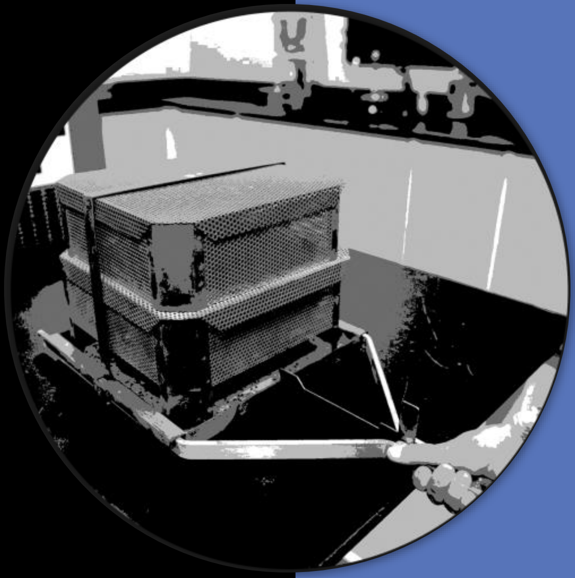
HMEC

Highway Materials Engineering Course

PARTICIPANT WORKBOOK



Quality Assurance

MODULE

A

U.S.Department of Transportation
Federal Highway Administration

Table of Contents

About This Workbook

This workbook has been developed as a resource for participants. This workbook can be used during the training session to follow along with the instructor and take notes, as well as for reference after the module has ended.

Course Overview

The Federal Highway Administration (FHWA) Highway Materials Engineering Course (HMEC) is a comprehensive multi-week training event that consists of eight content “modules” that provide students with the knowledge to develop materials specifications and guidance, make effective acceptance decisions, and design, construct, and maintain assets with a long service life. Modules range in duration for the number of days they take to complete. The modules are:

- Module A: Quality Assurance
- Module B: Soils and Foundations
- Module C: Steel, Welding, and Coatings
- Module D: Aggregates for Transportation Construction Projects
- Module E: Mechanistic Empirical Pavement Design Guide
- Module F: Asphalt Materials and Paving Mixtures
- Module G: Portland Cement Concrete
- Module H: Evaluating Recycled Materials for Beneficial Uses in Transportation

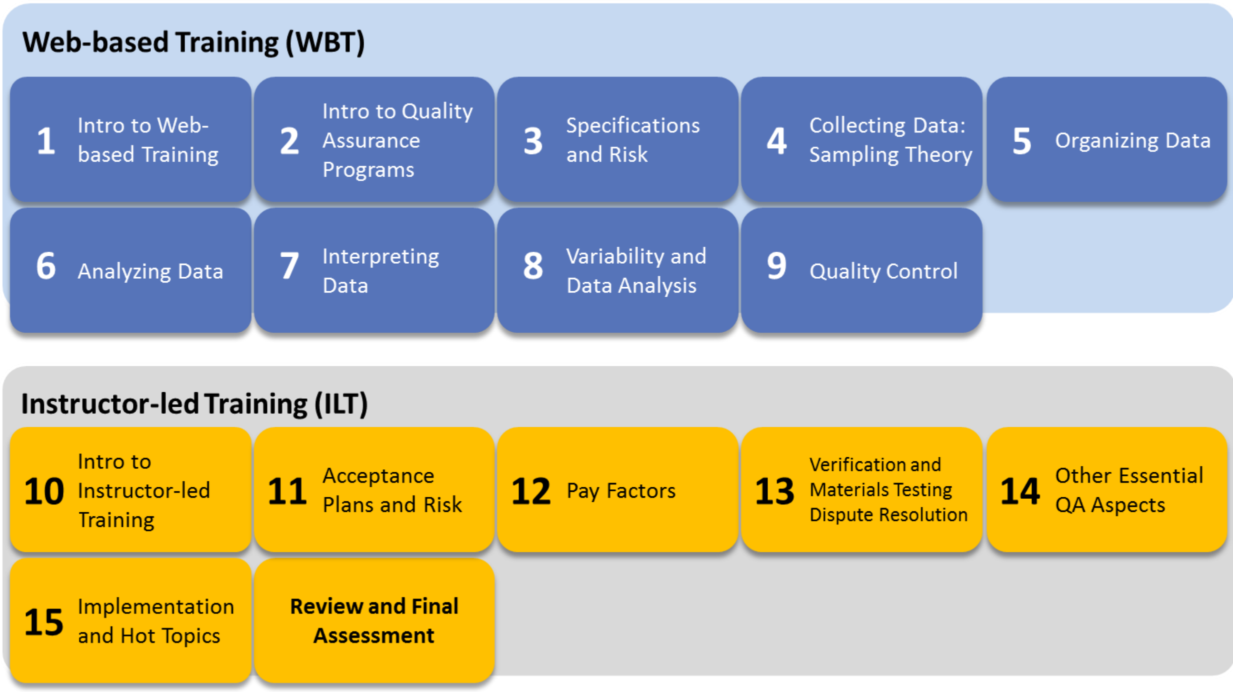
Introduction

This ILT module, Quality Assurance, is the first module in the FHWA HMEC. Quality Assurance is strategically placed as the first module in this course because all other modules will apply concepts from Quality Assurance. Establishing a relationship between quality assurance (QA) and each of the other modules is essential to the success of the HMEC.

Module A provides participants with an understanding of the basic elements of a statistically-based QA program and includes an introduction to quality assurance as well as techniques for collecting, organizing, and analyzing data. Participants will also have the opportunity to apply data using scenario-based learning activities and exercises. Participants will assess the strengths, weaknesses, and risks of process control and acceptance plans. Finally, participants will learn the steps for successful implementation of quality specifications.

Module A Overview

Below is a visual overview of all of the lessons covered in this module:



Module Goals

The goals for this module are as follows:

- Consistently apply fundamental quality assurance concepts, terminology, and definitions
- Organize data to support the analysis of the quality of materials
- Explain the relationship of sample data to the population
- Calculate population and sample means, standard deviation, and coefficient of variation
- Explain how sources of variability can affect the accuracy of data analysis
- Develop and use statistical process control charts to continually monitor for quality
- Develop and implement an effective acceptance plan
- Develop procedures for using contractor test results in the acceptance decision
- Apply the basic elements required for quality assurance specifications to a variety of materials
- Identify risks associated with a given acceptance plan

Learning Outcomes

Lesson 1: Introduction to Web-based Training (WBT)

Lesson 2: A Quality Assurance Program (WBT)

- LO 2.1: Define quality management and key terms related to quality assurance
- LO 2.2: Define the functions of a QA program
- LO 2.3: Explain the differences between quality assurance, quality control, and process control

Lesson 3: Specifications and Risk (WBT)

- LO 3.1: Compare and contrast fundamental differences among a variety of specification types
- LO 3.2: Explain why QA specifications are the preferred specification type for most products and processes
- LO 3.3: Explain that well-designed quality assurance specifications can balance the risk between the contractor and the agency

Lesson 4: Collecting Data: Sampling Theory (WBT)

- LO 4.1: Describe the basic phases of statistical analysis
- LO 4.2: Define sampling
- LO 4.3: Explain the importance of sampling and using all available data
- LO 4.4: Discuss security and documentation of random sample locations
- LO 4.5: Explain how a sample relates to a population
- LO 4.6: Apply random and stratified random sampling techniques to obtain valid data

Lesson 5: Organizing Data (WBT)

- LO 5.1: Using the guidelines provided, create a frequency table
- LO 5.2: Using a given data set, plot a frequency histogram

Lesson 6: Analyzing Data (WBT)

- LO 6.1: Calculate mean or average, population standard deviation, sample standard deviation, variance, and coefficient of variation

Lesson 7: Interpreting Data (WBT)

- LO 7.1: Define statistical interference, probability, and probability distributions
- LO 7.2: Identify various probability distributions
- LO 7.3: Explain the importance of the normal distribution
- LO 7.4: Relate the concept of probability to the calculation of risks

Lesson 8: Variability and Data Analysis (WBT)

- LO 8.1: List sources of variability in highway materials
- LO 8.2: Define precision, accuracy, and bias
- LO 8.3: Differentiate between information provided by split and independent samples, and state appropriate uses for each
- LO 8.4: Explain how precision and bias statements are measures of repeatability

Lesson 9: Quality Control (WBT)

- LO 9.1: Explain various ways in which a quality control plan can be developed and used to maintain the quality of a product or service
- LO 9.2: Describe a typical quality control plan
- LO 9.3: Relate variability to quality control
- LO 9.4: Describe different types of control charts and their applications
- LO 9.5: Establish and calculate control chart limits

Lesson 10: Introduction to Instructor-led Training (ILT)

Lesson 11: Acceptance Plans and Risk (ILT)

- LO 11.1: Describe the acceptance functions
- LO 11.2: List appropriate uses and the advantages and disadvantages of each type of acceptance plan
- LO 11.3: State the minimum criteria for an effective acceptance plan
- LO 11.4: Compare Percent Within Limits (PWL) with Average Absolute Deviation (AAD)
- LO 11.5: Evaluate data to establish Acceptable Quality Level (AQL) and Rejectable Quality Level (RQL) as well as an Operating Characteristic (OC) curve
- LO 11.6: Adjust an acceptance plan based on the buyer’s and seller’s risks (alpha/beta risks)
- LO 11.7: Explain how analysis of risks can address the development of an OC curve

Lesson 12: Pay Factors (ILT)

- LO 12.1: Explain the proper design of a pay equation
- LO 12.2: Evaluate a variety of composite pay factor criteria to select an appropriate combination

Lesson 13: Verification and Materials Testing Dispute Resolution (ILT)

- LO 13.1: Define "verification" and "validation" and the difference between the two
- LO 13.2: List characteristics of an appropriate dispute resolution process
- LO 13.3: Perform the *F*- and *t*-tests using equal variances
- LO 13.4: Perform *t*-test using unequal variances

Lesson 14: Other Essential QA Aspects (ILT)

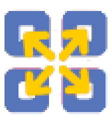
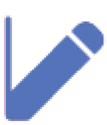
- LO 14.1: Explain the rationale for using Independent Assurance (IA) procedures
- LO 14.2: Compare and contrast project-basis and system-basis IA programs
- LO 14.3: Explain an effective laboratory qualification program
- LO 14.4: Explain an effective personnel qualification program
- LO 14.5: Explain the need for inspection

Lesson 15: Implementation and Hot Topics (ILT)

- LO 15.1: Identify potential political considerations of implementation
- LO 15.2: Generate a recommendation document based on best implementation practices from the United States
- LO 15.3: Develop a personal action plan for promoting a specific change in your organization
- LO 15.4: List the implementation steps
- LO 15.5: Identify emerging issues, technology, and trends that may affect the application of QA principles in the future

ILT Instruction Icons

The following icons are used in the lessons. These icons appear on the slides as a cue to the instructor and learners:

Icon	Icon Name	Typical Use
	Timer	▪ Call out the estimated time for the lesson
	Important Information	▪ Call out important information.
	Q & A	▪ Check for understanding or agreement. ▪ Survey participants. ▪ Solicit feedback.
	Breakout/Small Group Exercise	▪ Break participants into groups. ▪ Provide directions for exercise.
	Video/Sound	▪ Show a video.
	Reference	▪ Reference another document or resource.
	Links	▪ Share a Web link for additional resources.
	Whiteboard	▪ Draw or document something on a whiteboard or easel pad.

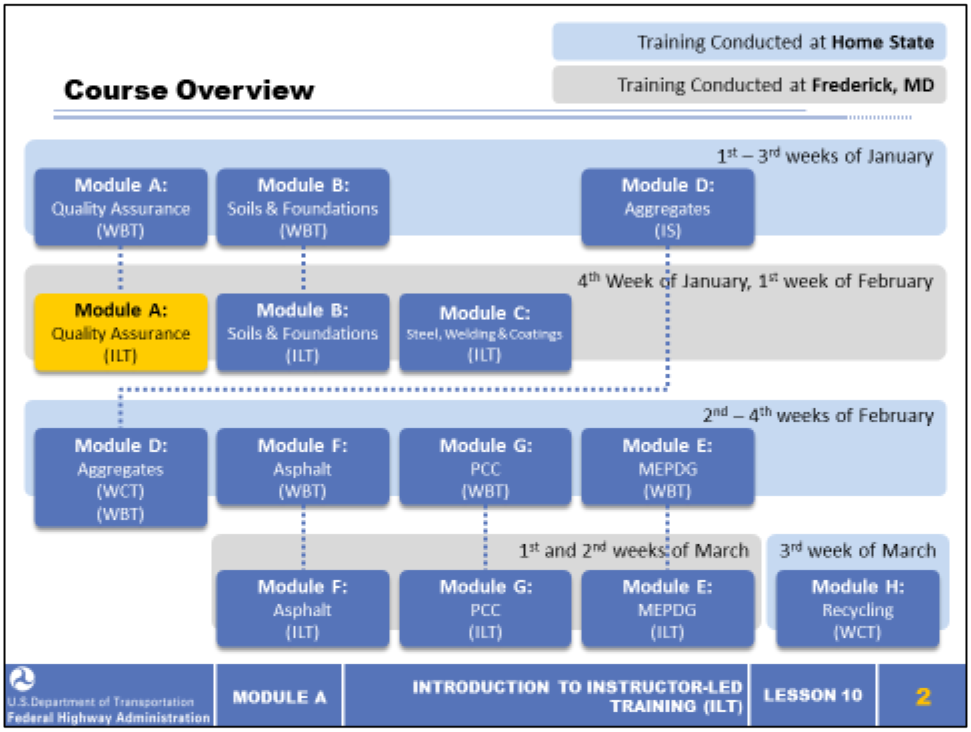
Icon	Icon Name	Typical Use
	Safety	▪ Call out important safety information.
	Common Error	▪ Call out a system or process that is often misused.

This page is intentionally blank.

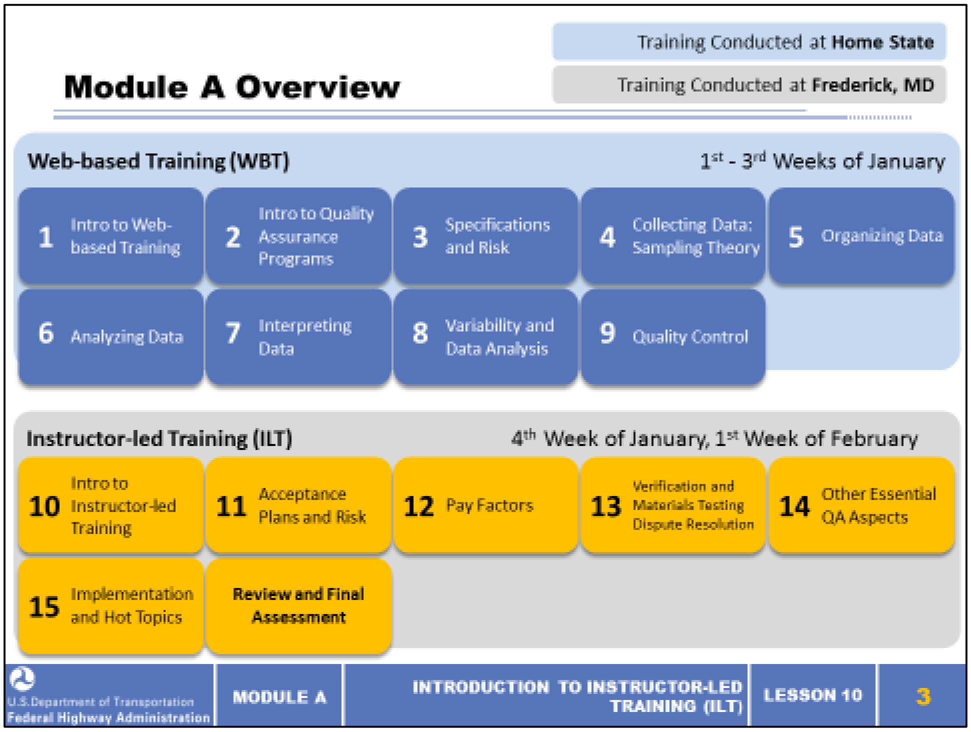
Slide 1


ILT
HMEC
Highway Materials Engineering Course
Lesson 10: Introduction to Instructor-led Training (ILT)
Quality Assurance
U.S.Department of Transportation
Federal Highway Administration
MODULE
A

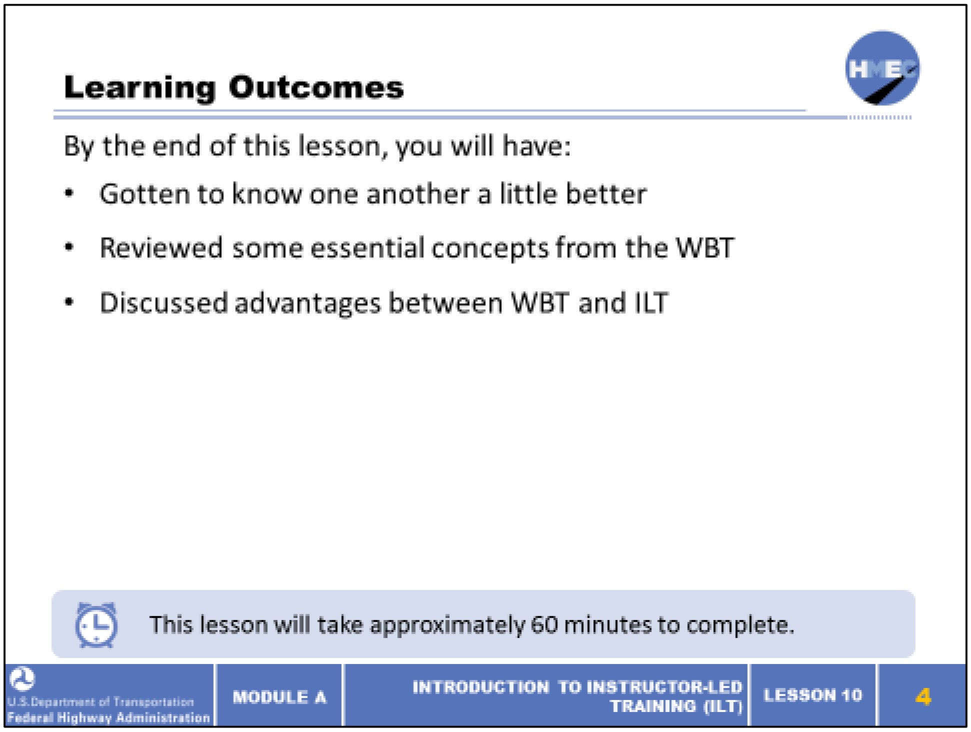
Slide 2


Training Conducted at Home State
Training Conducted at Frederick, MD
Course Overview
1st – 3rd weeks of January
Module A: Quality Assurance (WBT)
Module B: Soils & Foundations (WBT)
Module D: Aggregates (IS)
4th Week of January, 1st week of February
Module A: Quality Assurance (ILT)
Module B: Soils & Foundations (ILT)
Module C: Steel, Welding & Coatings (ILT)
2nd – 4th weeks of February
Module D: Aggregates (WCT) (WBT)
Module F: Asphalt (WBT)
Module G: PCC (WBT)
Module E: MEPDG (WBT)
1st and 2nd weeks of March
3rd week of March
Module F: Asphalt (ILT)
Module G: PCC (ILT)
Module E: MEPDG (ILT)
Module H: Recycling (WCT)
U.S. Department of Transportation
Federal Highway Administration
MODULE A
INTRODUCTION TO INSTRUCTOR-LED TRAINING (ILT)
LESSON 10
2

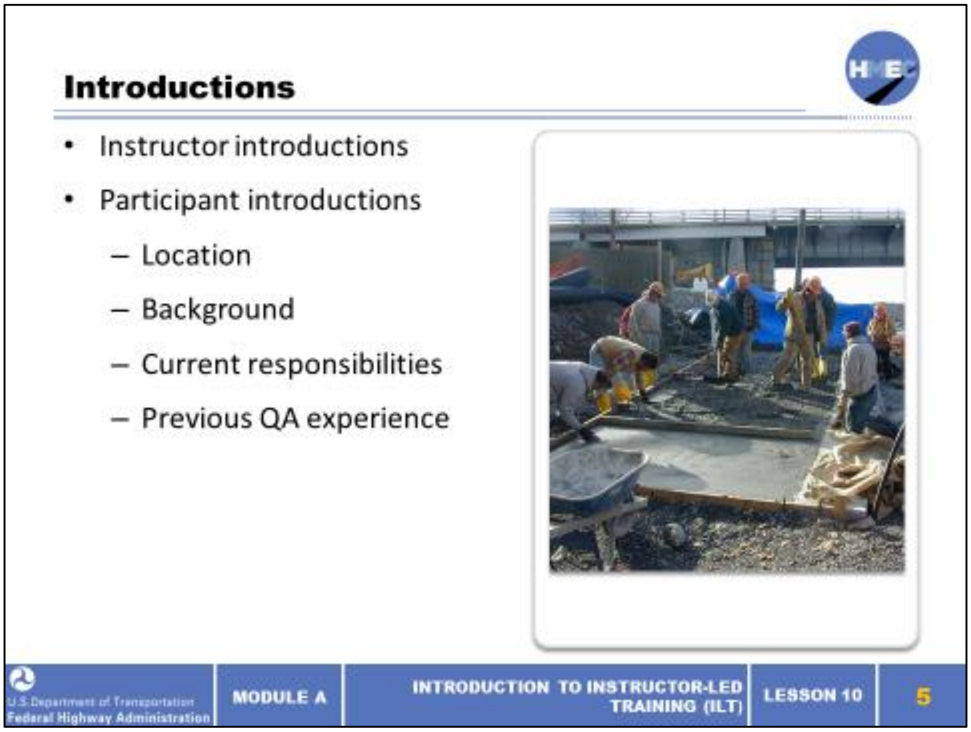
Slide 3


Training Conducted at Home State
Training Conducted at Frederick, MD
Module A Overview
Web-based Training (WBT)
1st - 3rd Weeks of January
1 Intro to Web-based Training
2 Intro to Quality Assurance Programs
3 Specifications and Risk
4 Collecting Data: Sampling Theory
5 Organizing Data
6 Analyzing Data
7 Interpreting Data
8 Variability and Data Analysis
9 Quality Control
Instructor-led Training (ILT)
4th Week of January, 1st Week of February
10 Intro to Instructor-led Training
11 Acceptance Plans and Risk
12 Pay Factors
13 Verification and Materials Testing Dispute Resolution
14 Other Essential QA Aspects
15 Implementation and Hot Topics
Review and Final Assessment
U.S. Department of Transportation Federal Highway Administration
MODULE A
INTRODUCTION TO INSTRUCTOR-LED TRAINING (ILT)
LESSON 10
3

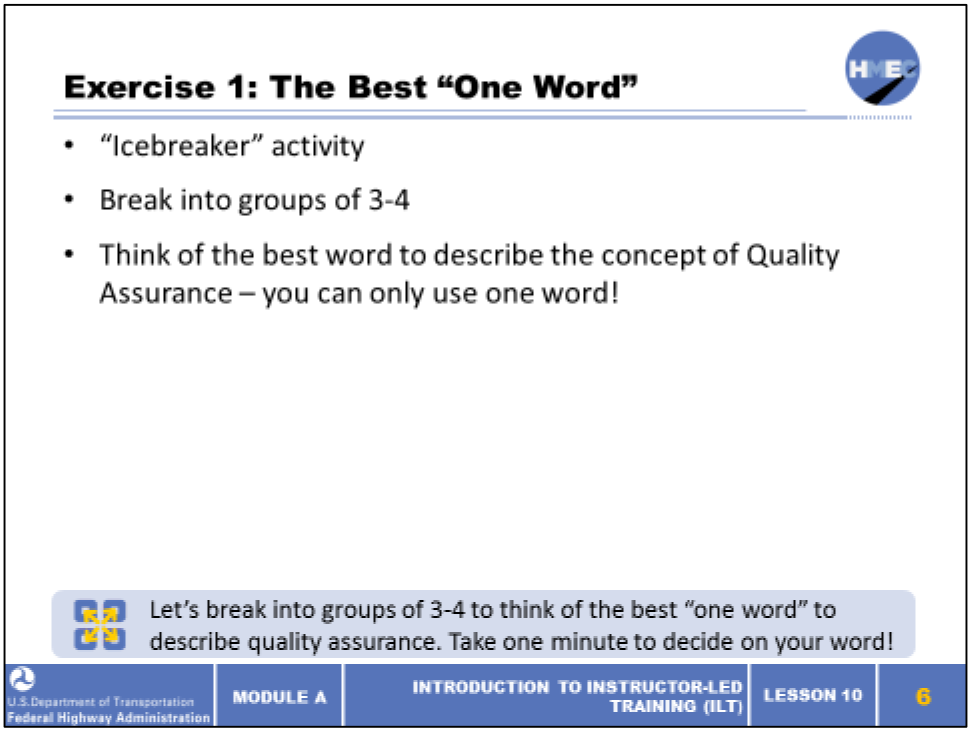
Slide 4


Learning Outcomes
HMEC
By the end of this lesson, you will have:
• Gotten to know one another a little better
• Reviewed some essential concepts from the WBT
• Discussed advantages between WBT and ILT
This lesson will take approximately 60 minutes to complete.
U.S. Department of Transportation
Federal Highway Administration
MODULE A
INTRODUCTION TO INSTRUCTOR-LED TRAINING (ILT)
LESSON 10
4

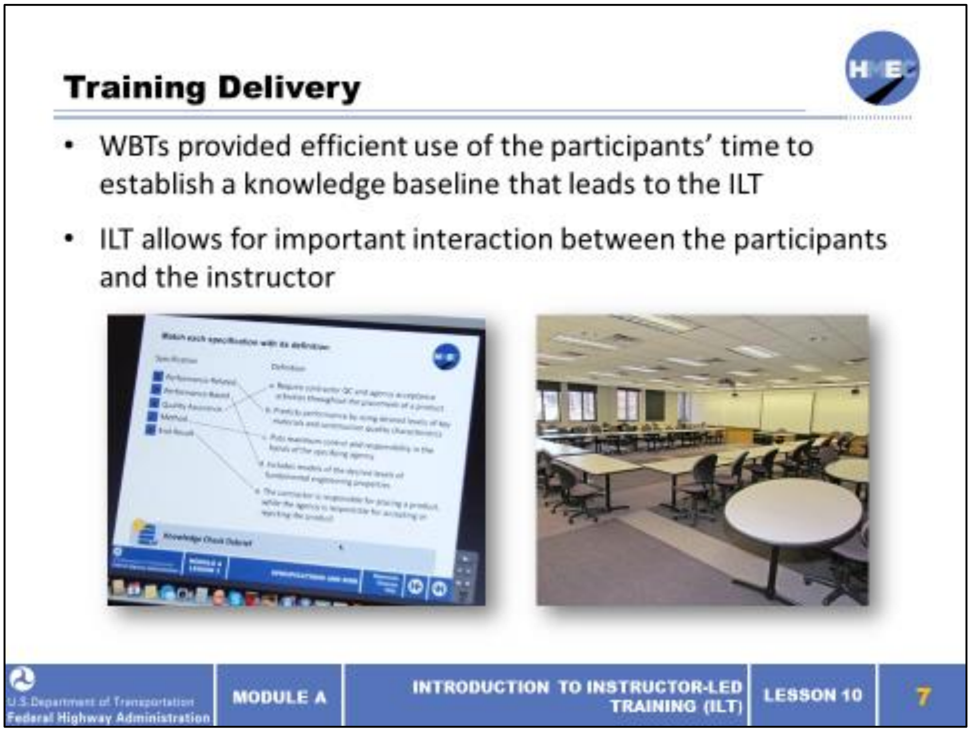
Slide 5


Introductions
Instructor introductions
Participant introductions
Location
Background
Current responsibilities
Previous QA experience
U.S. Department of Transportation
Federal Highway Administration
MODULE A
INTRODUCTION TO INSTRUCTOR-LED TRAINING (ILT)
LESSON 10
5

Slide 6


Exercise 1: The Best "One Word"
HMEC
• "Icebreaker" activity
• Break into groups of 3-4
• Think of the best word to describe the concept of Quality Assurance – you can only use one word!
Let's break into groups of 3-4 to think of the best "one word" to describe quality assurance. Take one minute to decide on your word!
U.S. Department of Transportation Federal Highway Administration
MODULE A
INTRODUCTION TO INSTRUCTOR-LED TRAINING (ILT)
LESSON 10
6

Slide 7



Advantages of WBT: The WBT is very efficient in allowing the participants to learn on their own and at their own pace. For some modules, the background for some participants are stronger in some areas than other participants. This allows those with a strong background to move along quickly while those with less experience gather the necessary knowledge at their own pace. This is particularly true for module A: Quality Assurance, thus starting with the WBT makes efficient use of the participants' time to establish a knowledge baseline that leads to the ILT.

Advantages of ILT: The ILT allows for important interaction between the participants and the instructor. It is difficult to cover all aspects of a subject in a WBT, whereas discussions on the subject allow the participants to ask questions or seek more information than may be presented on a slide. It also provides the instructor the opportunity to add anecdotal comments about subjects that may spark more interest than what can be put on a slide.

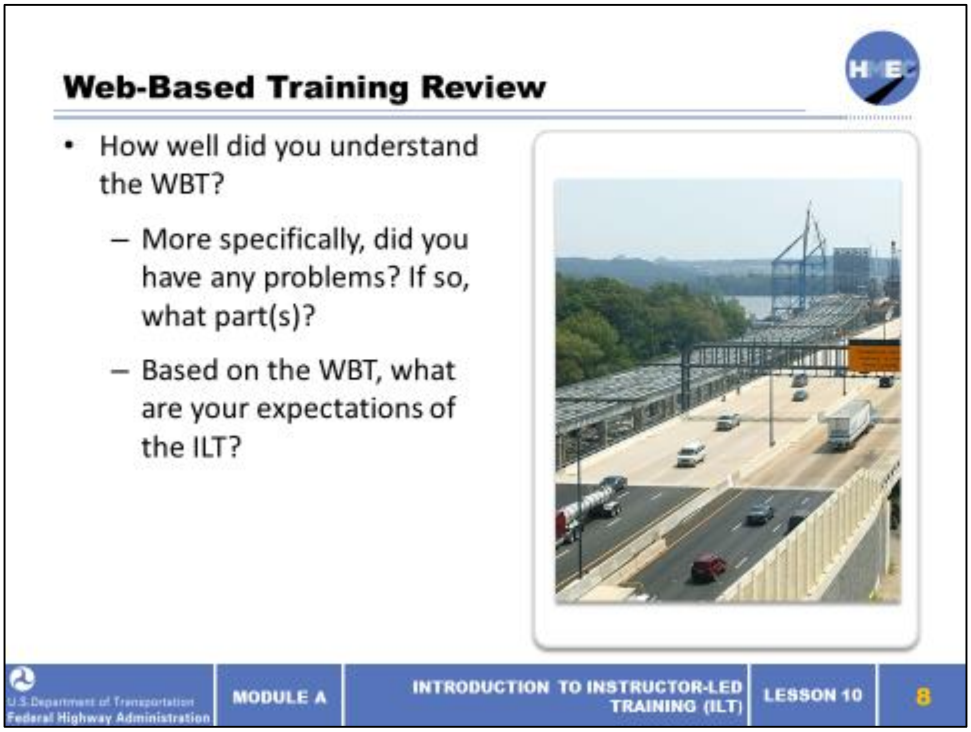
____

____

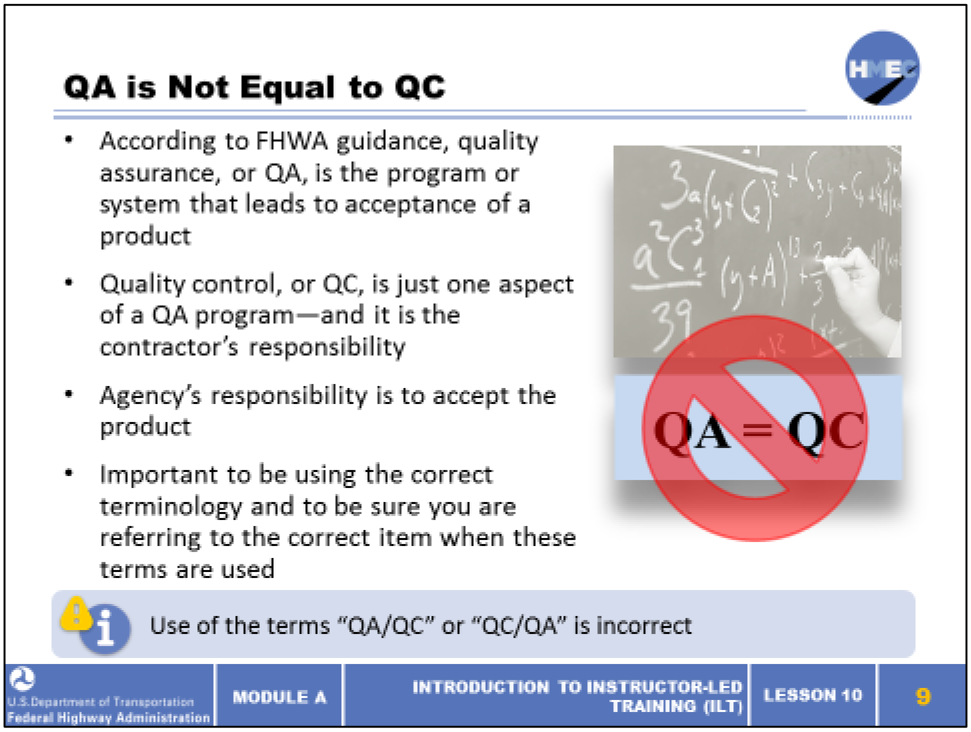
____

____

Slide 8

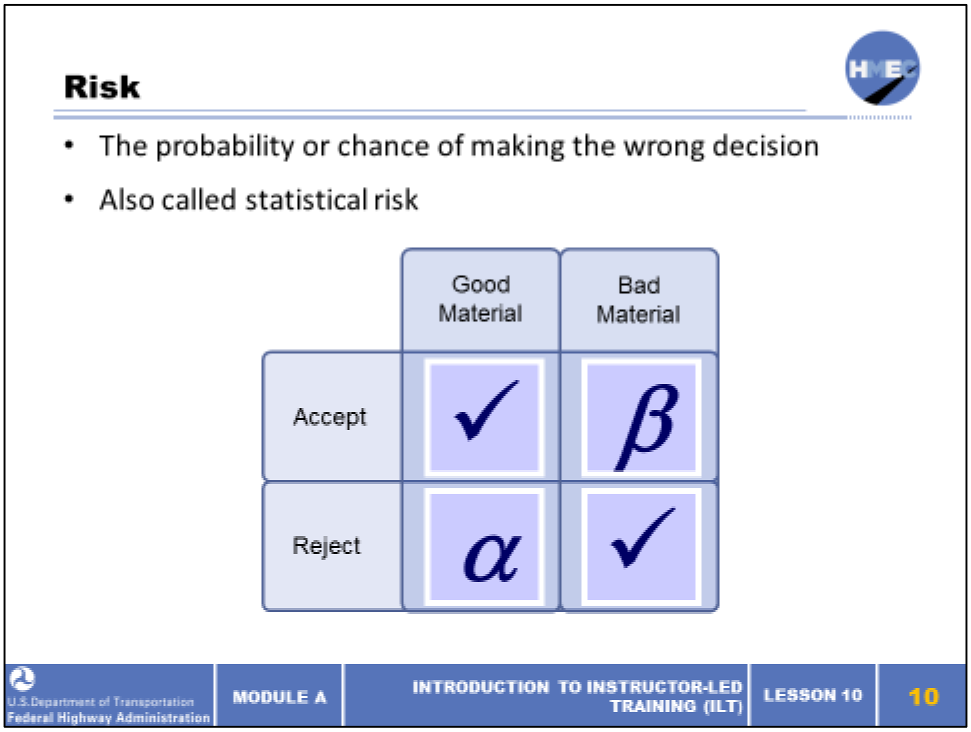

Web-Based Training Review
• How well did you understand the WBT?
– More specifically, did you have any problems? If so, what part(s)?
– Based on the WBT, what are your expectations of the ILT?
U.S. Department of Transportation
Federal Highway Administration
MODULE A
INTRODUCTION TO INSTRUCTOR-LED TRAINING (ILT)
LESSON 10
8

Slide 9



It is important to know that quality assurance (QA) is not equal to quality control (QC).

Slide 10



In all specification types, a decision must be made regarding compliance with the specifications, which involves risks. There are two types of risk. The seller's (contractor) risk, α, is the risk of rejecting "good" material. The buyer's (agency) risk, β, is the risk of accepting "bad" material. One of the greatest advantages of a QA specification is that, when properly developed, the risks can be quantified.

____

____

____

____

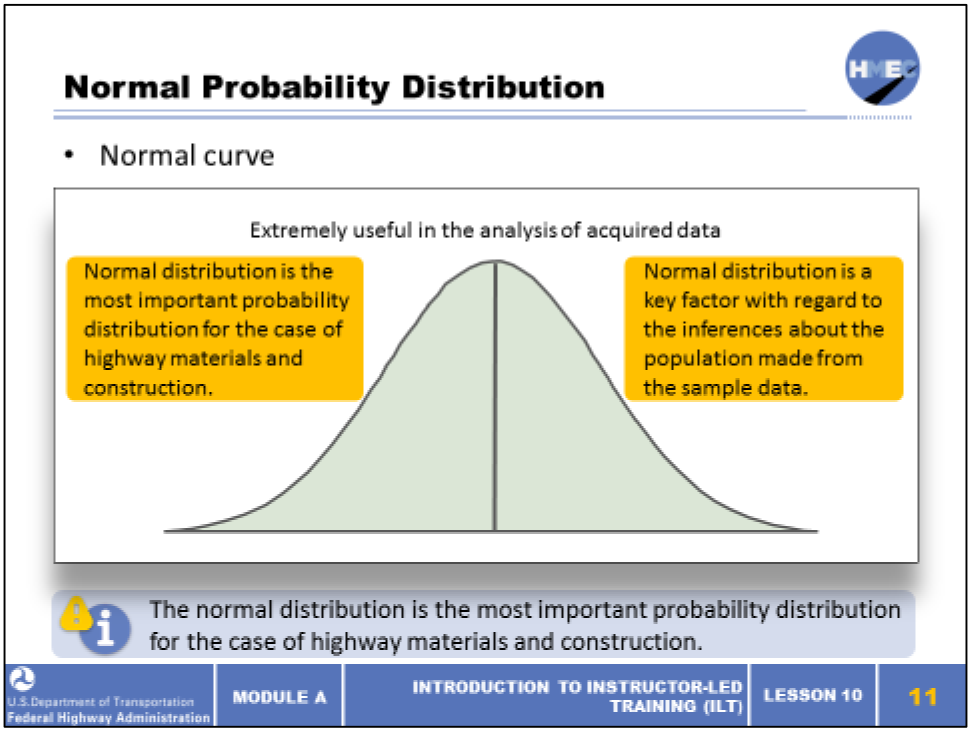
____

____

____

____

Slide 11



From extensive research, it has been concluded that numerous measurements that occur in highway construction, and in nature in general, distribute themselves about some average value with the majority of the measurements grouped near the mean and with progressively fewer results recorded as one proceeds away from the mean. The normal distribution, therefore, is the most important probability distribution for the case of highway materials and construction. Besides being extremely useful in the analysis of acquired data, the normal distribution is a key factor with regard to the inferences about the population that are made from the sample data.

____

____

____

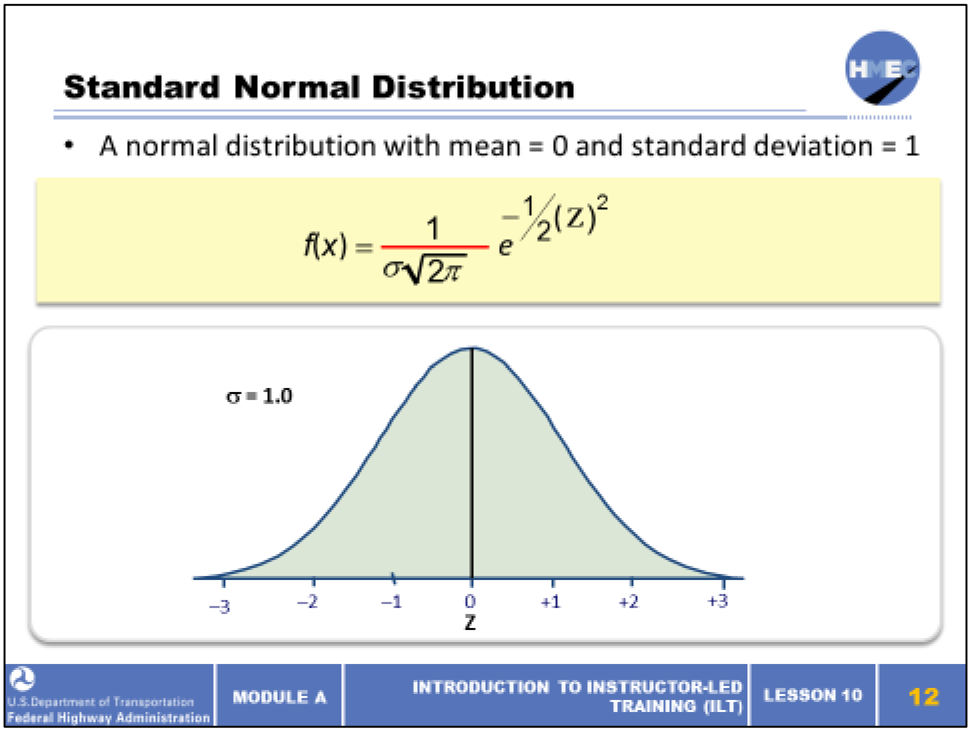
____

____

____

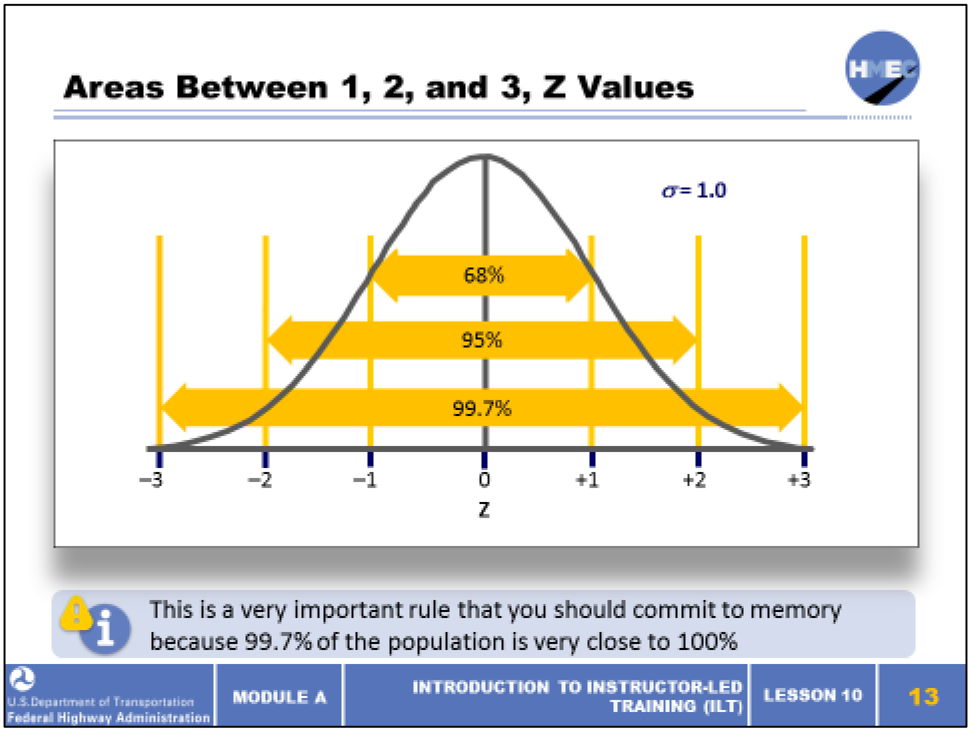
____

Slide 12



This may seem odd. Do you know of any highway materials that have a mean of zero and a standard deviation of one? There is likely not one, so what is the usefulness of a distribution such as this? The horizontal axis for the standard normal distribution is usually designated as the Z–axis. The values on the Z–axis are equal to the number of standard deviations above (positive Z values), or below (negative Z values) the mean of 0. The equation for the Z statistic is Z = X minus mu divided by sigma, so this simplifies the function that describes the distribution.

Slide 13



All normal distributions have the following areas.

1. The percentage of the area under the standard normal distribution is between Z = – 1.0 and Z = + 1.0 = 68.26%, rounded to 68%?
2. The percentage of the area under the standard normal distribution is between Z = –2.0 and Z = +2.0 = 95.44%, rounded to 95%?
3. The percentage of the area under the standard normal distribution is between Z = –3.0 and Z = +3.0 = 99.74%, rounded to 99.7%?

Remember because of the Z equation, Z is related to the number of standard deviations under a normal curve. This rule is a very important rule that you should commit to memory. Also note that because 99.7% of the population is very close to 100%, we often say three Z (or standard deviation) limits constitute "all" of the population.

