

BIOS 600.001 Exam 2

October 30, 2012

Honor Statement. I pledge that I have not used any reference materials (including electronic materials) during this examination. I pledge that I have neither given nor received any aid from any other person during this examination, and that the work presented here is entirely my own. I furthermore pledge that I will not reveal any of the material on this examination, either in the form of the exact question or the topics covered, to any person for any reason until after exams have been returned. I pledge that I will report all Honor Code violations observed by me. I understand that if I have committed any of the above, I have violated the UNC Honor Code.

Name: _____

Key

Signature: _____

Date: _____

INSTRUCTIONS: No electronic devices are permitted. Full credit will be given for correct answers that are unsimplified, e.g. $\frac{1+22}{\sqrt{333}}$ is an acceptable form. There are 100 points on this exam. Please ask if you have any questions. Note that this exam will be videorecorded.

EXAMS ARE DUE AT 10:45. At 10:45am you should be signing the honor statement and turning in the exam. Exams turned in after 10:45 am will be penalized one point for the first minute delay (10:46) and 5 points per minute thereafter (starting at 10:47).

1. (20 points) Severe stunting, indicative of chronic malnutrition, is defined by the World Health Organization as having a z-score of height-for-age below -3 and is a particular problem in developing countries. Researchers conducted a randomized controlled trial of Plumpy'nut[®], a fortified peanut butter paste, versus placebo in young children in Senegal with the hope that Plumpy'nut[®] would be associated with larger z-scores of height-for-age (and therefore less stunting in the population). They wish to test the null hypothesis that mean z-scores are the same in the placebo and Plumpy'nut[®] groups against the alternative that the mean z-scores are not the same in the two groups.

- (a) Suppose the truth is that there is no difference in mean z-scores between the two groups.

- i. Give an example of a type I error in this setting. If it is not possible to make a type I error in this setting, explain why not.

Rejecting the null hypothesis would always be a type I error because the null hypothesis is true.

Note: a type I error is different from a mathematical/calculation/procedural error.

- ii. Give an example of a type II error in this setting. If it is not possible to make a type II error in this setting, explain why not.

Not possible - type II errors only occur when the alternative hypothesis is true.

(b) Suppose the truth is that the mean z-score in the Plumpy'nut[®] group is 0.3 units higher than the mean z-score in the placebo group.

- i. Give an example of a type I error in this setting. If it is not possible to make a type I error in this setting, explain why not.

Not possible - type I errors occur only when the null hypothesis is true.

- ii. Give an example of a type II error in this setting. If it is not possible to make a type II error in this setting, explain why not.

Failing to reject the null hypothesis would always be a type II error because the alternative is true.

Again, we expect to make these errors $100(\beta)\%$ of the time for any given H_A - you do not have to make a mistake or miscalculation to have 1. Those are not type I or type II errors.

2. (30 points) Researchers in China were interested in estimating average fine particle organic carbon levels and took $n = 9$ random samples from Chinese cities. The mean of these samples was $\bar{X} = 24 \mu\text{g}/\text{m}^3$ with sample standard deviation $s = 6 \mu\text{g}/\text{m}^3$.

(a) What is the standard deviation of the mean (standard error) for these data?

$$SE = \frac{s}{\sqrt{n}} = \frac{6}{\sqrt{9}} = \frac{6}{3} = 2$$

- (b) Researchers wish to test whether this level is consistent with a population mean concentration of $\mu = 10 \mu\text{g}/\text{m}^3$. State the null and alternative hypotheses for a two-sided test.

$$H_0: \mu = 10 \mu\text{g}/\text{m}^3$$

$$H_A: \mu \neq 10 \mu\text{g}/\text{m}^3$$

Notes
- Need to specify value
(you can have $\mu_c = \mu_{pop} =$
any number, not just
10)
- \bar{X} should not be in hypothesis

- (c) Researchers wish to test whether this level is consistent with a population mean concentration of $10 \mu\text{g}/\text{m}^3$. Assuming that fine particle organic carbon levels are normally distributed, construct the appropriate test statistic for this hypothesis and provide the degrees of freedom.

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{24 - 10}{2} = \frac{14}{2} = 7$$

Compare this to the critical value that cuts off the upper 2.5% of the curve for a $t_{q-1} = t_8$ distribution.

- (d) Researchers wish to test whether this level is consistent with a population mean concentration of $10 \mu\text{g}/\text{m}^3$. Suppose you carry out the test and obtain $p < 0.001$. Interpret this test result in terms of the subject matter (that is, what does this tell you about fine particle organic carbon levels in these Chinese cities)?

Because $p < 0.05$ we reject the null hypothesis that these levels are consistent with a population mean concentration of $10 \mu\text{g}/\text{m}^3$. The mean in these cities is higher.

- (e) Now suppose you calculate the 95% confidence interval and obtain (19.4, 28.6). Circle all of the following statements that are correct.

- i. At the 5% level of significance, there is evidence to conclude that the sample comes from a population with mean different from $\mu = 10 \mu\text{g}/\text{m}^3$.
- ii. 95% of the interval (19.4, 28.6) covers the true population mean μ , and 5% does not.
- iii. There is a 95% chance the true mean lies in the interval (19.4, 28.6).
- iv. None of these statements are correct.

- (f) If the level of confidence were changed to 99%, what would happen to the confidence interval and the p-value? Circle the correct answer.

