Simple Linear Regression

Foundation

Dr. Maria Tackett

General form of model

$$Y = f(X) + \epsilon$$

Y: response variable

General form of model

$$Y = f(X) + \epsilon$$

Y: response variable

X: predictor variables

General form of model

$$Y = f(X) + \epsilon$$

Y: response variable

X: predictor variables

f: fixed but unknown function

General form of model

$$Y = f(X) + \epsilon$$

Y: response variable

X: predictor variables

f: fixed but unknown function

ϵ: random error

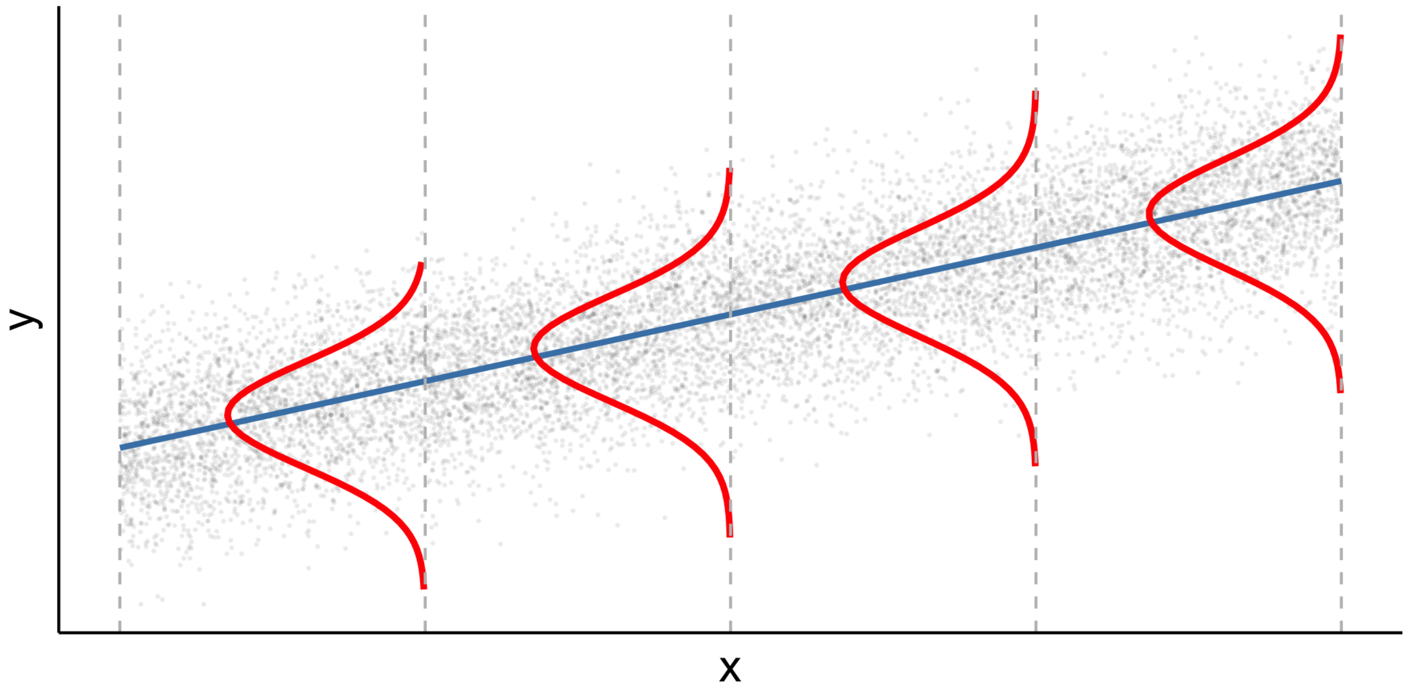
Simple linear regression

$$
\begin{aligned}
Y &= \textbf{Model} + \text{Error} \\
&= f(X) + \epsilon \\
&= \mu_{Y|X} + \epsilon \\
&= \beta_0 + \beta_1 X + \epsilon
\end{aligned}
$$

