# MMI-409 Exam Information

**Exam 2 is due**: Sunday, May 18, 2014 (11:55pm Central time)

**Description**: The exam will cover topics from sessions 4, 5, and 6.

**Resources:** The exam is completely open book. You may use the course textbooks, materials provided on Blackboard, a calculator, and SPSS to answer the questions. For questions that require calculations, all calculations should be shown, not just the final answer. For questions that require an answer using SPSS, the Output Panel screens should be copied to MS Word and attached to the exam.

**Restrictions**: All answers are to be your work only. You are not to receive assistance from any other person.

**To complete the exam:**

Answer all questions on the exam thoroughly and put your name on each page of the completed exam. Create an MS Word document, including the question number, the question, your typed answer, and SPSS screens if required. You may use Word's equation editor to complete your answers.

PLEASE NOTE: If you are scanning in handwritten work, you are responsible for the quality of the scan. If answers are illegible, spotty or too light to read, it may lead to your answer not being included in your final score.

1. Once you have completed your exam, return to the exam item where you downloaded the exam PDF, click View/Complete Assignment, and submit your document.

**Name: Kazi Russell

# MMI-409 Exam 2

Note:

- Some questions specify that SPSS output be shown as part of the answer. SPSS output can be copied and pasted into a Word document using the COPY and the PASTE SPECIAL commands. In the Paste Special dialog box, select "Picture (Windows Metafile)."

1. The 2006 General Social Survey contains information on the number of hours worked by a respondent each week. For a sample of 83 persons, about 30.1 percent of the sample work less than 40 hours per week. The mean number of hours worked per week is 39.04, with a standard deviation of 11.51.

1a. Calculate a 95 percent confidence interval for the mean data. Show your work. Explain, in plain English, what your results mean. (5 points)

**Standard error of mean,** $\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{\mathbf{11.51}}{\sqrt{\mathbf{83}}} = \mathbf{1.26}$

95% of confidence interval = $\overline{x} \pm 1.96\ \sigma_{\overline{x}}$ = 39.04$\pm$ 2.48 = **{36.56, 41.52}**

We are 95% confident that the respondents worked between 36.56 hours and 41.52 hours.

1b. If we increased the sample size to 10,000 and used a level of confidence of 99 percent, how would this impact the confidence interval for the mean? Calculate the confidence interval (showing your work) and describe your results. (5 points)

**By increasing the sample size and**

**The new Standard error of mean,** $\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{\mathbf{11.51}}{\sqrt{\mathbf{10000}}} = \mathbf{0.12}$

**99% of confidence interval =** = $\overline{x} \pm 2.576\ \sigma_{\overline{x}} = \mathbf{39.04} \pm 0.30 = \{38.74, 39.34\}$

**We are 99% confidence that for the sample size of 10000, the respondents worked between 38.74 and 39.74 hours.**

2. Using the normal distribution, we know that **95%** of all random sample means will fall within +/-1.96 standard errors of the true population mean. (5 points)

A) 68
**B) 95**
C) 99
D) 100

3. **Type II error** is the probability associated with failing to reject the null hypothesis when it is false; whereas, **type I error** is the probability associated with rejecting the null hypothesis when it is true. (5 points)

A) type I error; type II error
**B) type II error; type I error**
C) p-value; alpha
D) alpha; p-value

4. The 40-hour work week is generally considered as a standard in American society today. Using data from the 2006 General Social Survey, you wish to determine whether the mean number of hours worked per week by men in the sample differs from the 40-hour standard.

| Weekly Hours Worked Men | |
|---|---|
| Range | 60 |
| Minimum value | 5 |
| Maximum value | 65 |
| Mean | 42.31 |
| Variance | 100.00 |
| Standard deviation | 10.00 |
| Sum | 1,227 |
| Number of observations | 29 |

4a. State the null hypothesis ($H_0$) and alternative hypothesis ($H_{1)}$. (5 points)

**Null hypothesis** $\boldsymbol{H_0}$: $\boldsymbol{\mu = \mu_0}$, the mean number of hours worked per week by men is the same as the 40-hour standard.
**Alternate hypothesis** $\boldsymbol{H_a}$: $\boldsymbol{\mu \neq \mu_0}$**,** the sample mean differs from the 40-hour standard.

