STA 240/ENV 298.01 Midterm  Name:______________________
October 9, 2001

READ THESE DIRECTIONS:

1. This test is CLOSED BOOK; however, you may use one 8.5” × 11” sheet (front and back)
of notes.

2. With this exam, we will provide tables of t and z.

3. You may use a scientific calculator for this exam.

4. To receive partial credit, you must show all steps.

5. You have 1 hour and 15 minutes to complete the exam; you must finish by the end of the period.

Duke University
Nicholas School of the Environment

Honor Code

I HAVE NOT GIVEN OR RECEIVED, NOR WILL I GIVE, INFORMATION BEFORE, DURING OR AFTER THIS EXAM RELATING TO ITS CONTENTS OR COMPLETION THAT WOULD CONFER AN ADVANTAGE OVER STUDENTS NOT PRIVY TO THIS INFORMATION, UNLESS SPECIFICALLY TOLD BY THE INSTRUCTOR THAT SUCH CONDUCT IS PERMISSIBLE. I HAVE NOT USED DESIGNATED STUDY MATERIALS FOR THIS COURSE IN A MANNER THAT WOULD PREVENT THEM FROM BEING EQUALLY AVAILABLE TO OTHER STUDENTS. I WILL REPORT ANY VIOLATIONS OF THIS CODE TO THE INSTRUCTOR IN CHARGE OF THE COURSE AND TO THE DEAN.

Signature: ________________________  Date:______________________
1. If I perform an analysis and obtain a $p$-value of 0.05, then which of the following are true: **Circle all that apply. [5 points]**

(a) There is a 5% probability that my data occurred by chance.
(b) If I evaluate my $p$-value relative to a significance level of $\alpha=0.045$, I will have sufficient evidence to reject the null hypothesis.
(c) The $p$-value depends on the sample size.
(d) Under the null hypothesis, there is a 5% probability of obtaining a result at least as extreme as my result.
(e) The Type II error in this case is 0.95.

2. Monitors A and B measure hourly levels of carbon monoxide (CO) and are co-located along a heavily travelled highway. Their measurements are identical, except Monitor A’s readings are censored at 20 ppm. That is, any measured CO value equal to or greater than 20 ppm is recorded as 20 ppm by Monitor A.

We have 100 measurements each from Monitors A and B. In this 100 hour period, 2 of the hourly values of CO were equal to or greater than 20 ppm; these two values were both 37 ppm. Which values will be the same for the two monitors? **Circle all that apply. [5 points]**

(a) Medians
(b) Standard deviations
(c) Inter-quartile ranges
(d) Confidence intervals for the mean CO level for each monitor
(e) The sum of the ranks for each monitor’s measurements (assume tied values are given the average rank)

3. The nicotine content in a single cigarette of a particular brand has a normal distribution with mean 1.0 mg and known standard deviation 0.5 mg.

(a) **Approximately** how likely is it that a single cigarette will have a nicotine content less than 0.8 mg? **[5 points]**

(i) 0%  
(ii) 5% to 6%  
(iii) 33% - 34%  
(iv) 66%  
(v) 100%

(b) If 100 of these cigarettes are analyzed, what is the approximate probability that the resulting sample mean nicotine content will be less than 0.8 mg? **[5 points]**

(i) 0%  
(ii) 5% to 6%  
(iii) 33% - 34%  
(iv) 66%  
(v) 100%
4. In a study of the lizard *Sceloporis occidentalis*, researchers obtain a random sample of lizards from a field and test them for infection by the malarial parasite *Plasmodium*. To help assess the ecological impact of malarial infection, the researchers plan to test infected and non-infected lizards for stamina, as indicated by the distance (in meters) each animal can run in 15 minutes.

Let $X$ be the distance in meters run by an infected lizard. Let $Y$ be the distance in meters run by an uninfected lizard. A pilot study yielded $s_{pooled} = 10$ meters.

Researchers will apply a two-sample t-test to investigate the claim that infected lizards have less stamina than uninfected lizards. Assume a sample of 15 lizards in each group. Problems (c) through (f) deal with a power analysis of the proposed hypothesis test. Assume $s_{pooled}$ is known and equal to the value found in the pilot study.

(a) What are the assumptions of the proposed two-sample t-test? [5 points]

(b) The researchers choose a rejection region for their test, which is $\bar{Y} - \bar{X} > 8.51$ meters. Give the $\alpha$ level associated with this rejection region. [5 points]

(c) Using the same rejection region given in (b), find the Type II error for the test of the claim above if the true difference in stamina between the uninfected and infected groups is 9 meters. [10 points]
Again assume that the true difference in stamina between the uninfected and infected groups is 9 meters. For problems (d) and (e), consider the following two rejection regions:

**A:** \( \bar{Y} - \bar{X} > 8.51 \) meters \hspace{1cm} **B:** \( \bar{Y} - \bar{X} > 7.00 \) meters

(d) **True or False. Circle one.** [5 points] Region A is more desirable if we want to minimize Type I error.

(e) **True or False. Circle one.** [5 points] Region B is more desirable if we want to maximize power.