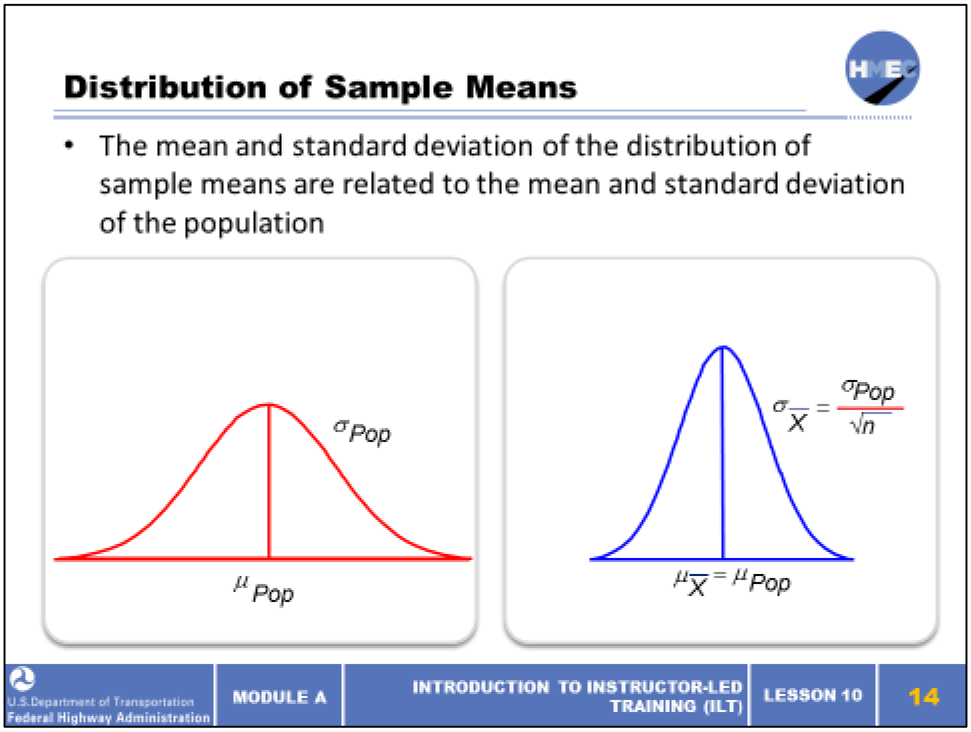
____

____

____

____

Slide 14



The distribution of means follows a normal distribution, but with a different shape (width) than one based on a sample size of n=1. It has been proven mathematically that the mean and standard deviation of the distribution of sample means are related to the mean and standard deviation of the population. The mean of the distribution of samples means, mu sub X-bar, will be equal to the mean of the population, mu sub Pop. We got an indication of that from the previous example of the chip sampling experiment. While it was not as obvious from the previous example, the standard deviation of the sample means sigma sub X-bar will be equal to the population standard deviation divided by the square root of the sample size, this is sigma sub Pop divided by the square root of 4. It is important to understand that the sample size used here is the number of values used to calculate the sample mean, and not the number of sample means.

____

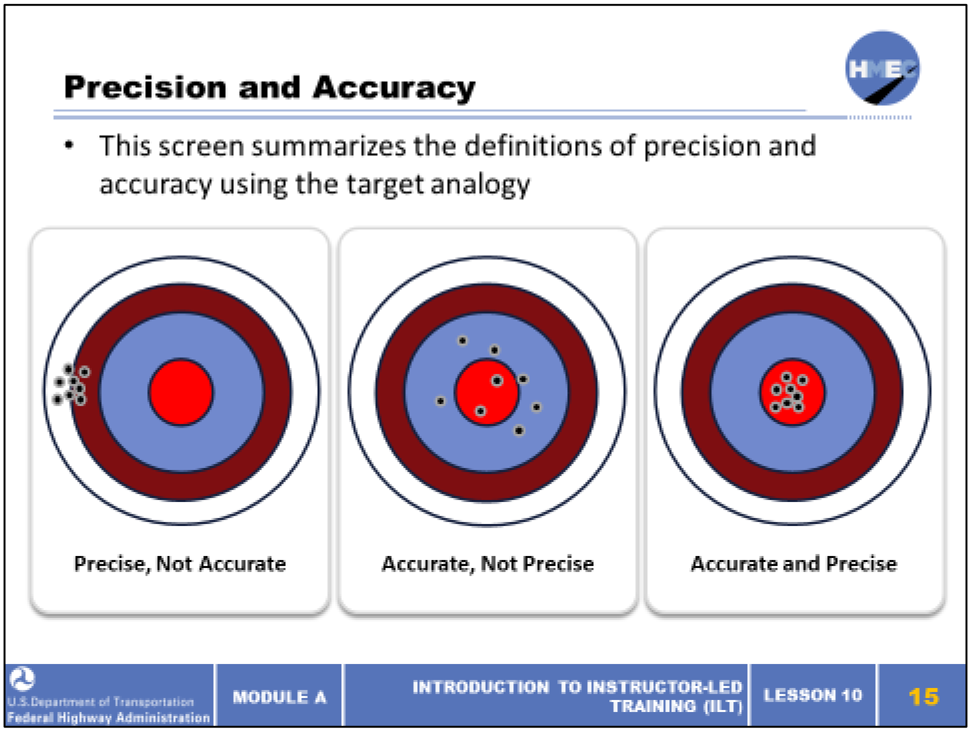
____

____

____

____

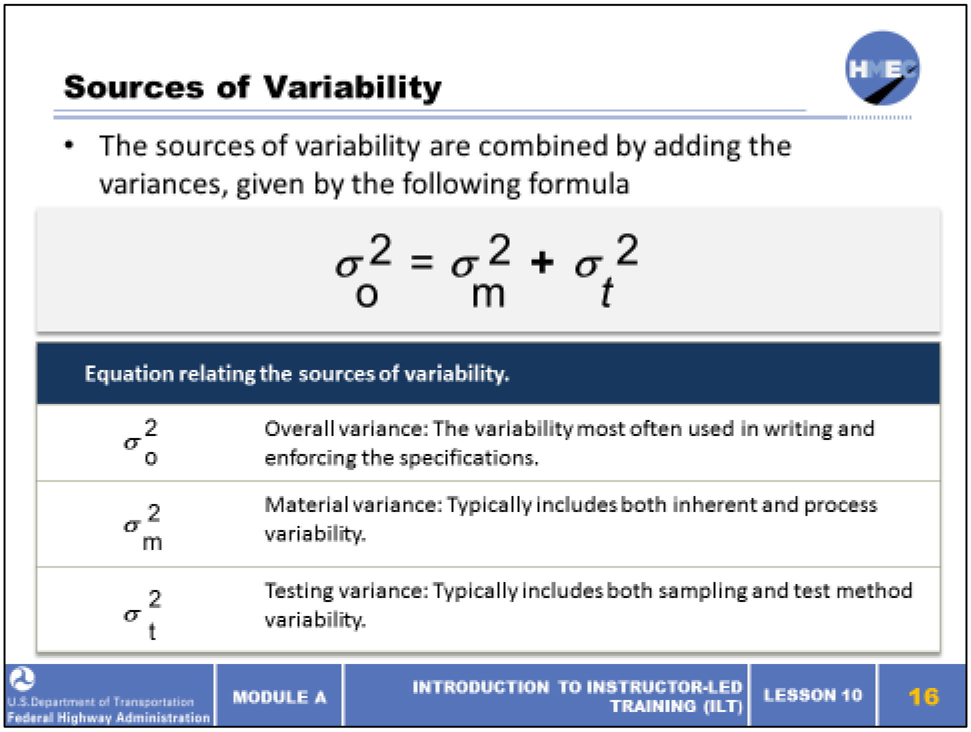
Slide 15



This screen shows the definitions of precision and accuracy using a target analogy. These targets illustrate that it is possible to be precise without being accurate (the target on the left), or accurate without being precise (the target in the middle). Ideally, we would like for a production or measurement process to be both accurate and precise (the target on the right).

____

Slide 16



This screen discusses the sources of variability, meaning that the variability of a sample can come from many different sources; statisticians call them "errors"—sampling error or testing error, for example. These terms mean sampling variability and testing variability, not mistakes. Sources of variability are combined by the use of the basic measure of variability, the variance, σ squared.

The sources of variability are combined by adding the variances given by the formula on the screen:

Sigma sub o squared equals sigma sub m squared plus sigma sub t squared. Where:

- Sigma sub o squared = overall variance, which is the variability most often used in writing and enforcing the specifications;
- Sigma sub m squared = material variance, which typically includes both inherent and process variability; and
- Sigma sub t squared = testing variance, which typically includes both sampling and test method variability.

As an example of these sources of variability, take aggregate production. When the aggregate is quarried, there is inherent variability in the rock in-place. When the aggregate is crushed, another source of variability is introduced. This also occurs when it is stockpiled, when it is hauled to the job site, and when it is placed. Lastly, when it is sampled and tested, two more sources of variability are introduced. Most likely, the only standard deviation identified will be the overall variability, σ_o, because of the time-consuming process that is required to identify the internal sources of variability.

____

____

____

____

____

____

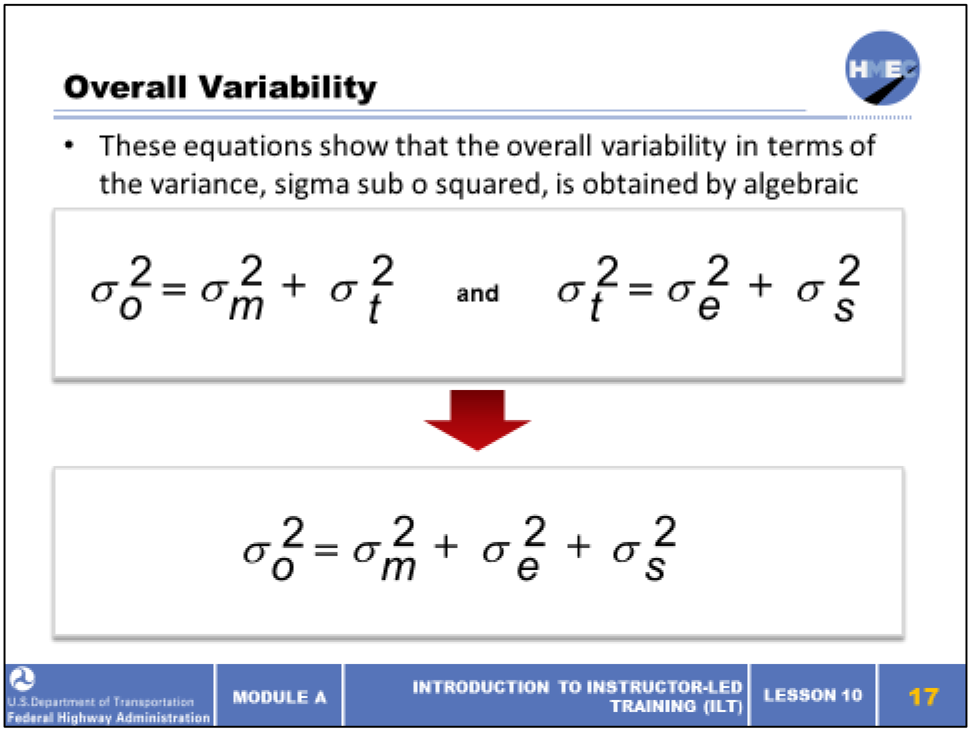
____

____

____

____

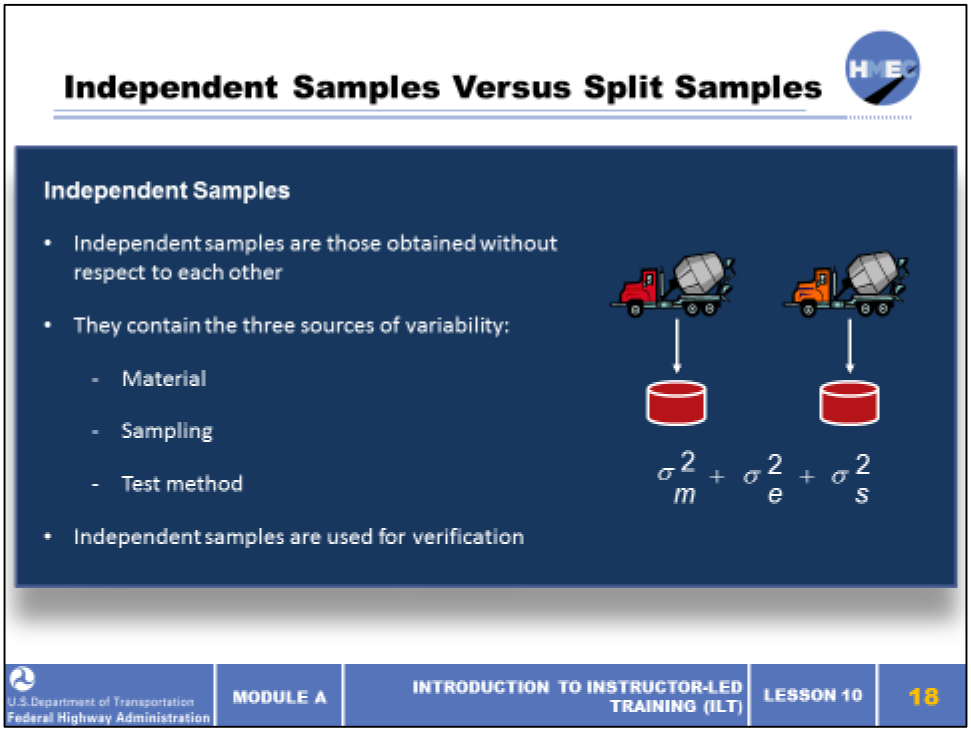
Slide 17



The relationship that relates the overall variability to material, test method, and sampling variability is given by the equation on the screen: Sigma sub o squared equals sigma sub m squared plus sigma sub t squared and sigma sub t squared equals sigma sub e squared plus sigma sub s squared; so by substitution, sigma sub o squared equals sigma sub m squared plus sigma sub e squared plus sigma sub s squared.

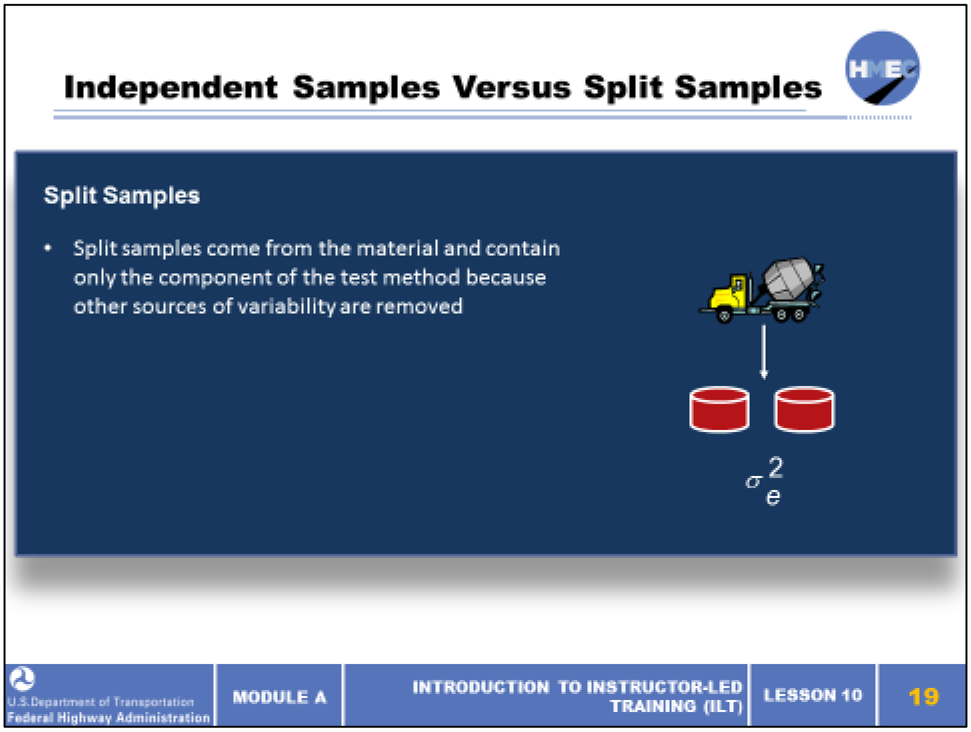
This equation shows that the variance of sampling and testing directly affects the overall variance. So any deviation from the sampling and testing protocol will affect the overall variability. If the overall variability is larger than what is expected, often the first thing that comes to mind is that the process has gotten more variable. But it may be equally likely that the sampling and testing is the reason for the increase.

Slide 18



Independent samples are those obtained without respect to each other. They contain the three sources of variability: material, sampling, and test method. Independent samples are used for verification. Split samples contain only the component of the test method because other sources of variability are removed.

Slide 19



Whereas as we just saw Independent samples contain the three sources of variability: material, sampling, and test method. On the other hand, Split samples come from the same material and contain only the component of the test method because other sources of variability are removed. So it is important to know what type of sample is being analyzed.

____

____

____

____

____

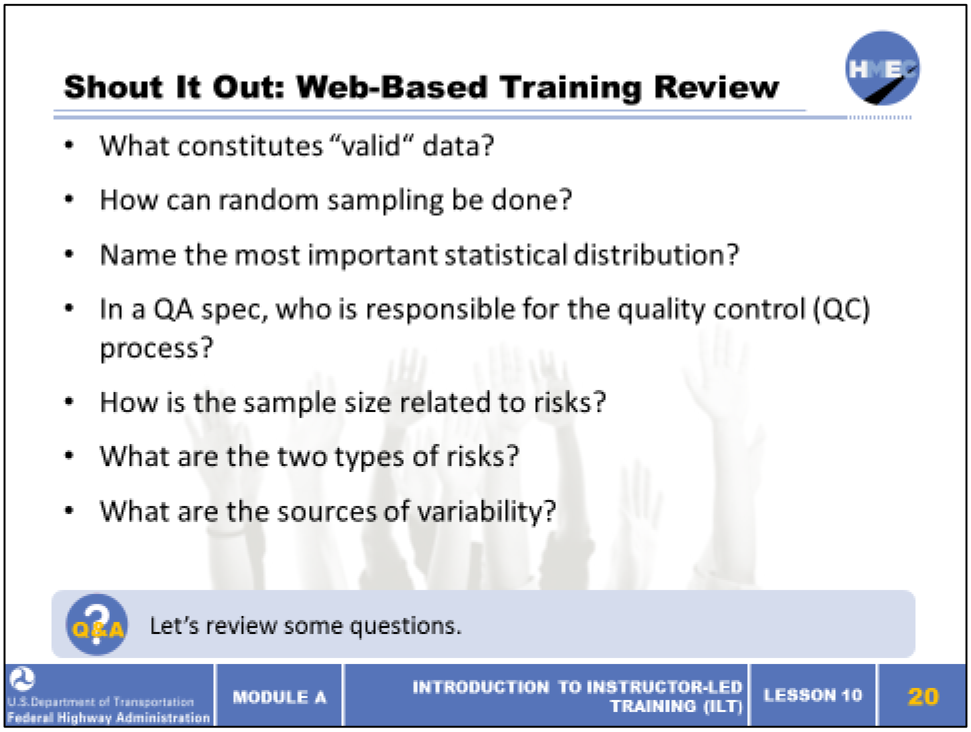
____

____

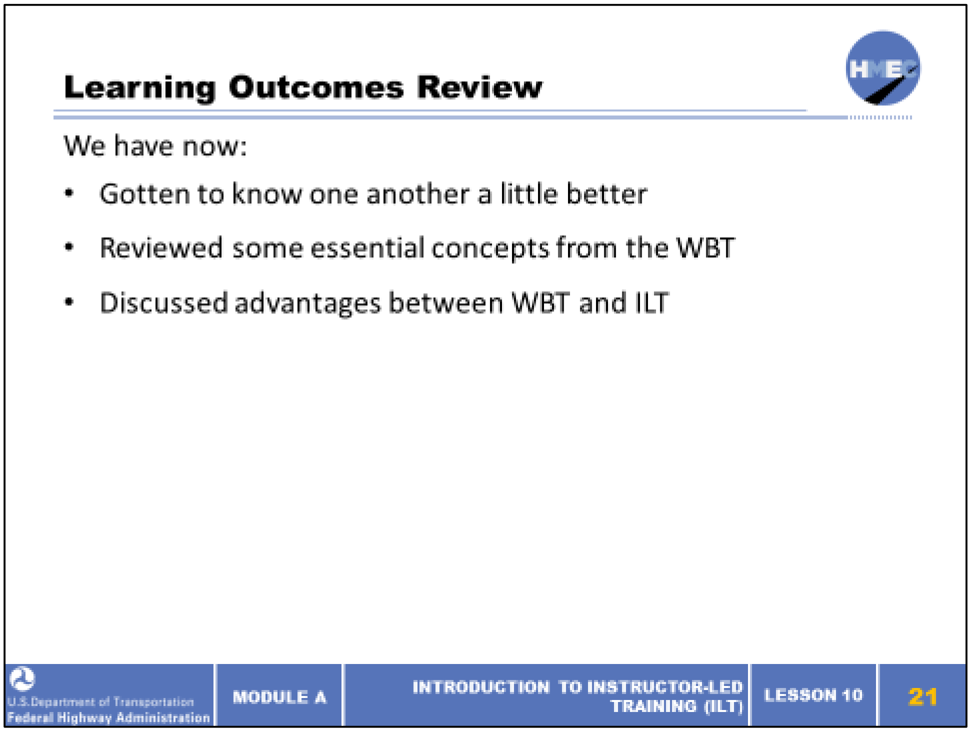
____

____

Slide 20


Shout It Out: Web-Based Training Review
HMEC
• What constitutes "valid" data?
• How can random sampling be done?
• Name the most important statistical distribution?
• In a QA spec, who is responsible for the quality control (QC) process?
• How is the sample size related to risks?
• What are the two types of risks?
• What are the sources of variability?
Q&A
Let's review some questions.
U.S. Department of Transportation
Federal Highway Administration
MODULE A
INTRODUCTION TO INSTRUCTOR-LED TRAINING (ILT)
LESSON 10
20

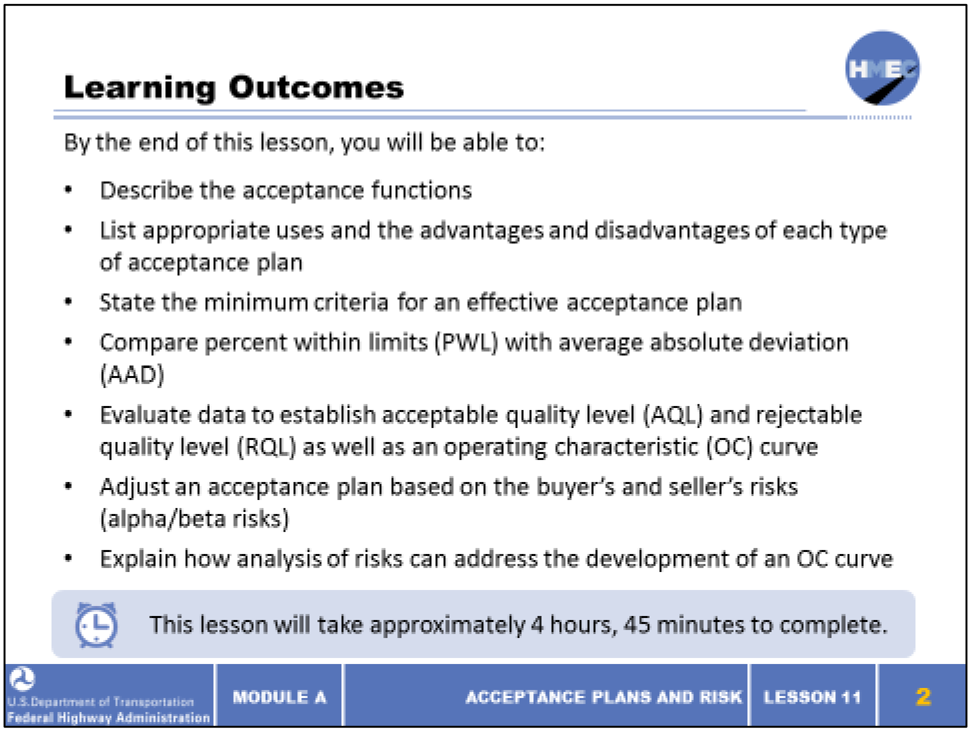
Slide 21


Learning Outcomes Review
We have now:
• Gotten to know one another a little better
• Reviewed some essential concepts from the WBT
• Discussed advantages between WBT and ILT
U.S. Department of Transportation
Federal Highway Administration
MODULE A
INTRODUCTION TO INSTRUCTOR-LED TRAINING (ILT)
LESSON 10
21

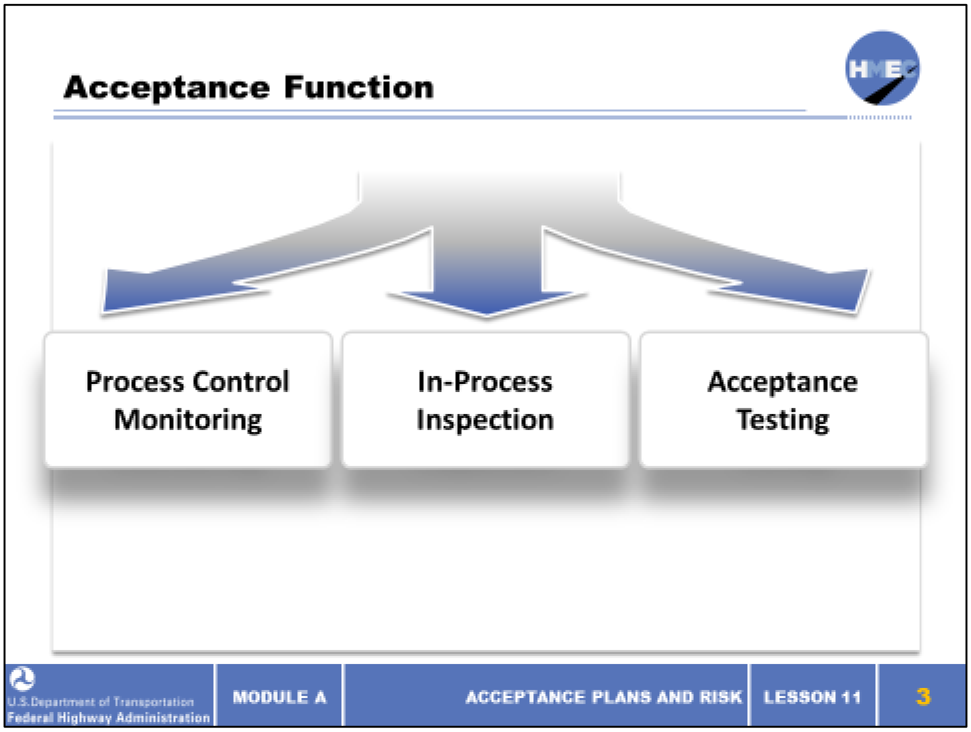
Slide 1


ILT
HMEC
Highway Materials Engineering Course
Lesson 11: Acceptance Plans and Risk
Quality Assurance
U.S.Department of Transportation
Federal Highway Administration
MODULE
A

Slide 2

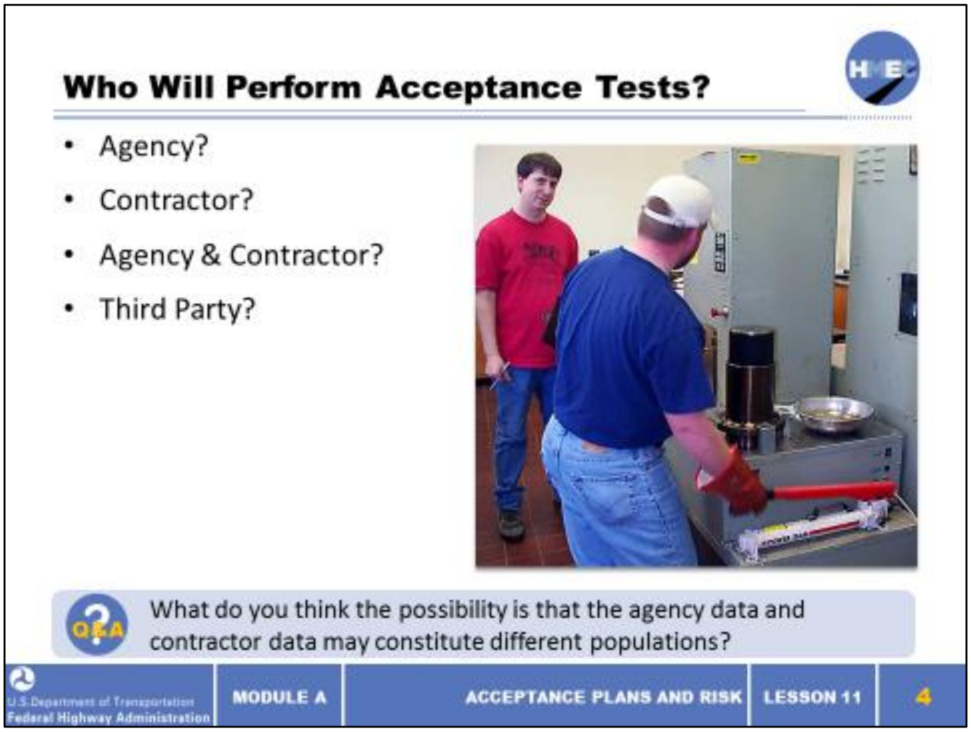

Learning Outcomes
HMEC
By the end of this lesson, you will be able to:
• Describe the acceptance functions
• List appropriate uses and the advantages and disadvantages of each type of acceptance plan
• State the minimum criteria for an effective acceptance plan
• Compare percent within limits (PWL) with average absolute deviation (AAD)
• Evaluate data to establish acceptable quality level (AQL) and rejectable quality level (RQL) as well as an operating characteristic (OC) curve
• Adjust an acceptance plan based on the buyer's and seller's risks (alpha/beta risks)
• Explain how analysis of risks can address the development of an OC curve
This lesson will take approximately 4 hours, 45 minutes to complete.
U.S. Department of Transportation
Federal Highway Administration
MODULE A
ACCEPTANCE PLANS AND RISK
LESSON 11
2

Slide 3



The acceptance testing part of the function is very important. It is performed to provide the specifying agency with information that relates the product to the specification limit(s). There are many factors and elements to consider in specification design. One of the principal elements is the acceptance plan.

Slide 4



The answer to this question will influence subsequent decisions and procedures in the acceptance plan. The agency may decide to do the acceptance testing, may assign the testing to the contractor, may have a combination of agency and contractor acceptance testing, or may require a third party to do the testing. Using third-party testing in the acceptance decision is often a consultant hired by the contractor. This may occur if the contractor doesn't have a certified quality control (QC) technician available. Under all conditions the agency is responsible for making the acceptance decision.

____

____

____

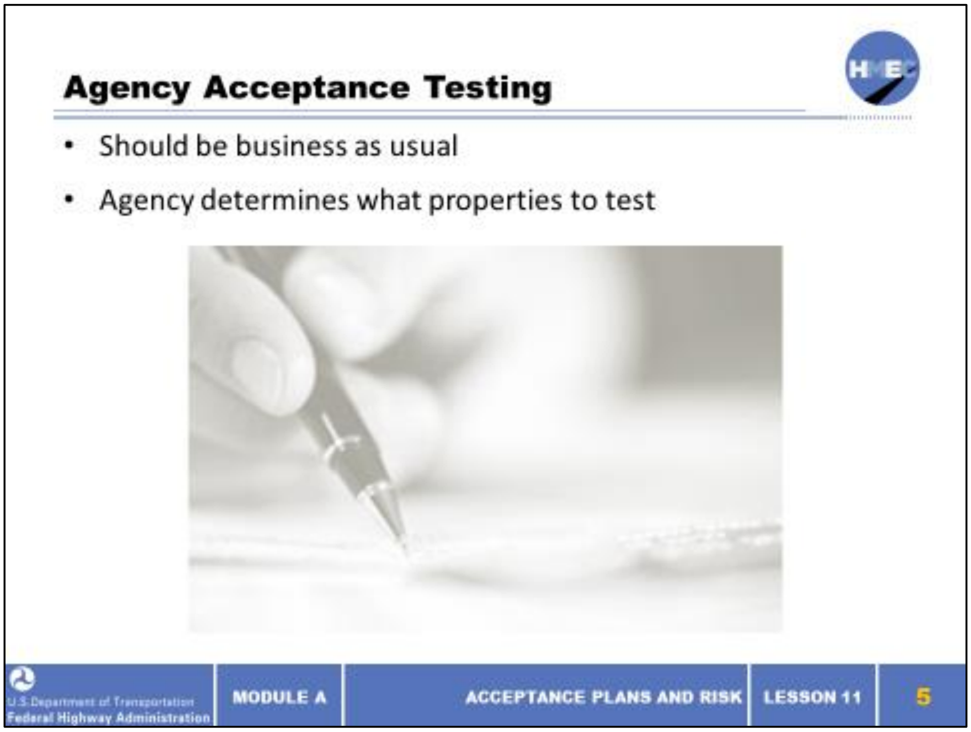
____

____

____

____

Slide 5

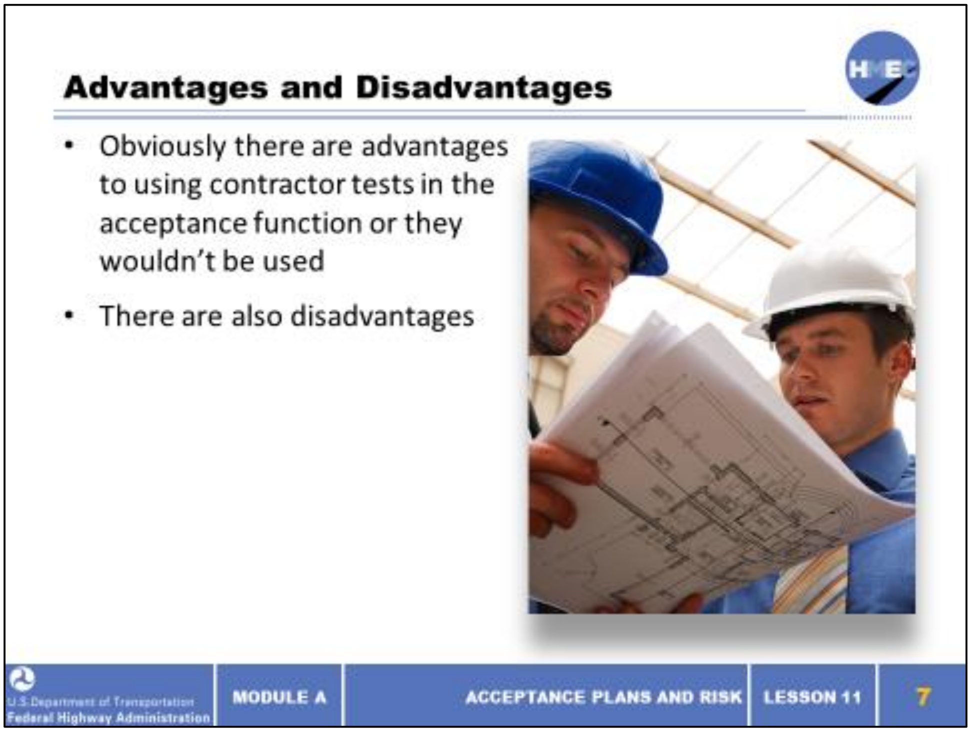


If the agency does the acceptance testing, “business as usual” will be the predominate theme and the next step is to determine what properties to measure. Typically, irrespective of who does the acceptance testing, the agency determines what properties to measure. However, if the contractor performs this function there are additional requirements.

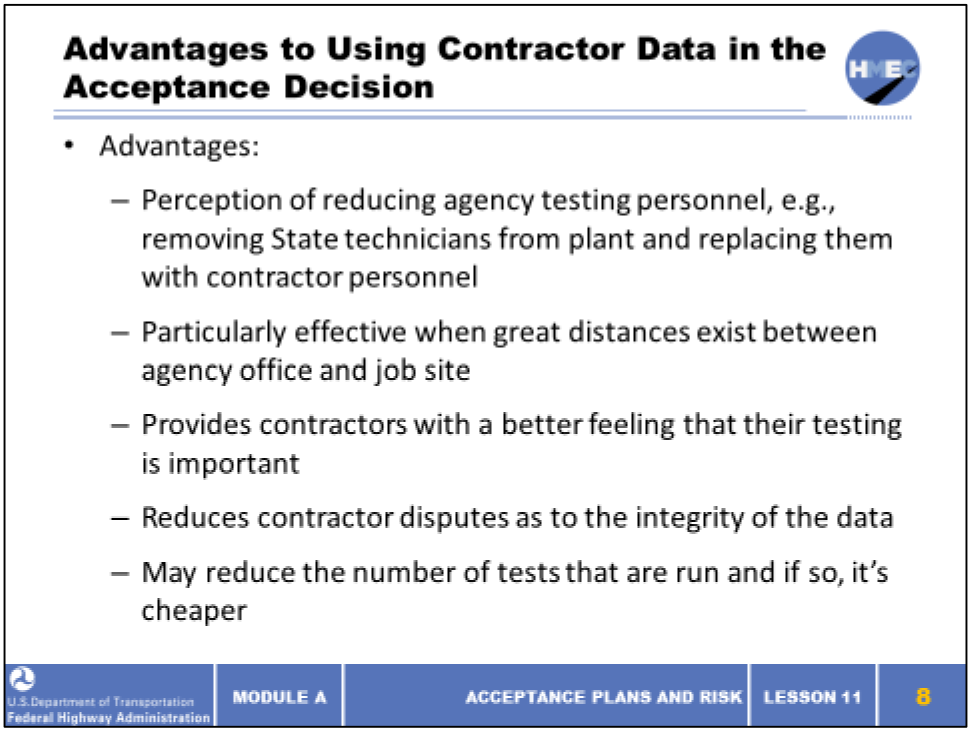
Slide 6

The decision as to who does the testing is usually related to the agency’s personnel assessment, particularly when there is agency downsizing. The lack of personnel availability by the agency should not be the main reason to go to contractor acceptance testing; however, this has, most often, proven to be the case. Furthermore, contractors should never be assigned the responsibility for acceptance testing without providing the industry sufficient time to prepare to assume this task, especially in terms of qualified personnel and facilities.

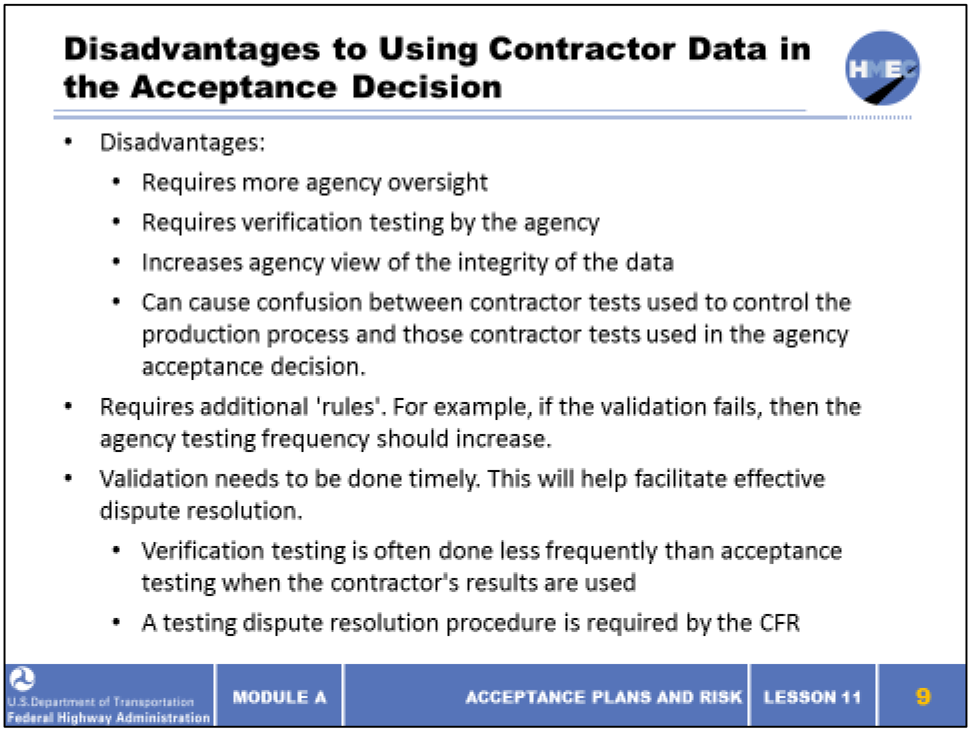
Slide 7


Advantages and Disadvantages
• Obviously there are advantages to using contractor tests in the acceptance function or they wouldn't be used
• There are also disadvantages
U.S. Department of Transportation
Federal Highway Administration
MODULE A
ACCEPTANCE PLANS AND RISK
LESSON 11
7

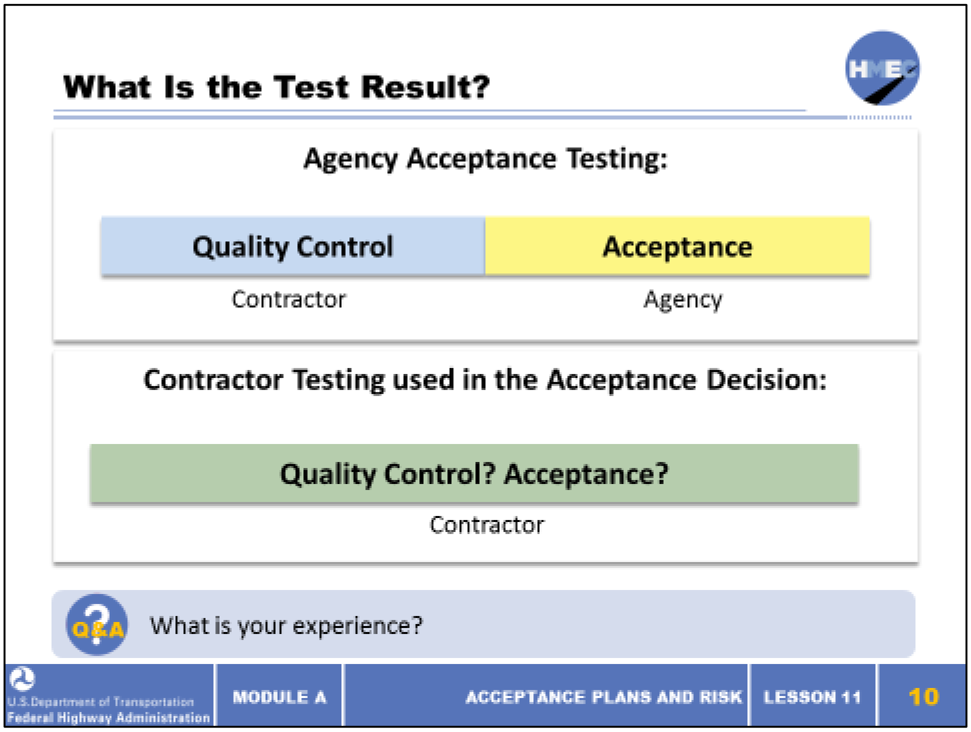
Slide 8


Advantages to Using Contractor Data in the Acceptance Decision
HMEC
• Advantages:
– Perception of reducing agency testing personnel, e.g., removing State technicians from plant and replacing them with contractor personnel
– Particularly effective when great distances exist between agency office and job site
– Provides contractors with a better feeling that their testing is important
– Reduces contractor disputes as to the integrity of the data
– May reduce the number of tests that are run and if so, it's cheaper
U.S. Department of Transportation
Federal Highway Administration
MODULE A
ACCEPTANCE PLANS AND RISK
LESSON 11
8

Slide 9


Disadvantages to Using Contractor Data in the Acceptance Decision
HMEC
- Disadvantages:
  - Requires more agency oversight
  - Requires verification testing by the agency
  - Increases agency view of the integrity of the data
  - Can cause confusion between contractor tests used to control the production process and those contractor tests used in the agency acceptance decision.
- Requires additional 'rules'. For example, if the validation fails, then the agency testing frequency should increase.
- Validation needs to be done timely. This will help facilitate effective dispute resolution.
  - Verification testing is often done less frequently than acceptance testing when the contractor's results are used
  - A testing dispute resolution procedure is required by the CFR
U.S. Department of Transportation
Federal Highway Administration
MODULE A
ACCEPTANCE PLANS AND RISK
LESSON 11
9

Slide 10



When agencies do the acceptance testing, there is a clear distinction between the QC and acceptance function. But when agencies require the contractor testing to be used in the acceptance decision, the acceptance function can become more complicated for several reasons. One reason is that the contractor is required to perform both QC and acceptance testing, so is the test a QC test or an acceptance test? These two functions can become intermingled if care is not taken to assure their separation.

____

____

____

____

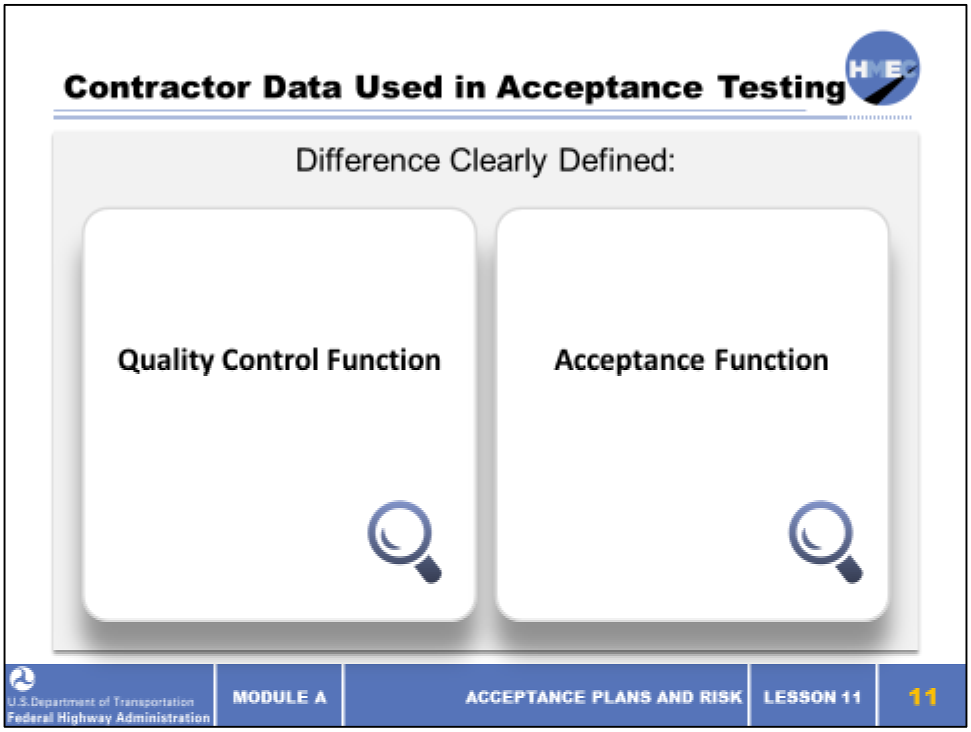
____

____

____

____

Slide 11



Agencies that have required contractor data to be used in the QC and acceptance have not always made this distinction clear and it has led to disputes.

