MA 262 Homework Problems

What follows is a list of suggested practice problems for this course. Do as many problems as you feel are necessary to understand the associated concept(s).

Basic Introduction

1. Give an example of a set of 5 data values that have the same mean, median, and mode.

2. Give an example of a set of 6 data values, no 2 the same, which have the same mean and median.

3. Suppose a set of 10 data values has mean 100 and median 110. Determine the new mean and median for each of the data sets below.

   a. Each value in the data set is doubled.
   b. Twenty is added to each value in the data set.
   c. Each value in the data set is doubled and then 20 is added to each resulting value.
   d. Twenty is added to each value in the data set and then each resulting value is doubled.

4. A set of 8 data values has a mean of 50 and a median of 52. To these data values, the additional values 60 and 40 are included. Can the sample mean and median of these 10 data values be determined? If so, determine the value(s).

5. A set of 8 data values has a mean of 50 and a median of 52. To these data values, the additional values 60 and 70 are included. Can the sample mean and median of these 10 data values be determined? If so, determine the value(s).

6. For the given list of data, find the following sums: $\sum x$, $\sum x^2$, and $\sum (x-\bar{x})^2$. Use these sums to calculate the variance of the data by two different methods.

   9 19 22 10 25 16 18 19 24 25

7. Determine the mean, median, variance, and standard deviation of the sample data

   123 116 122 110 162 134 125 113 118 117

   Find 3 other sets of 10 integers that have the same variance.

8. Can the variance ever be equal to the standard deviation? Under what circumstances is the variance equal to zero? Can the variance ever be less than the standard deviation?

9. Consider the sample data set 25, 45, and 65. The mean and median are equal. Can one additional number be added to this set so that the mean and median are still equal? Can one additional number be added that is not equal to one of the original 3 numbers, such that the mean and median are equal?

10. Suppose Data Set A is a set of values whose mean is 75 and whose variance is 100. For each of the following data sets, determine the mean, standard deviation, and variance.

    a. Data Set B, obtained from Data Set A by adding 12 to each value of Data Set A.
    b. Data Set C, obtained from Data Set A by subtracting 8 from each value of Data Set A.
    c. Data Set D, obtained from Data Set A by tripling each value of Data Set A.
    d. Data Set E, obtained from Data Set A by halving each value of Data Set A.
    e. Data Set F, obtained from Data Set A by first multiplying each value of Data Set A by 2.5 and then adding 10 to each result.
    f. Data Set G, obtained from Data Set A by first adding 10 to each value of Data Set A and then multiplying each result by 2.5.

11. Suppose a sample of 9 exam scores results in a sample mean of 82.3 and a sample standard deviation of 16.1. To this sample, an additional exam score of 90 is added.

    a. Will the sample mean of the 10 exam scores be larger or smaller than the mean of the 9 scores before the addition of the score of 90? Justify your response without performing any calculations.
    b. Will the sample standard deviation of the 10 exam scores be larger or smaller than the standard deviation of 9 scores before the addition of the score of 90? Justify your response without performing any calculations.
    c. What is the sample mean of the 10 exam scores? (You must show your calculations.)
    d. What is the sample standard deviation of the 10 exam scores? (You must show your calculations.)

12. A set of 8 sample data values has a sample mean of 40 and a sample standard deviation of 10. Two additional data values are added to the original 8 data values.

    a. What were the values of the 2 additional numbers if the new sample mean is 50 and the new sample standard deviation is 25?
    b. What were the values of the 2 additional numbers if the new sample mean is 50 and the new sample standard deviation is 10?
    c. Assuming that the new mean is 50, what is the smallest value the new standard deviation can assume?

This page will not be covered so don't bother.

Grouped Statistics

13. Determine the mean and variance of the grouped data given below:

Class	Frequency (f)
0 – 10	6
11 – 21	4
22 – 36	10
37 – 45	15
46 - 70	5

14. Find the mean and variance of the following grouped data:

Class	Frequency (f)
1 – 5	15
6 – 10	5
11 – 15	6
16 – 20	10
21 - 25	4

15. Find the mean and variance of the following grouped data:

Class	Frequency (f)
10 – 25	35
26 – 50	45
51 – 78	55
79 – 100	165

16. Find the mean and variance of the following grouped data:

Class	Frequency (f)
12.0 – 25.5	50
25.6 – 76.2	10
76.3 – 121.6	5
121.7 – 202.1	15
202.2 – 300.0	70

Binomial Distribution

17. For each of the following, find another binomial coefficient that is numerically equal to it:

    a. $\binom{395}{166}$

    b. $\binom{2847}{98}$

    c. $\binom{12047}{9455}$

18. Twenty cars are vying for 10 adjacent parking spaces. How many different selections of 10 cars can obtain parking in these spaces?

19. Assume that x is a binomial random variable for which p = 0.75. If an experiment is performed using n = 25 trials, determine the smallest value of the constant *k* such that $\Pr(x \geq k) \leq 0.10$.

20. The probability that a patient recovers from a delicate heart operation is 0.85. What is the probability that:

    a. exactly 5 of the next 7 patients having this operation survive?
    b. at least 8 of the next 10 patients survive?

    For the next 20 patients undergoing this operation, what is the maximum value of *k* such that $\Pr(x \geq k) \geq 0.5$?

21. Suppose that airplane engines operate independently and fail with probability 0.4. Assuming that plane makes a safe flight if at least one-half of its engines run, determine whether a 4-engine plane or a 2-engine plane has the higher probability for a successful flight.

