

---

# Regression: Simple Version

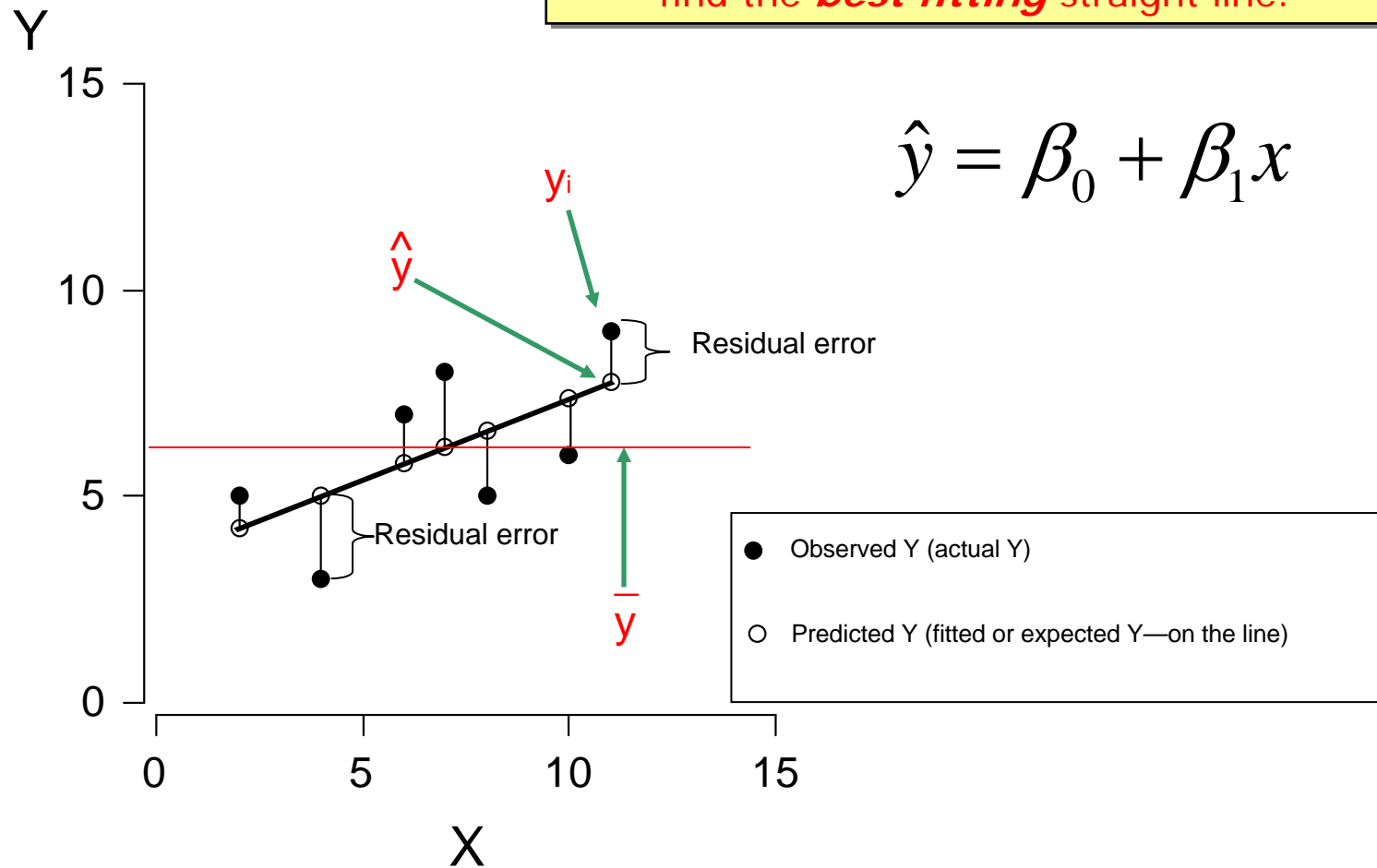


*The Lean Six Sigma Institute*

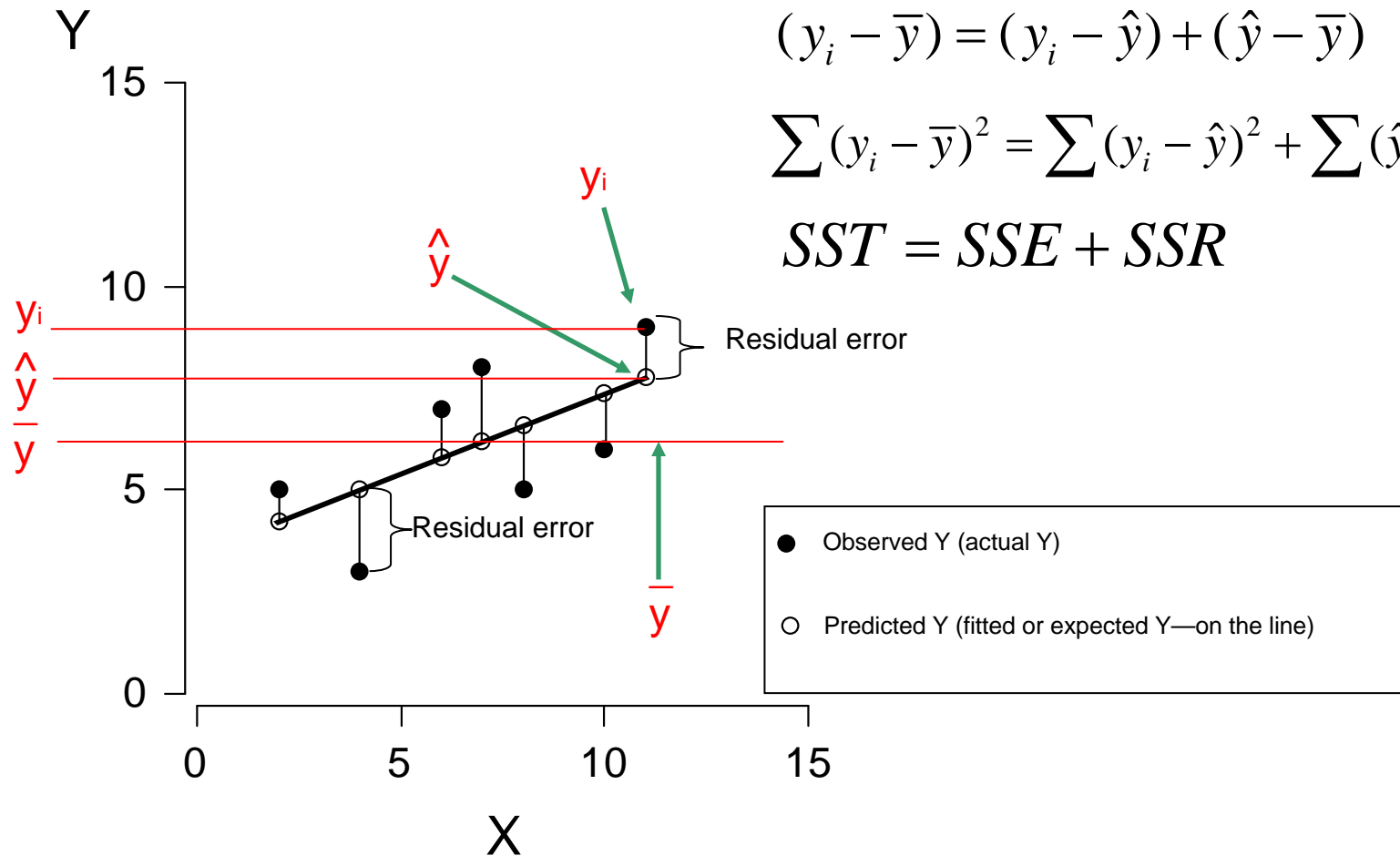
[www.lean-six-sigma.pro](http://www.lean-six-sigma.pro)

## REGRESSION LINE

We use the **method of least squares** to find the **best fitting** straight line.

