stress hormone cortisol in students' saliva, by assessments, and by psychological surveys.

Institutional Review Board

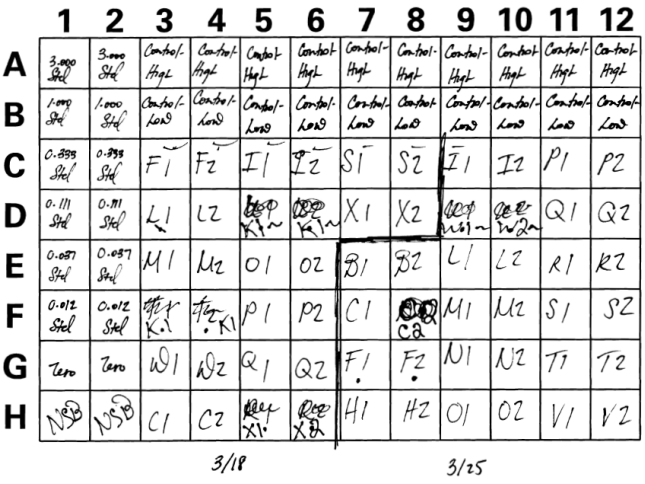
This study (Psychological and Physiological Stress in Intermediate Algebra Students: Relating Anxiety, Preparation, and Performance) was submitted and approved by the Henderson State University Institutional Review Board.

Introduction

The data for this study were obtained from students who took algebra from Ms. Morado in the spring of 2016. She described the design at the Oklahoma-Arkansas Mathematical Association of America meeting in the spring of 2017. The author also presented this paper dealing with the statistical analysis at that conference.

Methodology

The saliva samples were placed in plates containing $8 \times 12 = 96$ cells. The table on the right corresponded to Plate 1 and showed the layout for the saliva samples taken on March 18 and 25. The layout for Plates 2 and 3 corresponded to samples for the dates April 22, April 27, and April 8, but they are not shown here. Earlier dates are missing because we were waiting for IRB approval. We do not know where the samples for the later dates are. The standard (Std) concentrations were measured in micrograms per deciliter. In practice, all the standardized measurements appear on one side of the tray.

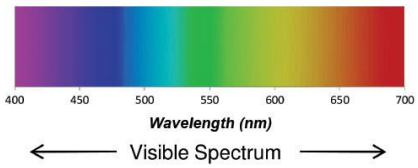


Some statisticians object to this violation of randomization, but this is what is done in practice. The variable Zero (B_0) corresponds to a sample that contained no cortisol. The nonspecific binding (NSB) samples did not appear on Plate 2. E. Beltzer said, "NSB wells are not required. In my cortisol assays, I usually do not include those because these assays are expensive to run, so I elect to run one more sample in place of the optional NSB wells." The author did not know how to use the Cortisol-High and Cortisol-Low values; they are probably used for quality control in checking the validity of the of the student saliva samples in the remaining cells. The letters A–W

represented the code randomly assigned to a student to preserve his or her identity, and the number refers to the sample time. For example, F1 and F2 were the before and after samples for student F, respectively. We had difficulty reading some of the labels for March 18 and do not know why there were four samples for Student K.

The website [2] summarizes the process of how the cortisol concentration of the student samples are determined: “The Free Cortisol in Saliva ELISA Kit is a solid phase enzyme-linked immunosorbent assay (ELISA), based on the principle of competitive binding. The microtiter wells are coated with a mouse monoclonal antibody directed towards an antigenic site on the Cortisol molecule. Endogenous Cortisol of a patient sample competes with a Cortisol-horseradish peroxidase conjugate for binding to the coated antibody. After incubation, the unbound conjugate is washed off. The amount of bound peroxidase conjugate is reverse proportional to the concentration of Cortisol in the sample. After addition of the substrate solution, the intensity of color developed is reverse proportional to the concentration of Cortisol in the patient sample.”