____

____

____

____

____

____

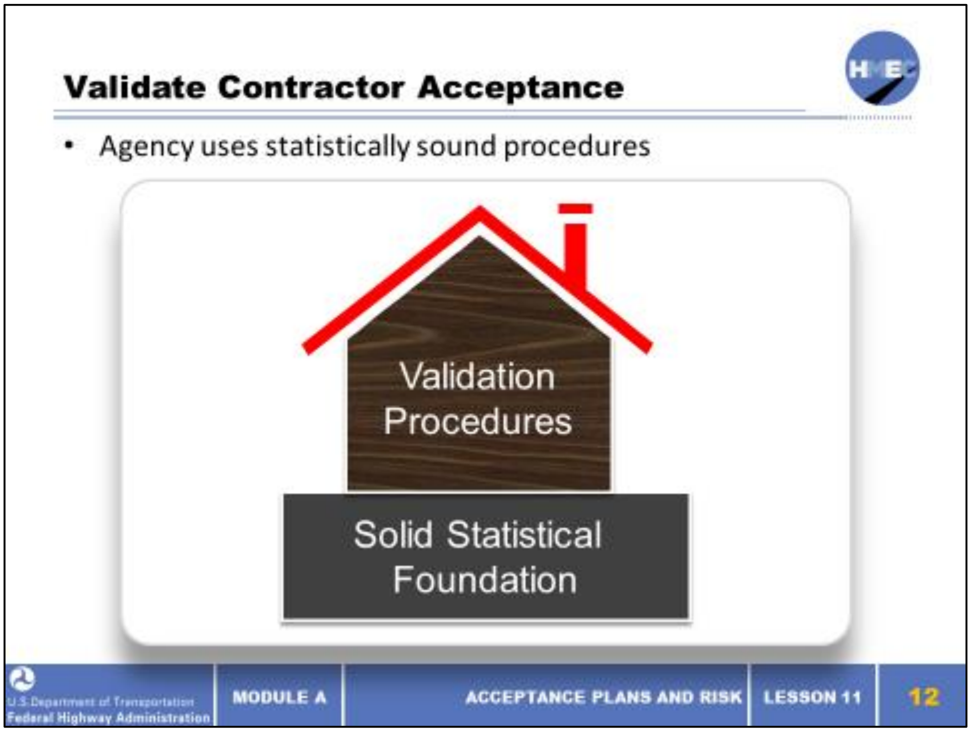
____

____

____

____

Slide 12



Different validation procedures are discussed in Lesson 14 where we'll get into more detail. For now, let it suffice to say the validation procedure should be based on sound statistical principles.

____

____

____

____

____

____

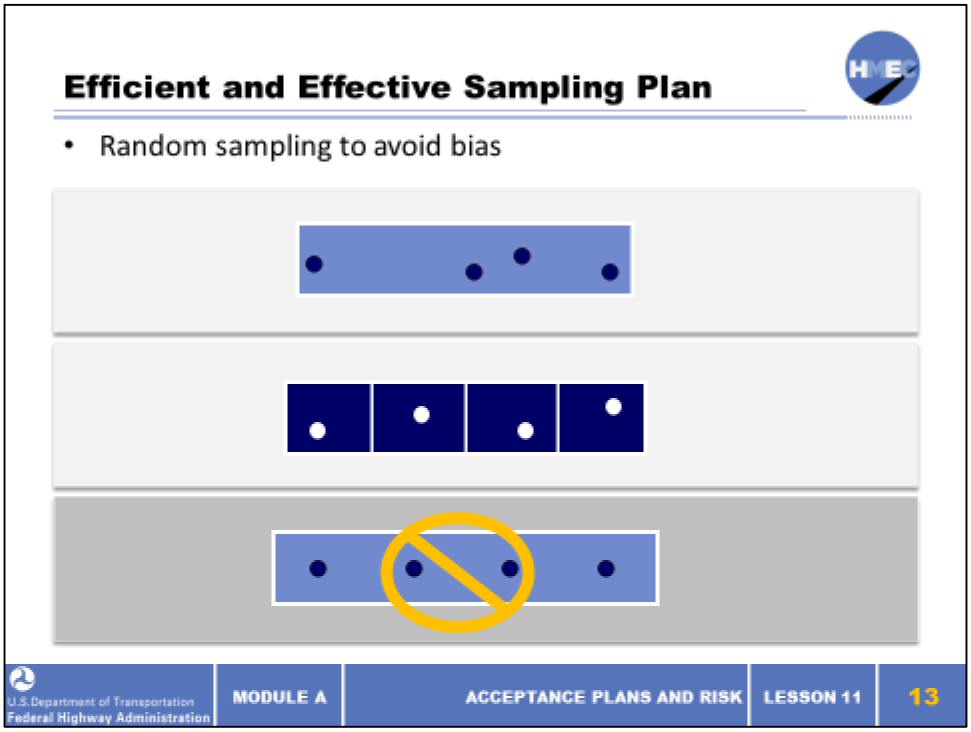
____

____

____

____

Slide 13



Does this message sound familiar? This is one of the necessities for obtaining valid data. If the sampling is not done in some random fashion, it is likely that bias will enter into the sampling program.

____

____

____

____

____

____

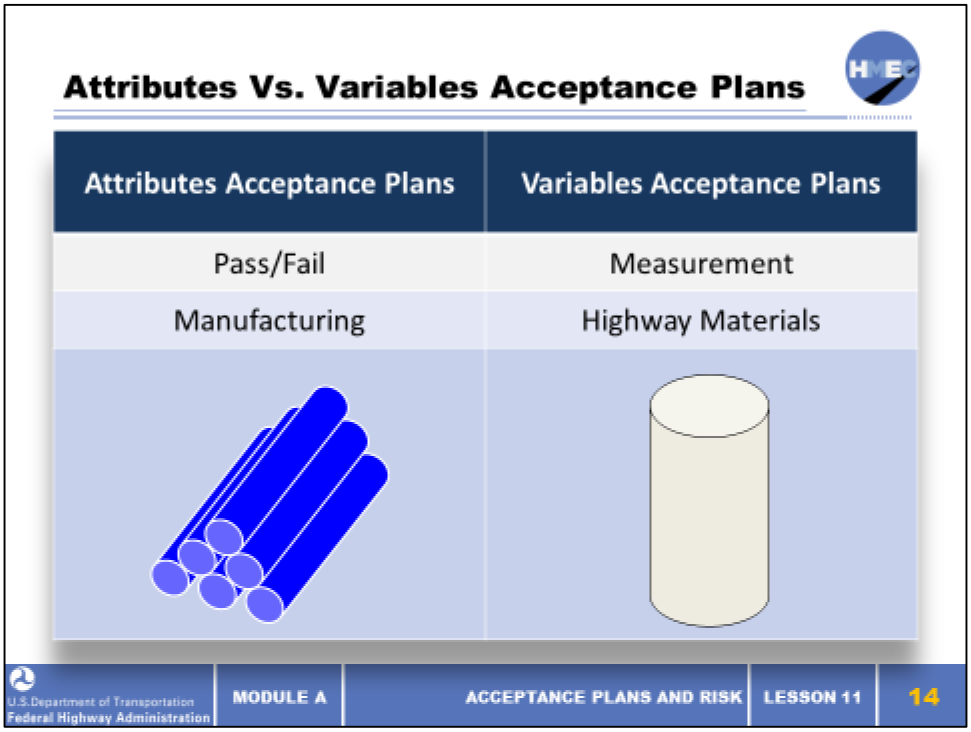
____

____

____

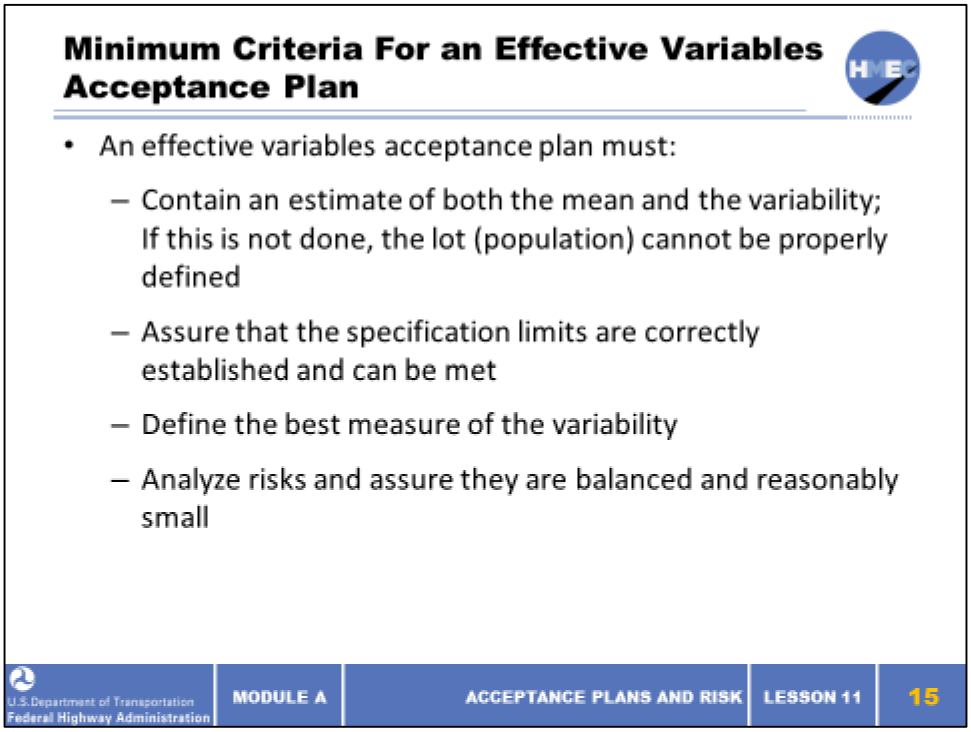
____

Slide 14



In an attributes acceptance plan, the sample either passes or fails and the lot is either accepted or rejected. In a variables plan, the quality characteristics (i.e., asphalt content, concrete pavement thickness) are measured and the values of the measurements are important pieces of information. Attributes acceptance plans are most appropriate in manufacturing processes. For example, a specification to accept asphalt coated corrugated pipe might use an attributes acceptance plan. This lesson concentrates on variables acceptance plans, which are more common to highway materials and construction specifications. These are also a more efficient use of testing time and personnel.

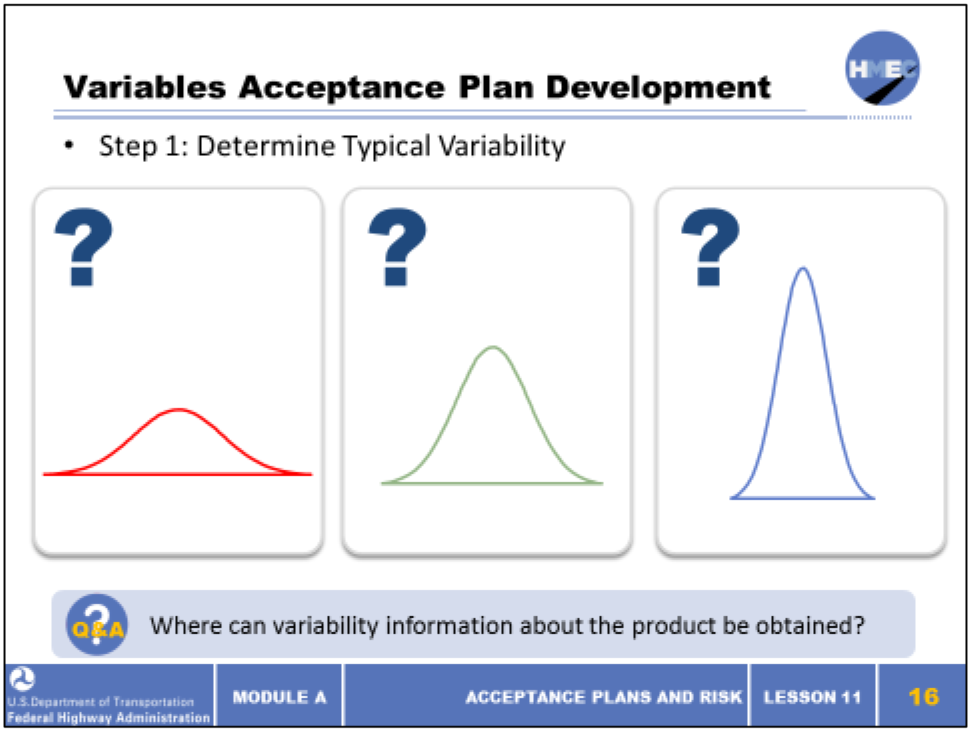
Slide 15



The criteria are that the variables acceptance plan must:

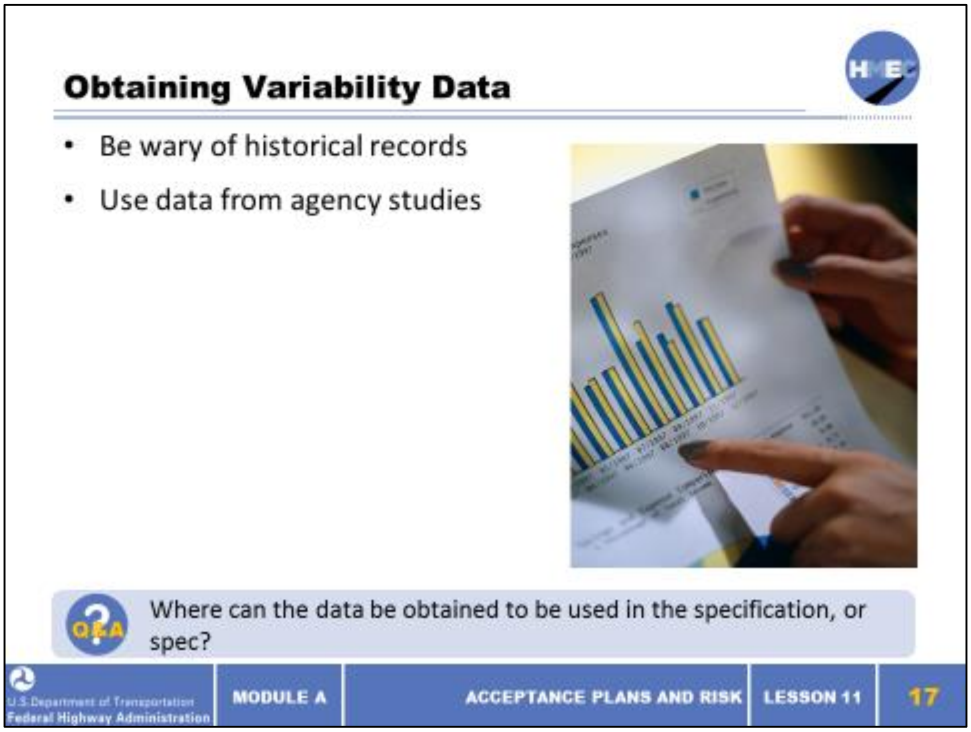
• Contain an estimate of both the mean and the variability (if this is not done, the population cannot be properly defined;
• Assure that the specification limits are correctly established and can be met;
• Define the best measure of the variability; and
• Analyze risks and assure they are balanced and reasonably small.

Slide 16



Prior knowledge of the product is essential. Factors such as typical inherent process variability and the level of variability that can be tolerated must be identified

Slide 17



Historical records are a potential source of variability data, and are a source that agencies tend to use. However, this can be a dangerous source due to potential unknown bias and lack of random sampling procedures that may have been used when the data were obtained. Better sources of variability data are studies conducted by the particular agency to determine the needed information.

Many studies of construction and materials processes, such as those presented in the WBT Lesson 8 by the West Virginia DOT, have been conducted to provide variability data. The results of these studies should be verified as applicable to the particular agency before they are used in developing specifications.

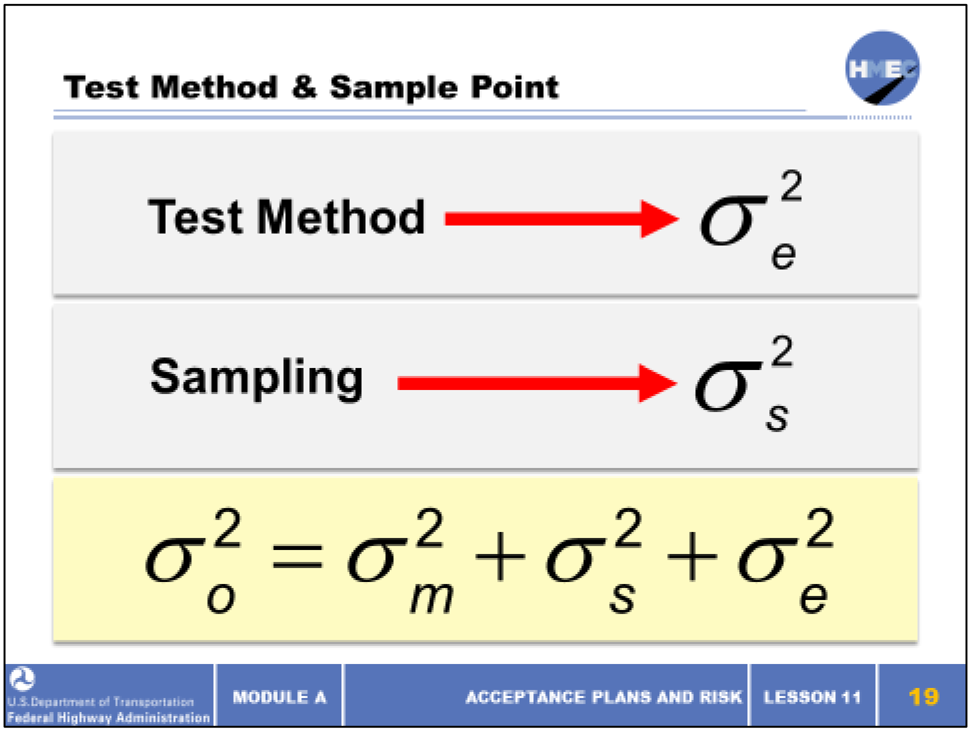
Studies conducted by the particular agency should use random sampling procedures and unbiased reporting of test results. Such studies may be necessary when initially developing the quality assurance (QA) acceptance plan and also in cases where new processes are introduced for which no measures of variability have been previously established.

Slide 18

A well-written acceptance plan should consider at least the following factors:

- Method of test and point of sampling;
- Lot size;
- Sample size;
- Acceptance limits;
- Hypothesis Testing and Risks; and
- Operating characteristics (OC) curve.

Slide 19



As discussed in Lesson 8, the test method impacts the overall variability and thus the specification limits. While there are often many choices for the point of sampling, a single point should be specified. Again, the point of sampling often influences the overall variability. Both of these elements should be the same as those used when establishing the acceptance limits of the specification. Thus, when gathering data in the studies mentioned above, it is best if the individual factors are decided upon prior to proceeding with the data collection.

____

____

____

____

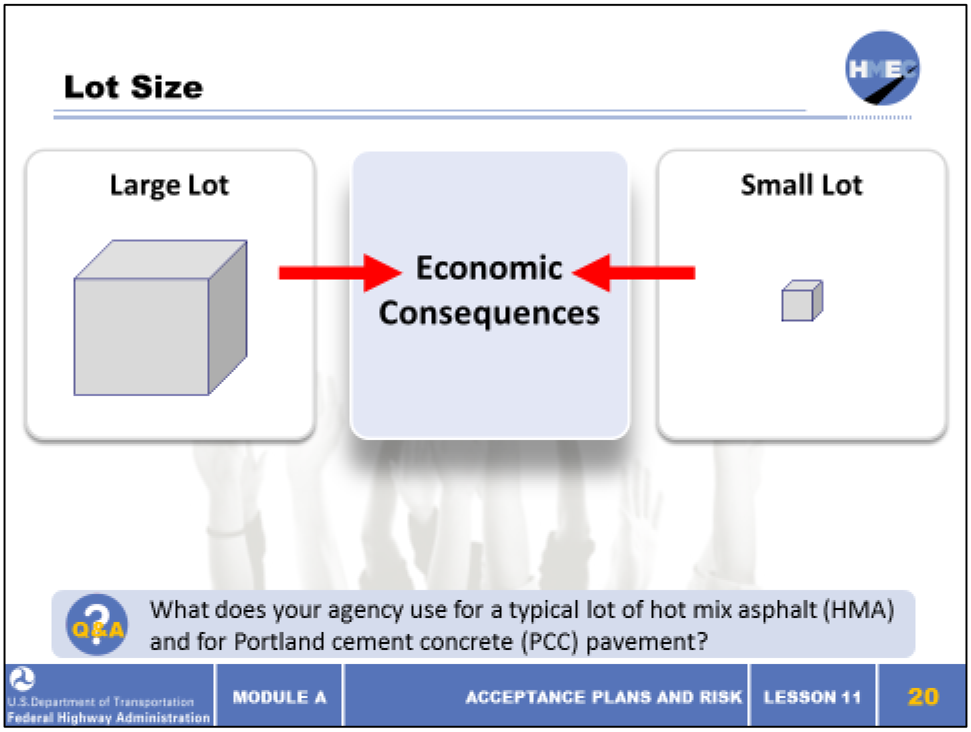
____

____

____

____

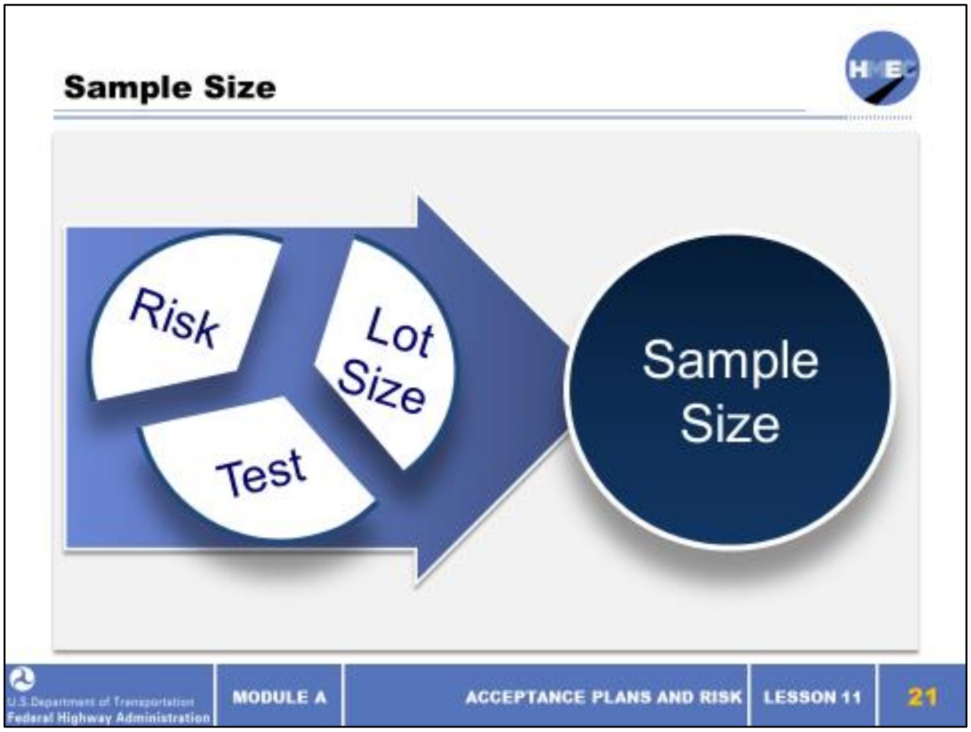
Slide 20



A lot is the amount of product that is to be judged acceptable or unacceptable on the basis of a sample comprised of a stated number of test results. As discussed in the WBT Lesson 4, stratified lots consisting of a number of equal size sublots are often used. Since the number of specimens in the sample may remain constant for a lot of a particular product, the determination of the most appropriate lot size is often economic. If the lot is very large (e.g., an entire project), the cost of rejecting the product or adjusting the payment can have severe negative consequences on the contractor. However, some agencies have been successful in using an entire project as a definition of a lot. If this definition is used, close communication between the owner and the contractor is imperative to ensure that a large negative price adjustment is not imminent.

On the other hand, the consequences of a small lot (e.g., a few loads of material) are that the cost of testing may be more than the benefits provided. Generally, a lot is defined in terms of time, amount of production, or area.

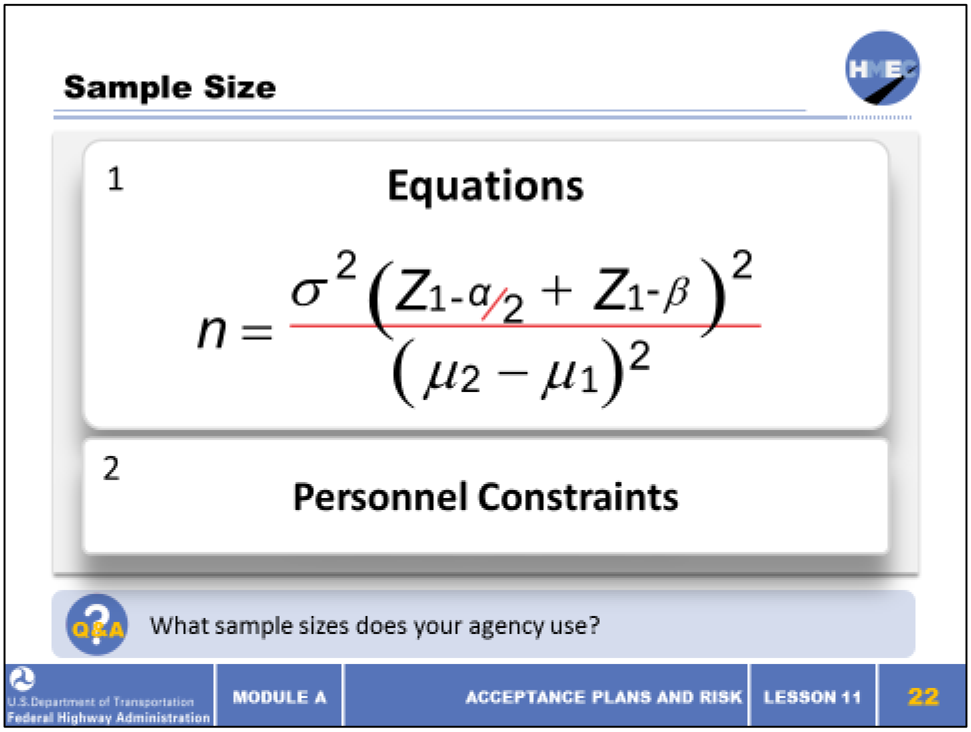
Slide 21



The proper sample size is associated with the risk that the specification writer uses in the specification development. The sample size is related to lot size, i.e., generally, the larger the lot size the larger the sample size. As will be seen when risks are discussed, generally the larger the sample size, the lower the buyer and seller risks. In most QA specifications, the sample size may be as few as three to five. Larger sample sizes typically exist for quick tests, such as using the nuclear density gauge or large lot sizes, such as an entire project.

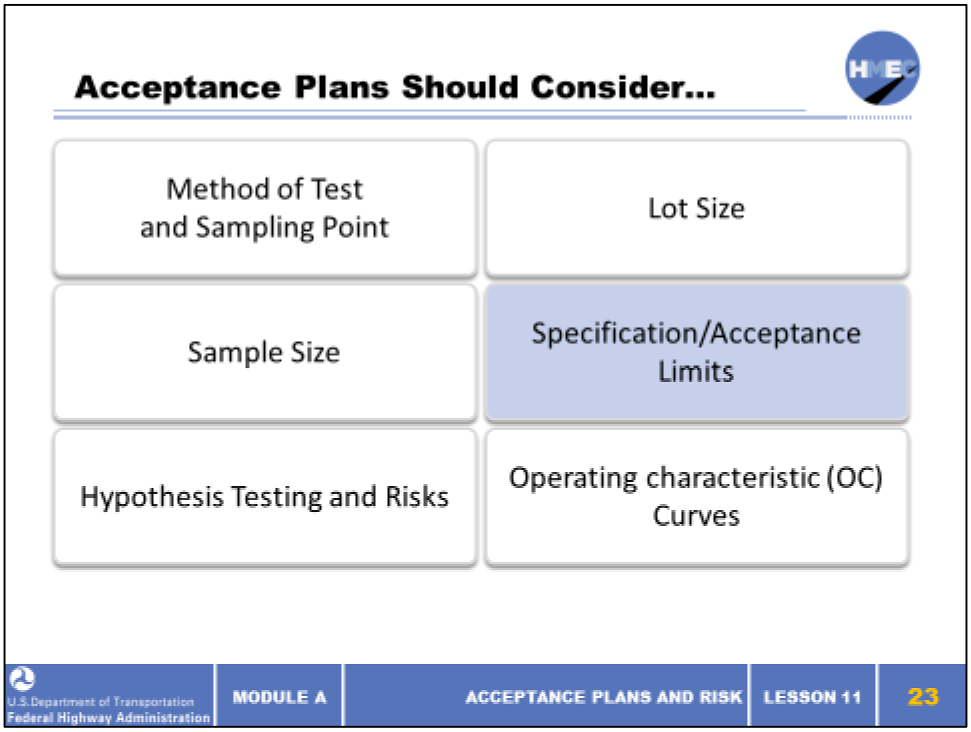
One weakness in many acceptance plans is the very small sample sizes used. As will be seen shortly when the magnitude of risks is analyzed, small sample sizes can result in larger than desirable risks, particular for the buyer. While intuitively, some would like to accept product on a single sample, n=1, it will be seen that the risks can be inordinately high.

Slide 22

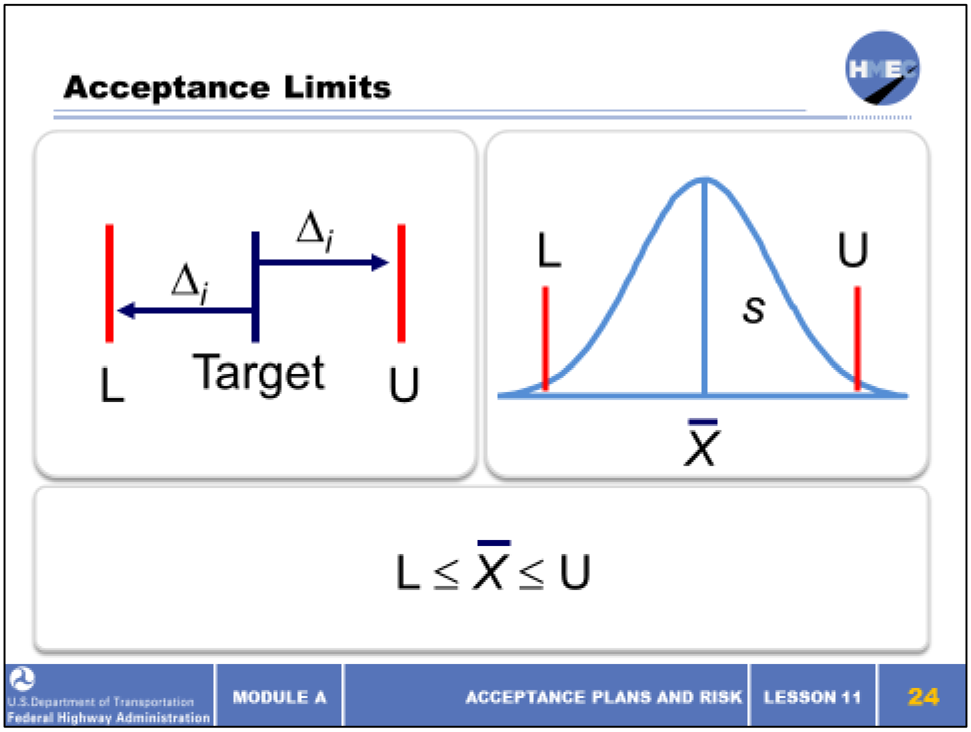


For illustration purposes, one of these equations is shown here. In this equation, the specification developer must know the risks, α and β, and the error term, $\mu_2 - \mu_1$, that the agency is willing to accept. Although equations for calculating sample size can be used, most often the sample size is determined by the personnel constraints.

Slide 23


Acceptance Plans Should Consider…
HMEC
Method of Test and Sampling Point
Lot Size
Sample Size
Specification/Acceptance Limits
Hypothesis Testing and Risks
Operating characteristic (OC) Curves
U.S. Department of Transportation Federal Highway Administration
MODULE A
ACCEPTANCE PLANS AND RISK
LESSON 11
23

Slide 24



Acceptance plans based on measuring only the mean of the sample are discussed in this lesson to illustrate risk determination. More appropriate acceptance plans measuring both the mean and standard deviation of the sample and the difference from the target value are discussed in the next lesson.

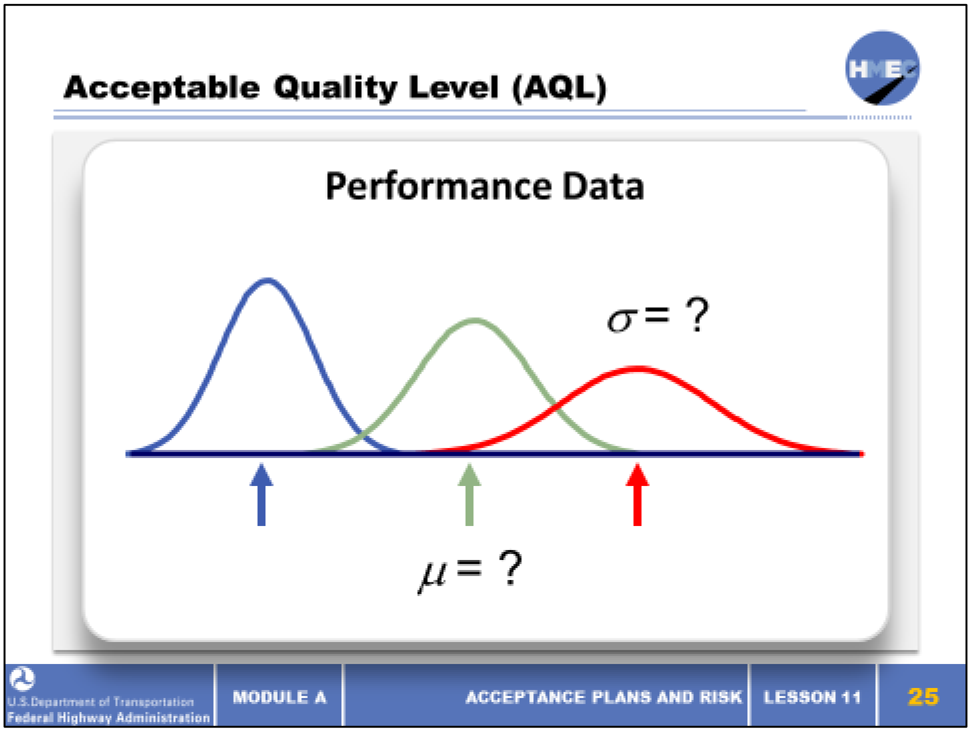
This discussion of acceptance limits based on only the mean is intended to ease into the understanding of statistical risks and, as will be emphasized is not an appropriate way of writing an effective acceptance plan. Both population parameters must be used to estimate the population but the standard deviation is "fixed" in all these calculations. This type of acceptance plan is called a “variability known” acceptance plan, which is a misnomer in reality. In highway construction it is rare, if ever, that the lot standard deviation is known. But it is a crucial step to determine and use the “typical” variability when determining the acceptance limits. Otherwise the acceptance limits may be either “too tight” or ineffective if “too loose”.

____

____

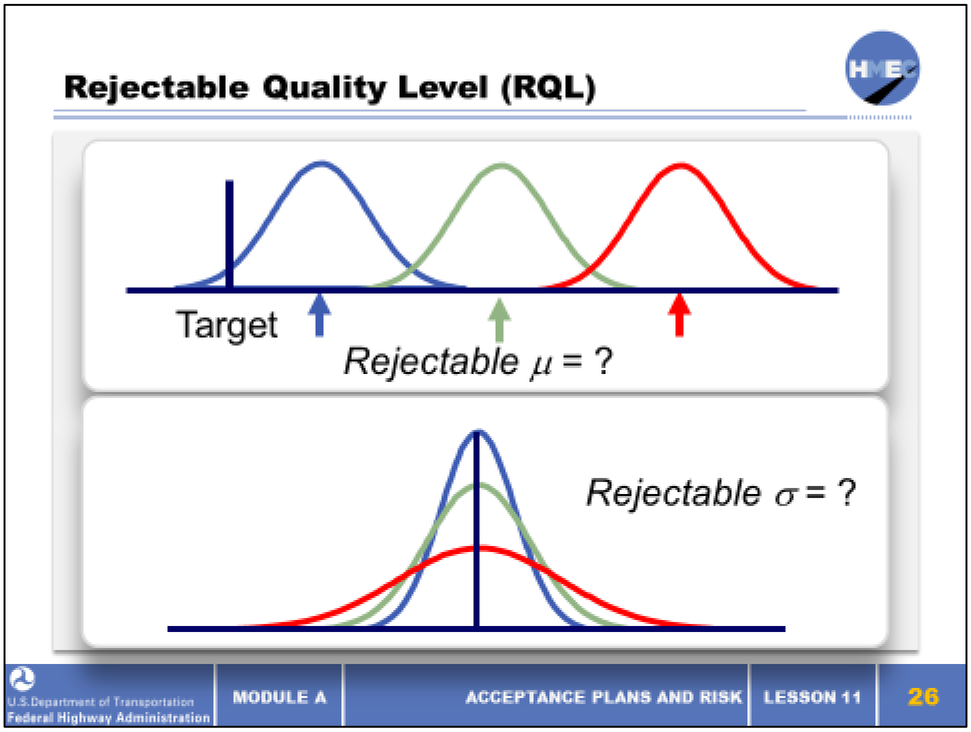
____

Slide 25



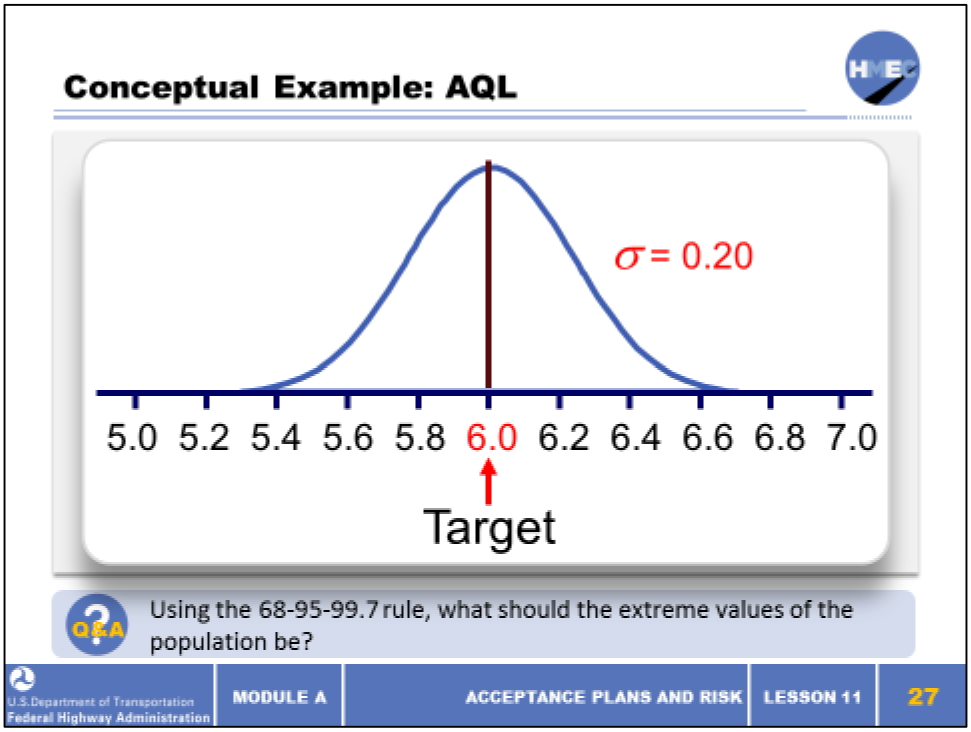
This decision defines acceptable material and should address the material that will provide satisfactory service at an affordable cost when used for the intended purpose. What constitutes acceptable material is often determined based on what has performed well in the past. However, it is preferable if performance data are available to quantify performance. The level at which the material is considered just acceptable is known as the acceptable quality level (AQL). Statistics has been a valuable tool in defining the population parameters (mean and standard deviation) of acceptable material. Caution should be exercised if a lower variability is chosen for the specification than has been shown to be readily achievable. Arbitrarily tightening the specs can increase the cost of the material above that which may be cost effective.

Slide 26



Unacceptable material should have a low probability of being accepted, or be material that will only be accepted under the conditions of a reduced payment schedule. This is the level at which the material is considered unacceptable and hypothetically requires removal and replacement.

Slide 27



While it is preferable to use an acceptance approach that considers both mean and standard deviation, control of only the mean is simpler to understand and is used here only as a simplified instruction example. It has been determined that for asphalt content, acceptable material has a standard deviation of about 0.20% when the mean is close to the target (job mix) value. If the job mix formula has established the target as 6.0% asphalt content, acceptable material could be defined as shown. In this example, the AQL is a lot (population) with a mean of 6.0% and a standard deviation of 0.20%.

____

____

____

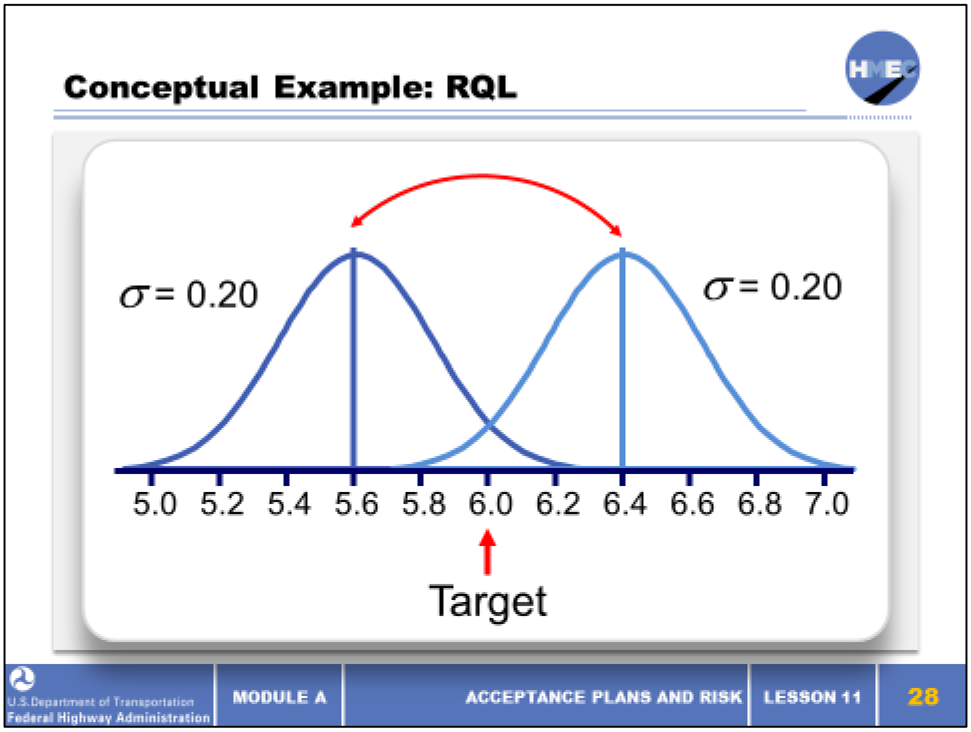
____

____

____

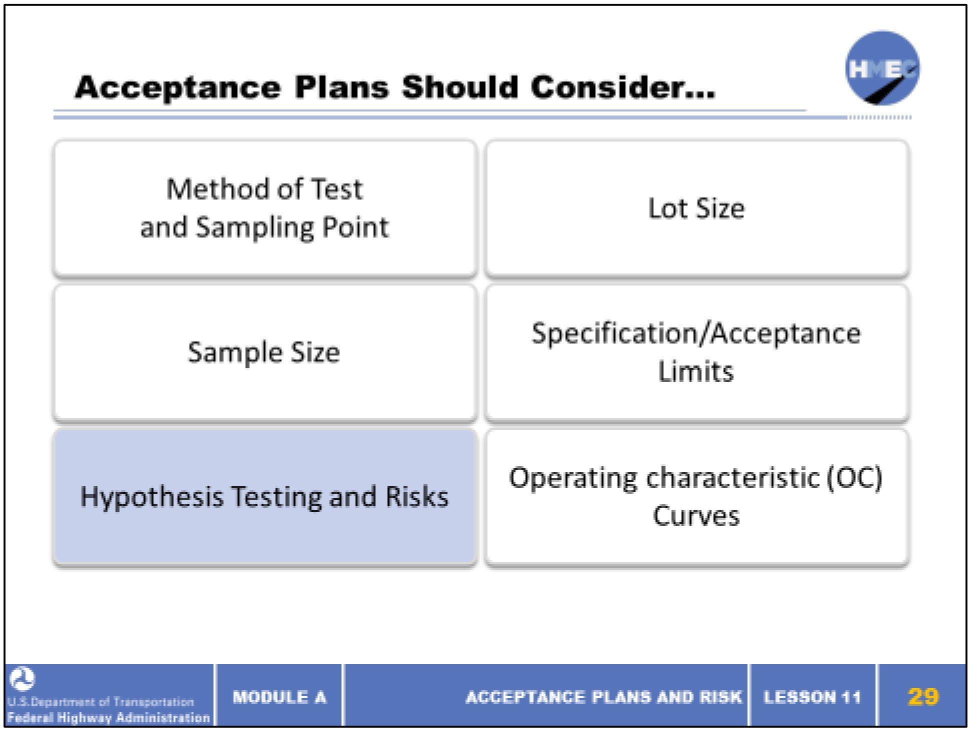
____

Slide 28

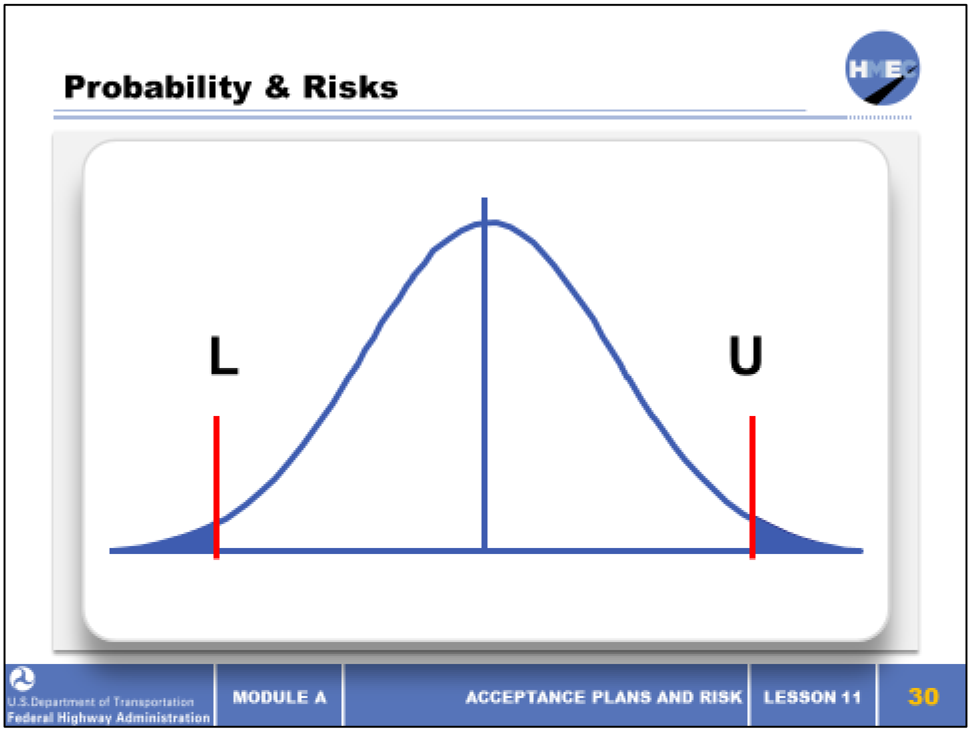


Unacceptable material might be defined as that for which the mean differs from the target value by 0.4% or more. That defines unacceptable material as shown. The RQL is a lot (population) with a standard deviation of 0.20% and a mean as low as 5.6% or as high as 6.4%. (Other definitions of unacceptable material would be equally valid.) But keep in mind, for the RQL material it can be either the population on the right or that on the left; it cannot be both.