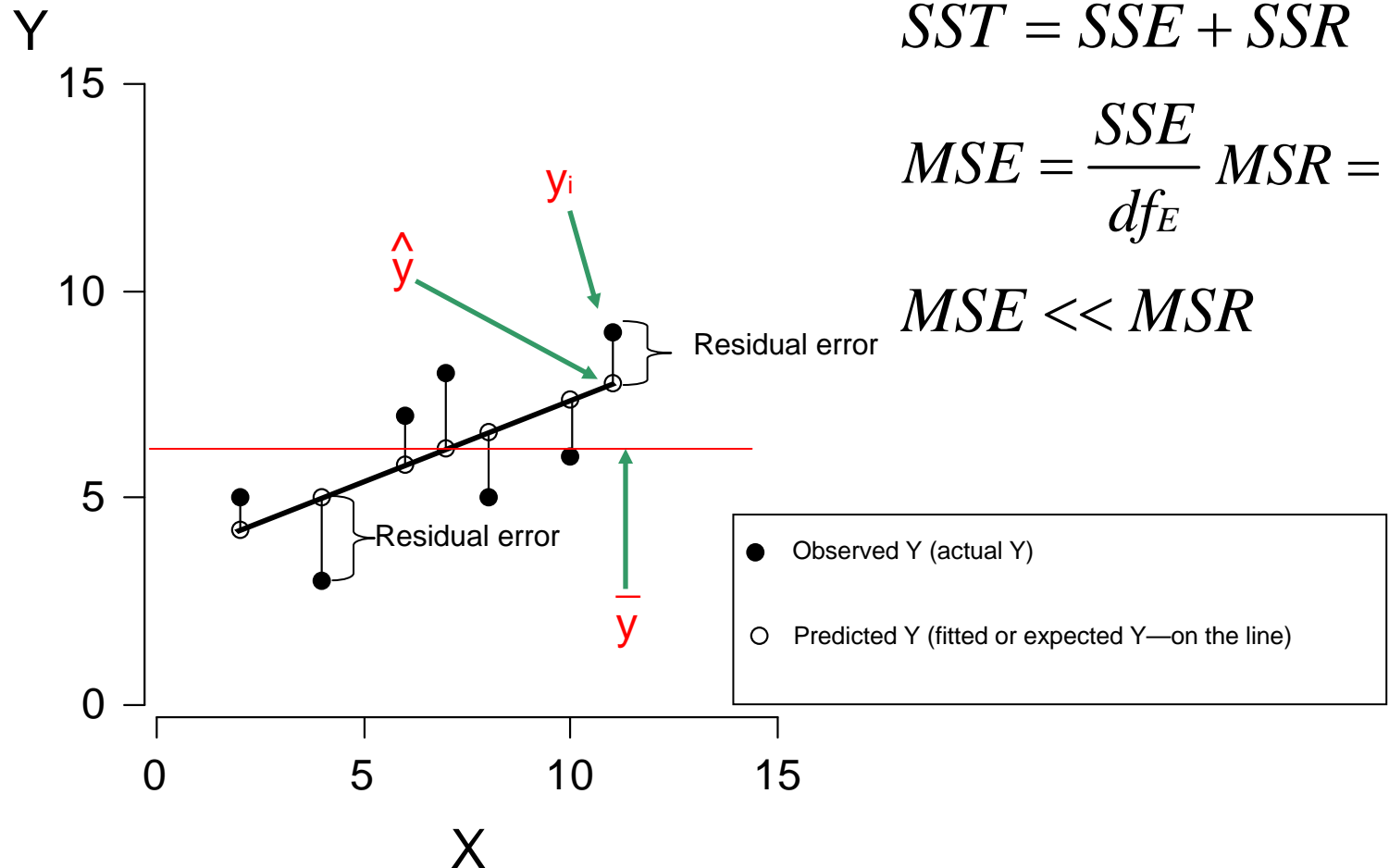


## REGRESSION SUMS OF THE SQUARES



We find the *regression sums of squares*.

## MEAN SQUARE ERROR AND MEAN SQUARE REGRESSION



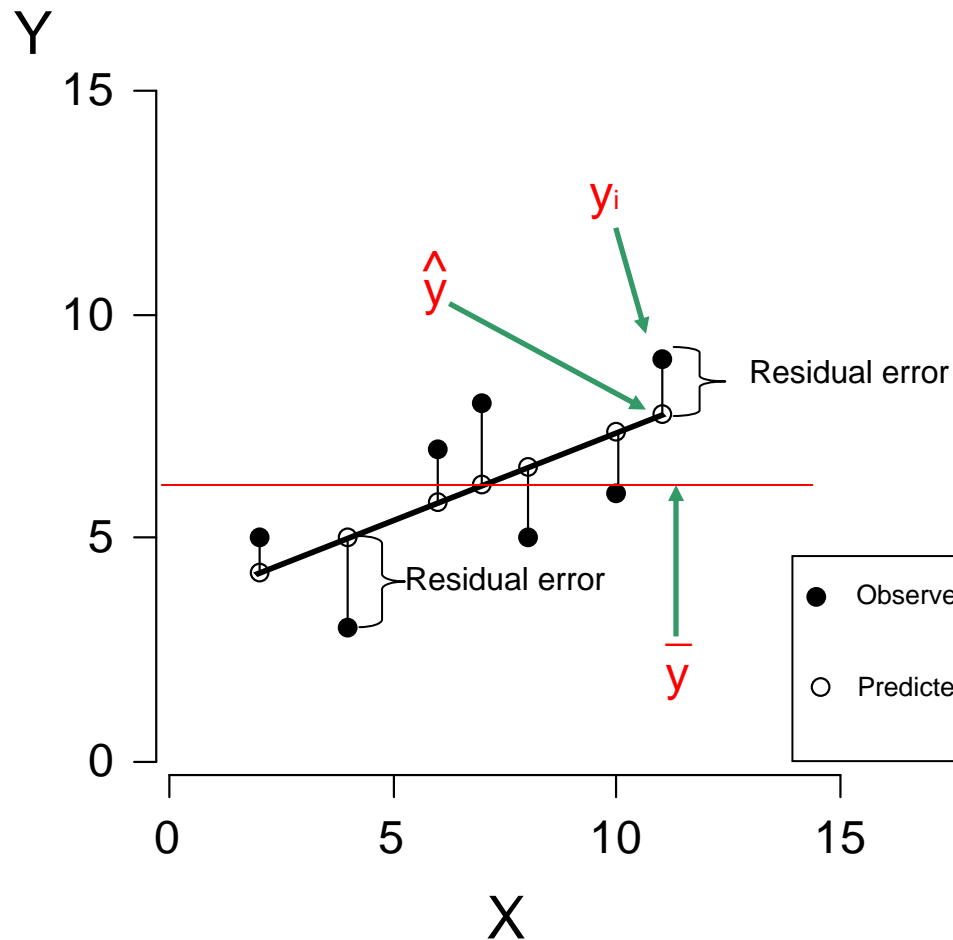
$$SST = SSE + SSR$$

$$MSE = \frac{SSE}{df_E} \quad MSR = \frac{SSR}{df_R}$$

$$MSE \ll MSR$$

To find the **best fitting** straight line, the variance of the error (**Mean Square Error**) about the Regression Line must be **significantly less than the variance of y** about the mean.

