TSS

T-TEST

Studying indepently the relation b/w input & and Output

Select a column and perform T-Test

Steps :

1. State the null & alternate hypothesis.

For slope (β_1)

Null hypothesis (H0): $\beta 1 = 0$ (no relation b/w x and Target)
Alternate hypothesis (H1): $\beta 1 \neq 0$ (relation exists)

For intercept

Null hypothesis (H0:) = $\beta 0 = 0$ (no relation ie line passes through origin)

Alternate hypothesis (H1) = $\beta 0 \neq 0$ (doesn't pass through origin)

② fit the regression & sample coef

$b_0, b_1 \rightarrow$ model coef

③ Calculate standard error for b_1 & b_0

$$SE(b_1) = \sqrt{\frac{\Sigma(y_i - \hat{y}_i)^2}{(n-2)\,\Sigma(x_i - \bar{x})^2}}$$

$$SE(b_0) = \sqrt{\frac{\Sigma(y_i - \hat{y}_i)^2}{(n-2)}\left[\frac{1}{n} + \frac{\bar{x}^2}{\Sigma(x_i - \bar{x})^2}\right]}$$

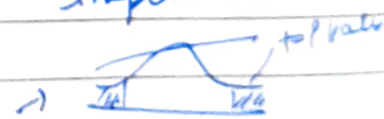
④ Compute T stat

$$T\text{stat}_{b_1} = \frac{b_1 - 0}{SE(b_1)}$$

(hypothetical value for b_1 which is 0)

$$T\text{stat}_{b_0} = \frac{b_0 - 0}{SE(b_0)}$$

To check if the x input is important or not.



Calculate P-value for $T\text{stat}_{b_1}$ (2 tail)

If P-value is less than 0.05 (significance) then we can reject null hypothesis

i.e the statement $\beta_1 = 0$ is false. Hence the associate column is important

If there are multiple columns we will perform T-test and even for intercept

CONFIDENCE interval for coef

Provides a interval for coef

Formula

→ calculated using significance level and df

$$CI_{b0} = b_0 \pm t\text{-value} \times SE(b_0)$$

$$CI_{b1} = b_1 \pm t\text{-value} \times SE(b_1)$$

Example → calculated using significance & degree of freedom

$$CI_{b1} = 0.0458 \pm 3.18 \times (0.001)$$

$$CI_{b1} = 0.0006$$