This competition includes 25 multiple-choice questions and three open-response questions that might be used as tie breakers. For questions 1 through 25 (the multiple-choice questions), mark your answer choice in the appropriate location on the sheet provided. After completing questions 1 through 25, answer each tie breaker question in sequential order (i.e., complete Tie Breaker #1 first, then Tie Breaker #2, and then Tie Breaker #3 last). Be sure that your name is printed on each of the tie breaker questions. When time is called, you will be asked to turn in your multiple-choice question answer sheet and your written responses to the tie breaker questions.

1. A study employs simple random sampling but does not utilize any treatment groups. Which of the following statements would be true?
   A. Researchers can infer causality and can generalize from sample at hand to larger population
   B. Researchers can infer causality but cannot generalize from sample at hand to larger population
   C. Researchers can generalize from sample at hand to larger population but cannot infer causality
   D. Researchers cannot generalize from sample at hand to larger population and cannot infer causality
   E. None of the above

2. Assuming the population standard deviation is known, which of the following specifications would result in the narrowest confidence interval for the mean:
   A. 90% Confidence level, $n = 100$
   B. 95% Confidence level, $n = 100$
   C. 90% Confidence level, $n = 250$
   D. 95% Confidence level, $n = 250$
   E. Cannot be determined

3. In hypothesis testing, the probability of rejecting H_0 when in fact H_0 is true is denoted by which of the following:
   A. α
   B. $1 - \alpha$
   C. β
   D. $1 - \beta$
   E. None of the above.

4. A right-tailed hypothesis test is conducted. If the corresponding 90% confidence interval is entirely below the parameter being estimated in H_0, then which of the following is a correct statement?
   A. H_0 cannot be rejected at the 0.05 level.
   B. H_0 can be rejected at the 0.05 level.
   C. H_0 cannot be rejected at the 0.10 level.
   D. H_0 can be rejected at the 0.10 level.
   E. Cannot be determined

5. The probability of observing a sample mean as extreme or more extreme than currently observed given that the null hypothesis is true is represented by which one of the following?
   A. α
   B. Level of Confidence
   C. p-value
   D. Test statistic value
   E. None of the above.

6. When testing the hypothesis presented in the following, at $\alpha = 0.05$ level of significance with the t-test, where is the reject region?

$$H_0: \mu = 100$$
$$H_1: \mu \neq 100$$

   A. The upper tail
   B. The lower tail
   C. Both the upper and lower tails
   D. The hypothesis is always rejected.
   E. Cannot be determined

7. The fundamental difference between the z test and t test for testing hypothesis about a population mean is which of the following?
   A. Only z assumes the population distribution be normal.
   B. z is a two-tailed test, whereas t is one-tailed.
   C. Only t becomes more powerful as sample size increases.
   D. Only z requires the population variance be known.
   E. There is no difference between these tests.

8. The mathematical ability of 30 pre-school children was measured when they entered their first year of pre-school and then again in the spring of their kindergarten year. The test for pre- to post-mean differences, which of the following tests would be used?
   A. Independent t-test
   B. Dependent t-test
   C. z test
   D. χ^2 test
   E. None of the above

9. A group of 25 females was compared to a group of 35 males with respect to intelligence. To test if the sample means are significantly different, which of the following tests would you use?
    A. Independent t-test
    B. Dependent t-test
    C. z test
    D. χ^2 test
    E. None of the above

10. The denominator of the test statistic of an independent t test is known as the standard error of the difference between two means, and may be defined as which one of the following?
    A. The difference between two group means
    B. The amount by which the difference between the two group means differs from the population mean
    C. The standard deviation of the sampling distribution of the difference between two means
    D. The square root of the sample size
    E. All of the above

11. In a linear regression analysis, if two individuals have the same predicted score, their residual scores will be which one of the following?
    A. Be necessarily equal
    B. Depend only on their observed scores on Y
    C. Depend only on their predictor scores on Y
    D. Depend only on the number of individuals that have the same predicted score
    E. Cannot be determined

12. The statement that "A 99% confidence interval obtained from a simple random sample of 500 people has a smaller chance of containing the true population parameter than a 99% confidence interval obtained from a simple random sample of 1000 people" is:
    A. Always True
    B. Never True
    C. Sometimes True
    D. Randomly True
    E. Cannot be determined

13. Which of the following frequency distributions will generate the same relative frequency distribution?

X	Freq.	Y	Freq.	Z	Freq.
100	2	100	6	100	8
99	5	99	15	99	18
98	8	98	24	98	28
97	5	97	15	97	18
96	2	96	6	96	8

