AP Statistics
Unit 02 – Bivariate Data
Day 05 Notes

Name________________________________
Period_______

Some scatter plots will not have a linear line that best describes the data. In this section, we see what other mathematical models we can provide that best fit the data.

Definition and Properties of Logarithms:

1. $\log(AB) = logA + logB$
2. $log(A/B) = logA - log\ B$
3. $logx^p = p \cdot logx$
4. $log_b y = x$ then $y = b^x$

Non-Linear Regression

1. **Plot data and check conditions if given (can do in calculator)**

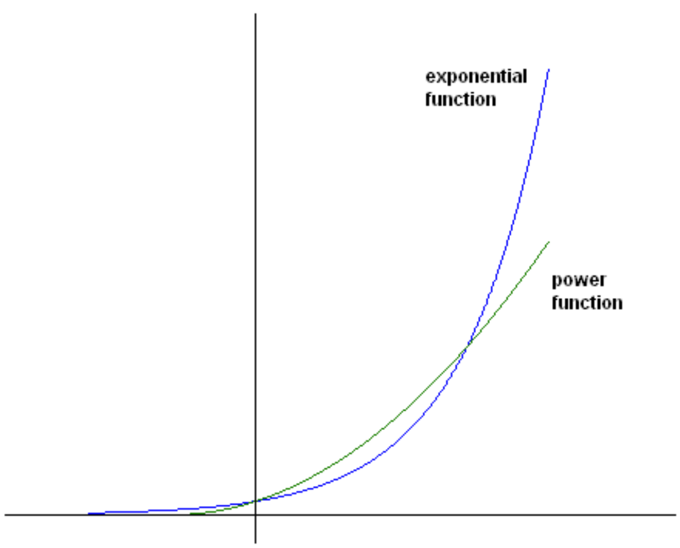
2. **Transform x and y using log or ln (in L3 and L4 in calculator)**

3. **Create two equations in LinReg (a +bx)**
   - **x (L1) and log y (L4), write down *r***
   - **log x (L3) and log y (L4), write down *r***

4. **Which has a better correlation, Exponential Model or Power Model**

5. **Determine LSRL of transformed data**

6. **Un-transform LSRL to obtain non-linear model**



Exponential Model:

Transformation of Exponential Model

$\boldsymbol{y = ab^x}$

Exponential models become linear when we apply the logarithm transformation to the response variable (y)

Therefore, if we <u>plot x against log(y)</u>, the plot should be linear

Un-transforming to create Exponential Model:

$$log\hat{y} = a + bx$$

$$\hat{y} = 10^{a+bx}$$

$$\hat{y} = 10^a \cdot 10^{bx}$$

Examples for Exponential Functions:

1. $\widehat{\log y} = 3.5 + .23x$

2. $\widehat{\log y} = -2.1 - 5.3x$

3. $\widehat{\log y} = 5 + .89x$

Power law model:

Transformation of Power Law Model:

$\boldsymbol{y = ax^b}$

Power law models become linear when we apply the logarithm transformation to both variables

Therefore, if we plot <u>log(x) against log (y)</u>, the plot should be linear

Un-transforming to create power low model:

$$\widehat{\log y} = a + b \cdot \log x$$

$$\hat{y} = 10^{a+b\cdot logx}$$

$$\hat{y} = 10^{a} \cdot 10^{b\cdot logx}$$

$$\hat{y} = 10^{a} \cdot (10^{logx})^{b}$$

$$\hat{y} = 10^{a} \cdot x^{b}$$

Examples for Power Functions:

1. $\widehat{\log y} = 1.2 + 0.75 \cdot logx$

2. $\widehat{\log y} = -2.6 - 1.3 \cdot logx$

3. $\widehat{\log y} = 0.5 + 2.8 \cdot logx$

To determine which transformation is better:

1. Check the value of r for both. The closer to -1 or 1, the better.
2. Look at the transformed plots. Which one is more straight? Can't tell? Then…
3. Look at the residual plots. We want no pattern, no curve.

Example:

On July 31, 2005, a team of astronomers announced that they had discovered what appeared to be a new planet in our solar system. They had first observed this object almost two years earlier using a telescope at Caltech's Palomar Observatory in California. Originally named UB313, the potential planet is bigger than Pluto and has an average distance of about 9.5 billion miles from the sun (for reference, Earth is about 93 million miles from the sun). Could this new astronomical body, now called Eris, be a new planet?

At the time of the discovery, there were nine known planets in our solar system. Here are the data on the distance from the sun and period of revolutions of those planets. Note that distance is measure in astronomical units (AU), which is the number of Earths distances an object is from the sun:

Planet	**Distance from Sun (AU)**	**Period of Revolutions (Earth Years)**
Mercury	0.387	0.241
Venus	0.723	0.615
Earth	1.000	1.000
Mars	1.524	1.881
Jupiter	5.203	11.862
Saturn	9.539	29.456
Uranus	19.191	84.070
Neptune	30.061	164.810
Pluto	39.529	248.530

a) Determine which function best represents this data: Exponential or Power and explain why.

b) Write the LSRL equation for the <u>transformed</u> data in context or identifying your variables.

c) Now perform an inverse transformation on your linear equation to obtain a model for the original federal debt data. Write the equation for this model.

d) Use your model to prove your function is accurate by determining the period of revolutions if the planet is Saturn.

e) If another mass were to be found at a distance of 25 AU's, what would be the period of revolutions?

Example:

Expose marine bacteria to X-rays for time periods from 1-15 minutes. Here are the number of surviving bacteria (in hundreds) on a culture plate after each exposure time:

Time (t)	Count (y)
1	355
2	211
3	197
4	166
5	142
6	106
7	104
8	60
9	56
10	38
11	36
12	32
13	21
14	19
15	15

a) Determine which function best represents this data: Exponential or Power and explain why.

b) Write the LSRL equation for the <u>transformed</u> data in context or identifying your variables.

c) Now perform an inverse transformation on your linear equation to obtain a model for the original federal debt data. Write the equation for this model.

d) Use your model to prove your function is accurate by determining the number of bacteria after 9 minutes.

e) Use your model to predict the number of surviving bacteria after 17 minutes. Show your work. Is it appropriate to use this equation to calculate the number of surviving bacteria after 17 minutes?

Overview

Model	Exponential Model	Power Model
Plot	x vs. log(y)	log(x) vs. log(y)
In Calculator	L1 vs. L4	L3 vs. L4
Transformed	$log\hat{y} = a + bx$	$\widehat{\log y} = a + b \cdot \log x$
Un-transformed	$\hat{y} = 10^a \cdot 10^{bx}$	$\hat{y} = 10^a \cdot x^b$
Calculations	1. Enter data into L1 and L2. 2. Plot data and inspect to confirm exponential shape. 3. Select L4 and enter log(L2) to compute log y values in L4. 4. Plot L1 and L4 to confirm linear shape. 5. Enter 8. LinReg(a + bx) L1,L4,Y1 to calculate LSRL of x and log y. 6. Inspect residual plot to confirm fit of linear model. 7. Write LSRL of x and log y. 8. Un-transform to get exponential model.	1. Enter data into L1 and L2. 2. Plot data and inspect to confirm power model shape. 3. Select L3 and enter log(L1) to compute log x values in L3. 4. Select L4 and enter log(L2) to compute log y values in L4. 5. Plot L3 and L4 to confirm linear shape. 6. Enter LinReg(a + bx) L3,L4,Y1 to get LSRL of log x and log y. 7. Inspect residual plot to confirm fit of linear model. 8. Write LSRL of log x and log y. 9. Un-transform to get power model.