Student Workbook

Fitting Exponential Models to Data

Name: ______________________ Date: ______________ Class: ________

ESSENTIAL QUESTION

When real-world data doesn't follow a perfect formula, how do we find the exponential or logarithmic model that best fits the data — and how do we know how good that fit is?

Lesson Overview

Real data is messy. Rather than finding an exact formula, we find the best-fit model — the curve that minimizes total error. For exponential data ($y = a \cdot b^x$), the key technique is linearization: take ln(y) to transform the curved data into a straight line, then apply linear regression. The slope gives ln(b) and the y-intercept gives ln(a). With only two points, the two-point method gives an exact model. R^2 measures how well the model fits: $R^2=1$ is a perfect fit; R^2 near 0 means the model explains little of the variation.

$$y = a \cdot b^x \to \ln(y) = \ln(a) + x \cdot \ln(b)$$
$$y = a + b \cdot \ln(x) \text{ (logarithmic model)}$$

Exp.	Linearize with $Y=\ln(y)$; regress on (x,Y); recover $b=e^{\text{slope}}$, $a=e^{\text{intercept}}$
Log	Substitute $X=\ln(x)$; regress on (X,y); slope and intercept give b and a directly
R^2	0 to 1 — measures the fraction of variation in y explained by the model

Two-Point Method & R^2 Scale

- Two-point method ($y=a \cdot b^x$): divide to get $b=(y2/y1)^{1/(x2-x1)}$, then $a=y1/b^{x1}$.
- R^2 scale: 0-0.5 poor, 0.5-0.7 weak, 0.7-0.85 moderate, 0.85-0.95 good, 0.95-1.0 excellent.
- Quick check: constant ratios between consecutive y-values -> exponential. Constant differences vs ln(x) -> logarithmic.

Worked Examples

EXAMPLE 1

Use the two-point method to find $y = a \cdot b^x$ through (1,6) and (4,48).

- Write two equations: $6=a \cdot b^1$ and $48=a \cdot b^4$
- Divide: $48/6 = b^4/b^1 \to 8 = b^3 \to b = 8^{1/3} = 2$
- Solve for a: $6 = a \cdot 2 \to a = 3$. Verify: $y(1)=6$ ✓ $y(4)=3 \cdot 16=48$ ✓

Answer: $y = 3 \cdot 2^x$

EXAMPLE 2

Bacteria (thousands) at x=0,1,2,3: y=2.1, 3.2, 4.8, 7.3. Find the best-fit exponential model.

- Compute Y=ln(y): 0.742, 1.163, 1.569, 1.988
- Linear regression on (x,Y): slope m ≈ 0.415, intercept c ≈ 0.742
- Recover: $b = e^{0.415} \approx 1.515$, $a = e^{0.742} \approx 2.10$

Answer: $y \approx 2.10 \cdot (1.515)^x$

EXAMPLE 3

Data (x,y): (1,5.2) (2,6.8) (4,8.9) (8,11.4) (16,14.1). Exponential or logarithmic model?

- Test exponential: Y=ln(y) vs x is NOT roughly linear (curve bends) -> poor exponential fit
- Test logarithmic: X=ln(x) vs y IS roughly linear -> good logarithmic fit
- Regression on (X,y): slope b ≈ 3.21, intercept a ≈ 5.2

Answer: Logarithmic model is better: $y \approx 5.2 + 3.21 \cdot \ln(x)$

EXAMPLE 4

A regression gives $y = 4.2 \cdot (0.87)^x$ with $R^2 = 0.982$. Interpret the model and R^2.

- b = 0.87 < 1, so this is exponential decay — about 13% decrease per unit x
- Initial value (x=0): $y = 4.2 \cdot (0.87)^0 = 4.2$
- $R^2 = 0.982$ means 98.2% of the variation in y is explained — an excellent fit

Answer: Decay; ~13% decrease per unit; R^2=0.982 -> excellent fit

EXAMPLE 5

Use the two-point method to find $y = a + b \cdot \ln(x)$ through (1,3) and $(e^2,9)$.

- First point: $3 = a + b \cdot \ln(1) = a + 0$ -> $a = 3$
- Second point: $9 = 3 + b \cdot \ln(e^2) = 3 + 2b$ -> $6 = 2b$ -> $b = 3$
- Verify: y(1)=3 ✓ $y(e^2)=3+3 \cdot 2=9$ ✓

Answer: $y = 3 + 3 \cdot \ln(x)$

Guided Practice

Work through each problem below. Use the hint if you get stuck — write your final answer on the last line.

