Midterm Exam MMI 409 Spring 2009 Gordon Bleil

Table of contents: (Hyperlinked to problem sections)

Problem 1

To investigate the effect of a new experimental drug on heart rates, a researcher randomly sampled 12 patients, and randomly assigned 6 patients to an experimental group and 6 patients to a placebo group. The drug therapy and placebo were administered. Two hours after the drug was administered, the number of beats per minute was recorded for each patient. The following results were obtained.

Placebo	Experimental Drug
63	81
66	76
85	81
69	87
72	72
71	84

Does the new experimental drug significantly affect heart rates?

Hypothesis:

Research Scenario:

From a larger population of patients, a small sample is divided randomly into 2 parts to compare the effects of a new drug against placebo. The question to be answered is whether the new experimental drug significantly affects heart rates.

Problem formulation:

This question requires the comparison of two test samples. There is no relationship between the measurements on the heart rate of any two patients, and no single patient is measured twice or given both drugs in order to compare effects.

Therefore they are independent samples.

Data Definitions/Comments:

Heart rate data are obtained two hours after the drug is taken and recorded as beats per minute.

Null Hypothesis: H_0: A □ B

There is no significant difference between the effects of the placebo and the experimental drug based on heart rate.

Alternative Hypothesis: H_1: A = B

There is a significant difference between the effects of the placebo and the experimental drug based on heart rate.

Statistical Procedure, Tests and Assumptions:

Test: T-Test; Independent Samples

Level of Significance: $\alpha = 0.05$

Assumptions: This is a normal population

Samples are independent

Patients are random selected

There is equality of variance

Results:

Group Statistics

	Therapy	N	Mean	Std. Deviation	Std. Error Mean
Heart Rate	Placebo	6	71.00	7.616	3.109
	Experimental Drug	6	80.17	5.419	2.212

Independent Samples Test

		t-test for Equality of Means		
		Sig. (2-tailed)	Mean Difference	Std. Error Difference
Heart Rate	Equal variances assumed	.037	-9.167	3.816

Independent Samples Test

		t-test for Equality of Means	
		95% Confidence Interval of the Difference	
		Lower	Upper
Heart Rate	Equal variances assumed	-17.669	-.664

Inference:

The null hypothesis is rejected, based on ρ=0.037 (<0.05). Therefore it there is a significant difference between the effects of the experimental drug and placebo, and **the new experimental drug significantly affects heart rates**. The mean heart rate for the experimental drug is higher. The specific goal of the drug was not defined; therefore no inference can be made regarding whether this information is helpful in determining the usefulness or safety of the drug. This is also a small sample, decreasing the likelihood of effective generalization to the larger population.

Problem 2

A researcher interested in how trauma impacts different parts of the brain, a random sample of five people were shown graphic images of trauma events. The following data represent levels of EEG activity for the five patients in two different locations of the brain. For each patient, EEG activity was recorded for each section of the brain.

Subject	EEG Location 1	EEG Location 2
1	3	5
2	4	8
3	2	3
4	4	5
5	7	9

Do the two locations show a difference in the level of EEG activity?

<u>Hypothesis:</u>

<u>Research Scenario</u>:

This is an experiment in which more than one measurement is made on randomly selected individuals with the goal of determining if the measurements vary significantly based on their location. The question to be answered is whether the two locations show a difference in the level of EEG activity.

<u>Problem formulation</u>:

This a single sample of patients each with two measurements based on the same environmental condition. The data are therefore paired based on their relationship (being measured on the same individual), but neither is completely dependent or independent of each other.

<u>Data Definitions/Comments</u>:

The measurements are made using the same basic device, located on standardized areas of the head, with the patient in the same basic environment. Measurements are provided in undefined levels between 2 and 9.

<u>Null Hypothesis</u>: H_0: A □ B

There is no significant difference between the measurements at Location 1 and Location 2.

