

Introduction to Statistics
Math 1040
Sample Final Exam - Chapters 1-11
6 Problem Pages
Time Limit: 1 hour and 50 minutes
Open Textbook Calculator Allowed: Scientific

Name: _____

The point value of each problem is in the left-hand margin. You must show your work to receive any credit, except in problem 1. Work neatly.

(20) 1. Fill in the blanks.

- (a) The probability of getting a HHT in three consecutive flips of a fair coin is _____ .
- (b) The median of the data 2, 4, 6, 7, 8, 9, 9, 10 is _____ .
- (c) The value of correlation coefficient r is between _____ and _____ .
- (d) If Z is a standard normal random variable and $P(Z < z_0) = 0.995$, then $z_0 =$ _____ .
- (e) For a t distribution with 32 degrees of freedom, $t_{0.01} =$ _____ .
- (f) If E and F are mutually exclusive events, then $P(E \text{ and } F) =$ _____ .
- (g) If a guy has a choice of 3 neck ties, 2 shirts and 5 pairs of pants for an outfit, he can wear _____ different outfits.
- (h) Of two normal distributions with the same mean, the one with a taller graph has a _____ standard deviation.
- (i) We can decrease the margin of error for a confidence interval for the population mean by _____ the sample size.
- (j) For a two-tailed test if the null hypothesis is $H_0 : p = 0.2$, then the alternative hypothesis is $H_1 :$ _____ .

(8) 2. According to study done by the Pew Research Center, 39% of adult Americans believe that marriage is now obsolete. Suppose a sample of 500 adult Americans are asked whether marriage is obsolete. Let \hat{p} be the proportion of adult Americans who believe marriage is obsolete. Compute the mean and standard deviation of the random variable \hat{p} .

(10) 3. (a) Use the random number table to choose eight numbers between 0 and 20 in random. In this process some numbers may be repeated. Explain how to get started and how to pick numbers.

(b) Find the sample mean of the numbers you have chosen in part (a). State the formula used.

(c) Find the sample variance of the numbers you have chosen in part (a). State the formula used.

(10) 4. Dr. Paul Oswiecinski randomly selectes 40 of his 20- to 29-year-old patients and obtains data regarding their serum HDL cholestrol. Construct a relative frequency histogram of the data listed below.

73	70	70	69	66	63	62	60
60	58	56	56	56	55	54	53
53	52	52	51	51	51	50	49
48	48	48	46	46	45	45	44
44	39	38	36	35	33	32	28

- (10) 5. Construct a box plot for the following data and identify any outliers. Show your work.

9, -2, 7, 13, 3, 9, 7, 9, 8, 6, 7

- (10) 6. Suppose the average time for an oil change in a national chain is 35 minutes with standard deviation of 12 minutes. What is the approximate probability that a random sample of 42 oil changes results in sample mean oil change time of less than 38 minutes? Note: Express this probability in mathematical symbols and show your work.

- (10) 7. For a simple random sample of 30 years, the mean amount of annual snowfall in Ogden was 18.3 inches, with a standard deviation of 2.5 inches. Find a 90% confidence interval for the Ogden's mean amount of snowfall. Note: State the formula before calculating the confidence interval.

- (10) 8. A researcher wishes to check whether the percentage of Americans who support abolishing the penny has changed from 15%. What size sample should be obtained if she wishes to limit the margin of error to 2% with probability of 98%? Note: State the formula before calculating the sample size.

- (10) 9. The muzzle velocity, in feet per second, from a gun is measured using two different measuring devices and listed below.

Observation	1	2	3	4	5	6	7	8	9	10
A	793.8	793.1	792.4	794.0	791.4	792.4	791.7	792.3	789.6	794.4
B	793.2	793.3	792.6	793.8	791.6	791.6	791.6	792.4	788.5	794.7

Assume the data is approximately normal with no outliers. Construct and interpret a 95% confidence interval for the mean difference of the measured muzzle velocity of the two devices.

