Name ______________________________
(Please ALSO write your name on the back of the last page.)

Ag 400 - Quiz 2
Makeup - Fall 2006

(20) A. Suppose that you are employed by the Mudville Nuttle Factory (which, as you must know, is a plant that manufactures nuttles). The company has been studying the nuttle production per worker for many years and, on the bases of these studies, has found that the mean number of nuttles produced per worker per day is normally distributed with a mean of 280 and a standard deviation of 40. Assume that these figures are parameters. Answer the following questions in terms of these data. Show your work.

1. What is the probability of selecting a random sample of 64 workers and obtaining a sample with a mean production per worker per day of 270 or fewer nuttles?

$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{N}} = \frac{270 - 280}{40/\sqrt{64}} = -2.00$

$p = .5000 - .4773 = .0227$

2. What proportion of all workers in the factory would produce more than 250 nuttles per day?

$Z = \frac{250 - 280}{40} = -.75$

$p = .5 + .2734 = .7734$

3. What proportion of all possible random samples of 100 workers that you might draw at random from this population would be expected to have mean scores between 270 and 290 nuttles per day?

$Z = \frac{270 - 280}{40/\sqrt{100}} = -2.50$

$Z = \frac{290 - 280}{40/\sqrt{100}} = +2.50$

$.4938 + .4938 = .9876$

4. How many workers in a random sample of 64 workers drawn at random from this population would you expect to produce between 280 and 290 nuttles per day?

$Z = \frac{290 - 280}{40} = \frac{10}{40} = .25$

$p = .0987 \times 64 = 6.3168$

5. What is the probability of selecting a worker at random from this population and obtaining one who produces between 200 and 220 nuttles per day?

$Z = \frac{200 - 280}{40} = -2.00$

$Z = \frac{220 - 280}{40} = -1.50$

$p = .4772 - .4332 = .0440$

B. Suppose that management at the Mudville Nuttle Factory introduces a new incentive program where workers are paid a bonus for every 100 nuttles over 280 they produce in a day. To determine whether the new incentive program has affected worker productivity, you draw a random sample of 30 workers in the factory who are working under the incentive program and compile the following data.

Number of workers sampled	30
Mean number of nuttles produced by the sampled workers per day (Productivity Score)	300
Total number of nuttles produced by the 30 workers in the sample	9,000
Sum of the squares of the number of nuttles produced by the 30 workers	2,751,156
Sum of the squares of the deviations of the individual Productivity Scores from the mean of the sample	51,156
Estimate of the population standard deviation	42

(20) 1. Test the statistical significance of the difference between the mean Productivity Score for workers before the incentive program (see Section A above) and the mean Productivity Score per worker for the sample of workers participating in the incentive program. Use a 1-tailed test and the .05 level to determine statistical significance.

a. State the Null Hypothesis and the Alternative Hypothesis.

$H_0: \mu_{After} = 280$

$H_A: \mu_{After} > 280$

b. Show the necessary calculations.

$$Z = \frac{300 - 280}{40/\sqrt{30}} = -2.74$$

c. Specify degrees of freedom, if appropriate, and indicate the critical value of the test statistic at the.05 level.

Not appropriate for Z-test

d. Reject or do not reject the Null Hypothesis.

Reject H_0

e. What is the probability that you have made a Type I error in (d) above?

$\alpha = .5000 - .4969 = \underline{\underline{.0031}}$

f. What is the probability that you have made a Type II error in (d) above?

$\beta = 0$

g. State your conclusion precisely in terms of the problem.

Productivity Scores are higher for workers on the incentive program

(10) 2. Indicate whether each of the following statements is true (T) or false (F) in terms of these data. If any part of a statement is untrue, it should be marked false (F). Add comments, if you wish, to clarify your answers.

_F_ a. The statistical unit is a nuttle.

_F_ b. The above analysis tests the significance of the difference between two independent sample means.

_F_ c. The figures presented in B above are parameters used to describe the productivity of workers who participated in the incentive program.

F d. A one-tailed test is justified in the above analysis because the mean number of nuttles produced by the 30 workers described above is greater than the 280 reported in Section A of this test.

