AP Statistics
Unit 02 REDO – HW #2

Name Key 2018
Period

Multiple Choice

1. If (12, 60) is an influential point for the regression line $\hat{y} = 7.908 + 4.098x$, then which of the following must be true?
   a) removal of (12, 60) will improve r
   b) removal of (12, 60) will not affect r
   c) removal of (12, 60) will change the value of the slope of the regression line
   d) (12, 60) has a large residual
   e) none of these

2. Suppose a data set is transformed using $(x, y) \to (x, \log y)$ and a least squares linear regression procedure is performed on the transformed data. If the residual plot of this regression shows a curved pattern, which of the following is an appropriate conclusion?
   a) A quadratic model should be used with the original data
   b) A square root transformation should be applied to the transformed data
   c) The correlation coefficient of the set of transformed data is zero
   d) The exponential transformation is not appropriate
   e) none of these are appropriate

3. After data are collected from an agricultural experiment, suppose a transformation is performed on the bivariate set (inches of water, total plant growth). If the linear regression for the transformed data has the equation:

$$\mathbf{\log(\widehat{growth}) = 0.7 + 1.93\log(water)}$$

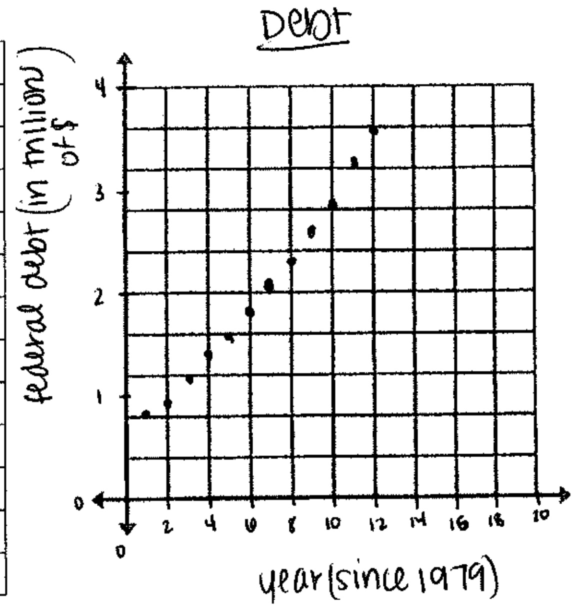
The regression model of the untransformed data is:
   a) $\widehat{growth} = 0.7 + 1.93(water)$
   b) $\widehat{growth} = 5.01 + 1.93(water)$
   c) $\widehat{growth} = (5.01)(1.93^{water})$
   d) $\widehat{growth} = 5.01(water^{1.93})$
   e) none of these

$\widehat{growth} = 10^{0.7}\, water^{1.93}$

Free Response:

1. The following table shows the federal debt for the years since 1979.

Year	Federal Debt (in trillions of $)
1 1980	0.909
2 1981	0.994
3 1982	1.1
4 1983	1.4
5 1984	1.6
6 1985	1.8
7 1986	2.1
8 1987	2.3
9 1988	2.6
10 1989	2.9
11 1990	3.2
12 1991	3.6



a) Construct a scatterplot on the grid provided.

b) Transform the data using the appropriate logarithms. Then, write the LSRL equation for the transformed data.

$\widehat{\log(\text{debt})} = -0.09429 + 0.05586(\text{year})$

$\widehat{\text{debt}} = 10^{-0.09429} \cdot 10^{0.05586(\text{year})}$

$\boxed{\widehat{\text{debt}} = 0.8048 \cdot 1.1373^{(\text{year})}}$

original: curved, $R^2 = 0.9838$
exp.: straight, $R^2 = 0.9918$
power: curved, $R^2 = 0.9156$

mostly random residual plot

c) Describe the scatterplot. What is the correlation coefficient?

The scatterplot of years since 1979 vs. log (federal debt in trillions of dollars) appears to be strong, positive, and linear.

$r = 0.9959$ which is close to 1.

d) What is the coefficient of determination? Interpret this value in context.

