ISE 315: Engineering Statistics

Lecture 15: Simple Linear Regression (Part 1)

Instructor: Mansur M. Arief, PhD
Industrial and Systems Engineering, KFUPM

Office: 22-219 — Email: mansur.arief@kfupm.edu.sa

Based on Montgomery & Runger, Applied Statistics and Probability for Engineers, 6th Ed.

Lecture 15

Simple Linear Regression and Correlation (Chapter 6)

Congratulations!

Major Exam 1 is Done

Alhamdulillah — Major Exam 1 is behind you!

Congratulations!

Major Exam 1 is Done

Alhamdulillah — Major Exam 1 is behind you!

You have worked through four chapters of statistical inference:

- Chapter 2: Point estimation & sampling distributions
- Chapter 3: Confidence intervals for μ, σ^2, and p
- Chapter 4: Hypothesis testing (one sample)
- Chapter 5: Two-sample inference

Congratulations!

Major Exam 1 is Done

Alhamdulillah — Major Exam 1 is behind you!

You have worked through four chapters of statistical inference:

- Chapter 2: Point estimation & sampling distributions
- Chapter 3: Confidence intervals for μ, σ^2, and p
- Chapter 4: Hypothesis testing (one sample)
- Chapter 5: Two-sample inference

It's a **major accomplishment**. Take a moment to appreciate how far you've come.

Grade Status Update

Exam grades are still being calculated and finalized.

Grade Status Update

Exam grades are still being calculated and finalized.

We are also updating **all components** accumulated so far:

- Quizzes 1–4
- Homeworks 1–5
- Attendance (Lectures 1–14)
- Major Exam 1

Grade Status Update

Exam grades are still being calculated and finalized.

We are also updating **all components** accumulated so far:

- Quizzes 1–4
- Homeworks 1–5
- Attendance (Lectures 1–14)
- Major Exam 1

Target: Your grades will be updated on Blackboard by **Tuesday**, in shaa Allah.

Grade Status Update

Exam grades are still being calculated and finalized.

We are also updating **all components** accumulated so far:

- Quizzes 1–4
- Homeworks 1–5
- Attendance (Lectures 1–14)
- Major Exam 1

Target: Your grades will be updated on Blackboard by **Tuesday**, in shaa Allah.

If you notice any discrepancies, please email me *after* the grades are posted.

Where Are We in the Course?

Part	Topic	Status
Part 1	Estimation & sampling distributions (Ch. 2)	✓ Done
Part 1	Confidence intervals (Ch. 3)	✓ Done
Part 2	Hypothesis testing — one sample (Ch. 4)	✓ Done
Part 2	Hypothesis testing — two samples (Ch. 5)	✓ Done
Part 3	**Simple linear regression (Ch. 6)**	**Starting today**
Part 3	Software tools (Ch. 10)	Next lecture
Part 3	Multiple linear regression (Ch. 7)	Upcoming
Part 4	ANOVA & designed experiments (Ch. 8–9)	Upcoming

Where Are We in the Course?

Part	Topic	Status
Part 1	Estimation & sampling distributions (Ch. 2)	✓ Done
Part 1	Confidence intervals (Ch. 3)	✓ Done
Part 2	Hypothesis testing — one sample (Ch. 4)	✓ Done
Part 2	Hypothesis testing — two samples (Ch. 5)	✓ Done
Part 3	**Simple linear regression (Ch. 6)**	**Starting today**
Part 3	Software tools (Ch. 10)	Next lecture
Part 3	Multiple linear regression (Ch. 7)	Upcoming
Part 4	ANOVA & designed experiments (Ch. 8–9)	Upcoming

We are entering the **modeling** phase of the course. Instead of testing whether a parameter equals a value, we now ask: *how are two variables related?*

Lecture 15 Outline

- Why regression? From hypothesis testing to modeling

Lecture 15 Outline

- Why regression? From hypothesis testing to modeling
- Empirical models and scatter diagrams (§6.1)

Lecture 15 Outline

- Why regression? From hypothesis testing to modeling
- Empirical models and scatter diagrams (§6.1)
- The simple linear regression model (§6.2)

Lecture 15 Outline

- Why regression? From hypothesis testing to modeling
- Empirical models and scatter diagrams (§6.1)
- The simple linear regression model (§6.2)
- Method of least squares (§6.3, Procedure 6.1)

Lecture 15 Outline

- Why regression? From hypothesis testing to modeling
- Empirical models and scatter diagrams (§6.1)
- The simple linear regression model (§6.2)
- Method of least squares (§6.3, Procedure 6.1)
- Properties of the least squares estimators (§6.5)

Lecture 15 Outline

- Why regression? From hypothesis testing to modeling
- Empirical models and scatter diagrams (§6.1)
- The simple linear regression model (§6.2)
- Method of least squares (§6.3, Procedure 6.1)
- Properties of the least squares estimators (§6.5)
- Estimating σ^2 and worked examples

Why Regression?

So far, our statistical questions have been:

- "Is the mean equal to some value?" (hypothesis testing)
- "What is a plausible range for the parameter?" (confidence intervals)

Why Regression?

So far, our statistical questions have been:

- "Is the mean equal to some value?" (hypothesis testing)
- "What is a plausible range for the parameter?" (confidence intervals)

Now we ask a **different kind of question**:

"How does one variable **relate to** *another?"*

Why Regression?

So far, our statistical questions have been:

- "Is the mean equal to some value?" (hypothesis testing)
- "What is a plausible range for the parameter?" (confidence intervals)

Now we ask a **different kind of question**:

"How does one variable **relate to** *another?"*

Examples:

- Does more verification testing reduce an AI system's failure rate?
- How does truck load affect fuel consumption on a delivery route?
- Does more ad spend bring a startup more signups?

