# EPY 9213

**Advanced Analysis in Educational Research Assignment 8, Fall, 2015**

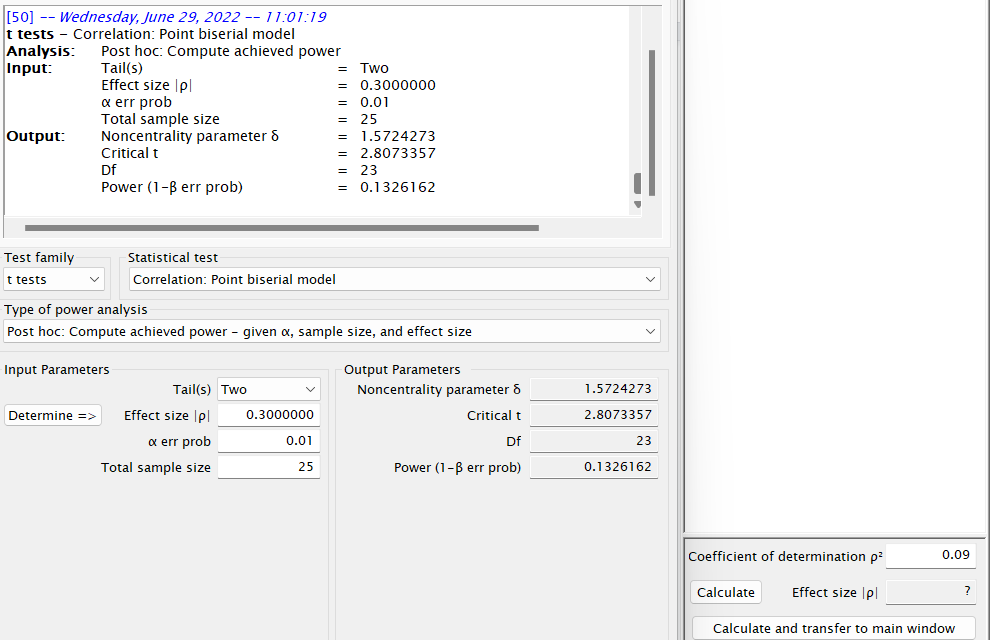
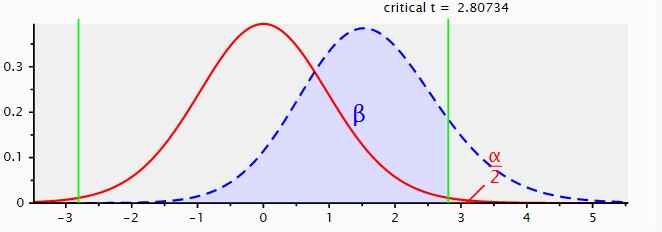
**Cindy Musselwhite**

1. An investigator administers a 30-item, semantic differential (9-point scale per response) instrument. She wants to compare four independent groups of respondents on the scale, treating each of the three semantic differential “factors” as a dependent variable, using MANOVA.
   1. **Using G\*Power**, how much power will the study have if the researcher uses groups of 25 persons each, tests differences at the 0.01 level, and would like to be able to detect an effect size (Cohen’s *f2*) of .09? (*Hint*: This is a **post hoc power analysis**, since you’re solving for power.)

Power (1-β err prob) = 0.1391403

Answer: Using G\*Power

Result:

