



Linear Regression Models

Dinesh K. Vishwakarma, Ph.D.

Learning Objectives

1. Describe the Linear Regression Model
2. State the Regression Modeling Steps
3. Explain Ordinary Least Squares
4. Compute Regression Coefficients
5. Understand and check model assumptions
6. Predict Response Variable

Learning Objectives...

7. Correlation Models
8. Link between a correlation model and a regression model
9. Test of coefficient of Correlation

What is a Model?

1. Representation of Some Phenomenon
2. Non-Maths/Stats Model



What is a Maths/Stats Model?

1. Often Describe Relationship between Variables
2. Types
 - **Deterministic Models (no randomness)**
 - **Probabilistic Models (with randomness)**

Deterministic Models

1. Hypothesize Exact Relationships
2. Suitable When Prediction Error is Negligible
3. Example: Body mass index (BMI) is measure of body fat based.

- **Metric Formula:** $BMI = \frac{\text{Weight in Kilograms}}{(\text{Height in Meters})^2}$

- **Non-metric Formula:** $BMI = \frac{\text{Weight (pounds)} \times 703}{(\text{Height in inches})^2}$

Probabilistic Models

1. Hypothesize 2 Components

- Deterministic
- Random Error

2. **Example:** Systolic blood pressure of newborns is **6 Times** the Age in days + Random Error

- $SBP = 6 \times age(d) + \varepsilon$
- Random Error may be due to factors other than age in days (e.g. Birth weight)

Bivariate & multivariate models

(Education) x \longrightarrow y (Income) **Bivariate**

(Education) x_1
(Sex) x_2
(Experience) x_3
(Age) x_4

\longrightarrow y (Income) **Multivariate**

Model with simultaneous relationship

Price of wheat



Quantity of wheat produced

Regression Modeling Steps

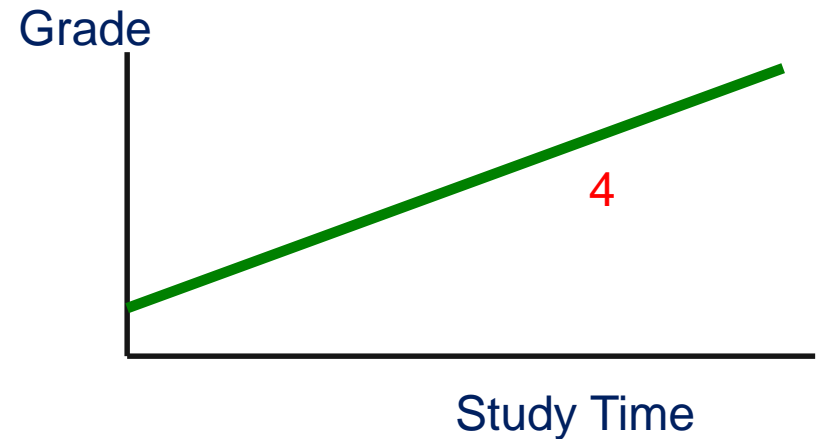
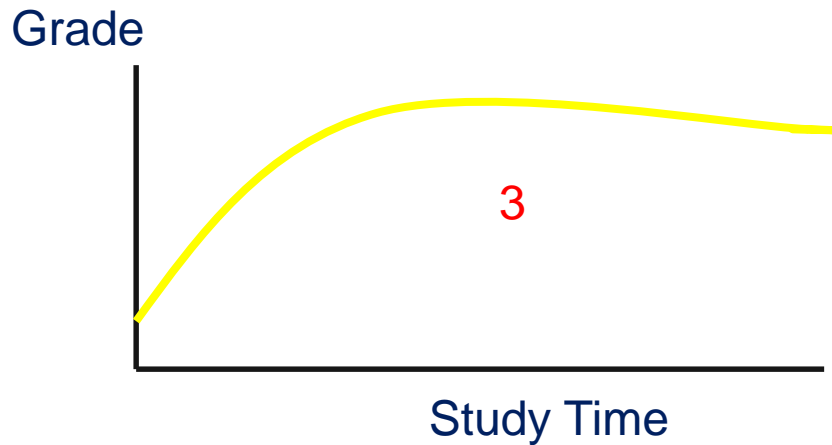
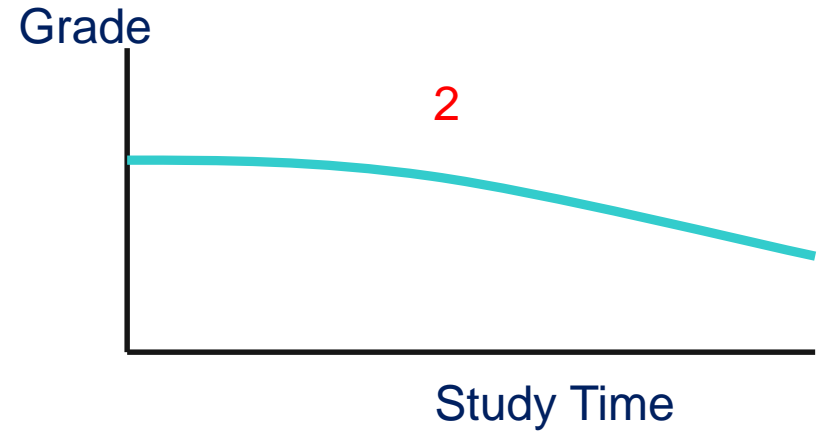
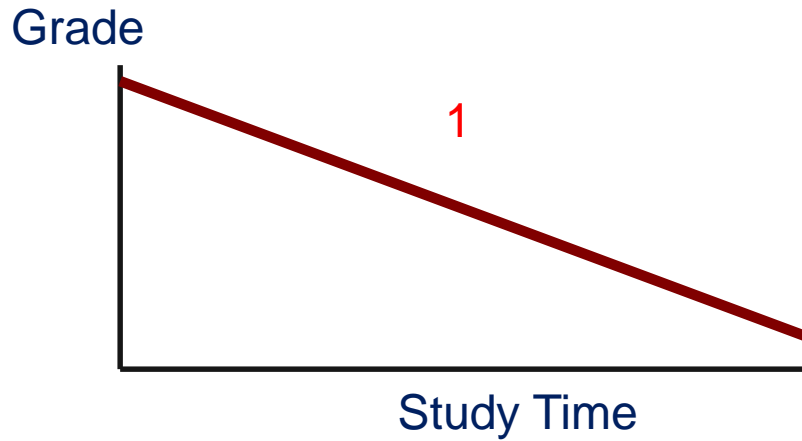
- 1. Hypothesize Deterministic Component
 - Estimate Unknown Parameters
- 2. Specify Probability Distribution of Random Error Term
 - Estimate Standard Deviation of Error
- 3. Evaluate the fitted Model
- 4. Use Model for Prediction & Estimation

Models Facts

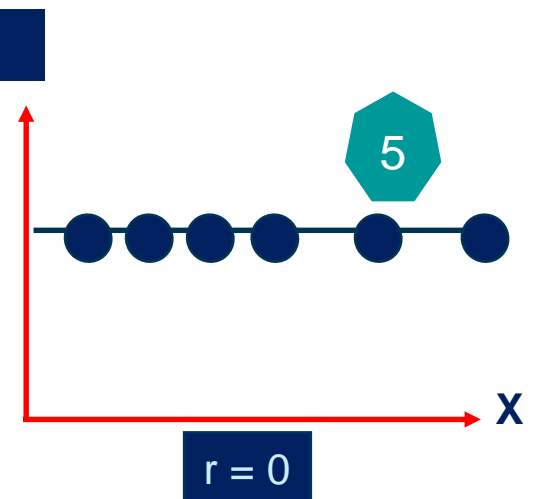
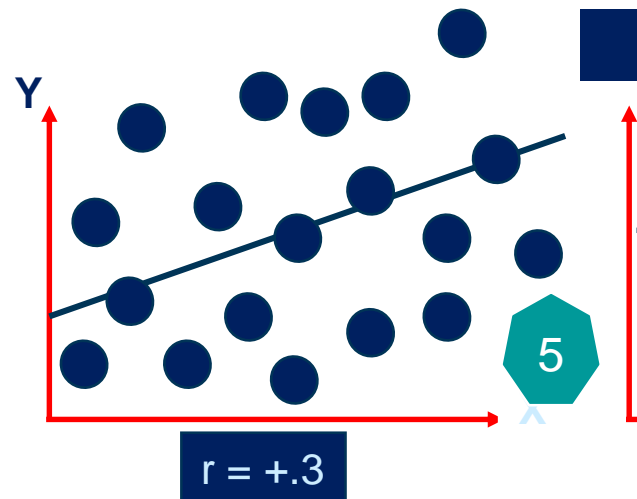
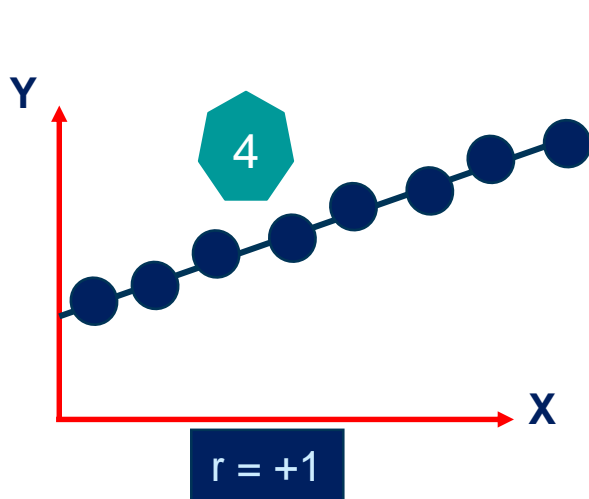
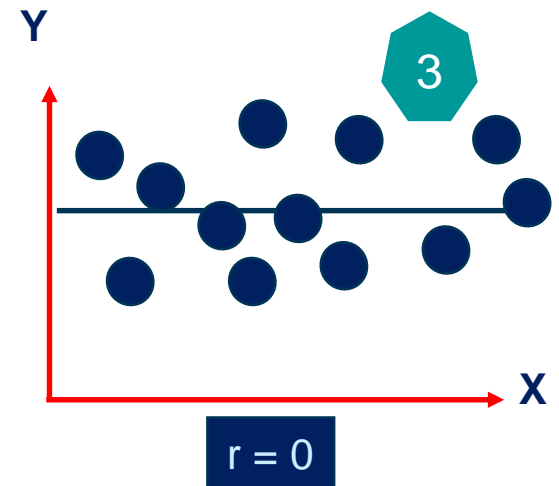
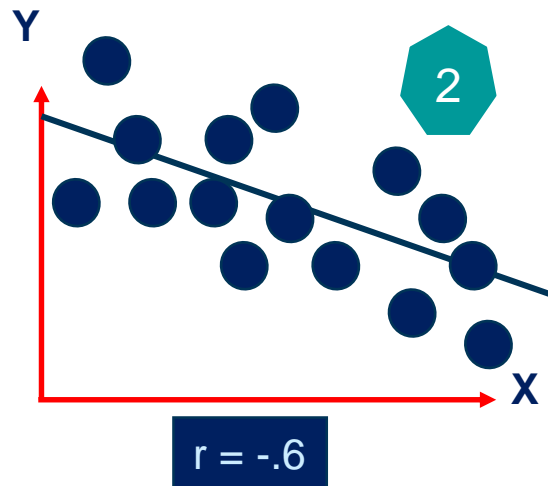
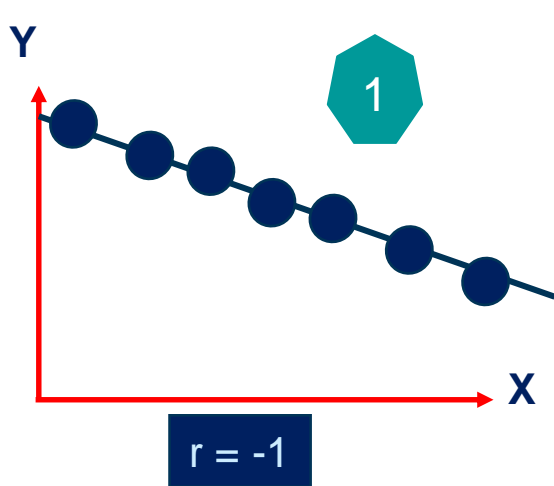
- 1. Theory of Field (e.g., Epidemiology)
- 2. Mathematical Theory
- 3. Previous Research
- 4. ‘Common Sense’



Thinking Challenge: Which is more logical?

