Let Master of Public Health Students Experience Statistical Reasoning
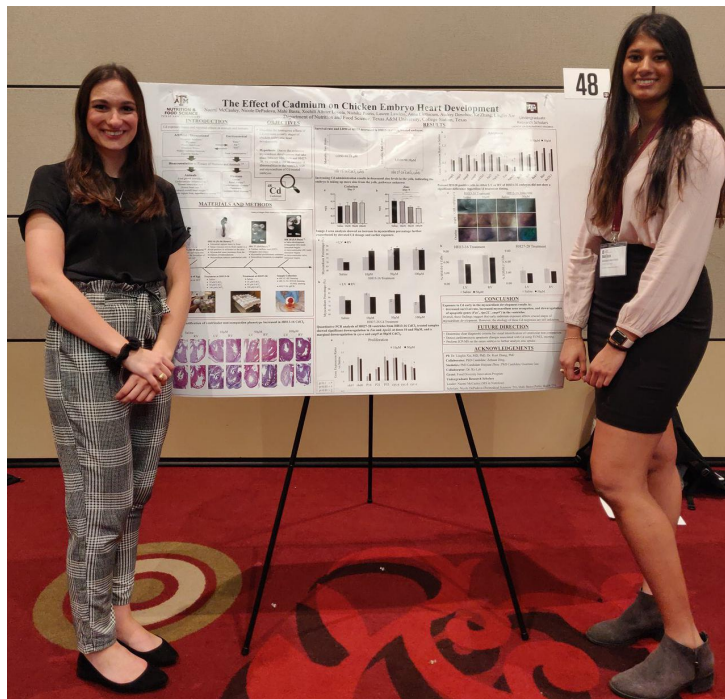
Athens, Greece, May 18, 2020

Qi Zheng

Department of Epidemiology and Biostatistics
Texas A&M University School of Public Health

Do these two groups differ in cognitive abilities?

- Styreotype I: The left can't do math

- Stereotype II: The left can't do computing

Stereotype is like an evil spell

- The witch is often invisible

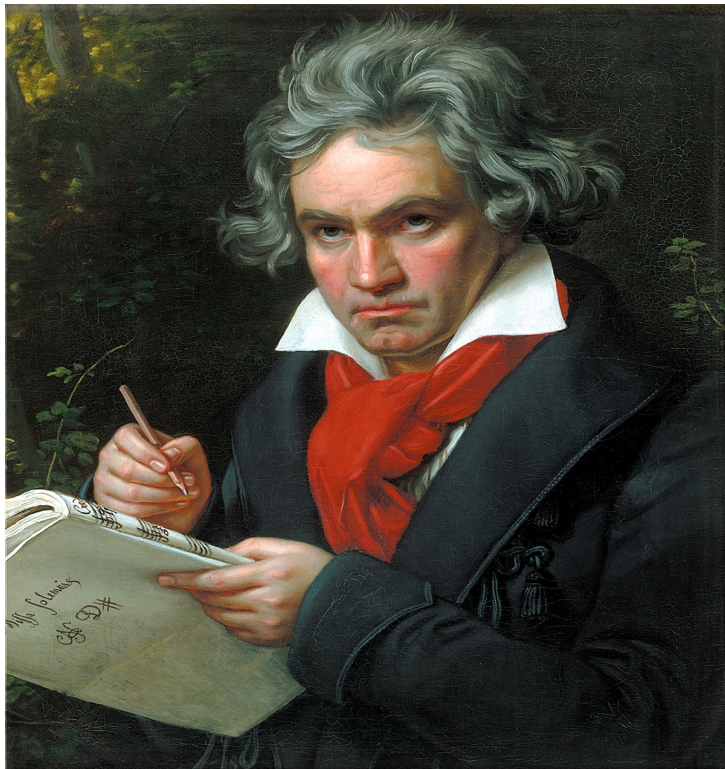- Sometimes the instructor acts as a witch as well