Why Regression?

So far, our statistical questions have been:

- "Is the mean equal to some value?" (hypothesis testing)
- "What is a plausible range for the parameter?" (confidence intervals)

Now we ask a **different kind of question**:

"How does one variable **relate to** *another?"*

Examples:

- Does more verification testing reduce an AI system's failure rate?
- How does truck load affect fuel consumption on a delivery route?
- Does more ad spend bring a startup more signups?

Regression analysis is the tool for building these *empirical models*.

Textbook: Chapter 6, §6.1 ((6.1) onward)

Regression in ISE and AI

Real Applications You Will Encounter

In supply chains and transportation:

- Predict fuel cost from load, distance, and route
- Model port dwell time as a function of arrival volume
- Relate warehouse picking time to order size

Regression in ISE and AI

Real Applications You Will Encounter

In supply chains and transportation:

- Predict fuel cost from load, distance, and route
- Model port dwell time as a function of arrival volume
- Relate warehouse picking time to order size

In AI safety testing:

- Scaling laws: predicting model performance from training compute
- Verification: does increasing test coverage reduce failure rate?
- Calibration: is a sensor's reported value linearly related to the true value?

Regression in ISE and AI

Real Applications You Will Encounter

In supply chains and transportation:

- Predict fuel cost from load, distance, and route
- Model port dwell time as a function of arrival volume
- Relate warehouse picking time to order size

In AI safety testing:

- Scaling laws: predicting model performance from training compute
- Verification: does increasing test coverage reduce failure rate?
- Calibration: is a sensor's reported value linearly related to the true value?

In your own startup:

- Ad spend vs. signups, price vs. conversion, usage vs. churn

Empirical Models

An **empirical model** is built from *observed data*, not from first principles.

Empirical Models

An **empirical model** is built from *observed data*, not from first principles.

Suppose we observe pairs $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$:

- $x =$ **regressor** (predictor, independent variable)
- $y =$ **response** (dependent variable)

Empirical Models

An **empirical model** is built from *observed data*, not from first principles.

Suppose we observe pairs $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$:

- $x =$ **regressor** (predictor, independent variable)
- $y =$ **response** (dependent variable)

A **scatter diagram** plots each observation as a point (x_i, y_i) and helps us visualize the relationship.

Empirical Models

An **empirical model** is built from *observed data*, not from first principles.

Suppose we observe pairs $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$:

- $x =$ **regressor** (predictor, independent variable)
- $y =$ **response** (dependent variable)

A **scatter diagram** plots each observation as a point (x_i, y_i) and helps us visualize the relationship.

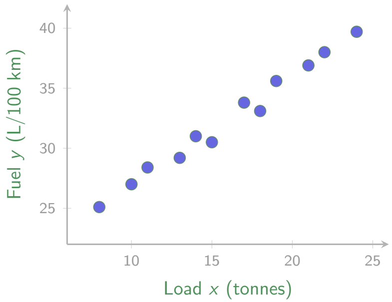
Key question: Does the scatter suggest a straight-line (linear) relationship? If so, we can fit a **simple linear regression** model.

Scatter Diagram: What Pattern Do You See?

Example: A logistics analyst records the load x (tonnes) and fuel consumption y (L/100 km) for 12 line-haul truck trips between Dammam and Riyadh.

Scatter Diagram: What Pattern Do You See?

Example: A logistics analyst records the load x (tonnes) and fuel consumption y (L/100 km) for 12 line-haul truck trips between Dammam and Riyadh.

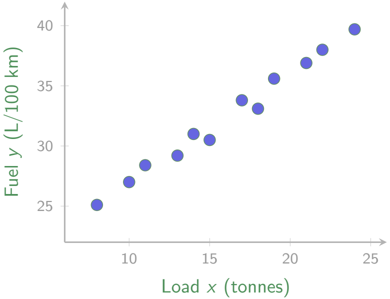


Scatter Diagram: What Pattern Do You See?

Example: A logistics analyst records the load x (tonnes) and fuel consumption y (L/100 km) for 12 line-haul truck trips between Dammam and Riyadh.

The scatter diagram suggests:

- A **positive linear** trend
- Heavier load $\Rightarrow$ more fuel per 100 km
- Some scatter around the trend



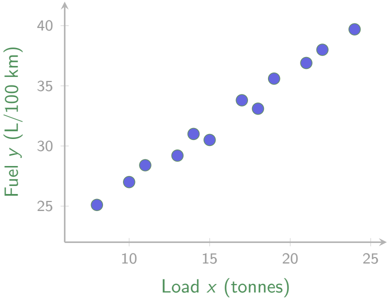
Scatter Diagram: What Pattern Do You See?

Example: A logistics analyst records the load x (tonnes) and fuel consumption y (L/100 km) for 12 line-haul truck trips between Dammam and Riyadh.

The scatter diagram suggests:

- A **positive linear** trend
- Heavier load $\Rightarrow$ more fuel per 100 km
- Some scatter around the trend

Can we **quantify** this relationship?



