Categorical Data Analysis Homework Assignments, Spring 2019

- The total length of your assignment should be no longer than about 5 pages, not including graphs, tables, and an appendix containing any computer code. Problems marked with an A are taken from Agresti (2007).
- Odd numbered problems from Agresti have brief answers in the back of the book. You should have more detail than what is presented there to receive credit!
- Due dates are approximate and may be altered depending on how things are going.
- I generally assume you will have questions of various sorts during class. Please ask them!
- The take home final will have questions from the last material we do in class.

Homework 1[1]: Due 2/11/19

Consider the following data, which involve a small study of mechanical ability. Researchers recruited 21 participants. These participants had final grades in formal classroom instruction (Gr), which was used as a covariate, which is assumed continuous. They were randomized into three practical training conditions. Conditions A and B are novel, while C is the control. After being trained their aptitude (Apt) for mechanical work was measured. Higher scores are better. You may use your software's factor coding for grouping and to generate any relevant interactions.

A		B		C	
Apt	Gr	Apt	Gr	Apt	Gr
6	3	8	4	6	3
4	1	9	5	7	2
5	3	7	3	7	2
3	1	9	4	7	3
4	2	8	3	8	4
3	1	5	1	5	1
6	4	7	2	7	4

1. What kind of design is this?
2. Set up the data appropriately. Show this in your answer.
3. Check (via one-way ANOVA) that the groups appear equivalent in the relevant sense.
4. Perform an appropriate linear regression analysis, including any necessary checks of the validity of the models (e.g., normal quantile plots of residuals or standardized residuals using qnorm). Be sure to check for possible interaction between group and covariate.
5. Summarize your findings.

(10 points)

Homework 2: Due 2/25/19

A1.3, A1.12 (also compute the Agresti-Coull interval), and A1.13. (5 points each)

Plot the likelihood function in A1.13. (5 points)

Homework 3: Due 3/4/19

A2.2, A2.5, A2.13, A2.21, A2.36. (3 points each)

A2.33 (5 points)

---

[1] This question assesses your prerequisite knowledge. You may find it challenging but it should not be out of reach. If it is, this class is probably not for you.

Homework 4: Due 3/18/19

A3.9, A3.15. (Enter the data and verify in both cases except for 3.15d, each worth 5 points.)

Consider the following 2×2 table, which lists counts of 1000 students with diagnosed learning disabilities (LD or NLD) cross-classified with diagnosed anxiety disorder (AD or NAD).

	NLD	LD	Margin
NAD	730	100	830
AD	130	40	170
Margin	860	140	1000

Compute AR, RR, and OR, focusing on the use of AD as a marker for LD. Generate appropriate 95% confidence intervals for each. Which measure appears the most reasonable for the question? Check your work for OR using an appropriate GLM. (10 points)

Homework 5: Due 4/1/19

A3.19, A3.20 (3 points each)

A3.21-A3.22 (4 points total)

Homework 6: Due 4/29/19

A4.3, A4.5, A4.15 (3 points)

A4.31 (5 points)

A4.16-A4.17, A5.4 (10 points)

Homework 7: Due 5/13/19

A8.2 (5 points)

A4.20. Explain why pooling is necessary. For this problem use both `cc` to do the MH procedure and `meglm` (or `melogit`) to do mixed effects analysis. Compare the results. In particular, compare the resulting odds ratios and confidence intervals. (10 points).