22. Evaluate $\sum_{k=0}^{n}\binom{n}{k}$ by generalizing the results of the $\sum_{k=0}^{2}\binom{2}{k}$, $\sum_{k=0}^{3}\binom{3}{k}$, $\sum_{k=0}^{4}\binom{4}{k}$, etc.

Poisson Distribution

23. Assume that *x* is a Poisson random variable having parameter $\mu = 8.5$. Determine the <u>smallest</u> value of the constant *k* such that $\Pr(x \geq k) \leq 0.10$.

24. A certain area of the Florida is, on average, hit by 8 hurricanes a year. Find the probability that for a given year that area will be hit by:

    a. fewer than 6 hurricanes.
    b. anywhere from 6 to 10 hurricanes.
    c. more hurricanes than average.

25. Suppose that, on average, 1 person in 125 makes a numerical error in preparing his or her income tax. If 1000 filed tax returns are selected at random and examined, what is the probability that 2, 3, 4, or 5 of the returns contain a numerical error?

26. The likelihood of having a certain genetic marker is 1 in 2000. Should one consider it unlikely if at least 12 of 10000 patients randomly examined exhibit this genetic marker?

27. Suppose X is a Poisson random variable. What is the parameter of this distribution if $\Pr(1 \leq X \leq 8) \approx 0.96$?

Normal Distribution

28. Determine the following values related to the normal and standard normal distributions:

a. $\Pr(z > -1.15)$
b. $\Pr(-2.04 \le z < -0.52)$
c. $\Pr(z \le 1.884)$
d. $\Pr(z \ge 0.997)$
e. $\Pr(-1.256 \le z \le 2.012)$
f. $\Pr(|z| < 1.35)$
g. $\Pr(|z| > 1.77)$
h. $\Pr(|z| < 0.663)$
i. $\Pr(|z| > 2.253)$
j. Below what value do 88% of the standard normal values lie?
k. Above what value do 34% of the standard normal values lie?
l. What is the 75^{th} percentile of the standard normal distribution?
m. What is the 18^{th} percentile of the standard normal distribution?
n. Below what value do 71% of the values lie in the distribution that is $n(120,20)$?
o. Above what value do 41% of the values lie in the distribution that is $n(52.6,12.8)$?
p. $z_{0.64}$
q. $z_{0.22}$
r. $\Pr(z < z_{0.81})$
s. $\Pr(z > z_{0.79})$
t. $\Pr(z_{0.32} \le z < z_{0.81})$
u. $\Pr(|z| < z_{0.75})$
v. $\Pr(|z| \ge z_{0.14})$

29. Let the notation $\Pr(a < X < b \mid X \sim n(\mu,\sigma))$ mean "the probability that X lies between a and b given that X is distributed normally with mean μ and standard deviation σ." With this notation, determine the following:

a. $\Pr(20 < X < 50 \mid X \sim n(30,10))$
b. $\Pr(42 < X < 51 \mid X \sim n(40,5))$
c. $\Pr(X < 62.5 \mid X \sim n(55.3,12.2))$
d. $\Pr(X > 72.8 \mid X \sim n(75.2,5.7))$
e. The 95^{th} percentile of the distribution that is $n(80,12)$
f. The 12^{th} percentile of the distribution that $n(42.7,30.6)$
g. $\Pr(X_{0.325} < X < X_{0.944} \mid X \sim n(181.6,55.9))$

Normal Applications

30. In an industrial process the diameter of a ball bearing is an important component part. The buyer sets specifications on the diameter to be 3.00 ± 0.01 cm. The implication is that no part falling outside these specifications will be accepted. It is known that the manufacturing process for these ball bearing produces ball bearing that are distributed n (3.000, 0.005) cm. On average, how many of the manufacturer ball bearings will be scrapped by the buyer?

31. Gauges are used to reject all components where a certain dimension is not within the specification $1.50 \pm x$. It is known that this measurement is n (1.50, 0.20). Determine the value x such that the specifications "cover" 95% of the measurements.

32. From experience a professor knows that the term grades for the class will be $n(450.0, 60.0)$. Suppose that the top 8% of the numerical grades will be given a letter grade of A, the next 28% will be given a letter grade of B, the next 35% will be given a letter grade of C, the next 16% will be given a letter grade of D, and the remainder will be given a letter grade of F. To 1 decimal place, determine the numerical range for each letter grade.

33. A professor is teaching 2 sections of the same course. Each section is given an exam. A student from the first section misses the exam and is allowed to take the exam with the other section. The exams cover the same material, but are decidedly different. The grades in the first section are $n(66.4, 13.2)$ and the grades in the other section are $n(72.8, 10.1)$. The student in question obtains a grade of 76.1 when taking the exam with the other class. However, the professor wants to record a grade that will be reflective of their performance on the exam given to the first section. What grade should be recorded?

34. Suppose that you produce a product for which you wish to offer a limited warranty. Your experience indicates to you that the average life of this product is 9.1 years with a standard deviation of 2.5 years. You are going to offer to replace any product that fails during the warranty period free of charge. To 1 decimal place, how long should the warranty be if you do not wish to have to replace more than 5% of the products under the warranty? Assume that the product lives are normally distributed.