Slide 29


Acceptance Plans Should Consider…
HMEC
Method of Test
and Sampling Point
Lot Size
Sample Size
Specification/Acceptance
Limits
Hypothesis Testing and Risks
Operating characteristic (OC)
Curves
U.S. Department of Transportation
Federal Highway Administration
MODULE A
ACCEPTANCE PLANS AND RISK
LESSON 11
29

Slide 30



This concept is the primary basis for understanding why QA specifications are so much more powerful than other specifications for deciding whether or not a product meets a specification.

____

____

____

____

____

____

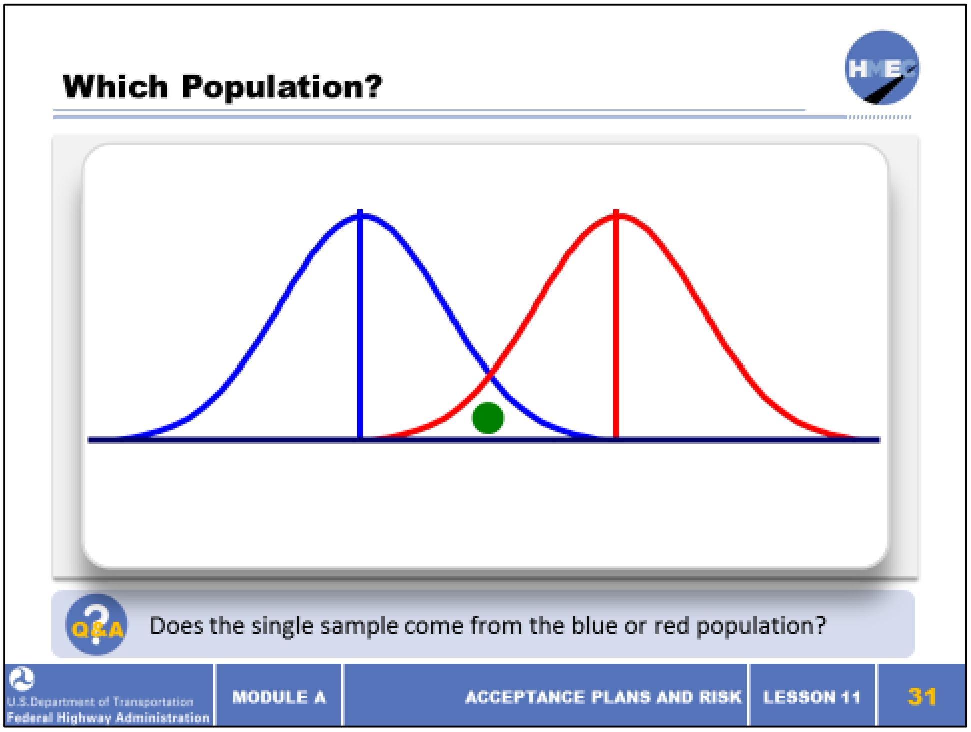
____

____

____

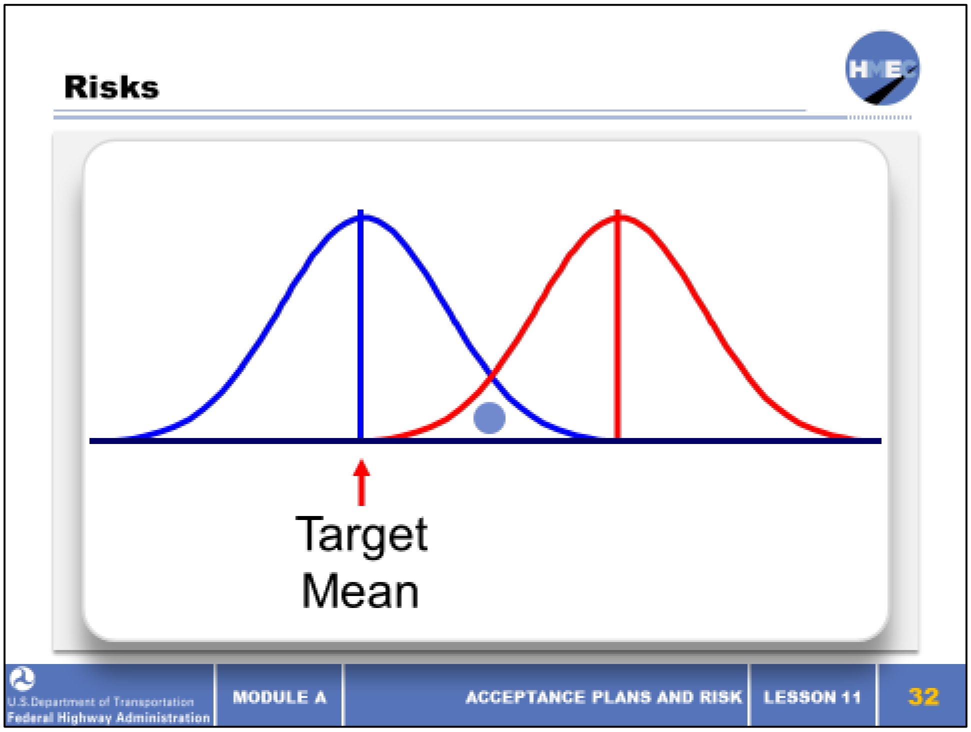
____

Slide 31



Does it matter? Remember, the purpose of a sample is to estimate the population. Since we never know from what population a sample actually comes, using sample estimates involves risks—risks of making the wrong decision.

Slide 32



However, does it represent "good" or “bad” material? It could come from either of the distributions shown, or from many others for that matter. If it is from the blue distribution, the population centered on the target value, which represents “good” material, we want to accept the result. Whereas if it is from the red distribution, the population is centered far from the target value and we want to reject that population.

____

____

____

____

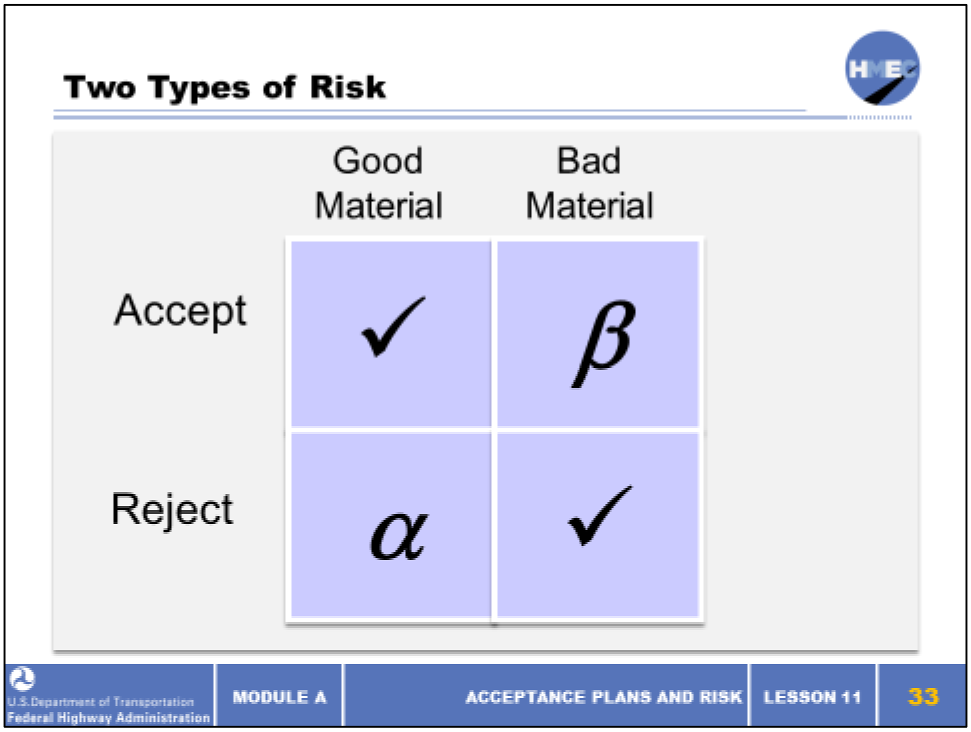
____

____

____

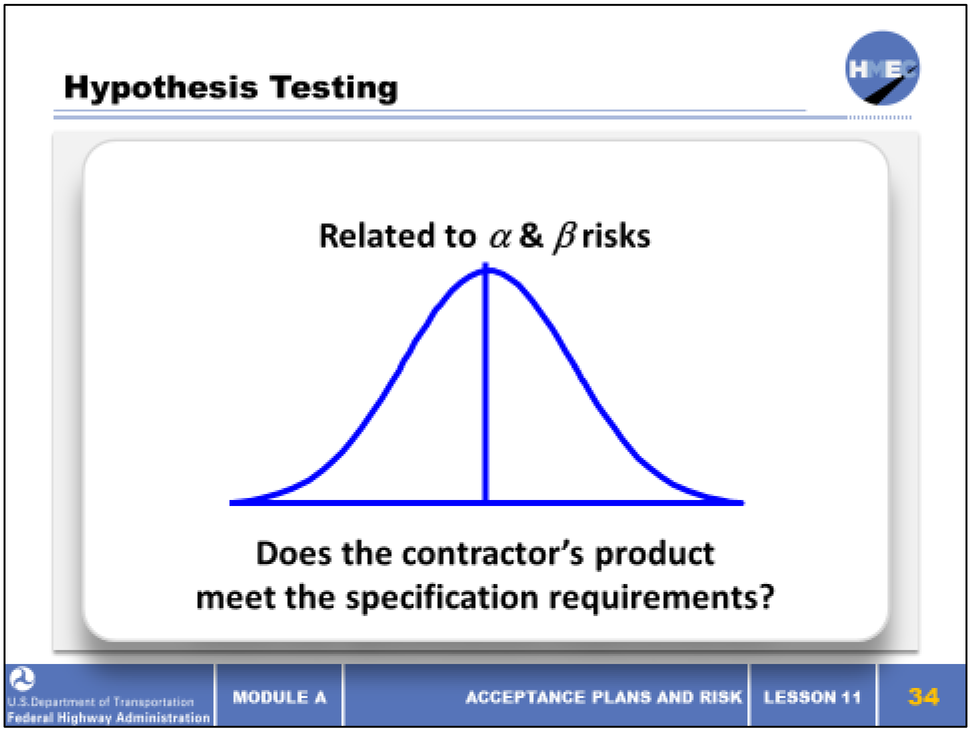
____

Slide 33



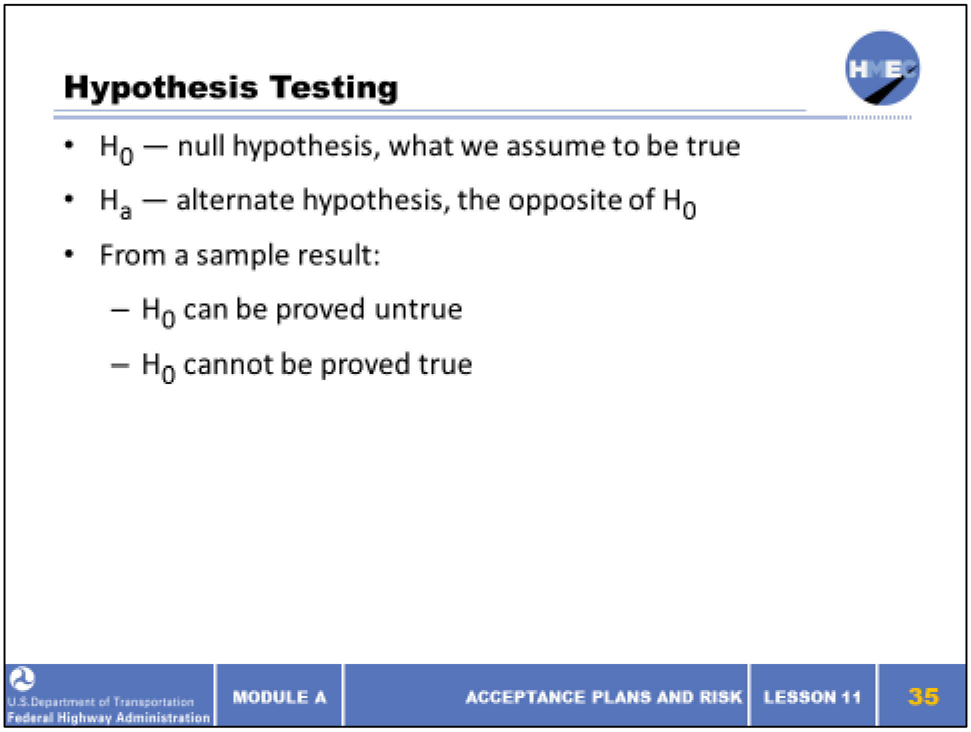
The seller's risk, α, is the probability of rejecting a lot when the lot is really acceptable. The buyer's risk, β, is the probability of accepting a lot when the lot is really unacceptable. One way to help remember which risk is which, β is associated with buyers. Remember a risk is the probability of making the wrong decision.

Slide 34



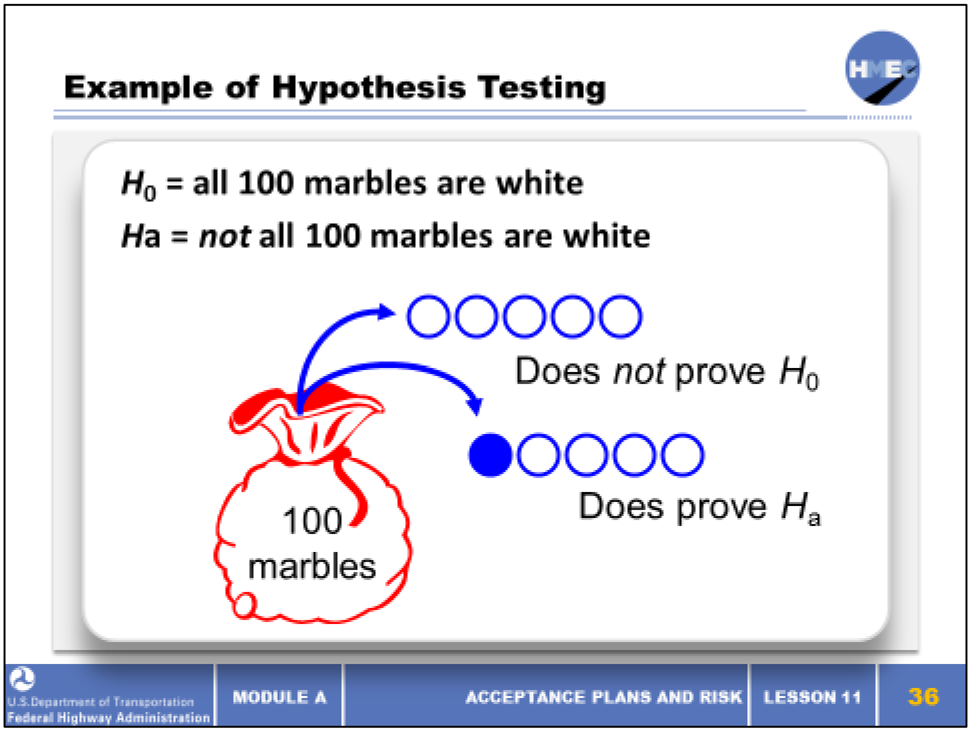
It is often necessary to test whether or not it is reasonable to accept an assumption about a set of data, such as a group of test results. For example, the agency in effect assumes that the contractor's product meets the specification requirements, and then conducts acceptance tests to determine if this assumption should be rejected.

Slide 35



H_0 is assumed to be true, and the hypothesis test is conducted to see if the assumption can be disproved. From sample results, H_0 cannot be proved, only disproved.

Slide 36



Suppose that you bought a cloth bag of 100 white marbles. H_0 would be that all of the marbles in the bag are white, and H_a would be that not all of the marbles are white. If you draw a sample of 5 marbles, it would be possible to prove H_a, but not H_0. If even one of the five marbles in the sample is not white, then H_a is true. However, even if all 5 of the marbles in the sample are white, this does not prove H_0, it merely fails to disprove it. The only way to prove H_0 is to look at all 100 marbles, i.e., the whole population.

____

____

____

____

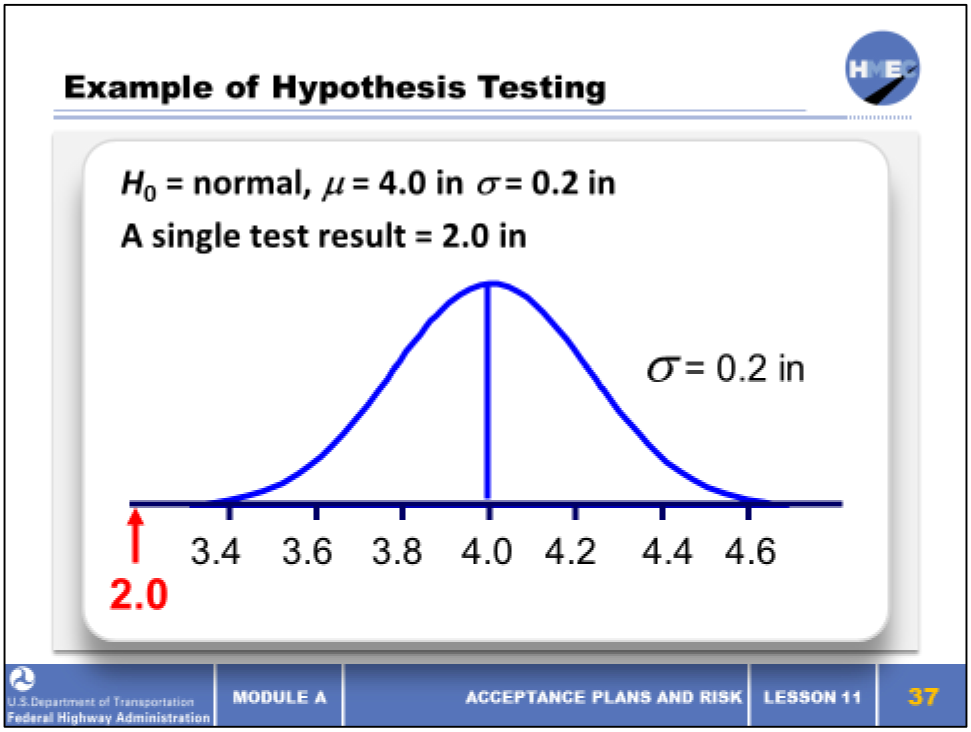
____

____

____

____

Slide 37



Consider the H_0 that a certain paving process is normally distributed with a mean thickness of 4 in, and a standard deviation of 0.2 in. The H_a would be that the process was not normally distributed with μ = 4 in and σ = 0.2 in. To test the hypothesis, a single core is obtained and the pavement thickness is measured to be 2.0 in. From the 68-95-99.7 rule, we know that it is very unlikely for a value to be more than 3σ below the mean. Since 4.0 – (3 × 0.2) = 3.4 in, it is extremely unlikely that a value of 2.0 in would be obtained from a normal distribution with μ = 4.0 in and σ =0.2 in. The H_0 would therefore be rejected.

____

____

____

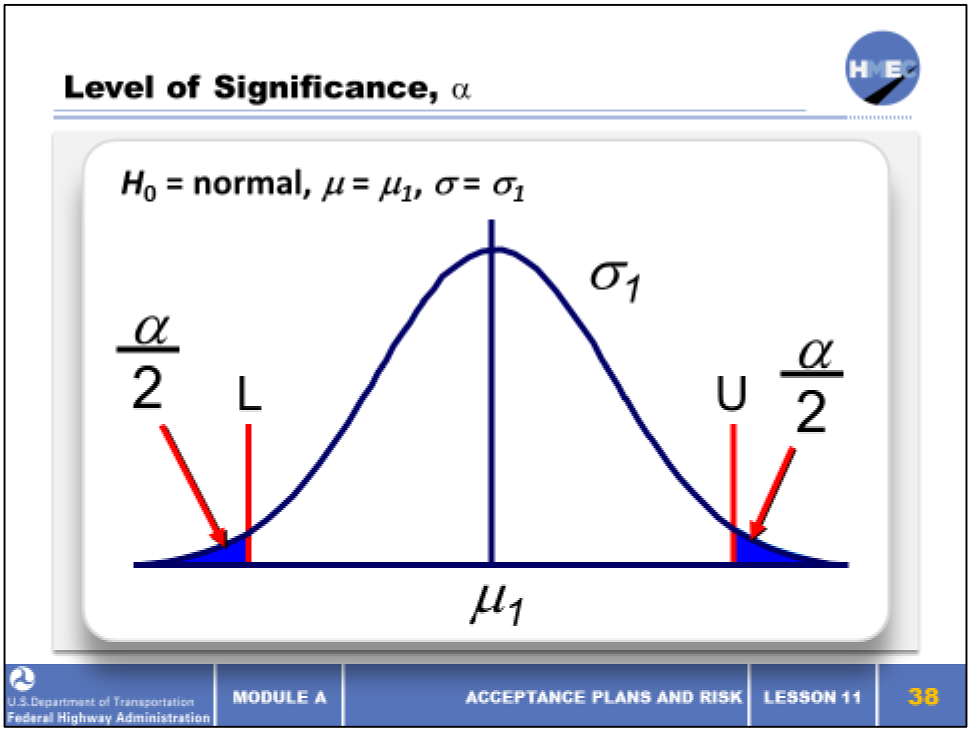
____

____

____

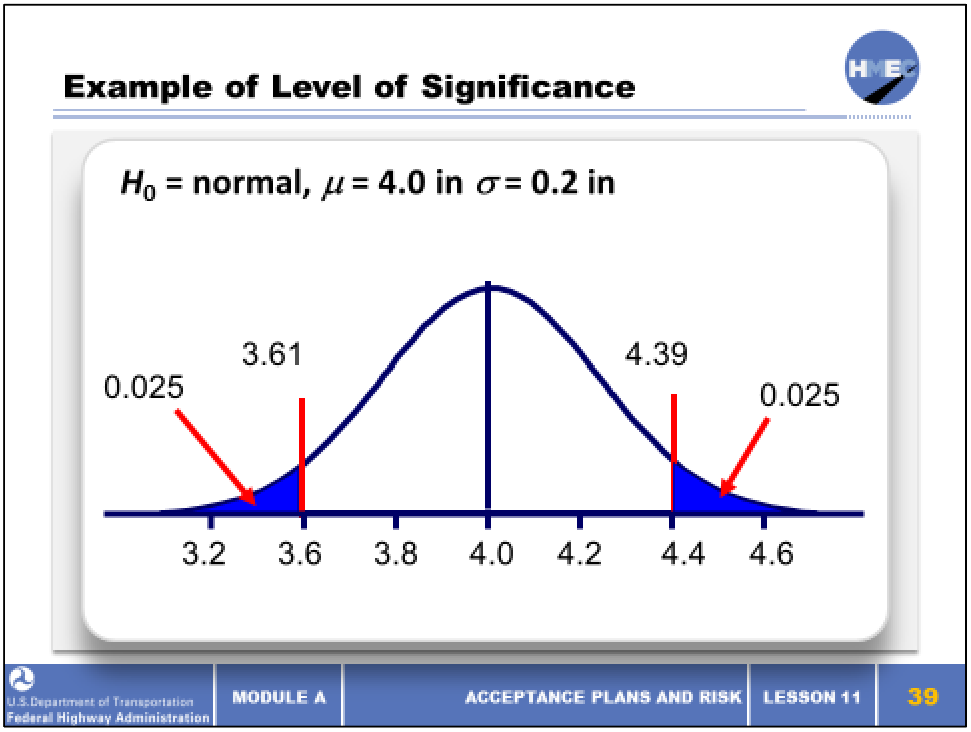
____

Slide 38



In the previous example, the test result was so far from the assumed mean that the decision to reject H_0 was obvious. However, one could ask the question, "How unlikely must the result be to justify rejecting H_0?" This relates to the concept of level of significance, α. Hypothesis tests are conducted at a selected level of significance, α, where α is the probability of incorrectly rejecting H_0 when it is actually true. The value of α is typically selected as 0.10, 0.05 or 0.01. If, for example, $\alpha = 0.01$ is chosen and the null hypothesis is rejected, then there is only 1 chance in 100 that H_0 is true and was rejected in error.

Slide 39



Refer to the Z-Table of areas under standard normal distribution located in the Resources folder of the tablet.

Consider the previous example where H_0 was that the population is normally distributed with μ = 4.0 in and σ = 0.2 in. How far from the assumed μ must a single test be to conclude that H_0 is false? The answer depends upon the value selected for α. Assume that α = 0.05 is selected. For a normal distribution, there is a 0.05 probability of obtaining a value that is more than $\pm 1.96\sigma$ from the population mean. For our example, $\pm$ (1.96 $\times$ 0.2 in) = approximately $\pm$ 0.39 in. So, if we decide to reject the H_0 if the single test is more than 0.39 in above or below 4.0, there is only a 0.05 probability (5% chance) that we will reject H_0 when it is actually true.

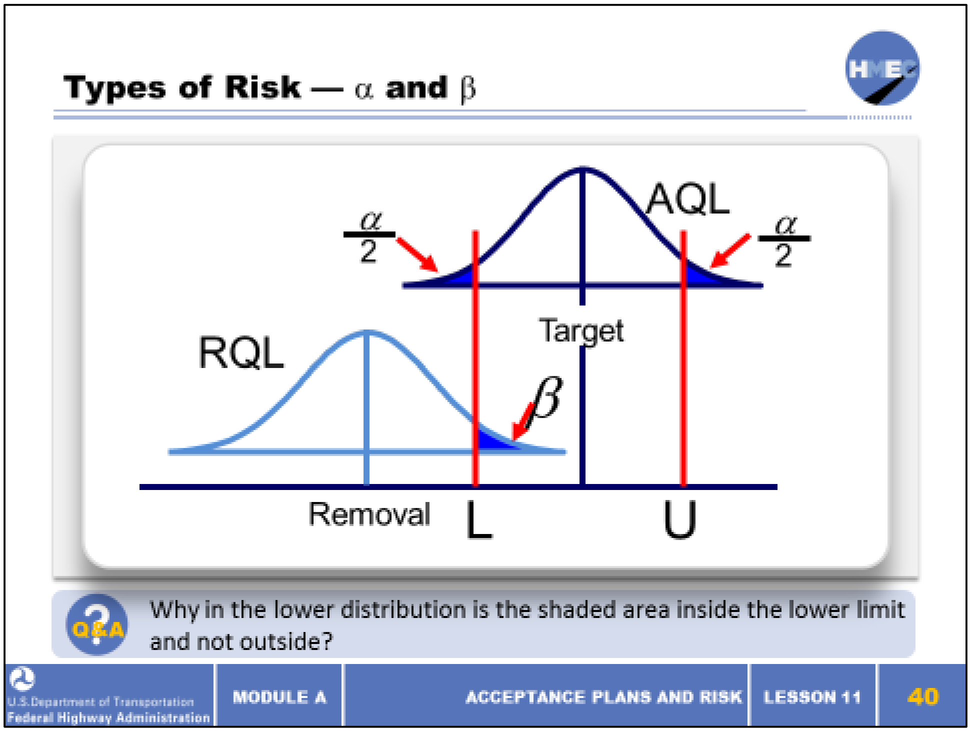
The 68-95-99.7 rule uses 2 σ' to approximate 95%, but the Z-table uses 1.96 as the more exact value for 95%. So 4.0 $\pm$ 0.39 determines the limits to be 3.61 in and 4.39 in.

____

____

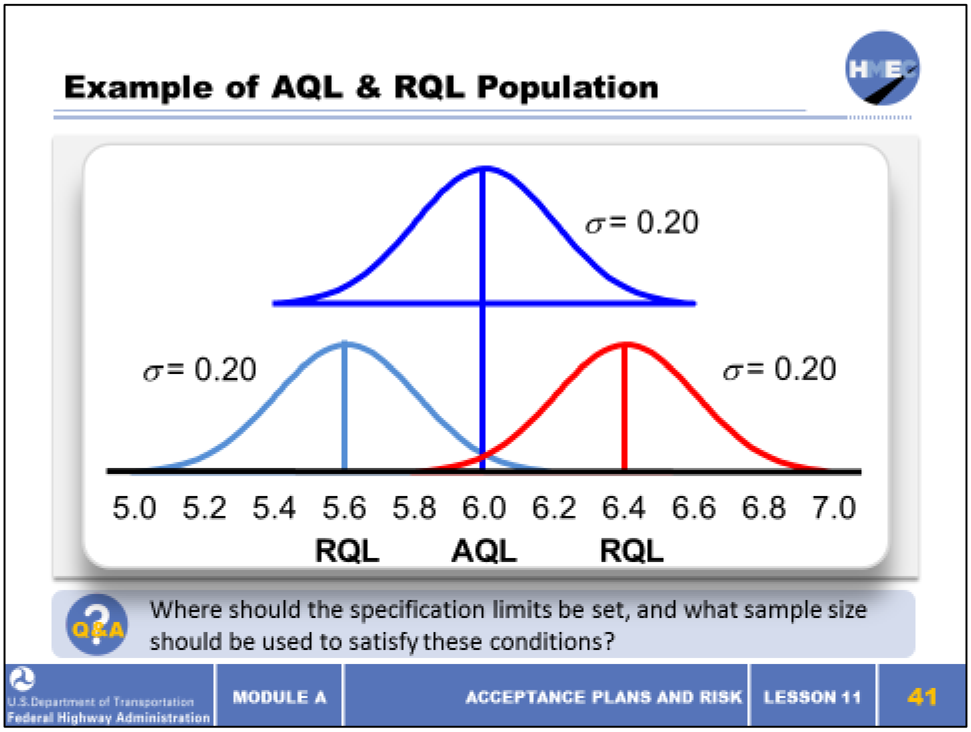
____

Slide 40



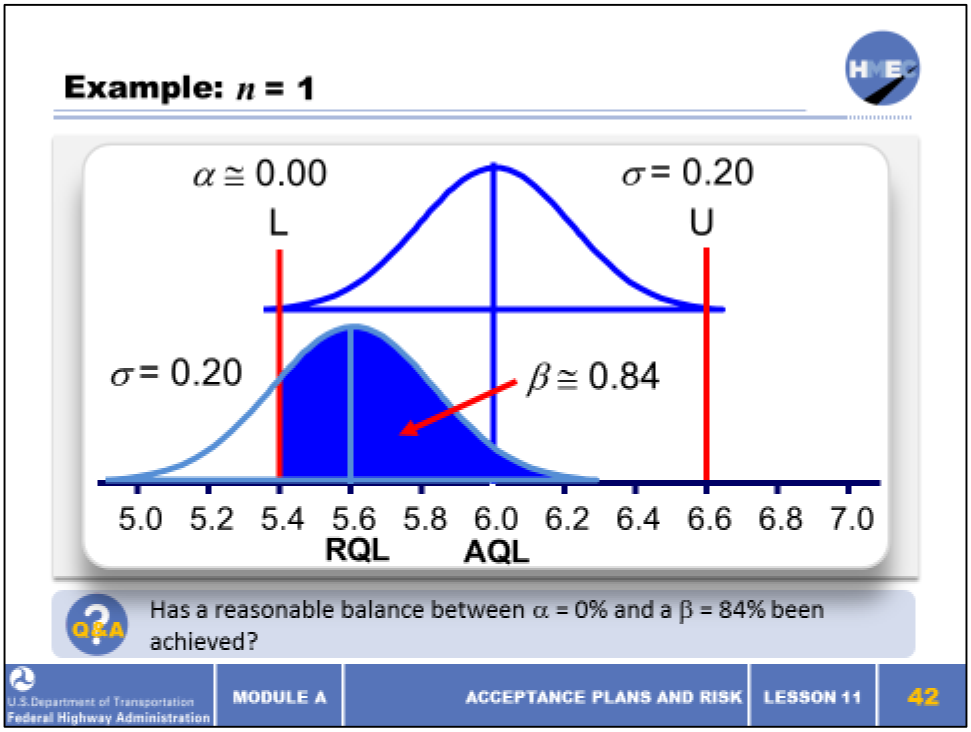
The two types of risk are illustrated on the slide. The normal distribution on the top represents a lot that is acceptable (the mean is at the target), while the normal distribution on the bottom represents a lot that is unacceptable (the mean is so far from the target as to require action). If acceptance decisions are based on the results of a single test, then the α risk, the risk that an AQL lot will be rejected, is the probability that a single test result from the AQL lot will be outside of the specification limits. This probability is represented by the sum of the two shaded areas in the top distribution. The β risk, the risk that an RQL lot will be accepted, is the probability that a single test result from the RQL lot will be within the specification limits. The shaded region in the bottom distribution represents this probability.

Slide 41



If possible, we would like to always accept a lot that is acceptable, but we do not want to accept a lot defined as unacceptable. To demonstrate the concept of α and β risks, we previously defined an AQL lot as a μ = 6.0% and σ = 0.20%, and defined an RQL lot as $\mu \leq$ 5.6% or $\mu \geq$ 6.4% with a σ = 0.20%.

Slide 42



Remember that n=1.

We also want to start with the contractor having no risk of acceptable product being rejected. To ensure essentially a 100% chance of accepting our definition of an AQL lot, the limits would be set at $\pm 3\sigma$, or 5.4% and 6.6%. Thus, α is essentially 0%, which is one of our objectives. However, β, the probability of accepting an RQL lot (i.e., one with $\mu \leq$ 5.6% or $\mu \geq$ 6.4%) would be 84% (see table of areas under normal curve). It is obvious that there is only a 16% chance that an RQL lot would correctly be rejected (i.e., 84% probability of incorrectly being accepted). The slide shows the probability of accepting a lot with an average of 5.6%. The same probability exists if the average is 6.4%.

____

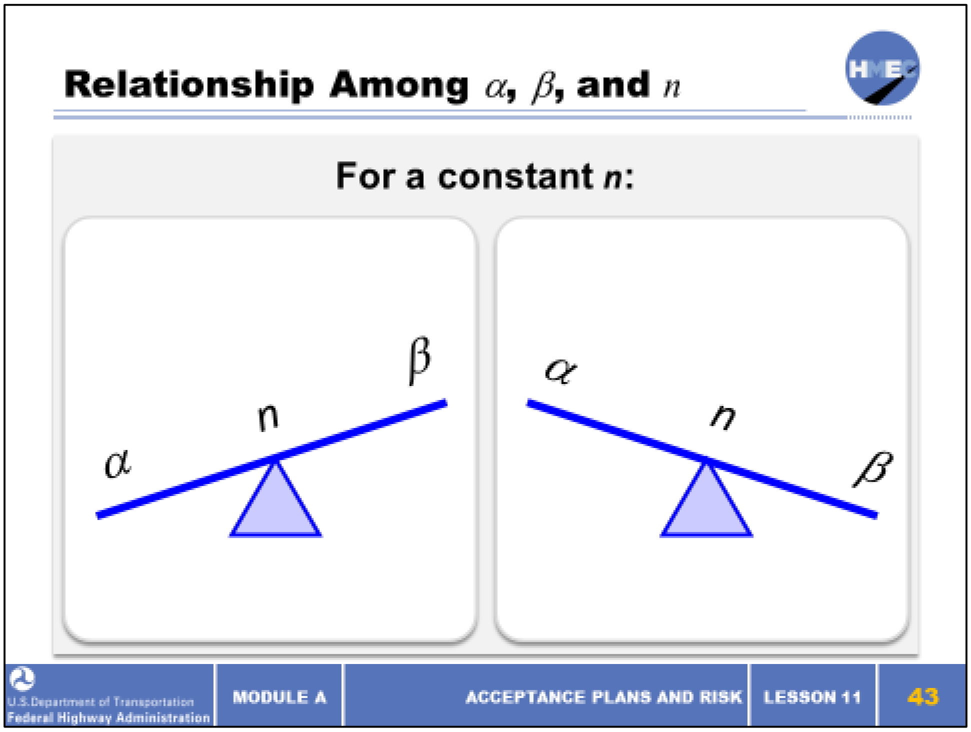
____

____

____

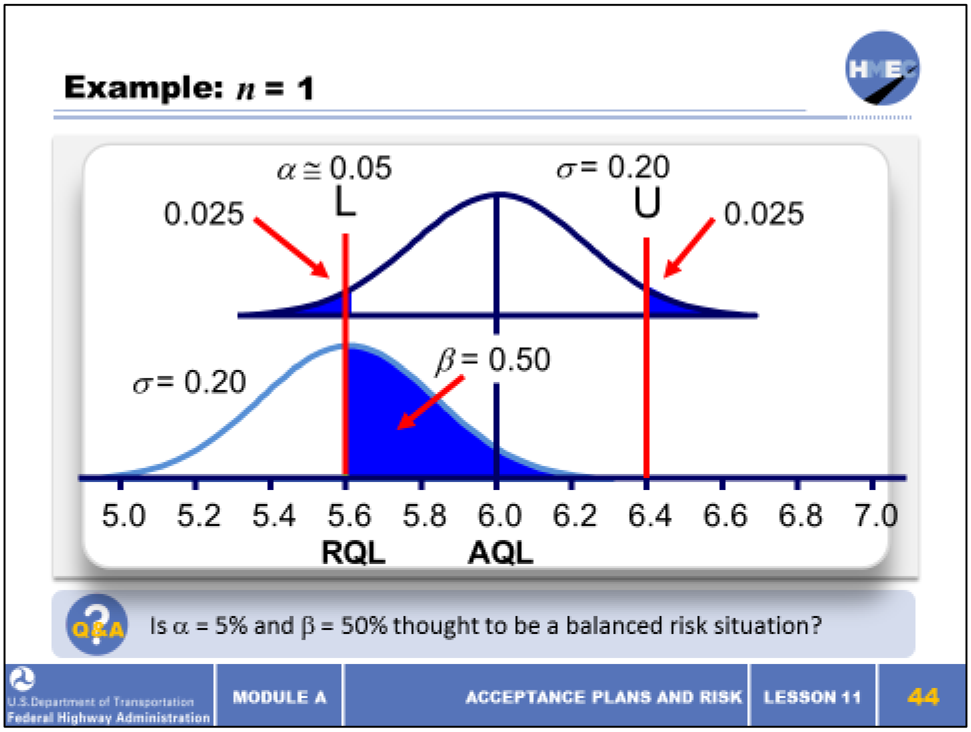
____

Slide 43



For a constant *n*, if α goes up, β goes down and vice versa. So in this case if it is desirable to reduce β, the simplest way is to increase α. And this is done by tightening the spec limits.

Slide 44



Suppose the ±3σ limits used in the first example are tightened to ±2σ limits. What does this do to α and β? With ±2σ limits, α has been increased to approximately 5% (remember the 68-95-99.7 rule) and β has been reduced to 50%. This still is likely to be considered too high a risk of accepting an unacceptable lot. Tightening the limits further to increase α and decrease β is an option, but not a very realistic one. This essentially does what we do not want to do, as it rejects more acceptable lots. Tightening limits beyond those that are considered practical is not likely to be cost effective.

____

____

____

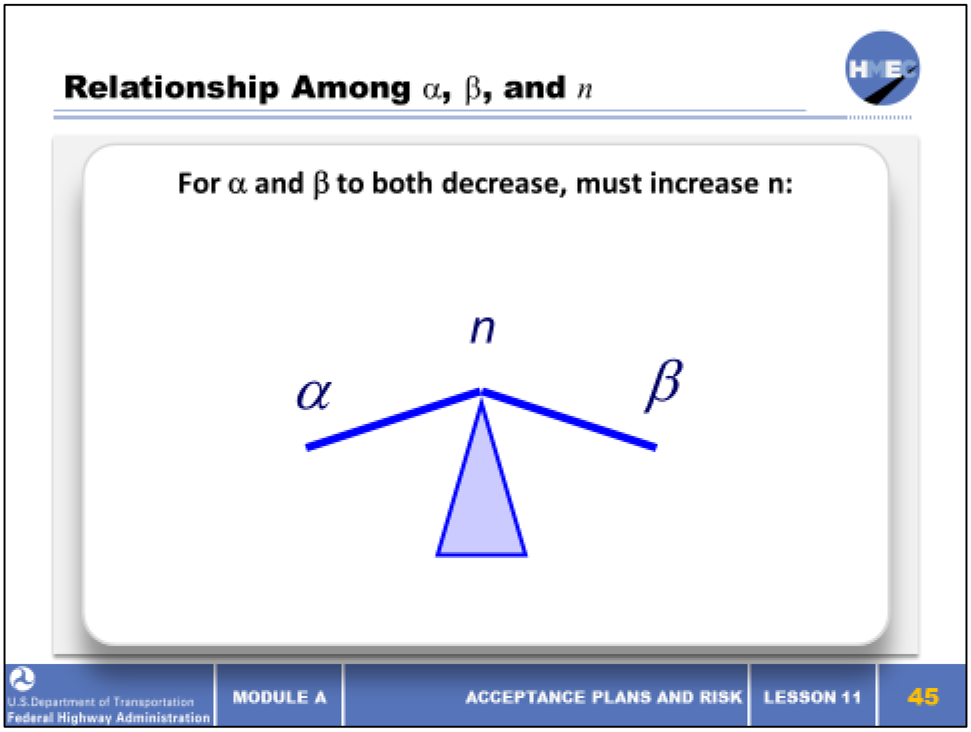
____

____

____

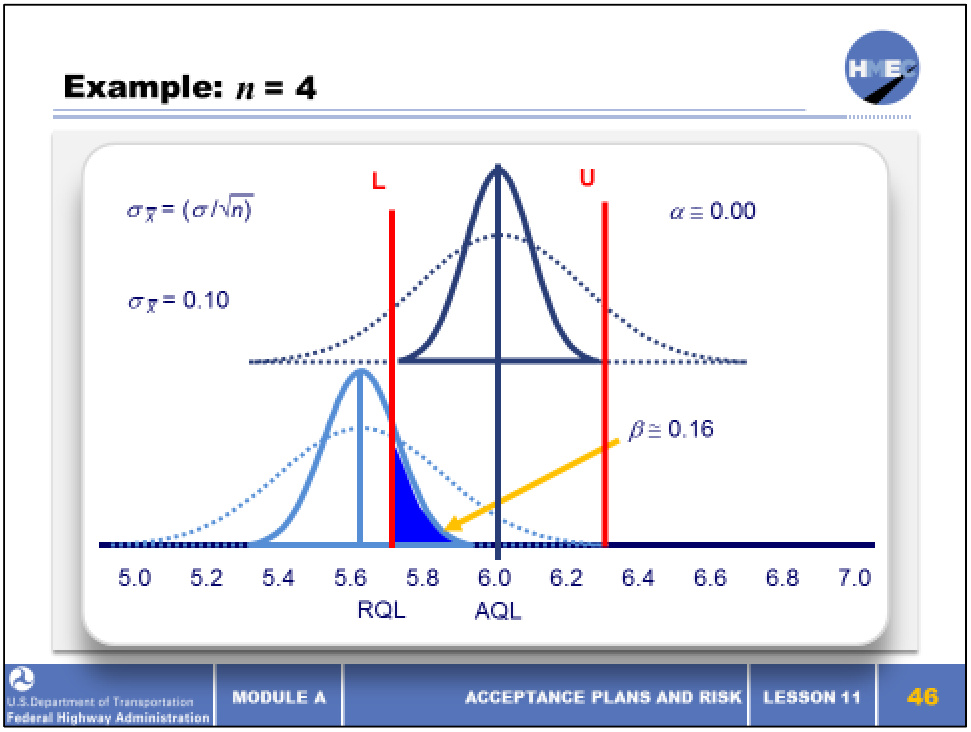
____

Slide 45



If we want to decrease both α and β, we must increase n.

