

Provide an appropriate response.

- 1) A meteorologist constructs a graph showing the total precipitation in Phoenix, Arizona in each of the months of 1998. Does this involve descriptive statistics or inferential statistics?
- 2) A political researcher wishes to gauge political sentiment regarding a proposed tax cut. He obtains a list of 1000 email addresses from an internet provider, uses a random number table to select a random sample of 100 of these addresses, emails the people in the sample and requests that they respond to his questions by email. Do you think that the group of people who respond is likely to be representative of all registered voters? Explain your answer.
- 3) At a college there are 120 freshmen, 90 sophomores, 110 juniors, and 80 seniors. A school administrator selects a simple random sample of 12 of the freshmen, a simple random sample of 9 of the sophomores, a simple random sample of 11 of the juniors, and a simple random sample of 8 of the seniors. She then interviews all the students selected. Identify the type of sampling used in this example.

Identify the study as an observational study or a designed experiment.

- 4) A researcher wished to assess the importance of exercise in weight-loss programs. 412 people, all considered to be at least 20 pounds overweight, volunteered to participate in a study. The participants were randomly assigned to one of two groups. Over a two-month period, the first group followed a particular diet but were instructed to perform no exercise other than walking. The second group followed the same diet but also performed aerobic exercise for one hour each day. At the end of the two months, the weight loss of each participant was recorded. The average weight loss was calculated for each group and it was found that the average weight loss for the second group was significantly greater than the average weight loss for the first group.

Identify each of the following variables as either qualitative or quantitative.

- 5) Outcome of tossing a coin.

Classify the data as either discrete or continuous.

- 6) What type of data is provided by the statement "Helen finished in 5th place in the ice dancing competition"?

For the given data, identify the variable under consideration.

- 7) Melissa has five siblings and Andrew has two siblings.

Construct a grouped–data table for the given data. Use the symbol \leq to mean "up to, but not including".

- 8) A medical research team studied the ages of patients who had strokes caused by stress. The ages of 34 patients who suffered stress strokes were as follows.

29 30 36 41 45 50 57 61 28 50 36 58
60 38 36 47 40 32 58 46 61 40 55 32
61 56 45 46 62 36 38 40 50 27

- Construct a frequency table for these ages. Use 8 classes beginning with a lower class limit of 25 and a class width of 5.
- Construct a relative frequency table.
- Use the raw data to find the mean, median, mode and the standard deviation.
- Use the frequency table to find the mean and the standard deviation. Why are the answers to (d) different to the answers in part (c)?
- Use the raw data and the calculator to find the five–number summary and draw a box–plot.
- Use your answers to (c) to find the z– score corresponding to 41.

Obtain the probability distribution of the random variable.

- 9) When a coin is tossed four times, sixteen equally likely outcomes are possible as shown below:

HHHH HHHT HHTH HHTT
HTHH HTHT HTTH HTTT
THHH THHT THTH THTT
TTHH TTHT TTTH TTTT

Let X denote the total number of tails obtained in the four tosses.

- Find the probability distribution of the random variable X . Leave your probabilities in fraction form.
- Find the mean and standard deviation for this probability distribution.

Find the indicated probability.

- 10) A multiple choice test has 30 questions, and each has five possible answers, of which one is correct. If a student guesses on every question, find the probability of getting exactly 12 correct.

Find the indicated probability.

- 11) A company purchases shipments of machine components and uses this acceptance sampling plan: Randomly select and test 25 components and accept the whole batch if there are fewer than 3 defectives. If a particular shipment of thousands of components actually has a 3% rate of defects, what is the probability that this whole shipment will be accepted?

Find the mean and the standard deviation of the binomial random variable.

- 12) A company manufactures batteries in batches of 19 and there is a 3% rate of defects. Find the mean and standard deviation for the random variable X , the number of defects per batch.

Find the specified probability distribution of the binomial random variable.

- 13) In one city, 22% of the population is under 25 years of age. Three people are selected at random from the city. Find the probability distribution of X , the number among the three that are under 25 years of age.

Use a table of areas to find the specified area under the standard normal curve.

