AP STATISTICS

UNIT 7

Inference for Quantitative Data: Means

10–18%
AP EXAM WEIGHTING

~14–16
CLASS PERIODS

UNIT 7 | 10–18% AP EXAM WEIGHTING | ~14–16 CLASS PERIODS

Inference for Quantitative Data: Means

BIG IDEA 1
Variation and Distribution VAR
- How do we know whether to use a t-test or a z-test for inference with means?

BIG IDEA 2
Patterns and Uncertainty UNC
- How can we make sure that samples are independent?

BIG IDEA 3
Data-Based Predictions, Decisions, and Conclusions DAT
- Why is it inappropriate to accept a hypothesis as true based on the results of statistical inference testing?

Developing Understanding

In this unit, students will analyze quantitative data to make inferences about population means. Students should understand that t^* and t-tests are used for inference with means when the population standard deviation, σ, is not known. Using s for σ in the formula for z gives a slightly different value, t, whose distribution, which depends on sample size, has more area in the tails than a normal distribution. The boundaries for rejecting a null hypothesis using a t-distribution tend to be further from the mean than for a normal distribution. Students should understand how and why conditions for inference with proportions and means are similar and different.

Building Course Skills

1.E 1.F 4.C 4.E

Unit 7 focuses on means, which has many similarities to the conditions and procedures for proportions. Since students sometimes confuse t-tests with z-tests, it will help to review the underlying rationales each time conditions come up. This will help students develop understanding through repeated practice in new situations. Teachers can encourage students to be mindful of notation and use the formula sheet as a reference.

Teachers can reinforce that inference testing requires careful selection of a procedure based on specific conditions for a given problem. Common errors include mislabeling conditions (e.g., incorrectly associating the large sample condition with independence), relying upon vague references to the normal distribution, or applying an inappropriate large sample condition. The null and alternative hypotheses must be clearly stated in terms of population parameters, not sample statistics. A formal decision compares the p-value to the level of significance. Students should also practice providing a numerical reference to support their claim (e.g., "Because $p < 0.05$, we reject the null hypothesis.") and interpreting findings within the context of the question.

Preparing for the AP Exam

It is critical for students to recognize that free-response questions asking whether data provide convincing evidence of some finding are asking for a significance test, not just a descriptive analysis. When using statistical inference for significance tests, students should identify the correct parameter and hypotheses, identify an appropriate test procedure and check conditions, calculate a test statistic and p-value, and provide a conclusion in context, along with a justification based on linkage between the p-value and the conclusion. For inference with means, the appropriate test will often be a t-test, but if σ is known, a z-test would be appropriate (see **2018 FRQ 6(a)**). For a t-test, conditions are (1) random sample and (2) large sample (e.g., $n > 30$). When sampling without replacement, students should also verify that the sample size is at most 10% of the population.

UNIT AT A GLANCE

Enduring Understanding	Topic	Skills	Class Periods ~14–16 CLASS PERIODS
VAR-1	**7.1 Introducing Statistics: Should I Worry About Error?**	1.A Identify the question to be answered or problem to be solved *(not assessed)*.	
VAR-7, UNC-4	**7.2 Constructing a Confidence Interval for a Population Mean**	3.C Describe probability distributions. 1.D Identify an appropriate inference method for confidence intervals. 4.C Verify that inference procedures apply in a given situation. 3.D Construct a confidence interval, provided conditions for inference are met.	
UNC-4	**7.3 Justifying a Claim About a Population Mean Based on a Confidence Interval**	4.B Interpret statistical calculations and findings to assign meaning or assess a claim. 4.D Justify a claim based on a confidence interval. 4.A Make an appropriate claim or draw an appropriate conclusion.	
VAR-7	**7.4 Setting Up a Test for a Population Mean**	1.E Identify an appropriate inference method for significance tests. 1.F Identify null and alternative hypotheses. 4.C Verify that inference procedures apply in a given situation.	
VAR-7, DAT-3	**7.5 Carrying Out a Test for a Population Mean**	3.E Calculate a test statistic and find a p-value, provided conditions for inference are met. 4.B Interpret statistical calculations and findings to assign meaning or assess a claim. 4.E Justify a claim using a decision based on significance tests.	

