Biostat Exam 2 F04 (11/9/04)

Coverage: Units 5–7

Part A (Closed Book) - Please record your answers to this part of the exam on your scantron. Make certain you bubble in your name, *last name first!*

CHAPTER 5

1. Statistical inference is

    a. the derivation of cause and effect conclusions from data
    b. generalization about a population from a sample
    c. use of data to explore patterns and relations

2. The two main forms of *statistical inference* are estimation and

    a. descriptive statistics
    b. exploratory data analysis
    c. hypothesis testing

3. A numerical characteristic of a population is:

    a. a parameter
    b. a statistic
    c. an estimate
    d. all of the above

4. The standard error of the mean quantifies the

    a. precision of the sample mean
    b. validity of the sample mean
    c. normality of the sample distribution of the mean

5. When the sampling distribution of a mean is normally distributed, 95% of potential sample means will fall with 1.96 _____ of the population mean.

    a. standard deviations
    b. standard errors
    c. sample means
    d. none of the above

6. The confidence interval for a mean seeks to capture the

    a. population mean
    b. sample mean
    c. neither A nor B
    d. both A and B

7. A *t* distribution with infinite degrees of freedom is equal to a _____ distribution.

    a. binomial ($n = 0$, $p = 1$)
    b. binomial ($n = 10$, $p = 0.5$)
    c. Normal ($\mu = 0$, $\sigma = 1$)
    d. Normal ($\mu = 100$, $\sigma = 15$)

8. In estimating a mean, what varies from sample to sample, the confidence interval or the population mean?

    a. confidence interval
    b. population mean

CHAPTER 6

9. When conducting a statistical test, which hypothesis is assumed to be true until proven otherwise?

a. H_0
b. H_1
c. neither
d. both

10. Select the best definition of "alpha".

a. probability of a type I error
b. probability of a type II error
c. probability of avoiding a type I
d. probability of avoiding a type II

11. Select the best definition of "beta".

a. probability of a type I error
b. probability of a type II error
c. probability avoiding a type I error
d. probability avoiding a type II

12. Select the best definition of "power"

a. probability of a type I error
b. probability of a type II error
c. probability of avoiding a type I
d. probability of avoiding a type II

13. When conducting a statistical test using the decisional-based approach, the decision rule states:

a. reject H_0 when $p >$ alpha
b. reject H_0 when $p \leq$ alpha

14. The p value is the probability the null hypothesis is correct.

a. True
b. False

15. The duration of a particular acute illness is approximately normally distributed with a mean of 10 days and standard deviation of 5 days. A researcher hopes to show that a new treatment hastens recovery and will employ a statistical hypothesis test to do so. Select the appropriate one-sided alternative hypothesis for this research question.

a. H_1 $\bar{x} > 10$
b. H_1: $\bar{x} < 10$
c. H_1: $\mu < 10$
d. H_1: $\mu > 10$

16. In the following table, which cell represents a type I error?

	H_0 true	H_0 false
retain H_0	a	b
reject H_0	c	d

17. In the table above, which cell represents a type II error?

CHAPTER 7

18. Which of the following is a synonym for "paired samples".

a. independent samples
b. dependent samples
c. unrelated samples
d. none of the above

19. Which of the following is an example of a paired sample?

a. men and women selected at random from the population
b. diabetics and non-diabetics selected at random
c. husband and wife couples selected at random
d. all of the above

20. How many degrees of freedom does a paired *t* test statistic have when studying 15 pairs?

a. 14
b. 15
c. 16
d. none of the above

21. How many measurements are taken when studying 15 pairs?

a. 14
b. 15
c. 30
d. none of the above

22. Ninety-five percent confidence intervals for mean differences locate the *sample* mean difference with 95% confidence.

a. True
b. False

23. What is the null hypothesis for testing a paired difference for significance?

a. H_0: $\bar{x}_d = 0$
b. H_0: $\mu_d = 0$
c. $H_0 = 0$
d. $H_0 = \mu_0$

24. What is the value of $\phi(0)$?

a. .025
b. .05
c. .50
d. none of the above

25. Which symbol represents the sample mean difference based on paired samples?

a. μ
b. $\bar{x}$
c. μ_d
d. $\bar{x}_d$

Part B (Procedure Book) - Record your answers to this part of the exam in your blue book. Show all work and *please* remember to *number each response!*

(1) In Lab 5 you conducted an experiment in which the mean AGE in your random sample was studied in relation to the mean of other random samples in the class. Discuss the key findings from this experiment. Consider, the mean, standard deviation, and shape of the sampling distribution of the classes' means. [3 to 6 sentence response; proper English please; 3 pts.]

(2) Weschler IQ scores are normally distributed with a standard deviation of 15. A sample of 25 children between 11- and 13-years of age who have depression show a mean IQ score of 98.7. Assume these children are a random sample of depressed children. Calculate a 95% confidence interval for the mean IQ score of all depressed children. Show all work. Interpret your confidence interval.[5 pts]

(3) For the above data [question 2], how large a sample would be needed to derive a 95% confidence interval with a margin of error of 2? [3 pts]

(4) Test whether the data in question 2 were significant at $\alpha = .05$ one-sided. (Recall that the mean Weschler IQ score is normally 100.) Show all hypothesis testing steps, including statements of H_0 and H_1. Be certain you draw a graphical representation of normal curve, placing the test statistic on the curve and shading the *p* value region. Determine the *p* value. State the conclusion of the test. Do data provide statistically significant evidence that these IQ scores are lower than average? [8 pts]

(5) Data from pre-test/post-test scores are shown in the table to the right. [10 pts]

ID	POSTTEST	PRETEST
1	461	384
2	642	557
3	540	585
4	454	480
5	398	368
6	382	296
7	501	441
8	434	372
9	380	349

(A) Calculate DELTA values. List differences as an *ordered array*. What is the median difference?

(B) What percentage of values showed a *decrease*?

(C) The mean increase in the data set is 40.00. The standard deviation is 47.64. The sample size is 9. Calculate a 95% confidence interval for μ_d and interpret your results.

(6) A statistician calculates $t_{stat} = 1.56$ with df = 19. Draw the sampling distribution of the test statistic ("*t* curve"), while placing the test statistic in its approximate location and

shading the one-tailed *p* value region. Provide the values of the "wedgie" landmarks from the *t* table and determine the approximate *p* value in one-tail (i.e., *p* is less than so-and-so and more than so-and-so). Then, report the two-tailed *p* value. [6 pts]

(7) You want to calculate the power of the test of a paired difference at $\alpha = .05$ two-sided. You have a pretty good idea of the standard deviation of the variable, and you know the sample size. What additional piece of information do you need before you are able to calculate the power of your test?