

Name _____

Student Number _____

STA 441s 2024 Quiz 9

1. (6 points) In a study of factors related to successful brain surgery in adults, the explanatory variables were age of patient in years (x_1), and experience of the surgeon, measured in number of times he or she had carried out the procedure (x_2). The response variable was $Y = 1$ for a successful surgery, or $Y = 0$ for unsuccessful. Denoting the probability of success by π , the logistic regression model is

$$\ln\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

- (a) You want to test whether, controlling for patient's age, the surgeon's experience is related to whether the surgery is successful. In symbols, what is the null hypothesis?
- (b) You want to test whether, controlling for surgeon's experience, the patient's age is related to whether the surgery is successful. In symbols, what is the null hypothesis?
- (c) You want to test patient's age and surgeon's experience simultaneously. In symbols, what is the null hypothesis?
- (d) For fixed age of patient, each time the surgeon has carried out the procedure in the past multiplies the odds of success by ... what? Answer in terms of symbols from the logistic regression model.
- (e) For any fixed level of surgeon's experience, each additional year of the patient's age multiplies the odds of success by ... what? Answer in terms of symbols from the logistic regression model.
- (f) Would you expect β_1 to be positive or negative? Why? I am asking for a common sense argument (a *short* one), not mathematical argument.

$$\ln\left(\frac{\pi}{1-\pi}\right) = \ln\left(\frac{P(Y=1|\mathbf{X}=\mathbf{x})}{P(Y=0|\mathbf{X}=\mathbf{x})}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_{p-1} x_{p-1}$$

$$P(Y = 1|\mathbf{X} = \mathbf{x}) = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_{p-1} x_{p-1}}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_{p-1} x_{p-1}}}$$

2. This question is based on your analysis of the bird-keeping and cancer data.

(a) (2 points) You wish to test whether, controlling for the other variables, being a bird keeper is related to getting cancer.

i. Fill in the table below. On your printout, circle the chi-squared statistic and write “Question 2a” beside it. Either Wald or likelihood ratio is okay.

Chi-squared Statistic (a number)	<i>p</i> -value (a number)	Reject Null Hypothesis? (Yes or No)	Statistically Significant? (Yes or No)

ii. In plain, non-statistical language, what do you conclude? No marks for this without the first part right. Start your answer with “Allowing for other possible risk factors ...”

(b) (1 point) For a non-smoking, bird-keeping woman of average age and low socioeconomic status, what is the estimated probability of lung cancer? Write the answer in the space below. On your printout, circle the number (produced by `proc iml`) and write “Question 2b” beside it.

(c) (1 point) For a non-smoking, **non**-bird-keeping woman of average age and low socioeconomic status, what is the estimated probability of lung cancer? Write the answer in the space below. On your printout, circle the number (produced by `proc iml`) and write “Question 2c” beside it.

Attach your complete log file and your *COMPLETE* results file to the quiz paper. Make sure your name and student number are written clearly on both printouts.