continued on next page

UNIT AT A GLANCE *(cont'd)*

Enduring Understanding	Topic	Skills	Class Periods ~14–16 CLASS PERIODS
UNC-4	**7.6 Confidence Intervals for the Difference of Two Means**	1.D Identify an appropriate inference method for confidence intervals. 4.C Verify that inference procedures apply in a given situation. 3.D Construct a confidence interval, provided conditions for inference are met.	
UNC-4	**7.7 Justifying a Claim About the Difference of Two Means Based on a Confidence Interval**	4.B Interpret statistical calculations and findings to assign meaning or assess a claim. 4.D Justify a claim based on a confidence interval. 4.A Make an appropriate claim or draw an appropriate conclusion.	
VAR-7	**7.8 Setting Up a Test for the Difference of Two Population Means**	1.E Identify an appropriate inference method for significance tests. 1.F Identify null and alternative hypotheses. 4.C Verify that inference procedures apply in a given situation.	
VAR-7, DAT-3	**7.9 Carrying Out a Test for the Difference of Two Population Means**	3.E Calculate a test statistic and find a p-value, provided conditions for inference are met. 4.B Interpret statistical calculations and findings to assign meaning or assess a claim. 4.E Justify a claim using a decision based on significance tests.	
	7.10 Skills Focus: Selecting, Implementing, and Communicating Inference Procedures	N/A	

Go to **AP Classroom** to assign the **Personal Progress Check** for Unit 7. Review the results in class to identify and address any student misunderstandings.

SAMPLE INSTRUCTIONAL ACTIVITIES

The sample activities on this page are optional and are offered to provide possible ways to incorporate various instructional approaches into the classroom. They were developed in partnership with teachers from the AP community to share ways that they approach teaching some of the topics in this unit. Please refer to the Instructional Approaches section beginning on p. 207 for more examples of activities and strategies.

Activity	Topic	Sample Activity
1	7.2	**Predict and Confirm** After introducing the confidence interval formula for a population mean when sigma is known, $\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$, have students discuss in small groups what will happen if we substitute the sample standard deviation s into the formula when σ is unknown. Have students use the Simulating Confidence Intervals for Population Parameter applet to test their conjectures (see link for Rossman/Chance Applets on page 210).
2	7.2 7.3 7.5	**Team Challenge** Give each team of three to four students a copy of 2004 FRQ 6, which focuses on the connection between a one-sample t-interval, a one-sample t-test, and the unfamiliar concept of a one-sided confidence interval. Challenge teams to collaboratively produce a model solution in 30 minutes.
3	7.4	**Discussion Groups** Ask each group of three to four students to identify the conditions for performing a test about a population mean. For each condition, have them explain why the condition is required and what would go wrong with the test if the condition were violated. Have groups pair up and compare answers.
4	7.9	**Team FRQ** Give each team of four students copies of a free-response question that involves performing a two-sample t-test (e.g., 2011 FRQ 4). Have each team member take responsibility for writing one part of the model solution (hypotheses, procedure and conditions, calculations, conclusion) with group input.
5	7.10	**Graphic Organizer** Have students work in teams of two to three to develop a flowchart for determining which inference procedure from Units 6 and 7 to use in a given setting.

TOPIC 7.1
Introducing Statistics: Why Should I Worry About Error?

Required Course Content

ENDURING UNDERSTANDING

VAR-1

Given that variation may be random or not, conclusions are uncertain.

LEARNING OBJECTIVE

VAR-1.I

Identify questions suggested by probabilities of errors in statistical inference. **[Skill 1.A]**

ESSENTIAL KNOWLEDGE

VAR-1.I.1

Random variation may result in errors in statistical inference.

SKILL

Selecting Statistical Methods

1.A

Identify the question to be answered or problem to be solved.

SKILLS

Using Probability and Simulation

3.C
Describe probability distributions.

3.D
Construct a confidence interval, provided conditions for inference are met.

Selecting Statistical Methods

1.D
Identify an appropriate inference method for confidence intervals.

Statistical Argumentation

4.C
Verify that inference procedures apply in a given situation.

TOPIC 7.2
Constructing a Confidence Interval for a Population Mean

Required Course Content

ENDURING UNDERSTANDING

VAR-7
The t-distribution may be used to model variation.

LEARNING OBJECTIVE

VAR-7.A
Describe t-distributions. **[Skill 3.C]**

ESSENTIAL KNOWLEDGE

VAR-7.A.1
When s is used instead of σ to calculate a test statistic, the corresponding distribution, known as the t-distribution, varies from the normal distribution in shape, in that more of the area is allocated to the tails of the density curve than in a normal distribution.