## REGRESSION AND R<sup>2</sup>



$$SST = SSE + SSR$$

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

| REGRESSION | R <sup>2</sup> | SSE = |
|------------|----------------|-------|
| NONE       | 0              | SST   |
| PERFECT    | 1              | 0     |

- Observed Y (actual Y)
- Predicted Y (fitted or expected Y—on the line)

We use R<sup>2</sup> to **check how much of the variability is explained by the regression line.**

## REGRESSION F STATISTIC AND HYPOTHESIS TEST

We then use the F statistic with the ANOVA table to find the p value and test the Null Hypothesis

$$\hat{y} = \beta_0 + \beta_1 x$$

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0$$

$$SST = SSE + SSR$$

$$MSR = \frac{SSR}{df_R}$$

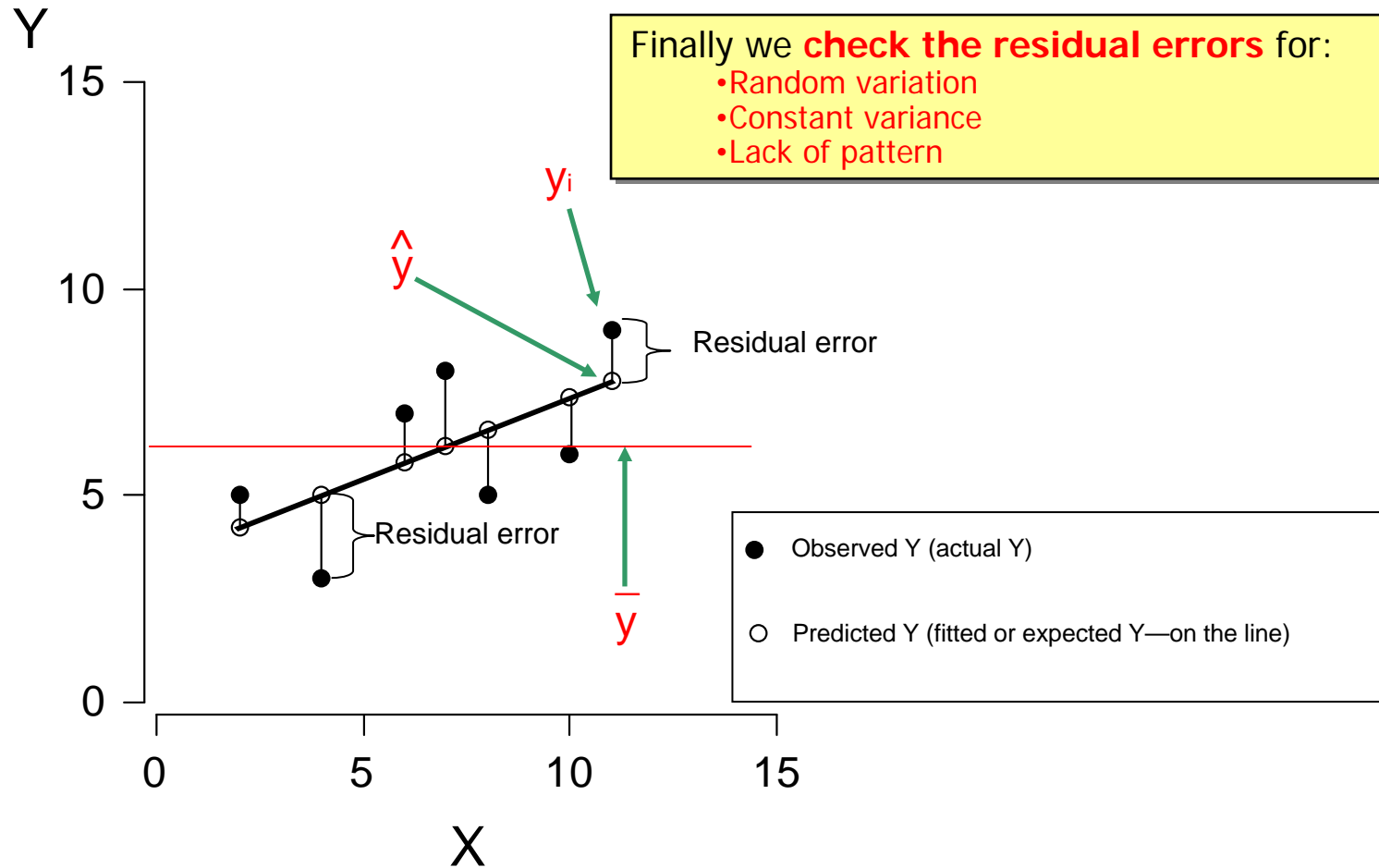
$$MSE = \frac{SSE}{df_E}$$

$$s_y^2 = \frac{SST}{df_T}$$

|            | Sum of Squares | df              | Mean Square | F value |
|------------|----------------|-----------------|-------------|---------|
| Regression | SSR            | df <sub>R</sub> | MSR         | MSR/MSE |
| Error      | SSE            | df <sub>E</sub> | MSE         |         |
| Total      | SST            | df <sub>T</sub> |             |         |