Slide 46



Since this try uses acceptance based on averages, we must use the concept of the distribution of means. For our example, the distribution of means will have a $\sigma_{\text{X-bar}} = \sigma$/square root n which is equal to σ *(*0.20*)*/square root 4 = 0.20/2*;* thus*,* $\sigma_{\text{X-bar}}$ = 0.10. For α = 0% and based acceptance on n = 4, the limits become $\mu \pm 3\sigma_{\text{X-bar}}$ = 6.0 ± 3(0.10) = 6.0 ± .30 , or 5.7 and 6.3, as shown in the slide. By increasing the sample size to n = 4 and setting α = 0%, β is reduced to about 16%.

____

____

____

____

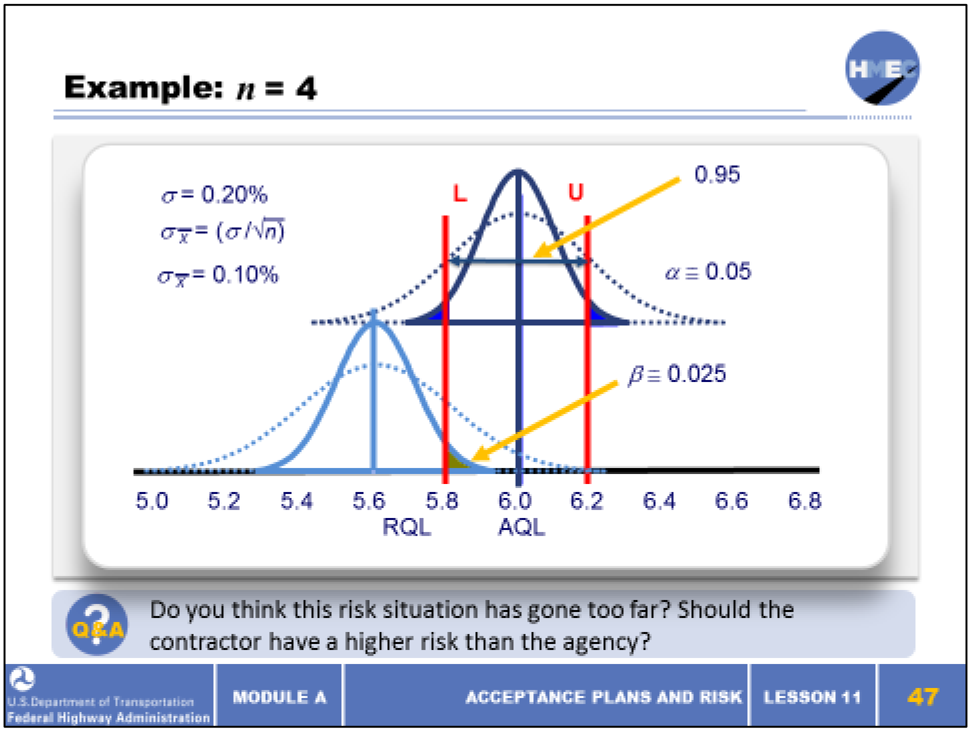
____

____

____

____

Slide 47



By increasing α to 5%, thereby using 2 $\sigma_{\textit{X-bar}}$ limits, and by using *n* = 4, the limits now become 5.80% and 6.20% as shown. Under these conditions, β is only 2.5%, very likely an acceptable level for risk of accepting an unacceptable lot. Calculate β: Z = (LSL – μ)/$\sigma_{\text{X-bar}}$ = (5.8 – 5.6)/ 0.10 = 2. 47.5% of the population is between 5.6 and 5.8. Thus, β = 50-47.5 = 2.5%. At this point, β is less than α and is not a particularly good situation. Contractors should not have a higher risk than the agency.

____

____

____

____

____

____

____

____

Slide 48

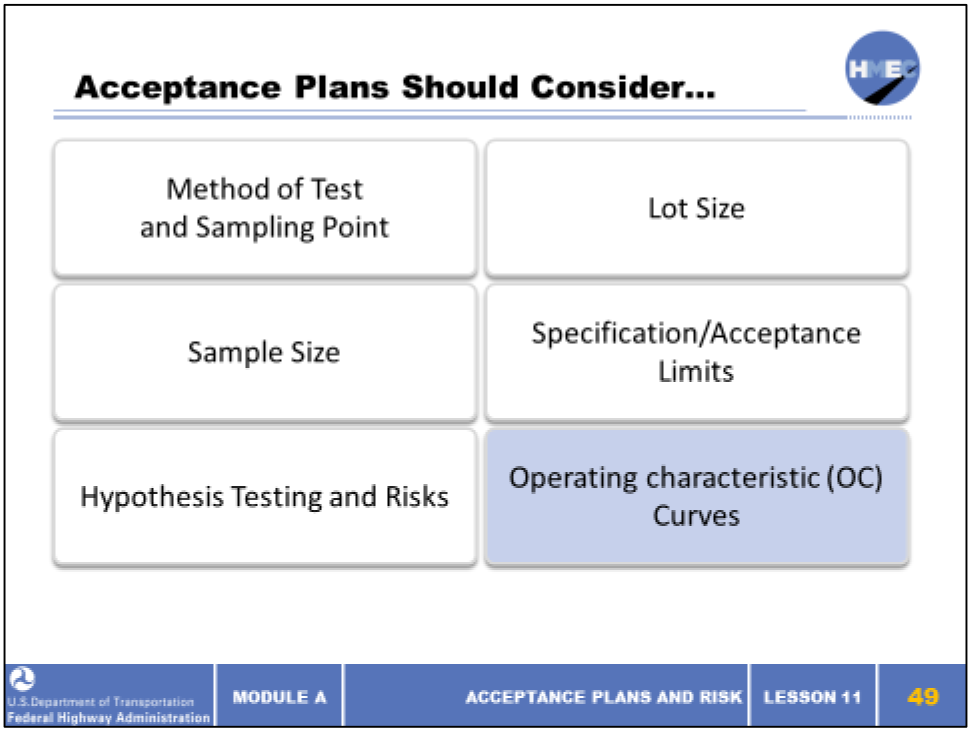
Summary of β Changed as α and n Changed

	Seller's Risk	Buyer's Risk
n	α	β
1	0%	84%
1	5%	50%
4	0%	16%
4	5%	2.5%

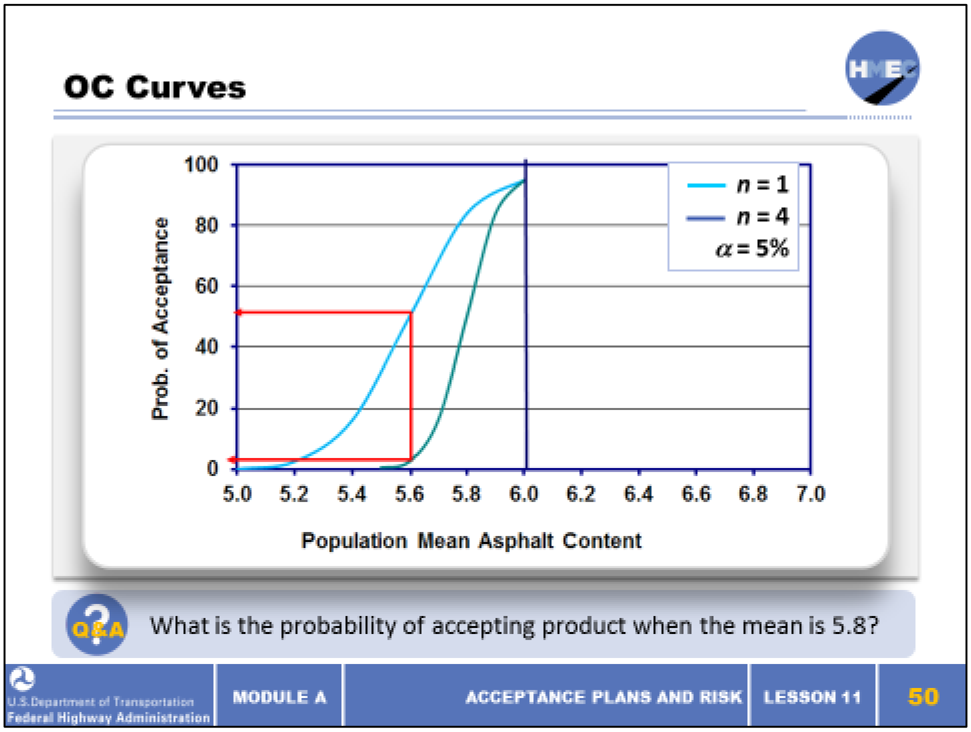
U.S. Department of Transportation Federal Highway Administration | MODULE A | ACCEPTANCE PLANS AND RISK | LESSON 11 | 48

Summarizing the changes that were found as we went through the process will aid us in the next step, which is a graphical representation between the actual quality and the probability of acceptance called an operating characteristic (OC) curve.

Slide 49


Acceptance Plans Should Consider…
Method of Test and Sampling Point
Lot Size
Sample Size
Specification/Acceptance Limits
Hypothesis Testing and Risks
Operating characteristic (OC) Curves
U.S. Department of Transportation
Federal Highway Administration
MODULE A
ACCEPTANCE PLANS AND RISK
LESSON 11
49

Slide 50



In the previous slide we saw the determination of α and β calculations for lots at the AQL and RQL levels. Similar values for the probability of accepting lots with other population means can also be calculated. The relationship between the lot mean and the probability that the lot will be accepted can be illustrated by the use of operating characteristic (OC) curves. For a given standard deviation and α value, and various sample sizes, a set of curves can be drawn to characterize the probability of accepting lots with various population means. On the OC curve, α is 100% minus the probability of accepting a lot with a mean at the AQL. And β is indicated as the probability of accepting a lot with a mean at the RQL. In this slide, α was chosen as 5% (0.05) so that when a lot has a mean of 6.0% (i.e., at the AQL), the probability of acceptance is 95%. When a lot has a mean of 5.6 % (i.e., at the RQL), β = 50% for *n* = 1 and 2.5% for *n* = 4. Obviously, OC curves can be extremely useful when developing a specification to verify that the risks are reasonably balanced.

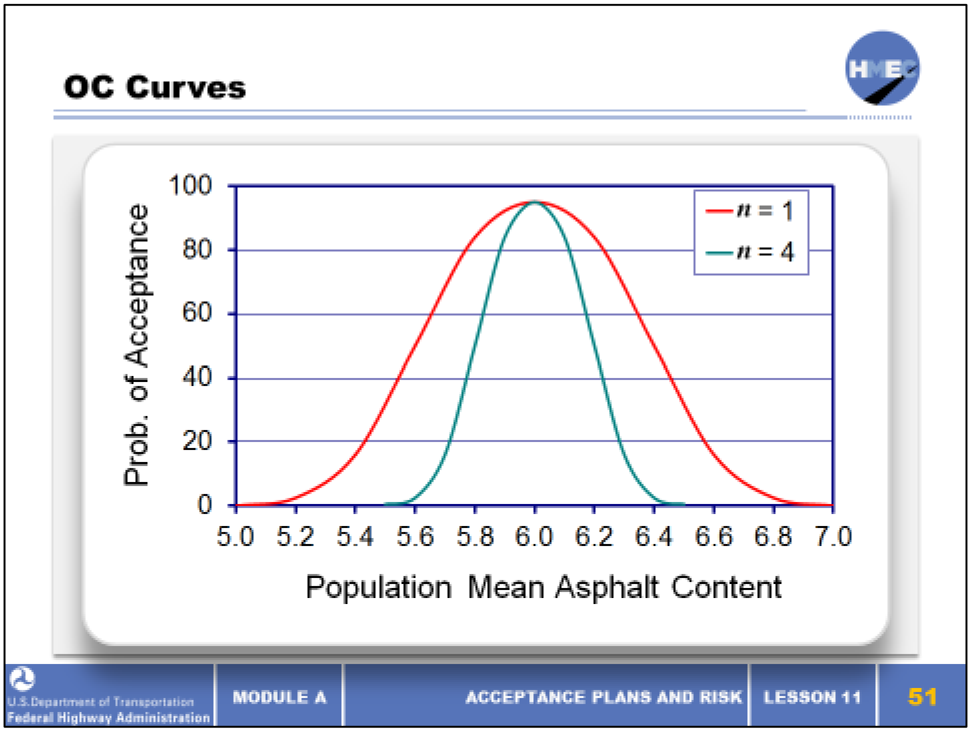
____

____

____

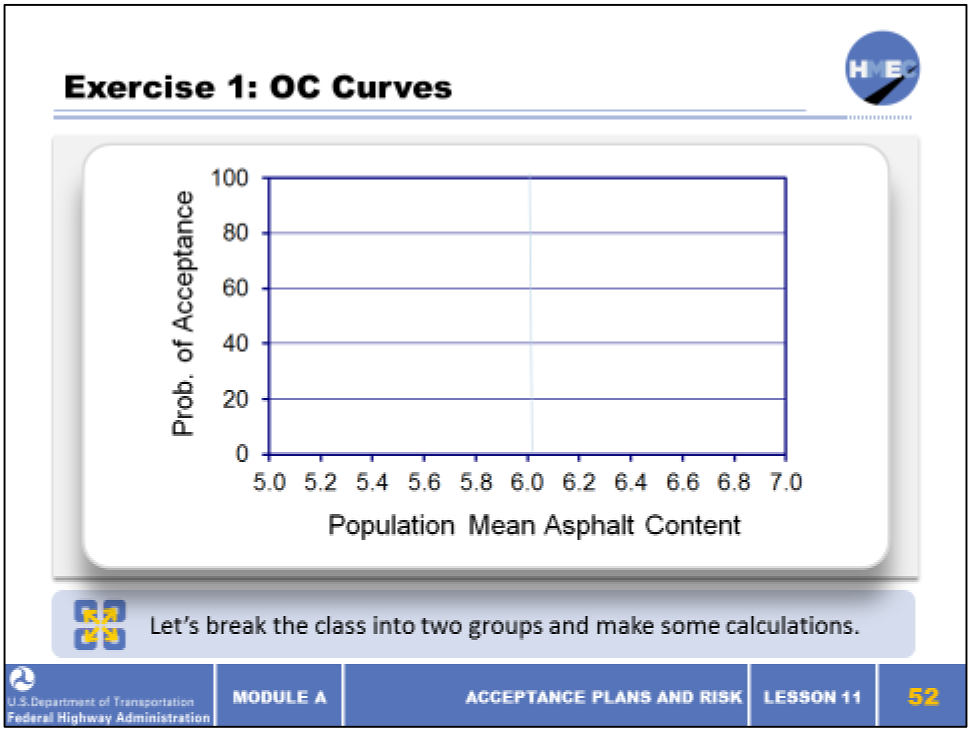
____

Slide 51



The calculations for this example were done for population mean asphalt contents on the low side of the target value of 6.0%. Similar calculations could be made for population means on the high side of the target. The complete OC curves for α = 5% (0.05) and n =1 and n = 4 are shown here.

Slide 52



The purpose of this activity is to calculate and plot points on the OC curve.

Refer to Figures 10-1 and 10-2 OC Curves for Paired t-test. This document is located in the Resources folder on the tablet.

____

____

____

____

____

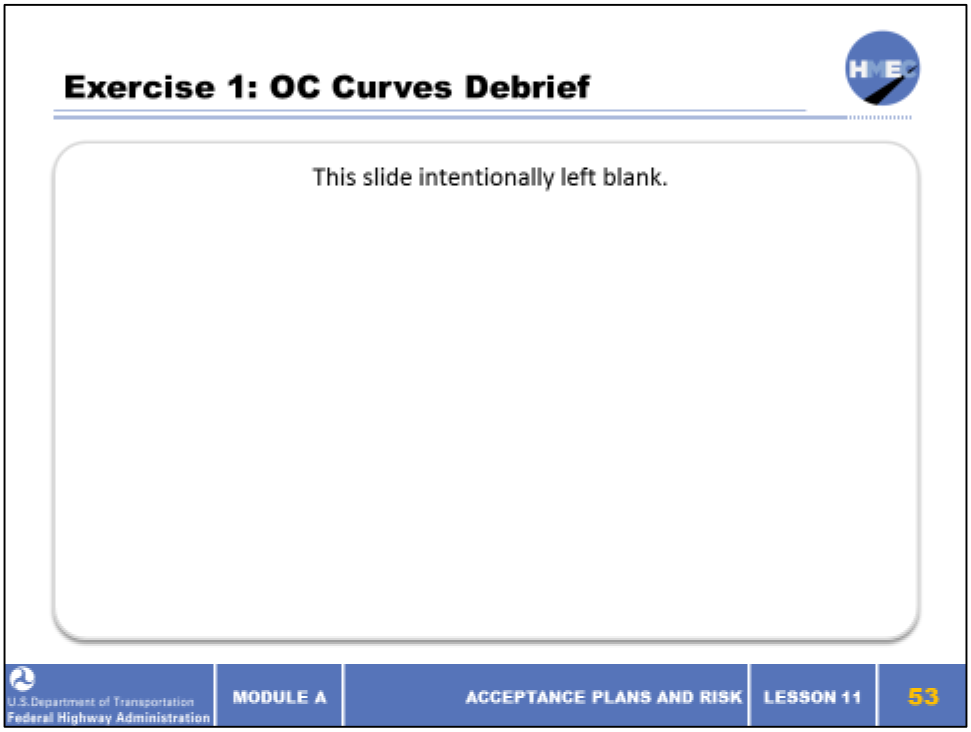
____

____

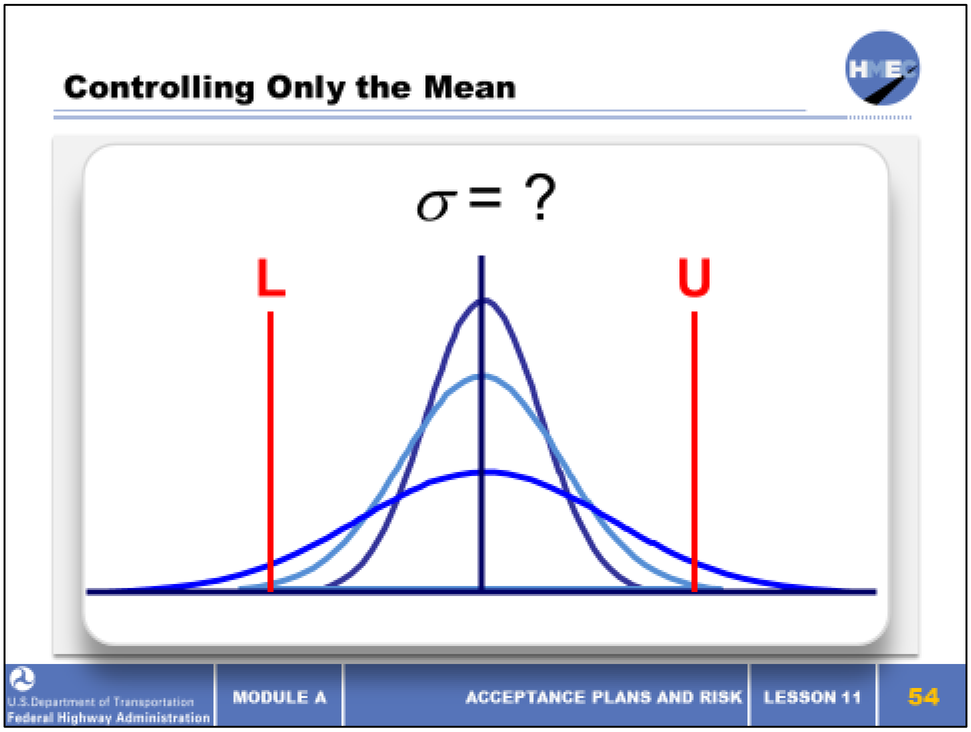
____

____

Slide 53


Exercise 1: OC Curves Debrief
This slide intentionally left blank.
U.S. Department of Transportation
Federal Highway Administration
MODULE A
ACCEPTANCE PLANS AND RISK
LESSON 11
53

Slide 54



Although it can be used for establishing limits for accepting the sample mean of a lot, it has a serious disadvantage. The standard deviation is assumed to be known and constant. If these acceptance limits are used, and the contractor has a standard deviation other than 0.2%, the risks are no longer those for which the specification was designed. One solution to this problem is to establish a tolerance on the variability and develop a specification to accept the standard deviation separately from the mean. However, another approach that uses an estimate of the percentage of the lot that is within the specification limits to determine acceptability is generally accepted as the best procedure for considering both mean and standard deviation.

____

____

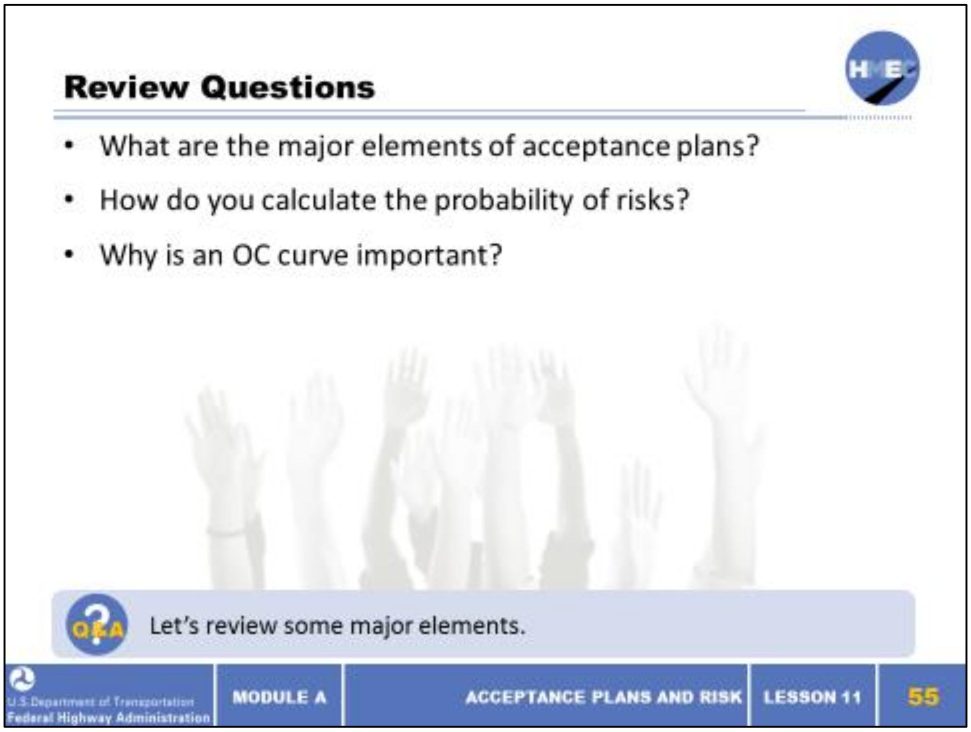
____

____

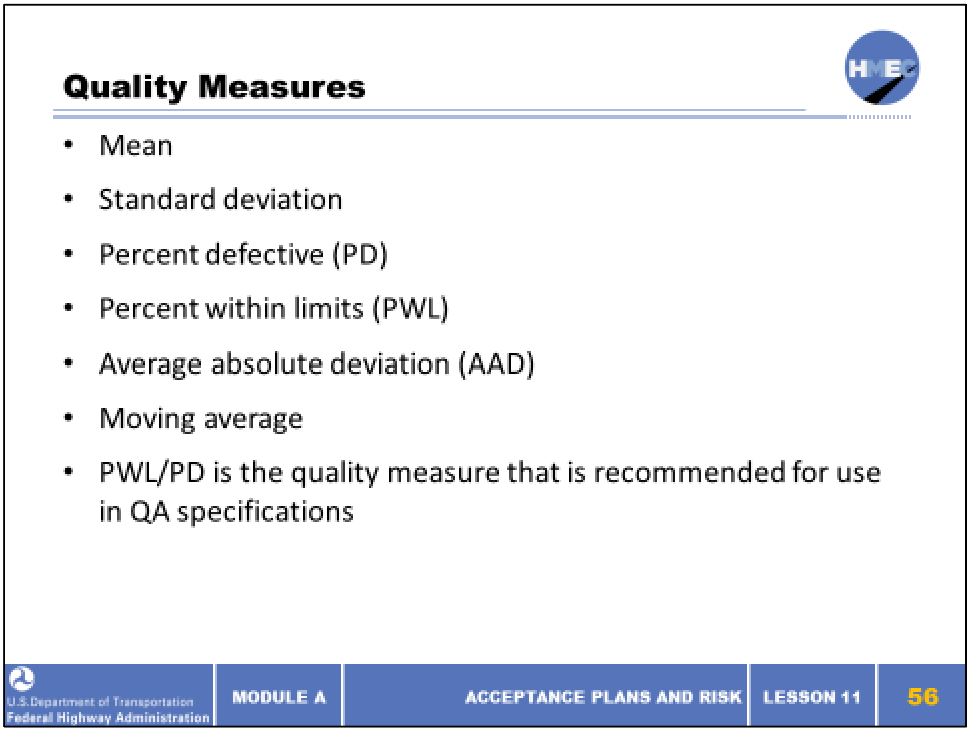
____

____

Slide 55

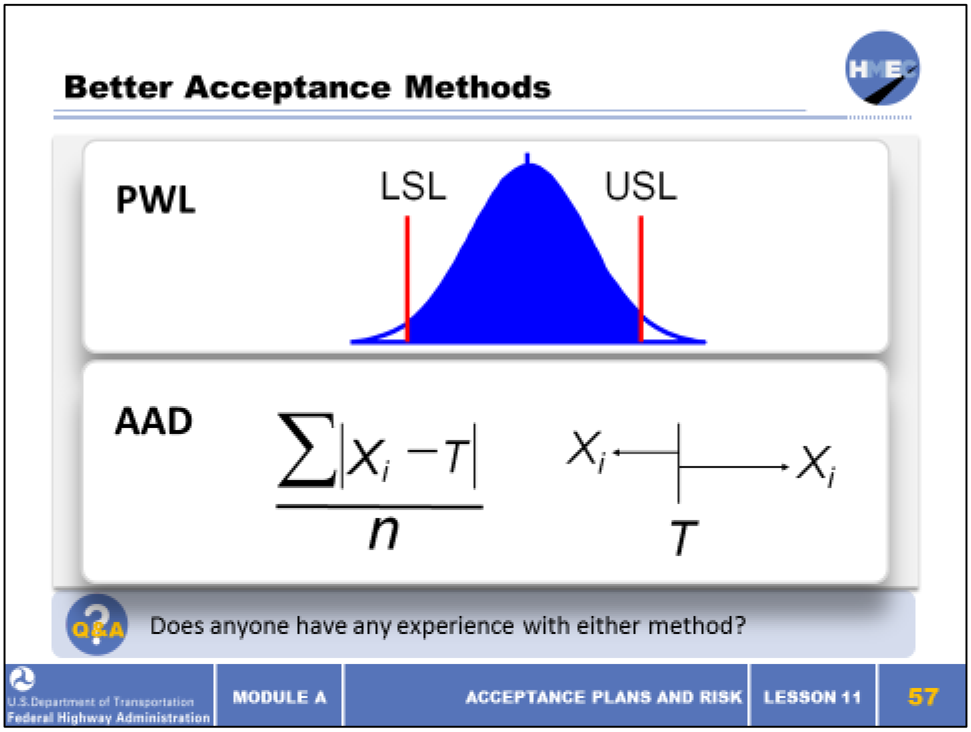

Review Questions
• What are the major elements of acceptance plans?
• How do you calculate the probability of risks?
• Why is an OC curve important?
Q&A
Let's review some major elements.
U.S. Department of Transportation
Federal Highway Administration
MODULE A
ACCEPTANCE PLANS AND RISK
LESSON 11
55

Slide 56



The Transportation Research Circular Glossary of Transportation Quality Assurance Terms defines quality measure as: "Any one of several mathematical tools that are used to quantify the level of quality of an individual quality characteristic." Typical quality measures used in QA are selected because they quantify the average quality, the variability, or both. Examples of quality measures that may be used include mean, standard deviation, percent defective (PD), percent within limits (PWL), average absolute deviation (AAD), and moving average. PWL or PD is the quality measure that is recommended for use in QA specifications.

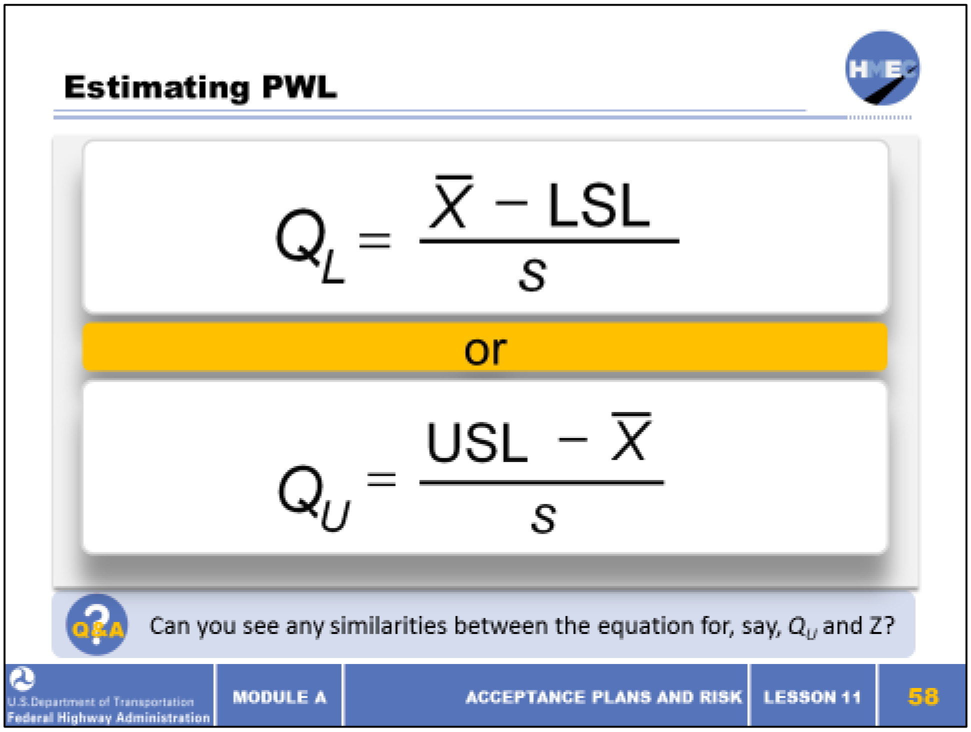
Slide 57



The first is called the percent within limits (PWL) method, and is similar in concept to determining the area under the normal curve. It uses the sample mean and the sample standard deviation to estimate the percentage of the population (lot) that is within the specification limits. The other is called the average absolute deviation (AAD) method and is a measure of the ability of a contractor to consistently hit the target value. It uses the average of the absolute difference between the target value and the test value.

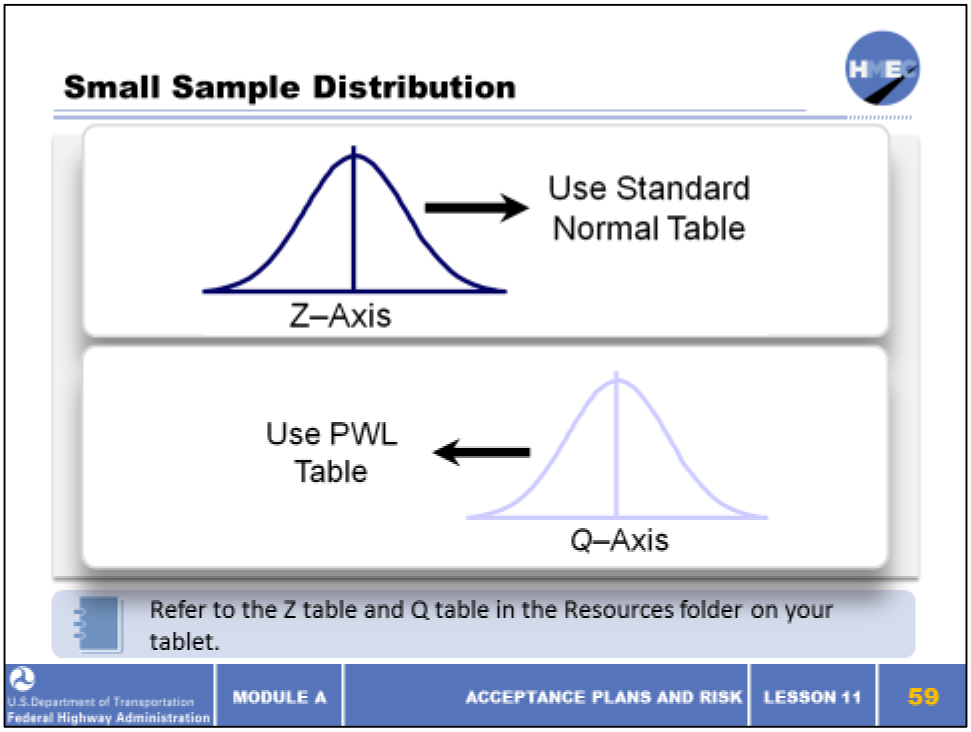
____

Slide 58



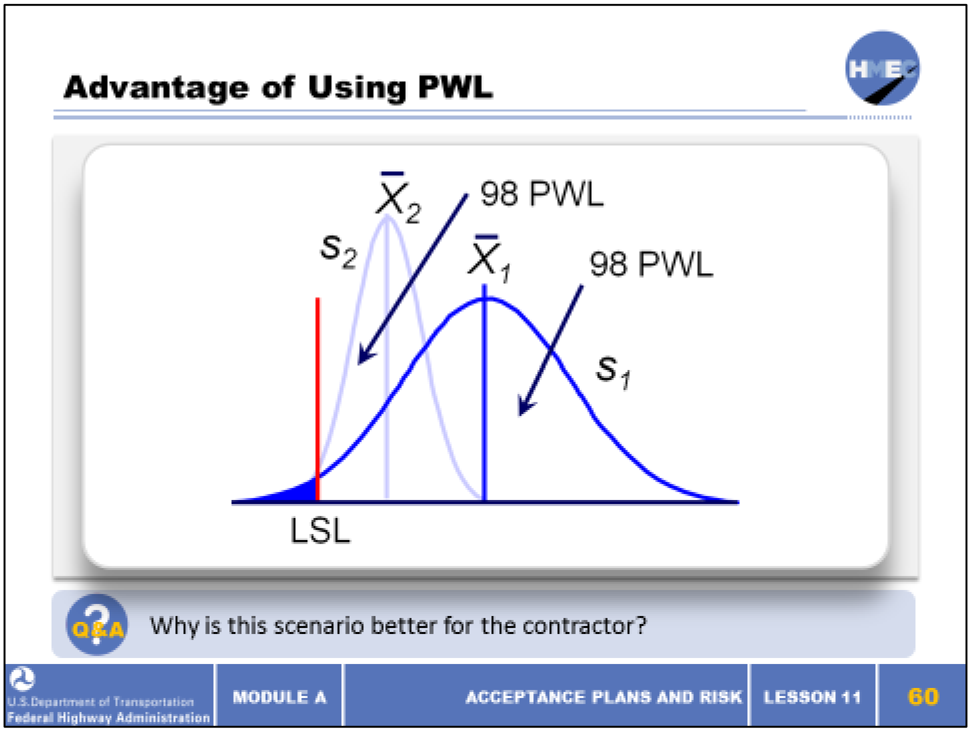
The area under the normal curve can be calculated (as shown in the WBT) to determine the percentage of the population that is within certain limits. Similarly, the percentage of the lot that is within the specification limits can be estimated. Instead of using the Z-value and the standard normal curve, a similar statistic, the quality index (*Q*), is used to estimate PWL.

Slide 59



Q is used instead of Z. The *Q* table is an abbreviated table that relates the values of *Q* and PWL for various sample sizes, *n*.

Slide 60



The contractor having the s_1 standard deviation must have a mean farther from the LSL than a contractor with the smaller s_2 standard deviation. As the standard deviation of the lot decreases, the lot mean can approach the specification limit and the lot can still be acceptable. This is often an incentive for a contractor to produce a more uniform product. Another way to look at a PWL specification is that the contractor's process variability (standard deviation) determines the required location of the process mean. This puts the contractor in control of his/her product and they must reconcile their variability with where their mean is located.

____

____

____

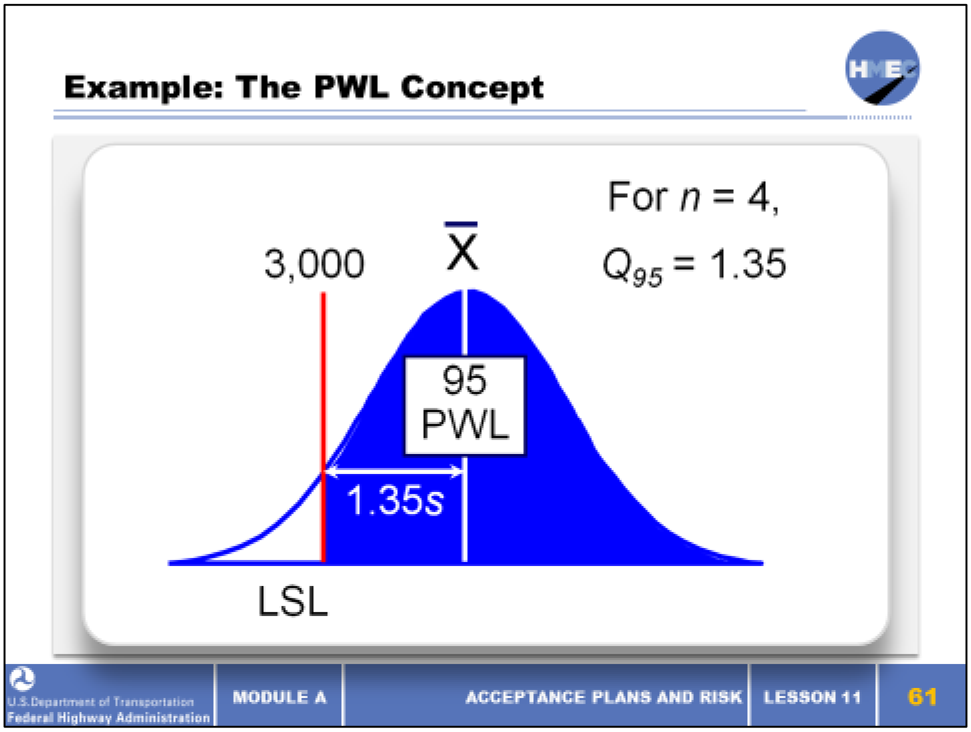
____

____

____

____

Slide 61



In this example, the lower specification limit (LSL) for concrete strength is 3,000 pounds per square inch (psi). One requirement of the PWL procedure is that the sample size must be greater than $n = 1$ since the sample standard deviation is necessary to estimate PWL. For this specification, the sample size has been chosen as $n = 4$. Furthermore, the specification requires that at least 95% of the lot exceed the minimum strength (i.e., PWL $\geq$ 95). The PWL table shows that the minimum Q value is 1.35 for 95 PWL and a sample size of $n = 4$. Whenever the mean is 1.35s above the specification limit, the lot is accepted. However, most frequently the Q value will be used to compute the PWL and that will, in turn, be used to determine a pay factor.

____

____

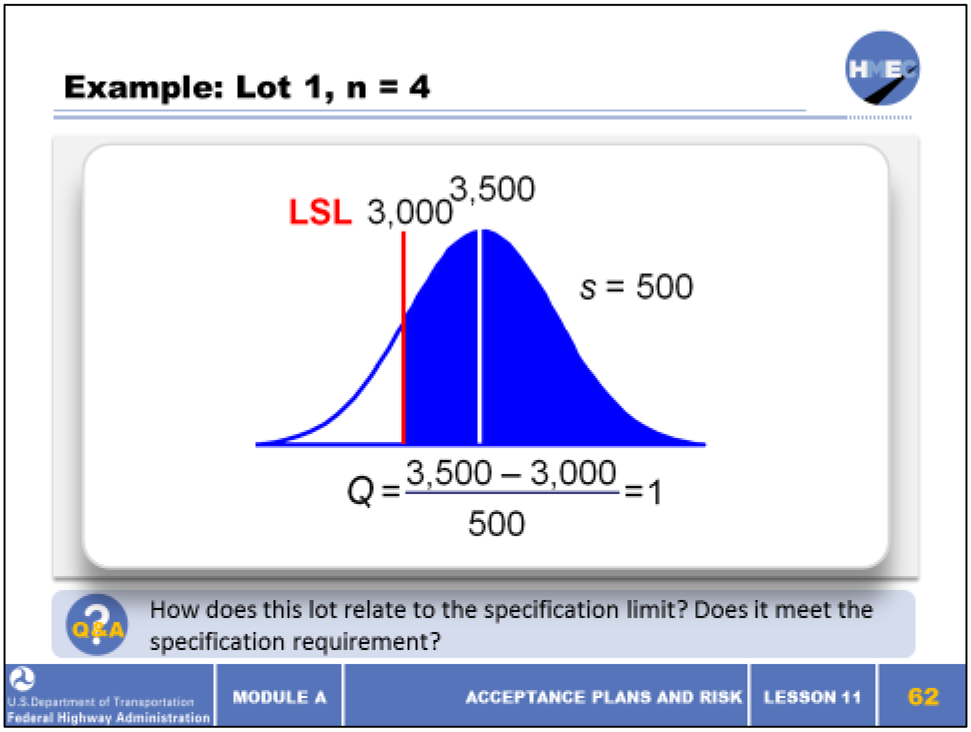
____

____

____

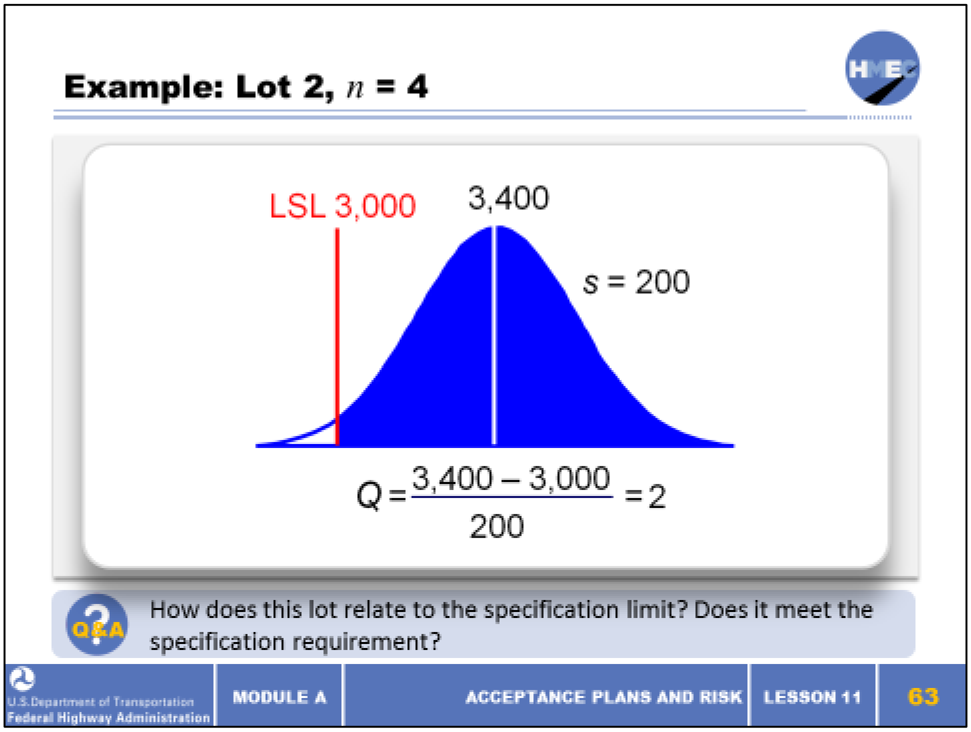
____

Slide 62



We want to calculate the *Q* value using the sample statistics on the slide. *Q* = (LSL – X-bar)/s = (3,500-3,000)/500 = 1.0.

Slide 63



We want to calculate the *Q* value using the sample statistics on the slide. *Q* = (LSL – X-bar)/s = (3,400-3,000)/200 = 2.0.