<u>Alternative Hypothesis</u>: H_1: A = B

There is a significant difference between the measurements at Location 1 and Location 2.

<u>Statistical Procedure, Tests and Assumptions:</u>

<u>Test</u>: T-Test; Paired Samples <u>Level of Significance</u>: α = 0.05

Assumptions: This is a normal population   Data are random

Samples are not independent   There is equality of variance

Results:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	EEG_Loc_1	4.00	5	1.871	.837
	EEG_Loc_2	6.00	5	2.449	1.095

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	EEG_Loc_1 & EEG_Loc_2	5	.873	.053

Paired Samples Test

		Paired Differences		
		Mean	Std. Deviation	Std. Error Mean
Pair 1	EEG_Loc_1 - EEG_Loc_2	-2.000	1.225	.548

Paired Samples Test

		Paired Differences	
		95% Confidence Interval of the Difference	
		Lower	Upper
Pair 1	EEG_Loc_1 - EEG_Loc_2	-3.521	-.479

Paired Samples Test

		t	df	Sig. (2-tailed)
Pair 1	EEG_Loc_1 - EEG_Loc_2	-3.651	4	.022

Inference:

The null hypothesis is rejected as ρ=0.022 (<0.05). Therefore **the two locations show a difference in the level of EEG activity** based on the sample data. Of note, there is no significant correlation between the pairs, and although the mean value for Location 1 is less than that of Location 2, neither the clinical accuracy (whether it truly reflects the patient condition) or the clinical significance of this determined based on this information alone.

Problem 3

Two groups of children, one with attention deficit disorder (ADD) and a control group of children without ADD, were randomly given either a placebo or the drug Ritalin. A measure of activity was made on all the children with the results shown in the table below (higher numbers indicate more activity).

Drug Key: 1=Placebo, 2=Ritalin Group Key: 1=ADD, 2=Control The following data were recorded:

Subject	Drug	Group	Activity
1	1	1	90
2	1	1	88
3	1	1	95
4	1	2	60
5	1	2	62
6	1	2	66
7	2	1	72
8	2	1	70
9	2	1	64
10	2	2	86
11	2	2	86
12	2	2	82

How is activity level related to Ritalin and ADD?

Hypothesis:

Research Scenario:

There are two distinct populations of patients, each being given two distinct treatments, looking for the effect of medication as measured by the activity level of the patients. The question to be answered is how activity level is related to Ritalin and ADD.

Problem formulation:

There are two patient populations and two treatments. There are 4 groups that need to be compared, with the measurement of activity dependent on both diagnosis (Normal and Attention Deficit Disorder [ADD]) and drug (Placebo and Ritalin).

Data Definitions/Comments:

Null Hypothesis: H_0: A □ B

There is no significant difference in the activity level of the patients, whether or not they have ADD, and whether or not they receive Ritalin.

Alternative Hypothesis: H_1: A = B

There is a significant difference in the activity level of the patients, depending on whether they have ADD, or whether they receive Ritalin.

Statistical Procedure, Tests and Assumptions:

Test: Two Way Univariate Analysis of Variance   Level of Significance: $\alpha = 0.05$

Assumptions: This is a normal population   Data are random

Samples are not independent   There is equality of variance

Results:

Post hoc tests are not performed for Drug because there are fewer than three groups.

Post hoc tests are not performed for Group because there are fewer than three groups.