- 14) The area that lies between -1.10 and -0.36

Use a table of areas for the standard normal curve to find the required z-score.

- 15) Find the z-score for having area 0.07 to its right under the standard normal curve, that is, find $z_{0.07}$.

Find the indicated probability or percentage for the normally distributed variable.

- 16) The variable X is normally distributed. The mean is $\mu = 15.2$ and the standard deviation is $\sigma = 0.9$. Find $P(X > 16.1)$.

Use the empirical rule to solve the problem.

- 17) The systolic blood pressure of 18-year-old women is normally distributed with a mean of 120 mmHg and a standard deviation of 12 mmHg. What percentage of 18-year-old women have a systolic blood pressure between 96 mmHg and 144 mmHg?

Find the specified percentile, quartile, or decile.

- 18) The amount of Jen's monthly phone bill is normally distributed with a mean of \$78 and a standard deviation of \$9. Find the first quartile, Q_1 .

Estimate the indicated probability by using the normal distribution as an approximation to the binomial distribution.

- 19) With $n = 20$ and $p = 0.60$, estimate $P(\text{fewer than } 8)$.

Identify the distribution of the sample mean. In particular, state whether the distribution of \bar{x} is normal or approximately normal and give its mean and standard deviation.

- 20) The heights of people in a certain population are normally distributed with a mean of 65 inches and a standard deviation of 3.4 inches. Determine the sampling distribution of the mean for samples of size 42.

Solve the problem.

- 21) Human body temperatures are normally distributed with a mean of 98.20°F and a standard deviation of 0.62°F . If 19 people are randomly selected, find the probability that their mean body temperature will be less than 98.50°F .

Find the indicated probability or percentage for the sampling error.

- 22) Scores on an aptitude test are normally distributed with a mean of 220 and a standard deviation of 30. What is the probability that the sampling error made in estimating the population mean by the mean of a random sample of 50 test scores will be at most 5 points?

Determine the margin of error in estimating the population mean, μ .

- 23) A telephone company wishes to estimate the mean duration of local calls. Given that $\sigma = 3.0$ minutes and that the sample size is 540, find the margin of error in estimating μ at the 88% level of confidence.

Find the necessary sample size.

- 24) Scores on a certain test are normally distributed with a variance of 77. A researcher wishes to estimate the mean score achieved by all adults on the test. Find the sample size needed to assure with 95.44 percent confidence that the sample mean will not differ from the population mean by more than 3 units.

Find the confidence interval specified.

- 25) 30 people are selected randomly from a certain town. If their mean age is 53.3 and $\sigma = 6.3$, find a 95% confidence interval for the true mean age, μ , of everyone in the town.

Find the confidence interval specified. Assume that the population is normally distributed.

- 26) Thirty randomly selected students took the calculus final. If the sample mean was 90 and the standard deviation was 5.5, construct a 99% confidence interval for the mean score of all students.

Perform a hypothesis test for the population mean. Assume that preliminary data analyses indicate that it is reasonable to apply the z-test or the t-test.

- 27) A newspaper in a large midwestern city reported that the National Association of Realtors said that the mean home price last year was \$116,800. The city housing department feels that this figure is too low. They randomly selected 66 home sales and obtained a sample mean price of \$118,900. Assume that the population standard deviation is \$3,700. Using a 5% level of significance, perform a hypothesis test to determine whether the population mean is higher than \$116,800.

Perform a one-sample z-test or t-test for a population mean using the P-value approach. Be sure to state the hypotheses and the significance level, to compute the value of the test statistic, to obtain the P-value, and to state your conclusion.

- 28) The forced vital capacity (FVC) is often used by physicians to assess a person's ability to move air in and out of their lungs. It is the maximum amount of air that can be exhaled after a deep breath. For adult males, the average FVC is 5.0 liters. A researcher wants to perform a hypothesis test to determine whether the average forced vital capacity for women differs from this value. The mean forced vital capacity for a random sample of 85 women was 4.8 liters. Do the data provide sufficient evidence to conclude that the mean forced vital capacity for women differs from the mean value for men of 5.0 liters? Perform the appropriate hypothesis test using a significance level of 0.05. Assume that $\sigma = 0.9$ liters.