- (10) 10. Suppose a score of more than 21 on the mathematics portion of the ACT means a student is ready for college level mathematics. A high school program is designed to prepare students for college. Suppose a random sample of 200 students who completed this program had a mean ACT Math score of 22.1 with a standard deviation of 4.5. Does this data suggest students who complete this program are ready for college level mathematics at the $\alpha = 0.01$ level of significance? Note: State the null and alternative hypotheses and the test statistic.
- (10) 11. Suppose 70% of parents rate their children as “above average”. Suppose 15 parents are randomly selected. Find the probability that between 11 and 13 parents, inclusive, rate their children as a “above average”. Note: You must state the probability distribution and/or formula(s).
- (10) 12. A standard 52-card deck contains four 10's and twelve face cards (jacks, queens and kings). What is the probability of getting three 10's and two face cards in a 5-card poker hand? Note: You must explain your work through words and/or formula(s).

- (10) 13. A double blind study is conducted to compare side effects of a new drug and the standard drug. 107 of 710 subjects who are given the new drug display a rash side effect. While, 67 of 611 subjects who are given the standard drug display a rash side effect. Does the evidence suggest that a higher proportion of people taking the new drug experience rash as a side effect at the $\alpha = 0.01$ level of significance?

- (12) 14. The following table represents the per capita gross domestic product x , in thousands of US dollars, and the average life expectancy y of the population for several countries.

	Austria	Belgium	Finland	France	Germany	Ireland	Italy	Netherlands	Switzerland	United Kingdom
x	21.4	23.2	20.0	22.7	20.8	18.6	21.5	22.0	23.8	21.2
y	77.48	77.53	77.32	78.63	77.17	76.39	78.51	78.15	78.99	77.37

Given that $\bar{x} = 21.52$, $\bar{y} = 77.754$, $s_x = 1.5317$ and $s_y = 0.7948$. Find the correlation coefficient r between x and y . Is there a linear relationship between x and y ? If yes, find the least-squares regression line to describe y as a function of x .

Introduction to Statistics

Math 1040

Additional (not on the Sample Exam) Review Problems for Chapters 1-11

This is not an exhaustive list of all possible types of problems.

1. Fill in the blanks.

- (a) The probability of obtaining three heads in a row when flipping a fair coin is _____ .
- (b) The 1st quartile of the data 2, 4, 6, 7, 8, 9, 9, 10 is _____ .
- (c) The value of ${}_9P_7$ is _____ .
- (d) If Z is a standard normal random variable, then $P(Z > 0.25) =$ _____ .
- (e) For a t distribution with 16 degrees of freedom, $t_{0.05} =$ _____ .
- (f) If an experiment has three disjoint outcomes A , B and C with $P(A) = 0.2$ and $P(B) = 0.5$, then $P(C) =$ _____ .
- (g) There are _____ way to choose a committee of size 5 from 8 people.
- (h) The type of error made when the null hypothesis is true and in the test it is rejected is _____ .
- (i) For a right-tailed test if the null hypothesis is $H_0 : p = 0.2$, then the alternative hypothesis is $H_1 :$ _____ .
- (j) A matched-pairs design is an experimental design in which the experimental units are _____ .
- (k) If E and F are independent events, then $P(E \text{ and } F) =$ _____ .

2. True or False.

- () (a) The level of confidence represents the expected proportion of intervals that will contain the parameter if a large number of different samples is obtained.
- () (b) A data set will always have exactly one mode.
- () (c) The 75th percentile is the value that separates the lower 75% of data from the upper 25%.
- () (d) It is possible that two variables x and y have a nonlinear relationship and the correlation coefficient r between them to be close to zero.
- () (e) If, based on empirical results, the probability of an event is approximated to be zero, then that event will never occur.
- () (f) $P(E) + P(E^c) = 1$.
- () (g) The area to the right of $t = 1.316$ in the t -distribution with 25 degrees of freedom is 0.10.
- () (h) If Z is a standard normal variable and $P(Z < z_0) = 0.747$, then $z_0 = 0.665$.
- () (i) Suppose a fair coin is flipped 101 times. Let A be the event that each of the first 100 tosses results in a head. Let B be the event that the 101st toss results in a tail. Events A and B are independent.

2. Continued

- () (j) For any population which is not normal, the sample mean of small samples, say size 10, is normally distributed.
- () (k) If the results of the sample are not representative of the population, then the sample has bias.
- () (l) Generally, the goal of an experiment is to determine the effect that treatments will have on the response variable.
- () (m) In a uniform distribution, the value of the probability density function is constant.
- () (n) If the scatter diagram of the residues has a pattern, then the least-squares regression line is not the appropriate model.
- () (o) The level of significance, α , is the probability of making Type I error.