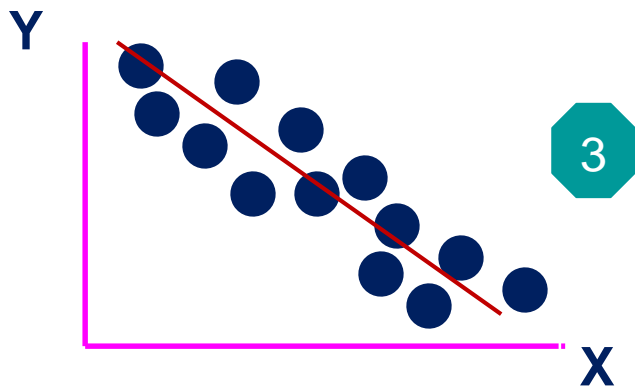
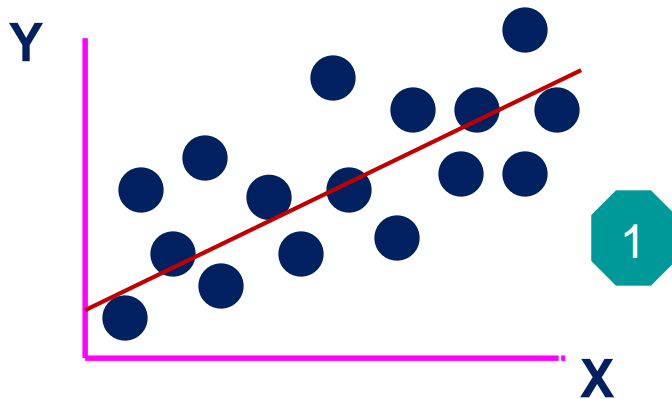


Scatter Plot of Data

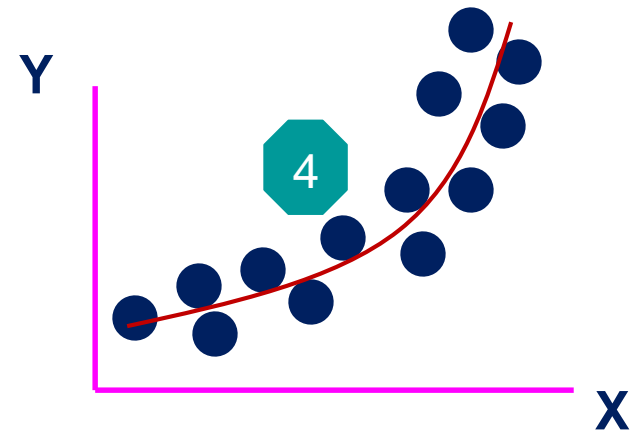
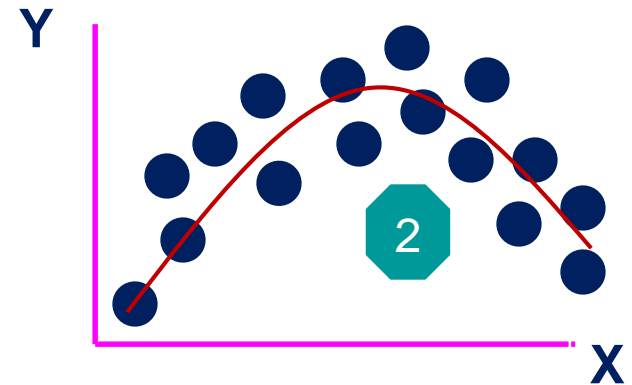


Types of Relationship

Linear relationships

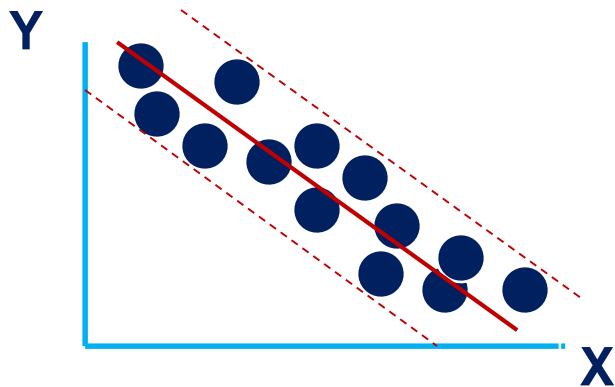
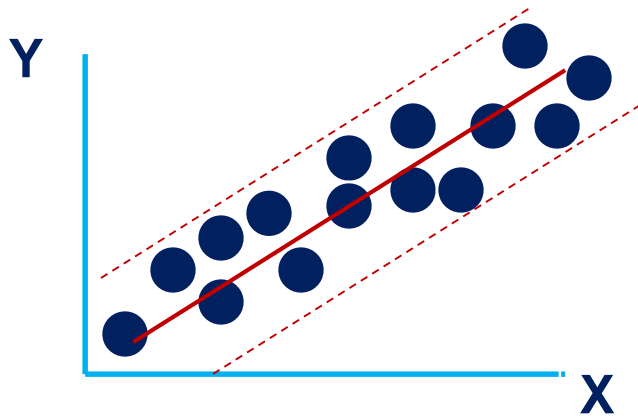


Curvilinear relationships

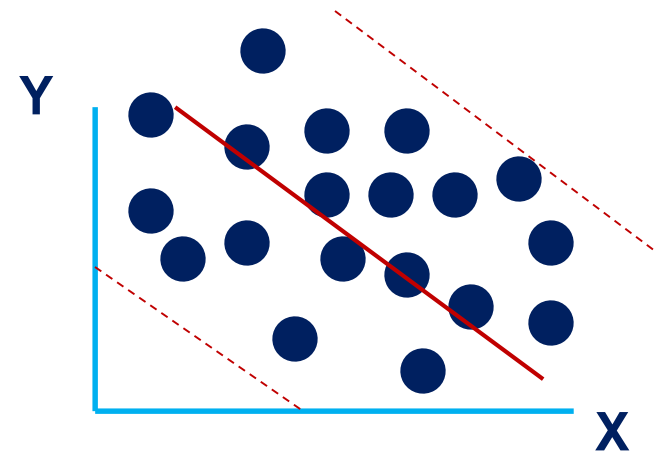
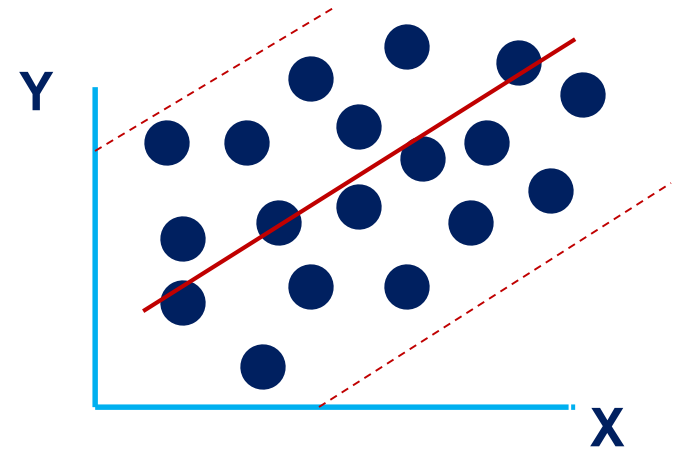


Types of Relationship...

Strong relationships

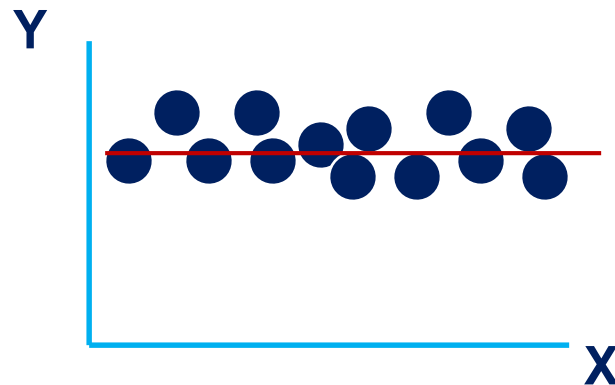
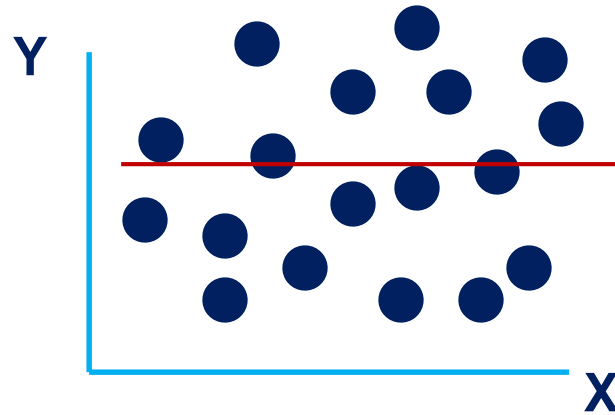


Weak relationships



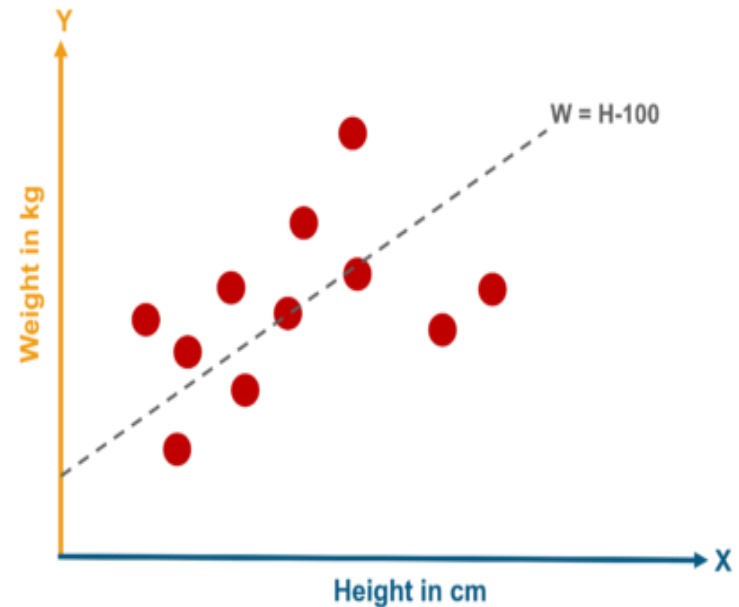
Types of Relationship...

No relationship



Linear Regression Models

- A linear regression is one of the easiest statistical models in machine learning.
- It is used to show the linear relationship between a dependent variable and one or more independent variables.
- Relationship between one dependent variable (y) and explanatory variable (s).
- Regression analysis is a form of predictive modelling technique which investigates the relationship between a dependent and independent variable



Types of Regression

➤ Types of Regression

➤ Linear Regression

➤ Logistic Regression

➤ Polynomial Regression

➤ Stepwise Regression

Basis	Linear Regression	Logistic Regression
Core Concept	The data is modelled using a straight line	The data is modelled using a sigmoid
Used with	Continuous Variable	Categorical Variable
Output/Prediction	Value of the variable	Probability of occurrence of an event
Accuracy and Goodness of Fit	Measured by loss, R squared, Adjusted R squared etc.	Measured by Accuracy, Precision, Recall, F1 score, ROC curve, Confusion Matrix, etc

Applications of LR

➤ Evaluating Trends and Sales Estimates

- A company sales analysis (monthly sales vs time)

➤ Analyzing the Impact of Price Changes

- If company changes the price of a product several times

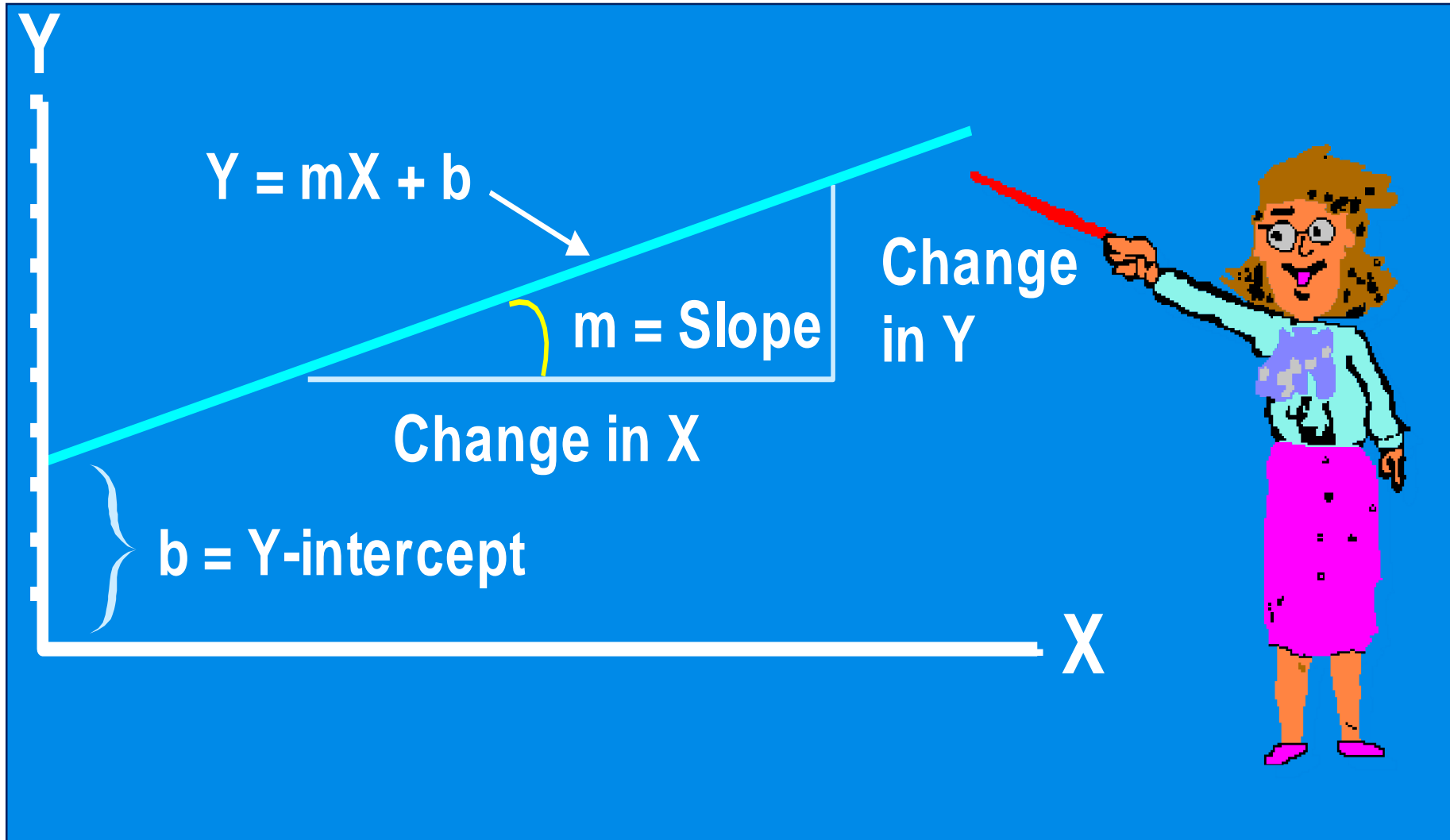
➤ Assessing Risk E.g. health care

- Number claims vs age

Benefits of LR Model

- *Ease*
- *Interpretability*
- *Scalability*
- *Deploys and Performs well on Online Settings*