Preliminary data analyses indicate that it is reasonable to use a z-test or a t-test to carry out the specified hypothesis test. Perform the using the critical-value approach.

- 29) A light-bulb manufacturer advertises that the average life for its light bulbs is 900 hours. A random sample of 15 of its light bulbs resulted in the following lives in hours.

995	590	510	539	739	917	571	555
916	728	664	693	708	887	849	

At the 10% significance level, do the data provide evidence that the mean life for the company's light bulbs differs from the advertised mean?

- 30) When bears were anesthetized, researchers measured the distance (in inches) around their chest and they weighed the bears (in pounds). The results are given below for eight male bears.

x Chest (in)	26	45	54	49	41	49	44	19
y Weight (lb)	90	344	416	348	262	360	332	34

- Determine whether there is a significant linear correlation.
- If so, find the regression equation.
- Find the best predicted weight of a bear with a chest size of 52 inches.

Perform a nonpooled t-test to compare the two population means. You may assume that the samples are independent and that the variable under consideration is normally distributed on both populations. The two population standard deviations cannot be assumed to be equal.

- 31) A researcher was interested in comparing the amount of time spent watching television by women and by men. Independent random samples of 14 women and 17 men were selected and each person was asked how many hours he or she had watched television during the previous week. The summary statistics are as follows.

Sample 1 (women)	Sample 2 (men)
$\bar{x}_1 = 12.6$	$\bar{x}_2 = 13.6$
$s_1 = 3.9$	$s_2 = 5.2$
$n_1 = 14$	$n_2 = 17$

Do the data provide sufficient evidence to conclude that the mean time for women is less than the mean time for men? Perform a nonpooled t-test at the 5% significance level.

Use the nonpooled t-interval procedure to obtain the required confidence interval for the difference between two population means. You may assume that the samples are independent and that the variable under consideration is normally distributed on both populations. The two population standard deviations cannot be assumed to be equal.

- 32) A paint manufacturer wishes to compare the drying times of two different types of paint. Independent random samples of 11 cans of type A and 9 cans of type B were selected and applied to similar surfaces. The drying times were recorded. The summary statistics are as follows.

Type A	Type B
$\bar{x}_1 = 76.0$	$\bar{x}_2 = 64.0$
$s_1 = 4.5$	$s_2 = 5.1$
$n_1 = 11$	$n_2 = 9$

Determine a 98% confidence interval for the difference, $\mu_1 - \mu_2$, between the mean drying time for type A and the mean drying time for type B.

Find the margin of error for the 95% confidence interval used to estimate the population proportion.

- 33) In a sample of 162 observations, there were 114 positive outcomes.

Use the given degree of confidence and sample data to construct a confidence interval for the population proportion p.

- 34) A survey of 865 voters in one state reveals that 408 favor approval of an issue before the legislature. Construct the 95% confidence interval for the true proportion of all voters in the state who favor approval.

Assume that you wish to estimate a population proportion, p. For the given margin of error and confidence level, determine the sample size required.

- 35) You wish to estimate the proportion of all voters in California who plan to vote in favor of a certain ballot measure. Obtain a sample size that will ensure a margin of error of at most 0.015 for a 99% confidence interval. Assume that it is reasonable to presume that of the voters sampled, the percentage in favor of the measure will be between 10% and 29%.

Solve the problem.

- 36) A pollster wishes to estimate the true proportion of U.S. voters that oppose capital punishment. How many voters should be surveyed in order to be 92 percent confident that the true proportion is estimated to within 0.02?

Perform a hypothesis test for a population proportion using the critical value approach.

- 37) A manufacturer considers his production process to be out of control when defects exceed 3%. In a random sample of 85 items, the defect rate is 5.9% but the manager claims that this is only a sample fluctuation and production is not really out of control. At the 0.01 level of significance, do the data provide sufficient evidence that the percentage of defects exceeds 3%?

Find the P-value for the indicated hypothesis test.

- 38) A medical school claims that more than 28% of its students plan to go into general practice. It is found that among a random sample of 130 of the school's students, 32% of them plan to go into general practice. Find the P-value for a test of the school's claim.

