

How can I use the LSMEANS statement in SAS? Can you provide an example?

Authored by
stats writer

June 23, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I use the LSMEANS statement in SAS? Can you provide an example?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=148155>

The LSMEANS statement in SAS allows users to calculate least squares means for a given model and perform multiple comparisons between the means. This statement is useful for analyzing and interpreting the results of a statistical model. An example of using the LSMEANS statement would be in a study comparing the effectiveness of three different medications for treating a specific condition. The statement can be used to calculate the average effectiveness of each medication and determine if there is a significant difference between them. This can provide valuable insights for decision making and further research.

Use LSMEANS Statement in SAS (With Example)

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

If the overall from the ANOVA table is less than some significance level, then we have sufficient evidence to say that at least one of the means of the groups is different from the others.

To find out exactly which group means are different, we must conduct a post hoc test.

You can use the LSMEANS statement in SAS to perform a variety of post-hoc tests.

The following example shows how to use the LSMEANS statement in practice.

Example: How to Use LSMEANS Statement in SAS

Suppose a researcher recruits 30 students to participate in a study. The students are to use one of three studying methods to prepare for an exam.

The exam results for each student are shown below:

Method A	Method B	Method C
78	81	84
81	83	88
82	83	88
82	85	89
85	86	90
88	88	93
88	90	95
90	91	98

We can use the following code to create this dataset in SAS:

```
/*create dataset*/  
data my_data;  
input Method $ Score;  
datalines;  
A 78  
A 81
```

A 82

A 82

A 85

A 88

A 88

A 90

B 81

B 83

B 83

B 85

B 86

B 88

B 90

B 91

C 84

C 88

C 88

C 89

C 90

C 93

C 95

C 98

;

run;

Next, we'll use proc ANOVA to perform the one-way ANOVA:

```
/*perform one-way ANOVA*/  
proc ANOVAdata=my_data;  
class Method;  
model Score = Method;  
run;
```

This produces the following ANOVA table:

The ANOVA Procedure
Dependent Variable: Score

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	175.5833333	87.7916667	5.26	0.0140
Error	21	350.2500000	16.6785714		
Corrected Total	23	525.8333333			

R-Square	Coeff Var	Root MSE	Score Mean
0.333914	4.698685	4.083941	86.91667

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Method	2	175.5833333	87.7916667	5.26	0.0140

From this table we can see:

The overall F Value: 5.26
The corresponding p-value: 0.0140
H₀: All group means are equal.
H_A: At least one

group mean is different from the rest.

Since the p-value from the ANOVA table (0.0140) is less than $\alpha = .05$, we reject the null hypothesis.

This tells us that the mean exam score is not equal between the three studying methods.

To determine exactly which group means are different, we can use the PROC GLIMMIX statement along with the LSMEANS statement and the option ADJUST=TUKEY to perform Tukey's post hoc tests:

```
/*perform Tukey post-hoc comparisons*/  
proc glimmixdata=my_data;  
class Method;  
model Score = Method;  
lsmeans Method / adjust=tukey alpha=.05;  
run;
```

The last table in the output shows the results of the Tukey post-hoc comparisons:

Differences of Method Least Squares Means Adjustment for Multiple Comparisons: Tukey												
Method	_Method	Estimate	Standard Error	DF	t Value	Pr > t	Adj P	Alpha	Lower	Upper	Adj Lower	Adj Upper
A	B	-1.6250	2.0420	21	-0.80	0.4351	0.7096	0.05	-5.8715	2.6215	-6.7719	3.5219
A	C	-6.3750	2.0420	21	-3.12	0.0052	0.0137	0.05	-10.6215	-2.1285	-11.5219	-1.2281
B	C	-4.7500	2.0420	21	-2.33	0.0301	0.0739	0.05	-8.9965	-0.5035	-9.8969	0.3969

We can look at the Adj P column to view the adjusted p-values for the difference in group means.

From this column we can see that there is only one row with an adjusted p-value less than .05: the row that compares the mean difference between group A and group C.

This tells us there is a statistically significant difference in mean exam scores between group A and group C.

Specifically, we can see:

The difference in mean exam scores of students in group A - students in group B was -6.375. (i.e. students in group A had an average exam score of 6.375 points less than students in group C) The adjusted p-value for the difference in means is 0.0137. The adjusted 95% confidence interval for the true difference in mean exam scores between these two groups is .

There are no statistically significant differences between any other group means.

Note: In this example we used ADJUST=TUKEY to perform Tukey post-hoc comparisons but you can also specify BON, BUNNET, NELSON, SCHEFFE, SIDAK, and SMM to perform other types of post-hoc comparisons.

The following tutorials provide additional information about ANOVA models:

[A Guide to Using Post Hoc Tests with ANOVA](#)