Linear Equations



Linear Regression Model

- Relationship Between Variables is a Linear Function

The diagram illustrates the Linear Regression Model equation: $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$. Each term in the equation is labeled with a red text label and an orange arrow pointing to it. The labels are: **Population Y-Intercept** for β_0 , **Population Slope** for β_1 , **Random Error** for ε_i , **Dependent (Response Variable)** for Y_i , and **Independent (Explanatory) Variable** for X_i . Below the labels, examples are provided: **E.g. Grade** for the dependent variable and **E.g. Study Time** for the independent variable.

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

Population Y-Intercept (points to β_0)

Population Slope (points to β_1)

Random Error (points to ε_i)

Dependent (Response Variable) (points to Y_i)

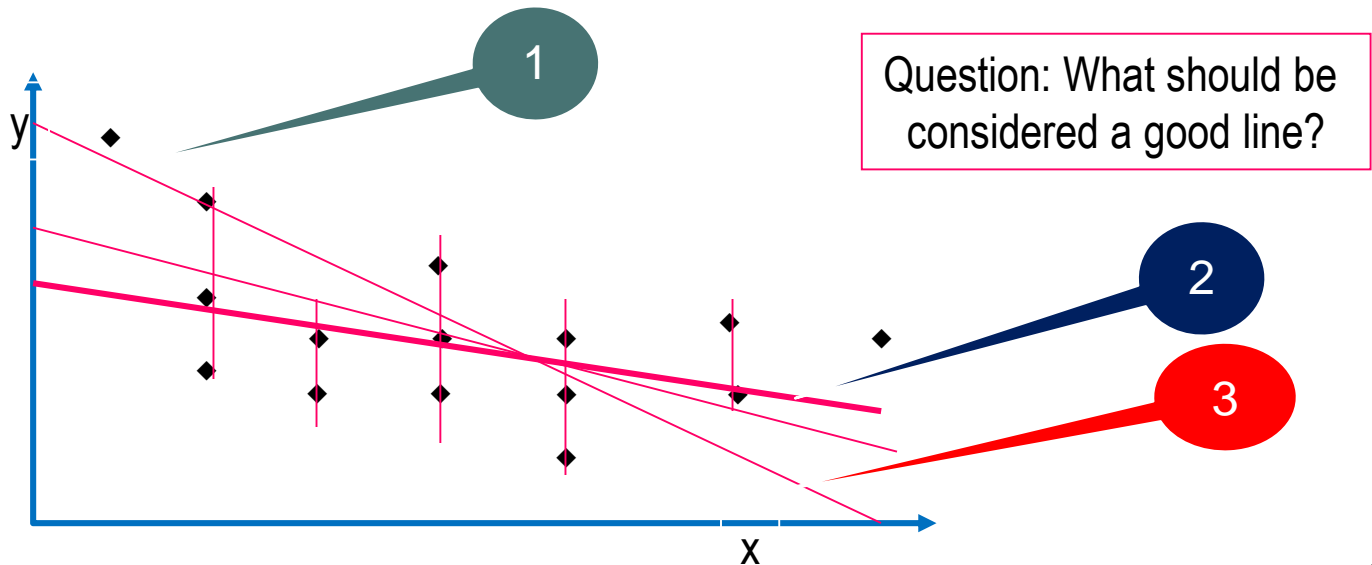
Independent (Explanatory) Variable (points to X_i)

E.g. Grade (under Dependent Variable)

E.g. Study Time (under Independent Variable)

Estimating the Coefficients

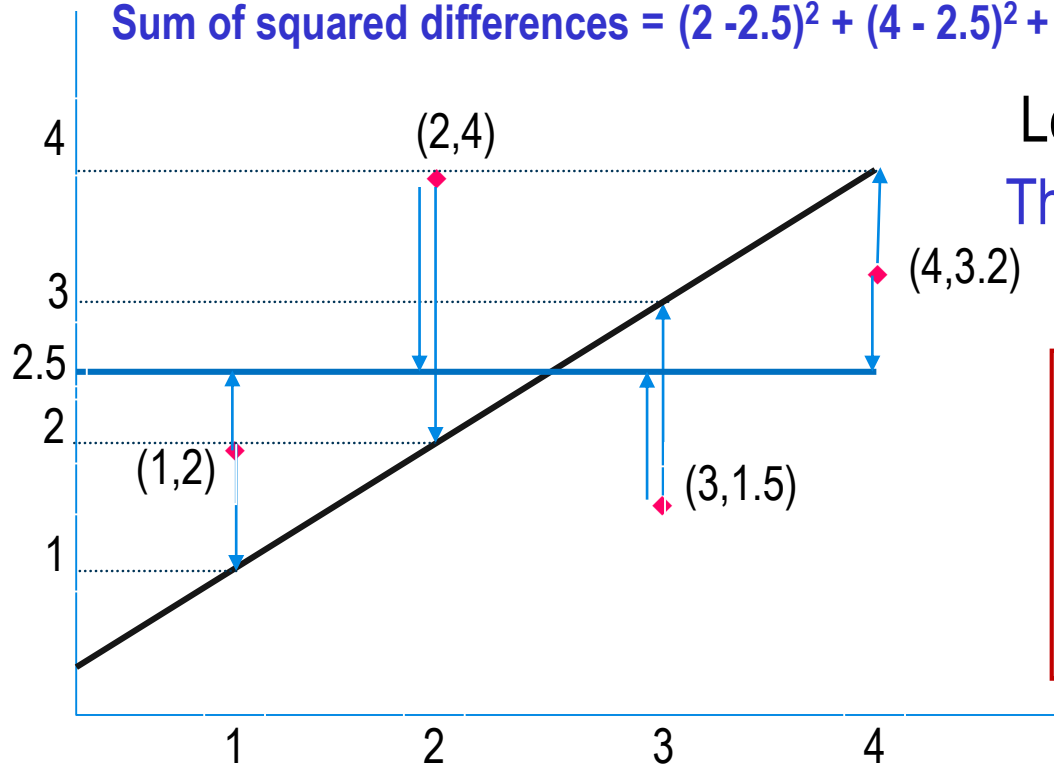
- The estimates are determined by
- drawing a sample from the population of interest,
 - calculating sample statistics.
 - producing a straight line that cuts into the data.



Sum Squared Difference

Sum of squared differences = $(2 - 1)^2 + (4 - 2)^2 + (1.5 - 3)^2 + (3.2 - 4)^2 = 6.89$

Sum of squared differences = $(2 - 2.5)^2 + (4 - 2.5)^2 + (1.5 - 2.5)^2 + (3.2 - 2.5)^2 = 3.99$