A microplate reader in the HSU Chemistry department shined a light wavelength of 405 nanometers through the plate and measured the absorption. The graph [1] on the right shows this purple color is near the end of the color spectrum visible to humans.



The accompanying table shows the raw molecular absorption spectrometry (MAS) values obtained from a comma separated file produced by the microplate reader. We only needed the top absorption value for each cell. (M is the median and R is the range of a single value.) The values for Plate 2 and Plate 3 are not shown.

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.105 M:0.105 R:0.000	0.140 M:0.140 R:0.000	0.342 M:0.342 R:0.000	0.241 M:0.241 R:0.000	0.295 M:0.295 R:0.000	0.371 M:0.371 R:0.000	0.318 M:0.318 R:0.000	0.292 M:0.292 R:0.000	0.352 M:0.352 R:0.000	0.392 M:0.392 R:0.000	0.438 M:0.438 R:0.000	0.442 M:0.442 R:0.000
B	0.247 M:0.247 R:0.000	0.316 M:0.316 R:0.000	1.054 M:1.054 R:0.000	1.154 M:1.154 R:0.000	1.311 M:1.311 R:0.000	1.082 M:1.082 R:0.000	1.174 M:1.174 R:0.000	1.331 M:1.331 R:0.000	1.208 M:1.208 R:0.000	1.204 M:1.204 R:0.000	1.185 M:1.185 R:0.000	1.284 M:1.284 R:0.000
C	0.504 M:0.504 R:0.000	0.613 M:0.613 R:0.000	0.868 M:0.868 R:0.000	1.458 M:1.458 R:0.000	1.306 M:1.306 R:0.000	1.365 M:1.365 R:0.000	0.762 M:0.762 R:0.000	0.928 M:0.928 R:0.000	1.234 M:1.234 R:0.000	1.234 M:1.234 R:0.000	1.030 M:1.030 R:0.000	0.684 M:0.684 R:0.000
D	0.901 M:0.901 R:0.000	1.208 M:1.208 R:0.000	2.013 M:2.013 R:0.000	0.957 M:0.957 R:0.000	0.708 M:0.708 R:0.000	0.777 M:0.777 R:0.000	0.874 M:0.874 R:0.000	1.020 M:1.020 R:0.000	1.297 M:1.297 R:0.000	0.767 M:0.767 R:0.000	1.184 M:1.184 R:0.000	0.933 M:0.933 R:0.000
E	1.556 M:1.556 R:0.000	1.418 M:1.418 R:0.000	1.018 M:1.018 R:0.000	1.461 M:1.461 R:0.000	1.360 M:1.360 R:0.000	1.147 M:1.147 R:0.000	1.239 M:1.239 R:0.000	1.231 M:1.231 R:0.000	2.014 M:2.014 R:0.000	1.151 M:1.151 R:0.000	0.836 M:0.836 R:0.000	1.120 M:1.120 R:0.000
F	1.531 M:1.531 R:0.000	1.613 M:1.613 R:0.000	0.905 M:0.905 R:0.000	0.881 M:0.881 R:0.000	1.046 M:1.046 R:0.000	1.161 M:1.161 R:0.000	1.903 M:1.903 R:0.000	0.633 M:0.633 R:0.000	0.590 M:0.590 R:0.000	0.707 M:0.707 R:0.000	0.776 M:0.776 R:0.000	0.474 M:0.474 R:0.000
G	1.912 M:1.912 R:0.000	1.931 M:1.931 R:0.000	1.076 M:1.076 R:0.000	1.276 M:1.276 R:0.000	0.686 M:0.686 R:0.000	0.853 M:0.853 R:0.000	0.943 M:0.943 R:0.000	1.080 M:1.080 R:0.000	0.997 M:0.997 R:0.000	0.984 M:0.984 R:0.000	0.925 M:0.925 R:0.000	1.130 M:1.130 R:0.000
H	1.916 M:1.916 R:0.000	1.877 M:1.877 R:0.000	1.120 M:1.120 R:0.000	1.142 M:1.142 R:0.000	0.932 M:0.932 R:0.000	0.707 M:0.707 R:0.000	0.570 M:0.570 R:0.000	1.189 M:1.189 R:0.000	0.566 M:0.566 R:0.000	0.563 M:0.563 R:0.000	0.389 M:0.389 R:0.000	0.085 M:0.085 R:0.000