3. The Gallup organization conducted a survey in December 2005 in which 1257 randomly selected adult Americans who use Internet were asked, "How often do you, yourself, use the Internet to buy or sell products in online auctions, such as e-Bay?" Construct a pie chart of the result of the survey shown below.

Response	Frequency
Frequently	541
Occasionally	123
Rarely	131
Never	462

4. Construct a stem and leaf plot for the following 50 measurements.

3.1 4.9 2.8 3.6 2.5 4.5 3.5 3.7 4.1 4.9
 2.9 2.1 3.5 4.0 3.7 2.7 4.0 4.4 3.7 4.2
 3.8 6.2 2.5 2.9 2.8 5.1 1.8 5.6 2.2 3.4
 2.5 3.6 5.1 4.8 1.6 3.6 6.1 4.7 3.9 3.9
 4.3 5.7 3.7 4.6 4.0 5.6 4.9 4.2 3.1 3.9

5. The following is a sample of salaries for nine players in NBA for the 2000-2001 season, in thousands of dollars.

\$19,290, \$10,130, \$6,500, \$4,800, \$3,400, \$3,380, \$2,250, \$1,760, \$1,200

Find the sample mean and standard deviation of this data.

6. How many different eight-letter passwords can be formed from the letters in the word ENGINEER?
7. Consider a group of 7 people.
 - (a) In how many ways can they be arranged in a row?
 - (b) In how many ways can we pick a committee of size 5 from this group?
 - (c) In how many ways can a president and a vice-president be chosen from this group?
8. (a) Use the random number table to choose six numbers between 0 and 9 in random. In this process some numbers may be repeated. Explain how to get started and how to pick numbers.
 (b) Find the sample mean of the numbers you have chosen in part (a). State the formula used.
 (c) Find the sample variance of the numbers you have chosen in part (a). State the formula used.
9. Twenty-five households are polled in a marketing survey, and table below lists the numbers of quarts of milk purchased during a particular week. Construct a relative frequency histogram of this data.

0	3	5	4	3	2	1	3	1	2	2	2	2
1	1	2	0	1	4	3	2	2	2	3	4	

10. According to the Current Population Survey, the mean household income in the United States in 1997 was \$45,127, with standard deviation of \$31,570. What is the approximate probability that a random sample of 49 homes results in a sample mean household of more than \$41,384? Note: Express this probability in mathematical symbols and show your work.
11. If individuals die randomly with respect to their birthdays, then we expect 25% to die during any given three-month period. However, a random sample of 747 obituaries in a SLC newspaper, we find that only 8.0% of decedents died in the three-month prior to their birthdays. Is this sufficient evidence to claim that individuals tend to die at a lower rate in the three-month before their birthdays, at the $\alpha = 0.05$ significance level?
12. Two years ago, 85 high school students were randomly selected to participate in new mathematics courses. That year, the average scores of all students on the SAT-I mathematics exam was 494, while the selected students had an average of 510 with the standard deviation of 124. Can it be claimed at the $\alpha = 0.05$ level of significance that the new mathematics courses resulted in an improvement on the SAT-I mathematics scores? Note: State the null and alternative hypotheses and the test statistics.
13. Suppose 60% of students rate themselves “good” students. Suppose 15 students are randomly selected. Find the probability that between 8 and 10 students, inclusive, rate themselves as a “good” student. Note: You must state the probability distribution and/or formula(s).