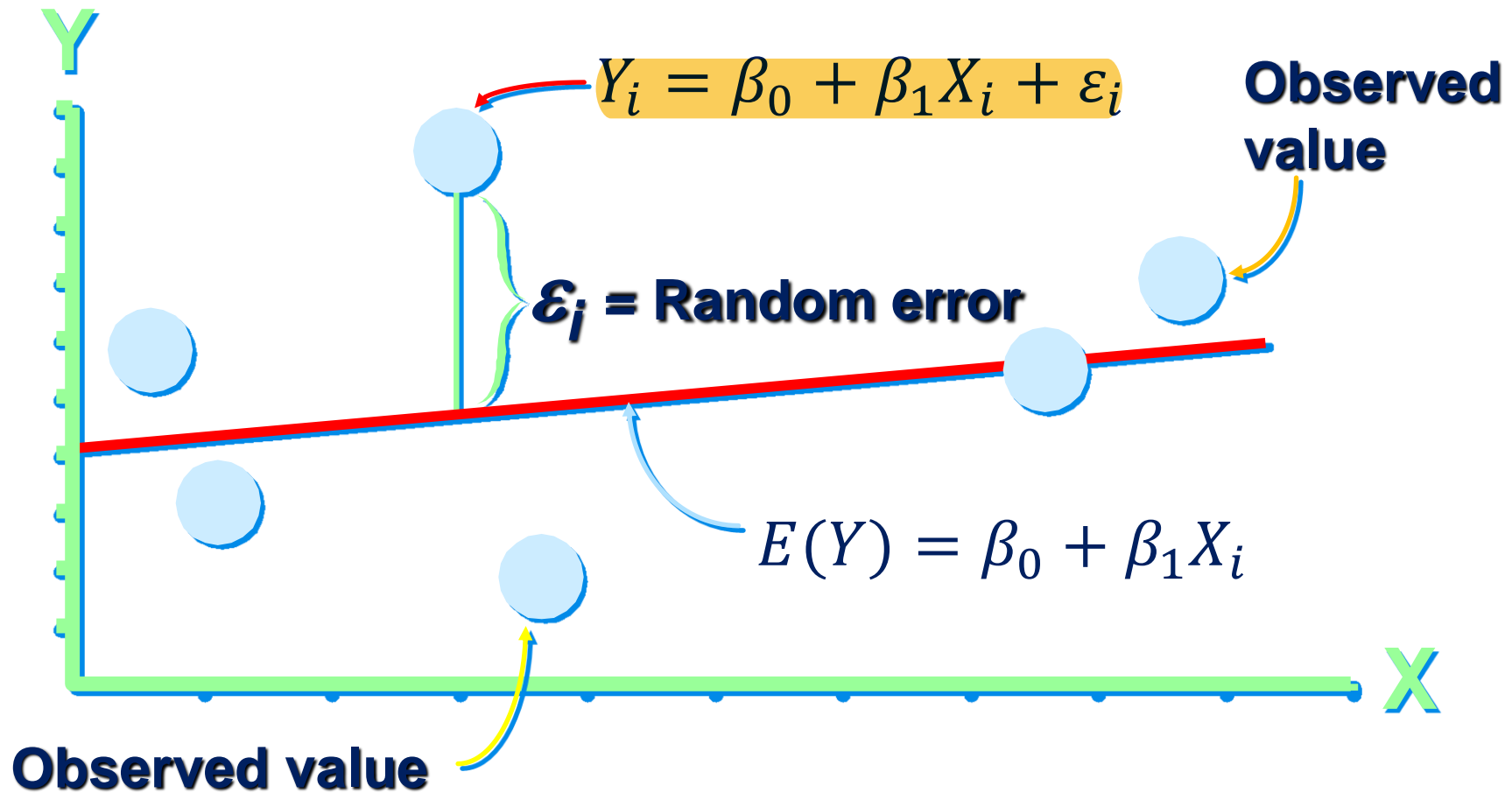


Let us compare two lines
The second line is horizontal

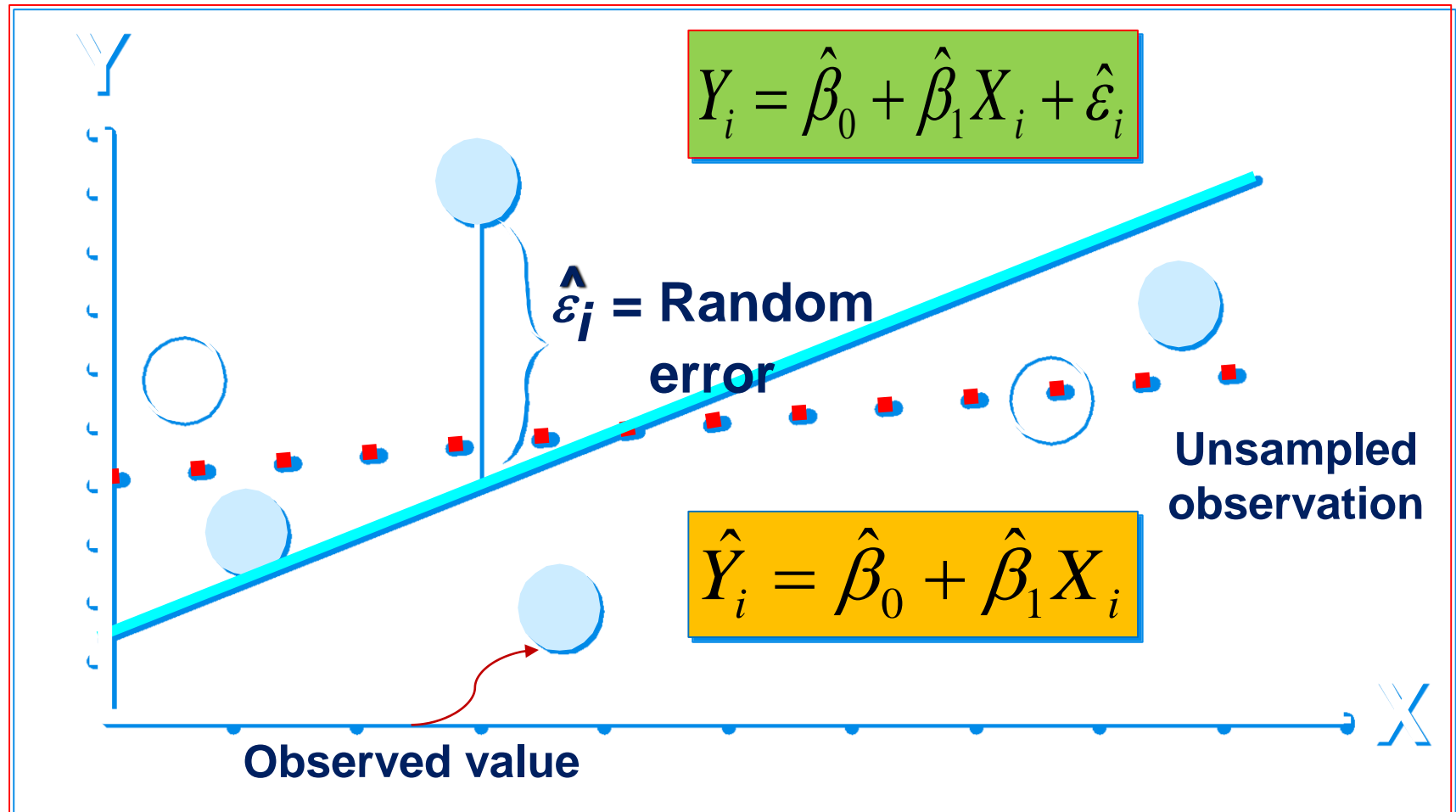
The smaller the sum of squared differences the better the fit of the line to the data.

A good line is one that minimizes the sum of squared differences between the points and the line.