The Simple Linear Regression Model

We model the **expected value** of Y at each level of x as a straight line:

$$E(Y \mid x) = \beta_0 + \beta_1 x \tag{6.1}$$

The Simple Linear Regression Model

We model the **expected value** of Y at each level of x as a straight line:

$$E(Y \mid x) = \beta_0 + \beta_1 x \tag{6.1}$$

Each individual observation is:

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, 2, \ldots, n \tag{6.2}$$

The Simple Linear Regression Model

We model the **expected value** of Y at each level of x as a straight line:

$$E(Y \mid x) = \beta_0 + \beta_1 x \tag{6.1}$$

Each individual observation is:

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, 2, \ldots, n \tag{6.2}$$

where:

- β_0 = **intercept** (value of $E(Y)$ when $x = 0$)
- β_1 = **slope** (change in $E(Y)$ per one-unit increase in x)
- ε_i = **random error** for observation i

The Simple Linear Regression Model

We model the **expected value** of Y at each level of x as a straight line:

$$E(Y \mid x) = \beta_0 + \beta_1 x \tag{6.1}$$

Each individual observation is:

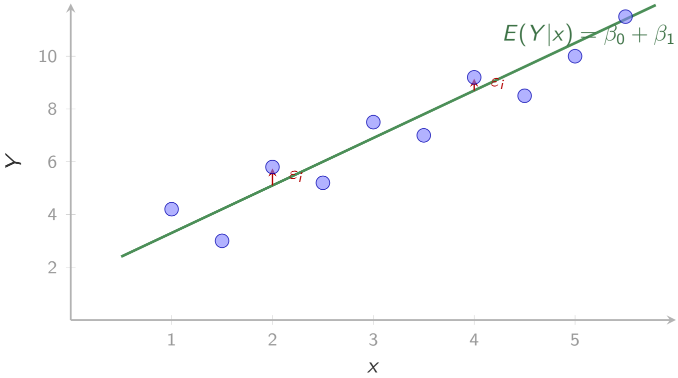
$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, 2, \ldots, n \tag{6.2}$$

where:

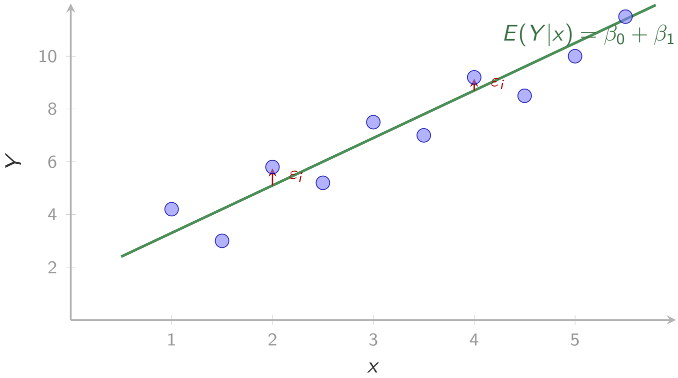
- β_0 = **intercept** (value of $E(Y)$ when $x = 0$)
- β_1 = **slope** (change in $E(Y)$ per one-unit increase in x)
- ε_i = **random error** for observation i

Assumptions on errors: $E(\varepsilon_i) = 0$, $\mathrm{Var}(\varepsilon_i) = \sigma^2$, and ε_i are uncorrelated.

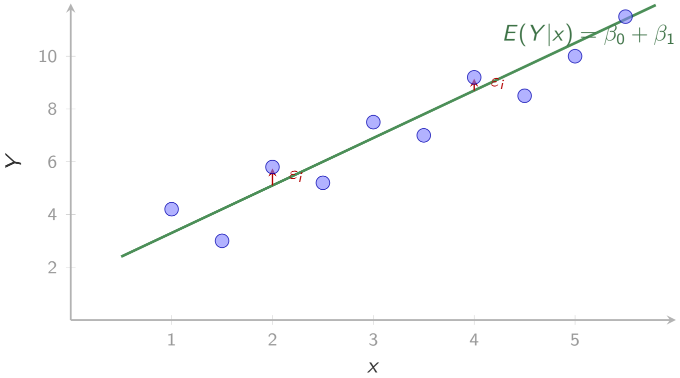
Visualizing the Model



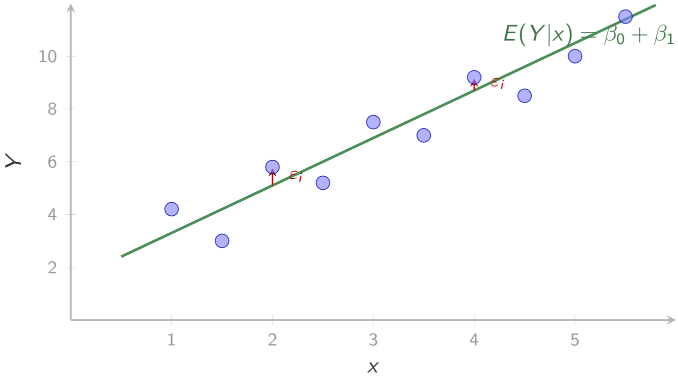
Visualizing the Model



Visualizing the Model



Visualizing the Model



The green line is the true (unknown) regression function (6.1). The blue points are the observed data. The red arrows show the errors ε_i.

Our Goal: Estimate the Unknown Parameters

We do not know β_0 and β_1. We need to **estimate** them from the data.

Our Goal: Estimate the Unknown Parameters

We do not know β_0 and β_1. We need to **estimate** them from the data.

Once we have estimates $\hat{\beta}_0$ and $\hat{\beta}_1$, the **fitted regression line** is:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x \tag{6.3}$$

Our Goal: Estimate the Unknown Parameters

We do not know β_0 and β_1. We need to **estimate** them from the data.

Once we have estimates $\hat{\beta}_0$ and $\hat{\beta}_1$, the **fitted regression line** is:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x \tag{6.3}$$

The **residual** for observation i is the difference between the actual value and the fitted value:

$$e_i = y_i - \hat{y}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i \tag{6.4}$$

Our Goal: Estimate the Unknown Parameters

We do not know β_0 and β_1. We need to **estimate** them from the data.

