Exam 1

Last Name: _____   **First Name:** _____

I hereby state that I have not communicated with or gained information in any way from other students or any outside resource during this exam. I agree to abide by the rules stated below, and to abide by the Wake Forest Honor Code. All work is my own. I understand that any violation of this agreement will be reported to the Honor Council and will result, at minimum, in a 0 on this exam.

Signature : _____

- -

All work on this exam must be your own.

1. You have 50 minutes to complete the exam.

2. Show all your work on the open ended questions in order to get partial credit. No credit will be given for open ended questions where no work is shown, even if the answer is correct.

3. You are allowed a calculator, however you may not share a calculator with another student during the exam. The calculator must be only a calculator, and may not be connected to the internet.

4. You are allowed to ask clarification questions to me, but you may not ask anyone else.

5. You are **<u>not</u>** allowed a cell phone, even if you intend to use it as a calculator or for checking the time. You are **<u>not</u>** allowed a music device or headphones, notes, books, or other resources.

6. You may **<u>not</u>** communicate with anyone other than myself during the exam.

7. Write clearly and be clear. Make it easy to find your answers.

Good luck!

- -

The Data We have a client who is interested in examining a recent advertising campaign for DoorDash, a popular food service in which food from restaurants is picked up and delivered to your door. The client wants to examine the relationship between exposure to advertising and the use of the service at a local university. They have data on 1500 randomly selected students from this university, and the data were collected during the month of September when the campaign was active.

We have information on X = the amount of exposure a student has to the ad campaign during the month (measured in minutes, between 0 to 6 minutes. Decimals are permitted for values less than a whole minute) and Y = whether or not the student makes a purchase using DoorDash during the same month (0 = no, 1 = yes).

1. Based on the information we have so far, write down a population model you would suggest the client consider to model the desired relationship. Use appropriate notation.

We create an empirical log odds plot to check the shape assumption, and you may assume all conditions for modeling are met. The client fits the model and obtains the following output:

Model 1

	Estimate	Std. Error
(Intercept)	-0.21	.12
AdExposure	0.26	.08

Null Deviance :	649.12	on 1499 degrees of freedom
Residual Deviance :	568.84	on 1498 degrees of freedom

2. Based on the output above, write down the equation of the fitted model.

3. Interpret the slope in terms of the odds.

Now a second client provides us data on whether or not the students in the study have a job (0 = no, 1 = yes). They build Model 2 by starting with Model 1 and then adding job status as explanatory variable. Call this model **Model 2**.

Model 1

	Estimate	Std. Error
(Intercept)	-0.21	.12
AdExposure	0.26	.08

Null Deviance :	649.12	on 1499 degrees of freedom
Residual Deviance :	568.84	on 1498 degrees of freedom

Model 2

	Estimate	Std. Error
(Intercept)	-0.31	.15
AdExposure	0.34	.11
Job	- .43	.17

Null Deviance :	649.12	on 1499 degrees of freedom
Residual Deviance :	528.23	on 1497 degrees of freedom

4. Suppose we want to answer the following: "After accounting for ad exposure, are students with jobs **more likely** to order from DoorDash?" Write down null and alternative hypotheses, in terms of one or more parameters in the model, which allow you to answer this research question.

5. What is the name of the hypothesis test you would use to answer this research question? Calculate the test statistic for this hypothesis test using the output above, and state the name of the distribution you would use to calculate a p-value from this test statistic. (You do not need to calculate the p-value).

Model 2

	Estimate	Std. Error
(Intercept)	-0.31	.15
AdExposure	0.34	.11
Job	- .43	.17

Null Deviance :	649.12	on 1499 degrees of freedom
Residual Deviance :	528.23	on 1497 degrees of freedom

6. Interpret the coefficient for job status (Job) in terms of the odds.

7. If a student with a job has been exposed to 2.5 minutes of advertising, what is the predicted probability that this student orders from DoorDash? Show all work and round your answer to 2 decimal places.

Suppose we are told that we have a certain random variable Y_i with four possible outcomes (A, B, C, D). We have the following data: D, A, B, C, C, A, B, C.

We are told that:

- $P(Y_i = A) = 2\pi_D$
- $P(Y_i = B) = 4\pi_D$
- $P(Y_i = C) = 1 - 7\pi_D$
- $P(Y_i = D) = \pi_D$.

---

8. Using the method of maximum likelihood, solve for the maximum likelihood estimate of π_D. Show your work.

You are done!!! Whooo!!!