AN INTRODUCTION TO THE GENERAL LINEAR MODEL

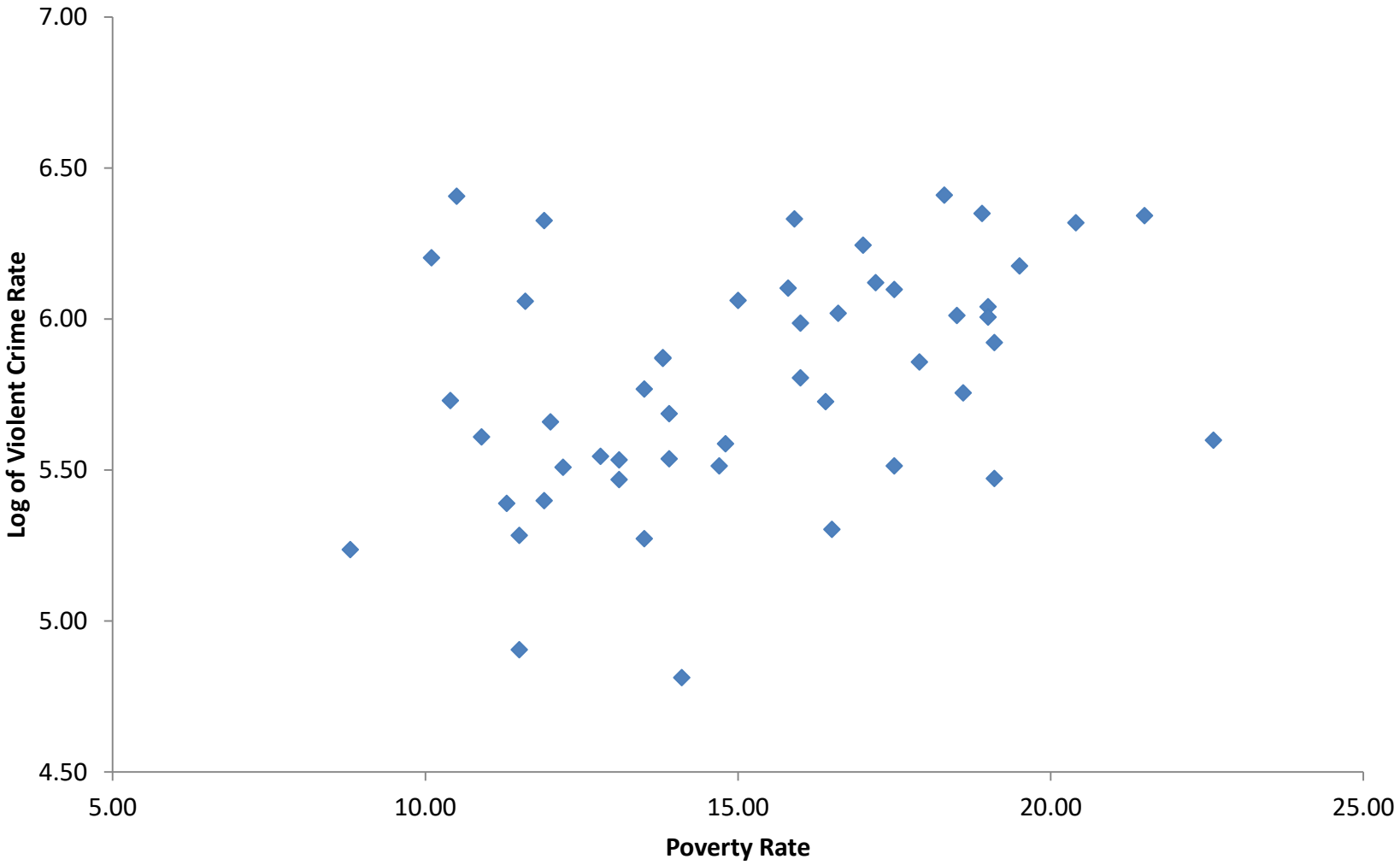
Richard Lee Rogers

About GLM

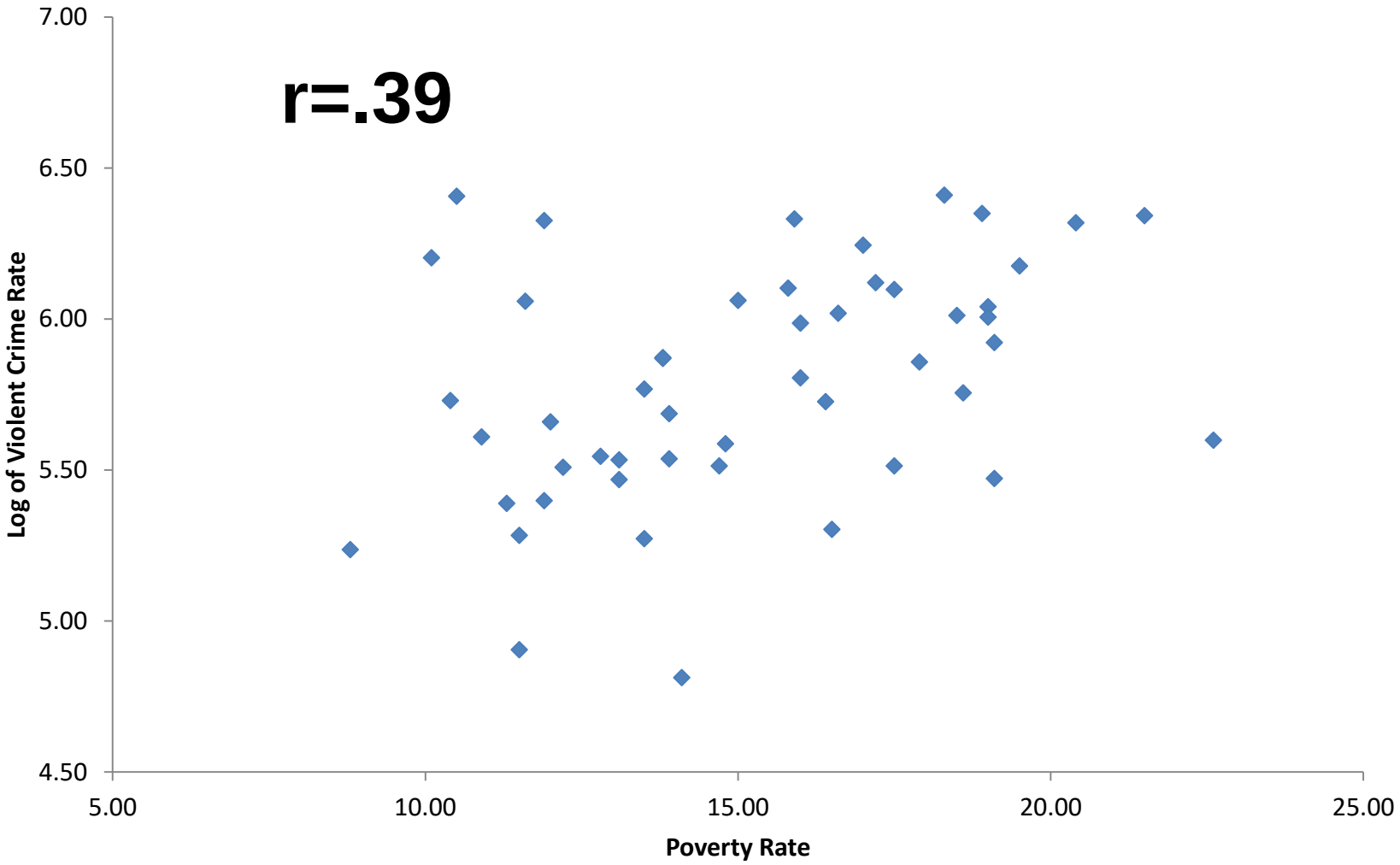
- Widely used in multivariate statistics
  - At least four additional videos in this series are about using GLM to determine relationships
  - Does not include development of diagnostic techniques and corrections based on diagnostics
  - Does not include the rise of classes of statistics to correct for specific situations that GLM does not handle well

- All four elements of an inferential statistic present: direction, magnitude, significance, performance of the model

