Print

Quiz, Lesson 4: Analysis of Covariance

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 type of predictors are always included in ANCOVA models?

   a. only categorical variables
   b. only contiguous variables
   c. both categorical and continuous variables
   d. only interactions of categorical and continuous variables

Your answer: c
Correct answer: c

Analysis of covariance is the regression between a continuous response (Y variable) and a combination of continuous and categorical predictors (X variables). Interactions might or might not be present in an ANCOVA model.

---

2. Suppose you fit an ANCOVA model using PROC GLM, where **Gender** is a categorical predictor variable and **Words1** is a covariate. Given the following parameter estimates table, what regression equation can you write for **Gender** = **M** (male)?

Parameter	Estimate		Standard Error	t Value	Pr > |t|
Intercept	17.17307433	B	3.84718936	4.46	<.0001
Gender F	-7.45084258	B	6.55582187	-1.14	0.2576
Gender M	0.00000000	B	.	.	.
Words1	1.65070369	B	0.16394259	10.07	<.0001
Words1*Gender F	0.62715889	B	0.26857250	2.34	0.0209
Words1*Gender M	0.00000000	B	.	.	.

   a. Y = 17.17307 + 1.65070 * Words1
   b. Y = 17.17307 + (1.66070 + 0.62715) * Words1
   c. Y = (17.17307 - 7.45084) + (1.65070 - 0.62715) * Words1
   d. Y = (17.17307 - 7.45084) + (1.65070 + 0.62715) * Words1

Your answer: a
Correct answer: a

In the parameter estimates table, the estimate 17.17307 represents the intercept for the last level of **Gender**, which is **M** (male) in this case. The parameter estimate for gender **M** is zero

because it estimates the difference between the intercept of gender **M** and the intercept of the last group (that is, gender **M**). The parameter estimate for **Words1*Gender M** is also zero because it estimates the difference between the overall slope and the slope of **Words1** for gender **M**. Hence, the regression equation for gender **M** can be written as Y = 17.17307 + 1.65070 * Words1.

3. In the following parameter estimates table (the same table shown in the previous question), what does the value -7.45084258 represent?

Parameter	Estimate		Standard Error	t Value	Pr > |t|
Intercept	17.17307433	B	3.84718936	4.46	<.0001
Gender F	-7.45084258	B	6.55582187	-1.14	0.2576
Gender M	0.00000000	B	.	.	.
Words1	1.65070369	B	0.16394259	10.07	<.0001
Words1*Gender F	0.62715889	B	0.26857250	2.34	0.0209
Words1*Gender M	0.00000000	B	.	.	.

a. estimate of the slope for gender *M*
b. estimate of the slope for gender *F*
c. difference between the intercept of gender *F* and the intercept of gender *M*
d. difference between the slope of gender *F* and the slope of gender *M*

Your answer: c
Correct answer: c

The intercept term estimates the intercept of the last level of the grouping variable **Gender**, which is **M** in this case. Therefore, the parameter estimate for gender **F** (-7.45084258) estimates the difference between the intercept of gender **F** and the intercept of the last level (that is, **M**).

4. Using PROC GLM, you are performing an analysis of covariance to determine the effect of age and gender on income. To determine whether there is a difference between the genders when age equals the average value, which of the following SAS programs can you use?

a.
```
proc glm data=sasuser.abc
   class Gender;
   model Income=Gender|Age;
   lsmeans Gender / at Age=mean pdiff adjust=tukey;
run;
quit;
```
b.
```
proc glm data=sasuser.abc
   class Gender;
   model Income=Gender|Age;
   lsmeans Gender / pdiff adjust=tukey;
run;
quit;
```
c.

```
proc glm data=sasuser.abc
   class Gender;
   model Income=Gender|Age;
   means Gender / pdiff adjust=tukey;
run;
quit;
```

d.

```
proc glm data=sasuser.abc
   class Gender;
   model Income=Gender|Age;
   means Gender / at Age=mean pdiff adjust=tukey;
run;
quit;
```

Your answer: b
Correct answer: b

By default, the LSMEANS statement gives the estimate of the average value of the dependent variable for each group at the mean value of the covariate (in this case, **Age**).

5. Consider the following Type III model ANOVA table. Assuming that α = 0.05, what do you conclude from the *p*-value of the **Words1*Gender** interaction term?

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Gender	1	838.9809	838.9809	1.29	0.2576
Words1	1	138976.2850	138976.2850	213.97	<.0001
Words1*Gender	1	3541.8299	3541.8299	5.45	0.0209

a. The slope parameters for the two levels of **Gender** are not equal.
b. The slope parameters for the two levels of **Gender** are not significantly different.
c. An equal slope model is needed for this data set.
d. There is insufficient information to draw a conclusion.

Your answer: a
Correct answer: a

The *F* statistic for the interaction term tests the null hypothesis that the slope parameters are equal. At the given value of α = 0.05, you reject the null hypothesis and conclude that the slope parameters for the two levels of **Gender** are not equal (that is, not significantly different from each other).