- i. The confidence interval would become longer, and the p-value would decrease.
- ii. The confidence interval would become shorter, and the p-value would increase.
- iii. The confidence interval would not change, and the p-value would not change.
- iv. The confidence interval would become shorter, and the p-value would decrease.
- v. The confidence interval would become shorter, and the p-value would not change.
- vi. The confidence interval would become longer, and the p-value would not change.
- vii. None of these statements are correct.

3. (20 points) In sub-Saharan Africa, malaria is a major cause of anemia during pregnancy. Several studies have investigated whether insecticide treated nets (ITNs) could reduce prevalence of malaria and anemia among pregnant women. A randomized controlled trial was conducted to compare hemoglobin levels at delivery between women randomized to receive ITNs during pregnancy and women in a placebo group. A total of 104 women were randomized to the ITN group. The average hemoglobin level at the time of recruitment (early in pregnancy) of these 104 women was $\bar{X}_{ITN-R} = 103.4g/L$ with standard deviation $s_{ITN-R} = 18.2g/L$, and the average hemoglobin level of these 104 women at the time of delivery was $\bar{X}_{ITN-D} = 108.4g/L$ with standard deviation $s_{ITN-D} = 19.3g/L$.
- (a) Which statistical test should be used to test the null hypothesis that hemoglobin levels were the same at recruitment and delivery in the ITN group? Be as specific as possible.

Paired t -test

- (b) The placebo group of 94 women had average hemoglobin levels at delivery of $\bar{X}_{PLACEBO-D} = 102.2g/L$ with standard deviation $s_{PLACEBO-D} = 22.4g/L$. Which statistical test should be used to test the null hypothesis that hemoglobin levels are the same in the ITN and placebo groups at delivery? Be as specific as possible.

Two-sample t -test with unequal variances unless a test shows variances are equal.

- (c) The study involved two other similarly-sized groups: one group who received oral sulfadoxine pyrimethamine (an oral malaria preventive medication), and one group who receive both oral sulfadoxine pyrimethamine and ITNs. Which statistical method would you use to test the null hypothesis that mean hemoglobin levels at delivery are the same across all four groups?

ANOVA, if variances equal.

If worried about unequal variances could
test medians are equal using
Kruskal-Wallis test

- (d) Now suppose that we wish to examine RBC folate in a subset of 10 women each from the placebo and ITN groups. The distribution of RBC folate is highly skewed. Which statistical test should be used to test the null hypothesis that RBC folate levels are the same in the ITN and placebo groups at delivery? Be as specific as possible.

Wilcoxon rank sum test

~~Wilcoxon rank sum test~~

4. (25 points) A researcher is interested in quantifying the difference in the body pesticide burden of individuals on diets using only conventional foods versus diets using only organic foods. The researcher measures baseline pesticide levels in a group of individuals currently eating conventional foods. Then these individuals are provided an all-organic diet for a one-month period, and pesticide levels are measured again. The power calculation indicates that a sample size of 200 is sufficient to detect a difference of 0.2 standard deviations in the pesticide level between baseline and follow-up, assuming a type I error rate of 5% and 80% power for a two-sided hypothesis test.

- (a) Researchers want to increase the power to 90%. How will increasing the power affect the required sample size if the other factors remain fixed?

 n (increases sample size = n required)

- (b) Researchers really would like to detect a difference of just 0.1 standard deviations. How will changing this minimum detectable difference affect the required sample size if the other factors remain fixed?

 n

- (c) Researchers feel strongly that pesticide concentrations on the organic food diet will not be greater than pesticide concentrations on a conventional diet and are considering a one-sided test. How will switching to a one-sided test affect the required sample size if the other factors remain fixed?

↓ sample size

- (d) Because they are testing 5 different pesticides, researchers decide to conduct this test at the 1% level of significance. How will this affect the power if the other factors remain fixed?

↓ power

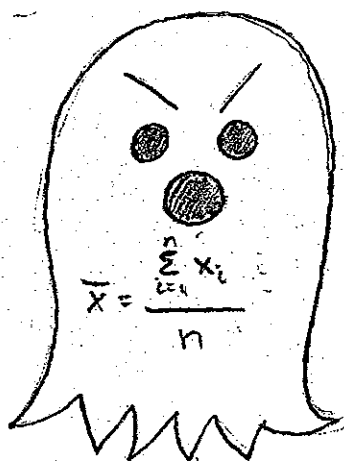
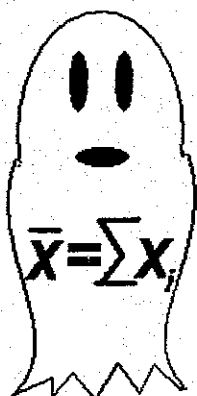
- (e) The investigators get a new grant and can increase the sample size to 350. How will this affect the power if the other factors remain fixed?

↑ power

5. (5 points) Suppose you conduct a hypothesis test with $\alpha = 0.05$. Circle the interpretations of $p > 0.05$ that are correct. (Straight from the p-value paper!)

- (a) The chances are greater than 1 in 20 that a difference would be found again if the study were repeated.
- (b) The probability is less than 1 in 20 that a difference this large or larger could occur by chance alone.
- (c) The probability is greater than 1 in 20 that a difference this large or larger could occur by chance alone.
- (d) The chance is 95% that the study is correct.
- (e) None of the interpretations are correct.

6. (2 POINT BONUS) What is wrong with this mean ghost?



NOW MEAN GHOST
IS MEAN!

HAPPY HALLOWEEN! :D