A. X and Y only
B. X and Z only
C. Y and Z only
D. X, Y, and Z
E. None of the above

14. Which of the following statements is correct for a continuous variable?
A. The proportion of the distribution below the 25^{th} percentile is 75%
B. The proportion of the distribution below 50^{th} percentile is 25%
C. The proportion of the distribution above the third quartile is 25%
D. The proportion of the distribution between 25^{th} and 75^{th} percentiles is 25%
E. All of the above

15. If in a distribution of 200 IQ scores, the mean is considerably above the median, then the distribution is which one of the following?
A. Negatively skewed
B. Symmetrical
C. Positively skewed
D. Bimodal
E. Cannot be determined

16. If a population distribution is highly positively skewed, then the distribution of the sample means for samples of size 100 will be,
A. Highly negatively skewed
B. Highly positively skewed
C. Approximately normally distributed
D. Tangentially skewed
E. Cannot be determined without further information

17. Which of the following distributions has the largest variance?

X	Freq.		Y	Freq.		Z	Freq.
15	6		15	4		15	2
16	7		16	7		16	7
17	9		17	11		17	13
18	9		18	11		18	13
19	7		19	7		19	7
20	6		20	4		20	2

A. X
B. Y
C. Z
D. The variance is the same for X, Y, and Z.
E. Cannot be determined

18. What does central limit theorem state?
A. The means of a sufficiently large number of random samples from a population will be normally distributed.
B. The raw scores of many natural events will be normally distributed.
C. Z scores will be normally distributed.
D. Sample mean trends to the population mean
E. None of the above.

19. Suppose we have a dataset containing the weights of 350 newborn puppies recorded in grams. Summary statistics were computed for this dataset. Subsequently, it was discovered that all of the original weight measurements were inaccurate by one gram. To correct this error, one gram was added to the weights of all 350 original values. Following this one-gram correction, which of the following summary statistics would change when recalculated with data from the new sample?

I. The median
II. The variance
III. Range
IV. The coefficient of variation

A. I only
B. I and II
C. I, II and IV
D. I and IV
E. I, II, III, IV

20. The following represent the results of a survey in which individuals were asked to disclose what they perceive to be the ideal number of children. If an individual who believes the ideal number of children is 4 is randomly selected, what is the probability the individual is male?

	0	**1**	**2**	**3**	**4**	**5**	**6**
Female	5	2	87	61	28	3	2
Male	3	2	68	28	8	0	0

A. 0.073
B. 0.222
C. 0.121
D. 0.149
E. Cannot be determined

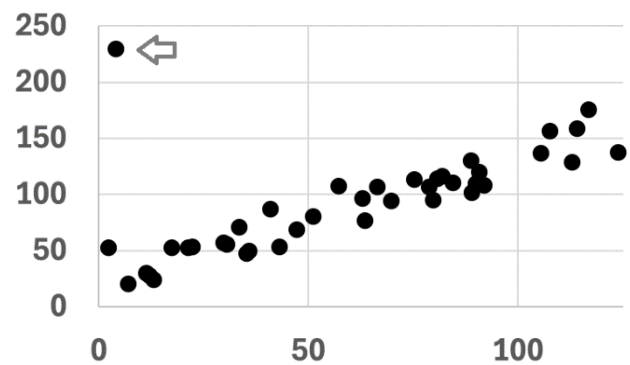
21. Suppose that Treadzilla Tires manufactures a tire with a lifetime that approximately follows a normal distribution with mean of 70,000 miles and standard deviation of 4400 miles. Suppose that Treadzilla Tires wants to warrant no more than 2% of its tires. What mileage should the company advertise as its warranty mileage?
A. 79,036
B. 60,963
C. 79,037
D. 60,964
E. Cannot be determined

22. A random sample of adult Americans were asked "Do you think televisions are a necessity or a luxury you could do without". A 95% confidence interval for the proportion of adult Americans who think televisions are a luxury is found to be $(0.428, 0.540)$. What is the point estimate of the population proportion?
A. 0.499
B. 0.477
C. 0.484
D. 0.540
E. Cannot be determined

23. Continue with the previous question. Construct a 95% confidence interval for the population proportion of Americans who think that the televisions are a necessity?
A. $(0.460, 0.572)$
B. $(0.480, 0.562)$
C. $(0.486, 0.522)$
D. $(0.410, 0.510)$
E. Cannot be determined

24. A television sports commentator wants to estimate the proportion of Americans who follow professional football. What sample size should be obtained if he wants to be within 3 percentage points with 95% confidence if he uses a 2010 estimate of 53% obtained from a Harris poll?
    A. 1068
    B. 1070
    C. 1064
    D. 1062

25. The scatter plot beside reveals that one data point lies very far from the rest of the data. What is the effect of including this data point when fitting a simple linear regression model?
    A. Including the point would decrease the intercept term and decrease the slope term
    B. Including the point would increase the intercept term and increase the slope term
    C. Including the point would increase the intercept term and decrease the slope term
    D. Including the point would decrease the intercept term and increase the slope term



This page left blank

Tie Breaker #1

Name: ________________

School: ________________

A history teacher believes that students in afternoon classes score higher than those in morning classes.
The teacher gives the two groups of students a randomly selected 100-point exam. The results are shown below.