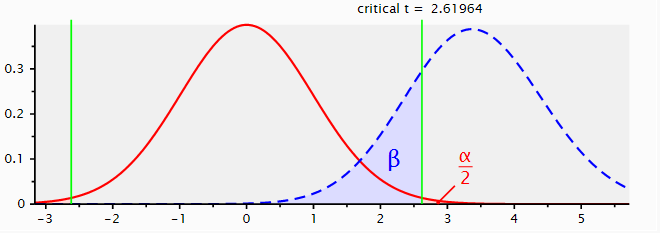


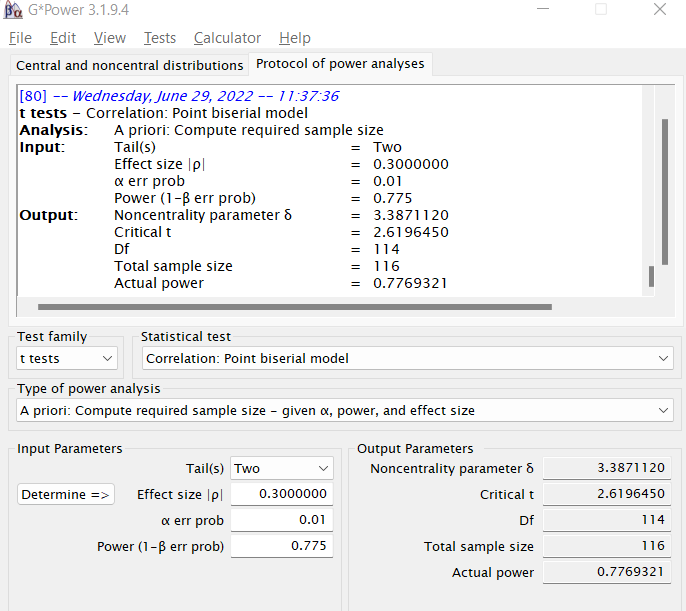
Explanation: The above test is based on the Post hoc to compute power i.e., 1-beta. The total sample size is 25 where the alpha error probability is 0.01. This shows us that we would have 1% power to detect the power. Thus, the power of analysis is 14%. Based on the evidence, we can conclude that the power is close to 0, the hypothesis test is not very good at detecting a false null hypothesis.

* 1. For this study, what total sample size would assure (again, using G\*Power, though as an a priori analysis) a power of .95 or better to detect an ES of .09, using an alpha of .01?

Total sample size = 116

Answer: Result:





Explanation: The above test is based on a priori test to compute required sample size i.e., n. After running the test, the total sample size is 116 where the alpha error probability is 0.01 and the power (1-beta error probability) is 0.775 that is closest to 1. Based on the power value, we can conclude that the power is close to 1, the hypothesis test is very good at detecting a false null hypothesis.

* 1. For this study, is the Mahalanobis *D*2 an appropriate effect size statistic to report? Explain your reasoning.

Answer: The Mahalanobis *D*² is an appropriate effect size statistic to report in this study because it takes into account correlations and outliers and this study is based on a 9-point scale response. All studies should be checked for outliers.

* 1. What are the names of the usual three factors said to account for scores on the semantic differential, as explained by Osgood, Suci, and Tannenbaum? (If you don’t know, look it up!)

Answer: According to Osgood, Suci, and Tannenbaum the three factors that account for scores on the semantic differential are evaluation (good-bad), potency (strong-weak), and activity (active-passive).

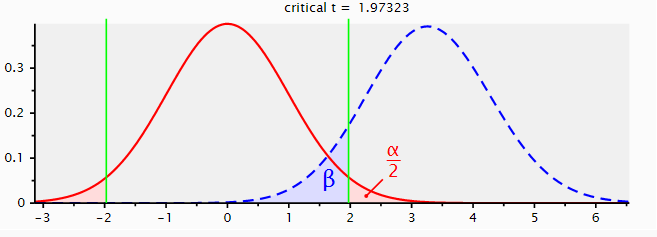
1. An investigator notes an estimated effect size (using Mahalanobis’ *D*2) of 0.20 from a previous two-group study that used the same set of six dependent variables that he proposes to use with a similar sample of subjects. Given this, and assuming equal group sizes, how many total cases are needed to yield a power of at least .90 using an alpha of .05?

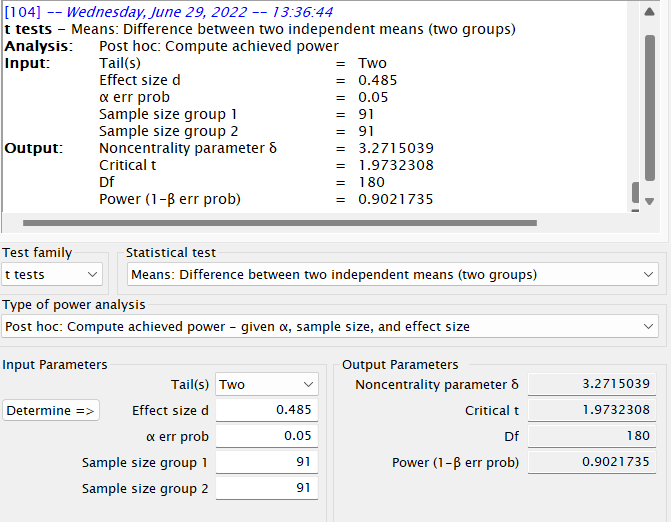
Use the **G\*Power program.**

(*Note*: See the handout on multivariate *ES* to convert *D*2 to *f2* or vice versa. Start with a guessed *N* = 200 for the conversion…you’ll find that the converted values don’t change very much unless *N* gets much smaller.)