4b. Use some or all of the information provided with this question (4) to calculate the value of the single sample t-test. (10 points)
Test Statistic: $\boldsymbol{t_{stat}}$ = $(\bar{x} - \mu_0)$ / (s/√n) = (42.31 – 40.00) / (10/√29) = 2.31/ 1.86 = **1.24**
**Critical t value,** $\boldsymbol{t_{critical}}$
Degrees of freedom, df = 29 -1 = 28
For α = 0.05, $\boldsymbol{t_{critical}}$ $(df = 28) = \mathbf{2.048}$

4c. Explain, in plain English, what your results mean. (5 points)
Since $\boldsymbol{t_{stat} < t_{critical}}$, the test statistics is inside the non-rejection region and inside the critical value. As a result, we failed to reject the null hypothesis that the number of hours worked by men is 40.

5. Suppose you are interested in comparing the mean number of hours worked by gender.

| Weekly Hours Worked Women | |
|---|---|
| Range | 59 |
| Minimum value | 11 |
| Maximum value | 70 |
| Mean | 37.28 |
| Variance | 142.88 |
| Standard deviation | 11.95 |
| Sum | 2013 |
| Number of observations | 54 |
| Weekly Hours Worked Men | |
| Range | 60 |
| Minimum value | 5 |
| Maximum value | 65 |
| Mean | 42.31 |
| Variance | 100.00 |
| Standard deviation | 10.00 |
| Sum | 1,227 |
| Number of observations | 29 |

5a. State the null and alternative hypotheses. (5 points)
**Null Hypothesis:** There is no statistical significant difference in the mean number of hours worked by women or men. $\boldsymbol{H_o}$: $\boldsymbol{\mu_1 = \mu_2}$
**Alternate Hypothesis:** There is a statistical significant difference in the mean number of hours worked by women or men. $\boldsymbol{H_a}$: $\boldsymbol{\mu_1 \neq \mu_2}$

5b. Use all or some of the information provided to calculate the value of the independent samples t- statistic. (10 points)

**Test statistic:** $\boldsymbol{t_{stat} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{42.31 - 37.28}{\sqrt{\frac{100}{29} + \frac{142.88}{54}}} = \frac{5.03}{\sqrt{3.44 + 2.64}} = \frac{5.03}{2.47} = 2.04}$

**Calculating $t_{critical}$ :**
The degrees of freedom, df = $n_1 + n_2 - 2 = 29 + 54 - 2 = 81$
At 95% confidence level ( α = 0.05) , $\boldsymbol{t_{critical}\,(81) = 1.99}$

5c. Explain, in plain English, your results. (5 points)
Since $\boldsymbol{t_{stat}}(2.04) > \boldsymbol{t_{critical}}\,(1.99)$; the null hypothesis is rejected. In other words, there is a difference in the mean number of hours worked by women or men.

6. For ANOVA, why is the F statistic also called an F ratio? (5 points)

B

A) it is the ratio of one group mean to another

**B) it is the ratio of the mean square between to the mean square within**

C) it is the ratio of the with group sum of squares to the between group sum of squares

D) it is the ratio of degrees of freedom for the within group sum of squares to the degrees of freedom for the between group sum of squares

6. Use the data file ALCOHOL.SAV (attached to the exam assignment) to conduct an ANOVA examining difference in alcohol use by age.

6a. State your null and alternative hypotheses. (5 points)

**Null hypothesis:** Alcohol consumption does not depend on age. That is,

$\boldsymbol{H_0: \mu_{14} = \mu_{15} = \mu_{16}}$

**Alternative hypothesis:** Not all the means of age are equal. In other words, the age effects alcohol consumptions.

6b. Attach your SPSS output. (10 points)

Please see the attachment.

6c. Interpret your SPSS output. (5 points)

The large F ratio (5.61) of ANOVA analysis rejects the null hypothesis. So there must be a relationship between the age and alcohol consumption. We can see that in the plot where the alcohol consumption increases with age.

6d. Run two more ANOVAs – one for males (examining age and alcohol use) and one for females (examining age and alcohol use). What can you conclude? (10 points)

Males consume more alcohol than their female counterparts. Males steadily increase their alcohol consumes as they age whereas females though at first increase their alcohol consumption as they age, their consumption level of between age 15 and 16.

7. Select the best response: Maximum power of a statistical test is achieved when the sample size for group 1 is greater than the sample size for group 2. (5 points)

**A) Greater than**

B) Less than

C) Equal to

[DataSet1] C:\Documents and Settings\Apple\Desktop\Alcohol(2).sav

**Descriptives**

ALCUSE

| | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | | Minimum | Maximum |
|---|---|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound | | |
| 14 | 82 | .6305 | .93891 | .10369 | .4242 | .8368 | .00 | 3.32 |
| 15 | 82 | .9636 | 1.04737 | .11566 | .7335 | 1.1938 | .00 | 3.61 |
| 16 | 82 | 1.1718 | 1.13535 | .12538 | .9223 | 1.4212 | .00 | 3.46 |
| Total | 246 | .9220 | 1.06311 | .06778 | .7884 | 1.0555 | .00 | 3.61 |

