



# RM4Es of Simple Linear Regression

The RM Institute  
Pasadena, CA 91801, USA

[www.ResearchMethods.org](http://www.ResearchMethods.org)



## – Equation

- $y = a + \beta x + \epsilon$  is the equation for any simple linear regression. Here,  $y$  is often called as a dependent variable or a response, while  $x$  is often called as an independent variable or a predictor.  $a$  is called as an intercept and  $\beta$  is called as a slope, while  $\epsilon$  is called the error term.  $a$  and  $\beta$  are the equation parameters to be estimated. Adapting this equation assumes the dependent variable  $y$  is linearly related to one and only one independent variable  $x$ .



## – Estimation

- After specifying our equation, we need to use available data to estimate the values of  $a$  and  $\beta$ . The ordinary least squares (OLS) method is the one employed most often, but the maximum likelihood method can also be used. When conducting OLS estimation, parameters  $a$  and  $\beta$  are chosen to minimize a quantity called as the residual sum of squares that is  $\sum [y - (a + \beta x)]^2$ . Under the assumption errors are uncorrelated and have the same variance, the OLS estimate is the best among all linear estimation methods.



## – Errors

- $\epsilon$  is the error term for simple linear regression that is the difference between the predicted values and the actual values of the dependent variable  $y$ . That is,  $\epsilon = y - (a + \beta x)$ .
- Errors can be used to evaluate the goodness of fit of your simple linear regression, and can also be used to diagnose your regression model in order to improve it.



## – Explanation

- $a$ ,  $\beta$  and  $R^2$  are what need to be explained for simple linear regression.
- Here,  $a$ , the intercept, is the value of  $y$  when  $x$  equals to 0. And,  $\beta$ , the slope, is the rate of change in  $y$  for a unit change in  $x$ .  $R^2 = 1 - \text{RSS}/\text{SYY}$  is called as coefficient of determination.  $R$  square tells us how much variability in  $y$  can be explained by our model. Simple linear regression can be represented by a straight line that graph is often used to help explanations.



Thank you!