VAR-7.A.2
As the degrees of freedom increase, the area in the tails of a t-distribution decreases.

continued on next page

ENDURING UNDERSTANDING

UNC-4

An interval of values should be used to estimate parameters, in order to account for uncertainty.

LEARNING OBJECTIVE	ESSENTIAL KNOWLEDGE
UNC-4.O Identify an appropriate confidence interval procedure for a population mean, including the mean difference between values in matched pairs. **[Skill 1.D]**	**UNC-4.O.1** Because σ is typically not known for distributions of quantitative variables, the appropriate confidence interval procedure for estimating the population mean of one quantitative variable for one sample is a one-sample t-interval for a mean. **UNC-4.O.2** For one quantitative variable, X, that is normally distributed, the distribution of $t = \frac{(\bar{x} - \mu)}{\frac{s}{\sqrt{n}}}$ is a t-distribution with $n-1$ degrees of freedom. **UNC-4.O.3** Matched pairs can be thought of as one sample of pairs. Once differences between pairs of values are found, inference for confidence intervals proceeds as for a population mean.
UNC-4.P Verify the conditions for calculating confidence intervals for a population mean, including the mean difference between values in matched pairs. **[Skill 4.C]**	**UNC-4.P.1** In order to calculate confidence intervals to estimate a population mean, we must check for independence and that the sampling distribution is approximately normal: a. To check for independence: i. Data should be collected using a random sample or a randomized experiment. ii. When sampling without replacement, check that $n \leq 10\%N$, where N is the size of the population. b. To check that the sampling distribution of $\bar{x}$ is approximately normal (shape): i. If the observed distribution is skewed, n should be greater than 30. ii. If the sample size is less than 30, the distribution of the sample data should be free from strong skewness and outliers.

continued on next page

UNIT 7

Inference for Quantitative Data: Means

LEARNING OBJECTIVE	ESSENTIAL KNOWLEDGE
UNC-4.Q Determine the margin of error for a given sample size for a one-sample *t*-interval. **[Skill 3.D]**	**UNC-4.Q.1** The critical value t^* with $n-1$ degrees of freedom can be found using a table or computer-generated output. **UNC-4.Q.2** The standard error for a sample mean is given by $SE = \frac{s}{\sqrt{n}}$, where s is the sample standard deviation. **UNC-4.Q.3** For a one-sample *t*-interval for a mean, the margin of error is the critical value (t^*) times the standard error (*SE*), which equals $t^*\left(\frac{s}{\sqrt{n}}\right)$.
UNC-4.R Calculate an appropriate confidence interval for a population mean, including the mean difference between values in matched pairs. **[Skill 3.D]**	**UNC-4.R.1** The point estimate for a population mean is the sample mean, $\bar{x}$. **UNC-4.R.2** For the population mean for one sample with unknown population standard deviation, the confidence interval is $\bar{x} \pm t^* \frac{s}{\sqrt{n}}$. **CLARIFYING STATEMENT:** *Formulas for interval estimates do not appear explicitly on the AP Statistics Formula Sheet provided with the AP Statistics Exam. However, these formulas do not need to be memorized, as they can be constructed based on the general test statistic formula and the relevant standard error formulas that are provided on the formula sheet.*

TOPIC 7.3
Justifying a Claim About a Population Mean Based on a Confidence Interval

Required Course Content

ENDURING UNDERSTANDING

UNC-4

An interval of values should be used to estimate parameters, in order to account for uncertainty.

LEARNING OBJECTIVE	ESSENTIAL KNOWLEDGE
UNC-4.S Interpret a confidence interval for a population mean, including the mean difference between values in matched pairs. **[Skill 4.B]**	**UNC-4.S.1** A confidence interval for a population mean either contains the population mean or it does not, because each interval is based on data from a random sample, which varies from sample to sample. **UNC-4.S.2** We are C% confident that the confidence interval for a population mean captures the population mean. **UNC-4.S.3** An interpretation of a confidence interval for a population mean includes a reference to the sample taken and details about the population it represents.
UNC-4.T Justify a claim based on a confidence interval for a population mean, including the mean difference between values in matched pairs. **[Skill 4.D]**	**UNC-4.T.1** A confidence interval for a population mean provides an interval of values that may provide sufficient evidence to support a particular claim in context.

continued on next page

SKILLS

Statistical Argumentation

4.B

Interpret statistical calculations and findings to assign meaning or assess a claim.

