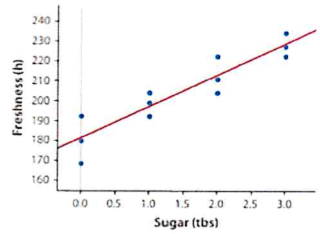
Name: ________________________ Class: ___________________ Date: __________

Probability & Statistics Chapter 2 Test Review Worksheet

1. Does adding sugar to the water in a vase help flowers stay fresh? To find out, two statistics sudents went to a flower shop and randomly selected 12 carnations. When they got home, the students prepared 12 identical vases with exactly the same amount of water in each vase. They put 1 tablespoon of sugar in 3 vases, 2 tablespoons of sugar in 3 vases, and 3 tablespoons in 3 vases. In the remaining 3 vases, they added no sugar. After the vases were prepared, the students randomly assigned 1 carnation to each vase and observed how many hours each flower continued to look fresh. Here is a scatterplot along with the regression line $\hat{y} = 180.8 + 15.8x$, where x = amount of sugar (in tablespoons) and y = hours of freshness.

$\hat{y} = 180.8 + 15.8(2)$

$\hat{y} = 212.4$



x=2 y

a.) Calculate and interpret the residual for the flower that had 2 tablespoons of sugar and looked fresh for 204 hours.

$y - \hat{y} = 204 - 212.4 = -8.4$

The actual value will be 8.4 lower than the predicted

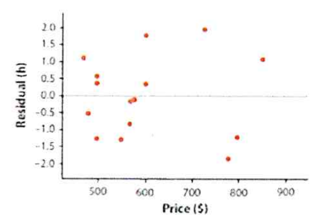
b.) Interpret the slope and y-intercept of the regression line.

15.8 as the sugar increases, your freshness increases by 15.8.

180.8 When you have 0 tablespoons of sugar, the flowers can last 180.8 hours.

2. Can you predict the battery life of a tablet using the price? Using data from a sample of 15 tablets, the least-squares regression line $\hat{y} = 4.67 + 0.0068x$ was calculated using x = ~~prince~~ price (in dollars) and y = battery life (in hours). A residual plot for this model is shown.

pattern X
no pattern ✓



a.) Use the residual plot to determine whether the regression model is appropriate.

Since there is a pattern, it is not appropriate.

b.) Interpret the value $s = 1.21$ for this model.

The actual value will usually be 1.21 away from the predicted.

c.) Interpret the value $r^2 = 0.342$ for this model.

34.2% of the battery life is based off the price.

3. People with diabetes measure their fasting plasma glucose (FPG; measured in units of milligrams per milliliter) after fasting for at least 8 hours. Another measurement, made at regular medical checkups, is called HbA. This is roughly the percent of red blood cells that have a glucose molecule attached. It measures average exposure to glucose over a period of several months. The table gives data on both HbA and FPG for 18 diabetics five months after they had completed a diabetes education class.

Subject	HbA (%)	FPG (mg/ml)	Subject	HbA (%)	FPG (mg/ml)
1	6.1	141	10	8.7	172
2	6.3	158	11	9.4	200
3	6.4	112	12	10.4	271
4	6.8	153	13	10.6	103
5	7.0	134	14	10.7	172
6	7.1	95	15	10.7	359
7	7.5	96	16	11.2	145
8	7.7	78	17	13.7	147
9	7.9	148	18	19.3	255

a.) Using technology, calculate the equation of the least-squares regression line relation y = FPG to x = HbA.

$\hat{y} = 66.429 + 10.408x$

b.) What effect do you think Subject 18 has on the equation of the least-squares regression line?

slope would increase & y-intercept would decrease

c.) Calculate the equation of the least-squares regression line without Subject 18.

$\hat{y} = 52.261 + 12.116x$

4. Many adults try to protect their families by buying life insurance. The policyholder makes regular payments (premiums) to the insurance company in return for the converage. When the insured person dies, a payment is made to designated family members or other beneficiaries.
How do insurance companies decide how much to charge for life insurance? They rely on a staff of highly trained actuaries (people with expertise in probability, statistics, and advanced mathematics) to determine premiums. If someone wants to buy life insurance, the premium will depend on the type and amount of the policy as well as on personal characteristics like age, gender, and health status. The table shows monthly premiums (in dollars) for a 10-year term-life insurance policy worth $1,000,000 for people of various ages (in years).

Age	Premium ($)
40	29
45	46
50	68
55	106
60	157
65	257

a.) Calculate a quadratic model for these data. $\hat{y} = 674.95 - 31.191x + 0.379x^2$

b.) Calculate an exponential model for these data. $\hat{y} = 0.939(1.09)^x$

more scatter is good!!

c.) Which model is more appropriate? Justify your answer.

exponential since most of the residual points are close to the line.

d.) Using your chosen model, calculate and interpret the residual for the 65-year-old.

$$\hat{y} = 0.939(1.09)^x = 0.939(1.09)^{65}$$

$$\hat{y} = 254.3243589$$

$$y - \hat{y} = 257 - 254.3243589 \approx 2.6756$$

The actual premium is 2.6756 more than predicted.