PRESS: PREDICTED ERROR SUM OF SQUARES

It explains us how my model is going to explain the variability i.e accuracy with future observations(New observations)

Yi = i^{th} observation is deleted from the model (Actual observation)
Yi hat = Predicted value of i^{th} observation and i^{th} observation is excluded from model

$$e_i = Y_i - Y_i \text{ hat}$$

Statistical software calculates predicted R-squared using the following procedure:

. It removes a data point from the dataset.
. Calculates the regression equation.
. Evaluates how well the model predicts the missing observation.
. And, repeats this for all data points in the dataset.

Let us work on the Cases, Distance, Delivery time example.

I removed the first observation from the data set and fitted the linear regression with remaining observations only.

DeliveryTime	Cases	Distance
16.68	7	560

Now the predicted model is **Y = 2.53838 + 1.5485 Cases + 0.01589 Distance**

Calculate the Y1 hat value using the above model and will get the value as 22.27797

Y1 hat = 22.27797
Y1 = 16.68
ei= 16.68 - 22.27797 = -5.59797

Similarly we are calculating for all the observations.

⟶ $\mathbf{R^2_{Press}}$ = 0.9206438. That means 92% of variability on new observations. This is nothing but predicted power as well

⟶ We are enough confident to say that our model is going to be work well not only in historical data but also it works well on future data as well.

⟶ That means, it is very good model.

→ R^2_{Press} is better measure than $R^2_{Adj.}$

→ Since R^2_{Adj} measures only on past data not sure of Future data.

→ R^2_{Press} is the predicted measure, So it is more accurate.