Total sample size = 182

Answer:





Explanation: The above test is based on a priori test to compute power i.e., n. After running the test, the power (1-beta error probability) is 0.90 where the alpha error probability is 0.05 and the number of sample sizes are divided equally (that is 91 for both sample sizes). Based on the power value, we can conclude that the power is close to 1, the hypothesis test is very good at detecting a false null hypothesis.

1. A study comparing women on four measures (DVs) of cognitive functioning was completed.

A 3 (cancer status: (a) no history,

(b) previous cancer treated via chemotherapy, or

(c) previous cancer treated only surgically) x 2 (over 60, under 60 years of age) balanced, factorial design was used (with six cells).

With a total of 216 participants, the multivariate *F*-ratio (from Wilks’ lambda) for the main effect of cancer status was reported as 3.88. From this information:

(a) was the result (main effect of cancer status) statistically significant at the .01 level (*show how you arrived at your decision*)?

(b) report and interpret the value of Cohen’s effect size (*f2*) for the IV of cancer status.

(*Hint*: Check the handout on multivariate ES, or you may use the MINSIZE program.)

MINSIZE2:

Minimum sample size for statistical significance.

======================================================================

\*\*\* Evaluate an ANOVA design (F-ratio) \*\*\*

Nondirectional alpha level was 0.0500

Observed F-ratio was 3.8800

Total sample size was 216

Number of groups compared was 4

Total cells in design was 4

Number of covariates was 1

Number of dependent variables was 4

----------------------------------------------------------------------

Original df was (12,551)

Wilks' lambda value was 0.8068

Partial eta-squared was 0.0779 ("medium")

Effect size (f-squared) was 0.0845 ("small")

Minimum N for significance at this alpha is 105 cases.

(With equal n's, that would call for 27 cases per group.)

With this size, the resultant F-ratio is 1.7956

with a nondirectional probability of 0.0491

The main effect of cancer status was not significant at the *p* < .01, *p* = .05. The effect size was small, *f*² = .08.

Answer: The rule of p-value is that if the p-value is less than the level of significance then the test/variable is statistically significant, otherwise insignificant. So, the result of the main effect of cancer status is insignificant because the p-value of is more than the level of significance (that is 0.05 > 0.01). On the other hand, the value of Cohen’s effect size (f-squared) was small i.e., 0.0845. Thus, the meaning of Cohen's f^2 is appropriate for calculating the effect size within a multiple regression model in which the independent variable of interest and the dependent variable are both continuous. Hence, we can conclude that the small Cohen’s effect size indicates that the limited practical significance of a research outcome i.e., cancer status.

1. The data set (available on the course web page) is a subset from a study on whether specific listening conditions influence divergent thinking production. Scores of 32 cases from one of the treatment conditions are given.

The three dependent variables are:

*Fluency (how many ideas one comes up with),*

*Flexibility (how many different categories of ideas are generated), and*

*Originality (how many novel ideas one generates).*

The usual pattern for scores like this is Fluency > Flexibility > Originality.

* 1. Analyze the data set and interpret the results from the univariate perspective (using simple repeated measures ANOVA), using alpha = 0.01. Is there a difference in means across score types, and if so, does it support the pattern typically reported? For this data set, do not omit any cases as being “outliers.” However, do check assumptions, fix any problems, and use and interpret follow-up tests as needed (no revision permitted if you do not, or if you incorrectly run the overall test).

|  | Mean |
| --- | --- |
| Fluency score | 19.37 |
| Flexibility score | 11.41 |
| Originality score | 4.66 |

There is a difference between the means across score types. This reports the typical pattern because a simple repeated measure ANOVA test scores the variables as independent.