____

____

____

____

____

____

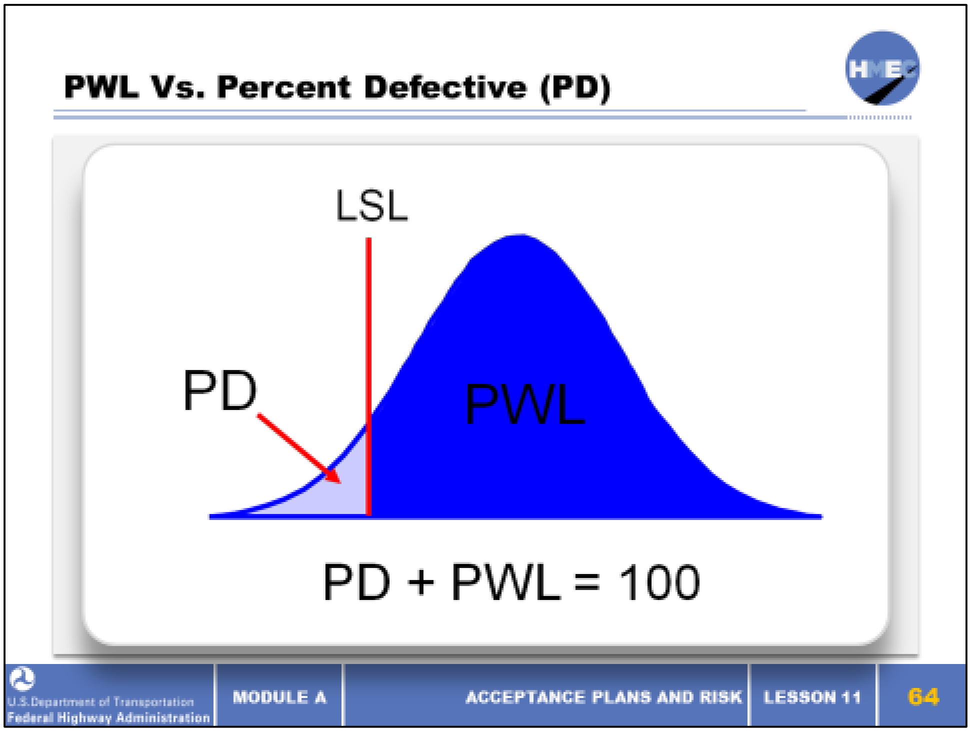
____

____

____

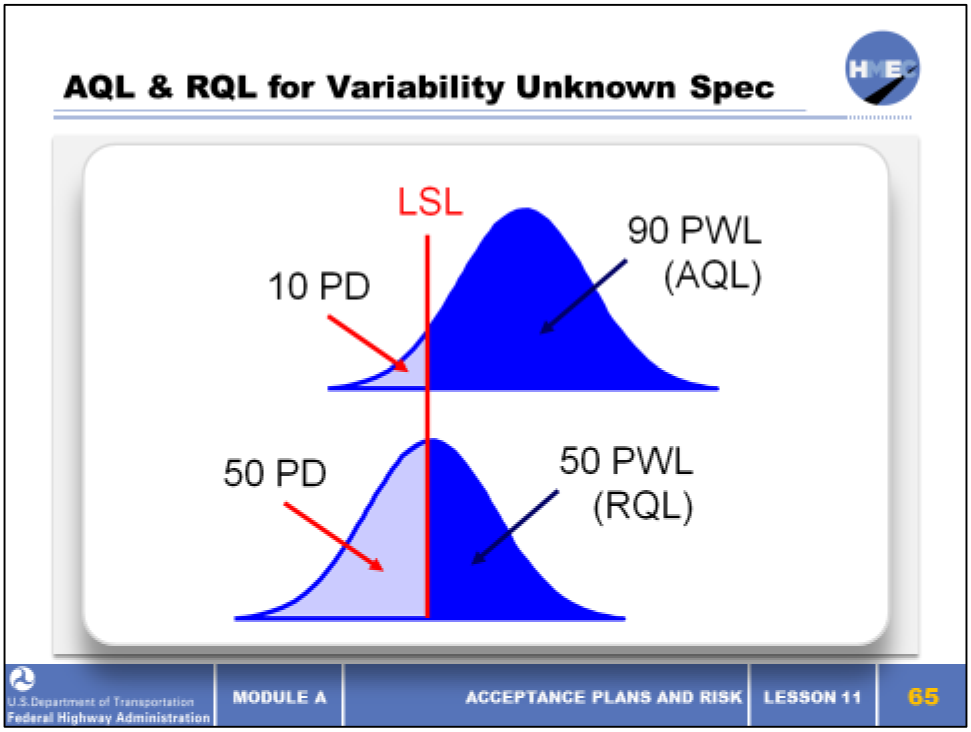
____

Slide 64



The PWL approach is based on Military Standard 414 that was published in 1958. That standard is based on the use of percent defective (PD) rather than percent within limits (PWL). Conceptually, they both estimate percentages of the population related to the specification limits. More agencies have chosen to base acceptance on PWL than PD.

Slide 65



Just as acceptable and unacceptable lots were defined in the acceptance plans based on a variability known specification, acceptable quality level (AQL) and rejectable quality level (RQL) must be defined for PWL or PD acceptance plans. The AQL is the minimum level of PWL at which the product is considered fully acceptable. The RQL is the maximum level of PWL at which the product is considered unacceptable (rejectable). When using PWL, the AQL and RQL both involve areas, not single points such as using only the mean.

AASHTO Quality Assurance Guide Specifications (1995) suggests that the AQL for paving materials be set at 90 PWL. The task force that wrote this document decided the definition of the RQL should be left up to each agency. This will be discussed in more detail later.

____

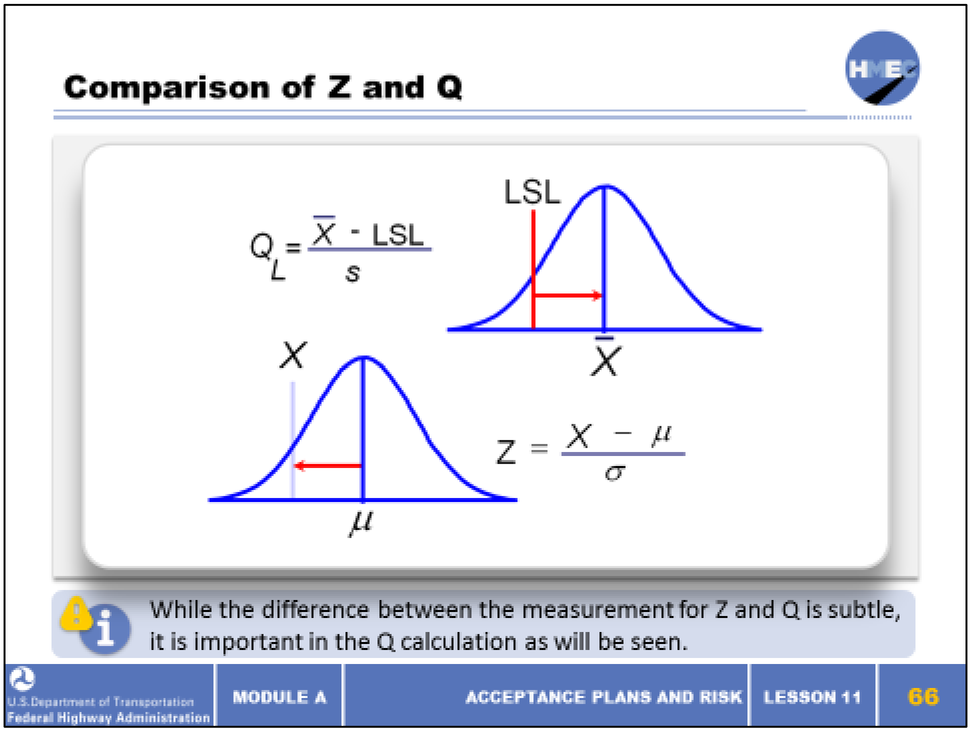
____

____

____

____

Slide 66



It was also mentioned that it is important to keep in mind the conceptual relationship between Z, for the standard normal distribution, and *Q*, for small sample sizes. What was not emphasized is that while Z is measured as the number of standard deviation units from the population mean, *Q* is a measure of the number of sample standard deviations that the sample mean is away from the specification limit.

____

____

____

____

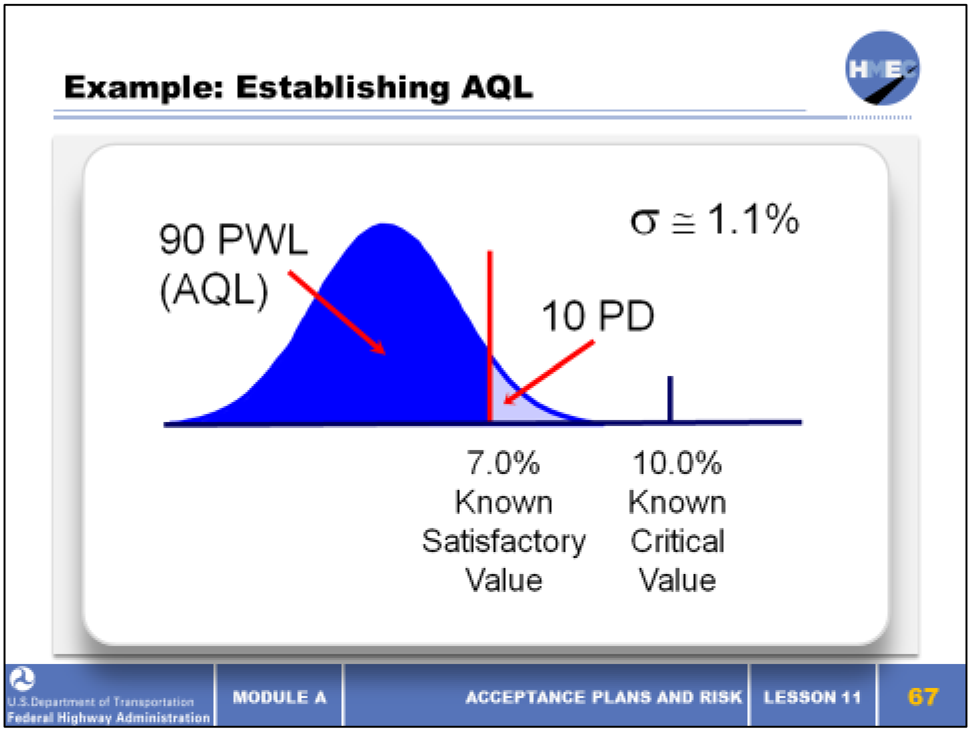
____

____

____

____

Slide 67



This example is of an acceptance plan for the percent passing the #200 sieve for an aggregate base course. Experience indicates that bases with 7% or less – #200 material perform well, and that they perform poorly if the amount of – #200 material exceeds 10%. A typical standard deviation for this material has been found to be about 1.1%. It is believed that the base will perform well as long as 90% or more (i.e., most) of the material has a – #200 value of 7% or less. Thus, the AQL is 90 PWL. This is a relatively conservative definition because, even if the standard deviation were considerably larger than the typical value, there is little chance that any of the material in the normal distribution representing AQL quality would reach the known critical value of 10% passing the #200 sieve.

____

____

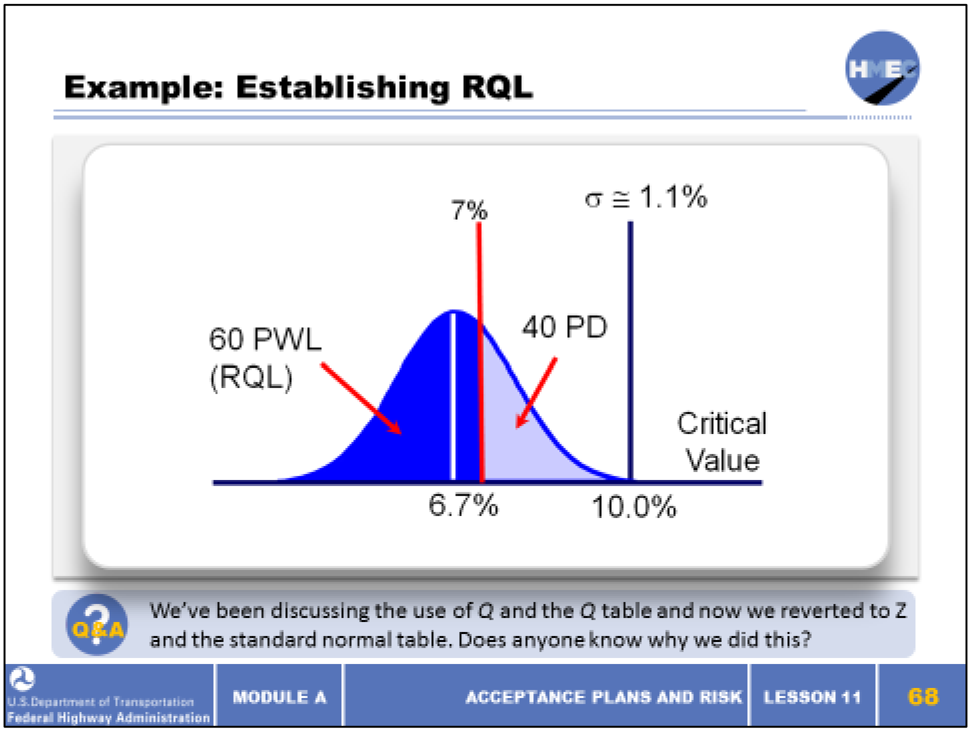
____

____

____

____

Slide 68



If we place the extreme upper tail of a normal distribution with a standard deviation of 1.1% at the known critical value of 10%, then the mean of that distribution will be at 10% – (3 x 1.1%) = 6.7%. The table of the standard normal curve can be used to determine that approximately 60% of the population lies below the known satisfactory value of 7% (or approximately 40% above the satisfactory value). On those occasions where the standard deviation was larger than the typical value of 1.1%, a relatively small portion of the distribution would extend above the critical value of 10%. As the amount of material below 7.0% – #200 material decreases below 60%, however, progressively more will exceed the critical value of 10% and performance problems might be expected to develop. Thus, the RQL is chosen as 60 PWL.

____

____

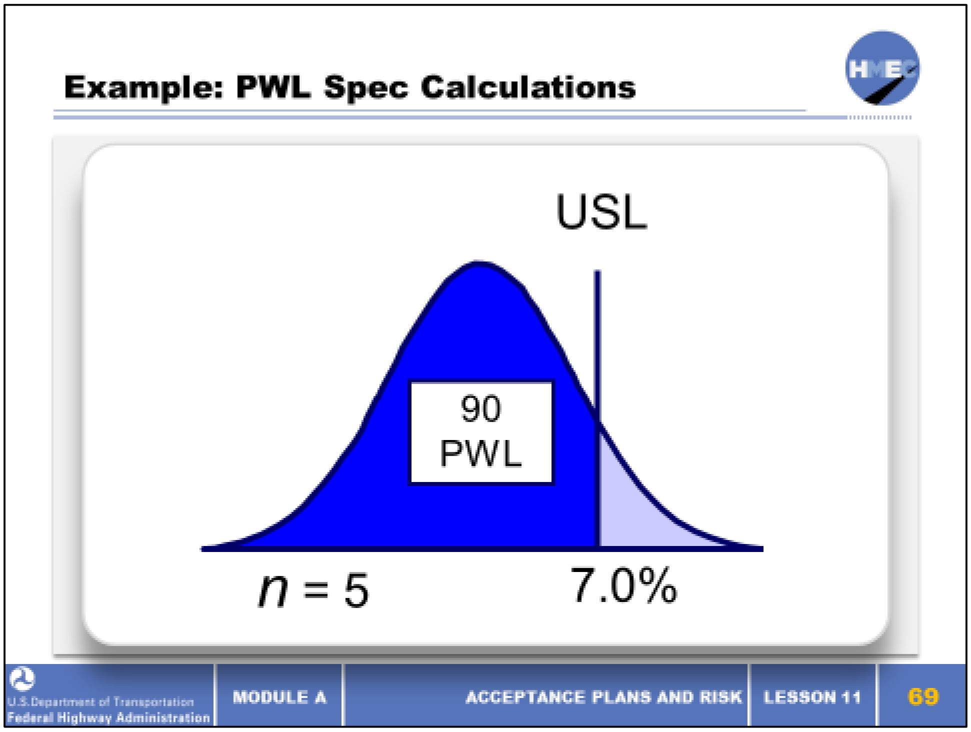
____

____

____

____

Slide 69



The specifying agency has chosen 7.0% as the upper specification limit and has decided to use a sample size of *n* = 5 for acceptance with an AQL of 90 PWL and a RQL of 60 PWL.

____

____

____

____

____

____

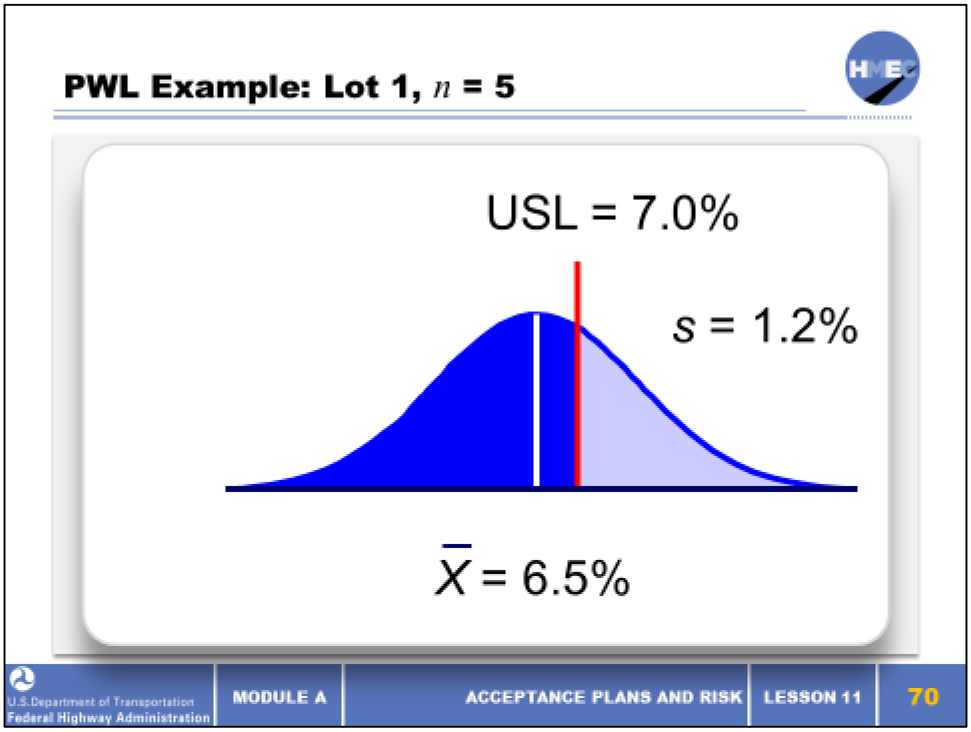
____

____

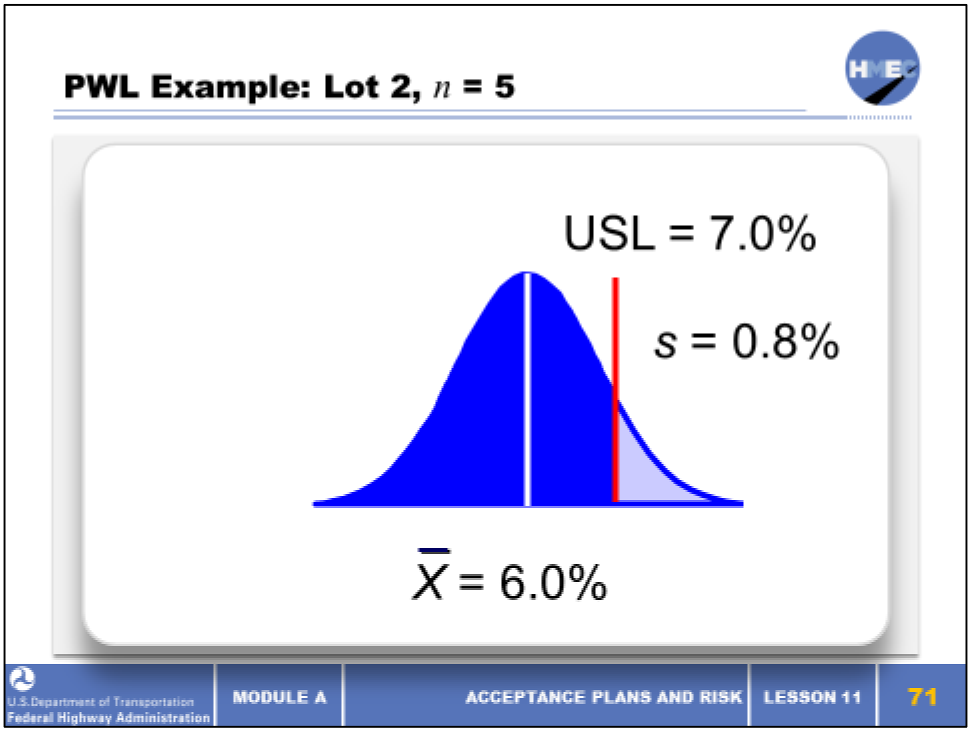
____

____

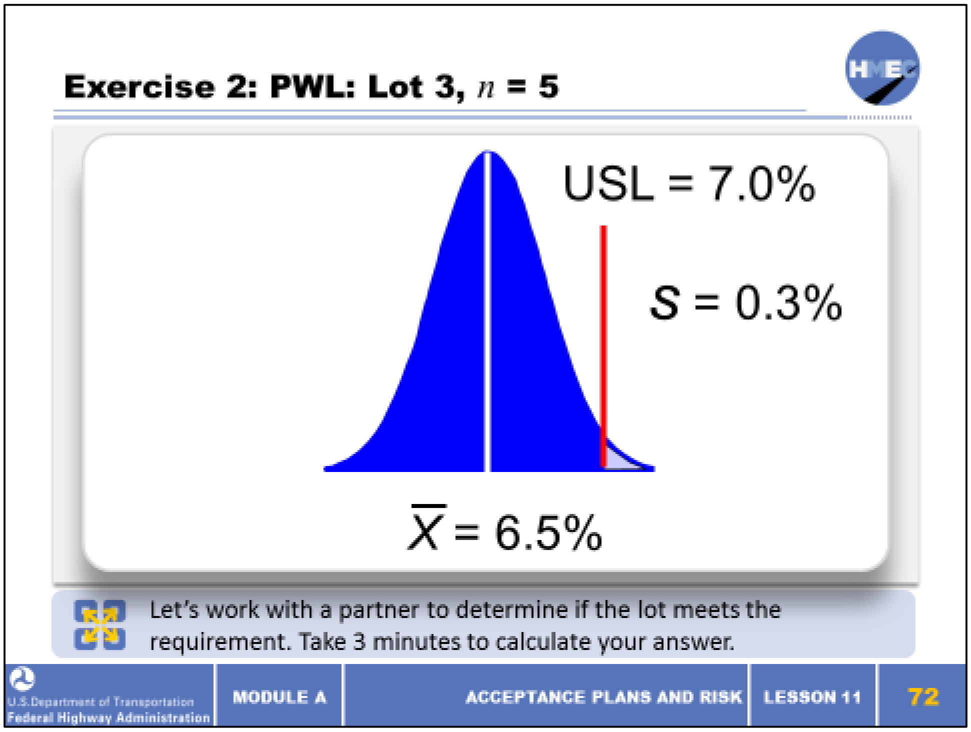
Slide 70


PWL Example: Lot 1, n = 5
USL = 7.0%
s = 1.2%
X̄ = 6.5%
U.S. Department of Transportation
Federal Highway Administration
MODULE A
ACCEPTANCE PLANS AND RISK
LESSON 11
70

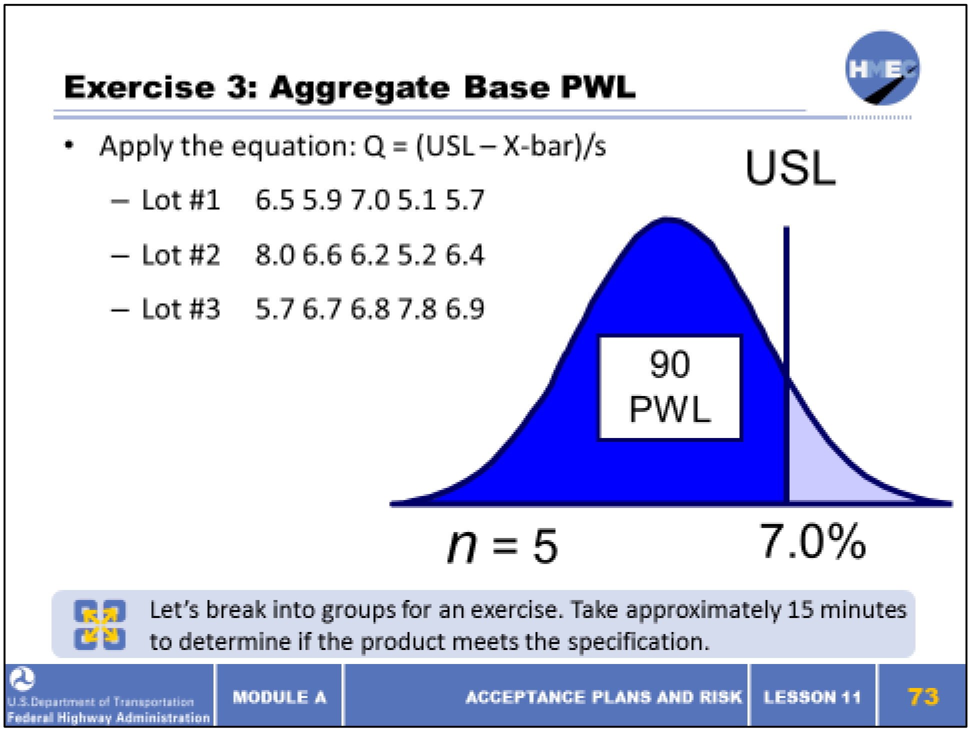
Slide 71


PWL Example: Lot 2, n = 5
HMEC
USL = 7.0%
s = 0.8%
X̄ = 6.0%
U.S. Department of Transportation
Federal Highway Administration
MODULE A
ACCEPTANCE PLANS AND RISK
LESSON 11
71

Slide 72


Exercise 2: PWL: Lot 3, n = 5
USL = 7.0%
S = 0.3%
$\overline{X}$ = 6.5%
Let's work with a partner to determine if the lot meets the requirement. Take 3 minutes to calculate your answer.
U.S. Department of Transportation
Federal Highway Administration
MODULE A
ACCEPTANCE PLANS AND RISK
LESSON 11
72

Slide 73



Using the data for the three lots that follow, determine whether or not each of the lots meets the specification for – #200 material. By solving the following problems, you will be able to determine if the product meets the specification.

Use calculators to compute average and standard deviation. Calculate *Q* and determine the PWL.

- Lot #1: 6.5, 5.9, 7.0, 5.1, 5.7
- Lot #2: 8.0, 6.6, 6.2, 5.2, 6.4
- Lot #3: 5.7, 6.7, 6.8, 7.8, 6.9

Then, determine if the lot meets specification.

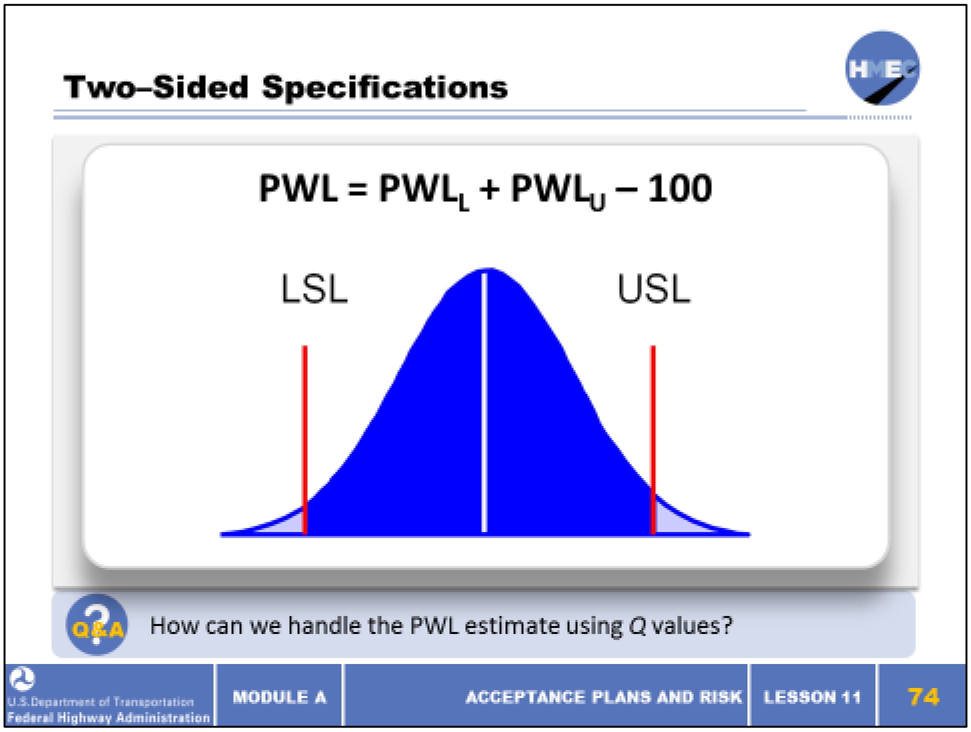
____

____

____

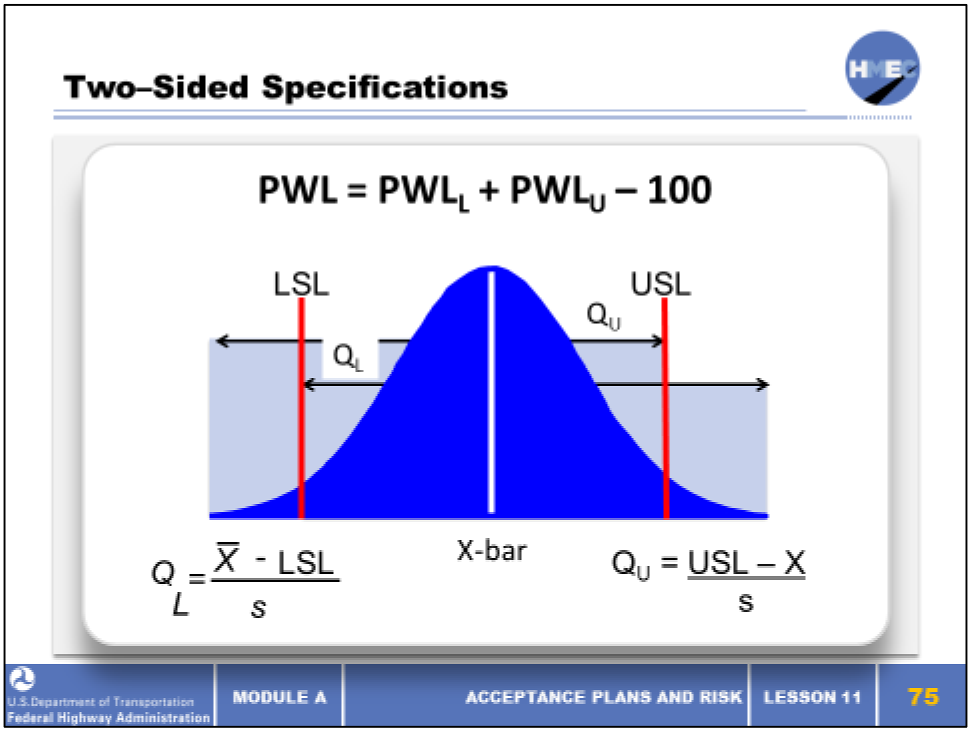
____

Slide 74



It is now possible for the lot to have material that is outside of both specification limits.

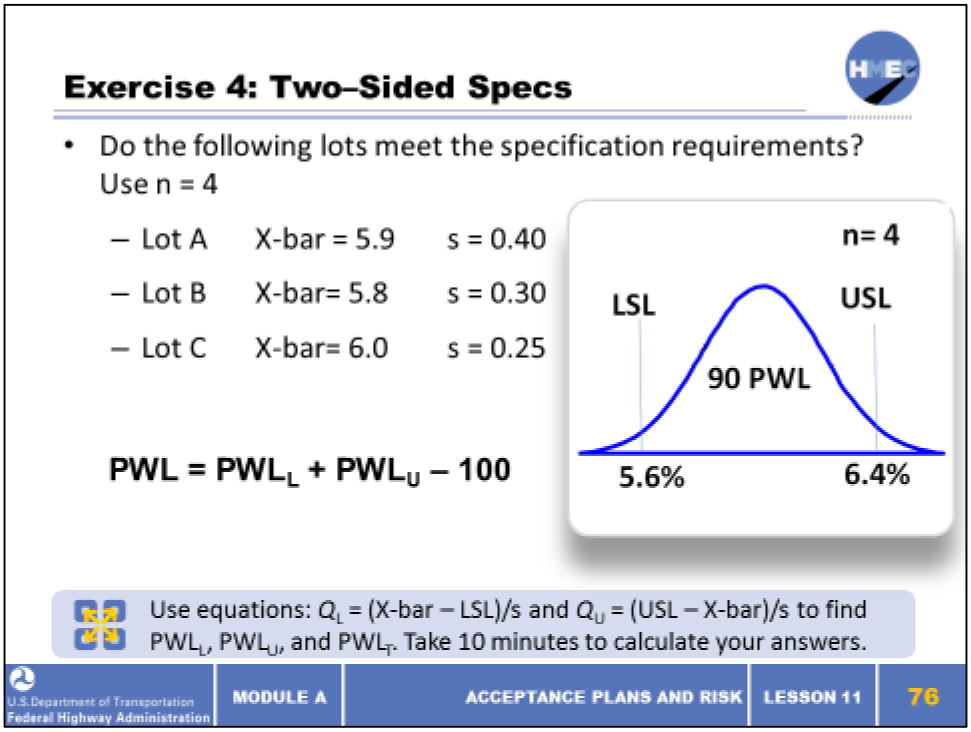
Slide 75



Since the *Q* values are measured from each specification limit, it is now possible for the lot to have material that is outside of both specification limits.

____

Slide 76



By solving these three problems, you will be able to determine if the product meets the specification.

Suppose we have an asphalt content specification that requires 90 PWL with specification limits of 5.6% to 6.4%.

Do the following lots meet the specification requirements? Use $n = 4$.

- Lot A: X-bar = 5.9; s = 0.40
- Lot B: X-bar= 5.8; s = 0.30
- Lot C: X-bar= 6.0; s = 0.25

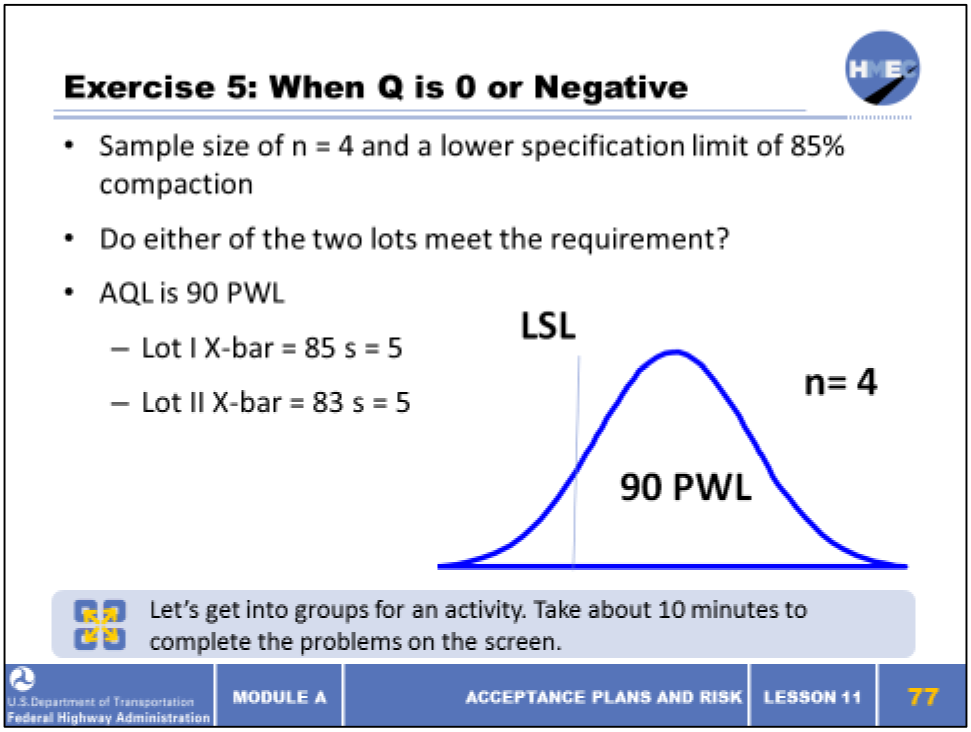
____

____

____

____

Slide 77



Up to now, the calculated *Q* values have been positive in all of our examples. What happens if *Q* is zero, or negative? Suppose we have a specification with a sample size of $n = 4$ and a lower specification limit of 85% compaction on a stabilized base course. Do either of the two lots meet the requirement? Assume the AQL is 90 PWL.

- Lot I : X-bar = 85 s = 5
- Lot II: X-bar = 83 s = 5

____

____

____

____

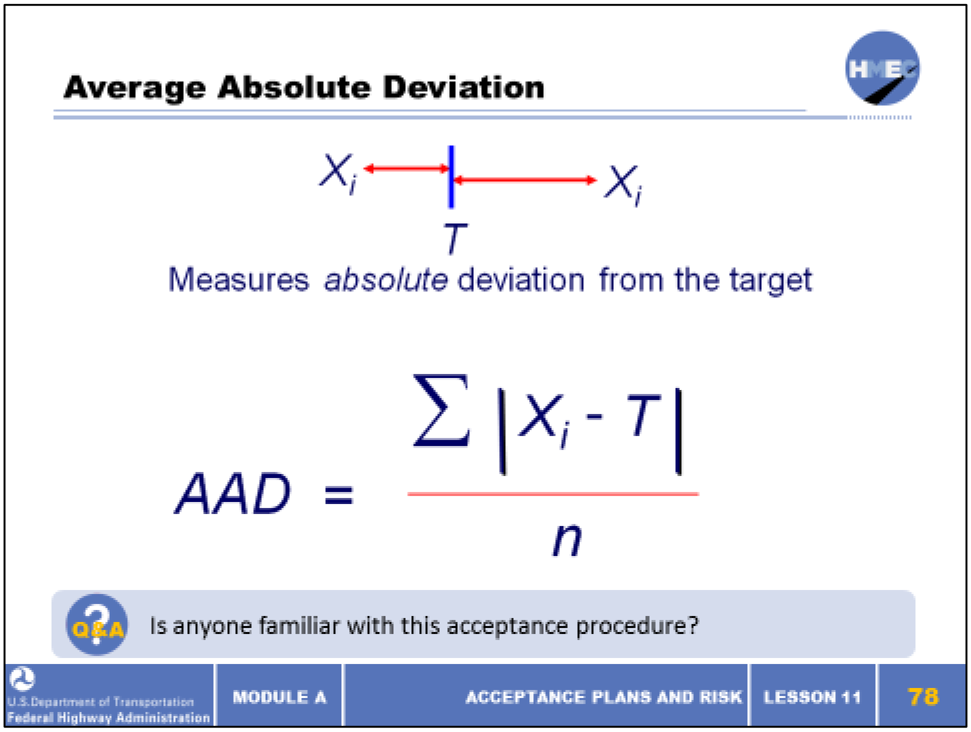
____

____

____

____

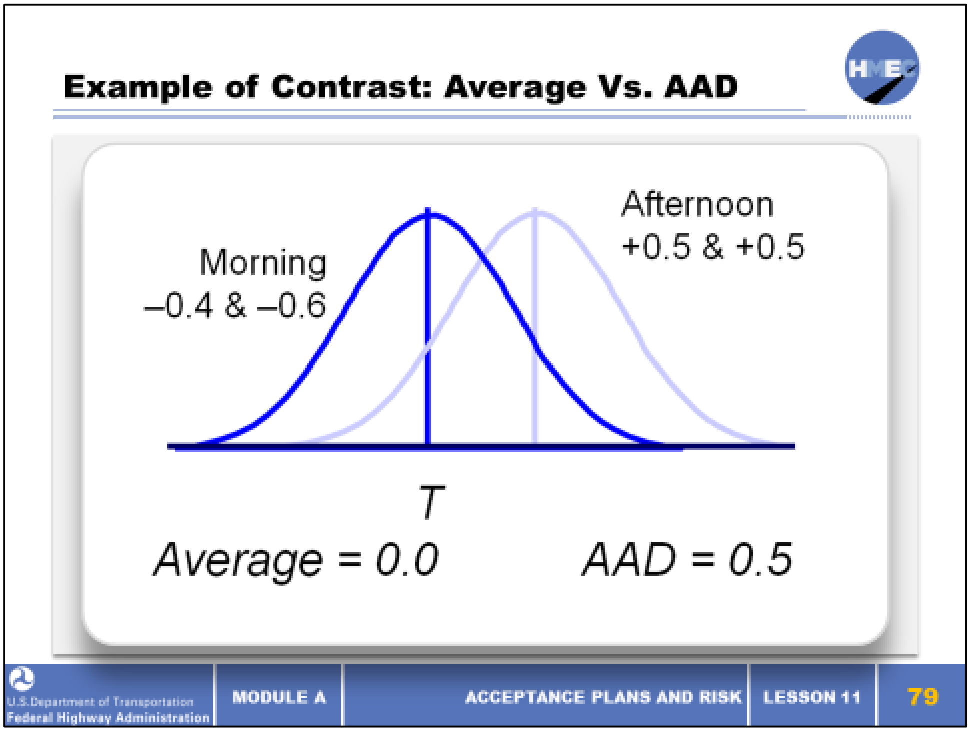
Slide 78



If the average absolute deviation (AAD) from the target is used for acceptance, the contractor cannot benefit by any strategy other than aiming for the target value. The equation for AAD is to sum all deviations from the target and divide by the sample size. This method allows the positive and negative values to accumulate.

____

Slide 79



In the example in the slide, if two early tests on the lot indicate values below the target by –0.4 and –0.6, using AAD the contractor cannot offset these low values with two later values of +0.5 and +0.5. This is true because, while the average of these four numbers is (–0.4 –0.6 +0.5 +0.5) ÷ 4 = 0.0, the average of the absolute values of these numbers is (0.4 + 0.6 + 0.5 +0.5) ÷ 4 = 0.5.

Average = (-0.4-0.6+0.5+.05)/4=0.0

AAD = (l-0.4l – 0.6l + 0.5l + 0.5l)/4=0.5

____

____

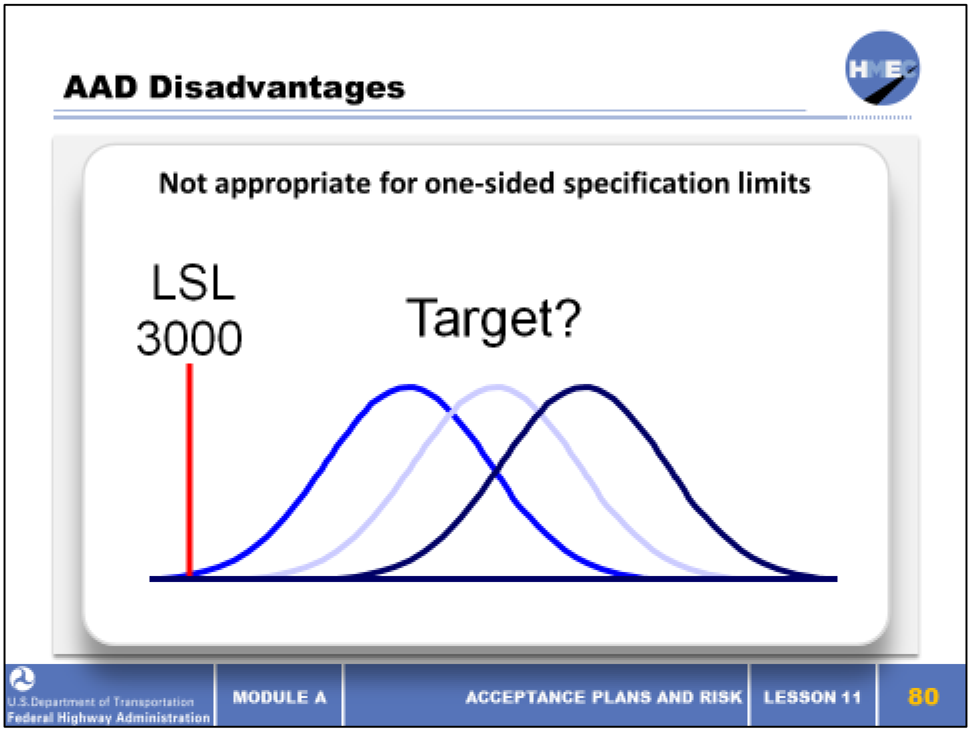
____

____

____

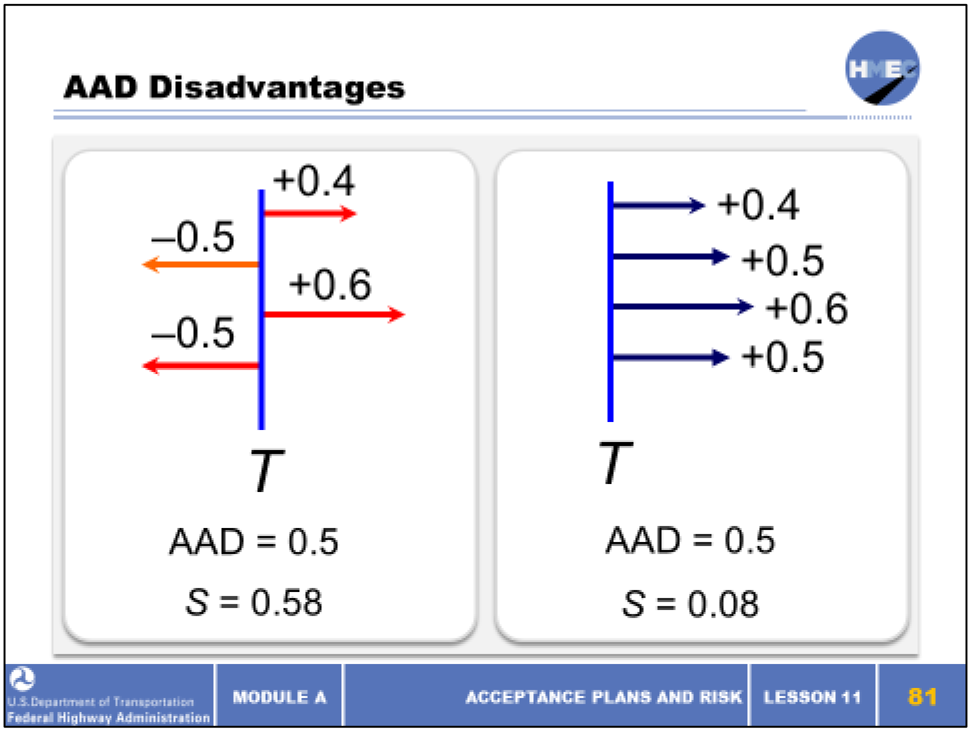
____

Slide 80



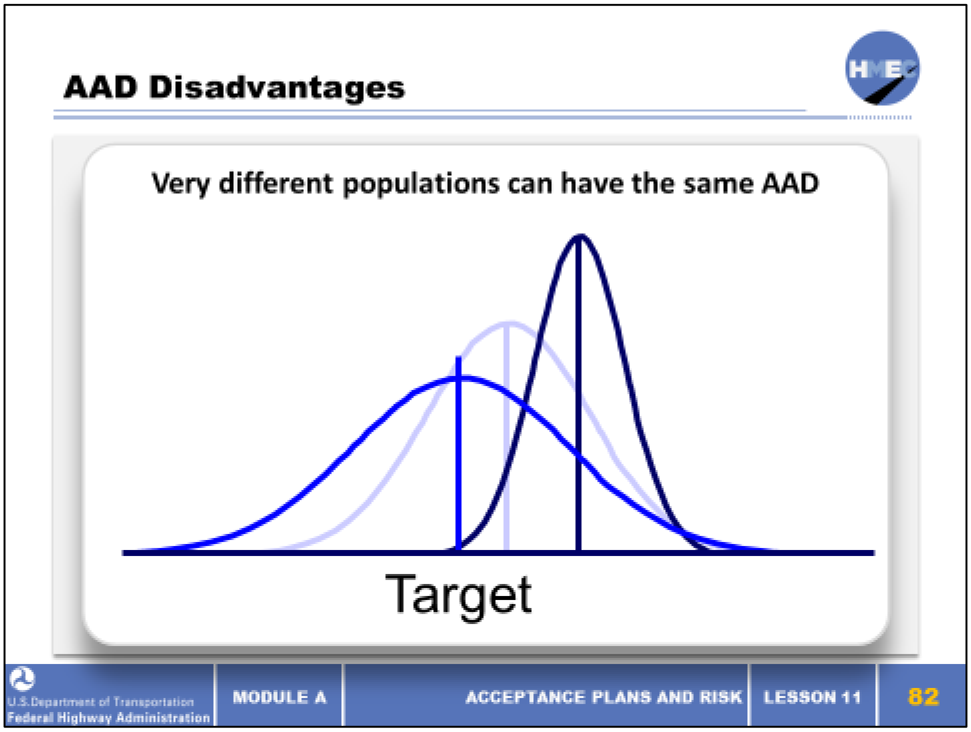
The primary drawback is that they can only be used when there is a target value. They cannot, therefore, be used when there is only one specification limit, as might be the case with concrete compressive strength.

