How can I utilize Fisher’s Least Significant Difference (LSD) in R for statistical analysis?

Authored by

stats writer

June 28, 2024

Fisher's Least Significant Difference (LSD) is a post-hoc statistical test used to compare the means of multiple groups in a dataset. In R, this method can be utilized by performing an ANOVA (Analysis of Variance) test to determine if there is a significant difference between the groups. If the ANOVA results show a significant difference, the LSD test can then be applied to determine which specific groups have significantly different means. This can be done by using the 'TukeyHSD' function in R, which provides a table of all pairwise comparisons between the groups. By utilizing Fisher's LSD in R, researchers can gain a better understanding of the differences between multiple groups in their data and make more accurate conclusions from their statistical analysis.

Use Fisher's Least Significant Difference (LSD) in R

A is used to determine whether or not there is a statistically significant difference between the means of three or more independent groups.

The used in a one-way ANOVA are as follows:

H0: The means are equal for each group.HA: At least one of the means is different from the others.

If the from the ANOVA is less than some significance level (like α = .05), we can reject the null hypothesis and conclude that at least one of the group means is different from the others.

But in order to find out exactly which groups are different from each other, we must conduct a post-hoc test.

One commonly used post-hoc test is Fisher's least significant difference (LSD) test.

You can use the LSD.test() function from the agricolae package to perform this test in R.

The following example shows how to use this function in practice.

Example: Fisher's LSD Test in R

Suppose a professor wants to know whether or not three different studying techniques lead to different exam scores among students.

To test this, she randomly assigns 10 students to use each studying technique and records their exam scores.

The following table shows the exam scores for each student based on the studying technique they used:

Technique 1	Technique 2	Technique 3
72	81	77
73	82	78
73	83	79
77	83	88
82	83	89
82	84	90
83	87	91
84	90	95
85	92	95
89	93	98

We can use the following code to create this dataset and perform a one-way ANOVA on it in R:

```
#create data frame
df <- data.frame(technique = rep(c("tech1", "tech2", "tech3"), each = 10),
score = c(72, 73, 73, 77, 82, 82, 83, 84, 85, 89,
81, 82, 83, 83, 83, 84, 87, 90, 92, 93,
77, 78, 79, 88, 89, 90, 91, 95, 95, 98))

#view first six rows of data frame
head(df)

technique score
1 tech1 72
2 tech1 73
```

3 tech1 73

4 tech1 77

5 tech1 82

6 tech1 82

#fit one-way ANOVA

model <- aov(score ~ technique, data = df)

#view summary of one-way ANOVA

summary(model)

Df Sum Sq Mean Sq F value Pr(>F)

technique 2 341.6 170.80 4.623 0.0188 *

Residuals 27 997.6 36.95

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Since the p-value in the ANOVA table (.0188) is less than .05, we can conclude that not all of the mean exam scores between the three groups are equal.

Thus, we can proceed to perform Fisher's LSD test to determine which group means are different.

library(agricolae)

```
#perform Fisher's LSD
print(LSD.test(model,"technique"))

$statistics
MSerror Df Mean CV t.value LSD
36.94815 27 84.6 7.184987 2.051831 5.57767

$parameters
test p.ajusted name.t ntr alpha
Fisher-LSD none technique 3 0.05

$means
score std r LCL UCL Min Max Q25 Q50 Q75
tech1 80.0 5.868939 10 76.05599 83.94401 72 89 74.00 82.0 83.75
tech2 85.8 4.391912 10 81.85599 89.74401 81 93 83.00 83.5 89.25
tech3 88.0 7.557189 10 84.05599 91.94401 77 98 81.25 89.5 94.00

$comparison
NULL

$groups
score groups
```

tech3 88.0 a
tech2 85.8 a
tech1 80.0 b

attr(,"class")
"group"

The portion of the output that we're most interested in is the section titled $groups. The techniques that have different characters in the groups column are significantly different.

From the output we can see:

Technique 1 and Technique 3 have significantly different mean exam scores (since tech1 has a value of "b" and tech3 has a value of "a")Technique 1 and Technique 2 have significantly different mean exam scores (since tech1 has a value of "b" and tech2 has a value of "a")Technique 2 and Technique 3 do not have significantly different mean exam scores (since they both have a value of "a")

Additional Resources

The following tutorials explain how to perform other

common tasks in R: