

Final Exam – Take Home

First Year College Student Nutrition Study

The data for your project comes from a nutritional study conducted at Youngstown State University during 1997-1998. Forty four subjects completed the study in which body measurements and nutrition data was collected at the beginning of the Fall semester and then again in the Spring semester. A portion of that data appears in the data file **nutri1a.xls** which is available at <http://csuohio.edu/holcombj/mth147/finalexam.htm>
Variable guide:

studynum	An identifying number to keep track of subjects
gender	0=male, 1=female
residenc	0=on-campus, 1=off-campus
athlete	0=non-athlete, 1=athlete
heightf	Height in cm in the Fall semester
weightf	Weight in kg in the Fall semester
weights	Weight in kg in the Spring semester
bmifall	Body Mass Index in the Fall semester
weightch	Change in Weight from Fall to Spring (weights-weightf)
wt10	0=no, 1=yes for weight change over 10lbs.
bmi25f	0=no, 1=yes, for bmi over 25 in the Fall

The main purpose of the study was to examine weight and nutrition characteristics in the college first year population. One the variables examined was the change in weight from the fall to the spring. The variable **weightch** above is a variable that indicates whether the student gained more than 10 lbs during the first year (**weightch=1**) or did not (**weightch=0**).

For the following tests of Hypothesis, be sure to state the hypotheses, the test statistic, the P-value or the P-value estimate, and your conclusion. Assume that the subjects are a random sample of first year students

1. Is there sufficient evidence to conclude that the mean amount of kilograms gained over the year is greater than 2 kg? Use the variable **weightch** for your test of hypotheses.
2. Is there sufficient evidence to conclude that over 20% of first year students gain over 10 lbs (Use variable **wt10**)?
3. Determine if **athlete** and **wt10** (gaining over 10lbs) are independent

Use the **heightf** to predict **weightf** to answer the following:

1. Show the scatterplot.
2. Give the model for the regression line.
3. Give the r^2 value, interpret its strength, and interpret its meaning.
4. Use the model to predict the **weightf** when the **heightf** is 171 cm.
5. Are there any influential observation? Why or why not?
6. Include the residual plot. Are there any outliers? If so, how many?
7. Is **heightf** a good predictor for **weightf**? Why or why not? Can you think of a way it could be improved?

Write a paragraph that describes your conclusions. Also, perform some kind of a test of hypothesis that I have not proposed. This could be a test involving a mean, a proportion, independence, or it could involve a regression analysis. Clearly state your null, alternative, test statistic, P-value, and conclusion.

Answers

<i>weightch</i>	
Mean	2.684091
Standard Error	0.401463
Median	2.2
Mode	3
Standard Deviation	2.663006
Sample Variance	7.091601
Kurtosis	-0.54494
Skewness	0.499786
Range	10.9
Minimum	-1.9
Maximum	9
Sum	118.1
Count	44

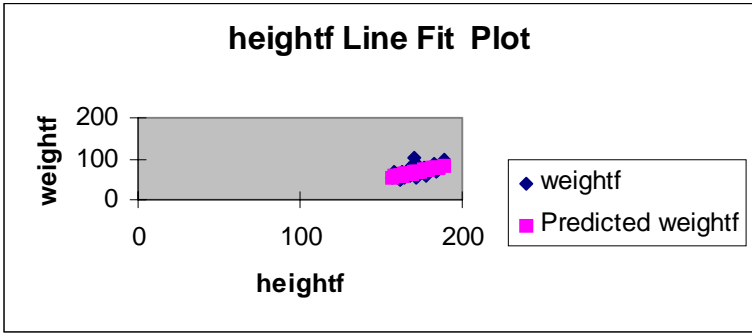
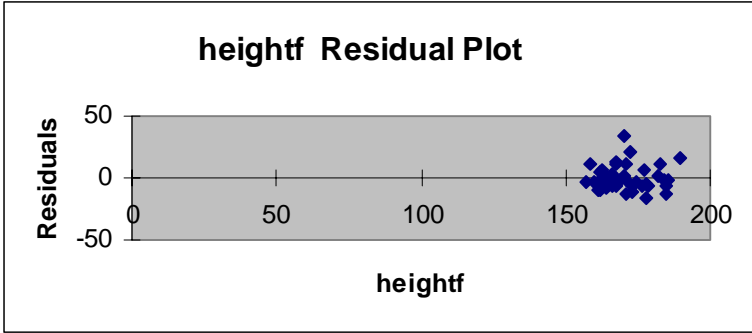
z 1.703994
pvalue 0.044191

Sum of count	
wt10	Total
0	33
1	11
Grand Total	44

phat=phat 0.25
z 0.829
pvalue 0.204

Sum of count	wt10		
athlete	0	1	Grand Total
0	23	8	31
1	10	3	13
Grand Total	33	11	44

Chi-Square 0.0364
df 1
p-value 0.849



<i>Regression Statistics</i>	
Multiple R	0.58204
R Square	0.33877
Adjusted R Square	0.323026
Standard Error	9.928697
Observations	44

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	-77.9865	31.32656	-2.48947	0.016839
heightf	0.851721	0.18361	4.638749	3.41E-05