Descriptive Statistics

Dependent Variable: Activity

Drug	Group	Mean	Std. Deviation	N
Placebo	ADD	91.00	3.606	3
	Control	62.67	3.055	3
	Total	76.83	15.804	6
Ritalin	ADD	68.67	4.163	3
	Control	84.67	2.309	3
	Total	76.67	9.266	6
Total	ADD	79.83	12.719	6
	Control	73.67	12.291	6
	Total	76.75	12.352	12

Tests of Between-Subjects Effects

Dependent Variable:Activity

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1588.250[a]	3	529.417	47.059	.000
Intercept	70686.750	1	70686.750	6283.267	.000
Drug	.083	1	.083	.007	.934
Group	114.083	1	114.083	10.141	.013
Drug * Group	1474.083	1	1474.083	131.030	.000
Error	90.000	8	11.250		
Total	72365.000	12			
Corrected Total	1678.250	11			

Results:

Post hoc tests are not performed for Drug because there are fewer than three groups.

a. R Squared = .946 (Adjusted R Squared = .926)

Estimated Marginal Means

Group

Dependent Variable:Activity

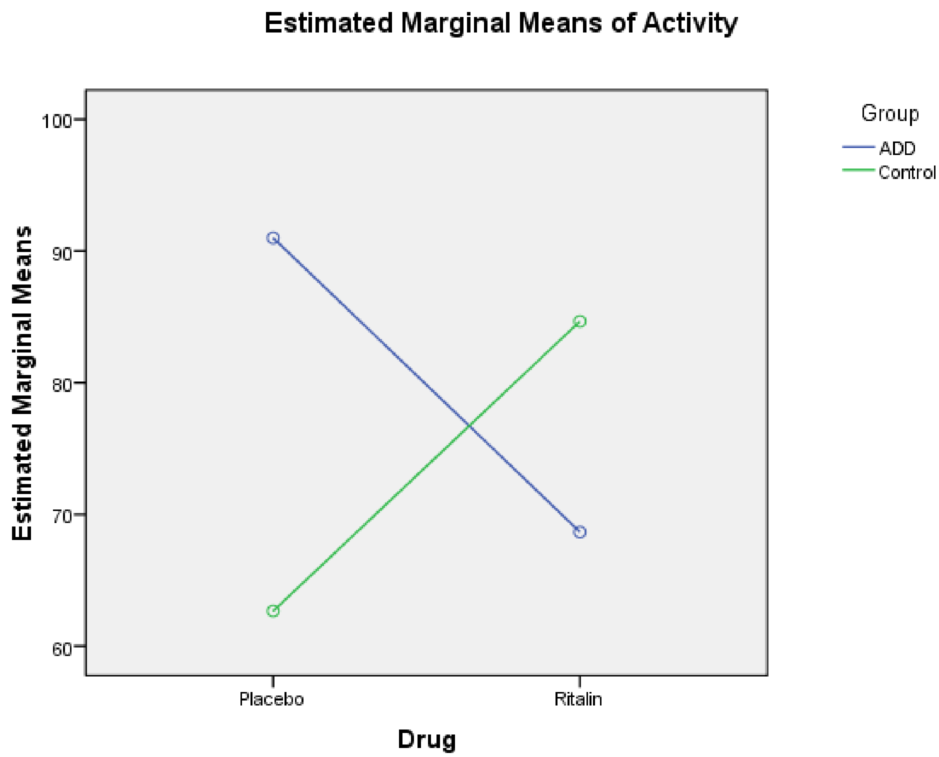
Group	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
ADD	79.833	1.369	76.676	82.991
Control	73.667	1.369	70.509	76.824

Drug * Group

Dependent Variable:Activity

Drug	Group	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Placebo	ADD	91.000	1.936	86.534	95.466
	Control	62.667	1.936	58.201	67.132
Ritalin	ADD	68.667	1.936	64.201	73.132
	Control	84.667	1.936	80.201	89.132

Profile Plots



Inference:

The corrected model ρ = .000 (<0.05), therefore the null hypothesis is rejected. There is a significant difference between the groups. Both the Group and the Group*Drug levels of significance are <0.05, therefore of significance. However, since the Group affects both of those values, and the Drug ρ>0.05, then only the Group difference is significant. This is consistent with the graph, which shows that the ADD marginal mean crosses the Normal mean when the Drug regimen is considered. **Activity level is related to Ritalin and ADD** in that patients with ADD show significantly lower levels of activity when given Ritalin in comparison to patients without ADD.