Use a two-sample z-test for two population proportions to perform the required hypothesis test. Use the critical value approach.

39) Use the given sample data to test the claim that $p_1 > p_2$. Use a significance level of 0.01.

<u>Sample 1</u>	<u>Sample 2</u>
$n_1 = 85$	$n_2 = 90$
$x_1 = 38$	$x_2 = 23$

A two-sample z-test for two population proportions is to be performed using the P-value approach. The null hypothesis is $H_0 : p_1 = p_2$ and the alternative is $H_a : p_1 \neq p_2$. Use the given sample data to find the P-value for the hypothesis test.

40) $n_1 = 100$ $n_2 = 100$
 $x_1 = 38$ $x_2 = 40$

Construct the indicated confidence interval for the difference between population proportions $p_1 - p_2$. Assume that the samples are independent and that they have been randomly selected.

41) $x_1 = 26$, $n_1 = 41$ and $x_2 = 34$, $n_2 = 44$; Construct a 90% confidence interval for the difference between population proportions $p_1 - p_2$.

Perform the indicated goodness-of-fit test. Be sure to state the hypotheses, to obtain the critical value, to compute the value of the test statistic, and to state your conclusion.

42) You roll a die 48 times with the following results.

Number	1	2	3	4	5	6
Frequency	3	14	2	13	2	14

Do the data provide sufficient evidence to conclude that the die is loaded (i.e., that the six numbers are not equally likely)? Perform the hypothesis test at the 0.05 level of significance.

Perform the indicated goodness-of-fit test using the P-value method. Be sure to state the hypotheses, to compute the value of the test statistic, to obtain the P-value, and to state your conclusion.

43) A die is rolled 180 times and the following data are obtained.

<u>Number</u>	<u>Frequency</u>
1	31
2	34
3	26
4	16
5	32
6	41

Do the data provide sufficient evidence to conclude that the die is loaded (i.e., that the six numbers are not equally likely)? Perform the hypothesis test at the 1% level of significance.

Use words or symbols, as indicated, to describe the event.

- 44) The following contingency table provides a joint frequency distribution for the popular votes cast in the 1984 presidential election by region and political party. Data are in thousands, rounded to the nearest thousand.

		Political Party			Total
		Demo. P_1	Repub. P_2	Other P_3	
Region	Northeast R_1	9046	11,336	101	20,483
	Midwest R_2	10,511	14,761	169	25,441
	South R_3	10,998	17,699	136	28,833
	West R_4	7022	10,659	214	17,895
	Total	37,577	54,455	620	92,652

Suppose a person who voted in the 1984 presidential election is selected at random. Describe in words the event P_2 .

Find the indicated probability.

- 45) The following contingency table provides a joint frequency distribution for the popular votes cast in the 1984 presidential election by region and political party. Data are in thousands, rounded to the nearest thousand.

		Political Party			Total
		Demo. P_1	Repub. P_2	Other P_3	
Region	Northeast R_1	9046	11,336	101	20,483
	Midwest R_2	10,511	14,761	169	25,441
	South R_3	10,998	17,699	136	28,833
	West R_4	7022	10,659	214	17,895
	Total	37,577	54,455	620	92,652

A person who voted in the 1984 presidential election is selected at random. Compute the probability that the person selected was in the West and voted Republican.

Find the conditional probability.

- 46) The following contingency table provides a joint frequency distribution for the popular votes cast in the 1984 presidential election by region and political party. Data are in thousands, rounded to the nearest thousand.

		Political Party			Total
		Demo. P ₁	Repub. P ₂	Other P ₃	
Region	Northeast R ₁	9046	11,336	101	20,483
	Midwest R ₂	10,511	14,761	169	25,441
	South R ₃	10,998	17,699	136	28,833
	West R ₄	7022	10,659	214	17,895
	Total	37,577	54,455	620	92,652

A person who voted in the 1984 presidential election is selected at random. Compute the probability that the person selected voted Democrat given that they were in the Northeast.

Perform the required hypothesis test for two population means. You may presume that the assumptions for using the pooled t-test are satisfied.