4.D

Justify a claim based on a confidence interval.

4.A

Make an appropriate claim or draw an appropriate conclusion.

ILLUSTRATIVE EXAMPLE

UNC-4.S.3:

For interpreting a 96% confidence interval for mean foot length for all footprints found in a cave based on a particular randomly selected sample of footprints in the cave:

"We are 96% confident that the mean foot length for all footprints found in the cave falls within the confidence interval" (based on **2000 FRQ 2**).

Inference for Quantitative Data: Means

LEARNING OBJECTIVE

UNC-4.U

Identify the relationships between sample size, width of a confidence interval, confidence level, and margin of error for a population mean. **[Skill 4.A]**

ESSENTIAL KNOWLEDGE

UNC-4.U.1

When all other things remain the same, the width of a confidence interval for a population mean tends to decrease as the sample size increases.

UNC-4.U.2

For a single mean, the width of the interval is proportional to $\frac{1}{\sqrt{n}}$.

UNC-4.U.3

For a given sample, the width of the confidence interval for a population mean increases as the confidence level increases.

TOPIC 7.4
Setting Up a Test for a Population Mean

Required Course Content

ENDURING UNDERSTANDING

VAR-7

The *t*-distribution may be used to model variation.

LEARNING OBJECTIVE

VAR-7.B

Identify an appropriate testing method for a population mean with unknown σ, including the mean difference between values in matched pairs. **[Skill 1.E]**

VAR-7.C

Identify the null and alternative hypotheses for a population mean with unknown σ, including the mean difference between values in matched pairs. **[Skill 1.F]**

ESSENTIAL KNOWLEDGE

VAR-7.B.1

The appropriate test for a population mean with unknown σ is a one-sample *t*-test for a population mean.

VAR-7.B.2

Matched pairs can be thought of as one sample of pairs. Once differences between pairs of values are found, inference for significance testing proceeds as for a population mean.

VAR-7.C.1

The null hypothesis for a one-sample *t*-test for a population mean is $H_0 : \mu = \mu_0$, where μ_0 is the hypothesized value. Depending upon the situation, the alternative hypothesis is $H_a : \mu < \mu_0$, or $H_a : \mu > \mu_0$, or $H_a : \mu \neq \mu_0$.

VAR-7.C.2

When finding the mean difference, μ_d, between values in a matched pair, it is important to define the order of subtraction.

continued on next page

SKILLS

Selecting Statistical Methods

1.E

Identify an appropriate inference method for significance tests.

1.F

Identify null and alternative hypotheses.

Statistical Argumentation

4.C

Verify that inference procedures apply in a given situation.

AVAILABLE RESOURCE

- Classroom Resource > Inference

Inference for Quantitative Data: Means

LEARNING OBJECTIVE

VAR-7.D

Verify the conditions for the test for a population mean, including the mean difference between values in matched pairs. **[Skill 4.C]**

ESSENTIAL KNOWLEDGE

VAR-7.D.1

In order to make statistical inferences when testing a population mean, we must check for independence and that the sampling distribution is approximately normal:

a. To check for independence:
   i. Data should be collected using a random sample or a randomized experiment.
   ii. When sampling without replacement, check that $n \leq 10\%N$.

b. To check that the sampling distribution of $\bar{x}$ is approximately normal (shape):
   i. If the observed distribution is skewed, n should be greater than 30.
   ii. If the sample size is less than 30, the distribution of the sample data should be free from strong skewness and outliers.

TOPIC 7.5
Carrying Out a Test for a Population Mean

Required Course Content

ENDURING UNDERSTANDING

VAR-7

The *t*-distribution may be used to model variation.

LEARNING OBJECTIVE

VAR-7.E

Calculate an appropriate test statistic for a population mean, including the mean difference between values in matched pairs. **[Skill 3.E]**

ESSENTIAL KNOWLEDGE

VAR-7.E.1

For a single quantitative variable when random sampling with replacement from a population that can be modeled with a normal distribution with mean μ and standard deviation σ, the sampling distribution of $t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$ has a t-distribution with $n - 1$ degrees of freedom.