Problem 4

Patients of a large community clinic with large call volumes on average experience a 94 second wait time before the call is answered. In an effort to decrease wait time for patients, the clinic hired an additional phone operator. The following wait times were collected on ten randomly selected phone calls on the first day that the additional operator was used. The following wait times in seconds were recorded. 94, 56, 23, 37, 116, 105, 78, 86, 75, 121
Do the data suggest that the additional operator has resulted in lower wait times for patients?

Hypothesis:

Research Scenario:

The current average length of waiting time is known, and the clinic wants to know if the desired effect of lower average call waiting time was accomplished by the intervention of hiring an additional phone operator. The question is whether the data suggest that the additional operator has resulted in lower wait time for patients.

Problem formulation:

A sample of caller waiting times is collected to compare against a known standard. There is a single test sample, to be compared to a standard sample mean

Data Definitions/Comments:

Null Hypothesis: H_0: A □ B

There is no significant difference between the current average call waiting time of 94 seconds and the new mean call waiting time.

Alternative Hypothesis: H_1: A = B

There is a significant difference between the current average call waiting time of 94 seconds and the new mean call waiting time.

Statistical Procedure, Tests and Assumptions:

Test: T-Test; Single Sample

Level of Significance: $\alpha = 0.05$

Assumptions: This is a normal population

Data are random

Samples are not independent

There is equality of variance

Results:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Times	10	79.10	32.539	10.290

One-Sample Test

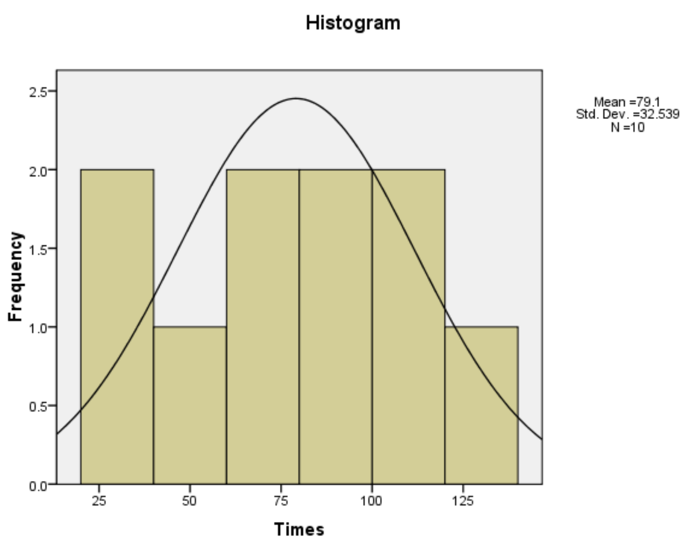
	Test Value = 94					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Times	-1.448	9	.182	-14.900	-38.18	8.38

Statistics

Times

N	Valid	10
	Missing	0
	Mean	79.10
	Median	82.00
	Mode	23[a]
	Std. Deviation	32.539
	Variance	1058.767

a. Multiple modes exist. The smallest value is shown



Inference:

The null hypothesis is <u>not</u> rejected as the two tailed ρ=.182 (>0.05). Therefore **the data do not suggest that the additional operator has resulted in a lower average phone call waiting time for patients**. Although the mean of 79 is 'apparently' different from 94, the wide variation and small sample size calculate to a large Standard Deviation 32, which is easily inclusive of the previous mean.

Problem 5

A researcher is interested in comparing the effect of breastfeeding and baby formula on the growth of newborn babies. 20 babies were randomly selected and randomly assigned to the following groups. After 14 days, the infants were weighed at the same time of day. The following results were recorded in ounces.