I might be among a minority of rebels

- Math ability is universal

- Interest in math is spread equally through the population
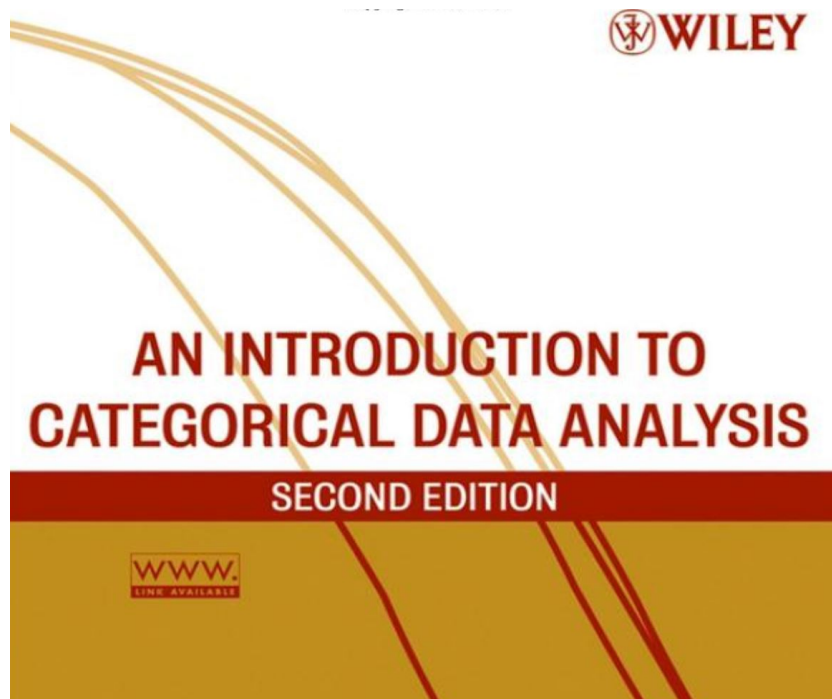
- Let us play a word substitution game

You may say "wait a minute?"

- It's the opportunity, stupid!

- Our job is to equalize the opportunities

- It's a conviction, not science

My experiments focused on PHEB 690

- It figures in 48.5% of public research

- It is heavily laden with abstract concepts

WILEY

AN INTRODUCTION TO CATEGORICAL DATA ANALYSIS

SECOND EDITION

WWW. LINK AVAILABLE

ĀM | PUBLIC HEALTH
TEXAS A&M UNIVERSITY

SYLLABUS

Course Information

Course Title and Number	Categorical Data Analysis: PHEB 609
Academic Term	Fall 2019
Meeting Times	On-line
Meeting Location	On-line
Instructor Name	Qi Zheng, PhD
Instructor Telephone Number	979-436-9398
Instructor Email Address	qzheng@tamu.edu
Instructor Office Hours	By Appointment
Instructor Office Location	SPHA 227
Teaching Assistant	None
Teaching Assistant Email Address	None

Course Description

PHEB 609 is an introduction to categorical data analysis that is tailored to the needs of students majoring in public health. The choice of topics and illustrative examples reflects common features in current public health research. A student taking this course is expected to possess basic high school mathematics knowledge, such as the exponential and logarithmic functions. The student is also assumed to have rudimentary computer skills, such as transferring and editing text files. The course is concept-driven and hands-on, and the course's motto is "comprehension before computation."

procedural skills v. conceptual knowledge

- level 1: production of computer output

- level 2: pronunciation of jargon

- level 3: understanding of underlying principles

Model Fit Statistics			
		Intercept and Covariates	
Criterion	Intercept Only	Log Likelihood	Full Log Likelihood
AIC	415.062	408.038	32.245
SC	418.759	415.432	39.640
-2 Log L	413.062	404.038	28.245

-2log(L)=404.038, hence log (L)=l= - 202.019

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.8097	1.6210	8.8037	0.0030
conc	1	0.3893	0.1315	8.7618	0.0031

Confirmed:
```
data beetle;
a=-4.81; b=0.39; lik=like1(a,b); output;
run;
proc print data=beetle;
run;
```

(-4.81, 0.39)= -202.021

Obs	a	b	lik
1	-4.81	0.39	-202.021

```
/* PHEB 609 HW2 */
proc fcmp outlib=work.hw2.a;
function like1(a,b);
l1=-15*log(1+exp(-(a+10.8*b)))-35*log(1+exp(a+10.8*b));
l2=-24*log(1+exp(-(a+11.6*b)))-25*log(1+exp(a+11.6*b));
l3=-26*log(1+exp(-(a+12.1*b)))-24*log(1+exp(a+12.1*b));
l4=-24*log(1+exp(-(a+12.6*b)))-26*log(1+exp(a+12.6*b));
l5=-29*log(1+exp(-(a+13.1*b)))-21*log(1+exp(a+13.1*b));
l6=-29*log(1+exp(-(a+13.5*b)))-20*log(1+exp(a+13.5*b));
return (l1+l2+l3+l4+l5+l6);
endsub;
run;
options cmplib=work.hw2;

data beetle;
a=2; b=3; lik=like1(a,b); output;
run;

proc print data=beetle;
run;
```

