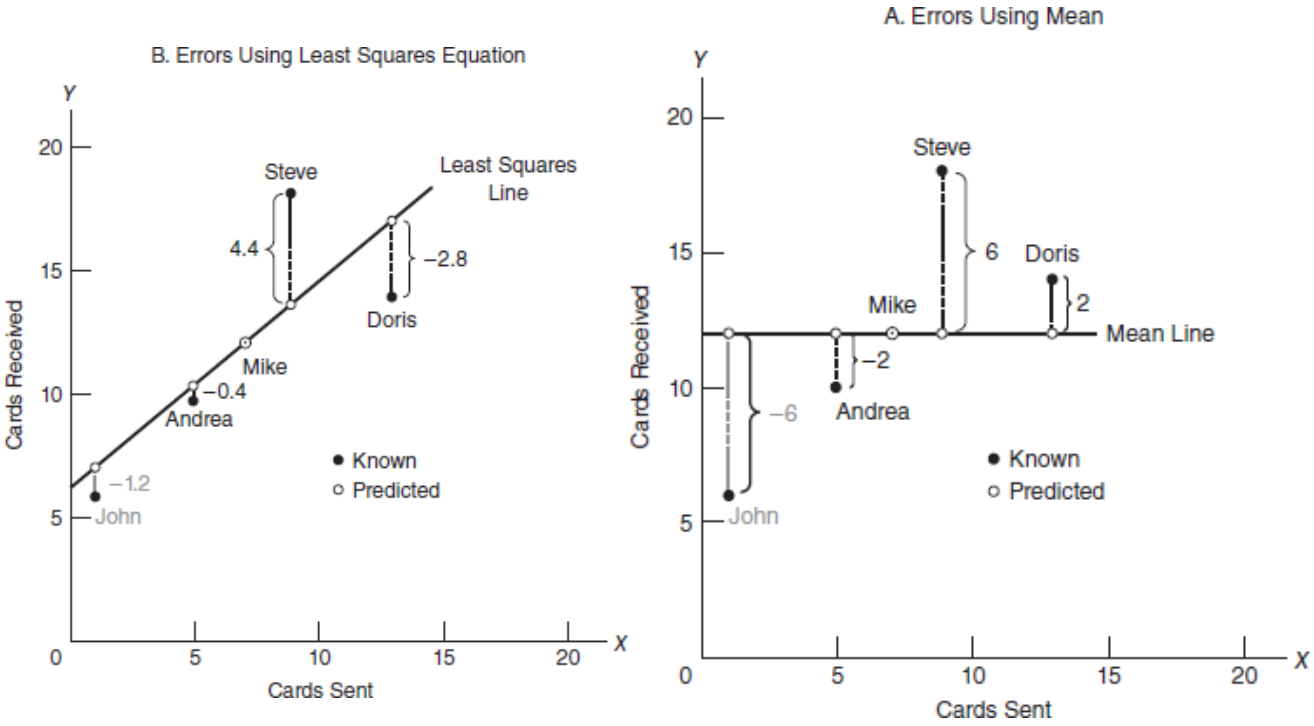
## STANDARD ERROR OF ESTIMATE ,*s y | x*

The standard error of the estimate is a measure of the accuracy of predictions. The regression line is the line that minimizes the sum of squared deviations of prediction (also called the sum of squares error), and the standard error of the estimate is the square root of the average squared deviation. The standard error of estimate and symbolized as ***s y | x***, this estimate of predictive error complies with the general format for any sample standard deviation, that is, the square root of a sum of squares term divided by its degrees of freedom.

$$SS_{y|x} = \Sigma(Y - Y')^2$$




**Example**

Calculate the standard error of estimate for the given X and Y values. X = 1,2,3,4,5 Y=2,4,5,4,5

Solution Create five columns labelled x, y, y', y – y', ( y – y')2 and N=5

| x | y | $x^2$ | xy | Y'= bx+a | y-y' | $( y - y')^2$ |
|---|---|---|---|---|---|---|
| 1 | 2 | 1 | 2 | 2.8 | -0.8 | 0.64 |
| 2 | 4 | 4 | 8 | 3.4 | 0.6 | 0.36 |
| 3 | 5 | 9 | 15 | 4.0 | 1 | 1 |
| 4 | 4 | 16 | 16 | 4.6 | -0.6 | 0.36 |
| 5 | 5 | 25 | 25 | 5.2 | -0.2 | 0.04 |
| **Σx:15** | **Σy:20** | **$Σx^2$:55** | **Σxy:66** | | | **$Σ( y - y')^2$ = 2.4** |

**Note: for finding b value we have to find xy and $x^2$, so add xy and $x^2$ column in table**

$$b = \frac{N\,\Sigma(xy) - \Sigma x\,\Sigma y}{N\,\Sigma(x^2) - (\Sigma x)^2}$$

$$b = \frac{5(66) - 15x20}{5(55) - (15)^2}$$

$$= \frac{330 - 300}{275 - 225}$$

**b= 30/50 = 0.6**

$$a = \frac{\Sigma y - b\,\Sigma x}{N}$$

$$= \frac{20 - (0.6 \times 15)}{5}$$

$$= \frac{20 - 11}{5}$$

**a= 9/5 = 2.2**

**$SSy/x = \sqrt{((y-y')^2 / n-2)}$**

$$= \sqrt{(2.4/3)}$$

**SSy/x** = 0.894

**INTERPRETATION OF *r* 2**

R-Squared ($R^2$ or the coefficient of determination) is a statistical measure in a regression model that determines the proportion of variance in the dependent variable that can be explained by the independent variable. In other words, r-squared shows how well the data fit the regression model (the goodness of fit). R-squared can take any values between 0 to 1. Although the statistical measure provides some useful insights regarding the regression model, the user should not rely only on the measure in the assessment of a statistical model.

In addition, it does not indicate the correctness of the regression model. Therefore, the user should always draw conclusions about the model by analyzing r-squared together with the other variables in a statistical model. The most common interpretation of r-squared is how well the regression model explains observed data.

$$r^2 = \frac{SS_{Y'}}{SS_Y} = \frac{SS_Y - SS_{Y|X}}{SS_Y}$$

$$SS_{y'} = \sum(Y' - \bar{Y})^2$$