Group 1: Exclusively breastfed.	Group 2: Exclusively with Formula A.	Group 3: Exclusively with Formula B.	Group 4: Exclusively with Formula C.
112	98	103	104
116	106	111	103
121	102	101	101
112	115	102	95
118	101	97	115

Does the data suggest that that there is a difference among the groups?

Hypothesis:

Research Scenario:

This group of babies is selected from an unknown pool and randomly assigned to a feeding type. Data given are identified as change in weight in ounces. No characteristics of the infants were otherwise described (such as maturity, age [newborn usually includes the first 30 days of life], co-morbid conditions). The question to be answered is whether the data suggest that there is a difference among 4 groups of babies.

Problem formulation:

There are 4 different formulas being fed to a randomly chosen group of infants. The weights are each related to a baby, but there is no relationship between babies and only one weight gain measure is given for each baby. No baby is fed more than one formula.

Data Definitions/Comments:

The formulas tested are not described as being the only formulas available, therefore this is not a discrete set of samples.

Null Hypothesis: H_0: A □ B

There is no significant difference between the weights of babies fed different formulas after 14 days.

Alternative Hypothesis: H_1: A = B

There is a significant difference between the weights of babies fed different formulas after 14 days.

Statistical Procedure, Tests and Assumptions:

Test: One-way ANOVA

Level of Significance: $\alpha = 0.05$

Assumptions: This is a normal population

Data are random

Samples are not independent

There is equality of variance

Results:

Descriptives

Weights

		N	Mean	Std. Deviation	Std. Error
	1	5	115.80	3.899	1.744
	2	5	104.40	6.580	2.943
	3	5	102.80	5.119	2.289
	4	5	103.60	7.266	3.250
	Total	20	106.65	7.659	1.713
Model	Fixed Effects			5.863	1.311
	Random Effects				3.067

Descriptives

Weights

		95% Confidence Interval for Mean				
		Lower Bound	Upper Bound	Minimum	Maximum	Between-Component Variance
	1	110.96	120.64	112	121	
	2	96.23	112.57	98	115	
	3	96.44	109.16	97	111	
	4	94.58	112.62	95	115	
	Total	103.07	110.23	95	121	
Model	Fixed Effects	103.87	109.43			
	Random Effects	96.89	116.41			30.762

ANOVA

Weights

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	564.550	3	188.183	5.474	.009
Within Groups	550.000	16	34.375		
Total	1114.550	19			

Post Hoc Tests

Multiple Comparisons

Weights Tukey HSD

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
1	2	11.400*	3.708	.033	.79	22.01
	3	13.000*	3.708	.014	2.39	23.61
	4	12.200*	3.708	.022	1.59	22.81
2	1	-11.400*	3.708	.033	-22.01	-.79
	3	1.600	3.708	.972	-9.01	12.21
	4	.800	3.708	.996	-9.81	11.41
3	1	-13.000*	3.708	.014	-23.61	-2.39
	2	-1.600	3.708	.972	-12.21	9.01
	4	-.800	3.708	.996	-11.41	9.81
4	1	-12.200*	3.708	.022	-22.81	-1.59
	2	-.800	3.708	.996	-11.41	9.81
	3	.800	3.708	.996	-9.81	11.41

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

Homogeneous Subsets

Weights

Tukey HSD[a]

Group	N	Subset for alpha = 0.05	
		1	2
3	5	102.80	
4	5	103.60	
2	5	104.40	
1	5		115.80
Sig.		.972	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 5.000.

<u>Inference:</u>

Based on the ρ = .009, (<0.05) the null hypothesis is rejected – **there is a difference among the groups** of baby weight gains based on the formula fed. In Post Hoc analysis, using the multiple comparisons table, it is evident that there is a significant difference between Group 1 and each of the other Groups, but no difference between Groups 2,3 & 4, consistent with the Homogenous Subsets table which shows this same grouping. The mean weight gain for Group 1 is greater than that of the other set of Groups, indicating that Breast Feeding results in significantly greater weight gain than any of the other formulas tests.