- 47) A researcher was interested in comparing the amount of time spent watching television by women and by men. Independent random samples of 14 women and 17 men were selected and each person was asked how many hours he or she had watched television during the previous week. The summary statistics are as follows.

Sample 1 (women)	Sample 2 (men)
$\bar{x}_1 = 11.3$	$\bar{x}_2 = 17.0$
$s_1 = 4.3$	$s_2 = 4.8$
$n_1 = 14$	$n_2 = 17$

Do the data provide sufficient evidence to conclude that the mean time for women is less than the mean time for men? Perform a pooled t-test at the 5% significance level.

Use the pooled t–interval procedure to obtain the required confidence interval for the difference between two population means. You may presume that the assumptions for using the pooled t–procedures are satisfied.

- 48) A researcher was interested in comparing the amount of time spent watching television by women and by men. Independent random samples of 14 women and 17 men were selected and each person was asked how many hours he or she had watched television during the previous week. The summary statistics are as follows.

Sample 1 (women)	Sample 2 (men)
$\bar{x}_1 = 11.3$	$\bar{x}_2 = 16.9$
$s_1 = 4.4$	$s_2 = 4.7$
$n_1 = 14$	$n_2 = 17$

Determine a 95% confidence interval for the difference, $\mu_1 - \mu_2$, between the mean weekly time for women and the mean weekly time for men.

Perform a chi–square independence test using the critical value approach, provided the conditions for using the test are met. Be sure to state the hypotheses, to obtain the expected frequencies, to obtain the critical value, to compute the value of the test statistic, and to state your conclusion.

- 49) A researcher performed a study to determine whether an association exists between sex and blood type. He obtained the following sample data.

		Blood Type				Total
		O	A	B	AB	
Sex	Female	152	148	40	20	360
	Male	148	122	35	15	320
	Total	300	270	75	35	680

At the 5% significance level, do the data provide sufficient evidence to conclude that an association exists between sex and blood type?

Perform the indicated goodness–of–fit test. Be sure to state the hypotheses, to obtain the critical value, to compute the value of the test statistic, and to state your conclusion.

- 50) You roll a die 48 times with the following results.

Number	1	2	3	4	5	6
Frequency	13	12	4	4	1	14

Do the data provide sufficient evidence to conclude that the die is loaded (i.e., that the six numbers are not equally likely)? Perform the hypothesis test at the 0.05 level of significance.

Answer Key

Testname: FIN-REV.ALL

- 1) Descriptive statistics
- 2) No; explanations will vary. Possible answer: the sample was obtained from among people who own a computer. That group is likely to include relatively wealthy people who are more likely to favor a tax cut. Furthermore, the group includes those who chose voluntarily to respond. People who respond voluntarily are likely to have stronger opinions than the average voter.
- 3) Stratified sampling
- 4) Designed experiment
- 5) Qualitative
- 6) Discrete
- 7) Number of siblings
- 8)

Age	Frequency	Relative Frequency
25-<30	3	.088
30-<35	3	.088
35-<40	6	.176
40-<45	4	.118
45-<50	5	.147
50-<55	3	.088
55-<60	5	.147
60-<65	5	.147

- c) mean = 44.9, median = 45, mode = 36, standard deviation 11
d) mean = 46.2, standard deviation 11.3
e) min = 27, Q1 = 36, Q2 = 45, Q3 = 56, max. = 62
f) z = -.35
- 9) a)

x	P(X = x)
0	1/16
1	1/4
2	3/8
3	1/4
4	1/16

- b) $\mu = 2, \sigma = 1$
- 10) A success is that the student answers a question correctly.
 $p = 0.2$
 $n = 30$

$$P(X = 12) = \binom{30}{12} (0.2)^{12} (0.8)^{18} = 0.0064$$

- 11) 0.9620
12) 0.57, 0.74
13)

