Homework 10

1. Have you ever wondered how tall Humpty Dumpty must have been before he fell from the wall? We have the tools to find his center of mass by constructing a confidence interval about μ. Let us find his height. Fortunately, STAT51 students took pity on poor Humpty Dumpty by scraping-up what was left of him. The aggregated bits and pieces weighed 175 pounds (79275 grams). Of course, we know that Humpty Dumpty was not spherical but egg shaped. In order to correlate length (mm) with weight (grams) of an egg, the following set of data for 29 species of bird eggs ranging from the Hummingbird to the extinct Elephant Bird is given below:

Bird	1	2	3	4	5	6	7	8	9	10
Weight (grams)	350	.5	7.8	3.2	18.18	20.76	17.93	22.6	19.01	37.95
log(weight), x,	5.86	-0.69	2.05	1.16	2.90	3.03	2.89	3.12	2.94	3.64
Length (mm)	113	13	28	21	30	32	31	35	32	41
log(length), y	4.73	2.56	3.33	3.04	3.40	3.47	3.43	3.56	3.47	3.71
Bird	11	12	13	14	15	16	17	18	19	20
Weight (grams)	7.55	4.47	3.78	5.02	3.09	3.01	5.38	7.1	4.54	4.41
log(weight), x,	2.02	1.50	1.33	1.61	1.13	1.10	1.68	1.96	1.51	1.48
Length (mm)	24	19	20	19	18	18	21	23	20	19
log(length), y	3.18	2.94	3.00	2.94	2.89	2.89	3.04	3.14	3.00	2.94
Bird	21	22	23	24	25	26	27	28	29	
Weight (grams)	11.95	7.34	8.62	4.19	2.72	4.43	6.30	1132.5	12231	
log(weight), x,	2.48	1.99	2.15	1.43	1.00	1.49	1.84	7.03	9.41	
Length (mm)	28	24	25	20	16	19	23	152.4	280	
log(length), y	3.33	3.18	3.22	3.00	2.77	2.94	3.14	5.03	5.63	

$\sum_{i=1}^{29} x_i = 71.0672$, $\sum_{i=1}^{29} x_i^2 = 283.7360$, $\sum_{i=1}^{29} x_i y_i = 274.7038$, $\sum_{i=1}^{29} y_i = 96.90692$, $\sum_{i=1}^{29} y_i^2 = 336.7251$, $SS_{xx} = 109.5792$, $SS_{xy} = 37.22434$, and $SSE = .2539898$.

Let w be the weight and l be the length of an egg. We know from our chemistry class that $l = \gamma w^{\frac{1}{3}}$ where γ is related to density. Since Humpty Dumpty was not perfectly spherical, let us propose the following model: $l_i = \eta_0 w_i^{\eta_1}$ to relate his height and weight. What looks like a non-linear model can in fact be transformed into a linear model by taking logarithms, i.e. $log(l_i) = log(\eta_0) + \eta_1 log(w_i) + \epsilon_i$ where $log(x) = log_e(x) = ln(x)$. By doing so, an equivalent model for explaining the data can be written as a linear fixed effects model by replacing $log(l_i)$ with y_i, $log(\eta_0)$ with β_0, η_1 with β_1, and $log(w_i)$ with x_i so that: $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ where $\epsilon_i \sim N(0, \sigma^2)$.

(a) Test the hypothesis that there is no inherent linear relation between the response variable and the explanatory variable at a level of significance, $\alpha = .05$.

(b) We are interested in finding Humpty Dumpty's height given that he weighed 79275 grams. Therefore, produce an estimate of y_p using your fitted model.

(c) Of course, we want a confidence interval of our estimate. Compute the 95% CI for the particular value of y_p.

(d) Compute Humpty Dumpty's height, i.e. calculate $\widehat{height} = e^{\widehat{y}_p}$. Use the `exp` key on your calculator. The answer will be in millimeters. To convert to inches, divide your answer by 25.4.

(e) Make a plot of residuals versus predicted values.

(f) Do you approve of the model and trust the estimate? Explain.

2. Based on used car auctions, the final auction prices of 100 randomly selected three year Fleming XX500's and the respective odometer readings were recorded and plotted in a graph. Each of these cars was in good condition with the standard features like automatic transmission, radio, and air conditioning. The only essential difference about them was the odometer reading and final auction price. The following statistics were derived from the data in which y=price and x=odometer reading in thousands of miles: $\sum_{i=1}^{100} x_i = 3600.945$, $\sum_{i=1}^{100} x_i^2 = 133977.4$, $\sum_{i=1}^{100} x_i y_i = 19351920$, $\sum_{i=1}^{100} y_i = 541141$, $\sum y_i^2 = 2934770709$, $SS_{xx} = 4309.34$, and $SS_{xy} = -134269.3$, and SSE=2251362.

The plot of the data suggests the linear fixed effects model:$y_i = \beta_0 + \beta_i x_i + \epsilon_i$ where $\epsilon_i \sim N(0, \sigma^2)$

(a) Find $\widehat{\beta}_0 =$

(b) Find $\widehat{\beta}_1 =$

(c) Write the fitted equation for $\widehat{E[y_i]}$=

(d) Test the hypothesis, $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$ at a level of significance, $\alpha = .05$ $F_{1,98,.05}$=3.938111

Source of Variation	df	Sum of Squares	Mean Sum of Squares	F statistic
Mean		2928335819		
Regression		4183528		
Residual Error				
Total				

(e) What is $s^2 =$

(f) What is the estimated price on the average of a Fleming XX500 having 27000 miles on its odometer?

(g) The 95% confidence interval of the expected price of a Fleming XX500 which has 27000 miles on its odometer is (5387.035,5997.213). Would a trade-in value of $5800 be a fair value to offer a customer wanting to buy a new car from you? Explain.

(h) Explain by referring to the plot of the data, the analysis of variance, and the residuals versus predicted value whether or not the model is good.