**Jaipuria Institute of Management, Indore**

*Statistics for Management*

End Term Examination, September, 2019

*Note: - Students need to enter data in MS Excel, use Data analysis tool pack and apply relevant test(s) to generate outputs.*

**Answer to Question 1**

Students will have to state Null and Alternate hypothesis; they need to select Level of significance.

H0 : µABS = µNon-ABS H1 : µABS < µNon-ABS (one tail, test), Significance Level (α = 0.05)

|  |  |
| --- | --- |
| t-Test: Paired Two Sample for Means |  |
|  |  |  |
|  | *ABS* | *non-ABS* |
| Mean | 5.4625 | 5.6375 |
| Variance | 1.51125 | 1.95410714 |
| Observations | 8 | 8 |
| Pearson Correlation | 0.993511614 |  |
| Hypothesized Mean Difference | 0 |  |
| df | 7 |  |
| t Stat | -2.19795039 |  |
| P(T<=t) one-tail | 0.031962061 |  |
| t Critical one-tail | 1.894578605 |  |
| P(T<=t) two-tail | 0.063924123 |  |
| t Critical two-tail | 2.364624252 |   |

Student need to compare t stat with t critical or compare p with α

Rejection region: or P (0.03196) < α. (0.05)

So, H0 is rejected and H1 is accepted. There is enough evidence to infer that ABS is better.

**Answer to Question 2**

Students will have to identify dependent and independent variables for regression model, state null and alternate hypothesis and need to select Level of significance.

Y = a + b1X1 + b2 X2 where,

Y = Sales, X1 is Advertisement Expenses and X2 = Event Sponsorship Expenses

H0: b1 = 0, b2 = 0, H1: b1 ≠ 0, b2 ≠ 0, α = 0.05

Students need to put the data in MS Excel, use Data Analysis tool pack and generate following output.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT - 1 |  |  |  |  |  |  |
| *Regression Statistics* |  |  |  |  |  |
| Multiple R | 0.968647 |  |  |  |  |  |
| R Square | 0.938277 |  |  |  |  |  |
| Adjusted R Square | 0.907415 |  |  |  |  |  |
| Standard Error | 2.577457 |  |  |  |  |  |
| Observations | 7 |  |  |  |  |  |
| ANOVA |  |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |
| Regression | 2 | 403.9468576 | 201.9734 | 30.40264122 | 0.003809774 |  |
| Residual | 4 | 26.57314242 | 6.643286 |  |  |  |
| Total | 6 | 430.52 |   |   |   |  |
|  |  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* |
| Intercept | -1.8934 | 3.764688142 | -0.50294 | 0.641439604 | -12.34584525 | 8.559054678 |
| Advertisement Expenses (Lacs) | 0.484978 | 0.498286963 | 0.97329 | 0.385521884 | -0.898488648 | 1.868444153 |
| Event Sponsorship (Lacs) | 1.805722 | 0.330131994 | 5.469697 | 0.005435366 | 0.889128638 | 2.722315353 |

Student need to analyze this output (i.e. compare p values for two tail test with chosen significance level 0.05) and infer that

Independent variable Advertisement Expenses do not have a significant relationship with sales, whereas, the second independent variable Event Sponsorship has a significant relationship with sales.

Students need to generate a modified output after discarding Advertisement Expenses as under: -

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT - 2 |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| *Regression Statistics* |  |  |  |  |  |
| Multiple R | 0.96107185 |  |  |  |  |  |
| R Square | 0.9236591 |  |  |  |  |  |
| Adjusted R Square | 0.908390921 |  |  |  |  |  |
| Standard Error | 2.563836347 |  |  |  |  |  |
| Observations | 7 |  |  |  |  |  |
|  |  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |
| Regression | 1 | 397.6537159 | 397.6537159 | 60.49569142 | 0.000562336 |  |
| Residual | 5 | 32.86628408 | 6.573256815 |  |  |  |
| Total | 6 | 430.52 |   |   |   |  |
|  |  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* |
| Intercept | 0.769843768 | 2.571949335 | 0.299323069 | 0.776736253 | -5.841562475 | 7.38125 |
| Event Sponsorship (Lacs) | 2.004808248 | 0.257757089 | 7.777897622 | 0.000562336 | 1.342222557 | 2.667394 |

Event sponsorship has a significant relationship with Sales.