x	P(X = x)
0	0.4746
1	0.4015
2	0.1133
3	0.0106

- 14) 0.2237
15) 1.48
16) 0.1587
17) 95.44%
18) \$71.97

Answer Key

Testname: FIN-REV.ALL

- 19) 0.0202
- 20) Normal, mean = 65 inches, standard deviation = 0.52 inches
- 21) 0.9826
- 22) 0.762
- 23) 0.201 minutes
- 24) 35
- 25) 51.05 to 55.55
- 26) 87.2 to 92.77
- 27) $H_0 : \mu = \$116,800$
 $H_a : \mu > \$116,800$
Test statistic: $z = 4.61$
Critical value: $z = 1.645$. Reject $H_0 : \mu = \$116,800$. There is sufficient evidence to support the claim that the mean is higher than \$116,800.
- 28) $H_0 : \mu = 5.0$ liters
 $H_a : \mu \neq 5.0$ liters
 $\alpha = 0.05$
 $z = -2.05$
P-value = 0.0404
Reject H_0 . At the 5% significance level, the data provide sufficient evidence to conclude that the mean forced vital capacity for women differs from 5.0 liters.
- 29) $H_0 : \mu = 900$ hours
 $H_a : \mu \neq 900$ hours
Test statistic: $t = -4.342$. Critical values: $t = \pm 1.761$. Reject $H_0 : \mu = 900$ hours. There is sufficient evidence to support the claim that the true mean life differs from the advertised mean.
- 30) a) $r = .993 > CV = .707$, significant.
b) $y = 11.3x - 187$
c) 400.6 lb
- 31) $H_0 : \mu_1 = \mu_2$
 $H_a : \mu_1 < \mu_2$
Test statistic: $t = -0.611$
Critical value: $t = -1.701$
Do not reject H_0 . At the 5% significance level, the data do not provide sufficient evidence to conclude that the mean time for women is less than the mean time for men.
- 32) 6.38 to 17.62
- 33) 0.0703
- 34) From 0.438 to 0.505
- 35) 6068
- 36) 1915
- 37) $H_0 : p = 0.03$ $H_a : p > 0.03$. Test statistic: $z = 1.57$. Critical value: $z = 2.33$.
Fail to reject the null hypothesis. There is not sufficient evidence to conclude that the percentage of defects exceeds 3%.
- 38) 0.1539
- 39) $H_0 : p_1 = p_2$ $H_a : p_1 > p_2$.
Test statistic: $z = 2.66$. Critical value: $z = 2.33$.
Reject the null hypothesis. There is sufficient evidence to support the claim that $p_1 > p_2$.
- 40) .7718
- 41) (-0.300 0.023)

Answer Key

Testname: FIN-REV.ALL

- 42) H_0 : The die is fair (all numbers occur with equal frequency).
 H_a : The die is loaded.
Test statistic: $\chi^2 = 24.25$. Critical value: $\chi^2 = 11.070$. Reject the null hypothesis.
At the 5% level of significance, there is sufficient evidence to conclude that the die is loaded.
- 43) H_0 : The die is fair.
 H_a : The die is loaded.
Test statistic: $\chi^2 = 11.8$.
 $0.025 < P\text{-value} < 0.05$.
Do not reject H_0 . At the 1% significance level, the data do not provide sufficient evidence to conclude that the die is loaded.
- 44) The person selected voted Republican.
- 45) 0.115
- 46) 0.442
- 47) $H_0: \mu_1 = \mu_2$
 $H_a: \mu_1 < \mu_2$
Test statistic: $t = -3.446$
Critical value: $t = -1.699$
Reject H_0 . At the 5% significance level, the data provide sufficient evidence to conclude that the mean time for women is less than the mean time for men.
- 48) -8.97 to -2.23
- 49) H_0 : Sex and blood type are not associated.
 H_a : Sex and blood type are associated.
Critical value: $\chi^2_{0.05} = 7.815$.
Test statistic: $\chi^2 = 1.256$
Do not reject H_0 . At the 5% significance level, the data do not provide sufficient evidence to conclude that an association exists between sex and blood type.
- 50) H_0 : The die is fair (all numbers occur with equal frequency).
 H_a : The die is loaded.
Test statistic: $\chi^2 = 19.75$. Critical value: $\chi^2 = 11.070$. Reject the null hypothesis.
At the 5% level of significance, there is sufficient evidence to conclude that the die is loaded.