35. Suppose you are told that, for positions comparable to the one you currently hold, the salary at the 25th percentile is \$38,210 and the salary at the 90th percentile is \$58,935. Assume that these salaries are normally distributed.

    a. To the nearest dollar, what is the median salary?
    b. Below what salary do three-quarters of the appropriate salaries lie?
    c. If your salary is \$48,424, between what 2 integer percentiles does your salary lie?

Approximating Distributions

36. It has long been supposed that there is a relationship between asbestos exposure and death due to chronic obstructive pulmonary disease (COPD). Among workers exposed to asbestos in a shipyard in 1980, 33 died over a 10-year period from COPD. Based on statewide mortality rates, only 24 should have been expected to die from COPD during that interval of time. Assuming the statewide rate is correct, with what probability could at least 33 deaths be expected?

37. A fair coin is tossed 2000 times. What is the probability of obtaining anywhere between 950 and 1025 heads, inclusive?

38. Suppose volunteers in a large national study are coded by the last 4 digits of their social security number. If 10,000 subjects are in this study what is the probability that at least 4 of them would share the same coding number?

39. A certain telemarketing firm has found that their employees average 45 calls per hour. Each employee is evaluated every 3 months and if they do not average at least 35 calls per hour, they are put on probation. Assuming that this firm has a large number of employees, what is the probability that a randomly selected employee will be placed on probation?

40. A recent study indicates that 1 of every 500 births in this country results in complications that cause the newborn to remain in the hospital for more than 2 days. Five of the most recent 1200 babies born in a specific hospital have had to remain in the hospital for more than 2 days. How likely is it that at least 5 of 1200 newborns must remain in the hospital?

41. Ten percent of the students at this university have GPAs of at least 3.35. The current graduating class has 350 students in it. What is the probability that at least 50 of them will have a GPA of at least 3.35?

Central Limit Theorem

42. A population is known to have a mean of 100 and a variance of 144. Suppose a very large number of samples of size n = 36 are taken from this population and, for each sample taken, the sample mean is calculated. Describe the expected distribution of these sample means.

43. Two hundred samples of size n = 25 are taken from a large population. The mean of these 200 sample means is 50 and the standard deviation of these 200 means is 10. Describe the population from which the 200 samples were taken.

44. One hundred samples of size n = 40 are taken from a population. The distribution of the sample means is found to have a mean of 55 and a standard deviation of 12. Describe the distribution of sample means taken from the same population if 300 samples of size n = 10 are taken.

45. A population with mean 200 and standard deviation 50 is sampled. What is the probability that a sample of size 100 taken from this population will have a sample mean:

   a. that exceeds 210?
   b. of no more than 195?
   c. lying between 192 and 205, inclusive?

46. If a certain machine makes electrical resistors having a mean resistance of 54 Ω and a standard deviation of 5 Ω, what is the probability that a random sample of 25 of these resistors will have a combined resistance of 1400 Ω?

47. If the standard error of the mean for the sampling distribution of random samples of size 64 from a large or infinite population is 5, how large must the size of the sample become if the standard deviation is to be reduced to 2?

48. If all possible samples of size 25 are drawn from a population that is $n(64.2, 8.4)$, what is the probability that a sample mean $\bar{x}$ will fall in the interval from $\mu_{\bar{x}} - 1.72\sigma_{\bar{x}}$ to $\mu_{\bar{x}} + 0.35\sigma_{\bar{x}}$? Assume that the sample means can be measured to any degree of accuracy.

Student-t Distribution

49. Determine the following distribution and/or probabilities. Interpolation might be necessary. Proper notation must always be used.

    a. The 40th percentile of the *t* distribution with respect to a sample of size 28.
    b. The 20th percentile of the *t* distribution with respect to a sample of size 63
    c. $\Pr\left(1.058 \le t \le 2.167 \mid df = 25\right)$
    d. $\Pr\left(-1.074 \le t < 1.753 \mid df = 15\right)$
    e. $\Pr\left(t \le 1.616 \mid df = 12\right)$
    f. $\Pr\left(t < -2.262 \mid df = 4\right)$
    g. $\Pr\left(t > -1.012 \mid df = 10\right)$
    h. $t_{0.995,75}$
    i. $t_{0.05,37}$
    j. $t_{0.98,350}$
    k. $t_{0.01,515}$
    l. The value of the constant *k* such that $\Pr\left(t \le k \mid df = 8\right) = 0.1$.
    m. The value of the constant *k* such that $\Pr\left(t > 1.057 \mid df = k\right) = 0.15$.
    n. The value of the constant *k* such that $\Pr\left(t > k \mid df = 22\right) = 0.01$.
    o. The value of the constant *k* such that $\Pr\left(|t| \le k \mid df = 30\right) = 0.50$.
    p. The value of the constant *k* such that $\Pr\left(|t| > k \mid df = 17\right) = 0.2$.

50. Five randomly selected scores from the last exam were found to be 75, 82, 78, 75, and 78. Construct a 95% confidence interval for the mean on this exam based on these scores. Suppose the instructor claimed that the actual average grade on this exam was 83.6. With respect to your analysis, do you believe the claim? That is, does the statistical evidence make it very unlikely that the claim is reasonable? Justify your response.

51. A maker of a low-fat granola bar claims that the average saturated fat content is 0.47 grams. In a random sample of 8 granola bars of this brand the saturated fat contents were 0.62, 0.74, 0.71, 0.35, 0.43, 0.63, 0.46, and 0.62 grams. Would you agree with the claim? Assume a normal distribution of saturated fat contents. Defend your decision statistically.