Slide 81



Here we see two very different sets of sample results. The first has +0.4, - 0.5, +0.6, and -0.5. The AAD of these four results is 0.5, but the standard deviation is 0.58. The second set has the same four values but they are all on the positive side of the target. The AAD is still 0.5 but now s= 0.08.

Slide 82



A drawback with AAD acceptance plans is that since variability is not directly measured, a given lot (population) AAD can come from a number of different populations. The population could be centered at the target, but have a relatively large standard deviation, i.e., larger than the one that was assumed when developing the AAD acceptance limits or payment equation such as the distribution on the left. Another population could have the same AAD by being centered off the target and having the same standard deviation that was assumed when developing the acceptance values, the distribution in the center. A third population could have the same AAD value by having a mean far from the target, but also having a relatively small standard deviation, the distribution on the right.

____

____

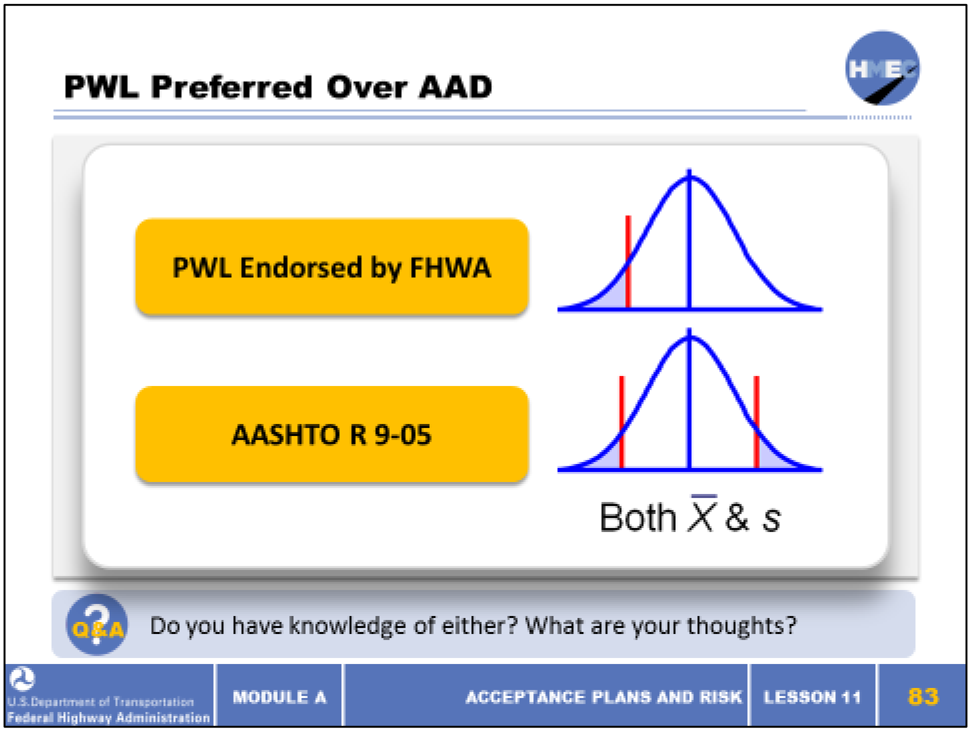
____

____

____

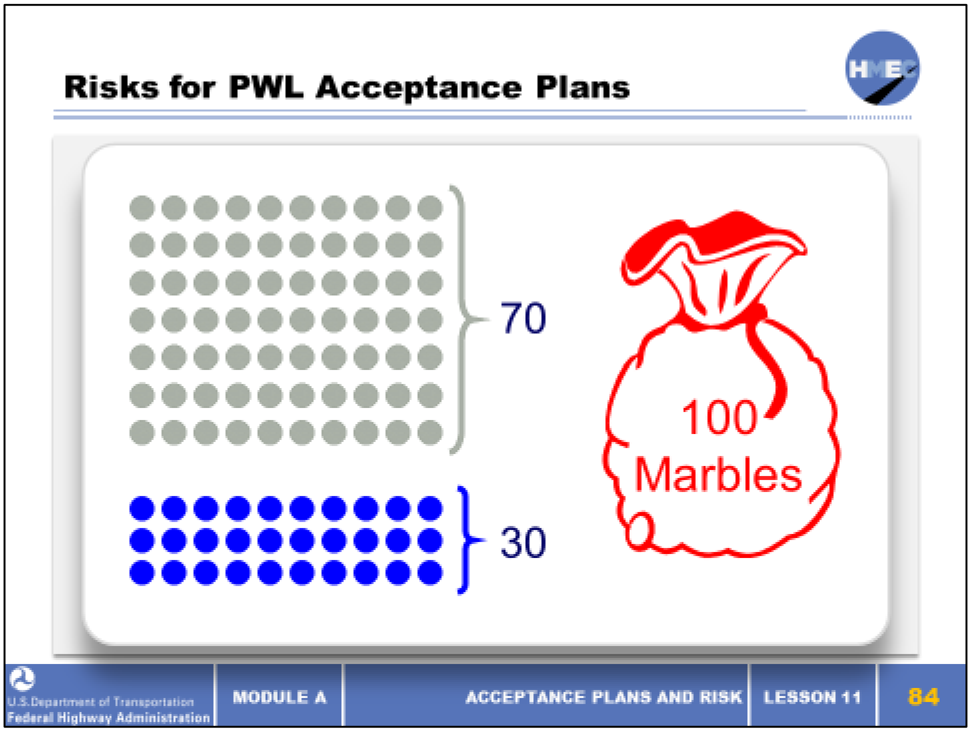
____

Slide 83



While some of the drawbacks for AAD may also apply to PWL acceptance plans, such as the fact that a given PWL can represent many different populations, there are fewer drawbacks due to the fact that both sample mean and standard deviation are determined in the PWL method. Also, since the PWL method can be used with both one-sided and two-sided acceptance properties, it is more versatile since it does not require separate approaches for one-sided and two-sided cases.

Slide 84



Earlier in this lesson, we learned how to calculate risks for acceptance plans based on an assumed known standard deviation, and with the acceptance decision based on only the sample mean. The risks associated with PWL acceptance plans cannot be calculated so easily. In fact, with PWL acceptance plans, the risks are almost always determined by means of computer simulation. A computer program called SPECRISK has been developed to measure risks with PWL acceptance plans. This will be discussed in more detail in Lesson 15. It is, however, possible to illustrate the risks associated with using a sample to estimate PWL by means of a simplified attributes example. Assume that we have a bag that has 100 marbles. Further assume that the bag has 70 gray marbles and 30 blue marbles. Also assume that we wish to take a sample of 10 marbles to estimate the percentage of the marbles in the bag that are blue. After each sample of 10 is recorded, those marbles are placed back in the bag prior to taking the next sample.

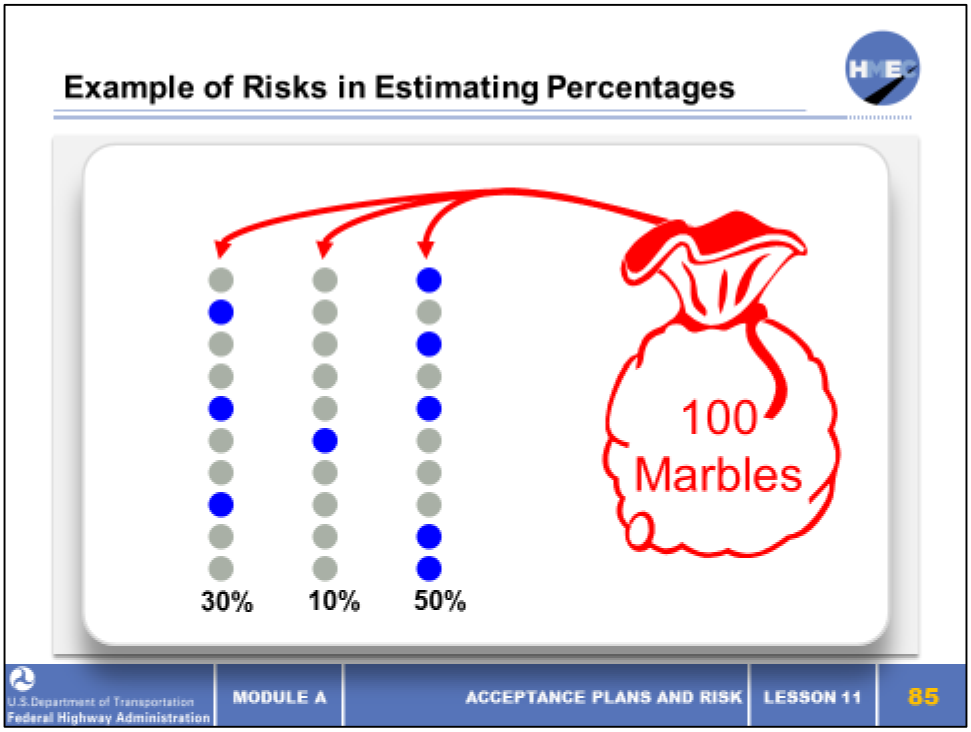
____

____

____

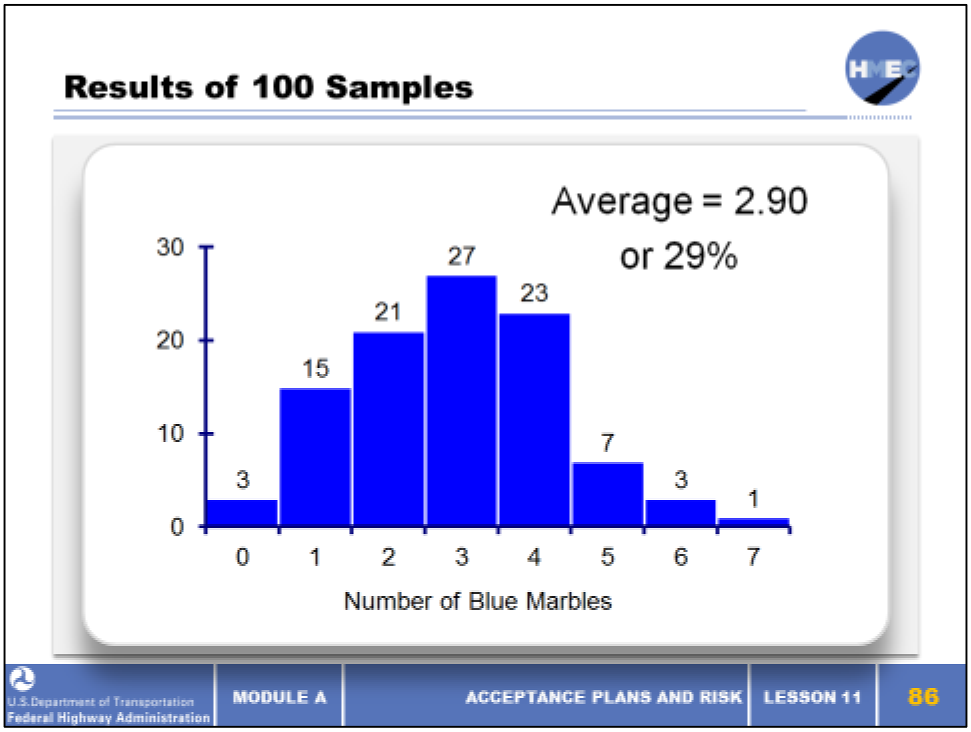
____

Slide 85



However, each sample of 10 marbles will not yield the same percentage of blue marbles. The slide shows the result of the first three samples of 10 marbles that were obtained. Note that an individual sample may provide the correct percentage of blue marbles, e.g., the first sample. However, it is also likely that the sample will not provide the correct estimate of the actual percentage of blue marbles.

Slide 86



The histogram in the slide shows the results of 100 samples, each with 10 marbles. While the individual sample results could be quite far from the actual percentage in the population, the average of the 100 samples is quite close to the true population value. Also, most of the sample values are close to the actual population percentage, with fewer values as the estimate becomes farther from the actual population percentage.

____

____

____

____

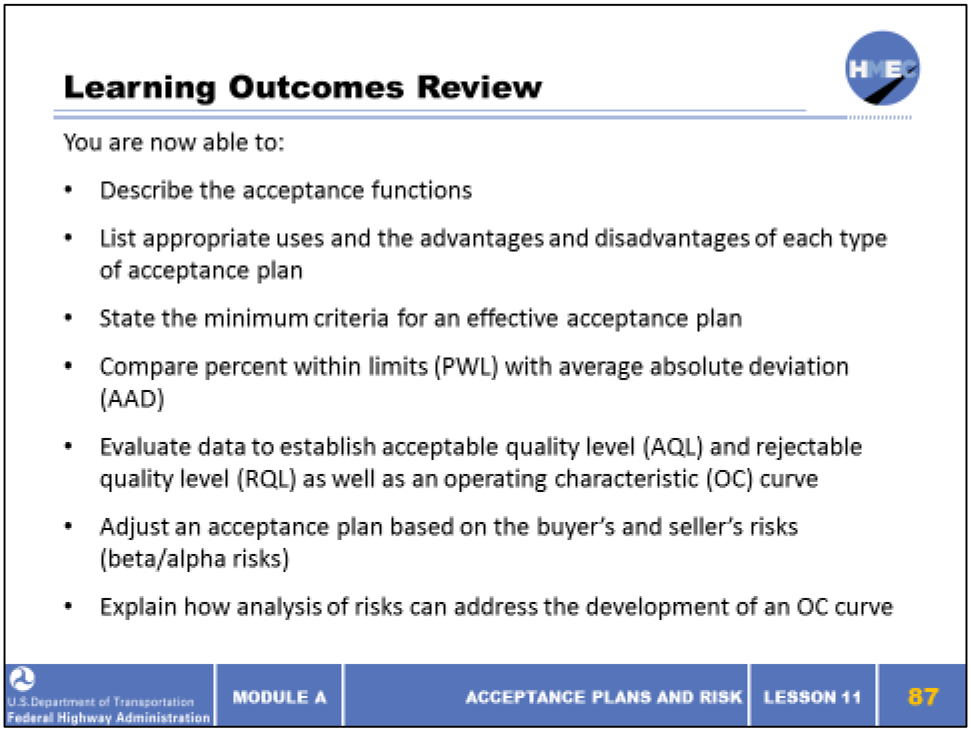
____

____

____

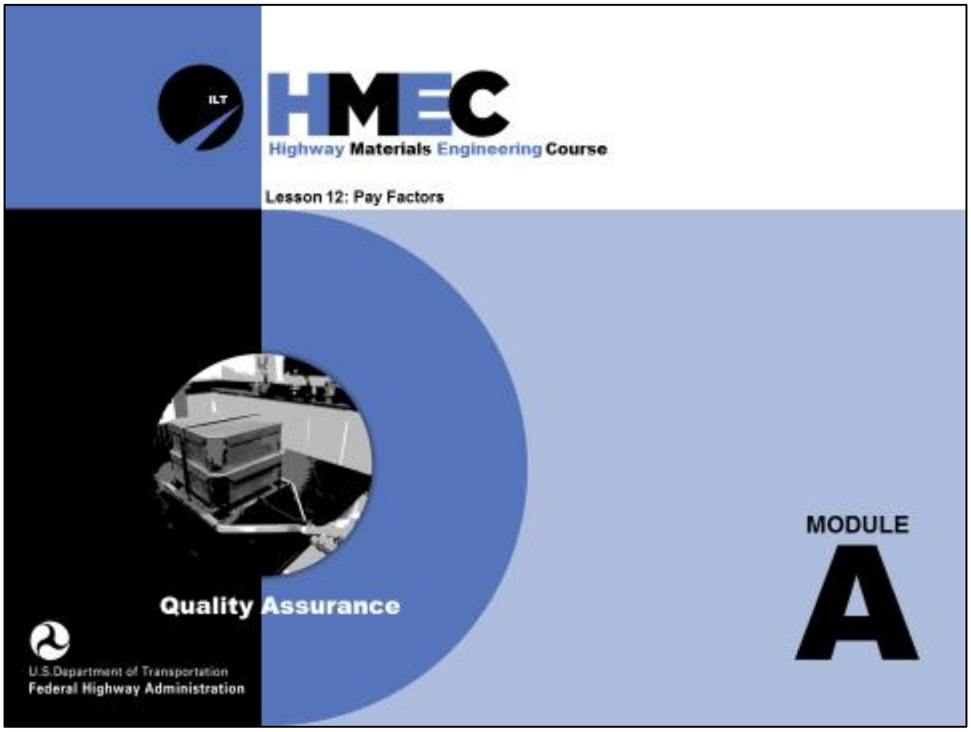
____

Slide 87

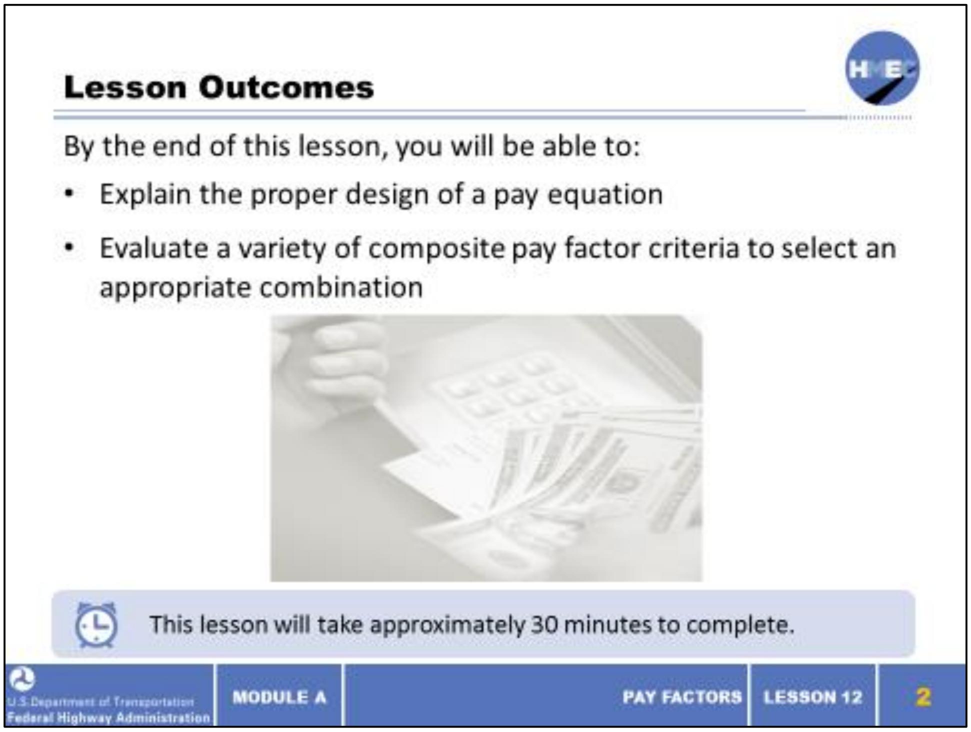

Learning Outcomes Review
You are now able to:
• Describe the acceptance functions
• List appropriate uses and the advantages and disadvantages of each type of acceptance plan
• State the minimum criteria for an effective acceptance plan
• Compare percent within limits (PWL) with average absolute deviation (AAD)
• Evaluate data to establish acceptable quality level (AQL) and rejectable quality level (RQL) as well as an operating characteristic (OC) curve
• Adjust an acceptance plan based on the buyer's and seller's risks (beta/alpha risks)
• Explain how analysis of risks can address the development of an OC curve
U.S. Department of Transportation
Federal Highway Administration
MODULE A
ACCEPTANCE PLANS AND RISK
LESSON 11
87

This page is intentionally blank.

Slide 1


ILT
HMEC
Highway Materials Engineering Course
Lesson 12: Pay Factors
Quality Assurance
U.S.Department of Transportation
Federal Highway Administration
MODULE
A

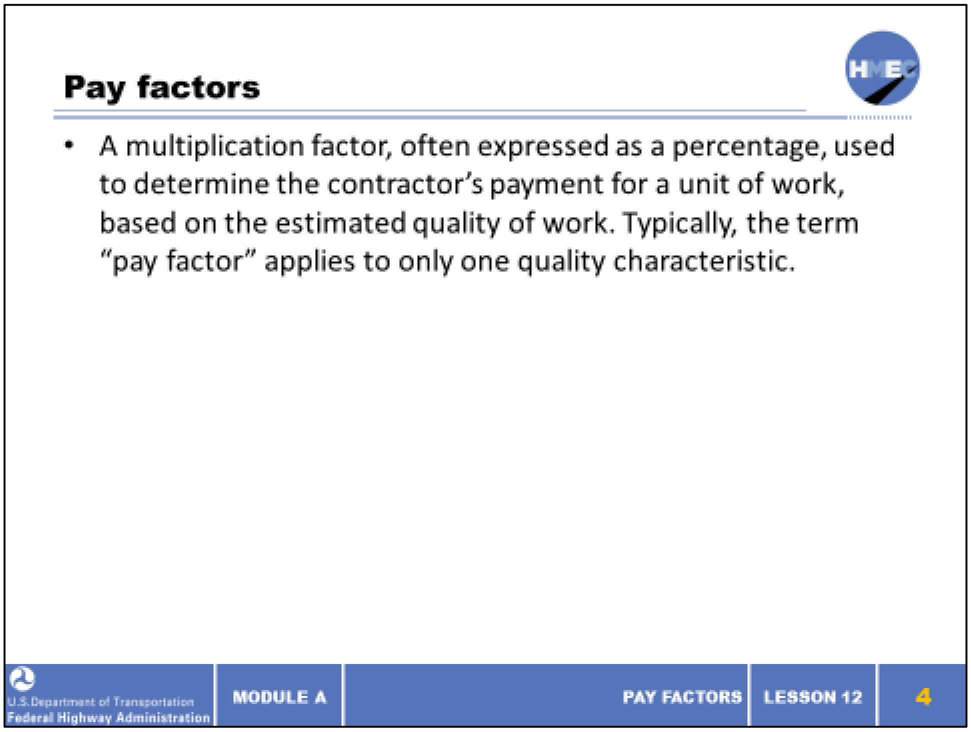
Slide 2


Lesson Outcomes
By the end of this lesson, you will be able to:
• Explain the proper design of a pay equation
• Evaluate a variety of composite pay factor criteria to select an appropriate combination
This lesson will take approximately 30 minutes to complete.
U.S. Department of Transportation
Federal Highway Administration
MODULE A
PAY FACTORS
LESSON 12
2

Slide 3

The procedure typically used in the past was to correct, if possible, or remove and replace (R&R) product that had less than acceptable quality. Another option has been to shut down the operation until the quality level could be improved. Such methods, typically, are costly and may not provide positive incentives to keep the process in control. More recently, price reduction schedules have become more commonplace.

Slide 4



Here are more terms related to pay factors:

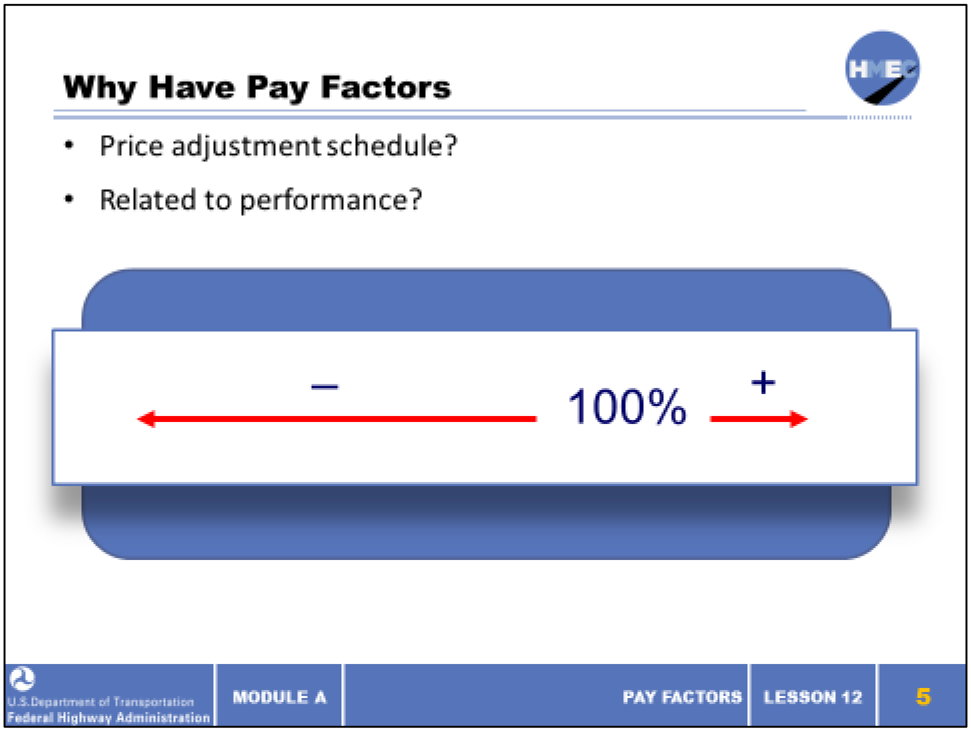
- A pay adjustment system refers to more than one schedule or to a schedule that considers several quality characteristics.
- Pay adjustment schedules may be categorized as
    - – Graduated (stepped) schedules versus continuous schedules.
    - – Tabular schedules versus schedules in equation form.
    - – Schedules that provide pay factors versus schedules that provide pay adjustment dollar amounts.
- Pay adjustment schedules, including those that allow pay increases, do not necessarily function as incentive or disincentive provisions.
- Pay adjustment schedules may or may not be based on liquidated damages.
- Pay adjustment. The actual amount, either in dollars or in dollars per area/weight/volume that is to be added or subtracted to the contractor's bid price or unit bid price.

____

____

____

Slide 5



The ideal state-of-the-art price adjustment philosophy is that the price should be adjusted commensurate with the estimated performance of the product. If the performance is estimated to be 10% below that desired, the price adjustment should consider this when determining the percentage of the bid price to pay. Likewise, if the estimated performance is greater than that specified, a positive price adjustment should be considered. Negative price adjustments are often called penalties and positive price adjustments have been called bonuses. Although often used, both terms are discouraged. The preferred terms are "incentives" and "disincentives". It should be understood that the quality measure, PWL, although used widely is often not truly performance-related.

____

____

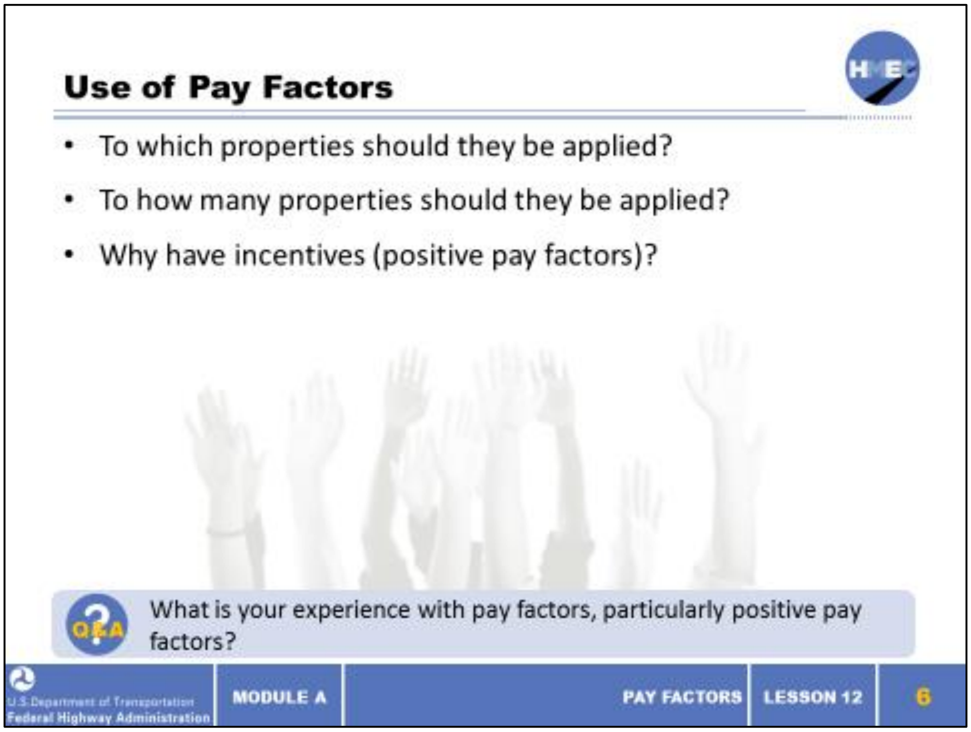
____

____

____

____

Slide 6



Pay factors should be applied to only properties that are considered important (e.g., asphalt content, concrete pavement thickness). They should be applied to the minimum that affect performance. Keep in mind, incentives are not a giveaway. They are a strong motivation for the contractor to meet the specs.

____

____

____

____

____

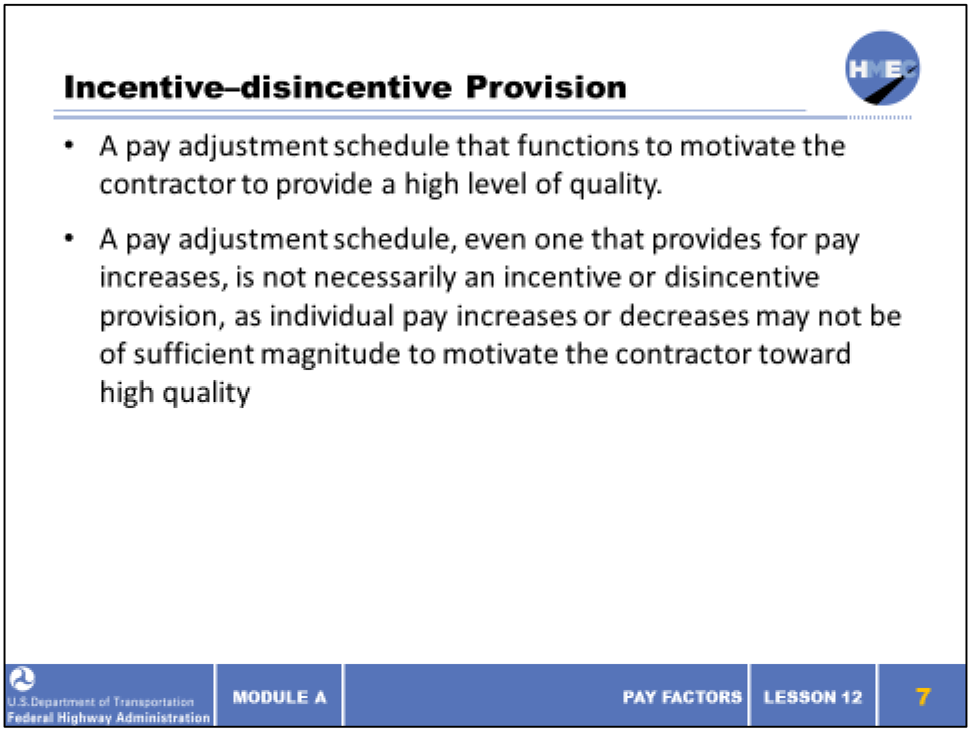
____

____

____

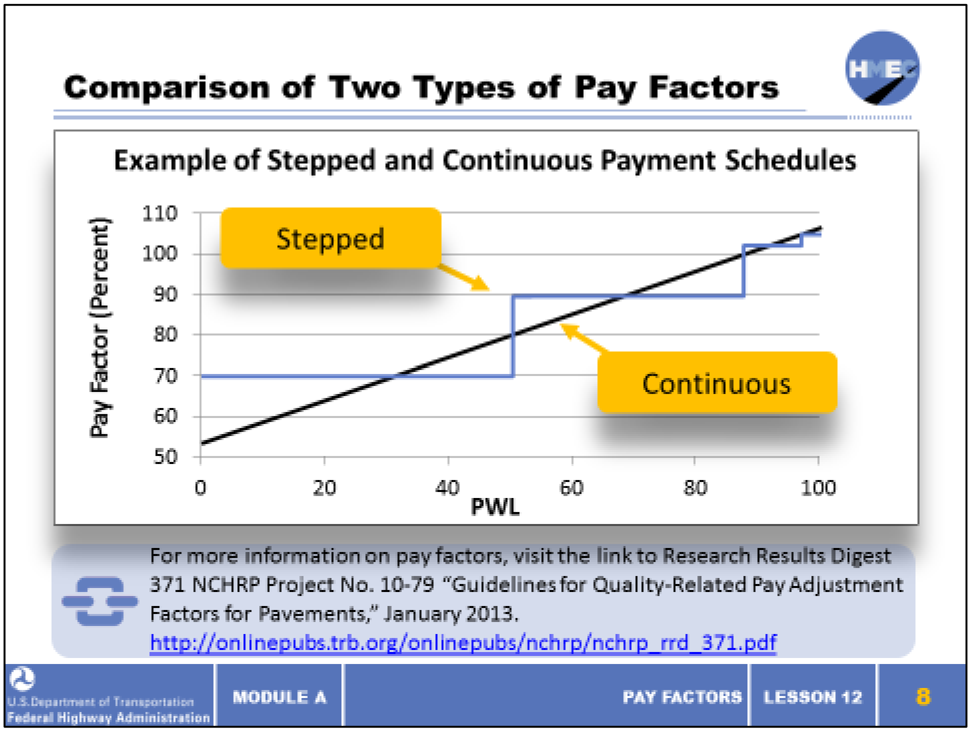
____

Slide 7



A pay adjustment schedule, even one that provides for pay increases, is not necessarily an incentive or disincentive provision, as individual pay increases or decreases may not be of sufficient magnitude to motivate the contractor toward high quality. However, they have proven to be important to the contractor because they provide a way to offset disincentives in the long run and have shown particularly in the case of ride quality incentives to provide smoother ride. From a statistical viewpoint it will be seen they are essential for balancing risks and providing reasonable expected pay (EP) and OC curves.

Slide 8



This example points out a weakness in stepped pay factor schedules that should be avoided. The disadvantage is that a small change in quality (as measured in this example by percent within limits, or PWL) can make a large difference in payment. For example, if the PWL estimate is 49.9, the payment is 70% but the payment would be 90% for a one-tenth increase in quality to a PWL of 50.0.

Slide 9

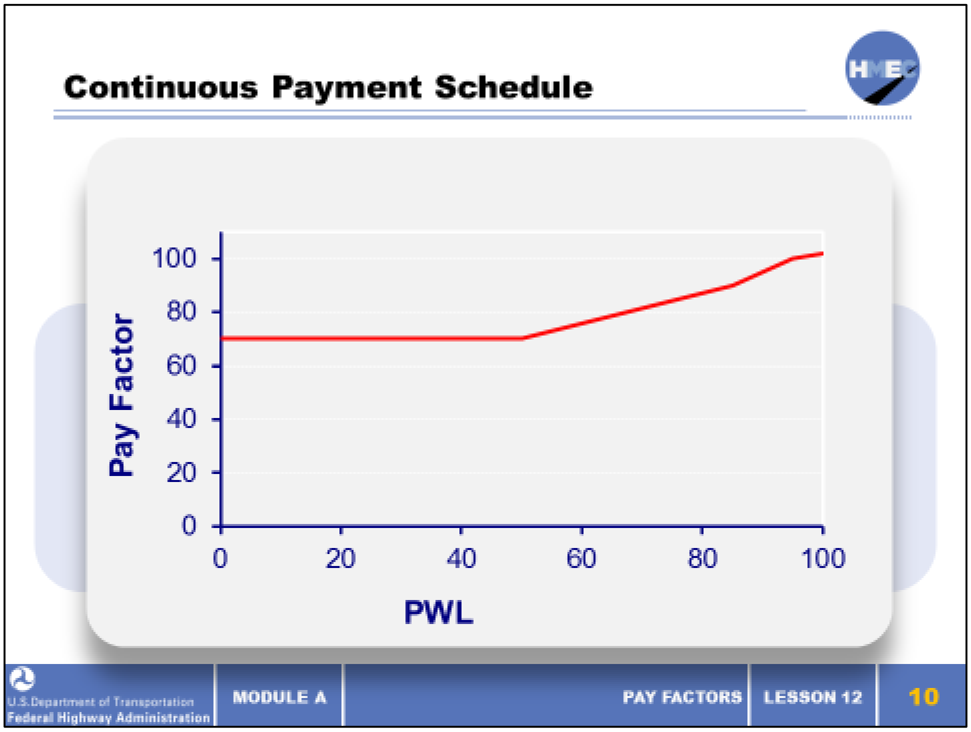
Typical Stepped Payment Schedule Based on PWL

Estimated PWL	Pay Factor, %
95.0 – 100.0	102
85.0 – 94.9	100
50.0 – 84.9	90
0.0 – 49.9	70

U.S. Department of Transportation Federal Highway Administration | MODULE A | PAY FACTORS | LESSON 12 | 9

This slide shows a stepped pay factor schedule for PWL. In this example a change in PWL from 84.9 to 95.0 can mean a pay factor change of 10%. Although the resultant pay factor may average out statistically in the long run, they can have impacts on payment levels on individual projects that may not be fair to either the contractor or the agency. They may also provide an incentive for “rounding” results. Also take note that a pay factor of 100% occurs when the PWL is 85.0 – 94.9. The importance of this will be discussed in more detail later.

Slide 10



This type of schedule is often based on an equation or series of equations and avoids the problem of relatively small changes in the measured property leading to relatively large differences in payment level.

____

____

____

____

____

____

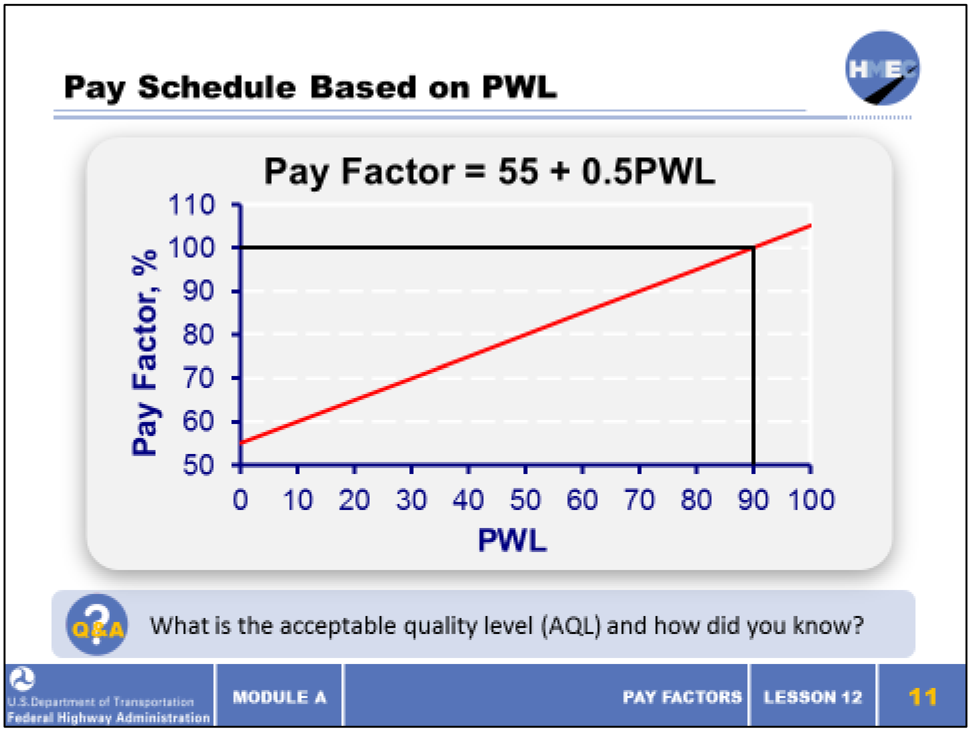
____

____

____

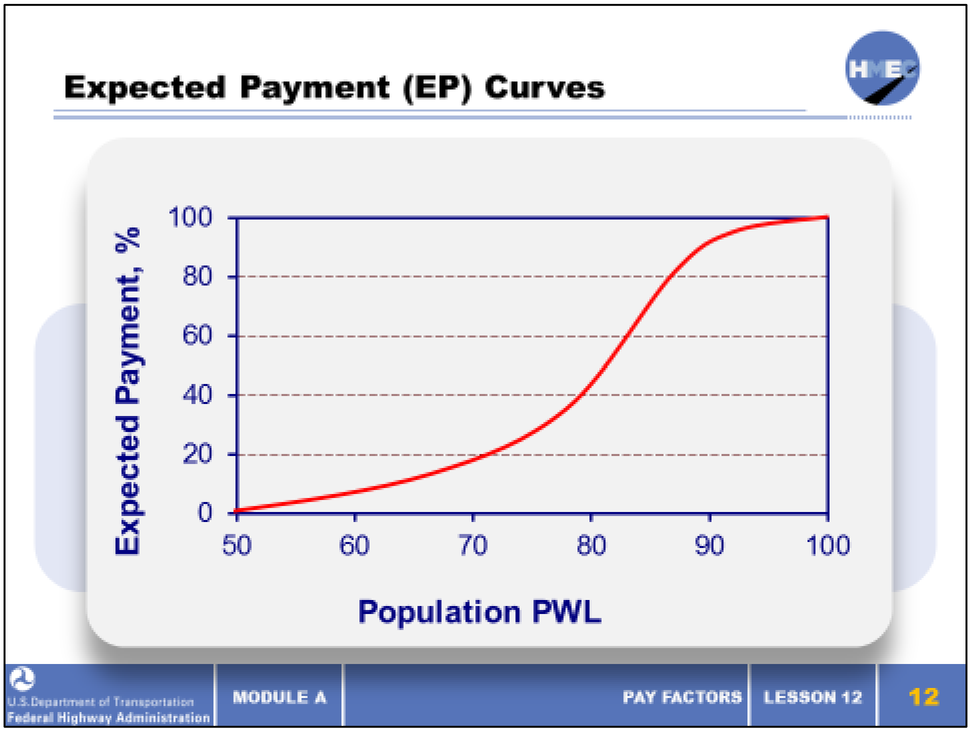
____

Slide 11



In this approach, the sample mean and standard deviation are used to calculate a *Q* value. The *Q* value is then used to determine the estimated PWL value. The PWL value is then used in an equation to determine the pay factor for the lot. The equation on the slide is from the AASHTO QA Guide specification.

Slide 12



A graphic representation of an acceptance plan that shows the relation between the actual quality of a lot and it's expected payment (i.e., mathematical pay expectation, or the average pay the contractor can expect to receive over the long run for submitted lots of a given quality). Both OC and EP curves should be used to evaluate how well an acceptance plan is theoretically expected to work. In the case of payment schedules, the appropriate curve plots population PWL on the horizontal axis and expected (or average) payment on the vertical axis.