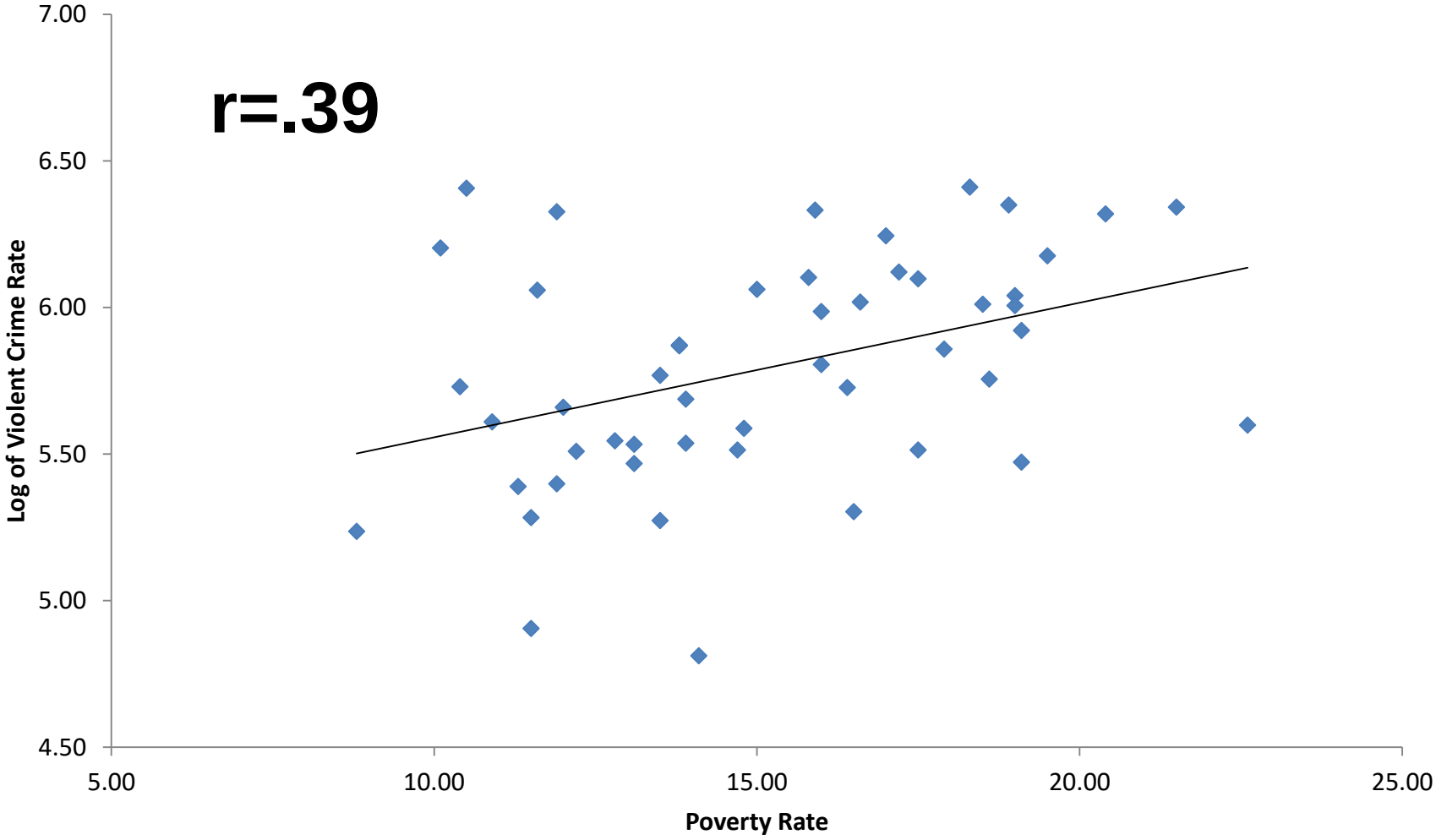
Log of Violent Crime Rate and Poverty Rate