Generate Regression or Estimation equation as under: -

 (Sales) = 0.7698 + 2.004 (Event Sponsorship)

Students need to comment of nature of relationship and strength of relationship.

Student need to comment on Adjusted r2 = 0.9083 and F value

The expected sale for 18.2 Lacs investment in Event Sponsorship expenses is

 (Sales) = 0.7698 + 2.004 x 18.2

 = 37.2426 Cr.

Estimated range for 95 % CI is

Estimated Sales = 37.2426 ± 1.96 x 2.5638 with 95 % Confidence Interval.

**Answer to Question 3**

a)



**EXCEL SOLUTION:**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| *Regression Statistics* |  |  |  |  |  |  |
| Multiple R | 0.876196547 |  |  |  |  |  |  |
| R Square | 0.767720389 |  |  |  |  |  |  |
| Adjusted R Square | 0.749852726 |  |  |  |  |  |  |
| Standard Error | 3.825235201 |  |  |  |  |  |  |
| Observations | 15 |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |  |
| Regression | 1 | 628.7118169 | 628.7118 | 42.96703 | 1.84E-05 |  |  |
| Residual | 13 | 190.2215165 | 14.63242 |  |  |  |  |
| Total | 14 | 818.9333333 |   |   |   |  |  |
|  |  |  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *Lower 95.0%* |
| Intercept | -24.70901151 | 4.748056533 | -5.20403 | 0.00017 | -34.9666 | -14.4515 | -34.9666 |
| Television | 0.9674474 | 0.147590938 | 6.554924 | 1.84E-05 | 0.648597 | 1.286298 | 0.648597 |

The sample regression line is

 (Overweight) = – 24.72 + 0.9675 X (Television)

Since p < α, Television has a significant relationship with Overweight.

Students need to comment of nature of relationship and strength of relationship. The slope coefficient indicates that for each additional hour of television, weight increases on average by 0.9675 pounds.

Student need to comment on Adjusted r2 = 0.7498 , and F value.

i.e. 74.98 % variation in dependent variable Y (Overweight) is explained by independent variable X (i.e. television)

Prediction Overweight for 30 hours of Television is as under :-

Estimated Overweight = (-24.72 + 0.9675 x 30) ± 1.645 x 3.82 i.e.

 = 4.305 ± 1.645 x 3.82 with 90 % Confidence Interval

**Answer to Question 4**

|  |
| --- |
| The test statistic is F = 4.06 whereas F critical = 2.866266 and the p-value = 0.0139 is less than α = 0.05. There is enough statistical evidence to infer that there are differences between some of the bumpers. Anova: Single Factor |
|  |  |  |  |  |  |  |
| SUMMARY |  |  |  |  |  |  |
| *Groups* | *Count* | *Sum* | *Average* | *Variance* |  |  |
| Bumper 1 | 10 | 3800 | 380 | 16924.2222 |  |  |
| Bumper 2 | 10 | 4859 | 485.9 | 8197.43333 |  |  |
| Bumper 3 | 10 | 4838 | 483.8 | 10426.1778 |  |  |
| Bumper 4 | 10 | 3482 | 348.2 | 14048.6222 |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |  |
| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Between Groups | 150883.875 | 3 | 50294.625 | 4.05630801 | 0.013947 | 2.866266 |
| Within Groups | 446368.1 | 36 | 12399.11389 |  |  |  |
|  |  |  |  |  |  |  |
| Total | 597251.975 | 39 |   |   |   |   |
|  |  |  |  |  |  |  |

The question is now; which bumpers differ?