This test shows that a difference exists between originality but not fluency nor flexibility.

| **Tests of Between-Subjects Effects** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Dependent Variable: case | | | | | | |
| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
| Corrected Model | 1682.180a | 3 | 560.727 | 15.012 | .000 | .617 |
| Intercept | 1252.445 | 1 | 1252.445 | 33.532 | .000 | .545 |
| flu | 1494.214 | 1 | 1494.214 | 40.005 | .000 | .588 |
| orig | 7.940 | 1 | 7.940 | .213 | .648 | .008 |
| flex | 556.712 | 1 | 556.712 | 14.905 | .001 | .347 |
| Error | 1045.820 | 28 | 37.351 |  |  |  |
| Total | 11440.000 | 32 |  |  |  |  |
| Corrected Total | 2728.000 | 31 |  |  |  |  |
| a. R Squared = .617 (Adjusted R Squared = .576) | | | | | | |

Using 70% of the sample instead of all 32 cases yielded similar results with a difference still existing between originality but not fluency nor flexibility and a slightly higher difference in score means.

|  | Mean |
| --- | --- |
| Fluency score | 20.04 |
| Flexibility score | 11.81 |
| Originality score | 5.04 |

| **Tests of Between-Subjects Effects** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Dependent Variable: case | | | | | | |
| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
| Corrected Model | 1675.928a | 3 | 558.643 | 17.353 | .000 | .703 |
| Intercept | 776.651 | 1 | 776.651 | 24.126 | .000 | .523 |
| flu | 1567.227 | 1 | 1567.227 | 48.684 | .000 | .689 |
| orig | .018 | 1 | .018 | .001 | .981 | .000 |
| flex | 565.700 | 1 | 565.700 | 17.573 | .000 | .444 |
| Error | 708.226 | 22 | 32.192 |  |  |  |
| Total | 8538.000 | 26 |  |  |  |  |
| Corrected Total | 2384.154 | 25 |  |  |  |  |
| a. R Squared = .703 (Adjusted R Squared = .662) | | | | | | |

Answer: Case 1- Sample Size = 32

The highlighted rows inform us whether our independent variables (the "fluency" rows and "flexibility" rows ) have a statistically significant effect on the dependent variable, "case". We can see if the variable is significant or not from **“Sig.”** column. The “orig” variable is insignificant because the p-value is more than the alpha level of significance and there is no effect on the “case” variable. On the other hand, the R-squared value of the model is 62%, which implies that the variability observed in the target variable “case” is explained by the independent model.

Case 2- Sample Size- 25.

The highlighted rows inform us whether our independent variables (the "fluency" rows and "flexibility" rows ) have a statistically significant effect on the dependent variable, "case". We can see if the variable is significant or not from **“Sig.”** column. The “orig” variable is insignificant because the p-value is more than the alpha level of significance and there is no effect on the “case” variable. On the other hand, the R-squared value of the model is 70%, which implies that the variability observed in the target variable “case” is explained by the independent model. Thus, the second model is better than the first model. The reason for r-square increases is that the sample size decreases from 32 to 25, so that the R-square is closer to 1.

* 1. Analyze the data set and interpret the results from the multivariate perspective, using alpha = .01. Is there a difference in means across score types, and if so, does it support the pattern typically reported? Check assumptions, and use and interpret follow-up tests as needed (no revision permitted if you do not, or if you incorrectly run the overall test). *Notes*: Refer to the handout on “profiles” for how to elicit the multivariate test on this. When you run the GLM- repeated measures analysis in SPSS, the first table you will get in your output summarizes the *multivariate* test, so be sure not to interpret the wrong table(s)!

|  | Mean |
| --- | --- |
| Fluency score | 20.04 |
| Flexibility score | 11.81 |
| Originality score | 5.04 |