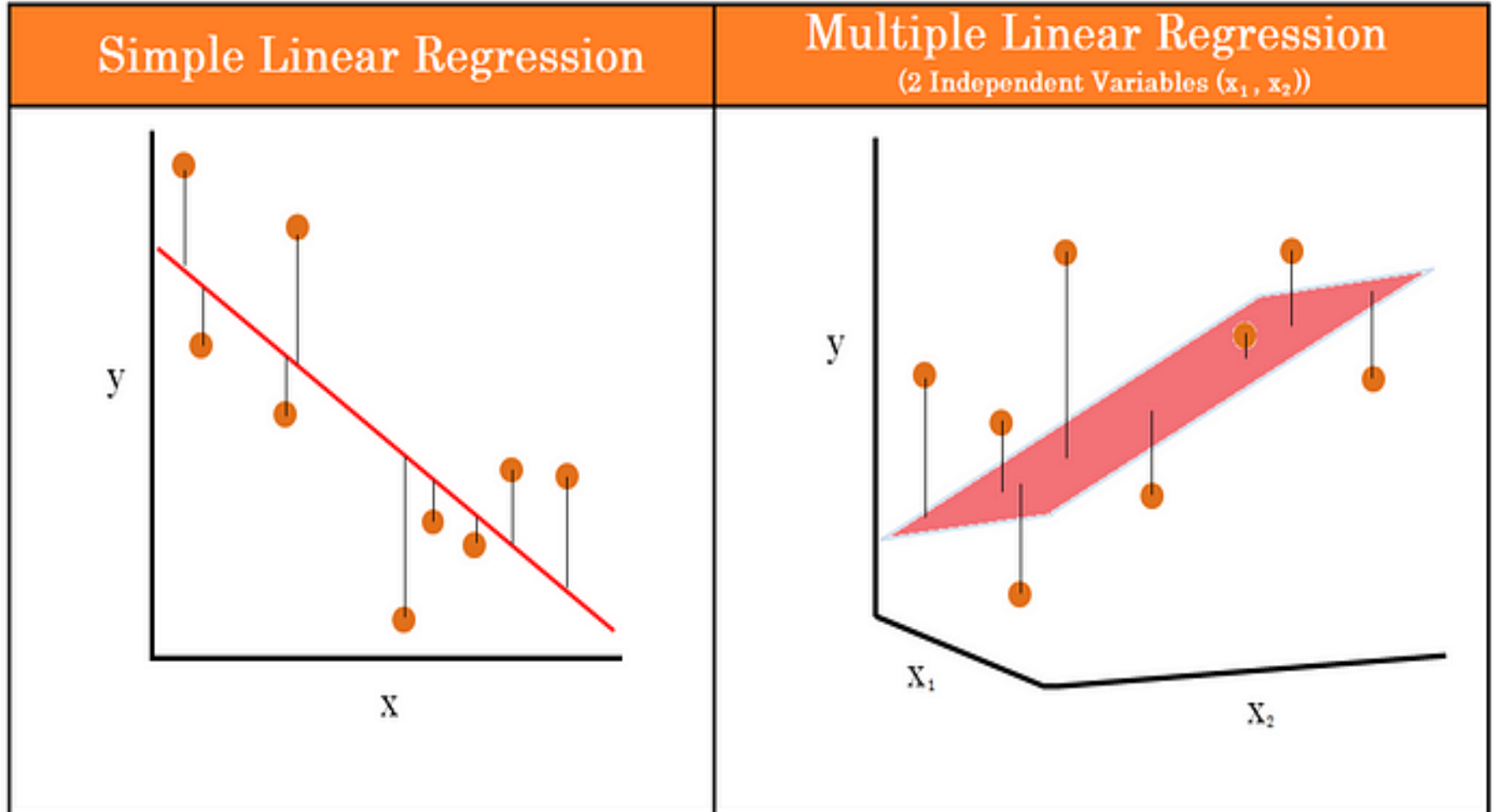
Population Linear Regression Model



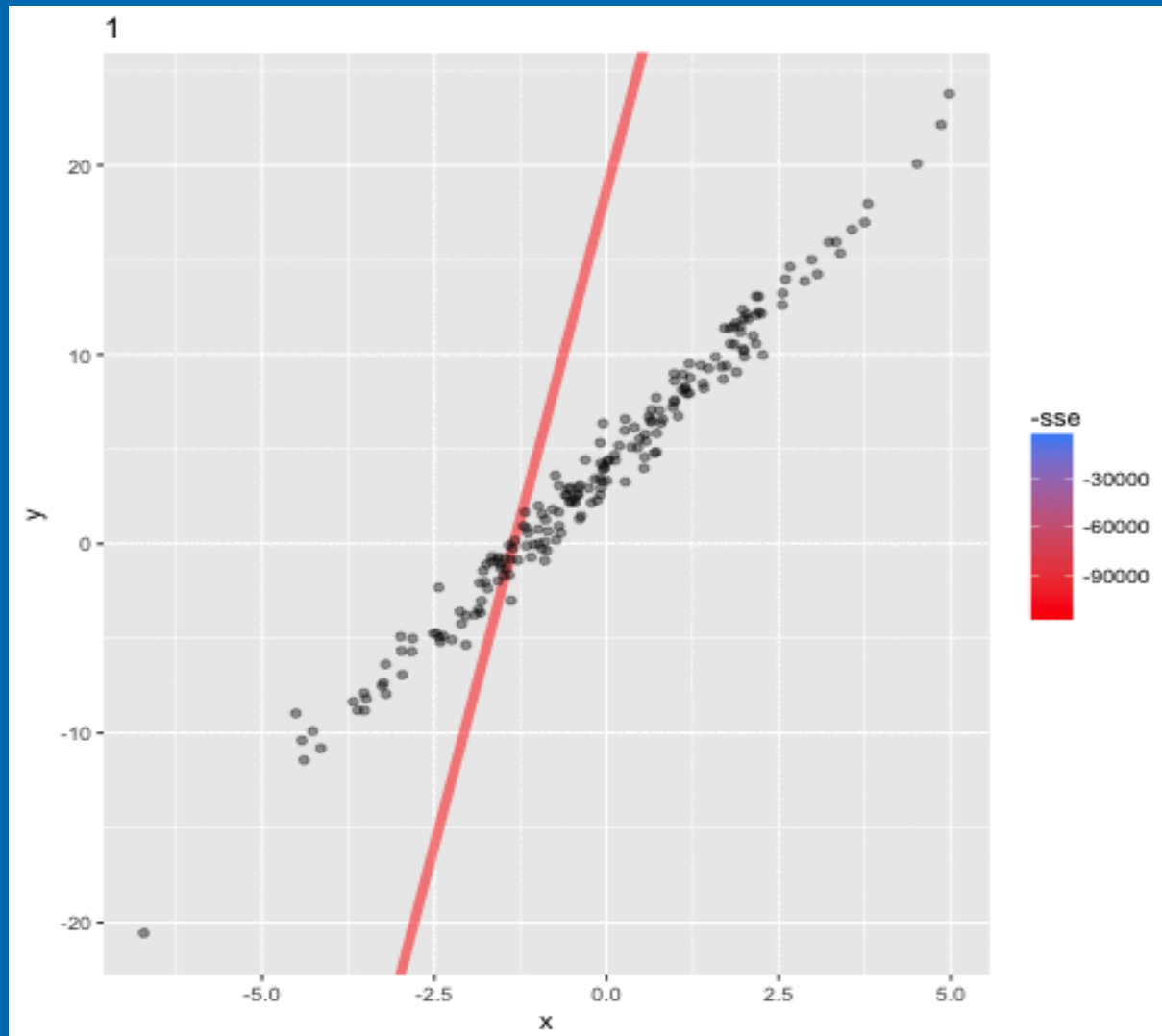
Simple Linear Regression Model



Simple LR Model vs Multiple LR

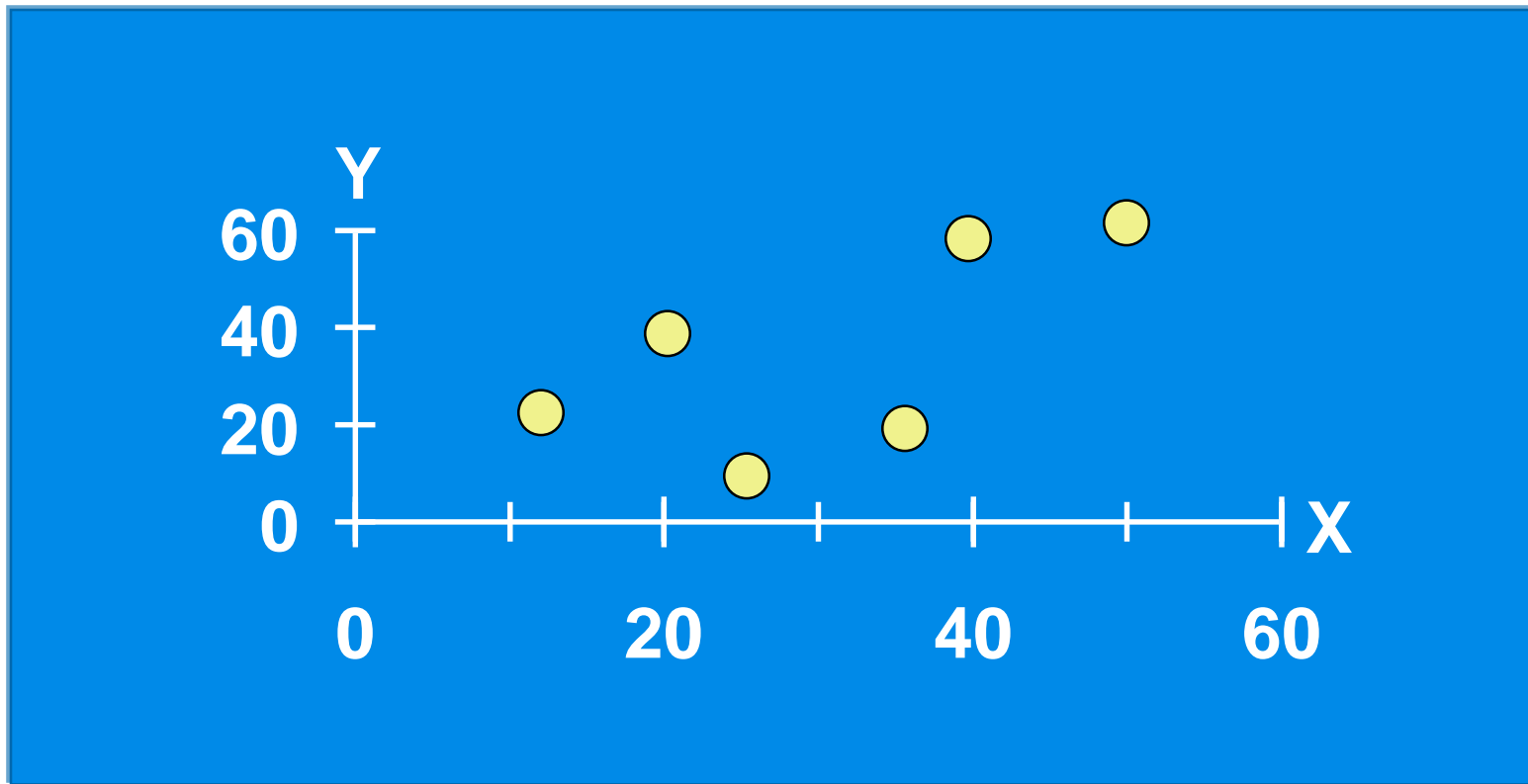


Estimating Parameters: Least Squares Method



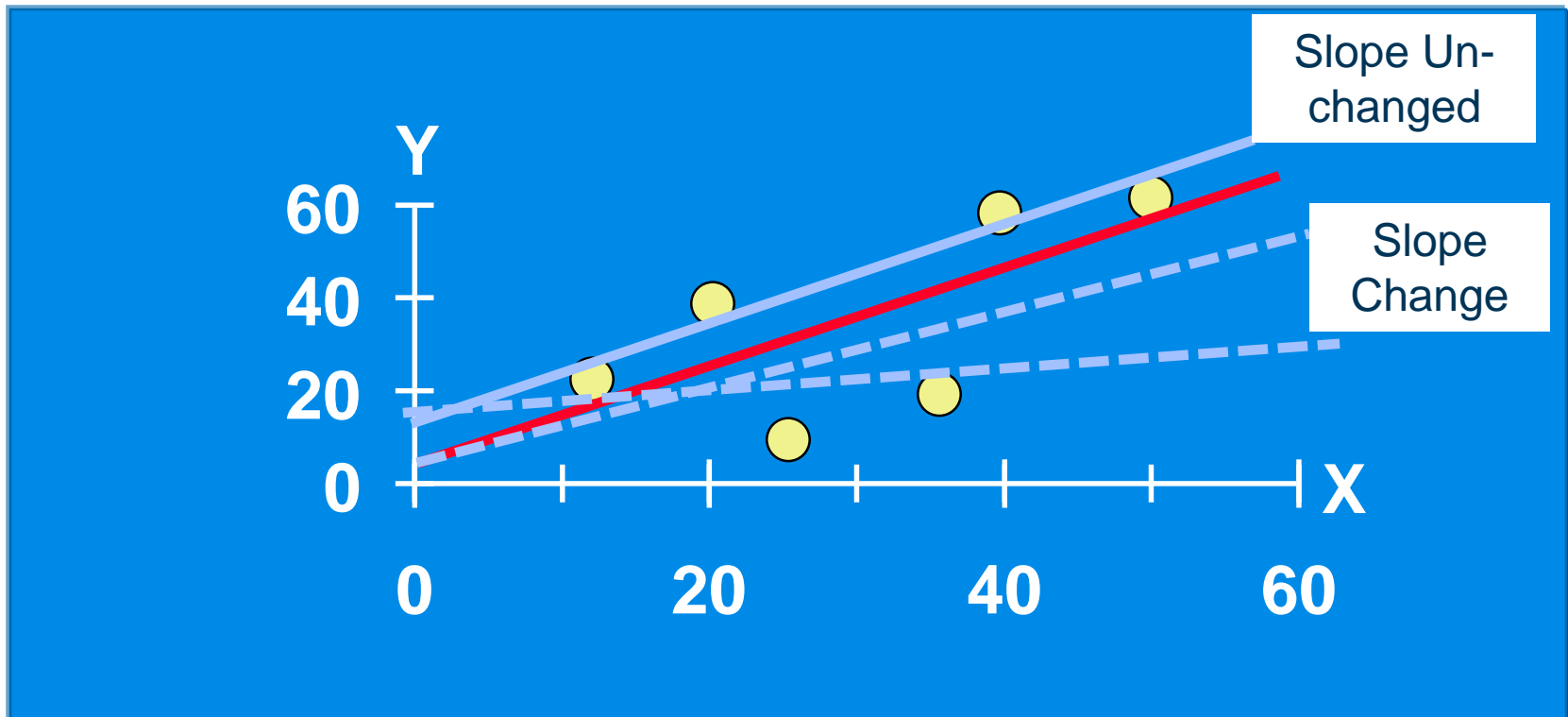
Scatter plot

- 1. Plot of All (X_i, Y_i) Pairs
- 2. Suggests How Well Model Will Fit



Thinking Challenge

How would you draw a line through the points?
How do you determine which line 'fits best'?

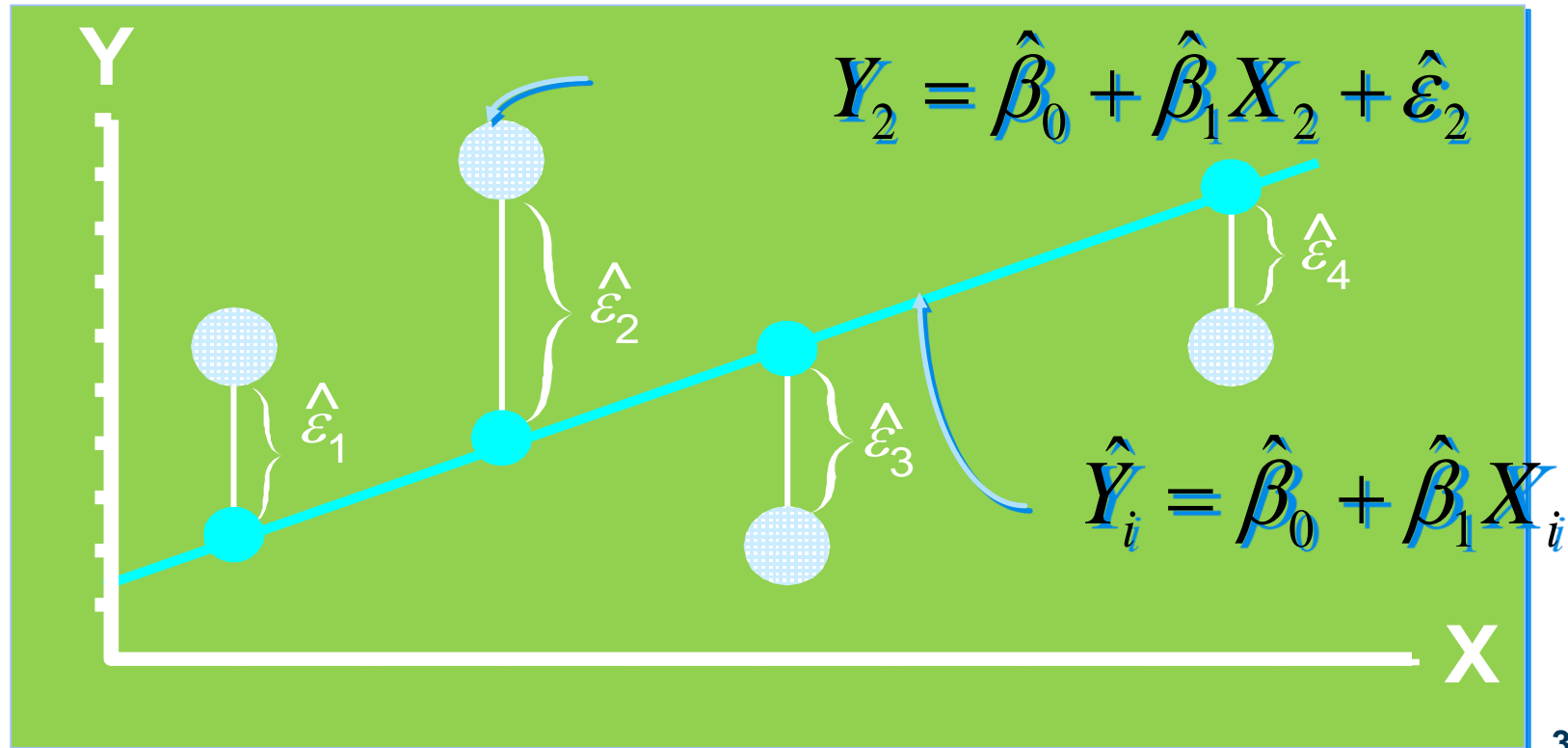


Least Squares Error

- ‘Best Fit’ Means Difference Between Actual Y Values & Predicted \hat{Y} Values are a Minimum. *But* Positive Differences Off-Set Negative ones
- So square errors!
- $\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^n \hat{\varepsilon}_i^2$
- LS Minimizes the Sum of the Squared Differences (errors) (SSE)

Least Squares Graphically

LS Minimizes $\sum_{i=1}^n \hat{\varepsilon}_i^2 = \hat{\varepsilon}_1^2 + \hat{\varepsilon}_2^2 + \hat{\varepsilon}_3^2 \dots \dots + \hat{\varepsilon}_n^2$



Coefficient Equations

Prediction equation

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

Sample slope

$$\hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

Sample Y - intercept

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

Finding β_0 MSE Method

- $\sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - \beta_0 + \beta_1 \cdot x_i)^2$
- $0 = \frac{\partial \sum_{i=1}^n (y_i - \beta_0 + \beta_1 \cdot x_i)^2}{\partial \beta_0}$
- $\frac{\partial \sum_{i=1}^n y_i^2 + \beta_0^2 + \beta_1^2 \cdot x_i^2 + 2\beta_0\beta_1 \cdot x_i - 2y_i\beta_0 - 2y_i\beta_1 \cdot x_i}{\partial \beta_0} = 0$
- $0 = 2 \sum_{i=1}^n \beta_0 + 2\beta_1 \sum_{i=1}^n x_i - 2 \sum_{i=1}^n y_i$
- $0 = 2(n\beta_0 + n\beta_1 \bar{x} - n\bar{y})$
- $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

Minimum Sum
Square Error
Method

Finding β_1 MSE Method

➤ $\sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = y_i^2 + \hat{y}_i^2 - 2y_i \cdot \hat{y}_i$

• Where $\hat{y}_i = \beta_0 + \beta_1 \cdot x_i$

➤ $\rightarrow \sum_{i=1}^n y_i^2 + (\beta_0 + \beta_1 \cdot x_i)^2 - 2y_i \cdot (\beta_0 + \beta_1 \cdot x_i)$

➤ \rightarrow

$$\sum_{i=1}^n y_i^2 + \beta_0^2 + \beta_1^2 \cdot x_i^2 + 2\beta_0\beta_1 \cdot x_i - 2y_i \cdot \beta_0 - 2y_i\beta_1 \cdot x_i$$

➤ Finding minimum error, the derivative must be zero:

Optimization

➤ $\frac{\partial(\sum_{i=1}^n \varepsilon_i^2)}{\partial \beta_1} = \frac{\partial \sum_{i=1}^n (y_i - (\beta_0 + \beta_1 \cdot x_i))^2}{\partial \beta_1} = 0$

➤ $0 = -2 \sum_{i=1}^n x_i (y_i - \beta_0 - \beta_1 x_i)$

➤ $0 = -2 \sum_{i=1}^n x_i (y_i - \bar{y} + \beta_1 \bar{x} - \beta_1 x_i)$ substituting the value of β_0

Minimum Square Error Method...

