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

AG 400 - Quiz 2
Fall 2006

NOTE: Unless otherwise noted, use the .05 level to determine statistical significance.

(24) A. Indicate whether each of the following statements 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. Sample means and standard deviations are descriptive statistics; t-values and z-values are inferential statistics.

T 2. If an intervally scaled variable is normally distributed, half of the cases have scores that are greater than the mean.

F 3. In making a statistical decision concerning whether an observed sample difference is "statistically significant," you will either make a Type I Error or a Type II Error. You can never make both a Type I and a Type II at the same time.

T 4. The t-distribution is symmetrical and unimodal, but more playtkurtic than a Normal Distribution when the degrees of freedom are small.

F 5. If a t-test for the difference between two sample means yields a t-value of 1.30 with 24 degrees of freedom, you should conclude that the two sample means are the same.

T 6. $\hat{\sigma}^2$, calculated using the following formula, is a point estimate for the variance of X in the population.

$$\hat{\sigma}^2 = \frac{\Sigma\left(X - \overline{X}\right)^2}{N - 1}$$

F 7. If in the population, an intervally scaled variable (X) has a mean (μ) of 50 and a standard deviation (σ) of 8, a z-test for the difference between 50 and the mean of a sample of 100 cases should not be used if the distribution of X in the population is skewed to either the right or the left.

F 8. The Null Hypothesis for a two tailed one sample t-test for the difference between a sample mean and a given population mean can be stated as follows:

> H_o: The mean for the sample is the same as the mean for the population.

F 9. If a calculated z-value is statistically significant at the .05 level using a one-tailed test, it would also be statistically significant at the .05 level using a two tailed test.

T 10. If the Null Hypothesis is true and you reject it, you have made a Type I Error.

T 11. A 99% confidence interval is *always* wider than a 95% Confidence Interval for the same data.

F 12. A 95% Confidence Interval estimating the mean contains 95% of the population means.

(20) B. The number of times (X) an adult breathes per minute when at rest varies from person to person. Suppose that the distribution of number of breaths/per minute for adults is normally distributed with a mean of 16 and a standard deviation of 4. Assume that these figures are parameters. Answer the following in terms of this information. $\mu = 16 \quad \sigma = 4$

1. What proportion of adults in the population breathe 22 times per minute or more?

$$Z = \frac{22-16}{4} = \frac{6}{4} = 1.25$$

$$\begin{array}{r} .5000 \\ -\ .4332 \\ \hline .0668 \end{array}$$

2. What is the probability of selecting a sample of 36 adults at random and obtaining a sample with a mean number of breaths per person of 17.5 or less?

$$Z = \frac{17.5-16}{4/\sqrt{36}} = \frac{1.5}{.6667} = 2.25$$

$$\begin{array}{r} .5000 \\ .4878 \\ \hline .9878 \end{array}$$

3. How many adults in a random sample of 25 adults would you expect to breathe 14 or more times per minute?

$$Z = \frac{14-16}{4} = \frac{-2}{4} = .5$$

$$\begin{array}{r} .5000 \\ +\ .1915 \\ \hline .6915 \end{array} \times 25 = 17.3$$

4. What is the probability of selecting an adult at random from this population and obtaining an individual whose breathing rate is between 20 and 24 breaths per minute?

$$Z = \frac{20-16}{4} = 1.00$$

$$Z = \frac{24-16}{4} = 2.00$$

$$\begin{array}{r} .4773 \\ -\ .3413 \\ \hline .1360 \end{array}$$

5. What proportion of all possible random samples of 36 adults drawn from this population would have breathing rates of between 15 and 18 breaths per minute.

$$Z = \frac{15-16}{4/6} = 1.50$$

$$Z = \frac{18-16}{4/6} = 3.00$$

$$\begin{array}{r} .4332 \\ .49865 \\ \hline .93185 \end{array}$$

C. To study the effect of noise level on people's ability to concentrate on a task, a sample of 24 students were randomly assigned to one of two groups. Each individual in the first group was placed in a quiet, well-lighted cubicle and given a puzzle to solve. Each person in the second group was also placed in a well-lighted cubicle and given the same puzzle to solve, but the sound of people talking was continuously broadcast into the setting.

The time required for each person working alone to solve the puzzle was recorded. The following data were obtained.

	Quiet Setting	Noisy Setting
	--------minutes needed to solve puzzle ----	
	4	2
	4	2
	3	2
	3	1
	5	1
	7	7
	5	3
	4	5
	5	3
	7	2
	9	5
	4	3
Number of cases	12	12
Sum of the scores (total # of minutes)	60	36
Sum of the squares of the scores	336	144
Sum of the squares of the deviations about the sample mean	36	36

(20) 1. Test the statistical significance of the difference between the mean time required to solve the puzzle in the quiet setting and the mean time required to solve the puzzle in the noisy setting. Use the .05 level and a two tailed test to determine statistical significance. Specify the following:

a. State the Null Hypothesis and the Alternative Hypothesis in words; not statistical symbols.