52. A manufacturer of car batteries guarantees that their batteries will last, on the average, 4.1 years with a standard deviation of 1.0 years. If 5 of these batteries have lifetimes of 1.9, 2.4, 3.0, 3.5, and 4.2 years, and assuming that battery lifetime follows a normal distribution, is the manufacturer still convinced that the batteries have the claimed mean lifetime? Defend your decision statistically.

53. A normal population with unknown variance has a mean of 28.4. Is one likely to obtain a random sample of size 9 from this population with a mean of 24.5 and a standard deviation of 3.9? If not, what conclusion(s) would you draw? Defend your decision statistically.

54. A sample of size n = 25 results in a 95% confidence interval for the population mean of $64.2 < \mu < 72.4$. What are the 2 possible variances that led to the construction of this confidence interval? Why are 2 different values for the variance possible?

55. An 88% confidence interval for the population mean based on sample data is $27.71 < \mu < 34.29$. Assuming that the population variance was known, with what level of confidence can it be stated that $26.50 < \mu < 35.50$?

56. Suppose an 85% confidence interval for the population mean is to be constructed. If the population variance is assumed known (or if the sample value will be used as a point estimate), what sample size will be needed to insure that the population mean differs from the sample mean by no more than d units? Generalize this result for any specified degree of confidence.

Chi-Square Distribution

57. Determine the following distribution and/or probabilities. Interpolation might be necessary. Proper notation must always be used.

    a. $\chi^2_{0.15,20}$

    b. $\chi^2_{0.975,16}$

    c. $\Pr\left(\chi^2 \geq 19.021 \mid df = 23\right)$

    d. $\Pr\left(8.547 \leq \chi^2 \leq 24.311 \mid df = 15\right)$

    e. $\chi^2_{0.95,112}$

    f. $\chi^2_{0.18,630}$

    g. $\Pr\left(\chi^2 \geq 20.386 \mid df = 28\right) - \Pr\left(\chi^2 < 19.406 \mid df = 14\right)$

58. Determine 98% confidence intervals for the population variance and population standard deviation based on the data values 4, 7, 12, 3, 6, and 10.

59. Five randomly selected scores from the last exam were found to be 75, 82, 78, 75, and 78. Suppose the instructor claimed that the actual standard deviation was 2.2. Is there sufficient statistical evidence to refute the claim? Justify your response.

60. A random sample of $n = 20$ taken from a normal population yields a sample mean of 52.5 and a sample standard deviation of 12.5. Determine the 90%, 95%, and 98% confidence intervals for the population variance.

61. Find the probability that a random sample of 28 observations, from a normal population with variance $\sigma^2 = 9$, will have a sample variance s^2:

    a. greater than 12.426;
    b. less than 7.573;
    c. between 5.203 and 6.901.

F Distribution

62. Determine the following distribution and/or probabilities. Interpolation might be necessary. Proper notation must always be used.

a. $F_{0.99,53,25}$

b. $F_{0.05,100,35}$

c. $F_{0.01,80,44}$

d. $\Pr\left(F \geq \frac{17}{4} \middle| df = 12,15\right)$

e. $\Pr\left(F \geq 0.450 \middle| df = 10,17\right)$

f. $\Pr\left(0.384 \leq F \leq 4.100 \middle| df = 13,12\right)$

63. Two samples are collected. The sample with the larger sample variance, call it Sample 1, is obtained from a sample of size 10 while the sample with the smaller sample variance, call it Sample 2, is obtained from a sample of size 28. Determine $\Pr\left(4s_1^2 > 9s_2^2\right)$.

64. If an interval estimate is needed for the comparison of 2 population variances, the comparison takes the form of the ratio of the population variances. The confidence interval for the ratio of the 2 population variances is given by the formula

$$\frac{s_1^2}{s_2^2}\frac{1}{F_{p,df_1,df_2}} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2}F_{p,df_2,df_1}$$

Suppose that the sample variance of the sample of size 12 is 60.2 and the sample variance of the sample of size 18 is 25.8. Determine 95% and 98% confidence intervals for the ratio of the 2 population variances. Based on these confidence intervals, can you refute a statement that claims that the variances of the populations are not different?

One-Sample Hypothesis Testing

65. Convert each of the following statements into a pair of appropriately labeled hypotheses:

    a. No less than 45% of the voters support the issue.
    b. The average grade on the last exam exceeded 75.
    c. The variance is no greater than 10.2.
    d. The standard deviation is at least 25.6.
    e. Four out of 5 dentists who recommend sugarless gum for their patients who chew gum recommend Trident.
    f. The typical waiting time is no longer 14.2 minutes.

66. With respect to each of the hypotheses pairs implied in Problem 65, determine the *p*-value given the following additional information, writing each using the appropriate notation:

    a. $z = -1.880$, $n = 45$
    b. $t = 1.583$, $n = 20$
    c. $\chi^2 = 33.196$, $n = 25$
    d. $\chi^2 = 7.120$, $n = 15$
    e. $z = -2.112$, $n = 23$
    f. $t = 1.721$, $n = 22$

67. With respect to each of the hypotheses pairs implied in Problem 65, determine the appropriate critical value(s) for the level of significance. Be sure to write the critical values in the correct subscripted form.

    a. $\alpha = 0.05$
    b. $\alpha = 0.10$
    c. $\alpha = 0.01$
    d. $\alpha = 0.02$
    e. $\alpha = 0.05$
    f. $\alpha = 0.10$

68. A survey sponsored by General Motors shows that women are the motivating force behind at least 80% of SUV purchases. A random sample of 75 SUV purchases indicates that 55 of them were either purchased by women or were purchased by families where the woman was the driving force behind the purchase. At the 0.05 level of significance, do these results contradict the findings of the GM study?