➤ $0 = -2 \sum_{i=1}^n x_i (y_i - \bar{y} + \beta_1 \bar{x} - \beta_1 x_i)$

- substituting the value of β_0

➤ $\beta_1 \sum_{i=1}^n x_i (x_i - \bar{x}) = \sum_{i=1}^n x_i (y_i - \bar{y})$

➤ $\beta_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})}$

➤ $\beta_1 = \frac{SS_{xy}}{SS_{xx}}$

Computation Table

X_i	Y_i	X_i^2	Y_i^2	$X_i Y_i$
X_1	Y_1	X_1^2	Y_1^2	$X_1 Y_1$
X_2	Y_2	X_2^2	Y_2^2	$X_2 Y_2$
\vdots	\vdots	\vdots	\vdots	\vdots
X_n	Y_n	X_n^2	Y_n^2	$X_n Y_n$
ΣX_i	ΣY_i	ΣX_i^2	ΣY_i^2	$\Sigma X_i Y_i$

Interpretation of Coefficients

- $\beta_1 > 0 \Rightarrow$ Positive Association
- $\beta_1 < 0 \Rightarrow$ Negative Association
- $\beta_1 = 0 \Rightarrow$ No Association

➤ Slope ($\widehat{\beta}_1$)

- Estimated Y Changes by $\widehat{\beta}_1$ for each 1 Unit Increase in X
 - If $\widehat{\beta}_1 = 2$, then Y Is Expected to Increase by 2 for each 1 Unit Increase in X

➤ Y-Intercept (β_0)

- Average Value of Y When $X = 0$
 - If $\beta_0 = 4$, then Average Y is expected to be 4 When X is 0

E.g. Parameter Estimation

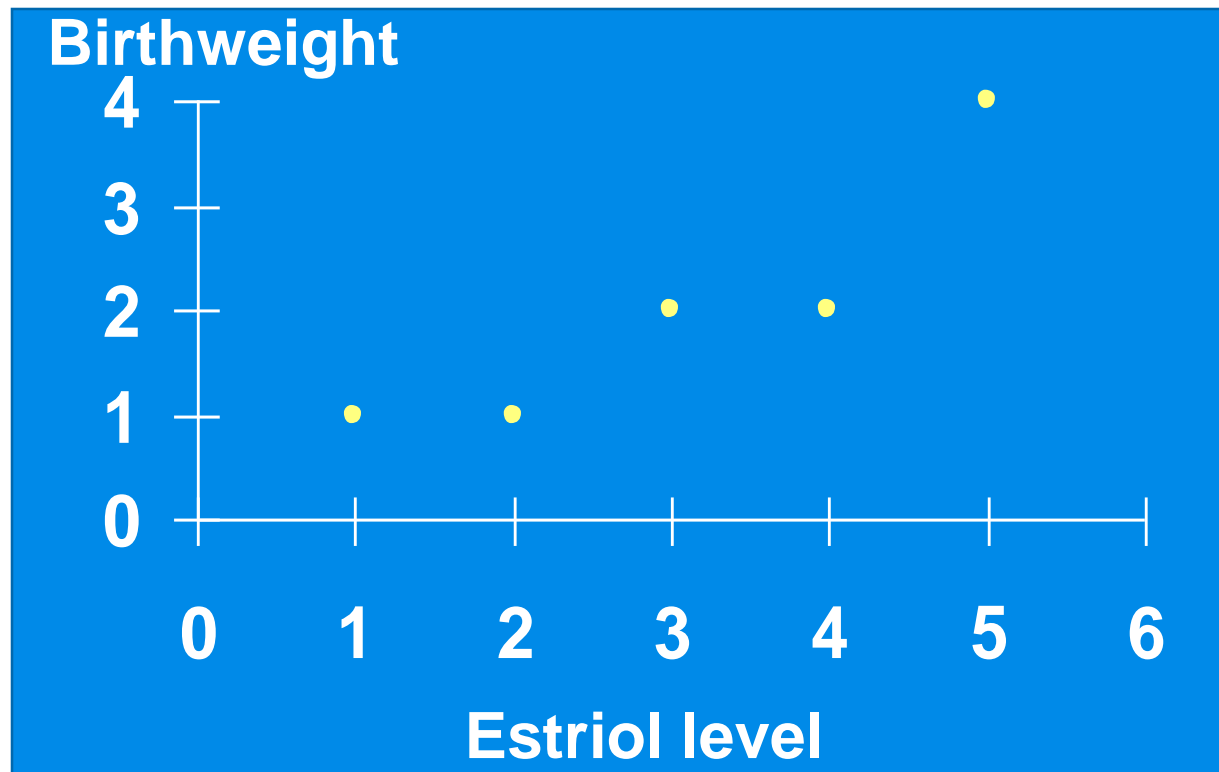
- What is the **relationship** between Mother's Estriol level & Birthweight using the following data?

<u>Estriol</u> (mg/24h)	<u>Birthweight</u> (g/1000)
1	1
2	1
3	2
4	2
5	4



Scatterplot

Birthweight vs. Estriol level



Parameter Estimation Solution Table

X_i	Y_i	X_i^2	Y_i^2	$X_i Y_i$
1	1	1	1	1
2	1	4	1	2
3	2	9	4	6
4	2	16	4	8
5	4	25	16	20
15	10	55	26	37

Parameter Estimation Solution

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n X_i Y_i - \frac{\left(\sum_{i=1}^n X_i\right)\left(\sum_{i=1}^n Y_i\right)}{n}}{\sum_{i=1}^n X_i^2 - \frac{\left(\sum_{i=1}^n X_i\right)^2}{n}} = \frac{37 - \frac{(15)(10)}{5}}{55 - \frac{(15)^2}{5}} = 0.70$$

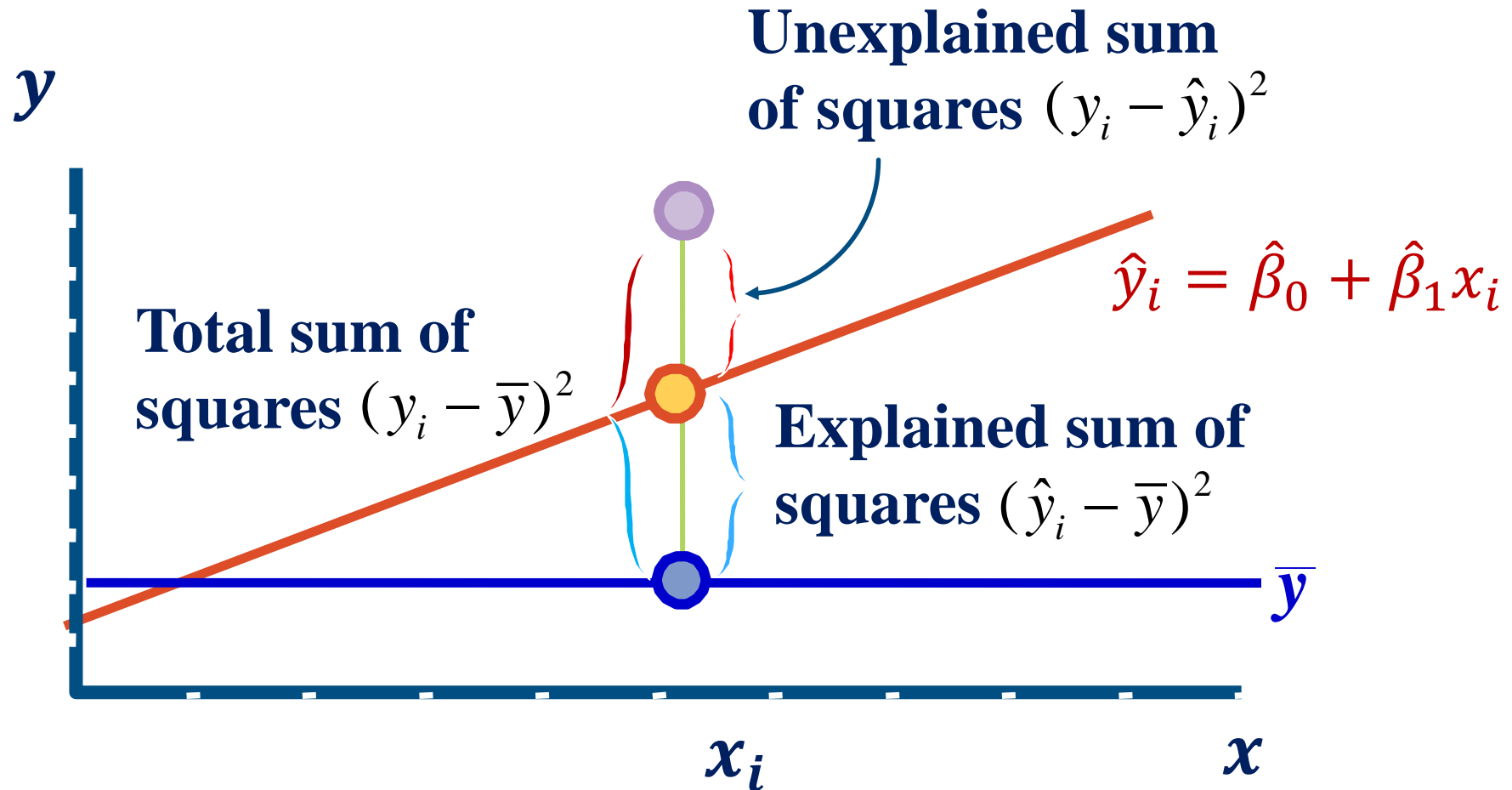
$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = 2 - (0.7)(3) = -0.1$$

$$\hat{y} = -.1 + .7x$$

Coefficient Interpretation Solution

- 1. Slope (β_1)
 - Birthweight (Y) is Expected to Increase by .7 Units for Each 1 unit Increase in Estriol (X).
- 2. Intercept (β_0)
 - Average Birthweight (Y) is -.10 Units When Estriol level (X) Is 0
 - Difficult to explain
 - The birthweight should always be positive

Goodness: Variation Measures



Estimation of σ^2

$$s^2 = \frac{SSE}{n - 2}$$

Where $SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$

$$s = \sqrt{s^2} = \sqrt{\frac{SSE}{n - 2}}$$

The subtraction of 2 can be thought of as the fact that we have estimated two parameters: β_0 and β_1

E.g. Compute SSE, s^2 , s

You're a marketing analyst for any Toys. You gather the following data:

<u>Ad (₹)</u>	<u>Sales (Qty)</u>
1	1
2	1
3	2
4	2
5	4

Find **SSE**, s^2 , and s .



E.g. Solution: SSE, s^2 , s

x_i	y_i	$\hat{y} = -.1 + .7x$	$y - \hat{y}$	$(y - \hat{y})^2$
1	1	.6	.4	.16
2	1	1.3	-.3	.09
3	2	2	0	0
4	2	2.7	-.7	.49
5	4	3.4	.6	.36
				SSE=1.1

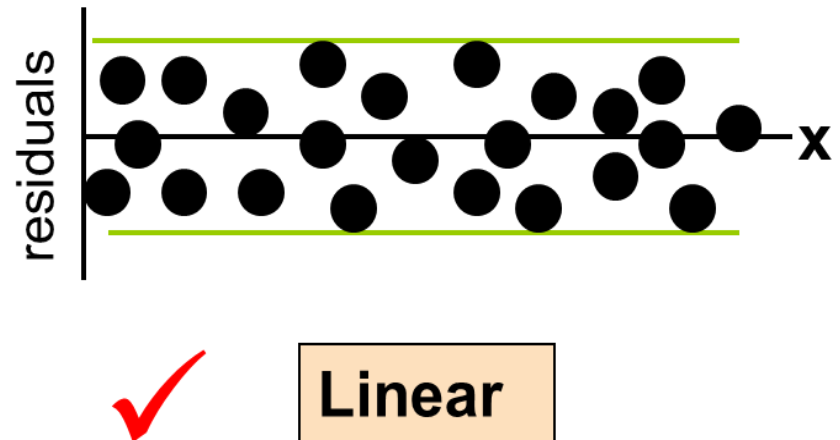
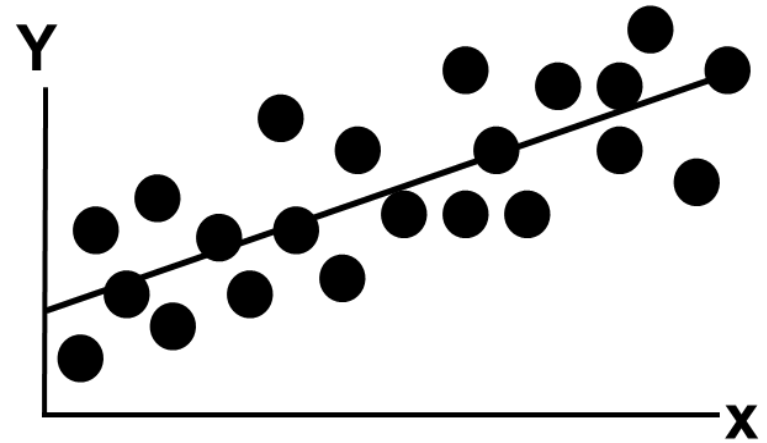
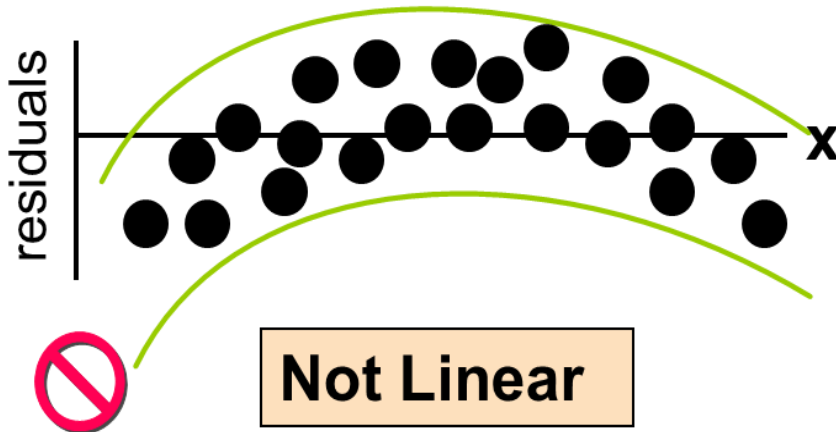
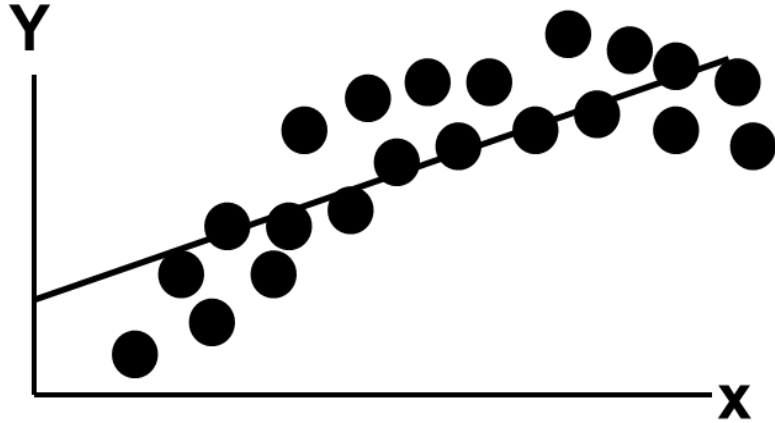
$$s^2 = \frac{SSE}{n - 2} = \frac{1.1}{5 - 2} = .366667 \quad s = \sqrt{.366667} = .6055$$