Log of Violent Crime Rate and Poverty Rate



Scatterplot with Line



The Equation of a Line

$y = mx + b$

The Equation of a Line

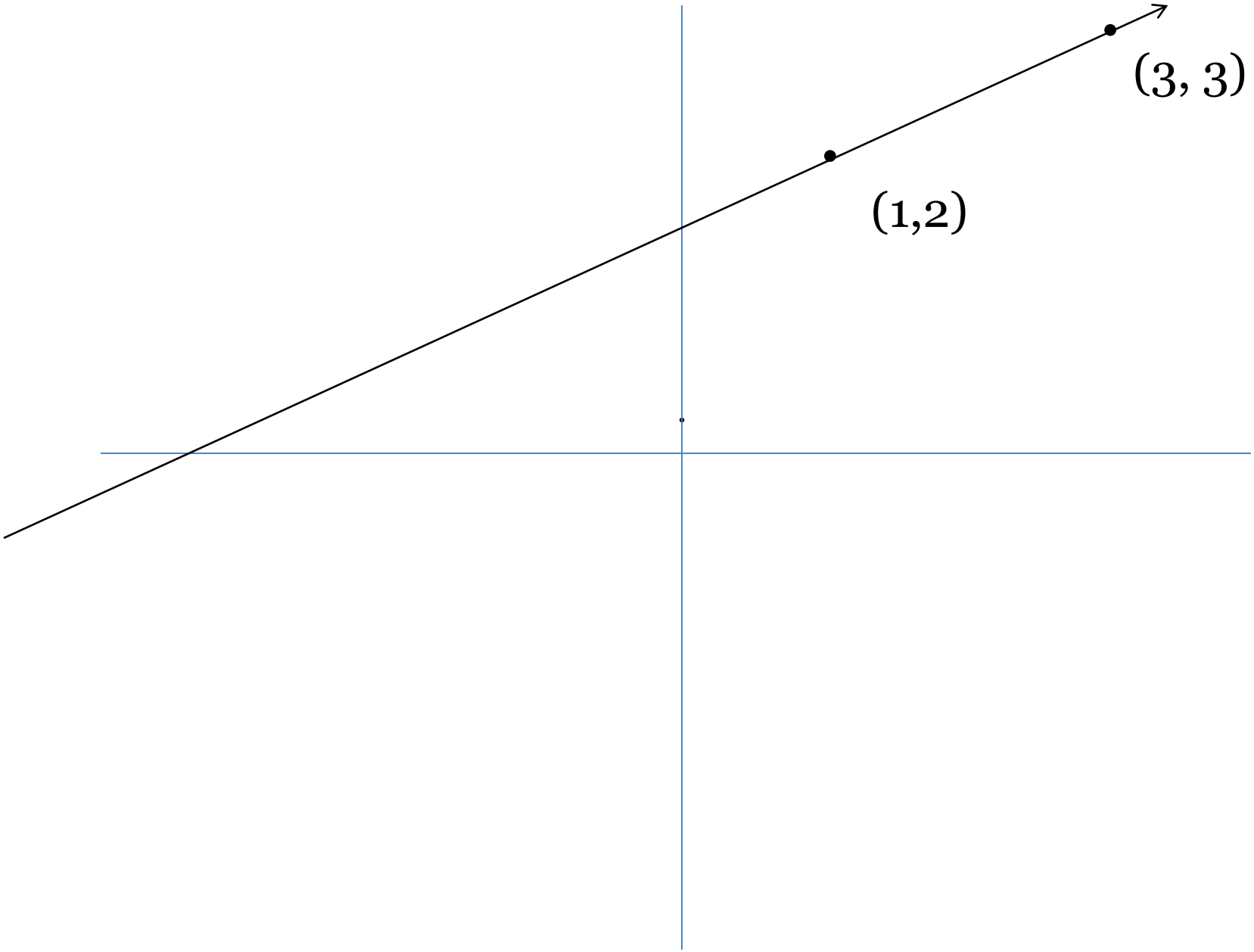