Standard Concentration Curve

The author wrote an R script for reading these files and storing the values in matrices. Here are the values for Plate 1:

```
> MAS1
      1     2     3     4     5     6     7     8     9    10    11    12
A 0.105 0.140 0.342 0.241 0.295 0.371 0.318 0.292 0.352 0.392 0.438 0.442
B 0.247 0.316 1.054 1.154 1.311 1.082 1.174 1.331 1.208 1.204 1.185 1.284
C 0.504 0.613 0.868 1.458 1.306 1.365 0.762 0.928 1.234 1.234 1.030 0.684
D 0.901 1.208 2.013 0.957 0.708 0.777 0.874 1.020 1.297 0.767 1.184 0.933
E 1.556 1.418 1.018 1.461 1.360 1.147 1.239 1.231 2.014 1.151 0.836 1.120
F 1.531 1.613 0.905 0.881 1.046 1.161 1.903 0.633 0.590 0.707 0.776 0.474
G 1.912 1.931 1.076 1.276 0.686 0.853 0.943 1.080 0.997 0.984 0.925 1.130
H 1.916 1.877 1.120 1.142 0.932 0.707 0.570 1.189 0.566 0.563 0.389 0.085
```

These absorption values will be referred to as optical densities (OD). Most of the standard concentration OD values were smaller than the Zero values on Plate 2, and one of the standard values had a smaller OD value than the Zero value on Plate 3. This is evidence that both of these plates were corrupted. In an attempt to use as much of the available saliva data as possible and treat all the plates the same, the NSB and B_0 values were not initially used.

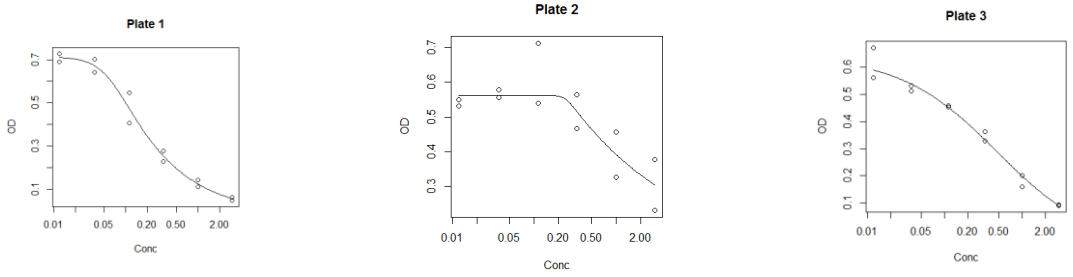
The 5-parameter (5PL) logistic model shown here was considered for the best-fitting standard-concentration curves. The parameter c is the slope of logit of the response ($\log_{10}\frac{OD}{1-OD}$) versus the explanatory variable ($\log_{10}$ concentration). The value of the asymmetric parameter is 1 in a 4-parameter logistic model.

$$OD = B + \frac{T - B}{[1 + 10^{b(C_{mid} - C)}]^s}$$

$OD =$ optical density
$C =$ concentration in µg/dl
$C_{mid} = C$ value of inflection point
$B =$ bottom asymptote
$T =$ top asymptote
$b =$ Hill slope
$s =$ asymmetric parameter

I scaled the OD values using 10% more than the maximum value of the values for each plate because the logistic model in the nplr R library [3] would not allow responses larger than 1, probably because it uses a maximum-likelihood fit. The 5PL model performed better than the simpler models in the same family when accounting for variable inflation.

The standard concentration fits for the three plates appear below. The concentration is shown on a logarithmic scale. Plate 2 had an outlier with OD above 0.7, and the shape of the 5PL fit was not sigmoidal; this is also evidence that Plate 2 was corrupted.



