Print

Quiz: ANOVA and Regression

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.

When you're finished, exit the lesson.

1. You can examine Levene's test for homogeneity to more formally test which of the following assumptions?

   a. the assumption of errors being normally distributed
   b. the assumption of independent observations
   c. the assumption of equal variances
   d. the assumption of treatments being randomly assigned

Your answer: c
Correct answer: c

You use Levene's test for homogeneity in PROC GLM to verify the assumption of equal variances in a one-way ANOVA model.

---

2. Given the following output, is there sufficient evidence to reject the assumption of equal variances?

Levene's Test for Homogeneity of Weight Variance					
ANOVA of Squared Deviations from Group Means					
Source	DF	Sum of Squares	Mean Squares	F Value	Pr > F
Brand	1	9.237E-7	9.237E-7	1.12	0.2942
Error	78	0.000065	8.283R-7		

   a. yes
   b. no

Your answer: b
Correct answer: b

The *p*-value of 0.2942 is greater than 0.05, so you fail to reject the null hypothesis

and conclude that the variances are equal.

---

3. Given the following SAS output, is there sufficient evidence to reject the hypothesis of equal means?

Source	DF	Sum of Squares	Mean Squares	F Value	Pr > F
Brand	1	0.03033816	0.03033816	51.02	<.001
Error	79	0.04638442	0.00059467		
Corrected Total	80	0.07672257			

a. yes
b. no

Your answer: a
Correct answer: a

The *p*-value of <.001 is less than 0.05, so you would reject the null hypothesis and conclude that the means between the two brands are significantly different.

---

4. Dunnett's method compares all possible pairs of means.

a. true
b. false

Your answer: b
Correct answer: b

The Tukey method compares all possible pairs of means. Dunnett's method compares all categories to a control group.

5. Which of the following phrases describes the model sums of squares, or SSM, in one-way ANOVA?

a. the variability between the groups
b. the variability within the groups
c. the variability explained by the error terms

Your answer: a
Correct answer: a

SSM is the variability explained by the predictor variable, and therefore, it measures the variability between the groups.

6. Based on the following correlation matrix, what type of relationship do **Performance** and **RunTime** have?

Pearson Correlation Coefficients, N = 31 Prob > |r| under H0: Rho=0			
	Performance	RunTime	Age
Performance	1.00000	-0.82049	-0.71257
		<.0001	<.0001
Error	-0.82049	1.00000	0.19523
	<.0001		0.2926
Corrected Total	-0.71257	0.19523	1.00000
	<.0001	0.2926	

a. a fairly strong, positive linear relationship
b. a fairly strong, negative linear relationship
c. a fairly weak, positive linear relationship
d. a fairly weak, negative linear relationship

Your answer: b
Correct answer: b

The correlation coefficient for the relationship between **Performance** and **RunTime** is -0.82049, which is negative. It's also close to -1, which makes it a relatively strong relationship.

7. In the simple linear regression model, what does β_1 represent?

$Y = \beta_0 + \beta_1 X + \varepsilon$

a. the intercept parameter
b. the predictor variable
c. the variation of X around the line
d. the variation of Y around the line
e. the slope parameter

Your answer: e
Correct answer: e

β_1 is the slope parameter, which is the average change in Y for a 1-unit change in X.

8. Which of the following statements describes a positive linear relationship between two variables?

   1. The more I eat, the less I want to exercise.
   2. The more salty snacks I eat, the more water I want to drink.
   3. No matter how much I exercise, I still weigh the same.

   a. 1 only
   b. 1 and 2
   c. 2 only
   d. 2 and 3
   e. 3 only

   Your answer: c
   Correct answer: c

   In statement 2, the amount of salty snacks eaten and thirst have a positive linear relationship. As the values of one variable (amount of salty snacks eaten) increase, the values of the other variable (thirst) increase as well.

---

9. What output does the following program produce?

```
proc corr data=stat1.bodyfat2 nosimple
   plots(only)=scatter(nvar=all);
   var Age Weight Height;
run;
```

   a. individual correlation plots and simple descriptive statistics
   b. a scatter plot matrix only
   c. a table of correlations and individual scatter plots for each variable in the VAR statement
   d. Not enough information is given.

   Your answer: c
   Correct answer: c

   By default, PROC CORR produces a table of correlations, which can be a correlation matrix, depending on your program. The NOSIMPLE option suppresses printing of the simple descriptive statistics for each variable. To request individual scatter plots, you specify the PLOTS=SCATTER option. After the keyword SCATTER, NVAR=ALL specifies that all the variables listed in the VAR statement be displayed in the plots

10. Given the following PROC REG output and assuming a significance level of 0.05, which of the following statements is true? Select all that apply.

Analysis of Variance					

Source	DF	Sum of Squares	Mean Squares	F Value	Pr > F
Model	1	119.72668	119.72668	2.00	0.1585
Error	250	14959	59.83716		
Corrected Total	251	15079			

Root MSE	7.73545	R-Square	0.0079
Dependent Mean	18.93849	Adj R-Sq	0.0040
Coeff Var	40.84511		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr >|t|
Intercept	1	32.16542	9.36350	3.44	0.0007
Height	1	-0.18856	0.13330	-1.41	0.1585

a. The model explains approximately 15% of the variation in the response variable.

b. You should reject the null hypothesis.

c. **Height** is statistically significant for predicting the values of the response variable.

d. The model explains less than 1% of the variation in the response variable.

Your answer: d
Correct answer: d

The R-square value indicates that the model explains less than 1% of the variation in the response variable. With a *p*-value of 0.1585, **Height** is not statistically significant for predicting the values of the response variable. Likewise, the *p*-value of 0.1585 for the model indicates that you should fail to reject the null hypothesis.

Close