$y = mx + b$

(x,y) = coordinates of a point along the line

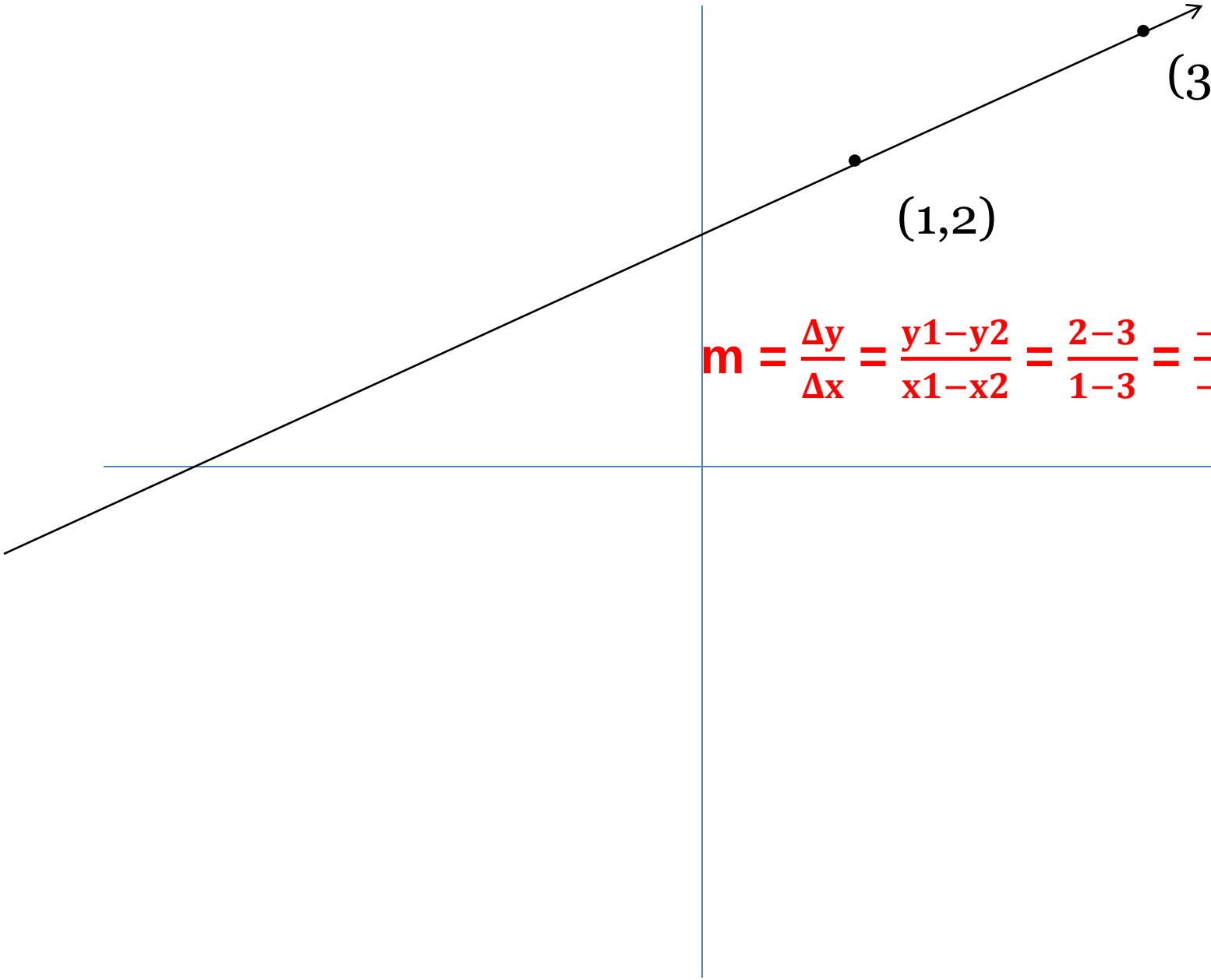
m = slope

b = y-intercept

Finding a Line