The accompanying table shows some assessment statistics. Over half the cells in the last plate were presumably used for another study because they were unlabeled.

	Plate 1	Plate 2	Plate 3
Date(s)	3/18, 3/25	4/22, 4/27	4/8
Number of samples	60	60	24
Goodness of fit (correlation)	0.98	0.74	0.97
Number samples too small	1	0	0
Number samples too high	3	28	10
Total fraction unusable	7%	47%	42%

A student OD value was deemed unusable if it was outside the range of the 5PL model response values over standard concentration ranges. Plate 2 had a much lower correlation (0.74) value than the other Plates (approximately 1) and a large fraction of unusable student OD values (47%). Although the standard curve for Plate 3 had a high correlation (0.97) and an acceptably-shaped concentration curve, it had a large fraction of unusable student OD values (42%).

To further assess the validity of the plate measurements, two-way tables of the variable Date by invalid student values were constructed:

Plate 1

Date	3/18	3/25
FALSE	27	29
TRUE	1	3

Plate 2

Date	4/22	4/27
FALSE	16	16
TRUE	16	12

Plate 3

Date	4/8
FALSE	14
TRUE	10

The fraction of invalid values does not appear to depend on the date strongly. This suggests that the problem is with the preparation of Plates 2 and 3, not the samples for any particular date. Thus, we were only able to use Plate 1.

Since only Plate 1 is viable, we will use the correct formulas for the standard concentration curve that takes into account both B_0 and NSB. The best estimates for these values were obtained by averaging the two respective cells as shown here.

$$NSB = \frac{NSB_1 + NSB_2}{2}$$

$$B_0 = \frac{B_{01} + B_{02}}{2}$$

The net OD was computed as the difference shown here; this transformation will reverse the direction of the graph as seen below.

$$NetOD = RawOD - NSB$$

The fraction bound (FB) was obtained by scaling the net OD to fall in the interval $(0,1)$ by dividing by the estimated Zero OD value.

$$FB = NetOD/B_0$$

Here is a graph of the 5PL standard concentration curve that accounts for the nonspecific binding and zero OD values. In addition to being a reflection, its shape is different than the previous model for Plate 1. Hence, it will predict slightly different values for the student concentrations.

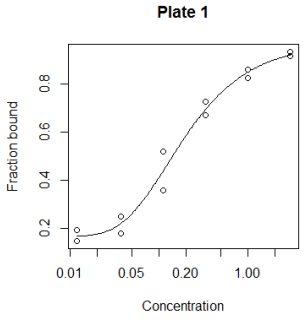


Plate 1 corresponded to Exam 4 in the spring 2016 semester. The tables on the right show the estimated student concentrations of cortisol in the students' saliva in micrograms per deciliter using the standard concentration curve. NA means the corresponding OD value was either missing or invalid because it fell outside the standard-concentration range. There were four more samples on the exam day than the lecture day. Nonresponse is a problem in this study because most of the variables are missing some values.

Since the only available saliva data were for Exam 4, the data that will be considered for this analysis will be for this period, the demographic data, or the surveys that appear to measure persistent personality characteristics.

Date 3/18
Lecture

Id	Conc1	Conc2
C	0.098	0.093
F	0.165	0.040
I	0.064	0.055
K	0.179	NA
L	NA	0.137
M	0.121	0.040
O	0.056	0.092
P	0.114	0.089
Q	0.246	0.170
S	0.207	0.145
W	0.107	0.069
X	0.153	0.166