$R^2 = 0.9959$ 99.59% of the variation in values of federal debt (in trillions of $) can be explained by the LSRL of federal debt on the value of the year since 1979.

e) Use the model to predict the federal debt in the year 2000. Is it appropriate to do so?

$\hat{y} = 0.8048 \cdot 1.1373^{(21)} = 11.9968$ $x = 21$

the predicted amount of federal debt, according to our model, is 11.9968 million dollars. However, this is not an appropriate calculation. It is extrapolation. $x = 21$ was not in the domain of x values used to create the model.

f) Calculate and interpret the residual value for the federal debt in 1988.

$\hat{y} = 0.8048 \cdot 1.1373^{(9)} = 2.5619$ million dollars $x = 9$

residual $= y - \hat{y} = 2.6 - 2.5619 = 0.03809$ $(9, 2.6)$

The federal debt in 1988 was 0.03809 million dollars more than it is predicted to be by our model.

4. Earthquakes are among the most damaging kinds of natural disasters. The size of an earthquake is generally reported as a rating on the Richter scale-usually a number between 1 and 9. That Richter scale rating indicates the energy released by the shaking of the ground and the height of the shock waves recorded on seismographs.

The data in the following table show Richter scale ratings and amounts of energy released for six earthquakes.

Earthquake Location	Richter Scale Rating	Energy (in sextillion ergs)
San Francisco, CA, 1906	8.25	1500
Yugoslavia, 1963	6.0	0.63
Alaska, 1964	8.6	5000
Peru, 1970	7.8	320
Italy, 1976	6.5	3.5
Loma Prieta, CA, 1989	7.1	28

original: curved $R^2 = 0.5654$

exponential: linear $R^2 = 0.9991$

power: linear $R^2 = 0.9977$ clear pattern in resid. plot

a) Use your calculator to make a scatter plot for this data. Explain which model would be a good model to use and why.

We will use an exponential model/transformation. The R^2 value is extremely close to 1. The scatterplot of x (Richter scale rating) and log(y) (energy) is linear, positive, and strong, and there is no pattern in the residual plot of this model.

b) Write a linear regression model for this data.

$\log(\widehat{energy}) = -9.2132 + 1.5017(\text{RS rating})$

$\widehat{energy} = 10^{-9.2132} \cdot 10^{1.5017(\text{RS rating})}$

$\boxed{\widehat{energy} = 10^{-9.2132} \cdot 31.7468^{(\text{RS rating})}}$

c) How confident would you be with predicting the Energy if the Richter Scale Rating was 6.3? Why?

I would be fairly confident. Although our model is based off of only 6 data points, the Richter Scale rating of 6.3 is within the domain of x values used to create the model. The scatterplot also shows a strong, positive, linear relationship. used to create this model

5. The femur is a large bone found in the leg or hind limb of an animal. Scientists use the circumference of an animal's femur to estimate the animal's weight. The ~~bale~~ table at the right shows the femur circumference C (in millimeters) and the weight W (in kilograms) of several animals.

a) Use your calculator to make a scatter plot for this data. Explain which model would be a good model to use and why.

original: curved
exponential: curved
power: linear

We will use a power model transformation. The scatterplot of log(circumference) vs. log(weight) appears linear, has a high r value (r = 0.9912), and there is random scatter in the residual plot.

Animal	C (mm)	W (kg)
Meadow mouse	5.5	0.047
Guinea pig	15	0.385
Otter	28	9.68
Cheetah	68.7	38
Warthog	72	90.5
Nyala	97	134.5
Grizzly bear	106.5	256
Kudu	135	301
Giraffe	173	710

▶ Source: Zoological Society of London

b) Write a linear regression model for this data.

$\log(\widehat{weight}) = -3.4574 + 2.8343\log(\text{circumference})$

$\boxed{\widehat{weight} = 10^{-3.4574}(\text{circumference})^{2.8343}}$

c) Predict the weight of an animal whose femur has a circumference of 21 mm.

$\widehat{weight} = 10^{-3.4574}(21)^{2.8343} = 1.9506 \text{ kg}$

An animal whose femur has a circumference of 21 mm is predicted to weight 1.9506 kg.