CLARIFYING STATEMENT:

The formulas for test statistics do not appear explicitly on the AP Statistics Formula Sheet provided with the AP Statistics Exam. However, these formulas do not need to be memorized, as they can be constructed based on the general test statistic formula and the relevant standard error formulas that are provided on the formula sheet.

continued on next page

SKILLS

Using Probability and Simulation

3.E

Calculate a test statistic and find a p-value, provided conditions for inference are met.

Statistical Argumentation

4.B

Interpret statistical calculations and findings to assign meaning or assess a claim.

4.E

Justify a claim using a decision based on significance tests.

AVAILABLE RESOURCE

- Classroom Resource > Inference

UNIT 7

Inference for Quantitative Data: Means

ENDURING UNDERSTANDING

DAT-3

Significance testing allows us to make decisions about hypotheses within a particular context.

LEARNING OBJECTIVE

DAT-3.E

Interpret the p-value of a significance test for a population mean, including the mean difference between values in matched pairs. **[Skill 4.B]**

DAT-3.F

Justify a claim about the population based on the results of a significance test for a population mean. **[Skill 4.E]**

ESSENTIAL KNOWLEDGE

DAT-3.E.1

An interpretation of the p-value of a significance test for a population mean should recognize that the p-value is computed by assuming that the null hypothesis is true, i.e., by assuming that the true population mean is equal to the particular value stated in the null hypothesis.

DAT-3.F.1

A formal decision explicitly compares the p-value to the significance α. If the p-value $\leq \alpha$, then reject the null hypothesis, $H_0 : \mu = \mu_0$. If the p-value $> \alpha$, then fail to reject the null hypothesis.

DAT-3.F.2

The results of a significance test for a population mean can serve as the statistical reasoning to support the answer to a research question about the population that was sampled.

TOPIC 7.6
Confidence Intervals for the Difference of Two Means

Required Course Content

ENDURING UNDERSTANDING

UNC-4

An interval of values should be used to estimate parameters, in order to account for uncertainty.

LEARNING OBJECTIVE

UNC-4.V

Identify an appropriate confidence interval procedure for a difference of two population means. **[Skill 1.D]**

ESSENTIAL KNOWLEDGE

UNC-4.V.1

Consider a simple random sample from population 1 of size n_1, mean μ_1, and standard deviation σ_1 and a second simple random sample from population 2 of size n_2, mean μ_2, and standard deviation σ_2. If the distributions of populations 1 and 2 are normal or if both n_1 and n_2 are greater than 30, then the sampling distribution of the difference of means, $\bar{x}_1 - \bar{x}_2$ is also normal. The mean for the sampling distribution of $\bar{x}_1 - \bar{x}_2$ is $\mu_1 - \mu_2$. The standard deviation of $\bar{x}_1 - \bar{x}_2$ + is $\sqrt{\frac{(\sigma_1)^2}{n_1} + \frac{(\sigma_2)^2}{n_2}}$.

UNC-4.V.2

The appropriate confidence interval procedure for one quantitative variable for two independent samples is a two-sample t-interval for a difference between population means.

continued on next page

SKILLS

Selecting Statistical Methods

1.D

Identify an appropriate inference method for confidence intervals.

Statistical Argumentation

4.C

Verify that inference procedures apply in a given situation.

Using Probability and Simulation

3.D

Construct a confidence interval, provided conditions for inference are met.

LEARNING OBJECTIVE	ESSENTIAL KNOWLEDGE
UNC-4.W Verify the conditions to calculate confidence intervals for the difference of two population means. **[Skill 4.C]**	**UNC-4.W.1** In order to calculate confidence intervals to estimate a difference of population means, we must check for independence and that the sampling distribution is approximately normal: a. To check for independence: i. Data should be collected using two independent, random samples or a randomized experiment. ii. When sampling without replacement, check that $n_1 \leq 10\%N_1$ and $n_2 \leq 10\%N_2$. b. To check that the sampling distribution of $(\bar{x}_1 - \bar{x}_2)$ should be approximately normal (shape): i. If the observed distributions are skewed, both n_1 and n_2 should be greater than 30.
UNC-4.X Determine the margin of error for the difference of two population means. **[Skill 3.D]**	**UNC-4.X.1** For the difference of two sample means, the margin of error is the critical value (t^*) times the standard error (*SE*) of the difference of two means. **UNC-4.X.2** The standard error for the difference in two sample means with sample standard deviations, s_1 and s_2, is $\sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}}$.
UNC-4.Y Calculate an appropriate confidence interval for a difference of two population means. **[Skill 3.D]**	**UNC-4.Y.1** The point estimate for the difference of two population means is the difference in sample means, $\bar{x}_1 - \bar{x}_2$.