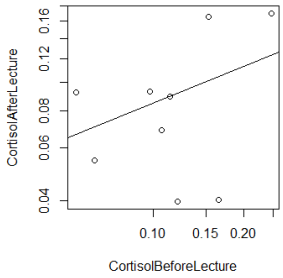
Date 3/25
Exam 4

Id	Conc1	Conc2
B	0.075	0.077
C	NA	0.280
F	0.141	0.106
H	0.330	0.084
I	0.076	0.076
L	NA	0.091
M	0.313	0.234
N	0.126	0.129
O	0.333	0.336
P	0.118	0.247
Q	0.085	0.144
R	0.176	0.098
S	0.201	0.435
T	0.146	0.096
V	0.580	NA
W	0.066	0.205

Analysis

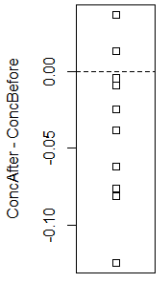
Nonparametric procedures [4] were used throughout this paper because the residuals often appeared to have outliers or be skewed left.

The saliva samples on March 18 were taken just before and after the lecture. There was no significant linear relationship between the $\log_{10}$ of the cortisol concentrations.



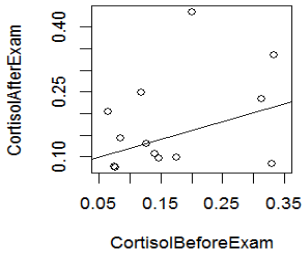
Rfit procedure
n=9
Overall Wald Test statistic =0.522
p-value = 0.61495

However, there was almost a significant difference between the cortisol concentrations between the beginning and end of the lecture on March 18. The students' stress hormones may have decreased on average during the lecture.



Wilcoxon Test
n = 9
Test statistic V = 8
p-value = 0.098

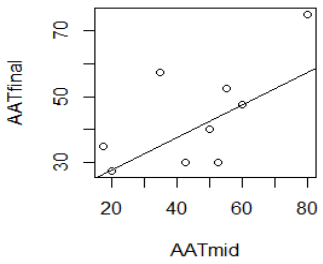
There was no significant linear relationship between the $\log_{10}$ of the cortisol concentrations taken before and after the exam on March 25.



Rfit procedure
n=13
Overall Wald Test statistic = 1.54
p-value = 0.26

Achievement Anxiety Test (AAT) [5] measured anxiety about academic achievement: The higher the score, the more confident a student feels about his or her academic performance when anxious. This survey was given on a lecture day (April 11) and on the day of the final exam (May 9). The original test had 19 questions, but we only used the ten debilitating ones. The scores could range from 0 to 100, but our students' scores ranged from 17.5 to 80 with a median of 45. Although not in the usage instructions, the score was adjusted for missing questions for four out of 26 subjects.

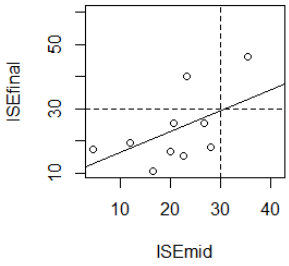
The smaller values of AAT correspond to anxiety causing problems with performance. There was not a significant linear relationship between the AAT values for the middle and end of the semester.



Rfit procedure
n = 9
Overall Wald Test statistic = 2.05
p-value = 0.20

The Index of Self-Esteem (ISE) [6] survey was given on the same lecture day and also during the final exam. ISE scores can range from 0 to 100, but ours ranged from 5 to 60 with a median of 23. Higher ISE scores indicate a problem with self-esteem.

ISE scores above 30 indicate the presence of a clinically significant problem. One student appeared to have poor self-esteem for both days, and another seemed to have low self-esteem only at the middle of the semester.



There was not a significant linear relationship between the mid and final ISE scores.

Rfit procedure
n = 10
Overall Wald Test statistic = 2.08
p-value = 0.19

We are interested in predicting a student's performance as measured by Exam 4 (n=20) based on the other variables. There were too many explanatory variables

compared to the number of valid responses, so some of these needed to be removed from consideration.

The surveys shown here were not used because they were only recorded for April and May dates, and it is plausible that the responses to these will change over time. Recall that the cortisol data were only available for March 18 and March 25. The last survey on the list measured the variables nervous, hours studied, and expected grade.

