WIKIPEDIA

Explained sum of squares

In statistics, the **explained sum of squares (ESS),** alternatively known as the **model sum of squares** or **sum of squares due to regression** (**"SSR"** – not to be confused with the residual sum of squares **RSS**), is a quantity used in describing how well a model, often a regression model, represents the data being modelled. In particular, the explained sum of squares measures how much variation there is in the modelled values and this is compared to the total sum of squares, which measures how much variation there is in the observed data, and to the residual sum of squares, which measures the variation in the modelling errors.

Contents

Definition

The **explained sum of squares (ESS)** is the sum of the squares of the deviations of the predicted values from the mean value of a response variable, in a standard regression model — for example, $y_i = a + b_1x_{1i} + b_2x_{2i} + ... + \varepsilon_i$, where y_i is the i th observation of the response variable, x_{ji} is the i th observation of the j th explanatory variable, a and b_i are coefficients, i indexes the observations from 1 to n, and ε_i is the i th value of the error term. In general, the greater the ESS, the better the estimated model performs.

If $\hat{a}$ and $\hat{b}_i$ are the estimated coefficients, then

$$\hat{y}_i = \hat{a} + \hat{b_1} x_{1i} + \hat{b_2} x_{2i} + \cdots$$

is the i^{th} predicted value of the response variable. The ESS is the sum of the squares of the differences of the predicted values and the mean value of the response variable:

$$\text{ESS} = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2.$$

In some cases (see below): total sum of squares = **explained sum of squares** + residual sum of squares.

Partitioning in simple linear regression

The following equality, stating that the total sum of squares equals the residual sum of squares plus the explained sum of squares, is generally true in simple linear regression:

$$\sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2.$$

Simple derivation

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

Square both sides and sum over all i:

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n} 2(\hat{y}_i - \bar{y})(y_i - \hat{y}_i).$$

Here is how the last term above is zero from simple linear regression[1]

$$\hat{y}_i = \hat{a} + \hat{b}x_i$$
$$\bar{y} = \hat{a} + \hat{b}\bar{x}$$
$$\hat{b} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

So,

$$\hat{y}_i - \bar{y} = \hat{b}(x_i - \bar{x})$$
$$y_i - \hat{y}_i = (y_i - \bar{y}) - (\hat{y}_i - \bar{y}) = (y_i - \bar{y}) - \hat{b}(x_i - \bar{x})$$

Therefore,

$$\sum_{i=1}^{n} 2(\hat{y}_i - \bar{y})(y_i - \hat{y}_i) = 2\hat{b}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \hat{y}_i)$$
$$= 2\hat{b}\sum_{i=1}^{n}(x_i - \bar{x})((y_i - \bar{y}) - \hat{b}(x_i - \bar{x}))$$
$$= 2\hat{b}\left(\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) - \sum_{i=1}^{n}(x_i - \bar{x})^2 \frac{\sum_{j=1}^{n}(x_j - \bar{x})(y_j - \bar{y})}{\sum_{j=1}^{n}(x_j - \bar{x})^2}\right)$$
$$= 2\hat{b}(0) = 0$$

Partitioning in the general ordinary least squares model

The general regression model with n observations and k explanators, the first of which is a constant unit vector whose coefficient is the regression intercept, is

$$y = X\beta + e$$

where y is an $n \times 1$ vector of dependent variable observations, each column of the $n \times k$ matrix X is a vector of observations on one of the k explanators, β is a $k \times 1$ vector of true coefficients, and e is an $n \times 1$ vector of the true underlying errors. The ordinary least squares estimator for β is

$$\hat{\beta} = (X^T X)^{-1} X^T y.$$

The residual vector $\hat{e}$ is $y - X\hat{\beta} = y - X(X^T X)^{-1} X^T y$, so the residual sum of squares $\hat{e}^T \hat{e}$ is, after simplification,

$$RSS = y^T y - y^T X(X^T X)^{-1} X^T y.$$

Denote as $\bar{y}$ the constant vector all of whose elements are the sample mean y_m of the dependent variable values in the vector y. Then the total sum of squares is

$$TSS = (y - \bar{y})^T (y - \bar{y}) = y^T y - 2y^T \bar{y} + \bar{y}^T \bar{y}.$$