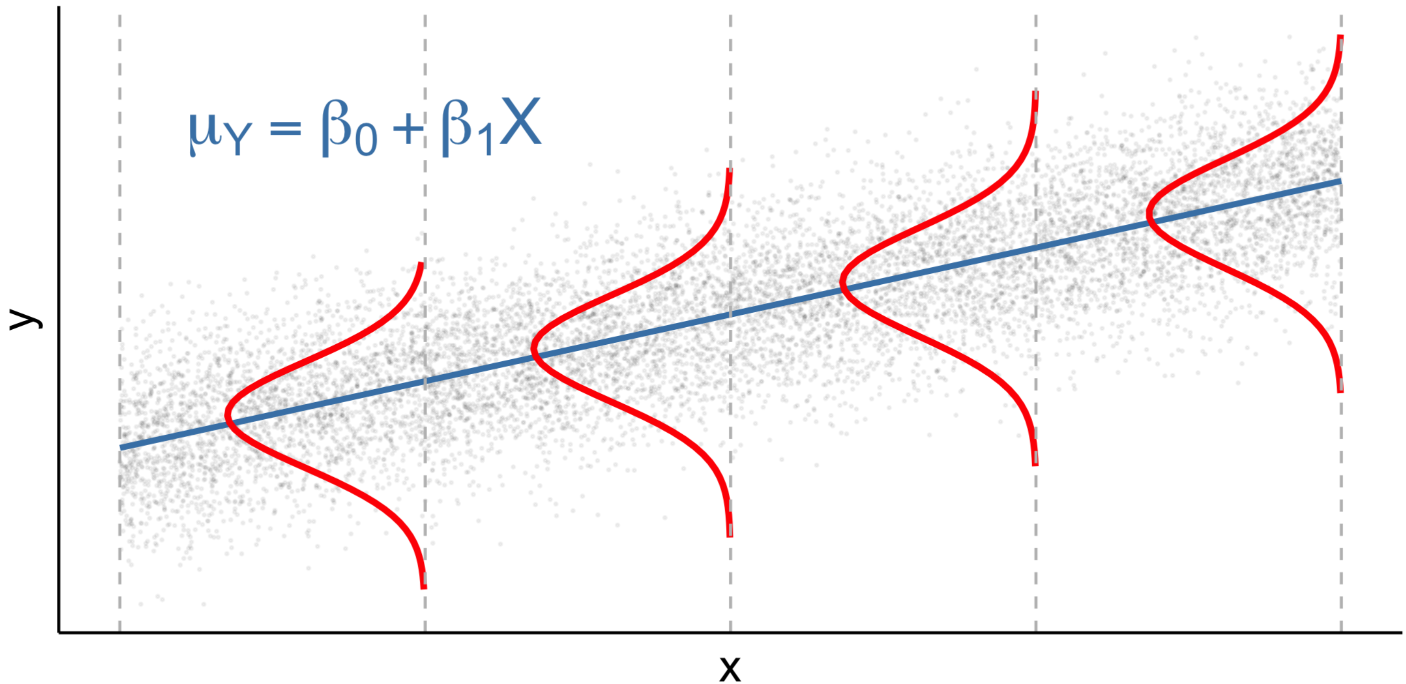
$$Y = \beta_0 + \beta_1 X + \epsilon$$

where the errors are independent and normally distribution, $\epsilon \sim N(0, \sigma_\epsilon)$

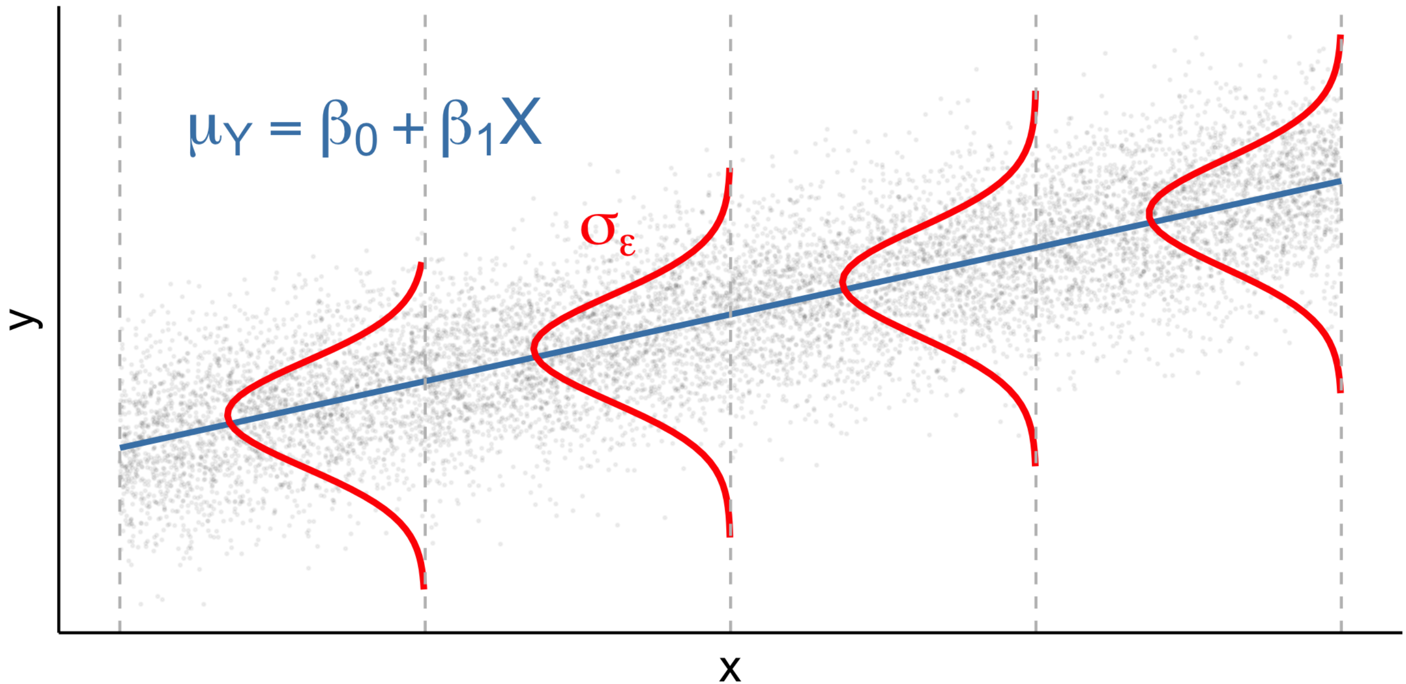
$$Y|X \sim N(\beta_0 + \beta_1 X, \sigma_\epsilon^2)$$



$$Y|X \sim N(\beta_0 + \beta_1 X, \sigma_\epsilon^2)$$



$$Y|X \sim N(\beta_0 + \beta_1 X, \sigma_\epsilon^2)$$



Regression standard error

Once we fit the model, we can use the residuals to estimate the regression standard error, $\hat{\sigma}_\epsilon$

$$\hat{\sigma}_\epsilon = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-2}} = \sqrt{\frac{\sum_{i=1}^{n} e_i^2}{n-2}}$$