$$MSE \ll MSR$$

# REGRESSION



---

# Regression: Version citing statistical formulae

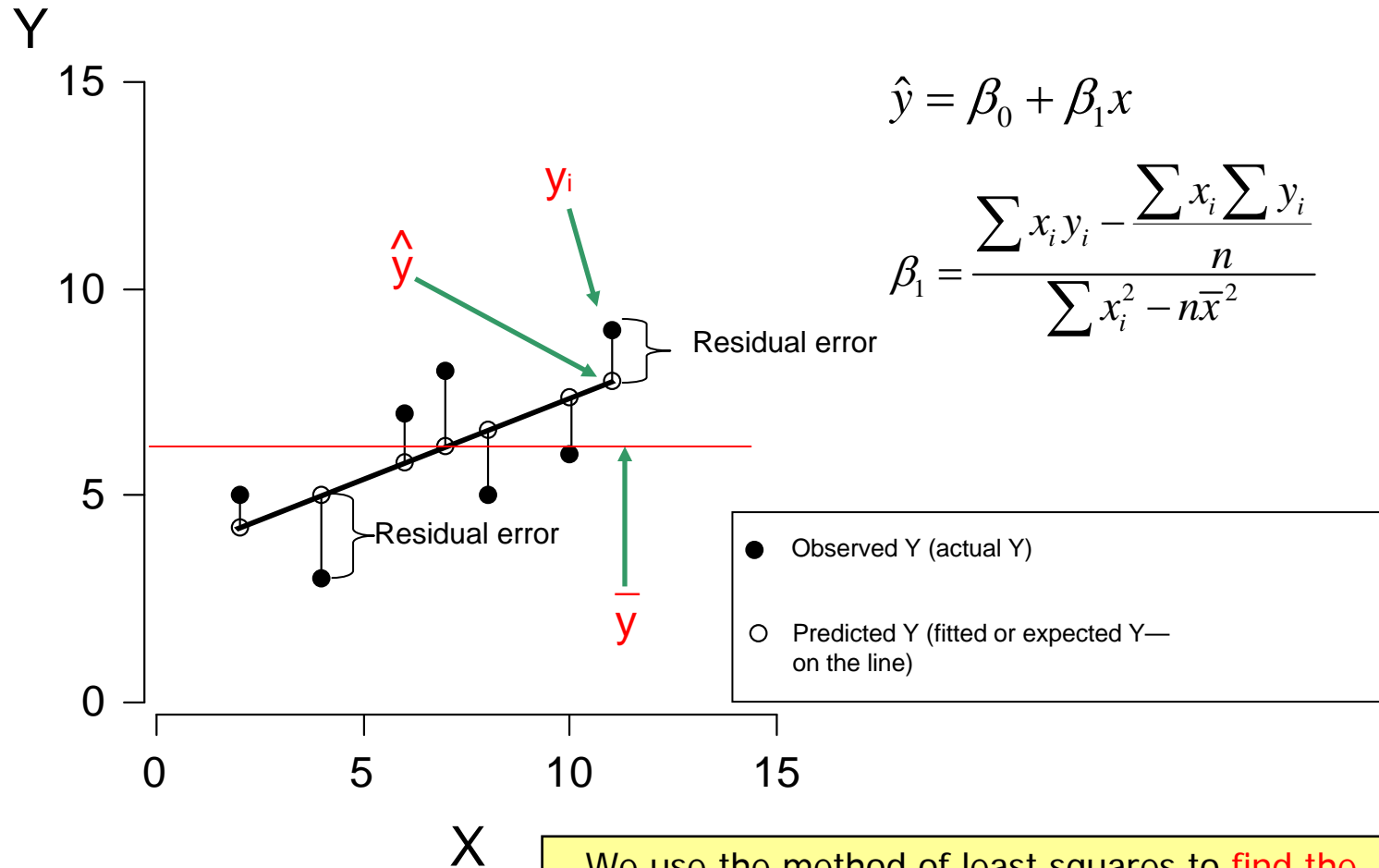


*The Lean Six Sigma Institute*

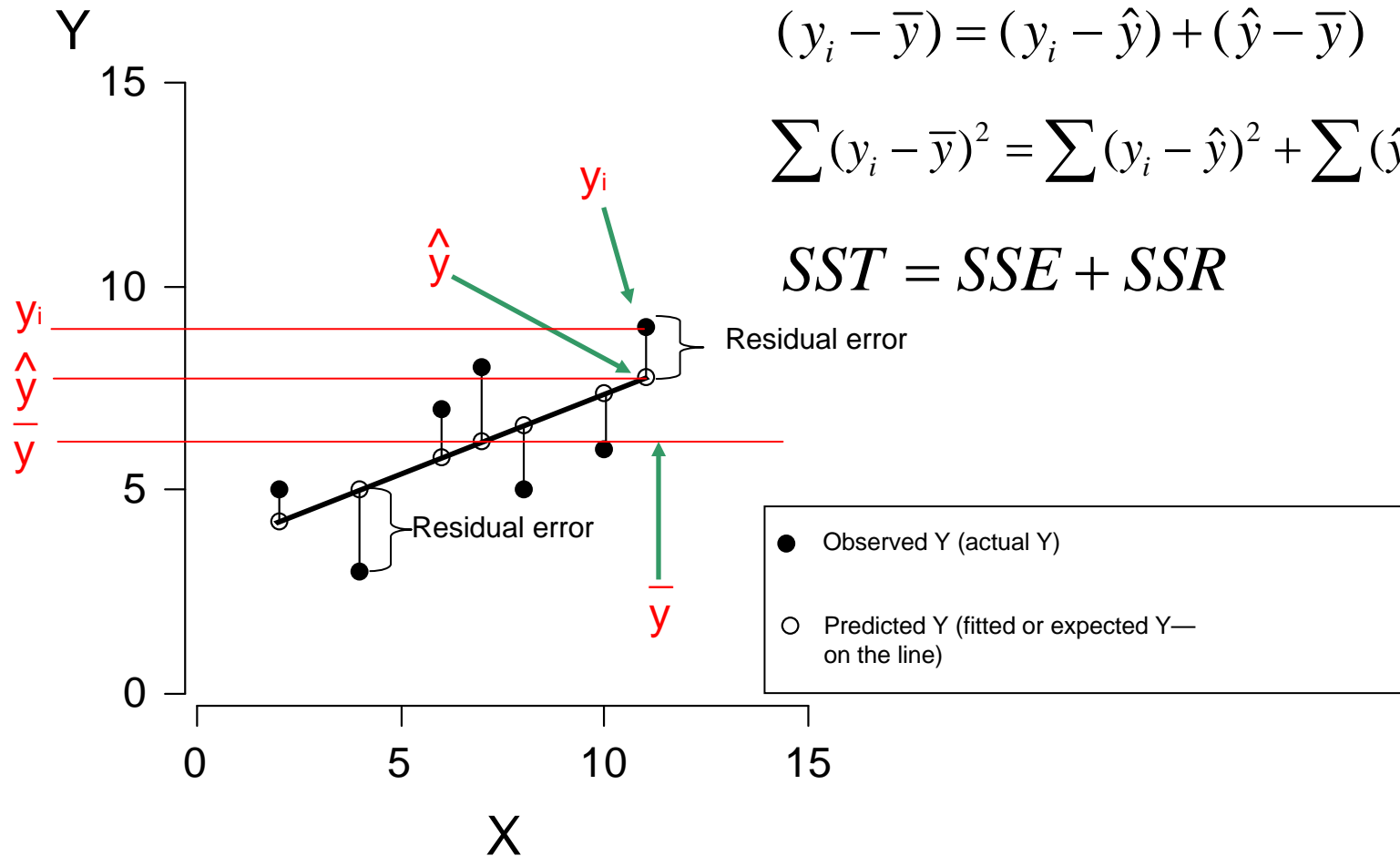
