

## STAT 435 Fall 2023 Test 1

---

1.(23 points) When the 2000 General Social Survey asked subjects whether they would be willing to accept cuts in their standard of living to protect the environment, 344 of 1170 subjects said “yes”

a)(5 points) Estimate the population proportion who would say “yes” .

b)(8 points) Conduct a hypothesis test to determine whether a majority or minority of the population would say “yes”.

c)(10 points) Find the 99% Wald, Agresti-Coull, and Wilson intervals. Why do you think the intervals are similar?

2.(24 points) The following table provides results from a study comparing radiation therapy with surgery in treating cancer of the larynx.

		Cancer	
		Controlled	Not controlled
Group	Surgery	189	10,845
	Radiation therapy	104	10,933

Let  $\pi_1$  be the probability that cancer was controlled for the surgery treatment, and let  $\pi_2$  be the probability that cancer was controlled for the radiation therapy treatment.

a)(6 points) Find and interpret the estimated relative risk and corresponding confidence interval.

b)(6 points) Find and interpret the estimated odds ratio and corresponding confidence interval.

c)(6 points) Perform a Pearson chi-square test to determine if there is a difference between the treatments?

d)(6 points) Show that  $\pi_1 = \pi_2$  is equivalent to  $OR = 1$  and  $RR = 1$ .

3.(33 points) The shuttle.csv file (available on the graded materials web page) contains information about the temperature at lift off for a space shuttle and whether there was thermal distress observed. Below is how I read the data into R:

```
> set1 <- read.csv(file = "C:\\data\\shuttle.csv")
> head(set1)
temp TD
1    66  0
2    70  1
3    69  0
4    68  0
5    67  0
6    72  0
```

Using temperature (temp) as the explanatory variable and thermal distress (TD) as the response variable, complete the following.

a)(8 points) Estimate the logistic regression model.

b)(8 points) Find the 95% profile LR confidence interval for the odds ratio involving temperature. Use a 5 degree DECREASE in temperature when calculating the interval. Interpret this interval with a phrase like “the odds of a success are ...”

c)(8 points) Estimate the probability of thermal distress when the temperature is 70.

(d)(9 points) Plot the estimated probability of thermal distress as a function of temperature on the graph below.

4.(20 points) Answer the following questions.

a) (5 points) When is a confidence interval referred to as being “liberal”?

b) (5 points) Describe how the formula for the Agresti-Coull confidence interval comes about in relation to the Wald confidence interval.

c) (10 points) Consider the logistic regression model of

$$\log\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_r x_r + \dots + \beta_p x_p \quad (1)$$

Derive the odds ratio expression that is used to interpret the effect that  $x_1$  has on the response variable. Simply stating the odds ratio expression will not receive any credit.