Once we have estimates $\hat{\beta}_0$ and $\hat{\beta}_1$, the **fitted regression line** is:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x \tag{6.3}$$

The **residual** for observation i is the difference between the actual value and the fitted value:

$$e_i = y_i - \hat{y}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i \tag{6.4}$$

Question: How should we choose $\hat{\beta}_0$ and $\hat{\beta}_1$ so that the line fits the data "best"?

Method of Least Squares

Minimizing the Sum of Squared Residuals

Choose $\hat{\beta}_0$ and $\hat{\beta}_1$ to **minimize** the sum of squared deviations:

$$S(\beta_0, \beta_1) = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 \tag{6.5}$$

Method of Least Squares

Minimizing the Sum of Squared Residuals

Choose $\hat{\beta}_0$ and $\hat{\beta}_1$ to **minimize** the sum of squared deviations:

$$S(\beta_0, \beta_1) = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 \tag{6.5}$$

Take partial derivatives, set them equal to zero, and solve. This gives the **normal equations**:

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i, \qquad \hat{\beta}_0 \sum_{i=1}^{n} x_i + \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i \tag{6.6}$$

Method of Least Squares

Minimizing the Sum of Squared Residuals

Choose $\hat{\beta}_0$ and $\hat{\beta}_1$ to **minimize** the sum of squared deviations:

$$S(\beta_0, \beta_1) = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 \tag{6.5}$$

Take partial derivatives, set them equal to zero, and solve. This gives the **normal equations**:

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i, \qquad \hat{\beta}_0 \sum_{i=1}^{n} x_i + \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i \tag{6.6}$$

Why "least squares"? We minimize the total squared distance from points to the line — the same core idea used in many machine learning methods!

Least Squares Estimators

The Formulas

First define the two building blocks:

$$S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum x_i\right)^2}{n} \tag{6.7}$$

$$S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i - \frac{\left(\sum x_i\right)\left(\sum y_i\right)}{n} \tag{6.8}$$

Least Squares Estimators

The Formulas

First define the two building blocks:

$$S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum x_i\right)^2}{n} \tag{6.7}$$

$$S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i - \frac{\left(\sum x_i\right)\left(\sum y_i\right)}{n} \tag{6.8}$$

Then the **least squares estimators** are:

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} \tag{6.9}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \tag{6.10}$$

Least Squares Estimators

The Formulas

First define the two building blocks:

$$S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum x_i\right)^2}{n} \tag{6.7}$$

$$S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i - \frac{\left(\sum x_i\right)\left(\sum y_i\right)}{n} \tag{6.8}$$

Then the **least squares estimators** are:

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} \tag{6.9}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \tag{6.10}$$

$\hat{\beta}_1$ measures the *co-movement* of x and y relative to the *spread* of x.

Procedure 6.1: Fitting a Simple Linear Regression

The Recipe We Follow Every Time

1. Compute the five sums: $\sum x_i$, $\sum y_i$, $\sum x_i^2$, $\sum x_i y_i$, $\sum y_i^2$.

Procedure 6.1: Fitting a Simple Linear Regression

The Recipe We Follow Every Time

1. Compute the five sums: $\sum x_i$, $\sum y_i$, $\sum x_i^2$, $\sum x_i y_i$, $\sum y_i^2$.
2. Compute $\bar{x}$, $\bar{y}$, then S_{xx} from (6.7) and S_{xy} from (6.8).

Procedure 6.1: Fitting a Simple Linear Regression

The Recipe We Follow Every Time

1. Compute the five sums: $\sum x_i$, $\sum y_i$, $\sum x_i^2$, $\sum x_i y_i$, $\sum y_i^2$.
2. Compute $\bar{x}$, $\bar{y}$, then S_{xx} from (6.7) and S_{xy} from (6.8).
3. Compute the slope $\hat{\beta}_1 = S_{xy}/S_{xx}$ from (6.9).

Procedure 6.1: Fitting a Simple Linear Regression

The Recipe We Follow Every Time

1. Compute the five sums: $\sum x_i$, $\sum y_i$, $\sum x_i^2$, $\sum x_i y_i$, $\sum y_i^2$.
2. Compute $\bar{x}$, $\bar{y}$, then S_{xx} from (6.7) and S_{xy} from (6.8).
3. Compute the slope $\hat{\beta}_1 = S_{xy}/S_{xx}$ from (6.9).
4. Compute the intercept $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ from (6.10).

Procedure 6.1: Fitting a Simple Linear Regression

The Recipe We Follow Every Time

1. Compute the five sums: $\sum x_i$, $\sum y_i$, $\sum x_i^2$, $\sum x_i y_i$, $\sum y_i^2$.
2. Compute $\bar{x}$, $\bar{y}$, then S_{xx} from (6.7) and S_{xy} from (6.8).
3. Compute the slope $\hat{\beta}_1 = S_{xy}/S_{xx}$ from (6.9).
4. Compute the intercept $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ from (6.10).
5. Write the fitted line $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ (6.3) and interpret the slope in context.

Procedure 6.1: Fitting a Simple Linear Regression

The Recipe We Follow Every Time

1. Compute the five sums: $\sum x_i$, $\sum y_i$, $\sum x_i^2$, $\sum x_i y_i$, $\sum y_i^2$.
2. Compute $\bar{x}$, $\bar{y}$, then S_{xx} from (6.7) and S_{xy} from (6.8).
3. Compute the slope $\hat{\beta}_1 = S_{xy}/S_{xx}$ from (6.9).
4. Compute the intercept $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ from (6.10).
5. Write the fitted line $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ (6.3) and interpret the slope in context.

This is Procedure 6.1 in the textbook. We will use it in every regression problem, including exams.