(f) The above figure corresponds to a power analysis of the test described above, with rejection region A and \( s_{pooled} \) fixed at 10 meters. The “effect size” we are measuring here is the difference in stamina, as measured by distance run by uninfected minus the distance run by infected lizards. Which of the following are true based on this figure? **Circle all that apply.** [5 points]

i. For a sample size of 10 per group and a maximum allowable Type II error of 20%, a difference in stamina between uninfected and infected groups of 13 meters or greater would be desirable.

ii. A sample size of 20 per group and a difference in stamina of 15 meters implies that \( \alpha = 0.05 \), since power would be nearly 95%.

iii. As \( s_{pooled} \) decreases, we can expect the curve corresponding to a stamina difference of 15 to flatten out and approach 1 more slowly for increasing values of \( n \).

iv. Doubling the sample size per group from 10 to 20 has a greater impact on power when the true stamina difference is 13 meters than when it is 9 meters.

v. If instead the plot were drawn for rejection region B, we would expect that all of the curves would be shifted down, reflecting lower power at each level of \( n \).
(g) **True or False. Circle one. [5 points]** (continued from prev. pg.) If the proposed analysis shows convincing evidence that the mean stamina levels were lower in the infected group, we can conclude that infection causes reduction in stamina since a random sample of lizards is taken.

5. The following situations all require inference about a mean or means. Identify each as (1) a single sample, (2) matched pairs or (3) two independent samples. [8 points]

(a) ______ Researchers use two methods to count the number of seeds in soil samples. For each of 27 soil specimens, the number of seeds was measured using Method A as well as Method B. Do the data provide evidence to conclude that the mean number of seeds detected differs for the two methods?

(b) ______ 322 subjects were selected at random from a group that successfully completed a quit-smoking program. The subjects were weighed at the beginning of the program and again one year later. Is there evidence to suggest that the true mean change in weight is positive?

(c) ______ Urinary fluoride concentrations were measured for 50 cows of a particular species that had been grazing in an area exposed to fluoride pollution. Urinary fluoride concentrations were also measured for 50 cows of the same species that had grazed in an adjacent unpolluted region. Is there evidence that the true average fluoride concentration for livestock grazing in the polluted region is larger than that for the unpolluted region?

(d) ______ The Environmental Protection Agency takes 27 soil samples in order to test whether the mean amount of radium-226 in soil in a Florida county exceeds the maximum allowable amount, 4 pCi/L. Is there evidence that the radium levels in the county are out of compliance?

6. **True or False. Circle one and explain why. [7 points]** The figure above shows measurements for two independent samples. A log transform would allow the two sample t-test to be applied to this dataset because after the transformation, spreads in the 2 groups would be approximately equal. In a sentence or two, explain your answer.
7. A randomized experiment was performed on a group of 24 female volunteers to compare how two forms of iron supplements, supplement A (Fe\(^{2+}\)) or supplement B (Fe\(^{4+}\)), were retained by the body. Twelve women received supplement A and twelve women received supplement B, and the amount of iron remaining in the body after a fixed time was measured (in percent) using a sophisticated test.

A researcher log-transformed these data and performed a two-sample t-test on the transformed data to determine whether median percent of supplement B retained was different from the median percent of supplement A retained. Using Splus, the following partial output was produced. Note for these log-transformed data, \(s_{pooled} = 0.53\).

**Standard Two-Sample t-Test**

```
data: log.SuppA and log.SuppB
t = _____, df = _____, p-value = _____
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  0.15      1.05
sample estimates:
mean of log.SuppA  mean of log.SuppB
   1.2       1.8
```

(a) Give the value of the test statistic and degrees of freedom. Using the tables, give an upper and lower bound for the \(p\)-value. [9 points]
(b) Which of the following are true? *Circle all that apply.* [5 points]

i. The median percent retained for supplement B is estimated to be approximately 1.8 times the median retained for supplement A.

ii. Because 0 is not included in the interval [1.16,2.86], we have sufficient evidence to reject the null hypothesis and conclude there is a difference in median percent retention for the two iron supplements.

iii. 95% of the values of $\frac{\text{median percent retained for supplement B}}{\text{median percent retained for supplement A}}$ are contained in the interval [1.16,2.86].

iv. The median percent of supplement A retained is estimated to be slightly more than half (.55) the median percent of supplement B retained.

v. With 95% confidence, it is estimated that median percent of supplement A retained is between 35% and 86% of the median percent of supplement B retained.

(c) Researcher A analyzes the same data, but instead performs a Wilcoxon rank sum test on the untransformed data. What assumptions does Researcher A make? [3 points]

Researcher B performs a Wilcoxon rank sum test on the log-transformed data.

(d) True or False. *Circle one.* [3 points] Researchers A and B will compute the same test statistics and make the same conclusion from their tests.

8. Researchers wish to address the question of whether the density of a plant species in a fertilized area is considerably larger than that in an non-fertilized area using a two sample t-test. A random sample of 1000 measurements of plant density is taken in each area. A statistically significant difference ($p$-value = 0.001) is detected. Can we always conclude that the fertilized areas have “significantly” larger plant densities than non-fertilized areas? Explain your answer. [5 points]