



Simple Linear Regression: John McGready



Simple Linear Regression

John McGready
Johns Hopkins University



Section A

Review: The Equation of a Line

The Equation of a Line

- Recall, from algebra, there are two values which uniquely define any line
 - Y-intercept—where the line crosses the y-axis (when $x = 0$)
 - Slope—the “rise over the run”—how much y changes for each one unit change in x

3

The Equation of a Line

- Recall, from algebra, there are two values which uniquely define any line
 - $y = mx + b$
 - $b = y$ -intercept
 - $m = \text{slope}$

4

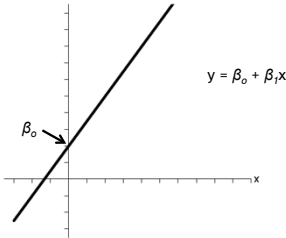
The Equation of a Line

- Of course statisticians must have their own notation!
- $y = b_0 + b_1x$
 - $b_0 = y$ -intercept
 - $b_1 = \text{slope}$
- $y = \beta_0 + \beta_1x$
 - $\beta_0 = y$ -intercept
 - $\beta_1 = \text{slope}$

5

The Intercept, β_0

- The intercept β_0 is the value of y when x is 0
 - It is the point on the graph where the line crosses the y (vertical) axis, at the coordinate $(0, \beta_0)$

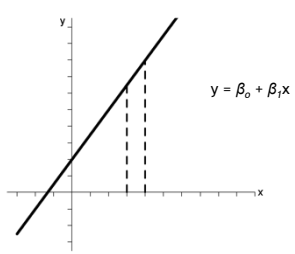


6

Simple Linear Regression: John McGready

The Slope, β_1

- The slope β_1 is the change in y corresponding to a unit increase in x

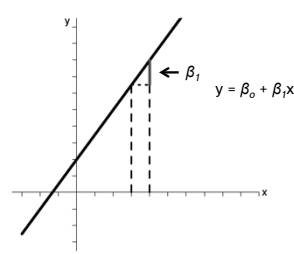


$y = \beta_0 + \beta_1 x$

7

The Slope, β_1

- The slope β_1 is the change in y corresponding to a unit increase in x



$y = \beta_0 + \beta_1 x$

8

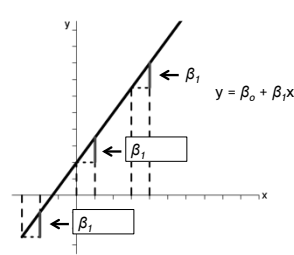
The Slope, β_1

- The slope β_1 is the change in y corresponding to a unit increase in x
- Another interpretation: β_1 is difference in y -values for $x+1$ compared to x
- This change/difference is the same across the entire line

9

The Slope, β_1

- The slope β_1 is the change in y corresponding to a unit increase in x



$y = \beta_0 + \beta_1 x$

10

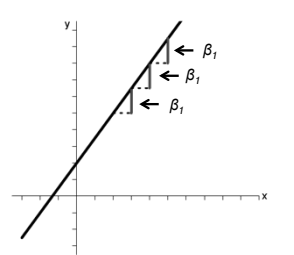
The Slope, β_1

- The slope β_1 is the change in y corresponding to a unit increase in x : β_1 is difference in y -values for $x+1$ compared to x
- All information about the difference in the y -value for two differing values of x is contained in the slope!
- For example: two values of x three units apart will have a difference in y values of $3 * \beta_1$

11

The Slope, β_1

- For example: two values of x three units apart will have a difference in y values of $3 * \beta_1$



12

Simple Linear Regression: John McGready

The Slope, β_1

- For example: two values of x three units apart will have a difference in y values of $3 \times \beta_1$ ($3\beta_1$)

13

The Slope, β_1

- The slope β_1 is the change in y corresponding to a unit increase in x ; β_1 is difference in y -values for $x+1$ compared to x
 - If slope $\beta_1 = 0$, this indicates that there is no association: (i.e., the values of y are the same regardless of the values of x)
 - If slope $\beta_1 > 0$, this indicates that there is a positive association: (i.e., the values of y increase with increasing values of x)
 - If slope $\beta_1 < 0$, this indicates that there is a negative association: (i.e., the values of y decrease with increasing values of x)

14

The Slope, β_1

- The slope β_1 is the change in y corresponding to a unit increase in x ; β_1 is difference in y -values for $x+1$ compared to x

15

The Equation of a Line

- In linear regression situations, points don't fit exactly to a line
- We estimate a line that relates the mean of an outcome y to a predictor x

$$E[y] = \hat{\beta}_0 + \hat{\beta}_1 x$$
 - $E[y]$ = estimated "expected" (mean) value of y
 - $\hat{\beta}_0$ = estimated y -intercept
 - $\hat{\beta}_1$ = estimated slope

16

The Equation of a Line

- $\hat{\beta}_0$ and $\hat{\beta}_1$ are called estimated regression coefficients
- These two quantities are estimated using the data
 - Line estimated is line that "fits the data best"
- Many times the equation is just written as the following:

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

or

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

17

The Equation of a Line

- $\hat{\beta}_0$ and $\hat{\beta}_1$ are called estimated regression coefficients
- We will see that in a regression context, $\hat{\beta}_1$ is nothing more than estimated mean difference in y between two groups who differ by one unit in x
 - i.e., how much the mean of y changes for a one-unit increase in x

18