T e. The best point estimate for the mean number of nuttles produced per workers for those who participated in the incentive program is 300 nuttles per day. The point estimate for the standard deviation for this distribution is 42.

(4) 3. Set up a 99% confidence interval to estimate the mean number of nuttles produced per worker per day by workers participating in the incentive program.

$$300 \pm t_{.01}\left(\frac{42}{\sqrt{30}}\right)$$

$$300 \pm (2.756)(7.668)$$

$$300 \pm 21.13$$

C. Suppose that a manufacturer wanted to compare the wearing quality of two different types of automobile tires. To make the comparison, a tire of Type A and one of Type B were mounted on the rear wheels of each of 8 automobiles, with Type A randomly assigned to the left rear on 4 of the automobiles and Type B assigned to the left rear of the other 4 automobiles. The automobiles were then operated for a specified number of miles and the amount of wear war recorded. These data are given below:

		Type A	Type B		
	Auto	------Amount of Wear------		D	D^2
	1	10	8	2	4
	2	8	7	1	1
	3	9	9	0	0
	4	7	5	2	4
	5	5	6	-1	1
	6	6	3	3	9
	7	4	6	-2	4
	8	7	4	3	9
				8	32
Mean		7.0	6.0		
Sum of the squares of the deviations about the mean	$\bar{D} = \frac{8}{8} = 1.0$	28	28		
Estimate of the population variance	$\Sigma d^2 = 32 - \frac{(8)^2}{8} = 24$	4	4		

(6) 1. Test the statistical significance of the difference between these two types of tires in regard to the amount of wear evidenced. Use a two-tailed test and the .05 level to determine significance.

$$t = \frac{1.0}{\sqrt{\frac{24}{(8)(7)}}} = \frac{1.}{.654654} = 1.53 \quad n.s. \ .05$$

(2) 2. Make your conclusion precisely in terms of the problem.

We cannot conc. that Type A and Type B tires differ in the amount of wear observed

(10) 3. Indicate whether each of the following statements is true (T) or false (F) in terms of these data. If any part of a statement is untrue, it should be marked false (F). Add comments, if you wish, to clarify your answers.

F a. The Null Hypothesis for the above test could be stated as follows:

H_o: There is no relationship between the mean amount of wear on Type A tires and Type B tires.

T b. If the Null Hypothesis is not rejected for the above analysis, you could not make a Type I error.

T c. If the Null Hypothesis is rejected in the above analysis, you cannot make a Type II error.

F d. Since the estimate of the population variance in amount of tire wear is the same for Type A and Type B tires in these data, you could use to the pooled formula t-test for the difference between independent sample means to test the statistical significance of the difference between the two types of tires in regard to wearability.

T e. The independent variable, Type of tire, is measured by a two-category nominal scale; the dependent variable is measured by an interval scale.

(28) E. Suppose you were involved in a study of a random sample of Mudville residents concerning their levels of expressed happiness with their current lives. Happiness was measured by a series of attitude questions that yields an interval scale. The higher the Happiness Score, the happier the individual indicated he/she was. Each subject reported his/her happiness at two points in time–once at the beginning of September and once at the end of September.

The following variables were used:

Gender of subject (gender)

1 Male
2 Female

Expressed Happiness at the beginning of September (Happy1)

Expressed Happiness at the end of September (Happy2)

The attached output reports analysis of these data:
Page 1 uses data from the entire sample.
Page 2 reports only data for the males in the sample.
Page 3 reports only data for the females in the sample.

Answer the following by indicating whether each statement is true (T) or false (F). If any part of a statement is untrue, it should be marked false (F). Add comments, if you wish, to clarify your answers.

T 1. The statistical unit is a Mudville resident.

T 2. In this sample females have higher mean happiness scores at the beginning of September than do males, and this difference is statistically significant at the.05 level.

T 3. The standard deviations on Page 1 are estimates of the standard deviations in happiness scores for males and for females in the population.

F 4. Based on this output on Page 1 you can conclude that the variances in Happiness Scores at the beginning of September are the same for males and females in the sample.

T 5. If you reject the following Null Hypothesis in terms of the following Alternative Hypothesis, the probability that you have made a Type I error is .032.