6. Which SAS procedure is best suited to perform all the required model diagnostics for ANCOVA models?

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

Your answer: d
Correct answer: d

PROC REG has all the diagnostics available in PROC GLM, plus additional diagnostics such as

multicollinearity diagnostics and plots of the DFBETA and DFFITS statistics. Hence, you can perform all of the diagnostics for ANCOVA models using PROC REG.

7. Which of the following is a possible remedy for multicollinearity in ANCOVA models?

   a. centering the continuous predictor (covariate)
   b. centering the categorical predictorr
   c. centering both the categorical and continuous predictors
   d. transforming the response variable

   Your answer: a
   Correct answer: a

   Multicollinearity is frequently seen in ANCOVA. One of the possible remedies is centering the continuous predictor variable. Centering a variable means that a constant is subtracted from every value of the variable. You could center a variable around its mean or around a constant that has some intrinsic meaning for the variable.

8. What makes PROC GLM more suitable than PROC REG for performing ANCOVA?

   a. PROC GLM produces more diagnostic statistics than PROC REG.
   b. PROC GLM supports the CLASS statement whereas PROC REG does not.
   c. PROC GLM enables you to list interaction terms directly in the MODEL statement, but when you use PROG REG, you must first create interaction terms in a DATA step.
   d. both *b* and *c*

   Your answer: d
   Correct answer: d

   Both PROG REG and PROC GLM use the OLS method to fit general linear models in which the response variable is continuous. However, PROC GLM is more convenient to use when you have categorical predictors because it supports the CLASS statement, which creates design variables based on the chosen parameterization. When you use PROC REG, you need to create design variables in a DATA step because PROC REG does not support the CLASS statement. Also, PROC GLM enables you to list interaction times directly in the MODEL statement whereas, in PROC REG, you must create interaction terms in a DATA step before referencing them in the MODEL statement.

9. Suppose you fit an ANCOVA model using PROC GLM by submitting the following SAS program:

```
proc glm data=sasuser.abc;
   class Eduqualification;
   model Income = Eduqualification Age;
         Eduqualification*Age;
run;
```

   Income is a continuous response variable, **Eduqualification** is a categorical variable with two levels (*Graduate* and *Postgraduate*), and **Age** is a continuous variable.

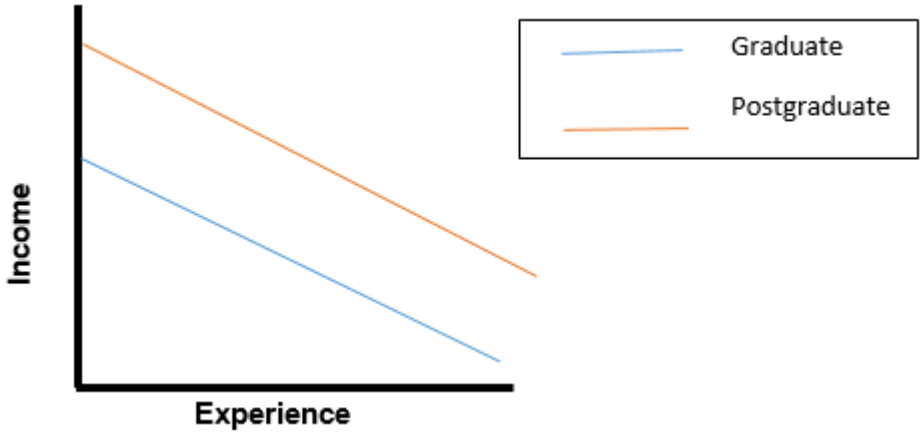
   How many parameters does PROC GLM estimate for this ANCOVA model?

   a. 3
   b. 4
   c. 5
   d. 6

Your answer: d
Correct answer: d

PROC GLM estimates six parameters in all, including the intercept, the effect of **Eduqualification1** (*Graduate*), the effect of **Eduqualification2** (*Postgraduate*), the slope of **Age**, the slope effect for the *Graduate* level, and the slope effect of the *Postgraduate* level (that is, the interaction term).

✔ 10. Consider an ANCOVA model with the continuous predictor **Experience** (years of experience), the categorical predictor **Eduqualification** (which has two levels), and the response variable **Income**. Which type of relationship between the variables is depicted by the following graph?



a. unequal slopes and intercepts
b. equal slopes and unequal intercepts
c. equal intercepts but unequal slopes
d. equal slopes and intercepts

Your answer: b
Correct answer: b

The graph shows the model in which the slopes are the same but the intercepts are different for **Eduqualification**. When **Experience** = 0, the income appears to be different for two of the **Eduqualification** levels—these are the unequal intercepts. However, **Income** increases or decreases at the same rate for the two levels of **Eduqualification** as the value of **Experience** changes—that is, all the slopes are equal.

Close