69. The average hourly pay rate at K-Mart is said to be at least $8.50 per hour. A random sample of 15 employees finds a sample mean of $8.15 and a sample standard deviation of $0.90. At the 0.05 level of significance, do the data refute the claim?

70. A coin is tossed 75 times, resulting in 30 heads. At the 0.05 level of significance, can we conclude that the coin is biased?

71. A population is assumed to be $n(50.0, 10.8)$. A random sample of size n = 28 yields a sample mean of 52.6 and a sample standard deviation of 14.3. At the 0.02 level of significance, can we conclude that the dispersion of the population is more than stated?

72. The scores on a placement test given to college freshmen for the past five years are approximately normally distributed with a mean of $\mu = 76$ and a variance of $\sigma^2 = 10$. Would you still consider $\sigma^2 = 10$ to be a valid value of the variance if a random sample of 20 students who take this placement test this year obtain a sample variance of $s^2 = 20$? Provide both a hypothesis test and a confidence interval justification.

73. An important manufacturing process produces cylindrical automotive component parts. The process must produce components with radius 5.0 mm. As chief engineer of the project, you decide to test the accuracy of the parts by measuring the radii of 100 randomly selected component parts. If it is known that the population standard deviation is $\sigma = 0.12$ mm and a mean of 5.035 mm is obtained from the sample, would you feel that the sampling has provided data in support or refutation of the process requirements? Provide both a hypothesis test and a confidence interval justification.

74. Suppose that, in the past, 40% of all adults favored capital punishment. Is there reason to believe that there has been a significant change in the support of capital punishment if a random sample of 50 adults finds that 14 favor capital punishment?

75. Five years ago, the running times of first-run movies in theaters were distributed $n(96.5, 10.3)$ minutes. Earlier this year, 25 first-run movies were sampled and were found to have a sample mean of $\bar{x} = 92.9$ minutes with a sample standard deviation of s = 13.7 minutes.

   a. At the 0.05 level of significance, do the data seem to indicate that the running times of first-run movies have changed significantly over the past 5 years?
   b. At the 0.08 level of significance, have the running times of first-run movies become significantly more variable over the past 5 years?

76. A machine that dispenses 8.0-ounce cups of coffee in a large office complex needs to be adjusted if the variance in the cup fill exceeds 1.15 ounces. If a random sample of 25 cups of coffee from this machine has a sample mean of $\bar{x} = 7.92$ ounces with a standard deviation of 1.28 ounces. At the 0.05 level of significance, is there statistical evidence that this machine needs adjustment?

77. A commonly prescribed drug for relieving nervous tension is believed to be 60% effective. Experimental results with a new drug administered to a random sample of 72 adults who were suffering from nervous tension show that 50 received relief. Is this sufficient evidence to conclude that the new drug is superior to the one commonly prescribed? Assume a 0.04 level of significance.

78. GEICO, in their television ads, claims that the average savings for new customers. when compared to their previous insurers, is $189. You randomly sample 30 new GEICO customers and find an average savings of $181.25 with a standard deviation of $20.50. At the 0.05 level of significance, do the data refute the GEICO claim?

79. A pH meter is suspected to need calibration. A sample of 10 readings for 10 neutral substances (pH = 7.00) yields the following results:

7.07	7.00	7.10	6.97	7.00	7.03	7.01	7.01	6.98	7.08

At the 0.06 level of significance, can we conclude that the pH meter is significantly out of calibration with respect to neutral substances?

Power of the Test and Sample Size Estimation

80. The instructor in a course you are taking states that the exams in the course are always distributed n (75.0, 10.0). You believe, after talking to a few friends, that the actual average this semester is significantly lower. You decide to collect a random sample of 25 current exam grades. How likely is it that you will be able to reject the instructor's claim if the actual average on the exam was 70.0 using a 1-sided test and a 0.02 level of significance?

81. A local candidate claims that he has the support of 58% of the voters. How likely is it that you will be able to show that his belief is in error with a sample size of 75 if his actual support is really 50%? Assume a 0.05 level of significance and both 1- and 2-sided tests.

82. What level of significance would be necessary to obtain 90% power using a sample of size 50 designed to detect a difference of 5.5 points in a placement test score if the variance in the scores of that test is 169? Assume both a 1- and 2-sided test.

83. A coin is suspected of being biased. A test is to be performed that will have 85% power. If the level of significance is 0.05 and the sample size is 80, determine the difference(s) that the test is designed to detect using a 1-sided test.

84. The instructor in a course you are taking states that the exams in the course are always distributed n (75.0, 10.0). You believe, after talking to a few friends, that the actual average this semester is significantly lower. You decide to collect a random sample of current exam grades. What sample sizes will be needed to have a probability of 0.9 of rejecting the instructor's claim if the actual average on the exam was 70.0 using both a 1-sided and 2-sided test with a 0.02 level of significance?

85. A local candidate claims that he has the support of 58% of the voters. What sample size will be needed if you want to have an 85% chance of being able to show that his belief is in error if his actual support is really 50%? Assume a 0.05 level of significance and both 1- and 2-sided tests.