| **Tests of Between-Subjects Effects** | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Source | | Dependent Variable | | | Type III Sum of Squares | | df | Mean Square | F | Sig. | Partial Eta Squared | |
| Corrected Model | | Fluency score | | | 782.208a | | 1 | 782.208 | 8.304 | .008 | .257 | |
| Flexibility score | | | 22.591b | | 1 | 22.591 | .618 | .440 | .025 | |
| Originality score | | | .651c | | 1 | .651 | .051 | .823 | .002 | |
| Intercept | | Fluency score | | | 915.013 | | 1 | 915.013 | 9.714 | .005 | .288 | |
| Flexibility score | | | 1285.280 | | 1 | 1285.280 | 35.155 | .000 | .594 | |
| Originality score | | | 203.372 | | 1 | 203.372 | 15.935 | .001 | .399 | |
| case | | Fluency score | | | 782.208 | | 1 | 782.208 | 8.304 | .008 | .257 | |
| Flexibility score | | | 22.591 | | 1 | 22.591 | .618 | .440 | .025 | |
| Originality score | | | .651 | | 1 | .651 | .051 | .823 | .002 | |
| Error | | Fluency score | | | 2260.753 | | 24 | 94.198 |  |  |  | |
| Flexibility score | | | 877.448 | | 24 | 36.560 |  |  |  | |
| Originality score | | | 306.311 | | 24 | 12.763 |  |  |  | |
| Total | | Fluency score | | | 13483.000 | | 26 |  |  |  |  | |
| Flexibility score | | | 4525.000 | | 26 |  |  |  |  | |
| Originality score | | | 967.000 | | 26 |  |  |  |  | |
| Corrected Total | | Fluency score | | | 3042.962 | | 25 |  |  |  |  | |
| Flexibility score | | | 900.038 | | 25 |  |  |  |  | |
| Originality score | | | 306.962 | | 25 |  |  |  |  | |
| a. R Squared = .257 (Adjusted R Squared = .226) | | | | | | | | | | | | |
| b. R Squared = .025 (Adjusted R Squared = -.016) | | | | | | | | | | | | |
| c. R Squared = .002 (Adjusted R Squared = -.039) | | | | | | | | | | | | |
| **Multivariate Testsa** | | | | | | | | | | | |
| Effect | | | Value | F | | Hypothesis df | Error df | Sig. | Partial Eta Squared | | |
| Intercept | Pillai's Trace | | .677 | 19.593b | | 3.000 | 28.000 | .000 | .677 | | |
| Wilks' Lambda | | .323 | 19.593b | | 3.000 | 28.000 | .000 | .677 | | |
| Hotelling's Trace | | 2.099 | 19.593b | | 3.000 | 28.000 | .000 | .677 | | |
| Roy's Largest Root | | 2.099 | 19.593b | | 3.000 | 28.000 | .000 | .677 | | |
| case | Pillai's Trace | | .617 | 15.012b | | 3.000 | 28.000 | .000 | .617 | | |
| Wilks' Lambda | | .383 | 15.012b | | 3.000 | 28.000 | .000 | .617 | | |
| Hotelling's Trace | | 1.608 | 15.012b | | 3.000 | 28.000 | .000 | .617 | | |
| Roy's Largest Root | | 1.608 | 15.012b | | 3.000 | 28.000 | .000 | .617 | | |
| a. Design: Intercept + case | | | | | | | | | | | |
| b. Exact statistic | | | | | | | | | | | |

Using the multivariate test displays a greater analysis of the between-subjects effects because multivariate tests are more sensitive because one is looking at more information. There is a difference between the means across score types. This reports the typical pattern similar to the repeated measure ANOVA test.

Answer: For p-value: The highlighted rows inform us whether our independent variables, that is the "fluency score" have a statistically significant effect on the dependent variable, "case". However, the p-value of both the independent variables, “flexi” and “orig” score, more than the alpha level of significance. So, these variables are insignificant, apart from that, the partial eta squared of these variables is too small which means that there is no effect on the dependent variable, “case”. On the other hand, the correct model gives three r-squared values where the better r-squared value is 0.257 or 26%.

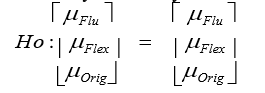
c. Which analysis seems to be better suited for the data, and why?

Answer: The multivariate seems better suited to this study because it yields more information as to the differences between variables.

d. What hypothesis is represented by each analysis? Write each in statistical form.

Answer: Hypothesis:

Null Hypothesis:



Alternative Hypothesis:

H1:

When the p-value of any variable is less than the alpha level of significance (5%), this means that the variable is statistically significant. So, based on that, we can reject the null hypothesis and in favor of an alternative hypothesis. Based on the above result, the p-value of the “orig” variable is insignificant because the p-value is more than the alpha level of significance, while the rest of the variables are significantly significant because p-value is less than the significance level.

e. Give the values of both Cohen’s f2 and eta-squared for each analysis (that’s four ES values in all); interpret one of the effect size measures for each analysis.

| Source | Cohen’s *f²* | Partial Eta Squared |
| --- | --- | --- |
| Intercept | 1.20 | .545 |
| flu | 1.33 | .588 |
| orig | .01 | .008 |
| flex | .53 | .347 |

Answer: Partial Eta Squared: The partial eta squared (η2) value of “flu” and “flex” is 0.6 and 0.35, respectively. So, the effect size of “flu” and “flex” is large. Thus, the above analysis results show that our independent variables (the "fluency" and "flexibility") have a large effect on the dependent variable, "case". However, the partial eta squared (η2) value of the “originality” independent variable is 0.008 which indicates that the effect size is very small.