1 Use the two-point method to find $y=a \cdot b^x$ through (0,5) and (3,40).

Hint: With x=0: $y=a \cdot b^0=a$, so a=5 directly. Then use (3,40) to find b.

2 Data: (0,1.0) (1,2.7) (2,7.4) (3,20.1). Compute Y=ln(y) and find the slope and intercept of the linearized data.

Hint: ln(1.0)=0, ln(2.7)≈0.993, ln(7.4)≈2.001, ln(20.1)≈3.000 — nearly 0,1,2,3. What does that tell you about the slope?

3 An exponential model $y=12\cdot(0.75)^x$ has $R^2=0.94$. (a) Growth or decay? (b) Percent decrease per unit? (c) Fit quality?

Hint: b=0.75<1 means decay. Percent decrease = (1-b)·100%. $R^2=0.94$ falls in the 'good' range.

4 Find a logarithmic model y=a+b·ln(x) through (e,7) and $(e^3,13)$.

Hint: ln(e)=1 and $\ln(e^3)=3$. Write two equations and solve the system for a and b.

5 Data (x, ln y): (1,1.2) (2,1.9) (3,2.6) (4,3.3). Find the exponential model $y=a\cdot b^x$.

Hint: The linearized data is given. Find slope m and intercept c, then $b=e^m$ and $a=e^c$.

Key Vocabulary

Regression	A statistical method for finding the curve of best fit by minimizing the sum of squared errors.
Exponential Regression	Best-fit model $y=a\cdot b^x$, found by linearizing with Y=ln(y) and regressing (x,Y).
Logarithmic Regression	Best-fit model y=a+b·ln(x), found by substituting X=ln(x) and regressing (X,y).
Linearization	Transforming a nonlinear relationship into a linear one by applying a function (like ln) to a variable.
Coefficient of Determination (R^2)	A value from 0 to 1 measuring how well the model explains the variation in the data.
Two-Point Method	Using exactly two data points to find the exact constants a and b in a model.
Residual	The difference between an observed y-value and the model's predicted value.
Extrapolation	Using a model to predict values outside the range of the original data — can be unreliable.

Check Your Understanding

Circle the letter of the correct answer for each question.

1 **An exponential model $y=a\cdot b^x$ passes through (0,7) and (2,63). What is b?** Multiple Choice

- **A** $b = 3$
- **B** $b = 9$
- **C** $b = 7$
- **D** $b = 2$

2 **After linearizing, the slope of the line through (x, ln y) is 0.5 and the y-intercept is 1.2. What is b?** Multiple Choice

- **A** $b = e^{0.5} \approx 1.649$
- **B** $b = e^{1.2} \approx 3.320$
- **C** $b = 0.5$
- **D** $b = 1.2$

3 **An exponential regression gives $R^2 = 0.73$. How would you describe the fit?** Multiple Choice

- **A** Excellent
- **B** Good
- **C** Moderate
- **D** Poor

4 **Which transformation linearizes the model $y = a + b\cdot\ln(x)$?** Multiple Choice

- **A** Let $Y = \ln(y)$
- **B** Let $X = \ln(x)$
- **C** Let $Y = e^y$
- **D** Let $X = e^x$

5 **A model $y=5\cdot(1.3)^x$ is used to predict y at x=20, but the original data only went up to x=8. This is:** Multiple Choice

- **A** Interpolation
- **B** Extrapolation
- **C** Linearization
- **D** Residual analysis

Common Mistakes to Avoid

✗ Recovering a and b from linearization using a=slope and b=intercept directly.

✓ Slope $m=\ln(b)$ and intercept $c=\ln(a)$. You must exponentiate: $b=e^m$ and $a=e^c$.

✗ Applying the two-point method without verifying the model with both original points.

✓ Always substitute both original points back into your model to confirm it fits exactly.

✗ Concluding $R^2=0.85$ means the model is '85% accurate.'

✓ $R^2=0.85$ means the model explains 85% of the variation in y — not point-by-point accuracy.

✗ Using an exponential model for data that grows fast then levels off (like a learning curve).

✓ Rapid-then-slow growth with no upper bound suggests a logarithmic model, not exponential.

Math Tips

★ Quick check for exponential data: ratios y2/y1, y3/y2,... roughly constant -> exponential.

★ Quick check for logarithmic data: differences in y roughly constant for equally spaced ln(x) -> logarithmic.

★ When x=0 is in the dataset, the two-point method is easiest: a=y(0) directly, then solve for b.

★ Extrapolation with exponential models is risky — real-world constraints eventually slow unbounded growth.

★ R^2 closer to 1 is better, but context matters: compare R^2 across candidate models on the SAME data.