Residual Analysis

$$e_i = Y_i - \hat{Y}_i$$

- The residual for observation i , e_i , is the difference between its observed and predicted value
- Check the assumptions of regression by examining the residuals
 - Examine for linearity assumption
 - Evaluate independence assumption
 - Evaluate normal distribution assumption
 - Examine for constant variance for all levels of X (homoscedasticity)

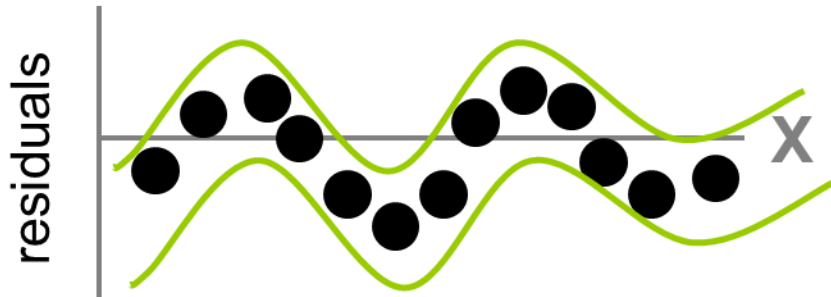
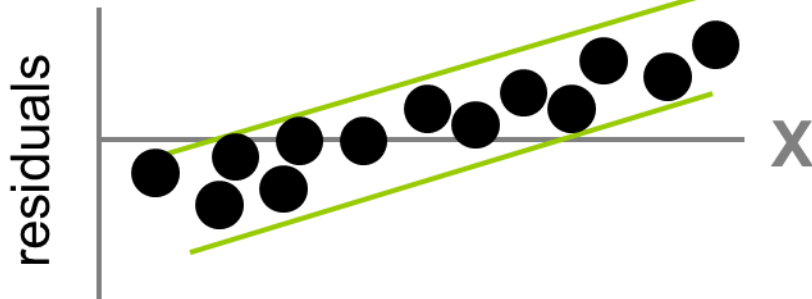
Residual Analysis for Linearity



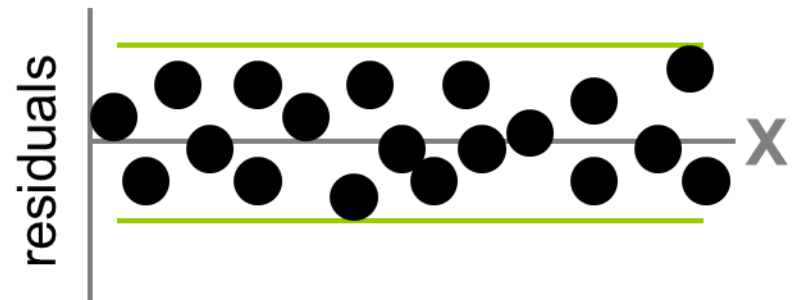
Residual Analysis for Independence



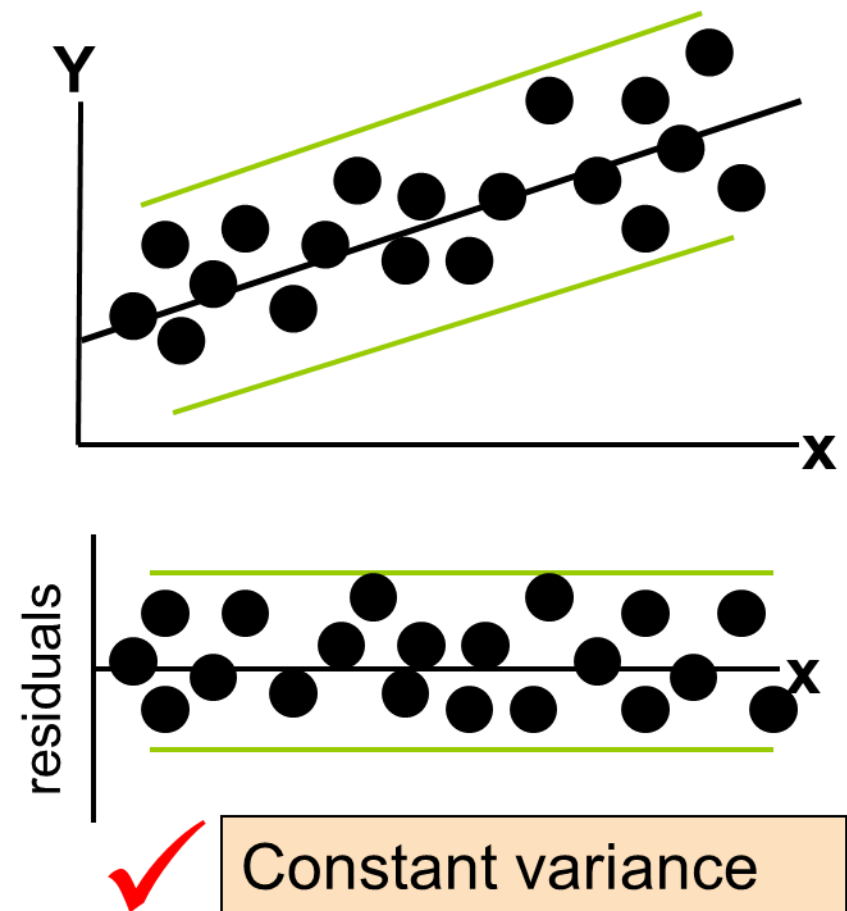
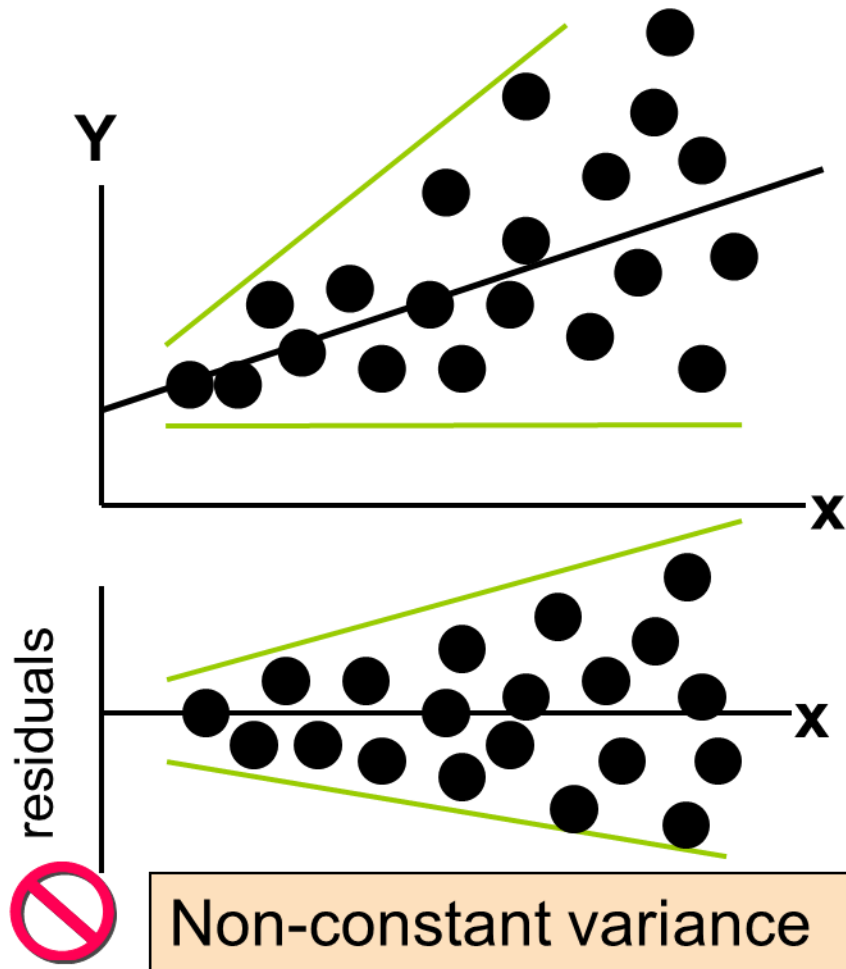
Not Independent



Independent



RA for Equal Variance



Evaluating the Model

Testing for Significance

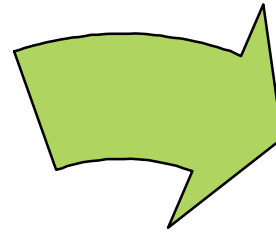
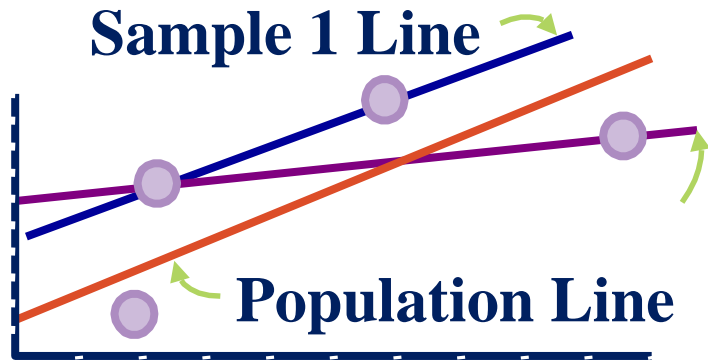
Regression Modeling Steps

1. Hypothesize deterministic component
2. Estimate unknown model parameters
3. Specify probability distribution of random error term
 - Estimate standard deviation of error
4. **Evaluate model**
5. Use model for prediction and estimation

Test of Slope Coefficient

- Shows if there is a linear relationship between x and y
- Involves population slope β_1
- Hypotheses
 - $H_0: \beta_1 = 0$ (No Linear Relationship)
 - $H_a: \beta_1 \neq 0$ (Linear Relationship)
- Theoretical basis is sampling distribution of slope

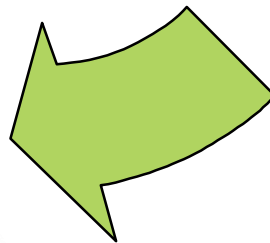
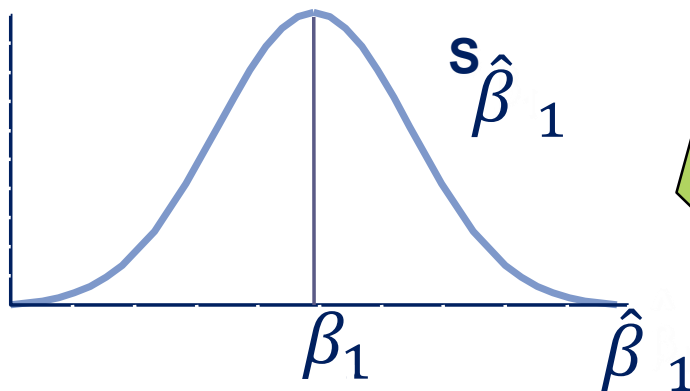
Distribution of Sample Slopes



All Possible Sample Slopes	
Sample 1:	2.5
Sample 2:	1.6
Sample 3:	1.8
Sample 4:	2.1
⋮	⋮

Very large number of
sample slopes

Sampling Distribution



Slope Coefficient Test Statistic

$$t = \frac{\widehat{\beta}_1}{S_{\widehat{\beta}_1}} = \frac{\widehat{\beta}_1}{\frac{S}{\sqrt{SS_{xx}}}}, \quad df = 2$$

$$SS_{xx} = \sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n}$$

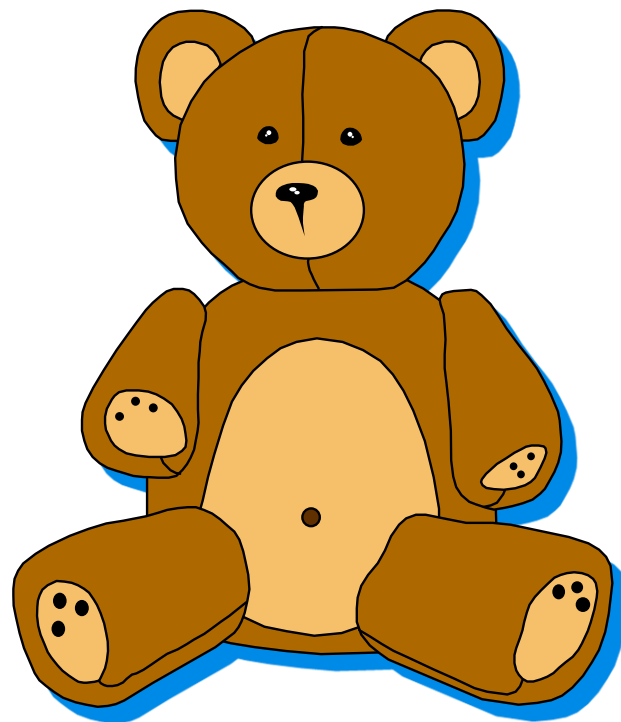
E.g. Test of Slope Coefficient

You're a marketing analyst for any Toys.

