**Interpretation Results**



Dependent variable :Tip(in dollars)

Independent variable:Totalbill(in dollars)

Number of obseravtions=215

Standard deviation of the observation = 0.73623506

How much tip should be expected on average?

With totalbill= 100

Tip= 0.761727239 + 0.112271819 \*(totalbill)

Therefore, the total tip on average would be 11.9889091

Tip can be calculated based on the numerical variable total bill. Since we got only one variable. our model is simple regression.

**Confidence Interval**

**Findings:**

Tip = b0 + b1(totalbill) +e (where b0 is intercept and b1 is slope)

The confidence Interval is given by where b0 is coefficient and SE is standard error. For 95 % confidence interval level t\* =1.985

We need t\* from t distribution with (n-p-1) degrees of freedom

We have n=215 and p=1 predictor

Df=n-p-1 =215-1-1 =213

CIβ0 = b0  ± t\* (SE b0) = 0.761727239±1.985\* 0.142950328= [1.04548364, 0.477827888]

CIβ1 = b1  ± t\* (SE b1) = 0.112271819±1.985\* 0.007466999 [0.127093811,0.097442359]

**Results:**



With 95% confidence interval β0 lies within the interval[0.47994872, 1.04350576]

With 95% confidence interval β1 lies within the interval[0.09755314, 0.1269905]

**Hypothesis Testing**

Hypothesis testing is used to find is the relation between two variables significant or not

**Findings:**

All tests are two sided

**Step.1:** Formulate HO and Ha

Null hypothesis: There is no significant relation between price and tip (since y=β0 + β1(x))

Alternative hypothesis: There is significant relation between price and tip

HO : βi = 0 vs Ha: βi  ≠ 0

**Step 2**: Compute test statistic

tstatistic  = bi - βi  / SE bi  = bi  / SE bi  (βi = 0)

For β1 = tstatistic  = 0.112271819/0.007466999= 15.03573536

**Step3**: Finding P value from t distribution with df=n-p-1 =217

t. sf (15.03573536, 213)

P= 1.320388305161084e-35=2\*(2 1.320388305161084e-35) = 2.64e-35

**Result:**

****

**Step 4**:

Since we know C=1-α(For 95% confidence interval)

Choose significance value α and make conclusions ( α = 5%)

Variable (Totalbill):

P= 2.64e-35< α =0.05

Reject H0 : β0 =0 in favor of Ha

We have a strong evidence that there is a linear relation between price and tip. Therefore the alternative hypothesis (Ha) says that the regression line is not flat .So relation between price and tip exists and line helps us to predict tip using price.

(Note: Since the intercept does not affect much relationship between price and tip .We are less concerned)

**Analysis of variance for regression model:**

****

**Step:1** We set the null and alternative hypothesis

Null hypothesis: slope is equal to zero

Alternative hypothesis: Slope is not equal to zero

HO : β1 = 0 vs Ha: β1 ≠ 0

**Step2:** Calculate the test statistic

Fstatistic = Regression MS/Residual MS =explained varinace/Unexplained variance 226.073338

Fstatistic= tstatistic2

From Fstatistic we find the P value using t distribution

P= 2.6408E-35

**Step 3:** Pvalue (Significance F)= 2.6408E-35 ≅ 0

**Step 4**: For α=0.05 we reject null hypothesis in favor of Ha.

In a simple regression the Fstatistic does not tell us much. It is same as the t test for the slope