

Please turn off your cell phone. For full credit, please show all of your work.

Name: _____ Section: _____ Student ID: _____

STA291 001-012

Final Exam

May 3, 2006

1. Consider the following sample drawn from some population:

62	63	63	64	65	67	68	69	73	74
79	80	86	92	99					

a) [3 points] Find the five number summary (min, Q_1 , median, Q_3 , and max) for this data.

b) [3 points] Give a stem-and-leaf plot for this data.

c) [3 points] The overall shape of the distribution in the above stem-and-leaf plot is
(Multiple Choice, Circle the correct answer)

- (i) clearly symmetric
- (ii) clearly right-skewed (skewed towards the higher values)
- (iii) clearly left-skewed (skewed towards the lower values)
- (iv) no clear shape

2. [3 points] In Canada in 1981, for the categories Catholic, Protestant, Eastern Orthodox, Jewish, None, Other for religious affiliation, the relative frequencies were 47.3%, 41.2%, 1.5%, 1.2%, 7.3%, 1.5%. Which of the following responses is true? (Multiple Choice, Circle the correct answer)

- (i) The median religion is Protestant.
- (ii) Only 2.7% of the subjects fall within one standard deviation of the mean.
- (iii) The "Other" response is an outlier.
- (iv) The mode is Catholic.

3. [3 points] A sample of size n is selected at random from a very large population. As n increases, which of the following statements is true. (Multiple Choice, Circle the correct answer)

- (i) The standard deviation of the sampling distribution increases.
- (ii) The population standard deviation increases.
- (iii) The standard deviation of the sampling distribution decreases.
- (iv) The population standard deviation decreases.

4. [3 points] A 99% confidence interval estimated for a population mean μ is determined to be 18.2 to 22.3. If the confidence level is reduced to 95%, the confidence interval for μ :

(Multiple Choice, Circle the correct answer)

- (i) becomes narrower
- (ii) remains the same
- (iii) becomes wider
- (iv) none of these

Please turn off your cell phone. For full credit, please show all of your work.

5. [6 points] Calculate the sample mean, the deviations, and the sample variance for the following five observations.

2 4 6 8 10

6. At the end of one STA 291 lecture, the students were asked about their experience with an online homework system called “Cyberstats”. Students were asked to choose from the answers given in the left column of the table below (additionally, there was the option to circle “no opinion”). We would like to find out if the opinion on Cyberstats depends on whether students used extra material provided with this online system. The table of joint probabilities is as follows.

	Student has used extra online material	Student has never used extra online material
“Cyberstats has been helpful in learning/understanding statistics”	.490	.255
“Cyberstats has not helped me in learning statistics”	.194	.061

a) [3 points] What is the probability of a student answering “Cyberstats has been helpful in learning/understanding statistics”? (Multiple Choice, Circle the correct answer)
(i) 0.255 (ii) 0.490 (iii) 0.745 (iv) 0.950

b) [3 points] If a student has used extra online material, what is the probability that this student answers “Cyberstats has been helpful in learning/understanding statistics”? (Multiple Choice, Circle the correct answer)
(i) 0.194 (ii) 0.490 (iii) 0.716 (iv) 0.745

c) [4 points] Is a favorable opinion regarding Cyberstats statistically independent from having used extra online material? (It is important that you show your calculations)

d) [4 points] Can we be sure that the result from c) can be generalized to all undergraduate students across campus? What factors are not included in the calculation in c)?

Please turn off your cell phone. For full credit, please show all of your work.

7. [3 points] Two dice are thrown. What is the probability that the results of the two dice have a sum of 5 or less? (Multiple Choice, Circle the correct answer)

- (i) $4/36$ (ii) $6/36$ (iii) $10/36$ (iv) $12/36$

8. [2 points] Given that Z is a standard normal random variable, determine $P(-1.52 \leq Z \leq 0.85)$.

9. [2 points] For some value of z , the probability that a standard normal variable is above z is 0.8810. Determine the value of z .

10. Assume that you want to test the hypothesis $H_0 : \mu = 1000$ against a two-sided alternative. For a sample of size $n=100$, the sample mean is 980 and the sample standard deviation is 200.

a) [4 points] Calculate the value of the test statistic and determine the p -value.

b) [4 points] Set up the rejection region for a test at significance level (alpha level) 0.01.

11. [3 points] When testing $H_0 : \mu = 0$ vs. $H_1 : \mu < 0$, the observed value of the z -score was found to be -2.15. The p -value for this test would be. (Multiple Choice, Circle the correct answer)

- (i) 0.0158 (ii) 0.0316 (iii) 0.9684 (iv) 0.9842

12. [3 points] When testing $H_0 : \mu = 0$ vs. $H_1 : \mu < 0$, the observed value of the z -score was found to be 2.15. The p -value for this test would be. (Multiple Choice, Circle the correct answer)

Please turn off your cell phone. For full credit, please show all of your work.