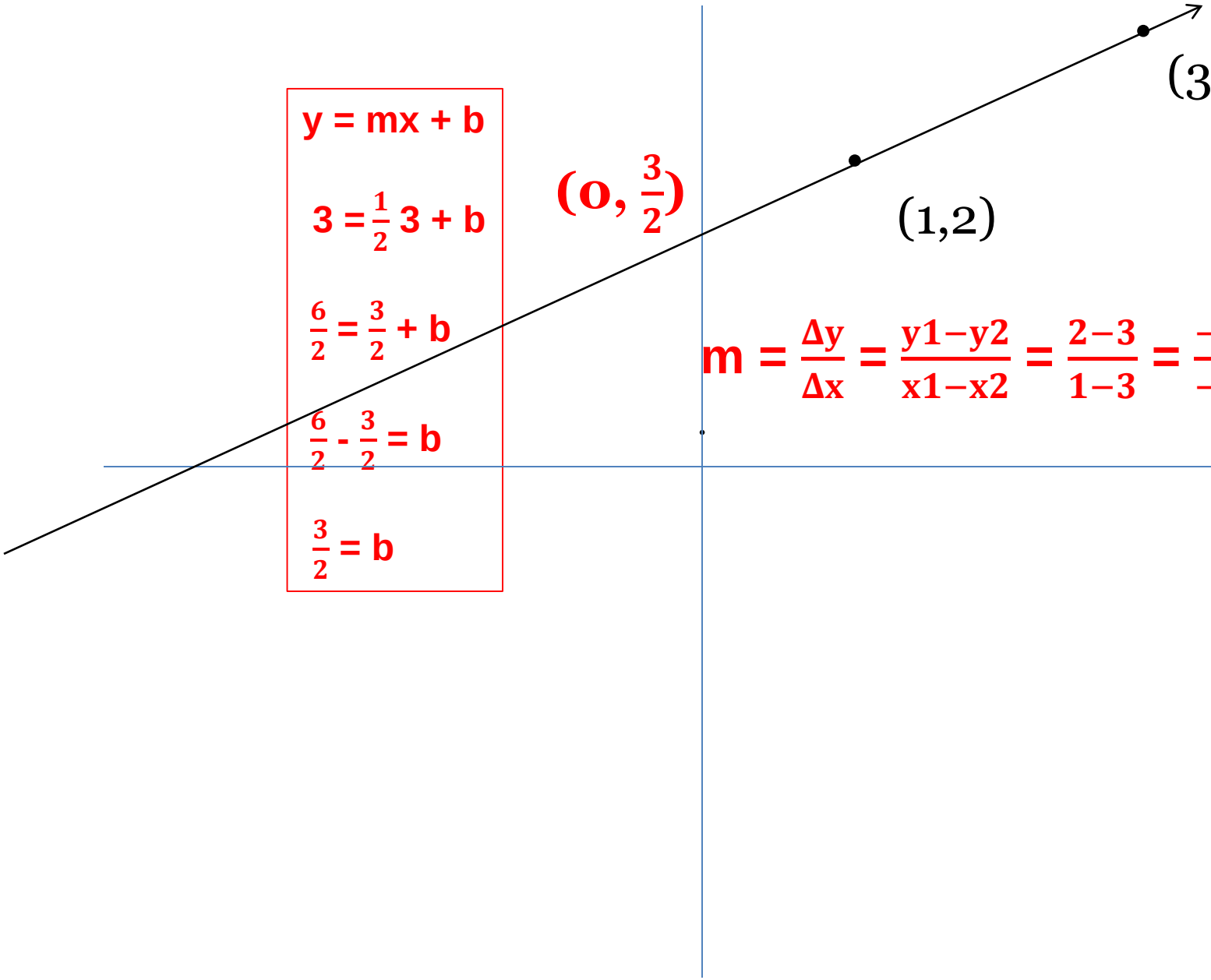


Finding a Line



$$m = \frac{\Delta y}{\Delta x} = \frac{y1-y2}{x1-x2} = \frac{2-3}{1-3} = \frac{-1}{-2} = \frac{1}{2}$$

