Print

Quiz, Lesson 3: Analysis of Variance

Your Score: 100%

Congratulations! Your score of 100% indicates that you've mastered the topics in this lesson. If you'd like, you can review the feedback for each question.

When you're ready to start the next lesson, exit this lesson and begin the next one.

1. Which program correctly performs multiple comparisons on the interactions?

a.
```
proc glm data=sasuser.disks;
   class technician brand;
   model time = technician | brand;
   lsmeans technician * brand / pdiff adjust=tukey cl;
run;
quit;
```

b.
```
proc glm data= sasuser.disks;
   class technician brand;
   model time = technician | brand;
   means technician * brand / pdiff adjust=tukey cl;
run;
quit;
```

c.
```
proc glm data= sasuser.disks;
   class technician brand;
   model time = technician  brand;
   lsmeans technician * brand / pdiff  adjust=tukey cl;
run;
quit;
```

d.
```
proc glm data= sasuser.disks;
   class technician brand;
   model time = technician brand;
   means technician * brand / pdiff  adjust=tukey cl;
run;
quit;
```

Your answer: a
Correct answer: a

The MEANS statement performs multiple comparisons only for main-effect means. In contrast to the MEANS statement, the LSMEANS statement performs multiple comparisons on interactions as well as main effects. Also, for the program to run successfully, it is important that the effects used in the MEANS / LSMEANS statement must have appeared previously in the MODEL statement. Notice that the interaction term is not in the model for choices c and d.

2. Below is the partial output from PROC GLM while comparing the average service time for technician Bob for servicing Brand 2 with the average service time for technician Justin for servicing Brand 2. What does the result indicate?

Least Squares Means Estimate						
Effect	**Label**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Technician*Brand	**Bob Brand 2 vs Justin Brand 2**	27.1250	6.5040	84	4.17	<.0001

a. The average servicing time is not significantly different for these two technicians.

b. The average servicing time is significantly different for these two technicians, and Bob takes a little more than 27 minutes longer on average than Justin to complete the repair.

c. The average servicing time is significantly different for these two technicians, and Justin takes a little more than 27 minutes longer on average than Bob to complete the repair.

d. none of the above

Your answer: b
Correct answer: b

The extremely small p-value (<0.0001) in the table shows that the average servicing time for the two technicians differs significantly, and the positive value of 27.1250 in the estimate column shows that Bob takes a little more than 27 minutes longer on average than Justin to complete the repair.

✔ 3. In a diffogram, if a line segment crosses the diagonal reference line, the comparison between the two factor levels (at a given alpha level) is considered to be which of the following?

a. significantly different from each other

b. not significantly different from each other

c. no such comparison can be done using a diffogram

Your answer: b
Correct answer: b

The diffogram provides a graphical test of significance for mean comparisons and displays all pairwise LS-means differences and their significance. Any line segments that cross the diagonal reference line indicate that these groups are not significantly different from each other at a given alpha level. Lines that do not cross the reference line represent pairings that are significantly different from one another.

✔ 4. Consider the two-way table below, which is based on the variables listed in a CLASS statement.

School	**Gender**		
	F	**M**	
Cottonwood			
Dogwood			
Maple			
Pine			

Which of the following statements is correct for testing the hypothesis of interest: "Is the **Reading3** value for Cottonwood the same as the **Reading3** value averaged across Maple and Pine"?

a. lsmestimate School*Gender 1 -0.5 -0.5 0;
b. lsmestimate School*Gender 1 0 -0.5 -0.5;
c. lsmestimate School 1 -0.5 -0.5 0;
d. lsmestimate School 1 0 -0.5 -0.5;

Your answer: d
Correct answer: d

This hypothesis refers to the reading3 scores of three schools (Cottonwood, Maple, and Pine) only and makes no mention of gender and the remaining school (Dogwood). Hence this problem concerns only main effects, and coefficients are needed only for the row margin with a zero in Dogwood because it was not mentioned. The body of the table (the interaction with Gender) will be completed automatically. So, the mathematical equation resembles this:

$\mu_{1.} - 0.5(\mu_{3.} + \mu_{4.}) = 0$

and coefficients: 1 0 -0.5 -0.5

✔ 5. Which of the following is incorrect in relation to ANOVA?

a. ANOVA is robust against departures from normality, with small sample sizes.
b. ANOVA is robust against unequal variances when sample sizes are equal.
c. both a and b
d. none of the above

Your answer: a
Correct answer: a

ANOVA is robust against departures from normality, especially with large enough sample sizes. Also, when the sample sizes are equal, ANOVA is robust against unequal variances.

✔ 6. Which procedure performs post-fitting statistical analyses for the contents of a SAS item store that were previously created with the STORE statement in some other SAS/STAT procedure?

a. PROC GLM
b. PROC GLMSELECT
c. PROC GLIMMIX
d. PROC PLM

Your answer: d
Correct answer: d

The PLM procedure, unlike most SAS/STAT procedures, does not operate primarily on an input data set. Instead, the procedure requires you to specify an item store with the RESTORE= option in the PROC PLM statement. The item store contains the necessary information and context about the statistical model that was fit when the store was created.

✔ 7. The GLIMMIX procedure can be used to model data that might be described in which of the following ways?

a. have correlations
b. have non-constant variances
c. not necessarily be from the normal distribution
d. all of the above

Your answer: d
Correct answer: d

PROC GLIMMIX provides a more general approach and does not require the assumptions of normality, independence, or constant variance. The only assumption is that the response variable belongs to the exponential family of distributions.

---

8. You are conducting a one-way ANOVA. While examining one of the assumptions of ANOVA, you get the following result:

Levene's Test for Homogeneity of Sales Variance ANOVA of Squared Deviations from Group Means					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Ad	3	60380.8	20126.9	9.75	<.0001
Error	132	272616	2065.3		

What would you do next?

a. use a Welch's ANOVA in PROC GLM
b. transform the independent variable and refit the model
c. use the PLM procedure to do post-fitting analysis
d. do nothing and continue interpreting the ANOVA table

Your answer: a
Correct answer: a

The table shows an extremely small p-value for Levene's test. Therefore, you reject the null hypothesis and conclude that the variances are not equal. Thus, an ANOVA assumption was violated. If the variances are not equal and it is a one-way ANOVA, then Welch's ANOVA can be used. Welch's ANOVA is specifically designed for unequal variance situations.

---

9. How do positively correlated observations affect the ANOVA results?

a. The standard errors might be underestimated.
b. The standard errors might be overestimated.
c. The standard errors are not at all affected.
d. The Type II error rate is inflated.

Your answer: a
Correct answer: a

Failure to satisfy the independence assumption can be serious because it affects the accuracy of the standard errors and thus the results of significance tests. If the observations are positively correlated, the standard errors might be underestimated. This can lead to test statistics that are too large. Significance might be detected when it is not truly there, which results in an inflated Type I error rate (rejecting the null hypothesis when it is true). On the other hand, if the observations are negatively correlated, then standard errors are overestimated, which can lead to test statistics that are too small, and true significance is not detected. In other words, the

Type II error rate (failing to reject the null hypothesis when it is false) is inflated, thereby affecting the power of the test. Consequently, the ability of the statistical test to detect a true difference is reduced.

10. Assume that A*B is significant and you only want to test for the effect of A within each level of B. What is the appropriate LSMEANS statement to use in this situation?

    a. lsmeans A*B / slice=A;
    b. lsmeans A*B / slice=B;
    c. lsmeans A*B / slice=A slice=B;
    d. lsmeans AB / slice=A;

    Your answer: b
    Correct answer: b

    When there is significant interaction, interpreting main effects might not be appropriate. It might be of interest to test the effect across one factor within each level of the other factor. To accomplish this, you can use the SLICE= option in the LSMEANS statement. The SLICE= option specifies effects within which to test for differences between interaction LS-mean effects.

Close
