Statistics 101: Section L - Laboratory 6

In today’s lab, we are going to look more at least squares regression, and interpretations of slopes and intercepts.

Activity 1: From lab 1, we collected data on the Total Weight (in grams) of Fun Size bags M&Ms and on the number of M&Ms in each bag. Can we use the number of M&Ms in each bag to predict the Total Weight of the bag? Go to the course web page and open the JMP data file that contains the data on the M&Ms. Use Analyze – Fit Y by X with Total Weight as the Y, Response and Number as the X, Factor. Fit a line to the data and have JMP plot the residuals. Turn in this output with the answers to the questions below.

a) Describe the general relationship between the number of M&Ms in each bag and the total weight of the bag. Does it appear this relationship is strong or weak?

b) Give the equation of the least squares regression line for predicting the total weight of the bag from the number of M&Ms in the bag.

c) Find the predicted total weight of a bag of M&Ms if the bag contains 21 M&Ms.

d) Calculate the residual for the bag with 21 M&Ms and a total weight of 20 grams.

e) Study the residual plot. Describe what you see and how this might affect (or not affect) the least squares regression of the two variables.

f) Give the value of the estimated y-intercept for this regression. Give the interpretation of the intercept value in the context of the problem.

h) Give the value of the estimated slope for this regression. Give the interpretation of the slope value in the context of the problem.

i) Look closely at the interpretation of the slope value. What physical quantity does the slope estimate?

j) How well do the intercept and slope values estimate these quantities?

k) Obviously, there is variability present in our data. Not all bags with the same number of M&Ms will weigh the same due to differences in the weight of each M&M and differences in the weight of the bag. In our data, more possible sources of variability were introduced when the data were collected. Name at least one of these sources.

Activity 2: Use Analyze – Fit Y by X with Contents Weight as the Y, Response and Number as the X, Factor. Fit a line to the data and have JMP plot the residuals. Turn in this output with the answers to the questions below.

a) Give the equation of the least squares regression line for predicting the contents weight of the bag from the number of M&Ms in the bag.

b) Give the value of the estimated y-intercept for this regression. Give the interpretation of the intercept value in the context of the problem.

b) Look closely at the interpretation of the intercept value. What should the true value of the intercept be?

c) Give the value of the slope for this regression. Give the interpretation of the slope value in the context of the problem.
e) How does this value compare to the value of the slope in Activity 1? What should be the true relationship between the slope in Activity 1 and the slope in Activity 2?

**Activity 3:** Use Analyze – Fit Y by X with Total Weight as the Y, Response and Contents Weight as the X, Factor. Fit a line to the data and have JMP plot the residuals. Turn in this output with the answers to the questions below.

a) Give the equation of the least squares regression line for predicting the total weight of the bag from the contents weight of M&Ms in the bag.
b) Give the value of the estimated slope for this regression. Give the interpretation of the slope value in the context of the problem.
c) Look closely at the interpretation of the slope value. What should the true value of the slope be?
d) Give the value of the estimated y-intercept for this regression. Give the interpretation of the intercept value in the context of the problem.
e) How does this value compare to the value of the intercept in Activity 1? What should be the true relationship between the intercept in Activity 1 and the intercept in Activity 3?
Stat 101 L: Laboratory 6 – Answer Sheet

Names: _________________________     _________________________

_________________________     _________________________

Activity 1:

a) Describe the general relationship between the number of M&Ms in each bag and the total weight of the bag. Does it appear this relationship is strong or weak?

b) Give the equation of the least squares regression line for predicting the total weight of the bag from the number of M&Ms in the bag.

c) Find the predicted total weight of a bag of M&Ms if the bag contains 21 M&Ms.

d) Calculate the residual for the bag with 21 M&Ms and a total weight of 20 grams.

e) Study the residual plot. Describe what you see and how this might affect (or not affect) the least squares regression of the two variables.

f) Give the value of the estimated y-intercept for this regression. Give the interpretation of the intercept value in the context of the problem.

g) Look closely at the interpretation of the intercept value. What physical quantity does the intercept estimate?
h) Give the value of the estimated slope for this regression. Give the interpretation of the slope value in the context of the problem.

i) Look closely at the interpretation of the slope value. What physical quantity does the slope estimate?

j) How well do the intercept and slope values estimate these quantities?

k) Obviously, there is variability present in our data. Not all bags with the same number of M&Ms will weigh the same due to differences in the weight of each M&M and differences in the weight of the bag. In our data, more possible sources of variability were introduced when the data were collected. Name at least one of these sources.

Activity 2:

 a) Give the equation of the least squares regression line for predicting the contents weight of the bag from the number of M&Ms in the bag.

 b) Give the value of the estimated y-intercept for this regression. Give the interpretation of the intercept value in the context of the problem.

 c) Look closely at the interpretation of the intercept value. What should the true value of the intercept be?

 d) Give the value of the slope for this regression. Give the interpretation of the slope value in the context of the problem.
e) How does this value compare to the value of the slope in Activity 1? What should be the true relationship between the slope in Activity 1 and the slope in Activity 2?

Activity 3:

a) Give the equation of the least squares regression line for predicting the total weight of the bag from the contents weight of M&Ms in the bag.

b) Give the value of the estimated slope for this regression. Give the interpretation of the slope value in the context of the problem.

c) Look closely at the interpretation of the slope value. What should the true value of the slope be?

d) Give the value of the estimated y-intercept for this regression. Give the interpretation of the intercept value in the context of the problem.

e) How does this value compare to the value of the intercept in Activity 1? What should be the true relationship between the intercept in Activity 1 and the intercept in Activity 3?