86. A placement test was designed so that the grades would be $n(75,15)$. It has been decided that the exam will need to be rewritten if the exam average drops by at least 5 points. What sample size would be necessary to obtain 90% power? Assume a 0.03 level of significance and a 1-sided test.

87. A coin is suspected of being biased. A test is to be performed that will have 88% power. If the level of significance is 0.05, determine the sample size needed to detect a difference of 0.06 in the probability of the occurrence of this face using a 2-sided test.

Paired-t Test

88. A test is given in a psychology class on memory-enhancement techniques. The test involves a list of 50 words. The subject is shown the list for 2 minutes. One minute after the list is removed from the subject's view, the subject is asked to recite as many of the words as he or she can remember. Shortly after the test is given, a 1-hour seminar on how to increase your recollection is offered to each subject. Those taking the seminar are then allowed to take the test again. Results from 6 participants are given below, where the scores given represent the number of words correctly remembered:

Subject	A	B	C	D	E	F
First Test Score	18	12	21	15	16	23
Second Test Score	33	21	30	26	26	32

The psychology department claims that this seminar will increase a subject's score by at least 12 words. At the 0.05 level of significance, do the data refute the claim? Construct a complementary confidence interval and explain how it supports the hypothesis decision.

89. A physician would like to know whether or not oral contraceptives (OC) affect the systolic blood pressure (BP) of women using them. The physician uses the records of 10 female patients who have used OC during the past year, seeking their systolic BP prior to, and during, the use of OC. The data are given below:

Woman	A	B	C	D	E	F	G	H	I	J
BP prior to OC	115	112	107	119	115	138	126	105	104	115
BP during use of OC	125	115	108	128	121	144	134	110	102	117

At the 0.02 level of significance, are the results significant? Construct a complementary confidence interval and explain how it supports the hypothesis decision.

90. A new weight-loss program guarantees that if you weigh more that 200 pounds, you will be able to lose at least 20 pounds during the first month of participation in this program. Sample data from 6 participants in this program are given below.

Subject	A	B	C	D	E	F
Weight Prior to Program	223	216	207	243	235	250
Weight One Month Into Program	214	199	186	230	221	228

a. At the 0.05 level of significance, do the collected data support the claim?
b. Construct the confidence interval complementary to the performed hypothesis test and state how it supports the hypothesis decision.

91. A topic of current interest in ophthalmology is whether or not spherical refraction is different between the left and right eyes. For this purpose, refraction is measured in both eyes of 17 people. The data are given below.

	Spherical Refraction (diopters)	
Person	Right Eye	Left Eye
1	+1.75	+ 2.00
2	- 4.00	- 4.00
3	- 1.25	- 1.00
4	+ 1.00	+ 1.00
5	- 1.00	- 1.00
6	- 0.75	+ 0.25
7	- 2.25	- 2.25
8	+ 0.25	+ 0.25
9	0	+ 0.50
10	- 1.00	- 1.25
11	+ 0.50	- 1.75
12	- 8.50	- 5.00
13	+ 0.50	+ 0.50
14	- 5.25	- 4.75
15	- 2.25	- 2.50
16	- 6.50	- 6.25
17	+ 1.75	+ 1.75

Carry out the implied hypothesis test.

92. To answer the question as to whether or not an environment containing 10% more carbon monoxide (CO) than normal increases a healthy male's breathing rate while under stress, nine subjects were first exposed to the CO-rich environment, and then after an adequate resting period, exposed to a CO-normal environment. In each of these environments, each subject was asked to perform the same strenuous task. The breathing rates, in breaths per minute, are given below:

Subject	1	2	3	4	5	6	7	8	9
With CO	30	45	26	25	34	51	46	32	30
Without CO	30	37	22	22	30	49	40	34	28

At the 0.05 level of significance, are breathing rates significantly higher in the presence of excess CO?

Two-Sample Hypothesis Testing

93. The hydrocarbon emissions are known to have decreased dramatically during the 1980s. A study was conducted to compare the hydrocarbon emissions at idling speed, in parts per million (ppm), for automobiles of 1980 and 1990. Twenty cars of each year model were randomly selected and their hydrocarbon emission levels were recorded.

1980 levels	141	359	247	940	882	494	306	210	105	880
	200	223	188	940	241	190	300	435	241	380
1990 levels	140	160	20	20	223	60	20	95	360	70
	220	400	217	58	235	380	200	175	85	65

Using a 0.05 level of significance and assuming the both populations are normally distributed:

a. Are the emission levels of the 1990 cars more consistent than those of the 1980 cars?
b. Construct the complementary confidence interval for the performed hypothesis test and explain how it supports the hypothesis decision.

94. Two types of instruments for measuring the amount of sulfur monoxide (SO) in the atmosphere are being compared in an air-pollution experiment. It is desired to determine whether the two types of instruments are comparable in the consistency of their measurements. The following readings were recorded for the two instruments:

	SO Levels								
Instrument A	0.86	0.82	0.75	0.61	0.89	0.64	0.81	0.68	0.65
Instrument B	0.87	0.74	0.63	0.55	0.76	0.70	0.69	0.57	0.53

a. Perform the indicated hypothesis test at the 0.05 level of significance. Assume that both populations are normally distributed.

b. Construct the complementary confidence interval for the performed hypothesis test and explain how it supports the hypothesis decision.