H_0: the time required to solve the puzzle is the SAME in both the quiet + the noisy settings

H_A: . . . . not the same . . .

b. Report the formula you would use and show your calculations.

$$t = \frac{5-3}{\sqrt{\left[\frac{36+36}{12+12-2}\right]\left[\frac{1}{12}+\frac{1}{12}\right]}} = \frac{2}{\sqrt{(3.272)(.1667)}} = \frac{2}{\sqrt{.5454}} = \frac{2}{.7385} =$$

2.71

$$\cancel{t = \frac{5-3}{\sqrt{\frac{36}{12}}}}$$

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

22 $\quad t_{.05} = 2.074$

d. Reject or do not reject the Null Hypothesis.

Reject

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

$.01 < \alpha < .02$

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

$\beta = 0$

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

It takes more time to solve the puzzle in a quiet setting than in a noisey one

(12) 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 want to clarify your answers.

F a. The statistical unit here is a minute.

T b. Based on the above data, you have no reason to question the Homogeneity of Variance assumption.

F c. The figures presented above are parameters.

F d. In the above analysis, a one- tailed test would be justified because one of the sample means is larger than the other sample mean.

F e. The test used to test the significance of the difference between the above means is a One Sample t-test since the data are obtained from a single sample.

F f. The Null Hypothesis for the test in (1) above can be stated as follows:

H_o: there is no difference between the mean number of minutes needed to solve the puzzle by the 12 persons in a quiet setting and the mean number of minutes needed to solve the puzzle by the 12 persons in the noisy setting.

(4) 3. Set up a 95% Confidence Interval estimating the mean number of minutes needed to solve the puzzle in the Quiet Setting for the population. Show your work.

$\hat{\sigma}^2 = \frac{36}{11} = 3.2727$

$5 \pm (2.201)(.5222)$ $\qquad \hat{\sigma} = 1.81$

$\hat{\sigma}_{\bar{X}} = \sqrt{\frac{3.2727}{12}}$

5 ± 1.15

$\hat{\sigma}^2_{\bar{X}} = .2727$

$\hat{\sigma}_{\bar{X}} = .5222$

(20) D. A survey of random samples of small (less than 100 acres) farms in Pennsylvania and Georgia provided data on farm income. The data file contains the following variables:

ID = Identification number

Location (state)
1 Pennsylvania
2 Georgia

Farm income in $1000 units (income)

The attached output was obtained using this data file. Answer the following by indicating whether each statement 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.

T 1. The statistical unit here is a farm.

F 2. Based on the above data, you should question the validity of the Homogeneity of Variance assumption for these data and hence you should use the "Equal variances not assumed" t-test.

F 3. The mean income of the Pennsylvania farms in the sample is greater than the mean income of Georgia farms, but this difference is not statistically significant at the .05 level using a two-tailed test.

T 4. On this output, degrees of freedom for the "Equal variances not assumed' test are less than the degrees of freedom for the "Equal variances assumed" test. Degrees of freedom for the "Equal variances not assumed' test are never greater than the degrees of freedom for the "Equal variances assumed" test.

F 5. In this sample, the dispersion in farm incomes is greater for the Pennsylvania farms than for the Georgia farms, and this difference is statistically significant at the .05 level.

F 6. If there were 100 farms in the Pennsylvania sample and 100 farms in the Georgia sample the statistical significance of the difference between the farm incomes of small farmers in Pennsylvania and Georgia could have been tested using a paired t-test.

T 7. If you reject the following Null Hypothesis in favor of the following Alternative Hypothesis, you could be wrong (α = .014)

> H_o: the mean incomes of small farms in Pennsylvania and Georgia are the same.
>
> H_A: the mean incomes of small farms in Pennsylvania and Georgia are different.

T 8. Suppose that past studies of small farms in Georgia and Pennsylvania have consistently found that small farms in Georgia have lower incomes than Pennsylvania farms, and you believe a one-tailed test of the difference would be appropriate. Using a one-tailed test and these data, you could reject the Null Hypothesis of no difference in farm incomes at the .01 significance level.

F 9. Based on Levene's Test for Equality of Variances you should conclude that the variances in Pennsylvania and Georgia small farm incomes are not the same.

F 10. The Null Hypothesis for testing the statistical significance of the relationship between "state" and "income" as shown on this output could be stated as follows:

> H_o: There is no relationship between the mean income of small farms in Pennsylvania and the mean income of small farms in Georgia.

```
T-TEST
  GROUPS = state(1 2)
  /MISSING = ANALYSIS
  /VARIABLES = income
  /CRITERIA = CI(.95) .
```

T-Test

Group Statistics

	state	N	Mean	Std. Deviation	Std. Error Mean
income	1 Pennsylvania	104	41.3462	4.95762	.48614
	2 Georgia	100	39.4800	5.73925	.57392

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
income	Equal variances assumed	2.492	.116	2.488	202	.014	1.86615	.74999	.38734	3.34497
	Equal variances not assumed			2.481	195.361	.014	1.86615	.75214	.38279	3.34951