- ParticipantAllDays
- Self-rating Anxiety Scale (SAS) [7]
- Stress Arousal Checklist (SACL) [8]
- SupplementAllExamDays

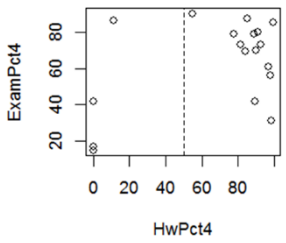
The ParticipantOnce variables shown here were not used. All the students were either 18 or 19 years old. There was only one Hispanic student. They were all freshmen. There were four psychology majors and a variety of other majors. There were only three smokers. Only four had a job. All but one had taken the course precisely once before.

- Age
- Hispanic
- Year in school
- Major
- Smoking
- Job
- NumRetakes

Here is a list of the relevant variables that were still being considered. Some of these still need to be removed because of the small sample sizes. The variable Relationship originally had three values: “No;” “Yes, cohabitating;” and “Yes, non-cohabitating.” I simplified this to a binary variable which only had the values “yes” and “no.”

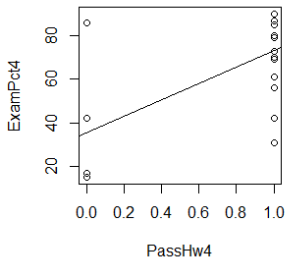
Academic	Cortisol	Demographic	Psychological
• HwPct4 (n=21) • TimeHours (n=21)	• CortisolAfterLecture (n=11) • CortisolBeforeLecture (n=11) • CortisolAfterExam (n=16) • CortisolBeforeExam (n=15)	• Class (n=8 noon, n=15 1 pm) • Race (12 black, 6 white) • Relationship (9 no, 6 yes) • Sex (8 female, 10 male)	• AATmid (n=12) • AATfinal (n=14) • ISEmid (n=9) • ISEfinal (n=11)

The scatter plot on the right showed the Exam 4 percentage (ExamPct4) versus the MyMathLab [9] homework average when the students took Exam 4. A linear model was inappropriate, so the logical variable PassHw4 was created. It was defined to be true if and only if the homework percent (HwPct) was at least 50%. Zero homework scores were not considered as missing data, but zero ExamPct4 grades were removed.



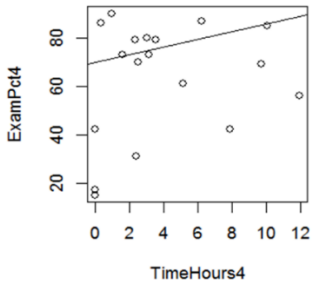
There was a highly significant linear relationship between the Exam 4 percentages and the PassHw4 indicator variable. (PassHw4 is 0 if false and 1 if true.)

Rfit procedure
n = 18
Overall Wald Test statistic = 9.12
p-value = 0.002

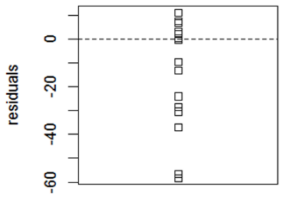


There was not a significant linear relationship between the Exam 4 percentage and time spent on homework during the Exam 4 period.

Rfit procedure
n = 18
Overall Wald Test statistic = 1.95
p-value = 0.17

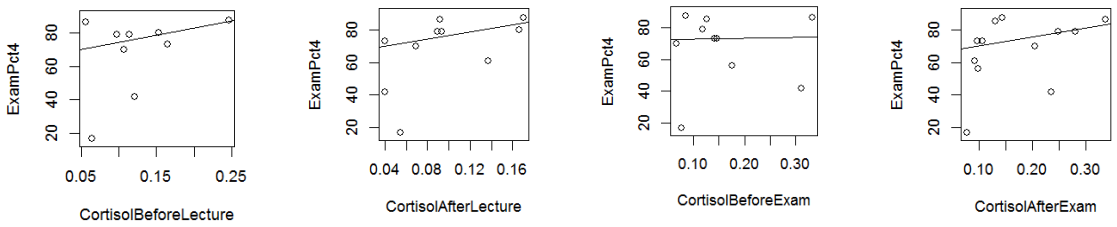


Although the residuals were skewed negatively, using the Bentscores3 option in the Rfit procedure made the p-value even larger.