95. A study is conducted to compare the length of time (in minutes) needed for men and women to perform a specified task. Past experience indicates that the distribution of times for both men and women is approximately normal. Sample results are given below:

	Men	Women
Sample mean	14.7	16.4
Sample s.d.	6.1	5.3
Sample size	11	14

a. At the 0.05 level of significance, do the data imply that the times exhibited by women are less variable?
b. Construct the complementary confidence interval for the performed hypothesis test and explain how it supports the hypothesis decision.

96. In a 1985 study of the effectiveness of streptokinase in the treatment of patients who have been hospitalized after myocardial infarction (MI), 20 of 199 males receiving streptokinase and 18 of 97 males in the control group died within 12 months. Test for significant differences in the 12-month mortality between the 2 groups.

97. Fifteen patients were recruited to test a new blood-clotting drug. Blood samples were taken from each patient, and to each sample, 1 of 2 blood-clotting drugs was added. At the 0.06 level of significance, does the new medication appear to clot blood significantly faster than the current standard medication?

Control Medication	9.9	8.8	11.1	9.6	8.6	10.4	9.5	
New Medication	8.7	8.8	8.6	7.9	8.7	9.1	9.2	8.6

98. A study of the relationship between salt intake and blood pressure of infants is in the planning stages. A pilot study is done, comparing five 1-year-old infants on a high-salt diet with seven 1-year-old infants on a low-salt diet. The results are given below.

	Systolic BP
High-salt diet	84 101 87 88 100
Low-salt diet	82 80 85 97 84 74 90

a. Test the claim that 1-year-old infants being fed a high-salt diet will have a higher systolic blood pressure than will 1-year-old infants being fed a low-salt diet. Use a 0.06 level of significance.

b. What would be the power of the test if a 5-point difference between the systolic blood pressures of the 2 diets if each sample size were to be 25? Assume a 0.05 level of significance and that the sample variances can be taken as reflective of the population values.

c. What equal sample sizes would be needed to achieve 95% power at the 0.05 level of significance? Assume that the sample variances can be taken as reflective of the population values.

99. Two types of gas additives were to be compared by measuring the distance one car traveled when the additive was added to 2 gallons of gas. Additive A was measured on 12 trials yielding an average distance traveled (to the nearest mile) of 85 miles with a standard deviation of 4 miles. Additive B was measured 10 times yielding an average distance traveled of 81 miles with a standard deviation of 5 miles. The makers of Additive A claim that their product, in a test such as this, will yield more than 2 additional miles beyond that obtained through Additive B. At the 0.05 level of significance do the data support the claim made by the manufacturer of Additive A?

100. Two sections of the same course use entirely different formats for the final exam. In Section 1, the exam is multiple-choice while in Section 2 the exam is essay. A random sample of times needed to complete the exam are given below:

Section	Time needed (in minutes)						
1	67	51	63	74	57		
2	81	165	97	134	92	87	114

At the 0.05 level, can we conclude that the time needed to do the exam given in Section 2 averages more than 30 minutes more than that for Section 1?

101. In 1979, two groups of childless wives, ages 25 to 29, were selected at random and each wife was asked if she eventually planned to have at least one child. One group was composed of women married less than 2 years and the other group was composed of women married at least 5 years. Of the 300 women married less than 2 years, 240 planned to have at least one child. Of the 400 women married at least 5 years, 288 planned to have at least one child. Can we conclude that the women in this age group married less than 2 years are significantly more likely to want children than are women married at least 5 years?

Analysis of Variance (ANOVA)

For each of the given problems, assume that the populations are approximately normally distributed and have approximately equal variances.

102. Three different metallurgical processing facilities are to be compared based on the percentage of impurities in their finished product. The sample data collected are given in the table the follows. At the 0.05 level, do the data suggest differences in the average percent impurities at the 3 facilities?

Facility		
1	2	3
1.06	1.58	1.19
0.79	1.45	0.96
0.82	0.57	1.34
0.89	1.16	1.45
1.05	1.12	1.51
0.95	0.91	1.33
0.65	0.83	1.41
1.15	0.43	1.00
1.12		1.43
		1.24
		0.96
		0.99
		1.07
		1.07
		1.50

103. Five sections of the same course are being taught this term on campus. You would like to know if there are any differences between them with respect to the average grades on the last exam. The following data have been collected:

	Section				
	1	2	3	4	5
sample size	8	5	9	11	7
sample mean	75.3	80.1	77.1	79.4	83.0
sample s.d.	5.2	6.1	6.7	5.8	7.4

At the 0.05 level of significance, do there appear to be any differences in the class averages?

104. An emergency room doctor would like to determine if the time needed (in minutes) before a non-critical patient is treated varies by the day of the week. She has collected the following data:

	Sun.	Mon.	Tues.	Wed.	Thurs.	Fri.	Sat.
Time Needed Before Treatment Began (in minutes)	15	16	22	17	14	14	18
	16	14	24	16	17	15	19
	12	12	26	17	10	22	21
	22	21	18	19	13	11	18
	15	20	23	12	17	13	20
	14	19	12	25	9	20	11

105. You wish to test the claim that the amount of running done by a person and their serum cholesterol level are related. To test this supposition, you have collected the following data:

	Weekly Miles Run			
	Less than 10	10 - 20	20 - 25	more than 25
sample size	6	6	6	6
sample mean cholesterol level	194	189	184	178
sample var. of cholesterol levels	175	185	203	220

At the 0.05 level of significance, is the serum cholesterol level related to the number of miles run weekly?