Finding a Line



$y = mx + b$

$3 = \frac{1}{2}\,3 + b$

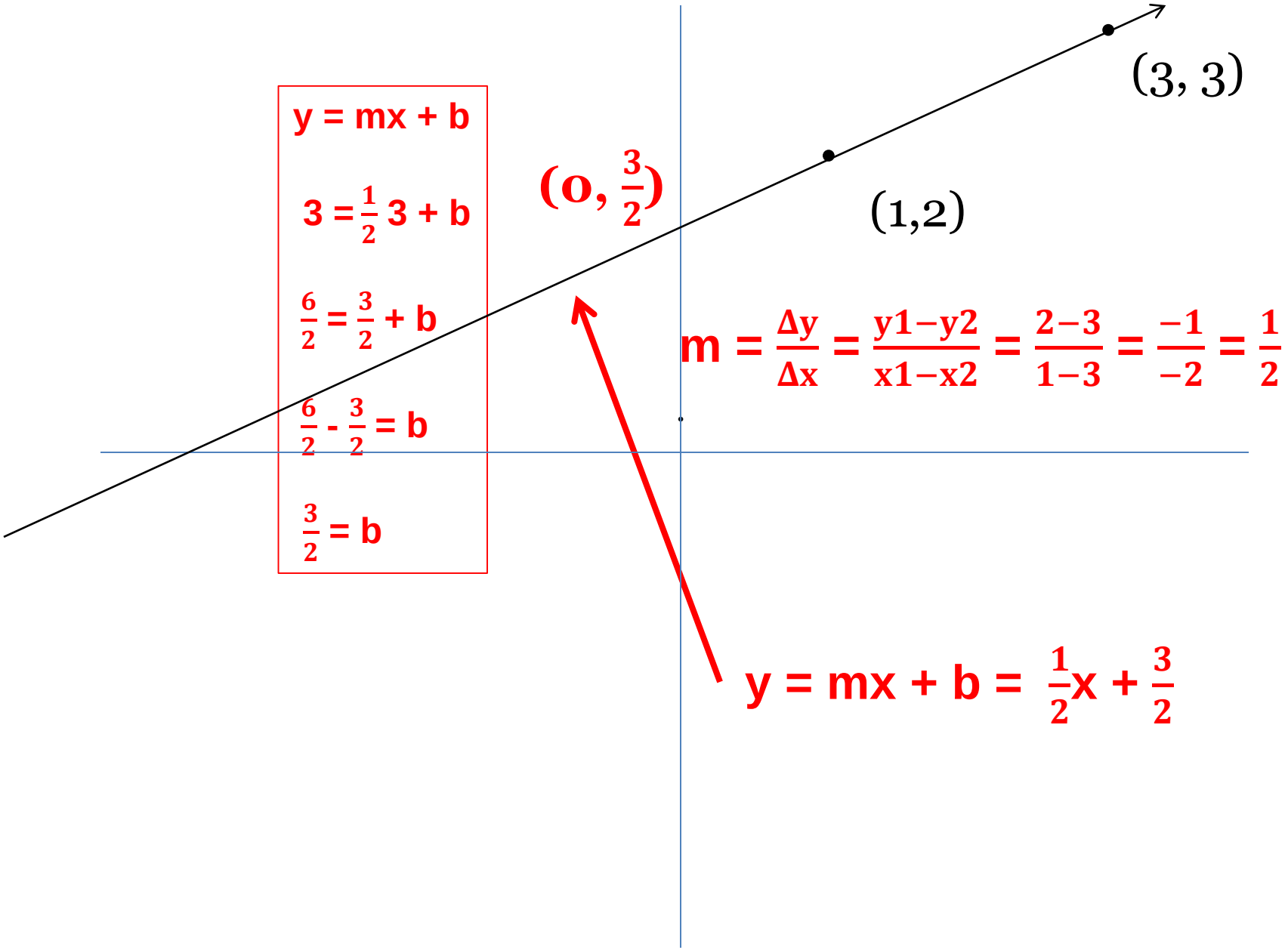
$\frac{6}{2} = \frac{3}{2} + b$

$\frac{6}{2} - \frac{3}{2} = b$

$\frac{3}{2} = b$

$$m = \frac{\Delta y}{\Delta x} = \frac{y1 - y2}{x1 - x2} = \frac{2-3}{1-3} = \frac{-1}{-2} = \frac{1}{2}$$