Can the teacher conclude that students in afternoon classes generally have a higher score? Use $\alpha = 0.01$.

Morning Students	**Afternoon Students**
$n_1 = 36$	$n_2 = 41$
$\bar{x}_1 = 78$	$\bar{x}_2 = 81$
$s_1 = 5.8$	$s_2 = 6.3$

This page left blank

Tie Breaker #2

Name: ___________________

School: __________________

A medical researcher is interested in determining if there is a relationship between adults over 50 who exercise regularly and low, moderate, and high blood pressure. A random sample of 236 adults over 50 is selected and the results are given below.

Test the claim that regular exercise and low, moderate, and high blood pressure are independent. Use $\alpha = 0.01$.

Blood Pressure	Low	Moderate	High
Regular Exercise	35	62	25
No Regular Exercise	21	65	28

This page left blank

Tie Breaker #3

Name: ________________

School: ________________

According to a study done by Nick Wilson of Otago University of Wellington, the probability a randomly selected individual will not cover their mouth when sneezing is 0.267.

What is the probability that among 10 randomly observed individuals fewer than 6 cover their mouth?

This page left blank

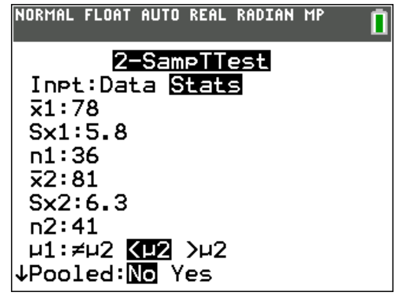
Multiple Choice Answers

1. C	11. B	21. B
2. C	12. B	22. C
3. A	13. A	23. A
4. B	14. C	24. C
5. C	15. C	25. C
6. C	16. C	
7. D	17. A	
8. B	18. A	
9. A	19. D	
10. C	20. B	

TB1:

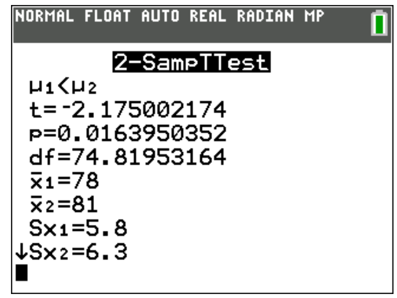
We will apply a two-sample t-test to this data. We assume that

- The samples are independently obtained, &
- The samples are independent, &
- We can confirm that both populations are normal ($n_2 > n1 > 30$).



Define our hypotheses:

Hypothesis:	Alternate Version:
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 < \mu_2$	$H_0: \mu_1 - \mu_2 = 0$ $H_1: \mu_1 - \mu_2 < 0$

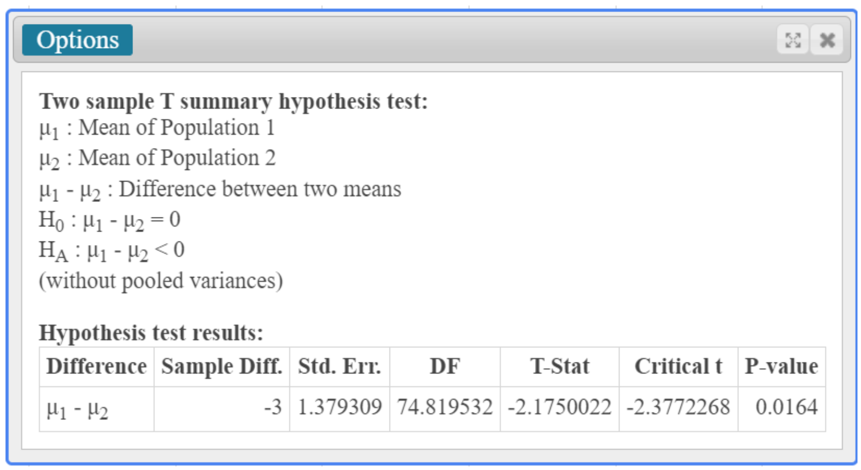


Run the test using technology.