continued on next page

LEARNING OBJECTIVE

UNC-4.Y

Calculate an appropriate confidence interval for a difference of two population means. **[Skill 3.D]**

ESSENTIAL KNOWLEDGE

UNC-4.Y.2

For a difference of two population means where the population standard deviations are not known, the confidence interval is $\left(\bar{x}_1 - \bar{x}_2\right) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ where $\pm t^*$ are the critical values for the central C% of a t-distribution with appropriate degrees of freedom that can be found using technology.

CLARIFYING STATEMENT:

Formulas for interval estimates do not appear explicitly on the AP Statistics Formula Sheet provided with the AP Statistics Exam. However, these formulas do not need to be memorized, as they can be constructed based on the general test statistic formula and the relevant standard error formulas that are provided on the formula sheet.

SKILLS

Statistical Argumentation

4.B
Interpret statistical calculations and findings to assign meaning or assess a claim.

4.D
Justify a claim based on a confidence interval.

4.A
Make an appropriate claim or draw an appropriate conclusion.

ILLUSTRATIVE EXAMPLE

UNC-4.Z.2:

For interpreting a confidence interval for a difference between mean response times for two fire stations (northern – southern): "Based on these samples, one can be 95 percent confident that the difference in the population mean response times (northern – southern) is between −2.37 minutes and 0.37 minutes" (2009 FRQ 4).

TOPIC 7.7
Justifying a Claim About the Difference of Two Means Based on a Confidence Interval

Required Course Content

ENDURING UNDERSTANDING

UNC-4
An interval of values should be used to estimate parameters, in order to account for uncertainty.

LEARNING OBJECTIVE	ESSENTIAL KNOWLEDGE
UNC-4.Z Interpret a confidence interval for a difference of population means. **[Skill 4.B]**	**UNC-4.Z.1** In repeated random sampling with the same sample size, approximately C% of confidence intervals created will capture the difference of population means. **UNC-4.Z.2** An interpretation for a confidence interval for the difference of two population means should include a reference to the samples taken and details about the populations they represent.
UNC-4.AA Justify a claim based on a confidence interval for a difference of population means. **[Skill 4.D]**	**UNC-4.AA.1** A confidence interval for a difference of population means provides an interval of values that may provide sufficient evidence to support a particular claim in context.
UNC-4.AB Identify the effects of sample size on the width of a confidence interval for the difference of two means. **[Skill 4.A]**	**UNC-4.AB.1** When all other things remain the same, the width of the confidence interval for the difference of two means tends to decrease as the sample sizes increase.

TOPIC 7.8
Setting Up a Test for the Difference of Two Population Means

Required Course Content

ENDURING UNDERSTANDING

VAR-7

The *t*-distribution may be used to model variation.

LEARNING OBJECTIVE

VAR-7.F

Identify an appropriate selection of a testing method for a difference of two population means. **[Skill 1.E]**

VAR-7.G

Identify the null and alternative hypotheses for a difference of two population means. **[Skill 1.F]**

ESSENTIAL KNOWLEDGE

VAR-7.F.1

For a quantitative variable, the appropriate test for a difference of two population means is a two-sample *t*-test for a difference of two population means.

VAR-7.G.1

The null hypothesis for a two-sample *t*-test for a difference of two population means, μ_1 and μ_2, is: $H_0 : \mu_1 - \mu_2 = 0$, or $H_0 : \mu_1 = \mu_2$. The alternative hypothesis is $H_a : \mu_1 - \mu_2 < 0$, or $H_a : \mu_1 - \mu_2 > 0$, or $H_a : \mu_1 - \mu_2 \neq 0$, or $H_a : \mu_1 > \mu_2$, or $H_a : \mu_1 < \mu_2$, or $H_a : \mu_1 \neq \mu_2$.

continued on next page

SKILLS

Selecting Statistical Methods

1.E

Identify an appropriate inference method for significance tests.

1.F

Identify null and alternative hypotheses.

Statistical Argumentation

4.C

Verify that inference procedures apply in a given situation.