None of the cortisol measurements had a significant linear relationship for predicting the Exam 4 scores. The bentscores3 option was used for all of these Rfit procedures because of the low outliers. All the best-fitting lines had positive slopes, despite our expectation that academic performance would have a negative relationship with cortisol.

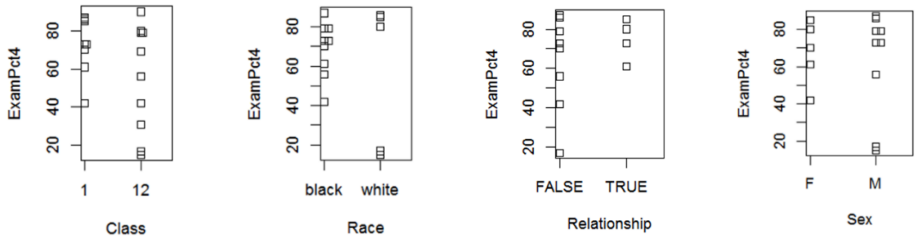
Variable	n	P-value	Keep
Cortisol-BeforeLecture	9	0.133	No
Cortisol-AfterLecture	11	0.232	No
Cortisol-BeforeExam	10	0.996	No
Cortisol-AfterExam	12	0.549	No



None of the four demographic variables had a significant effect on the average Exam 4 percentage as shown in the table on the right. The p-values are shown for 2-sample Wilcoxon tests, but they were also all large when the Rfit procedure was used.

Variable	n_1	n_2	P-value	Keep
Class	12	10	0.247	No
Race	9	5	0.894	No
Relationship	9	4	0.423	No
Sex	5	9	0.699	No

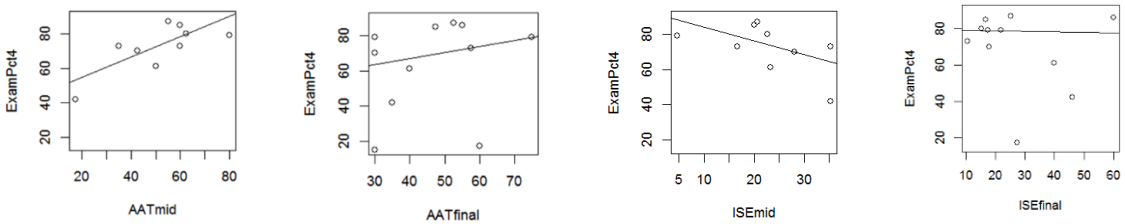
Here are side-by-side dot plots for the Exam 4 percentage verses the demographic variables.



The table on the right shows the results of applying Rfit tests to see if there is a linear relationship between the Exam 4 percentage and the psychological surveys. Only AATmid showed a significant relationship.

Variable	n	P-value	Keep
AATmid	9	0.034	Yes
AATfinal	11	0.645	No
ISEmid	9	0.216	No
ISEfinal	11	0.994	No

The following scatterplots show the Exam 4 percentage versus the psychological variables with their Rfit lines. We expected a positive relationship for the AAT graphs and a negative relationship for the ISE graphs.



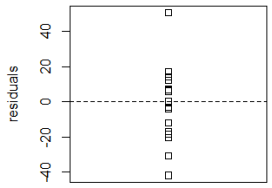
Conclusion

The accompanying table shows the variables for predicting ExamPct4 with p-values less than 0.20.

