

Lab Midterm #2: Review

Kin 304W

Week 12: July 23, 2013

Lab Midterm #2

- Thursday, July 25 in your regularly scheduled lab section
- 45 minutes, 4 questions
- Open book, open computer
- No emailing, no phone calls, no texting
- I'll post the data set on Wednesday, July 24
- We will look at the dataset briefly in lecture on Thursday
- Bring your questions to lecture on Thursday
- John will hold a review session on Wednesday, July 24 from 12-3 pm in the computer lab - thanks John!

Lab 6: Chi-Square

- **General objective:** Are the frequencies (percentages) of participants who report some characteristic/symptom different between groups?
- **Example:** Is there a difference in the percentage of employed and unemployed adults in Vancouver who report excellent health?
- **In SPSS,** choose Analyze/Descriptive Statistics/Cross-Tabs
 - Define **rows** (characteristics/symptom)
 - Define **columns** (groups that you are comparing)
 - Under Statistics, select Chi-square
 - Under Cells, select “observed counts” and “column percentages”
- **Pertinent statistics** are the Chi-square value, df, and p-value.

Lab 6: Chi-Square

Interpret these results

	Drivers		Cyclists				
	N=200		N=350				
	n	%	n	%	Chi-Square	df	P-value
Speed limits are too slow	125	62.5%	140	40.0%	17.842	1	0.008

Lab 6: Chi-Square

Interpret these results

	Drivers		Cyclists				
	N=200		N=350				
	n	%	n	%	Chi-Square	df	P-value
Speed limits are too slow	125	62.5%	140	40.0%	17.842	1	0.008

The percentage of drivers (62.5%) who reported that speed limits are too slow was significantly higher than the percentage of cyclists (40.0%) who reported that speed limits are too slow (Chi-square=17.842, df=1, $p=0.008$).

OR...Drivers (62.5%) were more likely than cyclists (40.0%) to report that speed limits are too slow (Chi-square=17.842, df=1, $p=0.008$).

Lab 6: Logistic Regression

- **General Objective:** Quantify how much more or less likely some binary outcome is in one group than another. You are able to take account of other covariates.
- **Example:** How much more likely are pregnant women to exercise 3x/week (yes/no) than non-pregnant women, when taking account of their age?
- **SPSS:** Analyze/Regression/Binary logistic
 - Define Dependent Variable (binary)
 - Define Covariates (Predictor + any covariates you want to take account of)
 - Under Method, select “Enter.”
 - Under Options, select CI for exp(B) 95 (this outputs the 95% CI for the OR)
- **Pertinent statistics** are the odds ratio (OR), which is labeled in SPSS as Exp(B), and the 95% confidence interval (CI) for the OR, which is labeled in SPSS at the 95% CI for Exp(B).

Lab 6: Logistic Regression

Interpret these results

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1 ^a	cigsnow	-1.070	.321	11.134	1	.001	.343	.183	.643
	Constant	.223	.224	.996	1	.318	1.250		

a. Variable(s) entered on step 1: cigsnow.

Dependent variable: Do you regularly have itchy eyes (1=Yes, 0=No)

Independent variable: cigsnow (Do you smoke cigarettes now? 1=Yes, 0=No)

Lab 6: Logistic Regression

Interpret these results

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1 ^a	cigsnow	-1.070	.321	11.134	1	.001	.343	.183	.643
	Constant	.223	.224	.996	1	.318	1.250		

a. Variable(s) entered on step 1: cigsnow.

Dependent variable: Do you regularly have itchy eyes (1=Yes, 0=No)

Independent variable: cigsnow (Do you smoke cigarettes now? 1=Yes, 0=No)

Among current cigarette smokers, the odds of having itchy eyes is 0.34 times that in non-smokers (OR=0.34; 95%CI=0.18-0.64, p=0.001).

OR...The odds of having itchy eyes is 66% lower in current smokers vs. non-smokers (OR=0.34; 95%CI=0.18-0.64, p=0.001).

Labs 7 & 8: Modeling

- Assume you are given some **experimental data** about a dependent variable.
- If you are also given a **mathematical equation** that relates that dependent variable to some combination of independent variables and constants (unknown parameters), know the steps you would take to fit the equation to the experimental data in Excel.

1. Identify the unknown equation parameters and specify their initial values
2. Calculate predicted values of the dependent variable from equation
3. Calculate squared residuals $(\text{observed} - \text{predicted})^2$
4. Sum the squared residuals.
5. Invoke EXCEL Solver and set the sum of squared residuals as the target cell to minimize. Identify the model parameters that Solver can modify.
6. Run Solver, twice to be sure.

Lab 9: Quantification of EMG

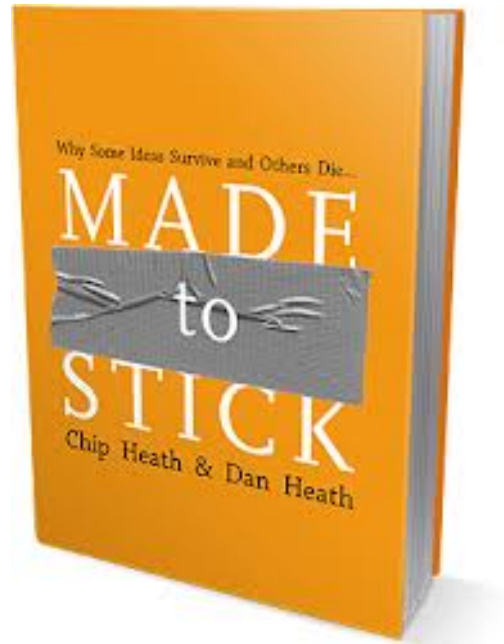
- If you are given **raw values for EMG** from a set of experimental conditions, list the steps you would take to quantify the amount of EMG in each condition.
 1. Calculate offset and subtract it from all cells
 2. Rectify, using absolute value function (ABS in Excel)
 3. Take the average rectified EMG for each condition
 4. Could calculate integrated EMG for each condition

Return of Marked Lab Midterm #2

- Lab midterm #2 will be returned by John Manning at 9:30 am on Tuesday, July 30.
- He will review the answers to each question and then collect the lab midterms from you.
- He will be available to meet with students on July 30 to go over individual questions/concerns.

A couple of loose ends to tie up...

23^{1/2} hours: 3.5 million views in 18 months



‘Stickiness’ Checklist*

- Simple
- Unexpected
- Concrete
- Credible
- Emotion
- Stories

*From Chip and Dan Heath, “Made to Stick: Why Some Ideas Survive and Others Die.”
New York: Random House, 2007.

'The Joy of Stats': Example of Data Visualization

- Hans Rosling, MD, PhD
- Expert in Global Health and Data Visualization
- 200 Countries in 4 Minutes: The Joy of Stats
 - <http://youtu.be/jbkSRLYSojo>
 - ~5.7 million views