Inference for Quantitative Data: Means

LEARNING OBJECTIVE

VAR-7.H

Verify the conditions for the significance test for the difference of two population means. **[Skill 4.C]**

ESSENTIAL KNOWLEDGE

VAR-7.H.1

In order to make statistical inferences when testing a difference between population means, we must check for independence and that the sampling distribution is approximately normal:

a. Individual observations should be independent:
   i. Data should be collected using simple random samples or a randomized experiment.
   ii. When sampling without replacement, check that $n_1 \leq 10\%N_1$ and $n_2 \leq 10\%N_2$.

b. The sampling distribution of $\bar{x}_1 - \bar{x}_2$ should be approximately normal (shape).
   i. If the observed distribution is skewed, both n_1 and n_2 should be greater than 30.
   ii. If the sample size is less than 30, the distribution of the sample data should be free from strong skewness and outliers. This should be checked for BOTH samples.

TOPIC 7.9
Carrying Out a Test for the Difference of Two Population Means

Required Course Content

ENDURING UNDERSTANDING

VAR-7

The *t*-distribution may be used to model variation.

LEARNING OBJECTIVE

VAR-7.I

Calculate an appropriate test statistic for a difference of two means. **[Skill 3.E]**

ESSENTIAL KNOWLEDGE

VAR-7.I.1

For a single quantitative variable, data collected using independent random samples or a randomized experiment from two populations, each of which can be modeled with a normal distribution, the sampling distribution of $t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ is an approximate *t*-distribution with degrees of freedom that can be found using technology. The degrees of freedom fall between the smaller of $n_1 - 1$ and $n_2 - 1$ and $n_1 + n_2 - 2$.

CLARIFYING STATEMENT:

The formulas for test statistics do not appear explicitly on the AP Statistics Formula Sheet provided with the AP Statistics Exam. However, these formulas do not need to be memorized, as they can be constructed based on the general test statistic formula and the standard error formulas for each of the relevant test statistics that are provided on the formula sheet.

continued on next page

SKILLS

Using Probability and Simulation

3.E

Calculate a test statistic and find a *p*-value, provided conditions for inference are met.

Statistical Argumentation

4.B

Interpret statistical calculations and findings to assign meaning or assess a claim.

4.E

Justify a claim using a decision based on significance tests.

AVAILABLE RESOURCE

- Classroom Resource > Inference

ILLUSTRATIVE EXAMPLE

VAR-7.I.1:

In a study comparing mean recovery times for two surgical procedures to repair a torn anterior cruciate ligament (ACL), the group receiving one procedure had a sample size of 110, while the group receiving the other procedure had a sample size of 100. The degrees of freedom fall between 100 (the smaller of 110 and 100) and 208 (110 + 100 − 2). The degrees of freedom may be determined using technology. If the test statistic for this study is $t \approx 7.13$, then the *p*-value is the area greater than 7.13 for a *t*-distribution with $df = 207.18$ (2018 FRQ 4).

UNIT 7

Inference for Quantitative Data: Means

ENDURING UNDERSTANDING

DAT-3

Significance testing allows us to make decisions about hypotheses within a particular context.

LEARNING OBJECTIVE

DAT-3.G

Interpret the p-value of a significance test for a difference of population means. **[Skill 4.B]**

DAT-3.H

Justify a claim about the population based on the results of a significance test for a difference of two population means in context. **[Skill 4.E]**

ESSENTIAL KNOWLEDGE

DAT-3.G.1

An interpretation of the p-value of a significance test for a two-sample difference of population means should recognize that the p-value is computed by assuming that the null hypothesis is true, i.e., by assuming that the true population means are equal to each other.

DAT-3.H.1

A formal decision explicitly compares the p-value to the significance α. If the p-value $\leq \alpha$, then reject the null hypothesis, $H_0 : \mu_1 - \mu_2 = 0$, or $H_0 : \mu_1 = \mu_2$. If the p-value $> \alpha$, then fail to reject the null hypothesis.

DAT-3.H.2

The results of a significance test for a two-sample test for a difference between two population means can serve as the statistical reasoning to support the answer to a research question about the populations that were sampled.

TOPIC 7.10

Skills Focus: Selecting, Implementing, and Communicating Inference Procedures

Required Course Content

This topic is intended to focus on the skill of selecting an appropriate inference procedure, now that students have a range of options. Students should be given opportunities to practice when and how to apply all learning objectives relating to inference involving proportions or means.