14. The packaging of an E.P.T. Pregnancy Test states that the test is “99% accurate at detecting typical pregnancy hormone levels.” Assume that the probability that a test will correctly identify a pregnancy is 0.99 and that 12 randomly selected pregnant woman with typical hormone levels are each given the test.
- What is the probability that all 12 tests will be positive?
 - What is the probability that at least one test will not be positive?
15. A bag contains 5 red balls numbered one through five and 6 blue balls numbered one through six. What is the probability that a ball chosen in random is a one or blue? Note: You must explain your work through words and/or formula(s).
16. The length of human pregnancies is normally distributed with mean $\mu = 266$ days and standard deviation $\sigma = 16$ days. Also, suppose an unusually long pregnancy is the one that is in the top 2%.
- What percent of pregnancies last between 242 and 262 days?
 - Determine the length of pregnancy that separates an unusually long pregnancy from one that is not unusually long (98th percentile).
17. Let X be a discrete random variable with the probability $P(X = x) = P(x) = \begin{cases} \frac{1}{8} & , \text{if } x = 1, 2 \\ \frac{1}{4} & , \text{if } x = 3 \\ \frac{1}{2} & , \text{if } x = 4 \end{cases}$.
- Draw the probability histogram of the given probability function.
 - Find the expected value or mean of X , μ_x .
 - Find the variance of X , σ_x^2 .
18. Suppose an AutoZone store receives a shipment of size 100 alternators from the manufacturer. The store manager checks 3 alternators in random and if one or more are defective, then shipment is rejected. If there are four defective alternators in the shipment, what is the probability of the shipment being rejected? Note: You must explain your work through words and/or formula(s).
19. Every Monday, the Energy Information Administration (EIA) determines the national average gasoline price by collecting retail prices for gasoline from a sample of 900 retail gasoline outlets from across the nation. On July 14, 2008, the EIA reported the national average retail price for regular-grade gasoline to be \$4.113 per gallon. Assuming that the population standard deviation is $\sigma = \$0.110$ per gallon, construct and interpret a 95% confidence interval for the national mean price per gallon for regular-grade gasoline on July 14, 2008.
20. Among children between the ages of 6 months and 5 years, blood lead levels of 16.0 mg/l are considered “normal”. The mean and standard deviation of blood lead level of 12 randomly chosen students in a school were 18.65 and 5.049, respectively. At the $\alpha = 0.05$ level of significance, can it be concluded that the chosen students have a higher than normal blood lead levels? Assume the blood lead level is approximately normally distributed. Note: State the null and alternative hypotheses and the test statistics.

21. Draw a scatter diagram of the following data.

x	-2	-1	2	4	6	6	7
y	-3	1	0	4	6	9	8

22. The following gives amount spent on groceries per week y and the number of household members x .

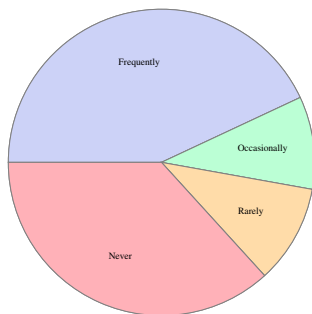
x	2	2	3	4	1	5
y	\$45.75	\$60.19	\$68.33	\$ 100.92	\$35.86	\$130.62

- (a) Given that $\bar{x} = 2.833$, $\bar{y} = 73.612$, $s_x = 1.472$, $s_y = 35.792$ and $r = 0.98$, find the least-squares regression line to describe y as a function of x .
 - (b) What is the predicted weekly grocery bill of a household of seven people?
 - (c) Should we use the value in part b to estimate the weekly grocery bill of a household of seven people? why or why not?
23. To predict the result of proposition A in the upcoming referendum, a poll was conducted and the 95% confidence interval of the proportion of likely voters in favor of it was (49.1%, 54.3%). Can we conclude that this proposition will pass? Explain!
24. A finite population consists of four elements: 1, 2, 6, 7. Samples of size two, without replacement, are taken from this population. Construct a sampling distribution for the sample means by first listing all possible samples of size two, calculating their means and then forming the probability distribution of the sample means.

Just Answers To Additional (not on the Sample Exam) Review Problems for Chapters 1-11

The following are just answers, not complete solutions!

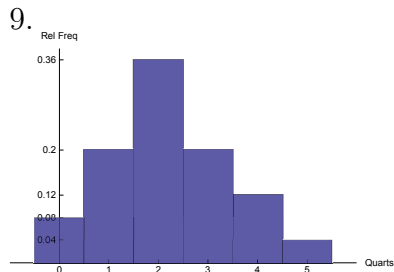
- 1. (a) $1/8$ (b) 5 (c) 181440 (d) 0.4013 (e) 1.746 (f) 0.3 (g) 56 (h) type I (i) $p > 0.2$ (j) paired (k) $P(E)P(F)$
- 2. (a) True (b) False (c) True (d) True (e) False (f) True (g) True (h) True (i) True (j) False (k) True (l) True (m) True (n) True (o) True
- 3. Note: Be sure to show calculations of angles!
- 4. Leaf Unit = 0.1