Example: AI Safety Verification

Does More Testing Reduce Failures? (Example 6.1)

An ISE engineer is evaluating an autonomous inspection system. She runs the system through different amounts of verification testing (hours) and records the failure rate (failures per 1000 inspections) in deployment:

Example: AI Safety Verification

Does More Testing Reduce Failures? (Example 6.1)

An ISE engineer is evaluating an autonomous inspection system. She runs the system through different amounts of verification testing (hours) and records the failure rate (failures per 1000 inspections) in deployment:

x (test hours)	10	15	20	25	30	35	40	45
y (failures/1000)	48	42	38	32	29	24	22	18

Example: AI Safety Verification

Does More Testing Reduce Failures? (Example 6.1)

An ISE engineer is evaluating an autonomous inspection system. She runs the system through different amounts of verification testing (hours) and records the failure rate (failures per 1000 inspections) in deployment:

x (test hours)	10	15	20	25	30	35	40	45
y (failures/1000)	48	42	38	32	29	24	22	18

Step 1 of Procedure 6.1 — the five sums ($n = 8$):

$$\sum x_i = 220, \quad \sum y_i = 253, \quad \sum x_i^2 = 7100, \quad \sum x_i y_i = 6070, \quad \sum y_i^2 = 8761$$

Example: AI Safety Verification

Does More Testing Reduce Failures? (Example 6.1)

An ISE engineer is evaluating an autonomous inspection system. She runs the system through different amounts of verification testing (hours) and records the failure rate (failures per 1000 inspections) in deployment:

x (test hours)	10	15	20	25	30	35	40	45
y (failures/1000)	48	42	38	32	29	24	22	18

Step 1 of Procedure 6.1 — the five sums ($n = 8$):

$$\sum x_i = 220, \quad \sum y_i = 253, \quad \sum x_i^2 = 7100, \quad \sum x_i y_i = 6070, \quad \sum y_i^2 = 8761$$

Example: AI Safety Verification

Step 2: S_{xx} and S_{xy}

Spread of x, using (6.7):

$$S_{xx} = \sum x_i^2 - \frac{(\sum x_i)^2}{n} = 7100 - \frac{(220)^2}{8}$$

Example: AI Safety Verification

Step 2: S_{xx} and S_{xy}

Spread of x, using (6.7):

$$S_{xx} = \sum x_i^2 - \frac{(\sum x_i)^2}{n} = 7100 - \frac{(220)^2}{8}$$

$$S_{xx} = 7100 - 6050 = 1050$$

Example: AI Safety Verification

Step 2: S_{xx} and S_{xy}

Spread of x, using (6.7):

$$S_{xx} = \sum x_i^2 - \frac{(\sum x_i)^2}{n} = 7100 - \frac{(220)^2}{8}$$

$$S_{xx} = 7100 - 6050 = 1050$$

Co-movement of x **and** y, using (6.8):

$$S_{xy} = \sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n} = 6070 - \frac{(220)(253)}{8}$$

Example: AI Safety Verification

Step 2: S_{xx} and S_{xy}

Spread of x, using (6.7):

$$S_{xx} = \sum x_i^2 - \frac{(\sum x_i)^2}{n} = 7100 - \frac{(220)^2}{8}$$

$$S_{xx} = 7100 - 6050 = 1050$$

Co-movement of x **and** y, using (6.8):

$$S_{xy} = \sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n} = 6070 - \frac{(220)(253)}{8}$$

$$S_{xy} = 6070 - 6957.5 = -887.5$$

Example: AI Safety Verification

Step 2: S_{xx} and S_{xy}

Spread of x, using (6.7):

$$S_{xx} = \sum x_i^2 - \frac{(\sum x_i)^2}{n} = 7100 - \frac{(220)^2}{8}$$

$$S_{xx} = 7100 - 6050 = 1050$$

Co-movement of x **and** y, using (6.8):

$$S_{xy} = \sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n} = 6070 - \frac{(220)(253)}{8}$$

$$S_{xy} = 6070 - 6957.5 = -887.5$$

Example: AI Safety Verification

Steps 3–5: Slope, Intercept, Fitted Line

Step 3: Slope, from (6.9):

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{-887.5}{1050} = -0.8452$$

Example: AI Safety Verification

Steps 3–5: Slope, Intercept, Fitted Line

Step 3: Slope, from (6.9):

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{-887.5}{1050} = -0.8452$$

Step 4: Intercept, from (6.10):

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 31.625 - (-0.8452)(27.5) = 54.87$$

Example: AI Safety Verification

Steps 3–5: Slope, Intercept, Fitted Line

Step 3: Slope, from (6.9):

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{-887.5}{1050} = -0.8452$$

Step 4: Intercept, from (6.10):

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 31.625 - (-0.8452)(27.5) = 54.87$$

Step 5: Write the fitted model (6.3):

$$\boxed{\hat{y} = 54.87 - 0.8452\,x}$$

Example: AI Safety Verification

Steps 3–5: Slope, Intercept, Fitted Line

Step 3: Slope, from (6.9):

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{-887.5}{1050} = -0.8452$$

Step 4: Intercept, from (6.10):

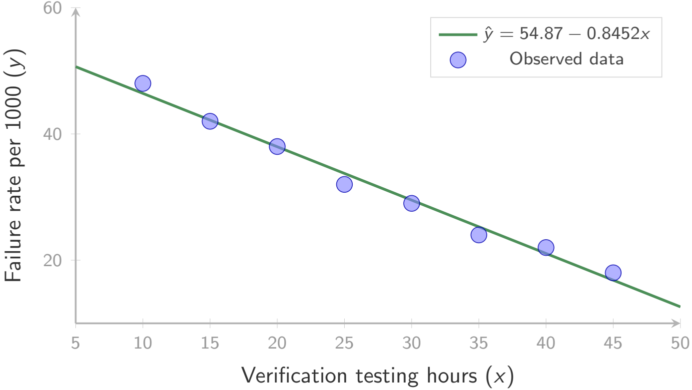
$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 31.625 - (-0.8452)(27.5) = 54.87$$