Finding a Line



The Equation of A Line in Statistics

$y = mx + b$

m = slope

x = y-intercept

The Equation of A Line in Statistics

$y = mx + b$

$y = b + mx$

m = slope

x = y-intercept

The Equation of A Line in Statistics

$y = mx + b$

$y = b + mx$

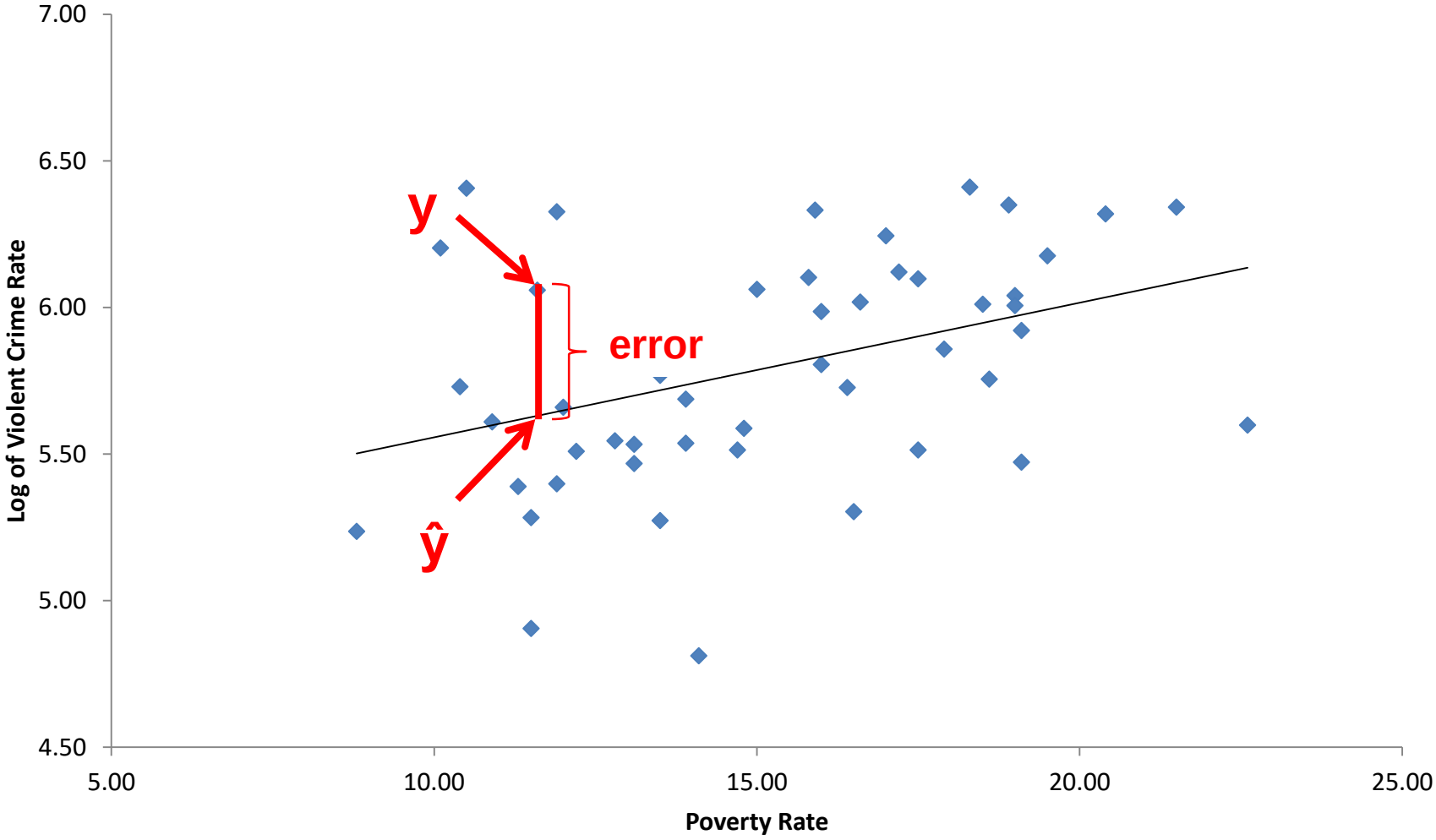
$\hat{y} = \alpha + Bx$

B = slope

α = y-intercept

$\hat{y}$ = y-hat = predicted value of y

The Concept


7.00
6.50
6.00
5.50
5.00
4.50
Log of Violent Crime Rate
5.00
10.00
15.00
20.00
25.00
Poverty Rate
y
ŷ
error

The Equation of A Line in Statistics

$y = mx + b$

$y = b + mx$

$\hat{y} = \alpha + Bx$

$\hat{y} + e = \alpha + Bx + e$

b = slope

a = y-intercept

$\hat{y}$ = y-hat = predicted value of y

e = error

The Equation of A Line in Statistics

$y = mx + b$

$y = b + mx$

$\hat{y} = \alpha + Bx$

$y = \alpha + Bx + e$

b = slope

a = y-intercept

$\hat{y}$ = y-hat = predicted value of y

e = error

The Equation of A Line in Statistics

$y = mx + b$

$y = b + mx$

$\hat{y} = \alpha + Bx$

$\hat{y} + e = \alpha + Bx + e$

$\hat{y} = B_o + B_1x$

$y = B_o + B_1x_1 + e$

B_1 = slope

B or B_0 = y-intercept

$\hat{y}$ = y-hat = predicted value of y

e = error

GLM Assumptions

- The dependent variable is numeric (most of the time)

GLM Assumptions

- The dependent variable is numeric (most of the time)
- Numeric variables have normal distributions

GLM Assumptions

- The dependent variable is numeric (most of the time)
- Numeric variables have normal distributions
- Error terms are homoskedastic

GLM Assumptions

- The dependent variable is numeric (most of the time)
- Numeric variables have normal distributions
- Error terms are homoskedastic
- If multiple X variables, then the variables are independent

Forms of the General Linear Model

- Simple OLS regression
- Multiple OLS regression
- One-way ANOVA
- Two-way ANOVAs
- ANCOVA