The SAS System

Obs	a	b	lik
1	2	3	-5795.3

An appetizer for conceptual knowledge

Simplified data from a beetle experiment						
Concentration	10.8	11.6	12.1	12.6	13.1	13.5
death	15	24	26	24	29	29
group size	50	49	50	50	50	49

- They can understand

$$\text{logit}[\text{P}(\text{death at concentration } C)] = \beta_0 + \beta_1 \times C$$

- They can see why they need

$$\left(\frac{1}{1 + e^{-(\beta_0 + 10.8\beta_1)}}\right)^{15} \left(\frac{1}{1 + e^{(\beta_0 + 10.8\beta_1)}}\right)^{35}$$

Simple math can go a long way towards comprehension

- Palpable joys of statistical reasoning

- Boosted confidence in statistical applications

→ a) likelihood function from the first group (death = 15, size = 50, c = 10.8)

$$L(p) = p^{15} \times (1-p)^{35}$$

given
$$\begin{cases} \eta = \text{logit}\left[P(\text{death at conc. } C)\right] = \beta_0 + \beta_1 C \\ C = 10.8 \end{cases}$$

$$\Rightarrow L(p) = \left(\frac{1}{1+e^{-\eta}}\right)^{15} \times \left(1 - \frac{1}{1+e^{-\eta}}\right)^{35}$$

$$= \left(\frac{1}{1+e^{-(\beta_0 + 10.8\beta_1)}}\right)^{15} \times \left(\frac{1}{1+e^{(\beta_0 + 10.8\beta_1)}}\right)^{35}$$

Students' first eureka moment

- Other key concepts can be similarly taught

```
/* PHEB 609 HW2 */
proc fcmp outlib=work.hw2.a;
function like1(a,b);
l1=-15*log(1+exp(-(a+10.8*b)))-35*log(1+exp(a+10.8*b));
l2=-24*log(1+exp(-(a+11.6*b)))-25*log(1+exp(a+11.6*b));
l3=-26*log(1+exp(-(a+12.1*b)))-24*log(1+exp(a+12.1*b));
l4=-24*log(1+exp(-(a+12.6*b)))-26*log(1+exp(a+12.6*b));
l5=-29*log(1+exp(-(a+13.1*b)))-21*log(1+exp(a+13.1*b));
l6=-29*log(1+exp(-(a+13.5*b)))-20*log(1+exp(a+13.5*b));
return (l1+l2+l3+l4+l5+l6);
endsub;
run;
options cmplib=work.hw2;

data beetle;
a=2; b=3; lik=like1(a,b); output;
run;

proc print data=beetle;
run;
```

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates	
		Log Likelihood	Full Log Likelihood
AIC	415.062	408.038	32.245
SC	418.759	415.432	39.640
-2 Log L	413.062	404.038	28.245

-2log(L)=404.038, hence log (L)=l= - 202.019

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.8097	1.6210	8.8037	0.0030
conc	1	0.3893	0.1315	8.7618	0.0031

Confirmed:
```
data beetle;
a=-4.81; b=0.39; lik=like1(a,b); output;
run;
proc print data=beetle;
run;
```

(-4.81, 0.39)= -202.021

The SAS System

Obs	a	b	lik
1	2	3	-5795.3

Obs	a	b	lik
1	-4.81	0.39	-202.021

requirements for imparting conceptual knowledge

- desire to know, to demystify

- common sense + first principles

- beginning high school math

$$\log(a \times b) = \log(a) + \log(b)$$

$$\log\left(\frac{a}{b}\right) = \log(a) - \log(b)$$

$$\log(e^x) = x$$

$$\log(a^b) = b \times \log(a)$$

$$\log\left(\frac{1}{a}\right) = -\log(a) \quad \text{with} \quad \log(1) = 0$$

$$\frac{e^a}{e^b} = e^{a-b} \quad \text{with} \quad e^0 = 1.$$

When students break the evil spell

Dr. Zheng,

I cried tears of joy when I saw my midterm grade! I did exactly as you recommended to prepare for the exam, and it served me well.

Thank you,

~~████████~~

On Mon, Oct 29, 2018 at 10:36 PM Qi Zheng - qzheng@sph.tamhsc.edu <do-not-reply@blackboard.com> wrote:

> Hello, on-line PHEB 609 students,
>
> About 58% of the students in this class scored 90 or above on the midterm exam, and four of them earned a perfect score. Allyson's work is attached to serve as key to the midterm exam. I was elated to see you making such impressive progress. I hope you will continue to work hard.
>
> Best wishes,
> QZ