Step 5: Write the fitted model (6.3):

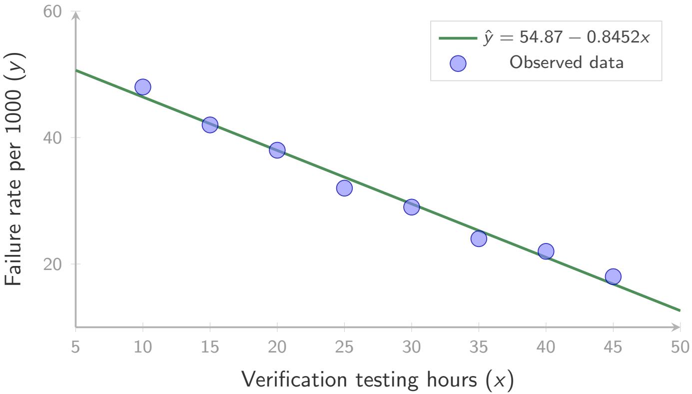
$$\boxed{\hat{y} = 54.87 - 0.8452\, x}$$

Interpretation: Each additional hour of verification testing is associated with a

Example: Fitted Regression Line



Example: Fitted Regression Line



Think About It

Based on the fitted model $\hat{y} = 54.87 - 0.8452\,x$:

Think About It

Based on the fitted model $\hat{y} = 54.87 - 0.8452\,x$:

Question 1: What failure rate would you predict for $x = 50$ hours of testing?

Think About It

Based on the fitted model $\hat{y} = 54.87 - 0.8452\,x$:

Question 1: What failure rate would you predict for $x = 50$ hours of testing?

$\hat{y} = 54.87 - 0.8452(50) = 54.87 - 42.26 = 12.61$ failures per 1000.

Think About It

Based on the fitted model $\hat{y} = 54.87 - 0.8452\,x$:

Question 1: What failure rate would you predict for $x = 50$ hours of testing?

$\hat{y} = 54.87 - 0.8452(50) = 54.87 - 42.26 = 12.61$ failures per 1000.

Question 2: Would you trust this prediction for $x = 200$ hours?

Think About It

Based on the fitted model $\hat{y} = 54.87 - 0.8452\,x$:

Question 1: What failure rate would you predict for $x = 50$ hours of testing?

$\hat{y} = 54.87 - 0.8452(50) = 54.87 - 42.26 = 12.61$ failures per 1000.

Question 2: Would you trust this prediction for $x = 200$ hours?

$\hat{y} = 54.87 - 0.8452(200) = -114.2 \quad$ (negative — nonsensical!)

Warning: Do not **extrapolate** far beyond the range of the observed data. The linear model may not hold outside $10 \leq x \leq 45$.

Properties of the Least Squares Estimators

§6.5

The least squares estimators $\hat{\beta}_0$ and $\hat{\beta}_1$ have nice statistical properties:

Properties of the Least Squares Estimators

§6.5

The least squares estimators $\hat{\beta}_0$ and $\hat{\beta}_1$ have nice statistical properties:

Both are unbiased: $E(\hat{\beta}_1) = \beta_1$ and $E(\hat{\beta}_0) = \beta_0$.

Properties of the Least Squares Estimators

§6.5

The least squares estimators $\hat{\beta}_0$ and $\hat{\beta}_1$ have nice statistical properties:

Both are unbiased: $E(\hat{\beta}_1) = \beta_1$ and $E(\hat{\beta}_0) = \beta_0$.

Their estimated standard errors (used for tests and CIs next lecture):

$$se(\hat{\beta}_1) = \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \tag{6.15}$$

$$se(\hat{\beta}_0) = \sqrt{\hat{\sigma}^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)} \tag{6.16}$$

Properties of the Least Squares Estimators

§6.5

The least squares estimators $\hat{\beta}_0$ and $\hat{\beta}_1$ have nice statistical properties:

Both are unbiased: $E(\hat{\beta}_1) = \beta_1$ and $E(\hat{\beta}_0) = \beta_0$.

Their estimated standard errors (used for tests and CIs next lecture):

$$se(\hat{\beta}_1) = \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \tag{6.15}$$

$$se(\hat{\beta}_0) = \sqrt{\hat{\sigma}^2\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)} \tag{6.16}$$

These are **exactly the same ideas** from Chapters 2–3: unbiasedness and standard error of estimators. Regression builds on what you already know!

Estimating σ^2

The Error Variance (§6.4)

We also need to estimate the error variance σ^2. The **error sum of squares** is:

$$SS_E = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{6.11}$$

Estimating σ^2

The Error Variance (§6.4)

We also need to estimate the error variance σ^2. The **error sum of squares** is:

$$SS_E = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{6.11}$$

A computational shortcut uses the total variability of y, S_{yy} (6.12):

$$SS_E = S_{yy} - \hat{\beta}_1 S_{xy}, \qquad S_{yy} = \sum_{i=1}^{n} y_i^2 - \frac{\left(\sum y_i\right)^2}{n} \tag{6.13}$$

Estimating σ^2

The Error Variance (§6.4)

We also need to estimate the error variance σ^2. The **error sum of squares** is:

$$SS_E = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{6.11}$$

A computational shortcut uses the total variability of y, S_{yy} (6.12):

$$SS_E = S_{yy} - \hat{\beta}_1 S_{xy}, \qquad S_{yy} = \sum_{i=1}^{n} y_i^2 - \frac{(\sum y_i)^2}{n} \tag{6.13}$$

