

Statistics 104 - Laboratory 4

Modeling linear relationships

1. M&Ms

On last week's lab we looked at data on a random sample of 10 Fun Size bags of M&Ms. In the table below are the Weight (g) (M&Ms plus bag) and the number of M&Ms for each of the 10 Fun Size Bags of M&Ms.

Bag	1	2	3	4	5	6	7	8	9	10
Number, x	18	19	17	19	20	17	17	19	19	21
Weight (g), y	17	18	16	17	18	15	16	19	17	20

$$s_x = 1.3499, s_y = 1.4944, r = 0.8923, \bar{x} = 18.6, \bar{y} = 17.3$$

- Calculate the least squares regression slope estimate.
- Give an interpretation of the slope estimate within the context of the problem.
- Calculate the least squares regression y-intercept estimate.
- The interpretation of the y-intercept estimate is the predicted value of y (Weight (g)) when x (Number) equals zero. What is the physical interpretation of this within the context of the problem?
- Does the value of the y-intercept estimate seem appropriate given the context of the problem? Explain briefly.
- Use your prediction equation to predict the Weight (g) of a Fun Size bag that contains 20 M&Ms. What is the residual for this prediction?
- Put the least squares regression line on the plot of the data. It should be clear that you have used the regression line to do the plot.
- What is the value of R^2 ? Give an interpretation of this value.

2. M&Ms: Using JMP with a larger sample

Go to the course web page and open the JMP data table that contains the Weight (g) and Number of a sample of 40 Fun Size Bags of M&Ms. Use Analyze – Distribution to summarize the variable Weight (g). Be sure to include a histogram with a count axis, a box plot, Quantiles, Moments and a Stem and Leaf. Print off a copy of this output and use it to answer the following questions.

- Describe the shape of the distribution.
- Does the sample mean and sample median values agree with your description of the shape of the distribution? Explain briefly.
- Report the values of appropriate summary statistic for the center and spread. Why did you choose these summaries?

Use Analyze – Fit Y by X to fit a straight line to the relationship between Number (x) and Weight (g) (y). Be sure to include a Plot of Residuals. Use the JMP output to answer the following questions.

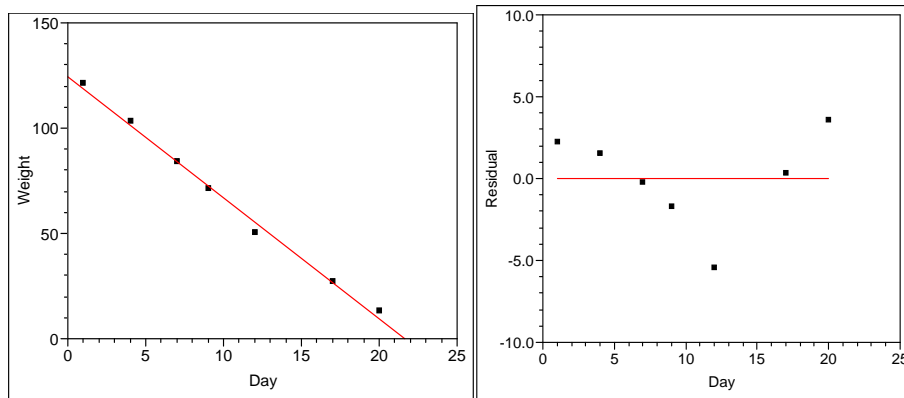
- d. Give the prediction equation for the line relating Number of M&Ms to the Weight (g).
- e. How does this equation compare to the one you found in Problem 1?
- f. Use this equation to predict the Weight (g) of a bag that contains 20 M&Ms. How does this prediction compare to an actual weight of 18 g?
- g. How much of the variation in Weight (g) can be explained by the linear relationship with Number?
- h. Describe the plot of residuals. What does this indicate about the predictions made using the prediction equation in d.?

3. A high school student in Australia collected data on the weight (grams) of a bar of soap and the number of days since the bar was first used.

Days in use	1	4	7	9	12	17	20
Weight (g)	121	103	84	71	50	27	13

We wish to be able to predict the weight of the bar given the number of days since the bar was first used. Use the JMP output provided below and your knowledge of regression analysis to answer the following questions.

- a. Give the prediction equation for the line relating days in use to weight.
- b. Give an interpretation, within the context of the problem, of the estimated slope.
- c. Give an interpretation, within the context of the problem, of the estimated y-intercept.
- d. Use the prediction equation to predict the weight of the bar after 7 days in use. Also calculate the residual for this prediction.
- e. Give the value of R^2 and an interpretation of this value.
- f. Describe the pattern in the plot of residuals vs. day in use. What does this indicate about the straight line model for these data?



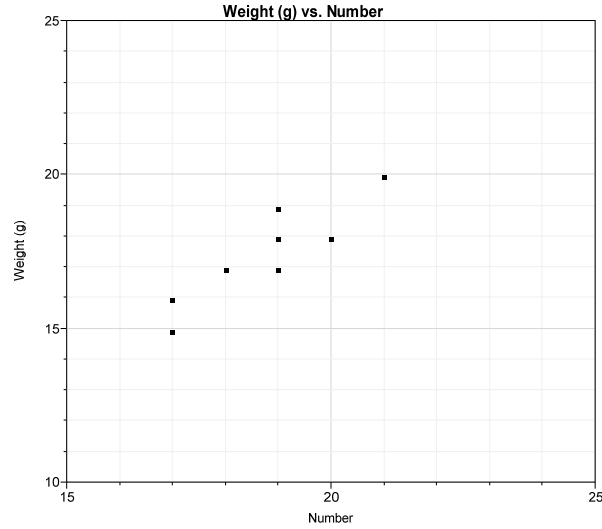
Linear Fit

$$\text{Weight} = 124.53571 - 5.7535714 \text{ Day}$$

Summary of Fit

RSquare	0.994315
RSquare Adj	0.993178
Root Mean Square Error	3.255654
Mean of Response	67
Observations (or Sum Wgts)	7

g. Plot regression line.



h. What is the value of R^2 ? Give an interpretation of this value.

2. M&Ms: Using JMP with a larger sample

- Describe the shape of the distribution of Weight (g).
- Does the sample mean and sample median values agree with your description of the shape of the distribution? Explain briefly.
- Report the values of appropriate summary statistic for the center and spread. Why did you choose these summaries?
- Give the prediction equation for the line relating Number of M&Ms to the Weight (g).

- e. How does this equation compare to the one you found in Problem 1?

- f. Use this equation to predict the Weight (g) of a bag that contains 20 M&Ms. How does this prediction compare to the actual weight of 18 g?

- g. How much of the variation in Weight (g) can be explained by the linear relationship with Number?

- h. Describe the plot of residuals. What does this indicate about the predictions made using the prediction equation in d.?

3. Bar of Soap

- a. Prediction equation.

- b. Interpretation of slope estimate.

- c. Interpretation of y-intercept estimate.

- d. Prediction and residual.

- e. Value of R^2 and interpretation.

- f. Describe pattern of residuals. What does this indicate?