**Test of Homogeneity of Variances**

ALCUSE

| Levene Statistic | df1 | df2 | Sig. |
|---|---|---|---|
| 2.925 | 2 | 243 | .056 |

**ANOVA**

ALCUSE

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 12.227 | 2 | 6.113 | 5.613 | .004 |
| Within Groups | 264.673 | 243 | 1.089 | | |
| Total | 276.900 | 245 | | | |

**Robust Tests of Equality of Means**

ALCUSE

| | Statistic[a] | df1 | df2 | Sig. |
|---|---|---|---|---|
| Welch | 5.835 | 2 | 161.006 | .004 |
| Brown-Forsythe | 5.613 | 2 | 237.455 | .004 |

a. Asymptotically F distributed.

## Post Hoc Tests

**Multiple Comparisons**

Dependent Variable: ALCUSE

Tukey HSD

| (I) AGE | (J) AGE | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| 14 | 15 | -.33316 | .16299 | .104 | -.7175 | .0512 |
| | 16 | -.54130* | .16299 | .003 | -.9257 | -.1569 |
| 15 | 14 | .33316 | .16299 | .104 | -.0512 | .7175 |
| | 16 | -.20814 | .16299 | .410 | -.5925 | .1762 |
| 16 | 14 | .54130* | .16299 | .003 | .1569 | .9257 |
| | 15 | .20814 | .16299 | .410 | -.1762 | .5925 |

*. The mean difference is significant at the 0.05 level.

## Homogeneous Subsets

**ALCUSE**

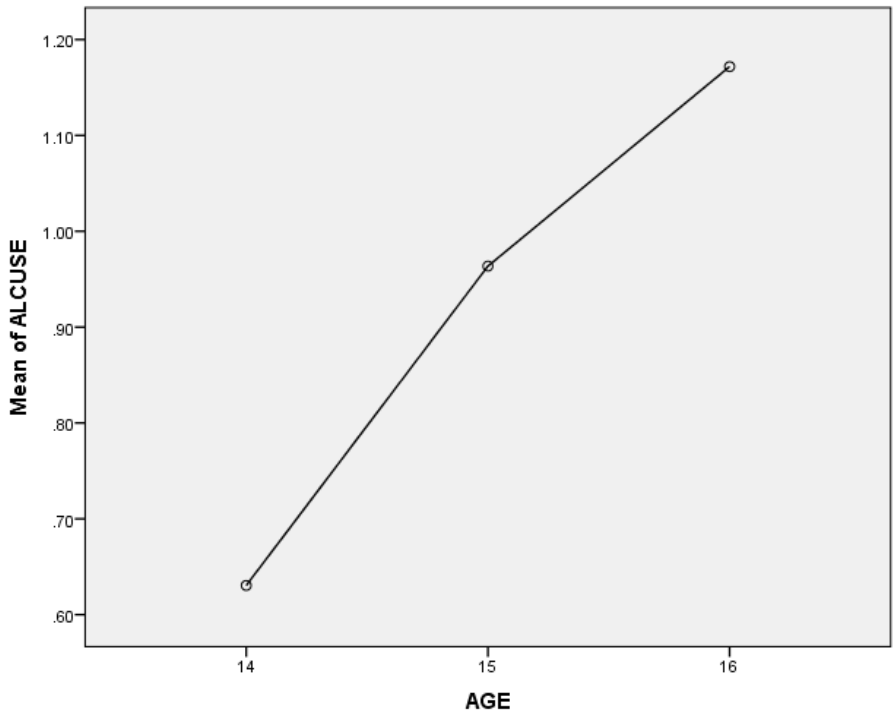
Tukey HSD[a]

| AGE | N | Subset for alpha = 0.05 | |
|---|---|---|---|
| | | 1 | 2 |
| 14 | 82 | .6305 | |
| 15 | 82 | .9636 | .9636 |
| 16 | 82 | | 1.1718 |
| Sig. | | .104 | .410 |

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 82.000.

**Means Plots**




```
SORT CASES BY AGE (A).
SORT CASES BY ALCUSE (A).
```

6d.