106. You wish to compare the gas mileage obtained from 3 different brands of gasoline, all of which have the same octane rating. The laboratory in which the testing takes place uses 15 identical motors running at identical speeds. The total mileage obtained using exactly 10 gallons of gasoline is given in the table below;

Brand of Gasoline	Mileages Obtained				
Brand A	220	251	226	246	260
Brand B	244	235	232	242	225
Brand C	252	272	250	238	256

At the 0.05 level of significance, are there significant differences in the mileages obtained from the 3 brands of gasoline?

MA 262 Homework Answers

Homework answers will be provided for most problems. (In general, those problems having no fixed answer(s) will not be listed here.) This list will be expanded as time permits.

Basic Introduction

Problem Number	Answer(s)
3a	200; 220
3b	120; 130
3c	220; 240
3d	240; 260
4	mean and median are unchanged
5	mean = 53; median cannot be determined
6	variance = 32.9
7	mean = 124; median = 120; variance = 224; s.d. = $4\sqrt{14}$
12	$90 \pm \frac{5}{2}\sqrt{74}$; no solution; $\frac{10}{3}\sqrt{47}$

Binomial Distribution

Problem Number	Answer(s)
18	184,756
19	22
20a	0.2097
20b	0.8202
21	2-engine plane
22	2^k

Poisson Distribution

Problem Number	Answer(s)
23	13
24a	0.1912
24b	0.6247
24c	0.4075
25	0.1872 (exact), 0.1882 (approximate)
26	0.0055
27	0.9600

Normal Distribution

Problem Number	Answer(s)
28a	0.8749
28b	0.2808
28c	0.9702
28d	0.1594
28e	0.8733
28f	0.8230
28g	0.0768
28h	0.4926
28i	0.0242
28j	1.175
28k	0.412
28l	0.674
28m	-0.915
28n	131.1
28o	55.52
28p	0.358
28q	-0.772
28r	0.81
28s	0.21
28t	0.49
28u	0.50
28v	1.000
29a	0.8185
29b	0.3307
29c	0.7224
29d	0.6631
29e	99.7
29f	6.75
29g	0.619

Normal Applications

Problem Number	Answer(s)
30	0.9544
31	0.392
32	$534.3 \le A$, $471.5 \le B < 534.4$, $416.8 \le C < 471.5$, $382.4 \le D < 416.8$, $F < 382.4$
33	70.7
34	5.0 years
35a	$45,351
35b	$52,493
35c	Between 61st and 62nd percentiles

Student-t Distribution (Confidence Interval Problems)

Problem Number	Answer(s)
50	$74.02 < \mu < 81.18$
51	$0.45 < \mu < 0.69$
52	$1.76 < \mu < 4.24$
53	$21.50 < \mu < 27.50$
54	$74.02 < \mu < 81.18$
55	96.66% confidence

Chi-Square Distribution (Confidence Interval Problems)

Problem Number	Answer(s)
58	$4.0 < \sigma^2 < 110.3$, $2.0 < \sigma < 10.5$
59	$1.72 < \sigma < 8.28$
60	$98.49 < \sigma^2 < 293.44$, $90.37 < \sigma^2 < 333.31$, $82.03 < \sigma^2 < 388.94$

F Distribution (Confidence Interval Problem)

Problem Number	Answer(s)
64	$0.81 < \frac{\sigma_1^2}{\sigma^2} < 7.66$, $0.66 < \frac{\sigma_1^2}{\sigma^2} < 9.75$

One-Sample Hypothesis Testing (Partial Answers)

Problem Number	Answer(s)
65a	$(H_0 : \pi \geq 0.45, H_1 : \pi < 0.45)$
65f	$(H_0 : \mu = 14.2, H_1 : \mu \neq 14.2)$
66a	0.0301
66d	0.0700
66e	0.0346
67a	$z_{0.05} = -1.645$
67d	$\chi^2_{0.02,14} = 5.368$
68	p = 0.0745
69	p = 0.0771
70	p = 0.0833
71	p = 0.0091
72	p = 0.0119
73	p = 0.0035
74	p = 0.0413
75a	p = 0.0805
75b	p = 0.0115
76	p = 0.0813
77	p = 0.0509
78	p = 0.0474
79	p = 0.1058

Paired-t Test (Partial Answers)

Problem Number	Answer(s)
88	$p = 0.0889$
89	$p = 0.0031$
90a	$p = 0.0532$
90b	$11.9 < \mu_d < 20.1$
91	$p = 0.1423$
92	$p = 0.0085$

Two-Sample Hypothesis Testing (Partial Answers)

Problem Number	Answer(s)
93a	$p = 0.0002$
93b	$2.19 < \sigma_{1989}^2 / \sigma_{1990}^2 < 14.00$
94a	$p = 0.8453$
94b	$0.260 < \sigma_B^2 / \sigma_A^2 < 5.112$
95a	$p = 0.3117$
95b	$0.50 < \sigma_W^2 / \sigma_M^2 < 3.54$
96	$p = 0.0400$
97	$p = 0.0115$
98a	$p = 0.0677$
98b	skip
98c	skip
99	$p = 0.1605$
100	$p = 0.0934$
101	$p = 0.0075$

Analysis of Variance (Partial Answers)

Problem Number	Answer(s)
102	$p = 0.0281$
103	$p = 0.1792$
104	$p = 0.0852$
105	$p = 0.2614$
106	$p = 0.1172$