$H_{o:}$ there is no difference in the population between males and females in regard to mean happiness scores at the beginning of September.

H_A: there is a difference between the mean happiness scores of males and females in the population at the beginning of September.

F 6. Levene's Test tests whether the data meet the Homogeneity of Variance assumption. Since the results of Levene's Test are not statistically significant at the .05 level, we conclude that the variances in the Happiness Scores for males and females reported on Page 1 are not homogeneous.

F 7. There are two t-values and two "df" values reported on the output on Page 1. For this analysis you should use t = -2.143 and df = 180.22 since the Homogeneity of Variance assumption is not true.

F 8. Based on the analysis on Page 2 of the attached output, you can conclude that the Happiness Scores of males in the sample are the same at the beginning of September as at the end of September.

F 9. Based on the analysis on Page 2, if you do not reject the following Null Hypothesis using a two tailed test, $\alpha = 0.00$; $\beta = \underline{.363}$.

$H_{o:}$ there is no difference between males' happiness scores at the beginning of September and males' happiness scores at the end of September.

F 10. Males in the sample have greater dispersion in happiness scores at the beginning of September than at the end of September, but this difference is shown to be not statistically significant by this output.

T 11. In the test reported on Page 2 of the output, the number of Happy1 scores is the same as the number of Happy2 scores. In this type of test, these two "Ns" for the two scores will always be the same.

T 12. The t-test presented on Page 2 tests the statistical significance of the difference between means for paired or matched data.

F 13. The total number of males in the sample for the analysis reported on Page 2 is 180 (90 in the Happy1 category and 90 in the Happy 2 category).

F 14. Based on the analysis on Page 3, you would reject following Null Hypothesis using a 2-tailed test ($\alpha = .018$) .

$H_{o:}$ there is no relationship between Happiness Scores at the beginning of September and Happiness Scores at the end of September for females.

```
SPLIT FILE
  OFF.
T-TEST
  GROUPS = gender(1 2)
  /MISSING = ANALYSIS
  /VARIABLES = happy1
  /CRITERIA = CI(.95) .
```

T-Test

[DataSet1] Y:\Willits\AAATeaching\AATeaching\Ag400\Quizzes400\Q2m_06.sav

Group Statistics

	gender	N	Mean	Std. Deviation	Std. Error Mean
happy1	1 Male	90	11.7889	3.77304	.39771
	2 Female	108	12.8889	3.37205	.32448

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
happy1	Equal variances assumed	2.657	.105	-2.165	196	.032	-1.10000	.50806	-2.10197	-.09803
	Equal variances not assumed			-2.143	180.422	.033	-1.10000	.51328	-2.11281	-.08719

```
SORT CASES BY gender .
SPLIT FILE
  SEPARATE BY gender .
T-TEST
  PAIRS = happy1  WITH happy2 (PAIRED)
  /CRITERIA = CI(.95)
  /MISSING = ANALYSIS.
```

T-Test

[DataSet1] Y:\Willits\AAATeaching\AATeaching\Ag400\Quizzes400\Q2m_06.sav

gender = 1

Paired Samples Statistics[a]

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	happy1	11.7889	90	3.77304	.39771
	happy2	11.5587	90	1.66401	.17540

a. gender = 1 Male

Paired Samples Correlations[a]

		N	Correlation	Sig.
Pair 1	happy1 & happy2	90	.899	.000

a. gender = 1 Male

Paired Samples Test[a]

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	happy1 - happy2	.23016	2.39024	.25195	-.27047	.73078	.913	89	.363

a. gender = 1 Male

gender = 2

Paired Samples Statistics[a]

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	happy1	12.8889	108	3.37205	.32448
	happy2	13.3326	108	1.63334	.15717

a. gender = 2 Female

Paired Samples Correlations[a]

		N	Correlation	Sig.
Pair 1	happy1 & happy2	108	.940	.000

a. gender = 2 Female

Paired Samples Test[a]

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	happy1 - happy2	-.44370	1.91848	.18461	-.80966	-.07774	-2.404	107	.018

a. gender = 2 Female