The **unbiased estimator** of σ^2 is:

$$\hat{\sigma}^2 = \frac{SS_E}{n-2} \tag{6.14}$$

Estimating σ^2

The Error Variance (§6.4)

We also need to estimate the error variance σ^2. The **error sum of squares** is:

$$SS_E = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{6.11}$$

A computational shortcut uses the total variability of y, S_{yy} (6.12):

$$SS_E = S_{yy} - \hat{\beta}_1 S_{xy}, \qquad S_{yy} = \sum_{i=1}^{n} y_i^2 - \frac{\left(\sum y_i\right)^2}{n} \tag{6.13}$$

The **unbiased estimator** of σ^2 is:

$$\hat{\sigma}^2 = \frac{SS_E}{n-2} \tag{6.14}$$

Example: Computing $\hat{\sigma}^2$

AI Safety Verification, Continued (Example 6.2)

From our AI verification example, using (6.12):

$$S_{yy} = \sum y_i^2 - \frac{(\sum y_i)^2}{n} = 8761 - \frac{(253)^2}{8} = 8761 - 8001.125 = 759.875$$

Example: Computing $\hat{\sigma}^2$

AI Safety Verification, Continued (Example 6.2)

From our AI verification example, using (6.12):

$$S_{yy} = \sum y_i^2 - \frac{(\sum y_i)^2}{n} = 8761 - \frac{(253)^2}{8} = 8761 - 8001.125 = 759.875$$

Then the shortcut (6.13):

$$SS_E = S_{yy} - \hat{\beta}_1 \cdot S_{xy} = 759.875 - (-0.8452)(-887.5) = 9.726$$

Example: Computing $\hat{\sigma}^2$

AI Safety Verification, Continued (Example 6.2)

From our AI verification example, using (6.12):

$$S_{yy} = \sum y_i^2 - \frac{(\sum y_i)^2}{n} = 8761 - \frac{(253)^2}{8} = 8761 - 8001.125 = 759.875$$

Then the shortcut (6.13):

$$SS_E = S_{yy} - \hat{\beta}_1 \cdot S_{xy} = 759.875 - (-0.8452)(-887.5) = 9.726$$

Finally, (6.14):

$$\hat{\sigma}^2 = \frac{SS_E}{n-2} = \frac{9.726}{6} = 1.621 \qquad \Rightarrow \qquad \hat{\sigma} = \sqrt{1.621} = 1.273 \text{ failures per 1000}$$

Example: Computing $\hat{\sigma}^2$

AI Safety Verification, Continued (Example 6.2)

From our AI verification example, using (6.12):

$$S_{yy} = \sum y_i^2 - \frac{(\sum y_i)^2}{n} = 8761 - \frac{(253)^2}{8} = 8761 - 8001.125 = 759.875$$

Then the shortcut (6.13):

$$SS_E = S_{yy} - \hat{\beta}_1 \cdot S_{xy} = 759.875 - (-0.8452)(-887.5) = 9.726$$

Finally, (6.14):

$$\hat{\sigma}^2 = \frac{SS_E}{n-2} = \frac{9.726}{6} = 1.621 \qquad \Rightarrow \qquad \hat{\sigma} = \sqrt{1.621} = 1.273 \text{ failures per 1000}$$

Mini-Exercise: Startup Ad Spend

Try Procedure 6.1 Yourself

Your startup runs a six-week marketing test. Each week you set the ad spend x (SAR thousands) and count new signups y:

Mini-Exercise: Startup Ad Spend

Try Procedure 6.1 Yourself

Your startup runs a six-week marketing test. Each week you set the ad spend x (SAR thousands) and count new signups y:

x (ad spend, SAR 1000s)	2	4	6	8	10	12
y (new signups)	55	70	92	110	118	135

Mini-Exercise: Startup Ad Spend

Try Procedure 6.1 Yourself

Your startup runs a six-week marketing test. Each week you set the ad spend x (SAR thousands) and count new signups y:

x (ad spend, SAR 1000s)	2	4	6	8	10	12
y (new signups)	55	70	92	110	118	135

Sums ($n = 6$):

$$\sum x_i = 42, \quad \sum y_i = 580, \quad \sum x_i^2 = 364, \quad \sum x_i y_i = 4622$$

Mini-Exercise: Startup Ad Spend

Try Procedure 6.1 Yourself

Your startup runs a six-week marketing test. Each week you set the ad spend x (SAR thousands) and count new signups y:

x (ad spend, SAR 1000s)	2	4	6	8	10	12
y (new signups)	55	70	92	110	118	135

Sums ($n = 6$):

$$\sum x_i = 42, \quad \sum y_i = 580, \quad \sum x_i^2 = 364, \quad \sum x_i y_i = 4622$$

Task: Fit the least squares line. What does the slope tell the founders?

