Statistics 300: Elementary Statistics

Section 10-4

1

For a single Point: (x,y)

Total Deviation $= (y - \bar{y})$

Explained Deviation $= (\hat{y} - \bar{y})$

Unexplained Deviation $= (y - \hat{y})$

$$(y - \bar{y}) = (\hat{y} - \bar{y}) + (y - \hat{y})$$

2

For All Points Together:

Total Variation in y $= \sum (y - \bar{y})^2$

Explained Variation in y $= \sum (\hat{y} - \bar{y})^2$

Unexplained Variation in y $= \sum (y - \hat{y})^2$

3

Important Relationship #1

Total Variation =

Explained Variation

+ Unexplained Variation

$$\sum (y-\bar{y})^2 = \sum (\hat{y}-\bar{y})^2 + \sum (y-\hat{y})^2$$

4

Important Relationship #2

r = correlation coefficient

$$r^2 = \frac{\text{Explained Variation}}{\text{Total Variation}}$$

$$r^2 = \frac{\sum (\hat{y}-\bar{y})^2}{\sum (y-\bar{y})^2}$$

5

$$r^2 = \frac{\sum (\hat{y}-\bar{y})^2}{\sum (y-\bar{y})^2}$$

so squaring the correlation gives the fraction, proportion, or percentage ($r^2 \times 100\%$) of the total variation that can be explained by the line.

6

$$1-r^2=\frac{\sum(y-\hat{y})^2}{\sum(y-\bar{y})^2}=\frac{\text{Unexplained}}{\text{Total}}$$

gives the fraction, proportion, or percentage ($(1-r^2)\times 100\%$) of the total variation that cannot be explained by the line.

7

Useful calculation shortcut

The total variation in $y=$

$$\sum(y-\bar{y})^2=s_y^2(n-1)$$

Practice doing this on your calculator.

8

"x" can be used the same way

The total variation in $x=$

$$\sum(x-\bar{x})^2=s_x^2(n-1)$$

Practice calculating this starting with s_x not s_y

9

Calculating "Explained Variation"

Approach #1:

Explained variation in y =

r^2 (total variation) =

$r^2 \times \sum (y - \bar{y})^2$

10

Calculating "Explained Variation"

Approach #2:

Explained variation in y =

(total) - (unexplained) =

$\sum (y - \bar{y})^2 - \sum (y - \hat{y})^2$

11

Calculating the "Unexplained Variation"

Approach #1:

Unexplained variation in y =

$(1 - r^2)$(total variation) =

$(1 - r^2) \times \sum (y - \bar{y})^2$

12

Calculating the "Unexplained Variation"

Approach #2:

Unexplained variation in y =

(total) - (explained) =

$$\sum (y-\bar{y})^2 - \sum (\hat{y}-\bar{y})^2$$

13

The "Standard Error of Estimate"

$$s_e = \sqrt{\frac{\text{Unexplained variation}}{n-2}}$$

$$s_e = \sqrt{\frac{\sum (y-\hat{y})^2}{n-2}}$$

14

CI($y|x_0$) uses the "Standard Error of Estimate"

$$CI(y \mid x_0) = \hat{y} \pm E$$

$$E = \sqrt{1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{\sum (x-\bar{x})^2}}$$

remember:

$$\sum (x-\bar{x})^2 = s_x^2 (n-1)$$

15