Variable	*Sample size*	*p-value*
AATmid	9	0.03
CortisolBeforeLecture	9	0.13
PassHw4	18	0.00
TimeHours4	18	0.17

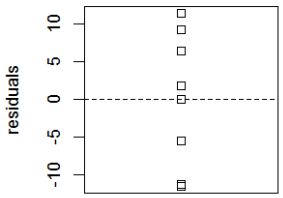
There were only six students who had values for all five variables and twice as many students with homework scores than the AATmid and CortisolBeforeLecture values. Hence, it would be inappropriate to attempt to find an optimal multilinear model starting with these four predictor variables.

Here is a graph of the residuals for predicting ExamPct4 only using PassHw4. The residuals are large, usually ranging from -42% to 17%. So although there is a highly significant linear relationship, PassHw4 would not likely be useful for accurately predicting ExamPct4.



ExamPct4 = 35.4 + 37.6 (PassHw4)

Here is a graph of the residuals for predicting ExamPct4 only using AATmid. The following is a more impressive model since the residuals only range from -12% to 11%:



ExamPct4 = 43.0 + 0.592 AATmid

Future Work

As of July 8, 2017, we are still waiting to process the cortisol data from the Fall 2016 semester. There were more students involved in that study (n=41 based on the ParticipantOnce survey). With this larger sample size, we might be able to predict student success based on the AAT score, cortisol, homework performance, or other variables.

Acknowledgements

The author acknowledges the contributions of Holly Morado (mathematics) as the primary investigator, David Bateman (chemistry), and Emilie Beltzer (psychology), student helpers, and administrative assistants.

References

[1] “Visible Spectrum” Available at http://www.physicsinfo.co.uk/?page=view&id=108.

[2] "MyBioSource: Your source for antibodies, proteins, ELISA kits and so much more!" Available at https://www.mybiosource.com/.

[3] "Package 'nplr'." Available at https://cran.r-project.org/web/packages/nplr/nplr.pdf.

[4] Kloke, K. and McKean J. W. (2014), *Nonparametric Statistical Methods Using R*, Chapman and Hall/CRC.

[5] Alpert, R., & Haber, R. N. (1960), "Anxiety in academic achievement situations," *The Journal of Abnormal and Social Psychology*, *61*(2), 207.

[6] Hudson W (1992), The WALMYR assessment scales scoring manual, Walmyr Publishing, Tempe.

[7] "Zung Self-Rating Anxiety Scale." Available at https://en.wikipedia.org/wiki/Zung_Self-Rating_Anxiety_Scale.

[8] Mackay, C., Cox, T., Burrows, G., & Lazzerini, T. (1978), "An inventory for the measurement of self-reported stress and arousal," *British Journal of Social & Clinical Psychology*.

[9] "MyMathLab." Available at https://www.pearsonmylabandmastering.com/northamerica/mymathlab/.

Biographical Sketch

Michael Lloyd graduated cum laude and in the honors program in Chemical Engineering with a B.S. in 1984. He accepted a position at Henderson State University in 1993 shortly after earning his Ph.D. in Mathematics (Probability Theory) from Kansas State University. He has presented papers at meetings of the Academy of Economics and Finance, the American Mathematical Society, the Arkansas Conference on Teaching, and the Southwest Arkansas Council of Teachers of Mathematics. He has been an active member of the Mathematical Association of America since 1993, earned 18 hours in computer science, and has been an Advanced Placement statistics consultant since 2002.

Progressive Team Home Run Leaders of the Los Angeles Dodgers, Boston Red Sox, and Pittsburgh Pirates

Fred Worth
Professor of Mathematics

Abstract. In this paper, we will look at which players have been the career home run leaders for the Los Angeles Dodgers, Pittsburgh Pirates, and Boston Red Sox since the beginning of the organizations.

Introduction

In the past, I published the progressive team home run leaders for the New York Mets, Chicago White Sox, Washington Nationals, Houston Astros, Los Angeles Angels and New York Yankees. Since I simply enjoy this kind of statistical amusement, I have done similar research and decided to publish three more this year.

I find this topic interesting for a variety of reasons. First, I simply enjoy baseball history. Of the four major sports (baseball, football, basketball, and cricket), none has had its history so