You find $\hat{\beta}_0 = -.1$, $\hat{\beta}_1 = .7$ and $s = .6055$.

<u>Ad (₹)</u>	<u>Sales (Qty)</u>
1	1
2	1
3	2
4	2
5	4

Is the relationship **significant**
at the **.05** level of significance?



Solution Table

x_i	y_i	x_i^2	y_i^2	$x_i y_i$
1	1	1	1	1
2	1	4	1	2
3	2	9	4	6
4	2	16	4	8
5	4	25	16	20
15	10	55	26	37

Slope Coefficient Test Statistic

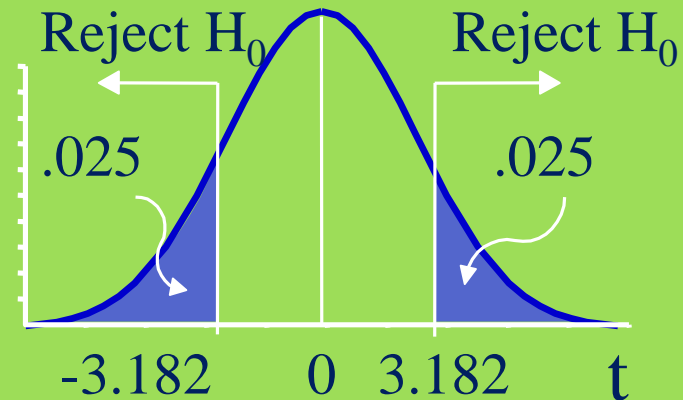
$$t = \frac{\widehat{\beta}_1}{S_{\widehat{\beta}_1}} = \frac{\widehat{\beta}_1}{\frac{s}{\sqrt{SS_{xx}}}}, \quad df = 2 \quad \Rightarrow \quad \text{Where } SS_{xx} = \sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n}$$

$$S_{\widehat{\beta}_1} = \frac{s}{\sqrt{SS_{xx}}} = \frac{0.6055}{\sqrt{55 - \frac{15^2}{5}}} = .1914$$

$$t = \frac{0.70}{.1914} = 3.657$$

Test of Slope Coefficient Solution

- $H_0: \beta_1 = 0$
- $H_a: \beta_1 \neq 0$
- $\alpha = .05$
- $df = 5 - 2 = 3$
- **Critical Value(s):**



Test Statistic: $t = \frac{\widehat{\beta}_1}{s_{\widehat{\beta}_1}} = \frac{0.70}{.1914} = 3.657$

Decision: Reject (H_0) at $\alpha = .05$,

Conclusion: There is evidence of a relationship

Correlation Coefficient

Correlation Models

- Answers ‘How strong is the **linear** relationship between two variables?’
- Coefficient of correlation
 - Sample correlation coefficient denoted r
 - Values range from -1 to $+1$
 - Measures degree of association
 - Does not indicate cause–effect relationship

Coefficient of Correlation

$$r = \frac{SS_{xy}}{\sqrt{SS_{xx} SS_{yy}}}$$

$$SS_{xx} = \sum x^2 - \frac{(\sum x)^2}{n}$$

$$SS_{yy} = \sum y^2 - \frac{(\sum y)^2}{n}$$

$$SS_{xy} = \sum xy - \frac{(\sum x)(\sum y)}{n}$$

Correlation Coefficient Values

**Perfect Negative
Correlation**

**No Linear
Correlation**

**Perfect Positive
Correlation**

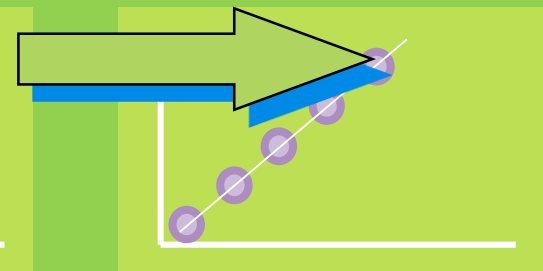
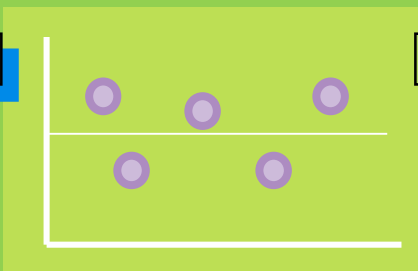
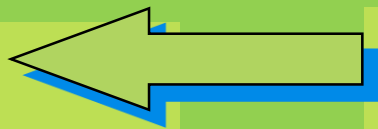
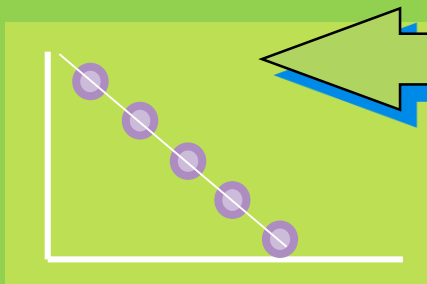
-1.0

-0.5

0

+0.5

+1.0



**Increasing degree of negative
correlation**

**Increasing degree of positive
correlation**

E.g. Coefficient of Correlation

You're a marketing analyst for any Toys.

<u>Ad (₹)</u>	<u>Sales (Qty)</u>
1	1
2	1
3	2
4	2
5	4

Calculate the **coefficient of correlation**.



Solution Table

x_i	y_i	x_i^2	y_i^2	$x_i y_i$
1	1	1	1	1
2	1	4	1	2
3	2	9	4	6
4	2	16	4	8
5	4	25	16	20
15	10	55	26	37

Coefficient of Correlation Solution

$$SS_{xx} = \sum x^2 - \frac{(\sum x)^2}{n} = 55 - \frac{(15)^2}{5} = 10$$

$$SS_{yy} = \sum y^2 - \frac{(\sum y)^2}{n} = 26 - \frac{(10)^2}{5} = 6$$

$$SS_{xy} = \sum xy - \frac{(\sum x)(\sum y)}{n} = 37 - \frac{(15)(10)}{5} = 7$$

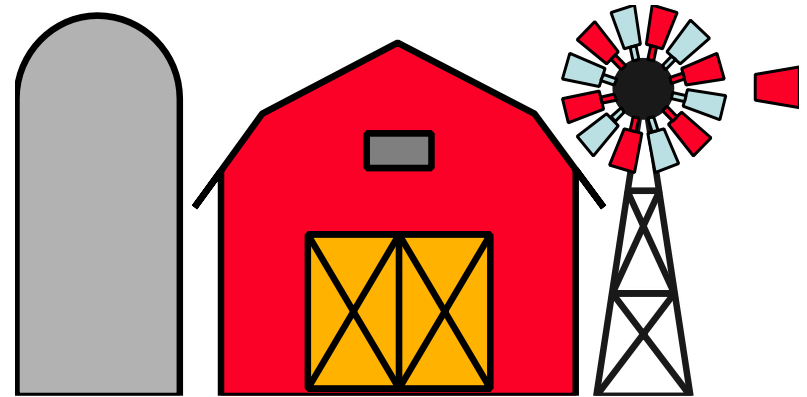
$$r = \frac{SS_{xy}}{\sqrt{SS_{xx}SS_{yy}}} = \frac{7}{\sqrt{10 \cdot 6}} = .904$$

It can be predicted using LR due
High value of Correlation Coefficient

Coefficient of Correlation Challenge

You're an economist for the county cooperative.
You gather the following data:

<u>Fertilizer (lb.)</u>	<u>Yield (lb.)</u>
4	3.0
6	5.5
10	6.5
12	9.0



Find the **coefficient of correlation**.

Solution Table*

x_i	y_i	x_i^2	y_i^2	$x_i y_i$
4	3.0	16	9.00	12
6	5.5	36	30.25	33
10	6.5	100	42.25	65
12	9.0	144	81.00	108
32	24.0	296	162.50	218

Coefficient of Correlation Solution*

$$SS_{xx} = \sum x^2 - \frac{(\sum x)^2}{n} = 296 - \frac{(32)^2}{4} = 40$$

$$SS_{yy} = \sum y^2 - \frac{(\sum y)^2}{n} = 162.5 - \frac{(24)^2}{4} = 18.5$$

$$SS_{xy} = \sum xy - \frac{(\sum x)(\sum y)}{n} = 218 - \frac{(32)(24)}{4} = 26$$

$$r = \frac{SS_{xy}}{\sqrt{SS_{xx}SS_{yy}}} = \frac{26}{\sqrt{40 \cdot 18.5}} = .956$$

Coefficient of Determination

Proportion of variation 'explained' by relationship between x and y

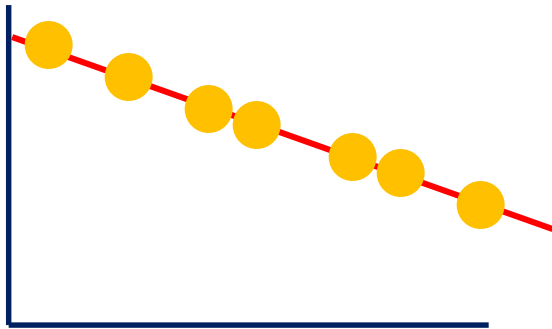
$$r^2 = \frac{\text{Explained Variation}}{\text{Total Variation}} = \frac{SS_{yy} - SSE}{SS_{yy}}$$

$$0 \leq r^2 \leq 1$$



$$r^2 = (\text{coefficient of correlation})^2$$

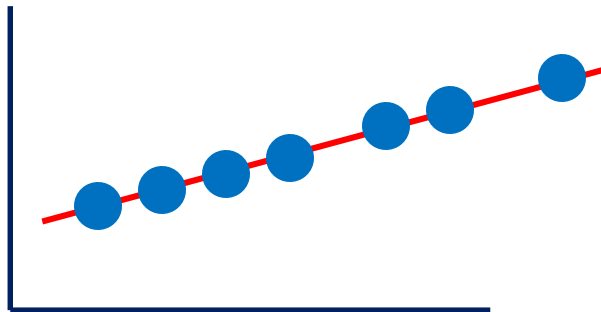
E.g. Approximate r^2 Values



$$r^2 = 1$$

$$r^2 = 1$$

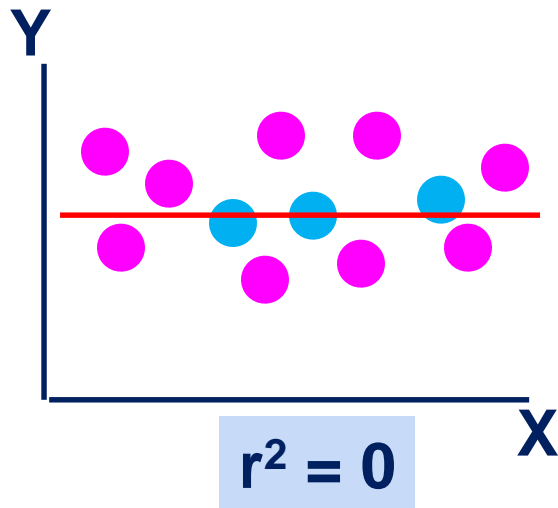
**Perfect linear relationship
between X and Y:**



$$r^2 = 1$$

**100% of the variation in Y is
explained by variation in X**

E.g. Approximate r^2 Values...



$$r^2 = 0$$

- No linear relationship between X and Y:
- The value of Y does not depend on X. (None of the variation in Y is explained by variation in X)

E.g. Determination Coefficient

You're a marketing analyst for any Toys. You know $r = .904$.

<u>Ad (₹)</u>	<u>Sales (Qty)</u>
1	1
2	1
3	2
4	2
5	4

Calculate and interpret the **coefficient of determination**.



E.g. Determination Coefficient

$$r^2 = (\text{coefficient of correlation})^2$$

$$r^2 = (.904)^2$$

$$r^2 = .817$$

Interpretation: About 81.7% of the sample variation in Sales (y) can be explained by using Ad ₹ (x) to predict Sales (y) in the linear model.

Other Evaluation Metrics

➤ Mean Squared Error (MSE)

- Most commonly used Metric
- Differentiable due to convex shape
- Easier to optimize.
- Penalizes large errors

$$MSE = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2$$

➤ Mean Absolute Error (MAE)

- Not preferred in cases where outliers are prominent
- MAE does not penalize large errors.
- **Small MAE** suggests the model is great at prediction, while a **large MAE** suggests that model may have trouble in certain areas.

$$MAE = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i|$$

Other Evaluation Metrics...

➤ Root Mean Squared Error (RMSE)

- RMSE measures the scatter of these residuals
- RMSE penalizes large errors.
- Lower RMSE indicates model is better for predictions.
- Higher RMSE indicates, large deviations between the predicted and actual value.

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2}$$

Conclusion

1. Described the Linear Regression Model
2. Stated the Regression Modeling Steps
3. Explained Least Squares
4. Computed Regression Coefficients
5. Explained Correlation
6. Predicted Response Variable