We note that the p-value from this test is $p = 0.016$.
We compare to the given $\alpha = 0.01$. Because $p > \alpha$, we fail to reject the null hypothesis.

Alternatively, the classical approach could be used stating that t=-2.175. The critical value is -2.326. t is not in the rejection region, so we fail to reject the null hypothesis.

We conclude that this is insufficient evidence at the $\alpha = 0.01$ level to conclude that the afternoon class performs better than the morning class.


Options

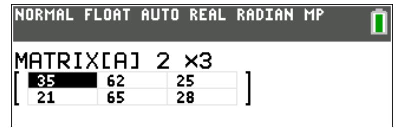
Two sample T summary hypothesis test:
μ_1 : Mean of Population 1
μ_2 : Mean of Population 2
$\mu_1 - \mu_2$: Difference between two means
$H_0 : \mu_1 - \mu_2 = 0$
$H_A : \mu_1 - \mu_2 < 0$
(without pooled variances)

Hypothesis test results:

Difference	Sample Diff.	Std. Err.	DF	T-Stat	Critical t	P-value
$\mu_1 - \mu_2$	-3	1.379309	74.819532	-2.1750022	-2.3772268	0.0164

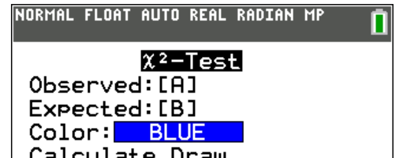
TB2:

We run a χ^2 (chi-square) test for independence on the provided data.
Requirement: We can confirm that all expected values are >5.

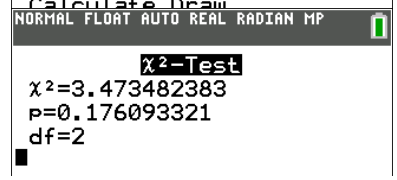


We want to know if blood pressure and exercise level are dependent or independent. We set up our hypotheses:

$$H_0: \text{blood pressure and exercise level are independent}$$
$$H_1: \text{blood pressure and exercise level are dependent}$$

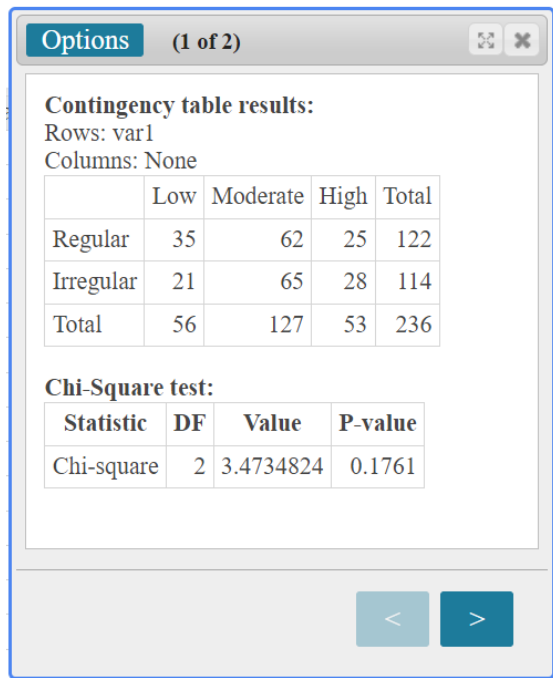


We use technology to calculate the test statistic.
(On a TI-84, we enter the data into matrix [A]. The Expected Matrix [B] is initially empty, it will be calculated automatically.)



The test statistic is p=0.176. We compare this to $\alpha = 0.01$, and note that $0.176 > 0.01$. Thus, we fail to reject the null hypothesis.

We conclude that there is insufficient evidence at the 0.01 level to say that blood pressure and exercise level are dependent.


Options (1 of 2)

Contingency table results:
Rows: var1
Columns: None

	Low	Moderate	High	Total
Regular	35	62	25	122
Irregular	21	65	28	114
Total	56	127	53	236

Chi-Square test:

Statistic	DF	Value	P-value
Chi-square	2	3.4734824	0.1761

TB3:

This problem gives us that $P(not\ sneezing) = 0.267$.

We calculate that $P(sneezing) = 1 - 0.267 = 0.733$.

Calculate a cumulative binomial distribution between 0 and 5 (inclusive) will cover their mouth. (Less than 6 means $0 \leq x \leq 5$).

Alternately we can sum the values between 0 and 5.

The probability is about 0.1, or about 10%.

Additional confirmation of binomial:

1. Fixed number of trials
2. Independent trials
3. Same probability of success for each trial
4. Two mutually outcomes possible