The explained sum of squares, defined as the sum of squared deviations of the predicted values from the observed mean of y, is

$$ESS = (\hat{y} - \bar{y})^T(\hat{y} - \bar{y}) = \hat{y}^T\hat{y} - 2\hat{y}^T\bar{y} + \bar{y}^T\bar{y}.$$

Using $\hat{y} = X\hat{\beta}$ in this, and simplifying to obtain $\hat{y}^T\hat{y} = y^TX(X^TX)^{-1}X^Ty$, gives the result that *TSS* = *ESS* + *RSS* if and only if $y^T\bar{y} = \hat{y}^T\bar{y}$. The left side of this is y_m times the sum of the elements of y, and the right side is y_m times the sum of the elements of $\hat{y}$, so the condition is that the sum of the elements of y equals the sum of the elements of $\hat{y}$, or equivalently that the sum of the prediction errors (residuals) $y_i - \hat{y}_i$ is zero. This can be seen to be true by noting the well-known OLS property that the $k \times 1$ vector $X^T\hat{e} = X^T[I - X(X^TX)^{-1}X^T]y = 0$: since the first column of X is a vector of ones, the first element of this vector $X^T\hat{e}$ is the sum of the residuals and is equal to zero. This proves that the condition holds for the result that *TSS* = *ESS* + *RSS*.

In linear algebra terms, we have $RSS = \|y - \hat{y}\|^2$, $TSS = \|y - \bar{y}\|^2$, $ESS = \|\hat{y} - \bar{y}\|^2$. The proof can be simplified by noting that $y^T\hat{y} = \hat{y}^T\hat{y}$. The proof is as follows:

$$\hat{y}^T\hat{y} = y^TX(X^TX)^{-1}X^TX(X^TX)^{-1}X^Ty = y^TX(X^TX)^{-1}X^Ty = y^T\hat{y},$$

Thus,

$$\begin{aligned}
TSS &= \|y - \bar{y}\|^2 = \|y - \hat{y} + \hat{y} - \bar{y}\|^2 \\
TSS &= \|y - \hat{y}\|^2 + \|\hat{y} - \bar{y}\|^2 + 2 < y - \hat{y}, \hat{y} - \bar{y} > \\
TSS &= RSS + ESS + 2y^T\hat{y} - 2\hat{y}^T\hat{y} - 2y^T\bar{y} + 2\hat{y}^T\bar{y} \\
TSS &= RSS + ESS - 2y^T\bar{y} + 2\hat{y}^T\bar{y}
\end{aligned}$$

which again gives the result that *TSS* = *ESS* + *RSS* if and only if $y^T\bar{y} = \hat{y}^T\bar{y}$.

See also

- Sum of squares (statistics)
- Lack-of-fit sum of squares
- Fraction of variance unexplained

Notes

1. Mendenhall, William (2009). *Introduction to Probability and Statistics* (13th ed.). Belmont, CA: Brooks/Cole. p. 507. ISBN 9780495389538.

References

- S. E. Maxwell and H. D. Delaney (1990), "Designing experiments and analyzing data: A model comparison perspective". Wadsworth. pp. 289–290.
- G. A. Milliken and D. E. Johnson (1984), "Analysis of messy data", Vol. I: Designed experiments. Van Nostrand Reinhold. pp. 146–151.

- B. G. Tabachnick and L. S. Fidell (2007), "Experimental design using ANOVA". Duxbury. p. 220.
- B. G. Tabachnick and L. S. Fidell (2007), "Using multivariate statistics", 5th ed. Pearson Education. pp. 217–218.