Week 3 Monday

Valentino Aceves

April 15, 2024

Measure of Fitness

One way to evaluate how well our model works is to look at R^2, i.e. how much of the changes in Y can be explained by our model.

TSS: Total Sum of Squares: total changes in the sample

$$TSS = \sum_{i=1}^{n} (Y_i - \bar{Y})^2$$

ESS: Explained Sum of Squares: changes our model can explain

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

SSR: Sum of Squared Residuals: changes our model cannot explain, or sum of OLS residuals

$$SSR = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{n} \varepsilon_i^2$$

Note that $TSS = ESS + SSR$

Measure of Fitness: how much of the variation in Y can be explained by the regression.

$$R^2 = \frac{ESS}{TSS} = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$
$$= 1 - \frac{SSR}{TSS} = r_{XY}^2 \in [0, 1]$$

R^2 does not measure how good or bad a model is, only how much variance the model explains.

Homework Help

Marketing researchers have determined that there is a relationship between sales of canned tuna and the price of canned tuna. Specifically,

$$SALES = 40,710 - 430PRICE,$$

where $SALES$ are cans sold per week and $PRICE$ is measured in cents per can. Suppose $PRICE$ is (approximately) a normal random variable with mean $\mu = 75$ cents and standard deviation $\sigma = 5$ cents. That is $PRICE \sim N(75, 25).\backslash$

(a) What is the expected value of $SALES$?

(b) What is the variance of $SALES$? Hint: What distribution is $SALES$?

(c) Find the probability that more than $6,300$ cans are sold in a week.