[www.lean-six-sigma.pro](http://www.lean-six-sigma.pro)



# REGRESSION LINE

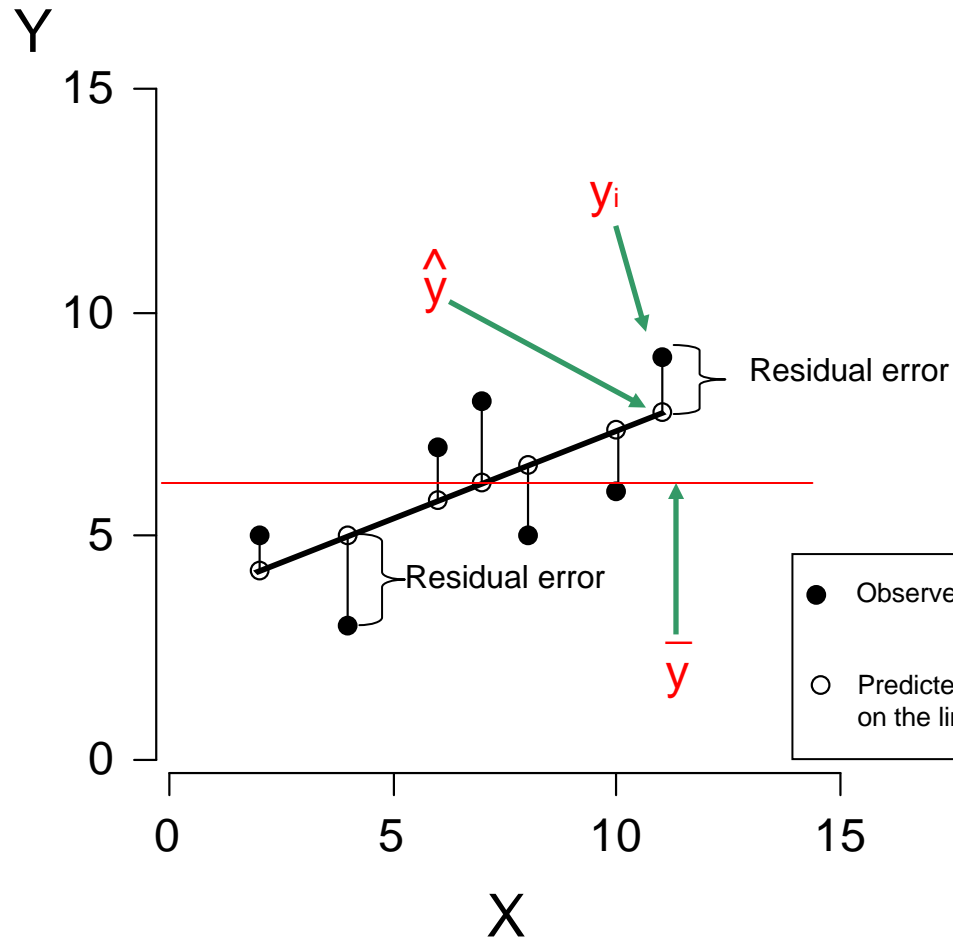


## REGRESSION SUMS OF THE SQUARES



We find the *regression sums of squares*.