1	68
2	1255578899
3	1145566677778999
4	000122345678999
5	11667
6	12

5. $\bar{x} = \$5,856.67$; $s = \$5,743.37$
 6. $\frac{8!}{3! \times 2!} = 3360$
 7. (a) $7! = 5040$ (b) ${}_7C_5 = 21$ (c) ${}_7P_2 = 42$

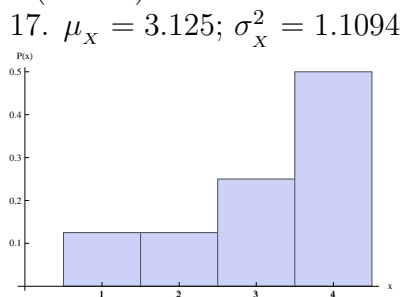
8. Start at a random place on the Random Numbers table, by dropping a pencil on it without looking. The starting place was row 13, column 20. We will read that digit and the 5 digits to its right. These numbers are 9, 7, 3, 6, 3 and 4. $\bar{x} = 16/3$, $s^2 = 88/15$.



10. $P(\bar{x} > 41384) = 1 - 0.2033 = 0.7967$.
 11. $H_0 : p = 0.25$, $H_1 : p < 0.25$; $z_0 = -10.7$;
 The data supports the claim since $z_0 = -10.7 < -1.645 = -z_{0.05}$.
 12. $H_0 : \mu = 494$, $H_1 : \mu > 494$; $t_0 \approx 1.19$;
 It can not be claimed that the improvement is due to the courses since, using nearest df = 80 value, $t_0 = 1.19 < 1.292 = t_{0.05}$.

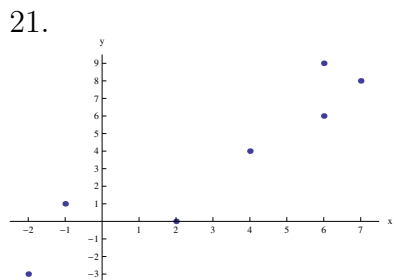
13. Let X be the number of students rate themselves “good”. X has a binomial distribution with $n = 15$ and $p = 0.6$. $P(8 \leq X \leq 10) = 0.5696$.

14. (a) $P(\text{All 12 positive}) = 0.99^{12} = 0.8864$ (b) $P(\text{At least one not positive}) = 0.1136$
 15. $P(\text{one or blue}) = P(\text{one}) + P(\text{blue}) - P(\text{one and blue}) = \frac{7}{11}$
 16. Let X be the length of a pregnancy. (a) $P(242 < X < 262) = 0.3345$. (b) Find x so that $P(X < x) = 0.98$. $x \approx 266 + 2.05 \times 16 = 298.8$



18. $P(\text{rejection}) = 1 - P(\text{all three not defective}) = 1 - \frac{{}_{96}C_3}{{}_{100}C_3} = 0.116$
 19. The bounds are $4.113 \pm 1.96 \frac{0.110}{\sqrt{900}}$.
 C.I. = (4.106, 4.120). We are 95% sure that the true mean price per gallon for regular-grade gasoline is in this interval. Or, that the 95% of all sample means for samples of size 900 for the price per gallon for regular-grade gasoline fall in this interval.

20. $H_0 : \mu = 16.0$, $H_1 : \mu > 16.0$; $t_0 = 1.82$; The chosen students do have a higher blood lead level since $t_0 = 1.82 > 1.7959 = t_{0.05}$.



22. (a) $\hat{y} = 23.826x + 6.106$ (b) 172.89
 (c) Household of seven people is outside the x values of data. So, the estimate can only be used with caution!
 23. No, since all proportions in the entire interval are not greater than 50%.

24.

Sample	\bar{x}	\bar{x}	$P(\bar{x})$
1, 2	1.5	1.5	1/6
1, 6	3.5	3.5	1/6
1, 7	4	4	1/3
2, 6	4	4.5	1/6
2, 7	4.5	6.5	1/6
6, 7	6.5		

