**Practical No. 8 :** Kruskal-Wallis H test

A study was conducted to see which of the four colors attracts cereal leaf beetles. Here are the data:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Color | Insects Trapped | | | | | | |
| Lemon yellow | 45 | 59 | 48 | 46 | 38 | 47 | 52 |
| White | 21 | 12 | 14 | 17 | 13 | 17 | 16 |
| Green | 37 | 32 | 15 | 25 | 39 | 41 | 35 |
| Blue | 16 | 11 | 20 | 21 | 14 | 7 | 10 |

Enter this data in Minitab and generate the following reports:

**Questions :**

1. What hypotheses does Kruskal-Wallis test?
2. Find the median number of beetles trapped by boards of each color. Which colors appear more effective?
3. Use the Kruskal Wallis test to see if there are significant differences among the colors. What do you conclude?

**Solution :**

**Step 1 :** Type your data into the data pane of a worksheet. Make sure you put your data into columns. Use column header for “Insect Trapped” and “Color”. Type the “Insect Trapped” data into column C1 and “Color” data into column C2-T.

**Step 2 :** To perform Kruskal-Wallis H test, under the drop-down menu “Stat”, choose “Nonparametrics” then “Kruskal-Wallis…”. A “Kruskal-Wallis” dialogue box will appear. Set the “Response:” as “C1 Insect Trapped” and “Factor:” as “C2 Color” from the table on the left. Click “OK”.

**Step 3 :** For the descriptive statistics, under the drop-down menu “Stat”, choose “Basic Statistics” then “Display Descriptive Statistics…”. A “Display Descriptive Statistics” dialogue box will appear. In the “Variables:” box, choose “C1 Insect Trapped” and in the “By variable (optional):“ box, choose “C2 Color” from the table on the left. Click the “Statistics…” option. A “Display Descriptive Statistics: Statistics” dialogue box will appear. Check the “Mean”, “Minimum”, “Maximum”, “First quartile”, “Median” and “Third quartile” checkboxes. Click “OK”.

**Step 4 :** Click the “Graphs…” option. Click “OK”. Check the “Boxplot of data” option. Click “OK”. Click “OK” again. The following boxplot will be generated.

A graph with blue boxes and text

AI-generated content may be incorrect.

Fig 1 : Boxplot of Insect Trapped

**Interpretation :**

From the boxplot, it is seen that in terms of variability, the most variable color category is green, and the least variable color category is white. In terms of shape, the color blue is almost symmetrical, while the color blue is highly left skewed. The color lemon yellow is right skewed, and color white is somewhat right skewed.

**Descriptive Statistics :**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Color** | **Mean** | **Minimum** | **Q1** | **Median** | **Q3** | **Maximum** |
| Insect Trapped | Blue | 14.14 | 7.00 | 10.00 | 14.00 | 20.00 | 21.00 |
| Green | 32.00 | 15.00 | 25.00 | 35.00 | 39.00 | 41.00 |
| Lemon yellow | 47.86 | 38.00 | 45.00 | 47.00 | 52.00 | 59.00 |
| White | 15.71 | 12.00 | 13.00 | 16.00 | 17.00 | 21.00 |

**Interpretation :**

The cereal leaf beetles are most attracted to the color lemon yellow (median = 47), the next to which is the color is green (median = 35). The least attractive colors are blue (median = 14) and white (median = 16). Clearly there are three groups or clusters: lemon yellow (most popular), green (moderately popular) and blue and white (least popular).

**Descriptive Statistics :**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Color** | **N** | **Median** | **Mean Rank** | **Z-Value** | |
| Blue | 7 | 14 | 7.1 | -2.76 |
| Green | 7 | 35 | 17.3 | 1.03 |
| Lemon yellow | 7 | 47 | 24.7 | 3.79 |
| White | 7 | 16 | 8.9 | -2.07 |
| Overall | 28 |  | 14.5 |  |

**Hypothesis :**

|  |  |
| --- | --- |
| Null hypothesis | H₀: All medians are equal |
| Alternative hypothesis | H₁: At least one median is different |

**Test :**

|  |  |  |  |
| --- | --- | --- | --- |
| **Method** | **DF** | **H-Value** | **p-value** |
| Not adjusted for ties | 3 | 20.52 | 0.000 |
| Adjusted for ties | 3 | 20.54 | 0.000 |

**Conclusion :**

The Kruskal-Wallis H test is a non-parametric statistical test that examines whether there are significant differences between the medians of three or more independent groups. Since the p-value of the test (0.0000) is much smaller than significance probability (0.05), we strongly reject the null hypothesis at 5 % level of significance. The median no. of insects trapped using different colors is not the same i.e. there is a significant difference in the median no. of insects trapped using various colors.

**Worksheet :**

**A table with numbers and letters

AI-generated content may be incorrect.**