## REGRESSION MEAN SQUARE ERROR AND VARIANCE OF Y



$$SST = SSE + SSR$$

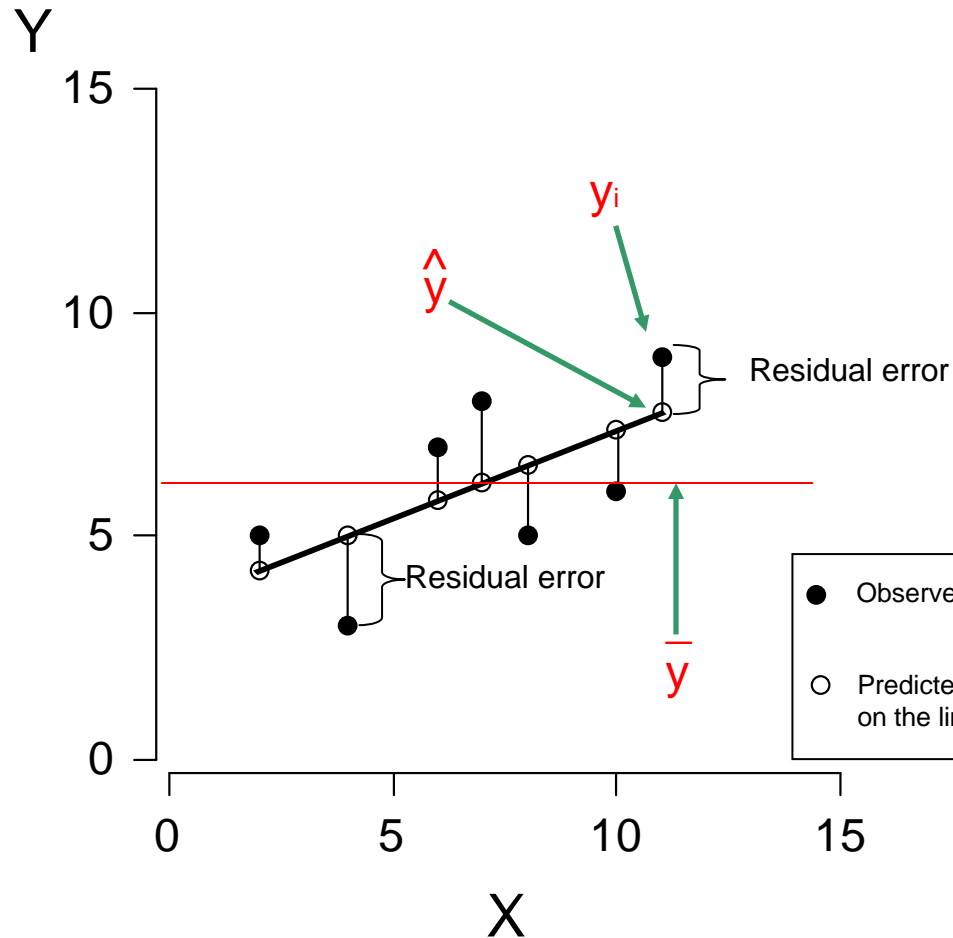
$$MSE = \frac{SSE}{n-2} \quad s_y^2 = \frac{SST}{n-1}$$

$$MSE \ll s_y^2$$

- Observed Y (actual Y)
- Predicted Y (fitted or expected Y—on the line)

To find the **best fitting** straight line, the variance of the error (**Mean Square Error**) about the Regression Line must be **significantly less than the variance of y** about the mean.

# REGRESSION AND R<sup>2</sup>



$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

| REGRESSION | R <sup>2</sup> | SSE = |
|------------|----------------|-------|
| NONE       | 0              | SST   |
| PERFECT    | 1              | 0     |

- Observed Y (actual Y)
- Predicted Y (fitted or expected Y—on the line)

We use R<sup>2</sup> to **check how much of the variability is explained by the regression line.**

$$R = \frac{\Sigma(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\Sigma(X_i - \bar{X})^2 \Sigma(Y_i - \bar{Y})^2}}$$

## REGRESSION F STATISTIC AND HYPOTHESIS TEST

We then use the F statistic with the ANOVA table to find the p value and test the Null Hypothesis