## Univariate Analysis of Variance

**Between-Subjects Factors**

| | | Value Label | N |
|---|---|---|---|
| AGE | 14 | | 82 |
| | 15 | | 82 |
| | 16 | | 82 |
| GENDER | 0 | females | 120 |
| | 1 | males | 126 |

**Descriptive Statistics**

Dependent Variable: ALCUSE

| AGE | GENDER | Mean | Std. Deviation | N |
|---|---|---|---|---|
| 14 | females | .6748 | .99726 | 40 |
| | males | .5883 | .88986 | 42 |
| | Total | .6305 | .93891 | 82 |
| 15 | females | 1.0086 | 1.00397 | 40 |
| | males | .9208 | 1.09754 | 42 |
| | Total | .9636 | 1.04737 | 82 |
| 16 | females | .9174 | .99146 | 40 |
| | males | 1.4141 | 1.21995 | 42 |
| | Total | 1.1718 | 1.13535 | 82 |
| Total | females | .8669 | .99922 | 120 |
| | males | .9744 | 1.12205 | 126 |
| | Total | .9220 | 1.06311 | 246 |

**Tests of Between-Subjects Effects**

Dependent Variable: ALCUSE

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Corrected Model | 17.592[a] | 5 | 3.518 | 3.256 | .007 | .064 |
| Intercept | 208.382 | 1 | 208.382 | 192.867 | .000 | .446 |
| AGE | 11.931 | 2 | 5.966 | 5.521 | .005 | .044 |
| GENDER | .710 | 1 | .710 | .657 | .418 | .003 |
| AGE * GENDER | 4.656 | 2 | 2.328 | 2.155 | .118 | .018 |
| Error | 259.307 | 240 | 1.080 | | | |
| Total | 486.000 | 246 | | | | |
| Corrected Total | 276.900 | 245 | | | | |

a. R Squared = .064 (Adjusted R Squared = .044)

## Estimated Marginal Means

**1. Grand Mean**

Dependent Variable: ALCUSE

| Mean | Std. Error | 95% Confidence Interval | |
|---|---|---|---|
| | | Lower Bound | Upper Bound |
| .921 | .066 | .790 | 1.051 |

## 2. AGE

**Estimates**

Dependent Variable: ALCUSE

| AGE | Mean | Std. Error | 95% Confidence Interval | |
|---|---|---|---|---|
| | | | Lower Bound | Upper Bound |
| 14 | .632 | .115 | .405 | .858 |
| 15 | .965 | .115 | .739 | 1.191 |
| 16 | 1.166 | .115 | .940 | 1.392 |

**Pairwise Comparisons**

Dependent Variable: ALCUSE

| (I) AGE | (J) AGE | Mean Difference (I-J) | Std. Error | Sig.[b] | 95% Confidence Interval for Difference[b] | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| 14 | 15 | -.333* | .162 | .041 | -.653 | -.013 |
| | 16 | -.534* | .162 | .001 | -.854 | -.214 |
| 15 | 14 | .333* | .162 | .041 | .013 | .653 |
| | 16 | -.201 | .162 | .217 | -.521 | .119 |
| 16 | 14 | .534* | .162 | .001 | .214 | .854 |
| | 15 | .201 | .162 | .217 | -.119 | .521 |

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

**Univariate Tests**

Dependent Variable: ALCUSE

| | Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Contrast | 11.931 | 2 | 5.966 | 5.521 | .005 | .044 |
| Error | 259.307 | 240 | 1.080 | | | |

The F tests the effect of AGE. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

## 3. GENDER

**Estimates**

Dependent Variable: ALCUSE

| GENDER | Mean | Std. Error | 95% Confidence Interval | |
|---|---|---|---|---|
| | | | Lower Bound | Upper Bound |
| females | .867 | .095 | .680 | 1.054 |
| males | .974 | .093 | .792 | 1.157 |

**Pairwise Comparisons**

Dependent Variable: ALCUSE

| (I) GENDER | (J) GENDER | Mean Difference (I-J) | Std. Error | Sig.[a] | 95% Confidence Interval for Difference[a] | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| females | males | -.107 | .133 | .418 | -.369 | .154 |
| males | females | .107 | .133 | .418 | -.154 | .369 |

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

**Univariate Tests**

Dependent Variable: ALCUSE

| | Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Contrast | .710 | 1 | .710 | .657 | .418 | .003 |
| Error | 259.307 | 240 | 1.080 | | | |

The F tests the effect of GENDER. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

## 4. AGE * GENDER

Dependent Variable: ALCUSE

| AGE | GENDER | Mean | Std. Error | 95% Confidence Interval | |
|---|---|---|---|---|---|
| | | | | Lower Bound | Upper Bound |
| 14 | females | .675 | .164 | .351 | .999 |
| | males | .588 | .160 | .272 | .904 |
| 15 | females | 1.009 | .164 | .685 | 1.332 |
| | males | .921 | .160 | .605 | 1.237 |
| 16 | females | .917 | .164 | .594 | 1.241 |
| | males | 1.414 | .160 | 1.098 | 1.730 |

**Profile Plots**