<u>Project 2</u> – Edith's Egg Farm

Edith Edwards owns many chickens and would like to own a business someday selling free range locally grown organic eggs. She has hired a statistics expert (you) to help her with some questions.

1
Edith likes extra large eggs and is interested in maximizing the weight of the eggs her chickens lay. She wonders which variables contribute to egg size and carefully collected data from 30 of her chickens during the month of July.

$Y =$ average weight among the eggs the chicken produced in July (grams)
$x_1 =$ age of the chicken (months)
$x_2 =$ weight of the chicken on July 15th (grams)
$x_3 =$ average amount of water consumed per day (liters)
$x_4 =$ average amount of chicken feed consumed per day (grams)
$x_5 =$ average amount of time spent outside per day (hours)

The data is in the file egg.csv

	A	B	C	D	E	F
	egg	age	weight	water	feed	time
2	52.6	55	1751	0.66	108	14.2
3	59.1	89	1881	0.72	118	12.1
4	63.3	85	1969	0.68	118	15
5	50.7	63	1593	0.57	119	15.9
6	62.1	97	1897	0.79	105	15.4

For example, the first line of data in the csv file describes a chicken who had an average egg weight of 52.6 grams during the month of July, was 55 months old, weighed 1751 grams, consumed an average of 0.66 liters of water per day, consumed an average of 108 grams of feed per day, and spent an average of 14.2 hours outside per day.

Edith says to you:

"I took an introductory statistics course in college and learned about simple linear regression. But now I have several potential explanatory variables and I don't know how to analyze the data. I would really like to know which variable(s) are useful while predicting average egg weight."

2

Edith is experimenting with different breeds of chickens. In particular, she is interested in White Leghorn chickens and Rhode Island Red chickens and wants to know if one breed typically lays more eggs than the other. She randomly selected 6 White Leghorns and 7 Rhode Island Reds and recorded how many eggs each of them laid last week. Here are the results:

White Leghorns = { 4, 5, 6, 7, 7, 9 }
Rhode Island Reds = { 5, 5, 6, 7, 7, 8, 10 }

Edith says to you:

"I was about to do a two-sample t-test with my data, but then I remembered from my intro stats course that I can only do a two-sample t-test if my sample sizes are large or if my data is normal. These sample sizes are small and the data is a discrete count of the number of eggs so the data is not normal. Is there a different hypothesis test we can do instead?"

3

Edith has read that chickens typically lay fewer eggs when the weather starts to get cold and wants to know if this claim is true. At Edith's farm, this typically happens at the end of September. She randomly selected 10 of her chickens and recorded how many eggs they laid in September and October. Here is her data:

name of chicken	eggs laid in September	eggs laid in October
Henrietta	34	25
Amelia Egghart	29	26
Dixie Chick	30	28
Big Bird	27	26
Eggy McEggFace	32	35
Cluck Norris	35	32
Chicken Little	26	28
General Tso	27	25
Foghorn Leghorn	36	26
Attila the Hen	38	26

Edith says to you:

"I was about to do a paired t-test with my data, but then I remembered from my intro stats course that I can only do a paired t-test if my sample size is large or if my data is normal. Is there a different hypothesis test we can do instead?"

Unlike Project 1, you are not required to submit a report to Edith. Instead, complete the problems below just like a regular homework assignment.

Problem-by-problem Instructions:

1 – use R

1a
First find the linear model that uses all five explanatory variables and report the intercept and coefficients ("slopes"). If you decide to eliminate any explanatory variables, justify your decision and report the relevant p-values.

1b
After you have determined which set of explanatory variables is appropriate to keep, find the linear model and report the intercept and coefficients ("slopes"). Compare the values of r^2 before and after removing the unnecessary explanatory variables.

1c
Finally, use your final model to predict the average egg weight for a chicken with the following values:

$x_1 =$ 70 months
$x_2 =$ 1800 grams
$x_3 =$ 0.80 liters
$x_4 =$ 110 grams
$x_5 =$ 15 hours

2 – use SAS

2a
What non-parametric hypothesis test is appropriate for this situation?

2b
Complete the hypothesis test by hand. Show work and report all relevant computations and information (H_0, H_a, test statistic, conclusion, p-value bounds, etc.)

2c
Write a SAS program that confirms your answers to part 2b. Print and turn in both the SAS program and the output.

3 – use SAS

3a
What non-parametric hypothesis test is appropriate for this situation?

3b
Complete the hypothesis test by hand. Show work and report all relevant computations and information (H_0, H_a, test statistic, conclusion, p-value bounds, etc.)

3c
Write a SAS program that confirms your answers to part 3b. Print and turn in both the SAS program and the output.