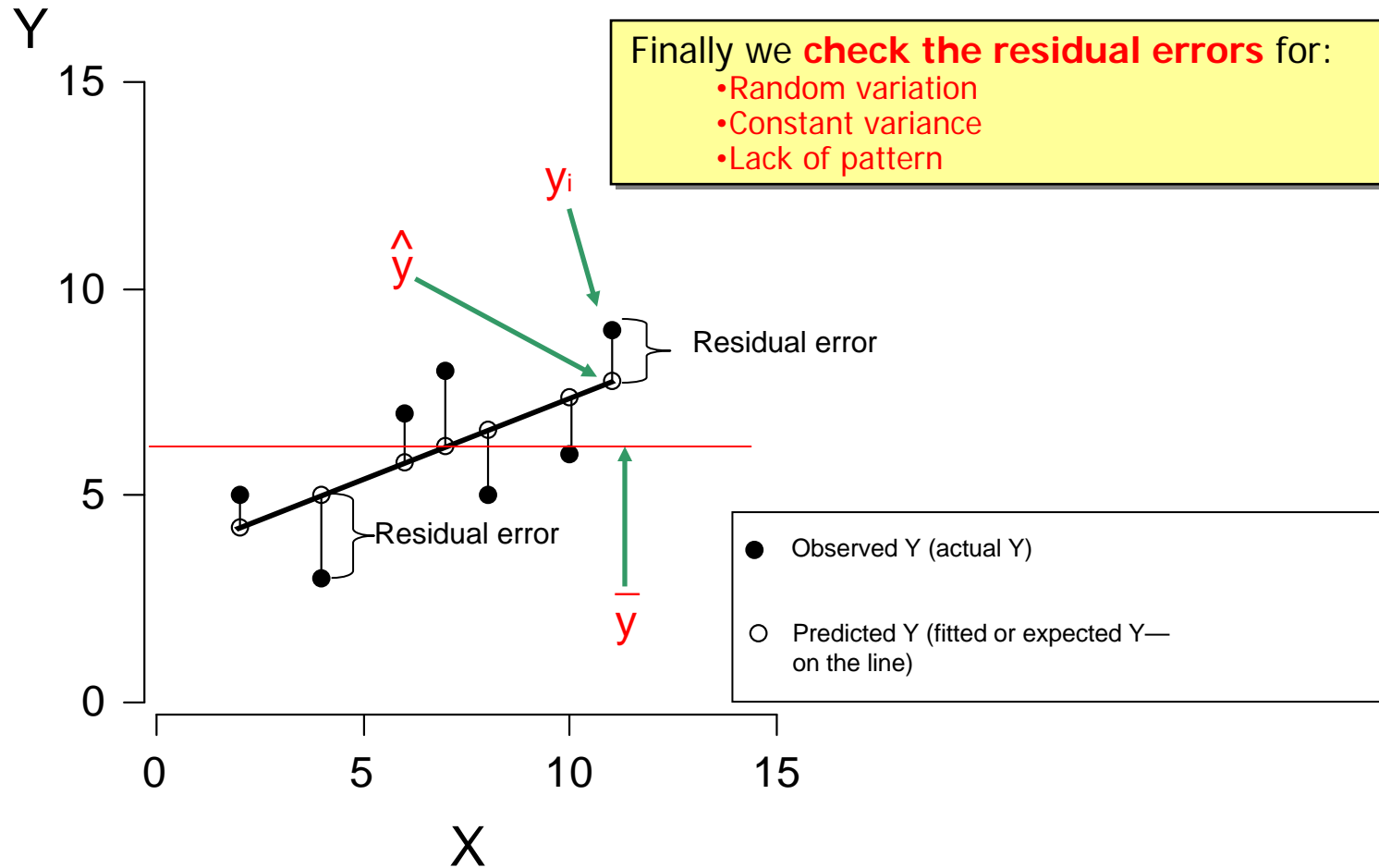
$$\hat{y} = \beta_0 + \beta_1 x$$

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0$$

|            | Sum of Squares                          | df    | Mean Square   | F value |
|------------|---|-------|---------------|---------|
| Regression | SSR=R <sup>2</sup> x SST                | 1     | MSR=SSR/1     | MSR/MSE |
| Error      | SSE=(1-R <sup>2</sup> ) x SST           | (n-2) | MSE=SSE/(n-2) |         |
| Total      | SST=S <sup>2</sup> <sub>y</sub> x (n-1) | (n-1) |               |         |

# REGRESSION



---

This Training Manual and all materials, procedures and systems herein contained or depicted (the "Manual") are the sole and exclusive property of La Rémige S.A.R.L.

The contents hereof contain proprietary information and materials that are the private property of La Rémige S.A.R.L. Unauthorized use, disclosure, or reproduction of any kind of any material contained in this Manual is expressly prohibited. The contents hereof are to be returned immediately upon termination of any relationship or agreement giving user authorization to possess or use such information or materials. Any unauthorized or illegal use shall subject the user to all remedies, both legal and equitable, available to La Rémige S.A.R.L. This Manual may be altered, amended or supplemented by La Rémige S.A.R.L. from time to time. In the event of any inconsistency or conflict between a provision in this Manual and any national, federal, provincial, state or local statute, regulation, order or other law, such law will supersede the conflicting or inconsistent provision(s) of this Manual in all properties subject to that law.

© 2008 by La Rémige S.A.R.L. All Rights Reserved.

*"Lean Six Sigma" is a registered mark with INPI in France*