Mini-Exercise: Solution

Using (6.7) and (6.8):

$$S_{xx} = 364 - \frac{(42)^2}{6} = 364 - 294 = 70, \qquad S_{xy} = 4622 - \frac{(42)(580)}{6} = 4622 - 4060 = 562$$

Mini-Exercise: Solution

Using (6.7) and (6.8):

$$S_{xx} = 364 - \frac{(42)^2}{6} = 364 - 294 = 70, \qquad S_{xy} = 4622 - \frac{(42)(580)}{6} = 4622 - 4060 = 562$$

Slope, from (6.9): $\hat{\beta}_1 = \dfrac{562}{70} = 8.029$

Mini-Exercise: Solution

Using (6.7) and (6.8):

$$S_{xx} = 364 - \frac{(42)^2}{6} = 364 - 294 = 70, \qquad S_{xy} = 4622 - \frac{(42)(580)}{6} = 4622 - 4060 = 562$$

Slope, from (6.9): $\hat{\beta}_1 = \dfrac{562}{70} = 8.029$

Intercept, from (6.10), with $\bar{x} = 7$, $\bar{y} = 96.67$:

$$\hat{\beta}_0 = 96.67 - 8.029(7) = 40.47 \qquad \Rightarrow \qquad \boxed{\hat{y} = 40.47 + 8.029\,x}$$

Mini-Exercise: Solution

Using (6.7) and (6.8):

$$S_{xx} = 364 - \frac{(42)^2}{6} = 364 - 294 = 70, \qquad S_{xy} = 4622 - \frac{(42)(580)}{6} = 4622 - 4060 = 562$$

Slope, from (6.9): $\hat{\beta}_1 = \dfrac{562}{70} = 8.029$

Intercept, from (6.10), with $\bar{x} = 7$, $\bar{y} = 96.67$:

$$\hat{\beta}_0 = 96.67 - 8.029(7) = 40.47 \qquad \Rightarrow \qquad \boxed{\hat{y} = 40.47 + 8.029\,x}$$

Interpretation: Each extra SAR 1000 of weekly ad spend is associated with about 8 more signups. The intercept (≈ 40) is the baseline signup level with no ads — inside the story, but $x = 0$ is outside the tested range, so treat it with care.

Connecting Old and New Ideas

Concept	Ch. 2–5	Ch. 6 (Regression)
Parameter	μ, σ^2, p	β_0, β_1, σ^2
Estimator	$\bar{X}$, S^2, $\hat{P}$	$\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\sigma}^2$
Unbiased?	Yes	Yes
Degrees of freedom	$n-1$	$n-2$
Next step	CI and hypothesis test	*Same — coming lectures!*

Connecting Old and New Ideas

Concept	**Ch. 2–5**	**Ch. 6 (Regression)**
Parameter	μ, σ^2, p	β_0, β_1, σ^2
Estimator	$\bar{X}$, S^2, $\hat{P}$	$\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\sigma}^2$
Unbiased?	Yes	Yes
Degrees of freedom	$n-1$	$n-2$
Next step	CI and hypothesis test	*Same — coming lectures!*

The structure is **identical**: estimate parameters, quantify uncertainty, then test hypotheses and build confidence intervals.

Lecture 15 Summary

- **Simple linear regression** models the relationship $Y = \beta_0 + \beta_1 x + \varepsilon$ (6.2).

Lecture 15 Summary

- **Simple linear regression** models the relationship $Y = \beta_0 + \beta_1 x + \varepsilon$ (6.2).
- The **method of least squares** minimizes $\sum e_i^2$ (6.5) to estimate β_0 and β_1.

Lecture 15 Summary

- **Simple linear regression** models the relationship $Y = \beta_0 + \beta_1 x + \varepsilon$ (6.2).
- The **method of least squares** minimizes $\sum e_i^2$ (6.5) to estimate β_0 and β_1.
- Key formulas: $\hat{\beta}_1 = S_{xy}/S_{xx}$ (6.9) and $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$ (6.10) — follow Procedure 6.1.

Lecture 15 Summary

- **Simple linear regression** models the relationship $Y = \beta_0 + \beta_1 x + \varepsilon$ (6.2).
- The **method of least squares** minimizes $\sum e_i^2$ (6.5) to estimate β_0 and β_1.
- Key formulas: $\hat{\beta}_1 = S_{xy}/S_{xx}$ (6.9) and $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$ (6.10) — follow Procedure 6.1.
- The error variance is estimated by $\hat{\sigma}^2 = SS_E/(n-2)$ (6.14).

Lecture 15 Summary

- **Simple linear regression** models the relationship $Y = \beta_0 + \beta_1 x + \varepsilon$ (6.2).
- The **method of least squares** minimizes $\sum e_i^2$ (6.5) to estimate β_0 and β_1.
- Key formulas: $\hat{\beta}_1 = S_{xy}/S_{xx}$ (6.9) and $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ (6.10) — follow Procedure 6.1.
- The error variance is estimated by $\hat{\sigma}^2 = SS_E/(n-2)$ (6.14).
- **Do not extrapolate** the model far beyond the range of observed x values.

Lecture 15 Summary

- **Simple linear regression** models the relationship $Y = \beta_0 + \beta_1 x + \varepsilon$ (6.2).
- The **method of least squares** minimizes $\sum e_i^2$ (6.5) to estimate β_0 and β_1.
- Key formulas: $\hat{\beta}_1 = S_{xy}/S_{xx}$ (6.9) and $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ (6.10) — follow Procedure 6.1.
- The error variance is estimated by $\hat{\sigma}^2 = SS_E/(n-2)$ (6.14).
- **Do not extrapolate** the model far beyond the range of observed x values.
- **Next lecture:** software tools (Chapter 10); then a full worked regression fit, step by step.

Reminders

- **Major Exam 1 grades** will be on Blackboard by **Tuesday**

Reminders

- **Major Exam 1 grades** will be on Blackboard by **Tuesday**
- If you have grade questions, please come next class or the office hour

Reminders

- **Major Exam 1 grades** will be on Blackboard by **Tuesday**
- If you have grade questions, please come next class or the office hour
- Office hours: Tuesdays 9–10 AM, Room 22-219 or Zoom.

Reminders

- **Major Exam 1 grades** will be on Blackboard by **Tuesday**
- If you have grade questions, please come next class or the office hour
- Office hours: Tuesdays 9–10 AM, Room 22-219 or Zoom.
- Start reading **Chapter 6** (§6.1 through §6.5 for now).