|  |  |
| --- | --- |
| t-Test: Two-Sample Assuming Equal Variances |  |
|  |  |  |
|  | *Bumper 1* | *Bumper 2* |
| Mean | 380 | 485.9 |
| Variance | 16924.22222 | 8197.433333 |
| Observations | 10 | 10 |
| Pooled Variance | 12560.82778 |  |
| Hypothesized Mean Difference | 0 |  |
| df | 18 |  |
| t Stat | -2.112865403 |  |
| P(T<=t) one-tail | 0.024421145 |  |
| t Critical one-tail | 1.734063607 |  |
| P(T<=t) two-tail | 0.048842289 |  |
| t Critical two-tail | 2.10092204 |   |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
| t-Test: Two-Sample Assuming Equal Variances |  |
|  |  |  |
|  | *Bumper 1* | *Bumper 3* |
| Mean | 380 | 483.8 |
| Variance | 16924.22222 | 10426.17778 |
| Observations | 10 | 10 |
| Pooled Variance | 13675.2 |  |
| Hypothesized Mean Difference | 0 |  |
| df | 18 |  |
| t Stat | -1.984794338 |  |
| P(T<=t) one-tail | 0.031311612 |  |
| t Critical one-tail | 1.734063607 |  |
| P(T<=t) two-tail | 0.062623223 |  |
| t Critical two-tail | 2.10092204 |   |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

|  |  |
| --- | --- |
| t-Test: Two-Sample Assuming Equal Variances |  |
|  |  |  |
|  | *Bumper 1* | *Bumper 4* |
| Mean | 380 | 348.2 |
| Variance | 16924.22222 | 14048.62222 |
| Observations | 10 | 10 |
| Pooled Variance | 15486.42222 |  |
| Hypothesized Mean Difference | 0 |  |
| df | 18 |  |
| t Stat | 0.571395182 |  |
| P(T<=t) one-tail | 0.287398955 |  |
| t Critical one-tail | 1.734063607 |  |
| P(T<=t) two-tail | 0.574797909 |  |
| t Critical two-tail | 2.10092204 |   |
|  |  |  |
|  |  |  |
|  |  |  |

|  |  |
| --- | --- |
| t-Test: Two-Sample Assuming Equal Variances |  |
|  |  |  |
|  | *Bumper 2* | *Bumper 3* |
| Mean | 485.9 | 483.8 |
| Variance | 8197.433333 | 10426.17778 |
| Observations | 10 | 10 |
| Pooled Variance | 9311.805556 |  |
| Hypothesized Mean Difference | 0 |  |
| df | 18 |  |
| t Stat | 0.048661708 |  |
| P(T<=t) one-tail | 0.480862388 |  |
| t Critical one-tail | 1.734063607 |  |
| P(T<=t) two-tail | 0.961724775 |  |
| t Critical two-tail | 2.10092204 |   |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

|  |  |
| --- | --- |
| t-Test: Two-Sample Assuming Equal Variances |  |
|  |  |  |  |
|  | *Bumper 3* | *Bumper 4* |  |
| Mean | 483.8 | 348.2 |  |
| Variance | 10426.17778 | 14048.62222 |  |
| Observations | 10 | 10 |  |
| Pooled Variance | 12237.4 |  |  |
| Hypothesized Mean Difference | 0 |  |  |
| df | 18 |  |  |
| t Stat | 2.740943692 |  |  |
| P(T<=t) one-tail | 0.006714615 |  |  |
| t Critical one-tail | 1.734063607 |  |  |
| P(T<=t) two-tail | 0.013429229 |  |  |
| t Critical two-tail | 2.10092204 |   |  |
| There is enough evidence to show that difference exist between Bumper 3 and 4, Bumper 2 and 4, and Bumper 1 and 2.  |

|  |  |
| --- | --- |
| t-Test: Two-Sample Assuming Equal Variances |  |
|  |  |  |
|  | *Bumper 2* | *Bumper 4* |
| Mean | 485.9 | 348.2 |
| Variance | 8197.433333 | 14048.62222 |
| Observations | 10 | 10 |
| Pooled Variance | 11123.02778 |  |
| Hypothesized Mean Difference | 0 |  |
| df | 18 |  |
| t Stat | 2.919492955 |  |
| P(T<=t) one-tail | 0.0045758 |  |
| t Critical one-tail | 1.734063607 |  |
| P(T<=t) two-tail | 0.0091516 |  |
| t Critical two-tail | 2.10092204 |   |