- (i) 0.0158 (ii) 0.0316 (iii) 0.9684 (iv) 0.9842

13. [3 points] Which of the following is true regarding the sampling distribution of the mean for a large sample size? *(Multiple Choice, Circle the correct answer)*
- (i) It has the same shape, mean, and standard deviation as the population distribution.
 - (ii) It has the same shape and mean as the population distribution, but has a smaller standard deviation.
 - (iii) It has a normal distribution with the same mean as the population distribution but with a smaller standard deviation.
 - (iv) It has a normal distribution with the same mean and standard deviation as the population.
14. [3 points] A 95% confidence interval estimate of the population mean : can be interpreted to mean that: *(Multiple Choice, Circle the correct answer)*
- (i) We are 95% confident that 5% of the values of the sample means \bar{X} will result in a confidence interval that includes the population mean .:
 - (ii) We are 95% confident that we have selected a sample whose range of values does not contain the population mean .:
 - (iii) There is a 95% probability that the population mean : will lie in this confidence interval.
 - (iv) If we repeatedly draw samples of the same size from the same population, then 95% of the times we will obtain a confidence interval that includes the population mean .:
15. [3 points] Suppose a 90% confidence interval for : turns out to be (32, 41). Give a definition of what it means to be “90% confident”. *(Multiple Choice, Circle the correct answer)*
- (i) 90% of the observations in the sample fall within the given interval.
 - (ii) 90% of the observations in the entire population fall within the given interval.
 - (iii) In repeated sampling, 90% of the intervals constructed would contain the population mean.
 - (iv) In repeated sampling, the population mean would fall 90% of the time within the interval (32, 41).
16. [3 points] Two samples of sizes 17 and 24 are independently drawn from two normal populations with equal population variances. The number of degrees of freedom of the equal-variances t -test statistic is: *(Multiple Choice, Circle the correct answer)*
- (i) 16 (ii) 23 (iii) 39 (iv) 40
17. [3 points] On UK’s campus, students are debating whether Chipotle’s or Qdoba’s offers the better burritos. In order to get some statistical information related to this controversy, researchers calculate a 95% confidence interval for the difference in the mean protein content of a standard burrito. The confidence interval for the difference Burrito 2 – Burrito 1 (the names of the franchises are not disclosed here) was from -1.3 to 3.0. Which of the following statements can the researcher make? *(Multiple Choice, Circle the correct answer)*
- (i) Burrito 1 has less protein content than Burrito 2.
 - (ii) Burrito 1 has a greater protein content than Burrito 2.
 - (iii) No comparison can be made. We need to perform a hypothesis test.
 - (iv) Burrito 1 and Burrito 2 do not differ significantly in protein content.
18. [3 points] If an economist wishes to determine whether there is evidence that the average family income in a neighborhood is below \$30,000, then: *(Multiple Choice, Circle the correct answer)*
- (i) Either a one-tailed or a two-tailed test could be used with equivalent results.
 - (ii) A one-tailed test should be utilized.

Please turn off your cell phone. For full credit, please show all of your work.

(iii) A two-tailed test should be utilized.

(iv) A test of proportions should be utilized.

19. [6 points] Do male or female college students spend more time on studying for their statistics class? Assume that random samples of undergraduate students yield the following summary statistics of study times: For 18 male students, the mean study time per week was 85.7 minutes with a standard deviation of 8.9 minutes. For 19 female students, the mean study time per week was 101.7 minutes with a standard deviation of 9.2 minutes. Set up null and alternative hypothesis, calculate the value of the test statistic, and determine the p -value for determining whether study times are different for male and female students.

Hint: In determining the p -value, you may not be able to give an exact answer using the table. In that case, provide a range in which the p -value is.

20. [3 points] In testing the hypothesis $H_0 : \mu = 100$ vs. $H_1 : \mu > 100$, the p -value is found to be 0.074, and the sample mean is 108. Which of the following statements regarding the p -value is true?

(Multiple Choice, Circle the correct answer)

- (i) The probability of observing a sample mean at least as large as 108 from a population whose mean is 100 is 0.074.
- (ii) The probability of observing a sample mean smaller than 100 from a population whose mean is 108 is 0.074.
- (iii) The probability that the population mean is larger than 100 is 0.074.
- (iv) The probability that the population mean is 100 is 0.074.

21. [3 points] The power of a test is measured by its capability of:

(Multiple Choice, Circle the correct answer)

- (i) Rejecting a null hypothesis that is true.
- (ii) Failing to reject a null hypothesis that is true.
- (iii) Rejecting a null hypothesis that is false.
- (iv) Failing to reject a null hypothesis that is false.

22. [4 points] A random sample of 15 STA 291 students has mean age equal to 23.2 years and the standard deviation is 6.9 years. Find a 90% confidence interval for the population mean age.

23. [4 points] A sample of 70 male college students was asked "How many times do you brush your teeth each day?". Out of the 70 students surveyed, 52 answered that they brushed their teeth more than once a day. Calculate a 99% confidence interval for the proportion of males who brush their teeth more than once a day.

24. [3 points] A type I error is committed if we make: (Multiple Choice, Circle the correct answer)

- (i) A correct decision when the null hypothesis is false.
- (ii) A correct decision when the null hypothesis is true.

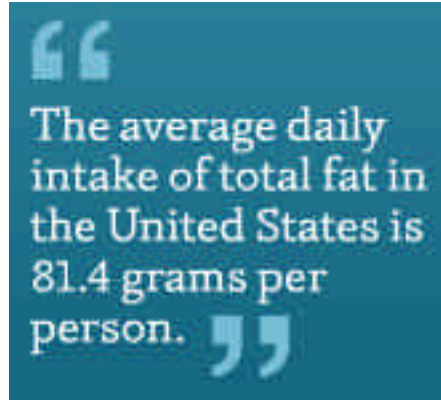
Please turn off your cell phone. For full credit, please show all of your work.

(iii) An incorrect decision when the null hypothesis is false.

(iv) An incorrect decision when the null hypothesis is true

Bonus Exercise.

a) [4 points] Consider the following statement that appears on Pfizer's website for Lipitor:



Briefly explain how you could test this claim using your statistical knowledge from this course.

b) [4 points] Consider the analogy between making a decision about a hypothesis in a significance test and making a decision about the innocence or guilt of a defendant in a criminal trial. Suppose " H_0 true" is "Defendant innocent", " H_0 false" is "Defendant guilty", "Reject H_0 " is "Convict defendant", and "Do not reject H_0 " is "Acquit defendant".

Explain what Type I error and Type II error mean in this trial setting.