____

____

____

____

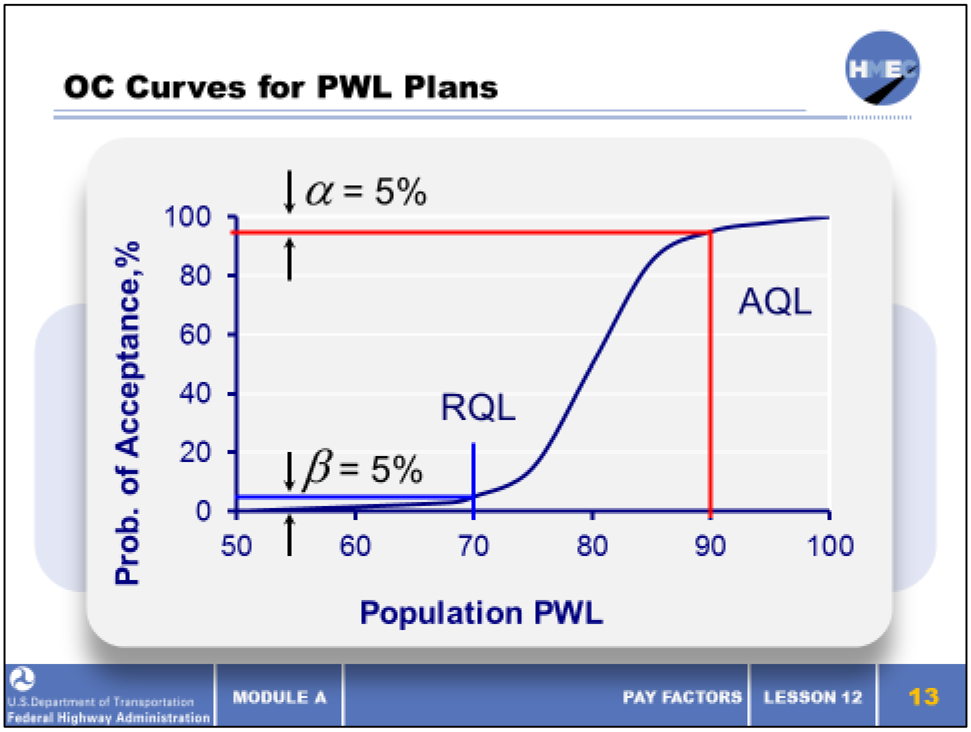
____

____

____

____

Slide 13



In the example shown, if 90 PWL is the AQL, then α = 5% (100% – 95% probability of acceptance at the AQL). If 70 PWL is the rejectable quality level (RQL), then β = 5% (the probability of acceptance at the RQL).

____

____

____

____

____

____

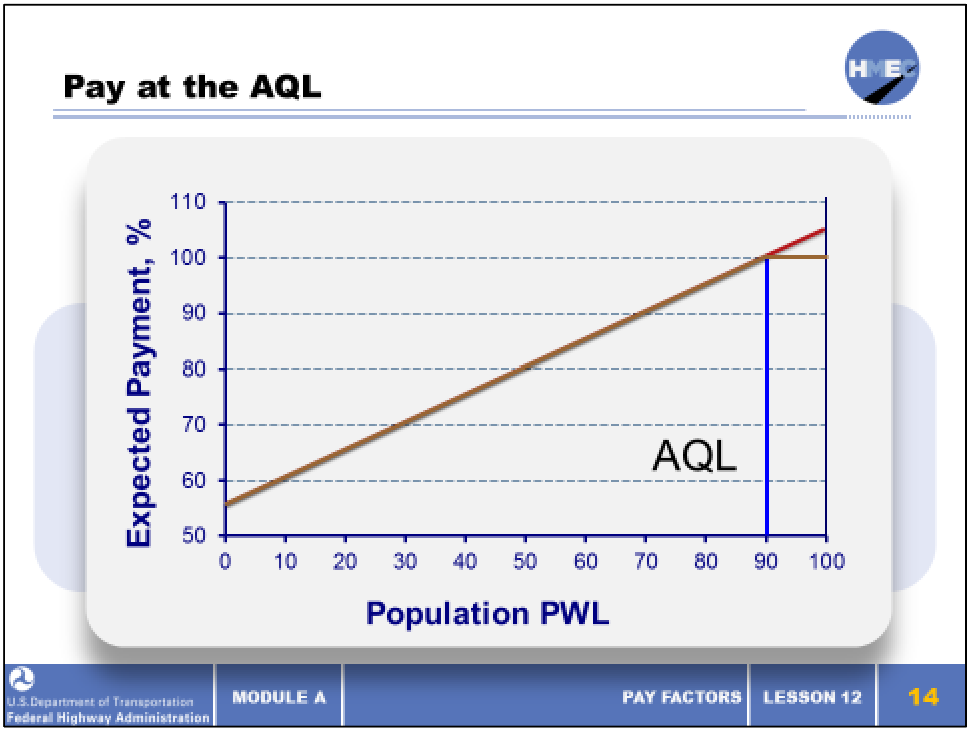
____

____

____

____

Slide 14



In this example, the brown curve has a pay factor of 100% at the AQL but since it's the maximum pay in the average pay factor, in the long run it will be less than 100%. This is because there is variability in the estimated PWL just as there is variability when the mean and standard deviation are estimated. When actually at the PWL, sometimes the estimate will be higher, and sometimes it will be lower.

____

____

____

____

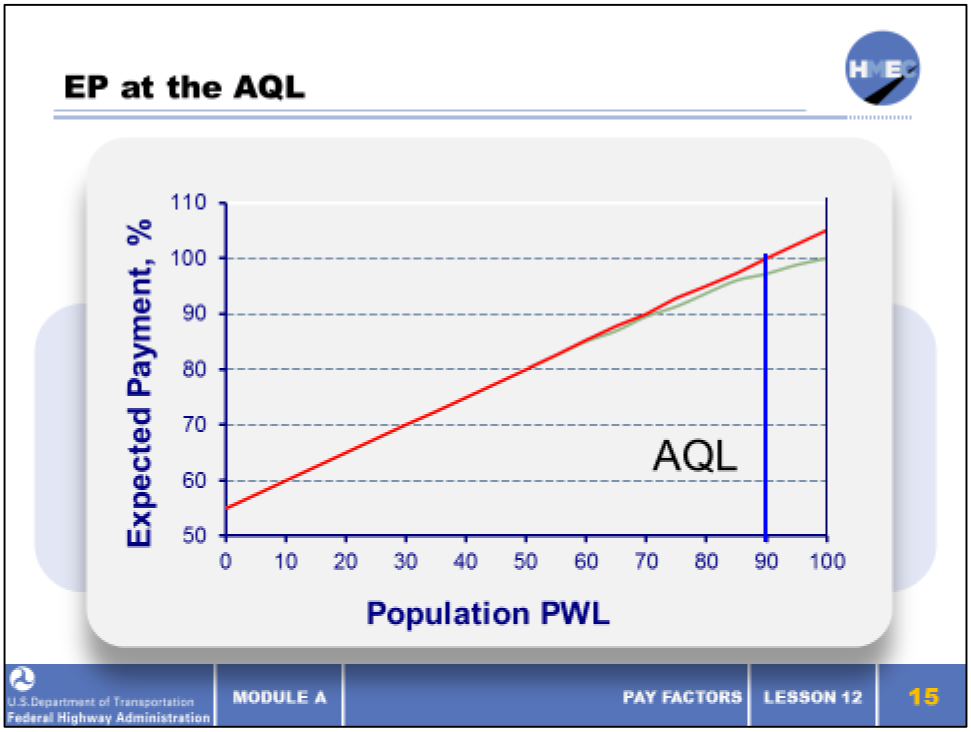
____

____

____

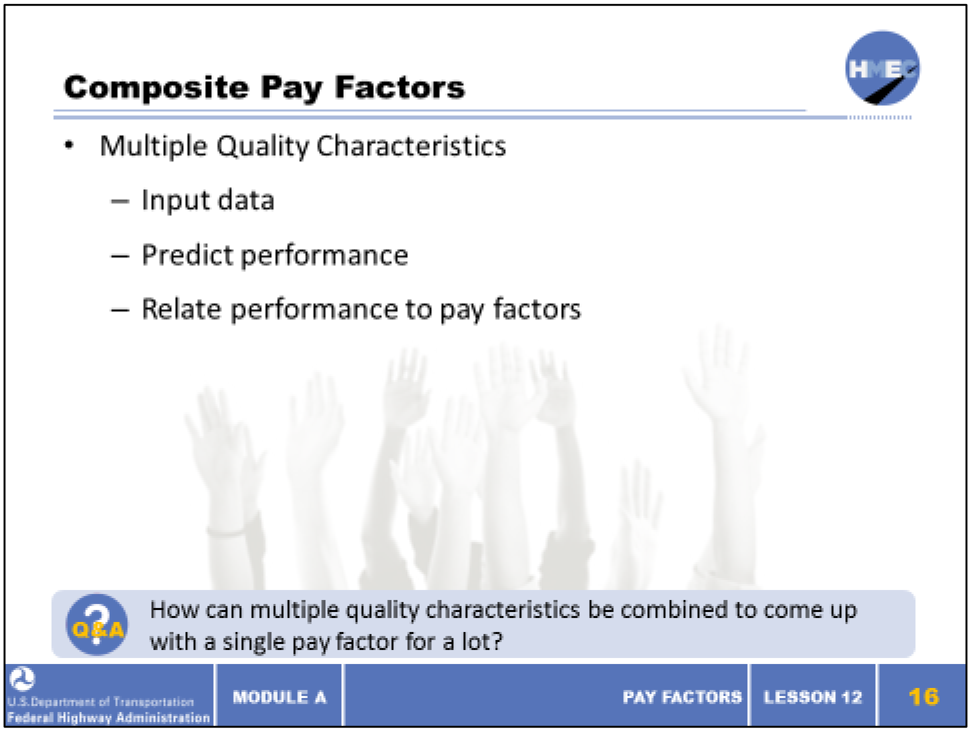
____

Slide 15



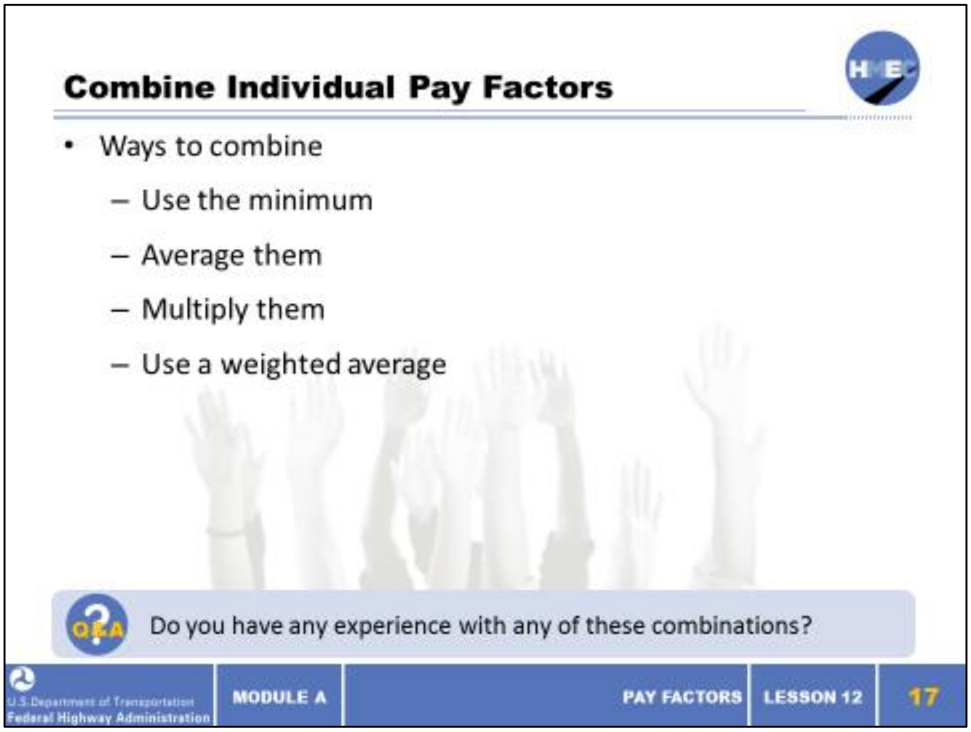
For the red curve, when the contractor produces exactly at the AQL and the estimated quality level is higher, the pay factor will exceed 100%. However, when production is at the AQL and the estimated quality level is lower, the pay factor will be less than 100%. This allows the lower payment to be offset by the higher payment and the average pay factor to be 100%. But for the green curve, since the maximum pay is 100%, the lower payment cannot be offset by a higher payment, so the average pay cannot be 100%.

Slide 16



The ideal situation is to have a performance model that can predict long-term performance of the pavement. In such a case, the quality characteristic measurements can be input into to the model and the payment factor can be based on the predicted performance of the in-place pavement as compared to the desired performance. Unfortunately, such models are either not available or are not widely accepted at this time. Therefore, other methods for determining the composite payment factor are currently in use.

Slide 17



Here are some of the approaches that have been used in combining individual pay factors:

- Use the minimum;
- Average them;
- Multiply them; or
- Use a weighted average.

____
____
____
____
____
____
____
____

Slide 18

Example of Combining Pay Factors

Strength	Permeability	Thickness	Min.	Avg.	Mult.
1.00	1.00	1.00	1.00	1.00	1.00
1.05	1.05	1.05	1.05	1.05	1.16
0.80	0.80	0.80	0.8	0.80	0.51
1.00	0.80	1.05	0.8	0.95	0.84

Q&A Which result do you think is the most appropriate?

U.S. Department of Transportation Federal Highway Administration | MODULE A | PAY FACTORS | LESSON 12 | 18

The composite pay factor determined by various methods are quite different depending on the magnitude of each individual pay factor. In this example, the average approach is the most stable and the approach using multiplication is the most variable.

____

____

____

____

____

____

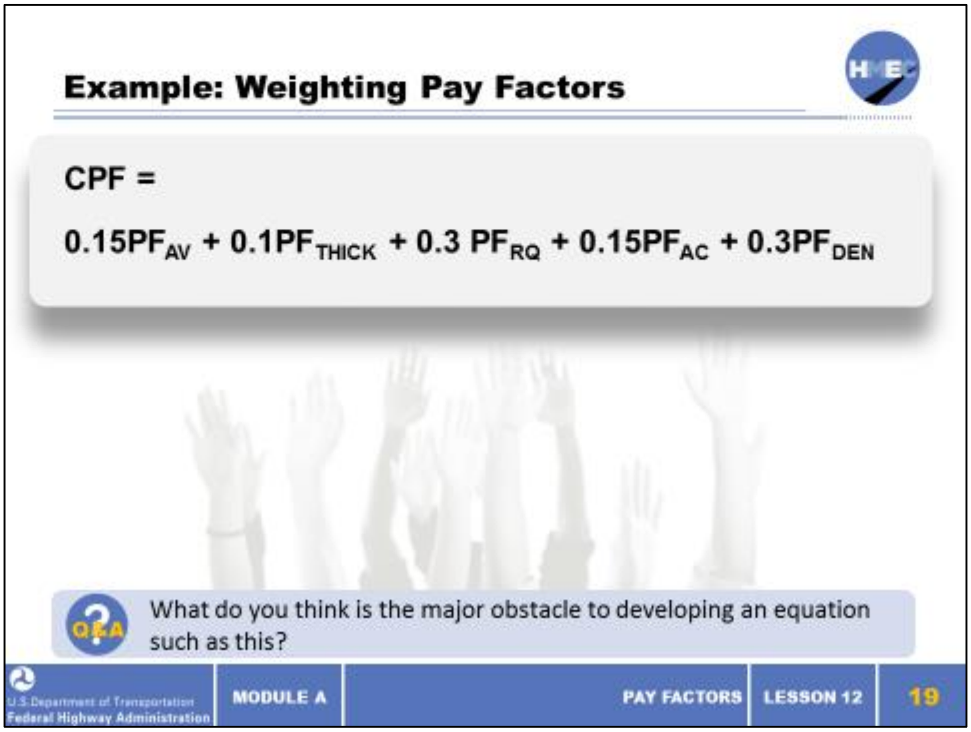
____

____

____

____

Slide 19

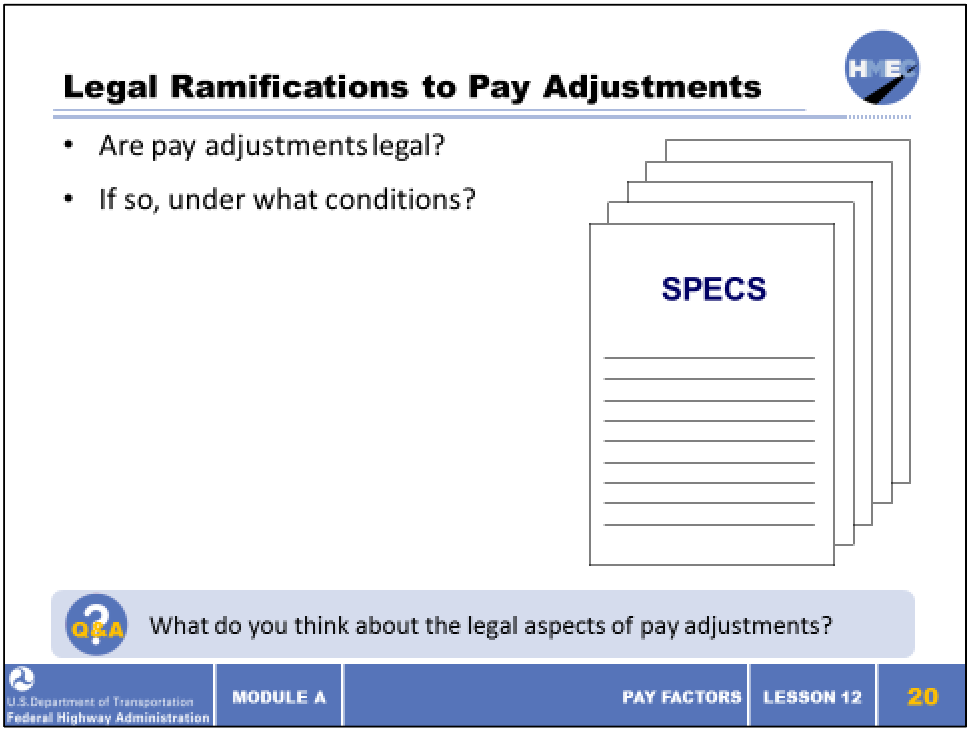


The recent NCHRP study referenced earlier provides examples of the weighted pay factor concept. Some quality characteristics are weighted as more important than others. For HMA, when mixture and construction properties are tied together for acceptance, the in-place air voids and smoothness (RQ) are weighted more heavily than the mixture properties.

$CPF = 0.15(PF_{AV}) + 0.1(PF_{THICK}) + 0.3(PF_{RQ}) + 0.15(PF_{AC}) + 0.3(PF_{DEN})$

Where:
CPF = Composite pay factor
- AV = Lab compacted air voids
- THICK= Voids in mineral aggregates
- RQ = Ride quality
- AC = Asphalt content
- DEN = Field in-place density

Slide 20



Because pay adjustments are viewed by the contracting industry as primarily penalties, their legality has been questioned. And if they are used, what conditions should prevail in their usage? A 2014 decision in a New Jersey Appellate Court upheld a judge's decision that the pay adjustments used in the specifications were valid. The contractor claimed that the smoothness requirements as measured by the International Roughness Index (IRI) were not obtainable on the urban streets with manholes, water valves, catch basins, and intersections. The pay schedule had both positive and negative pay adjustments, but the net result was a negative pay adjustment of approximately 5%. The court found that "the contracts permitted the negative pay adjustments" and that "bidders were required to take note of conditions that might be inconsistent or problematic and were required, prior to bid, to inform (the agency) of such problems that might affect performance of the contract."

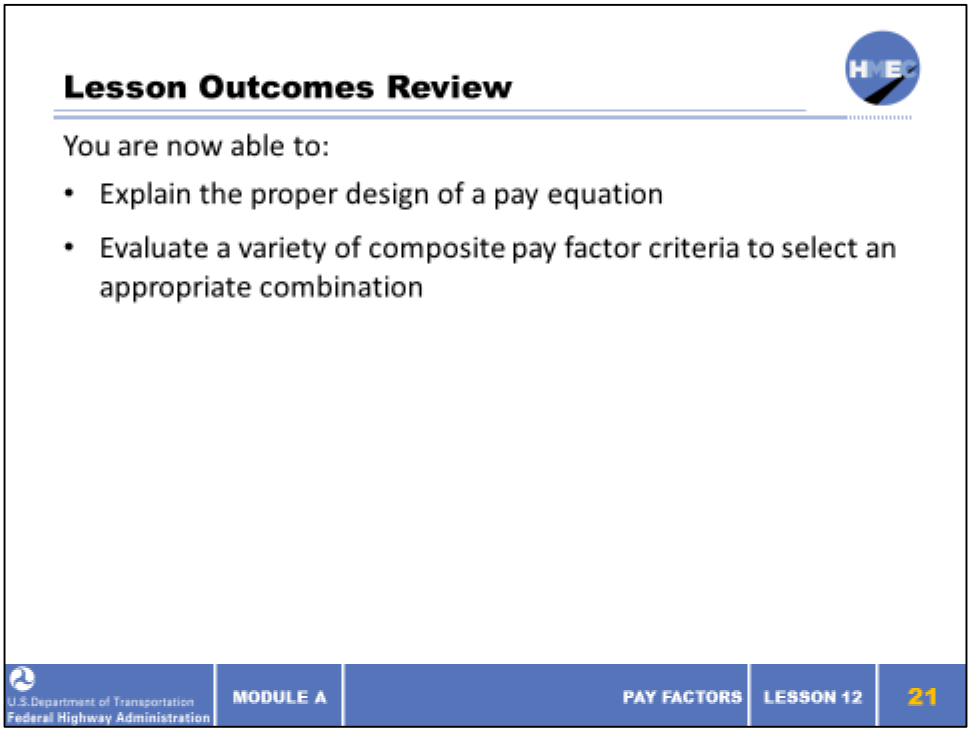
____

____

____

____

Slide 21

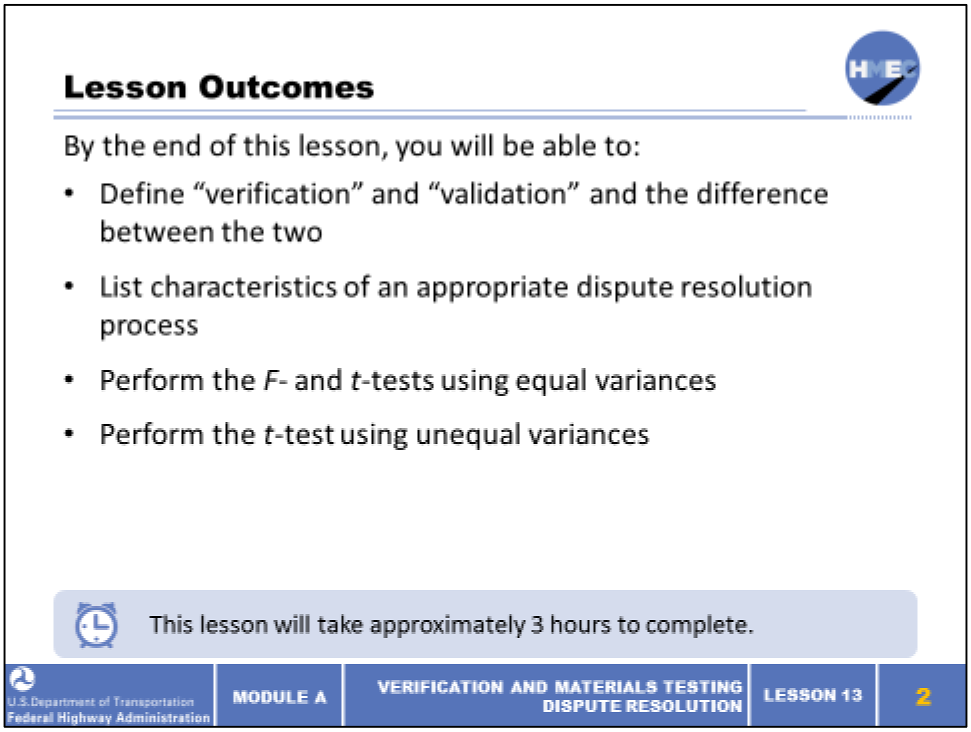

Lesson Outcomes Review
You are now able to:
• Explain the proper design of a pay equation
• Evaluate a variety of composite pay factor criteria to select an appropriate combination
U.S. Department of Transportation
Federal Highway Administration
MODULE A
PAY FACTORS
LESSON 12
21

This page is intentionally blank.

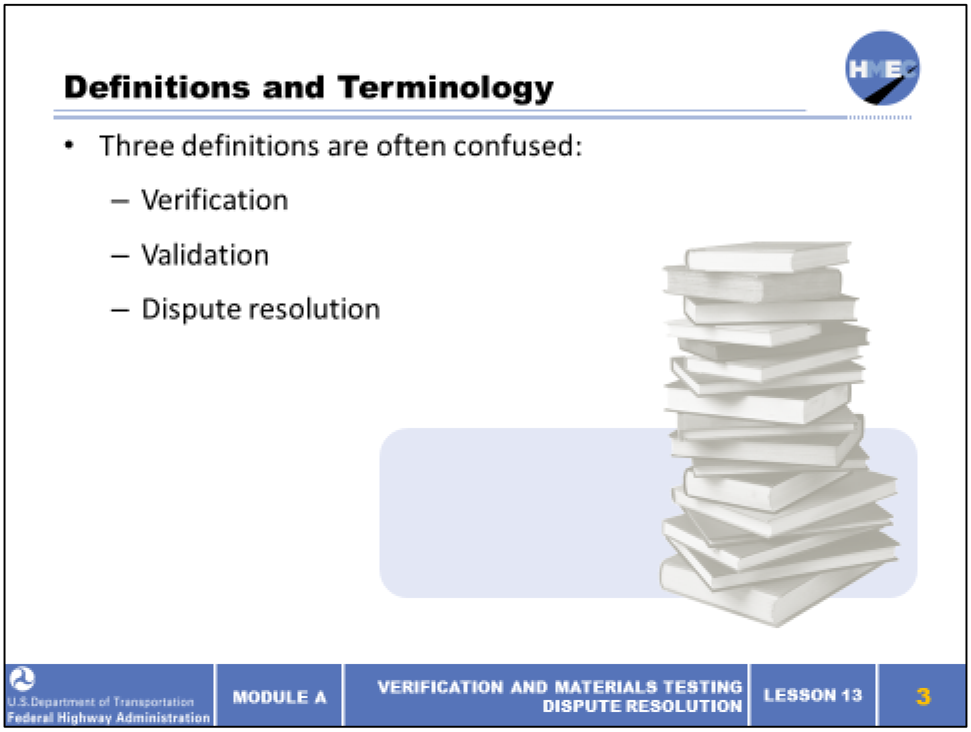
Slide 1


ILT
HMEC
Highway Materials Engineering Course
Lesson 13: Verification and Materials Testing Dispute Resolution
Quality Assurance
U.S.Department of Transportation
Federal Highway Administration
MODULE
A

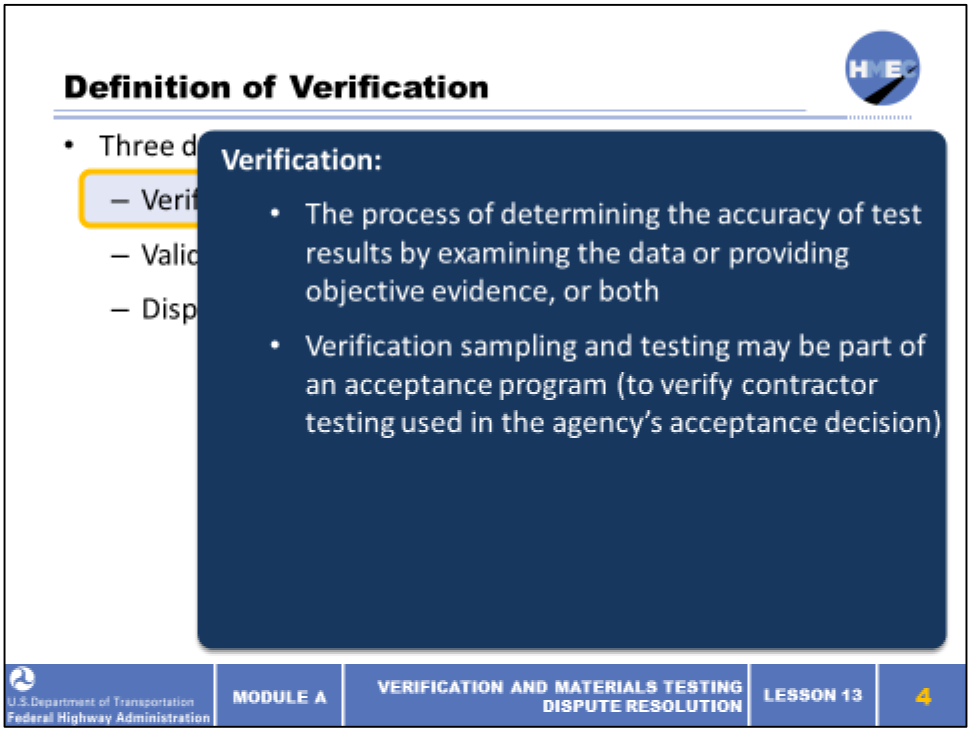
Slide 2


Lesson Outcomes
HMEC
By the end of this lesson, you will be able to:
• Define "verification" and "validation" and the difference between the two
• List characteristics of an appropriate dispute resolution process
• Perform the F- and t-tests using equal variances
• Perform the t-test using unequal variances
This lesson will take approximately 3 hours to complete.
U.S. Department of Transportation Federal Highway Administration
MODULE A
VERIFICATION AND MATERIALS TESTING DISPUTE RESOLUTION
LESSON 13
2

Slide 3

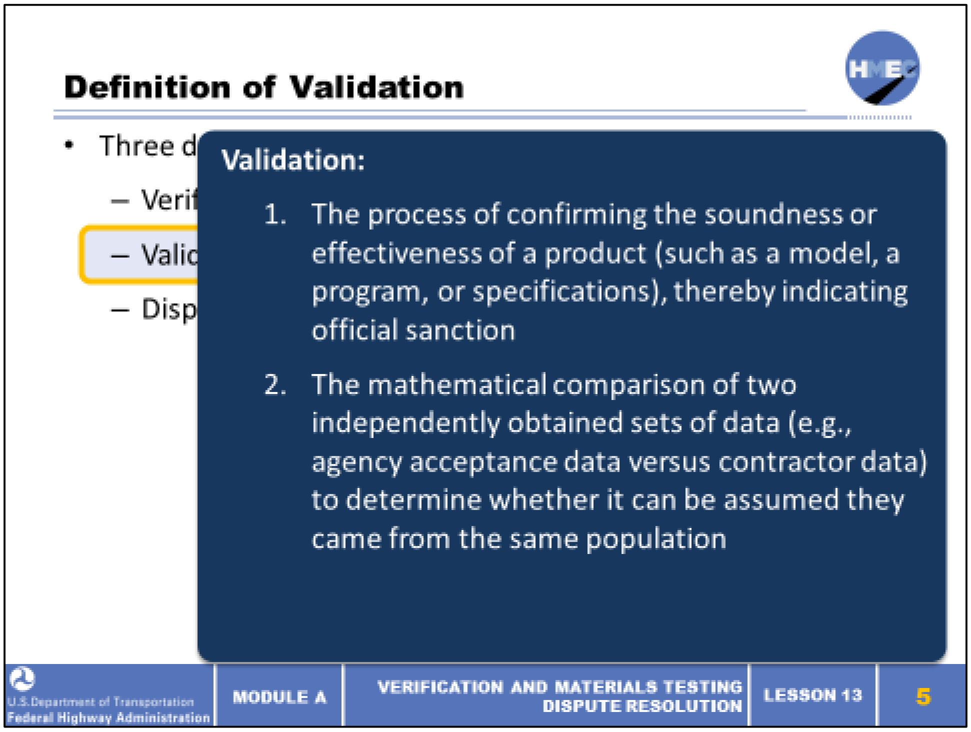

Definitions and Terminology
HMEC
• Three definitions are often confused:
– Verification
– Validation
– Dispute resolution
U.S. Department of Transportation
Federal Highway Administration
MODULE A
VERIFICATION AND MATERIALS TESTING DISPUTE RESOLUTION
LESSON 13
3

Slide 4



Verification is the process of determining the accuracy of test results by examining the data or providing objective evidence, or both. Verification sampling and testing may be part of an acceptance program (to verify contractor testing used in the agency's acceptance decision).

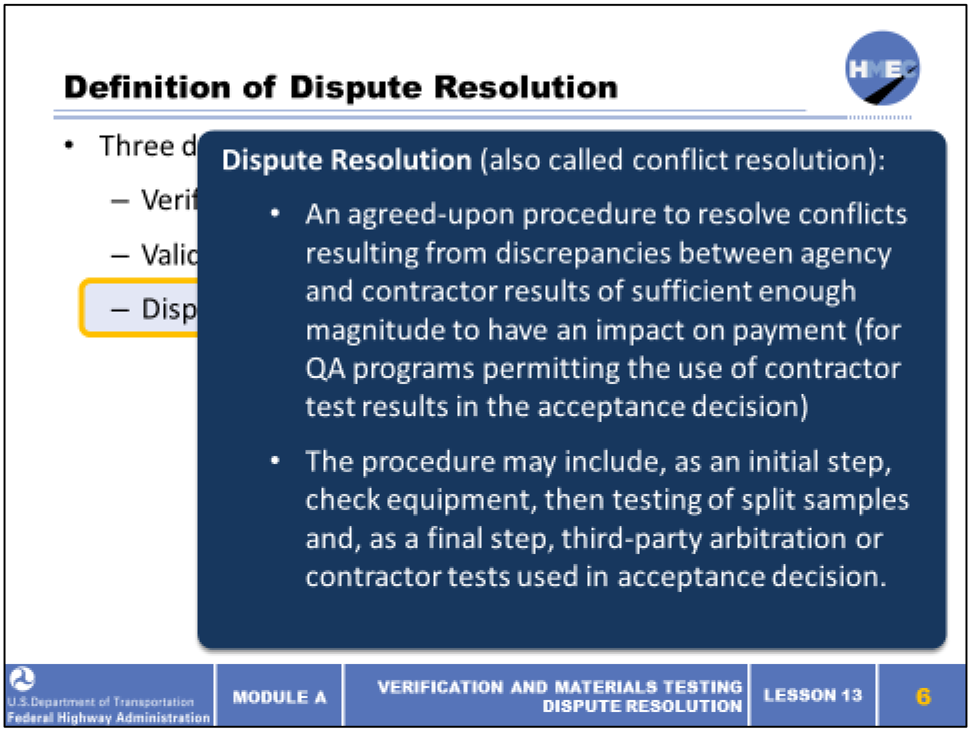
Slide 5



Validation has two definitions. (1) The process of confirming the soundness or effectiveness of a product (such as a model, a program, or specifications), thereby indicating official sanction; (2) The mathematical comparison of two independently obtained sets of data (e.g., agency acceptance data versus contractor data) to determine whether it can be assumed they came from the same population.

This lesson will primarily focus on the second definition.

Slide 6

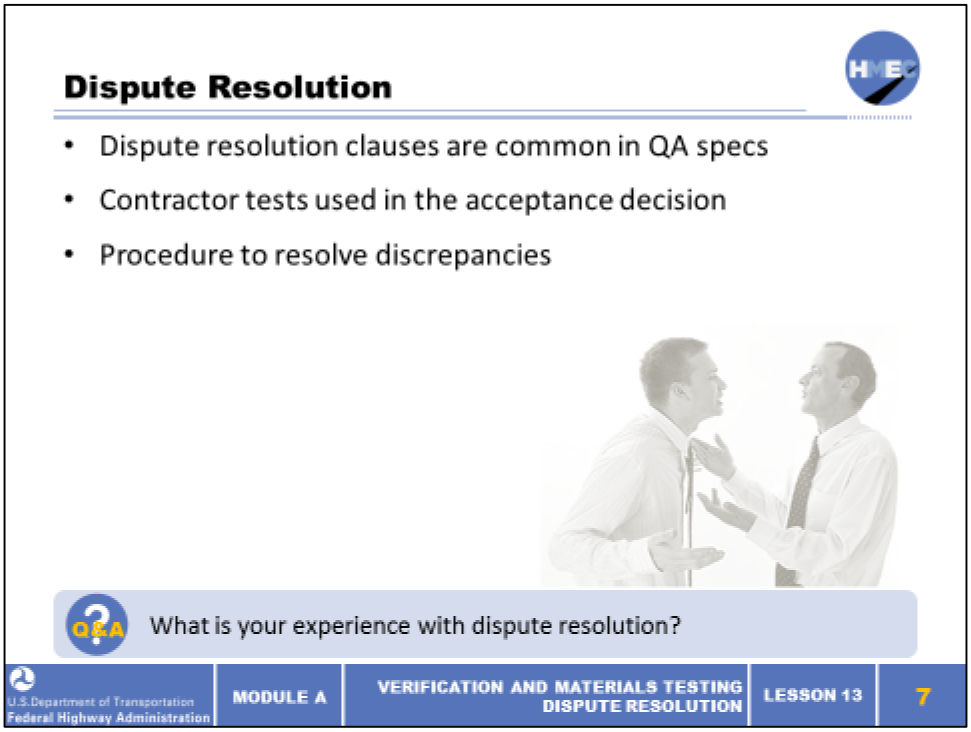


Dispute resolution, or conflict resolution, is an agreed-upon procedure to resolve conflicts resulting from discrepancies between agency and contractor test results of sufficient enough magnitude to have an impact on payment (for quality assurance, QA, programs permitting the use of contractor test results in the acceptance decision).

The procedure may include, as an initial step, check equipment, then testing of split samples and, as a final step, third-party arbitration or contractor tests used in acceptance decision.

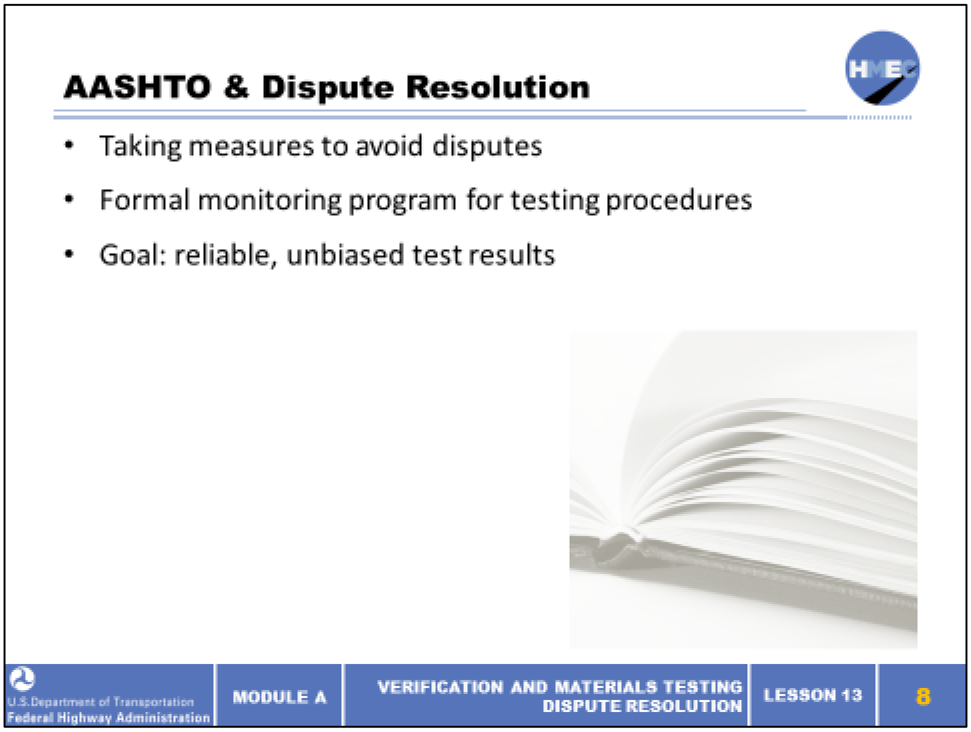
It is easy to see why these definitions may be confusing.

Slide 7



They provide a procedure to resolve conflicts resulting from discrepancies of sufficient magnitude between the agency's and contractor's test results that have an impact on payment, and are a vital part of FHWA Quality Assurance Program, 23 CFR 637b.

Slide 8



This document contains a section on dispute resolution that briefly suggests that appropriate measures should be taken to avoid disputes arising from differences in test results whenever possible. Also, a formal monitoring program should be established that will provide assurances to both the contractor and the agency that all data are reliable, unbiased, and truly indicate the product quality.

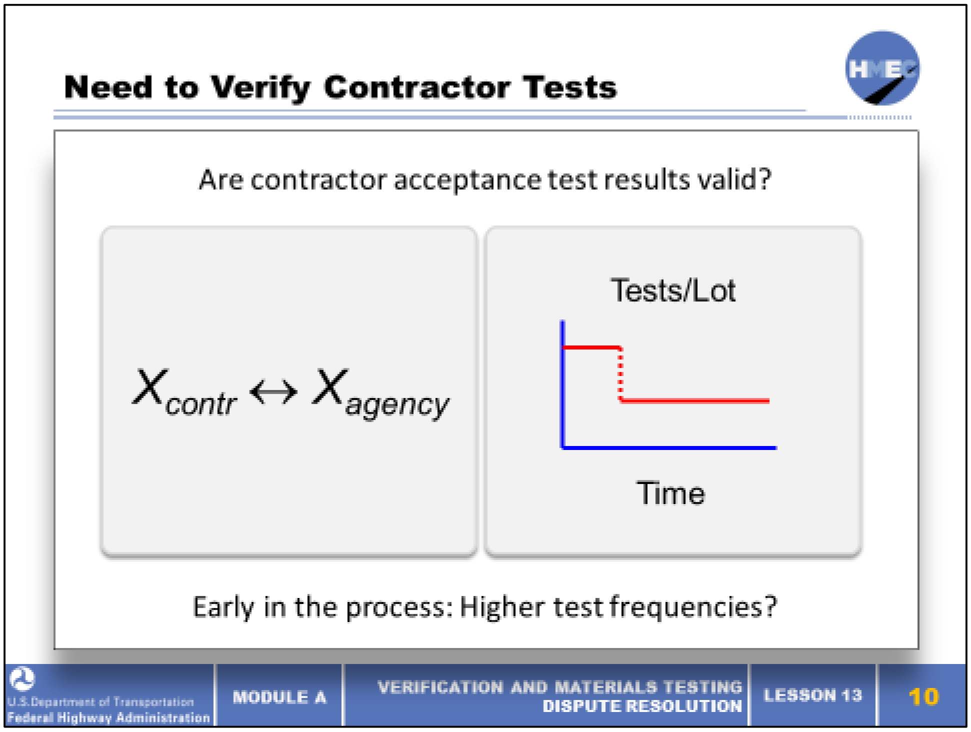
Some of the elements suggested are:

- All sampling and testing should be performed by qualified testing personnel and laboratories.
- Pre-production correlation testing involving comparison of split sample results between the contractor's and agency's laboratories should be conducted.
- All testing labs should participate in a national or agency-sponsored round robin testing program.
- All samples obtained by either party should be selected under an approved random sampling method.

Slide 9

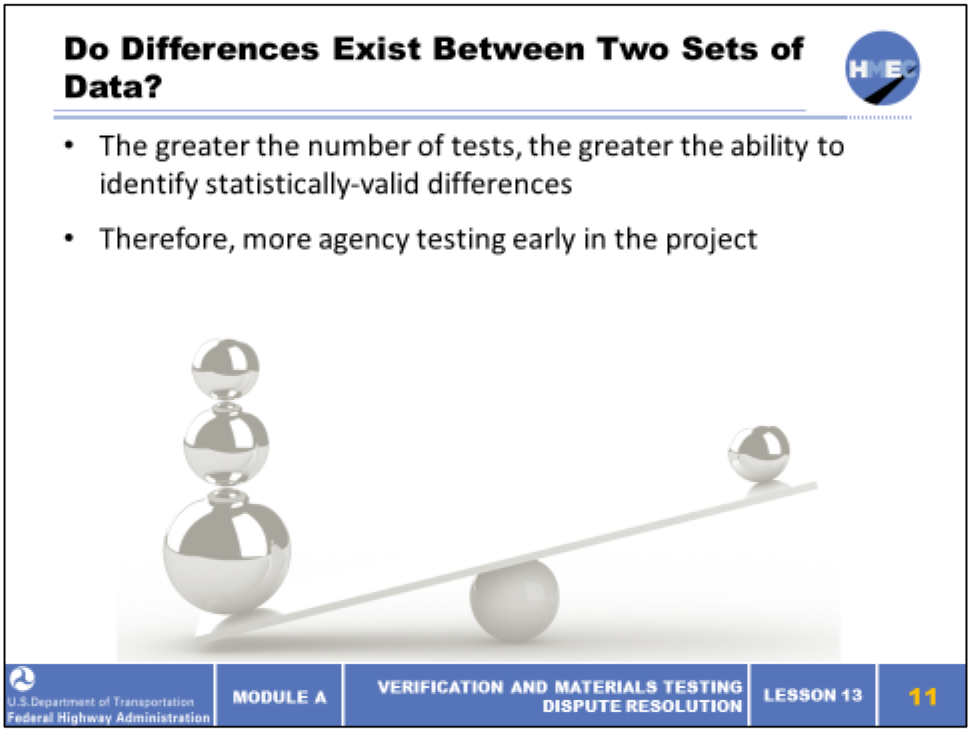
The agency may decide to do the acceptance testing, may assign the testing to the contractor, or may require a designated agent to do the testing. FHWA Quality Assurance Program, 23 CFR 637b states that agencies "will permit the use of data from contractors' quality control sampling and testing programs in acceptance programs if the results from the States' verification sampling and testing programs confirm the quality of the material." (Note that the CFR indicates that data from quality control, QC, results can be used in acceptance decisions. Some take issue with this wording, believing that the QC and acceptance functions should be separated regardless of who does the acceptance testing.) The verification procedure requires a comparison between contractor and agency test results and is another use for hypothesis testing.

Slide 10



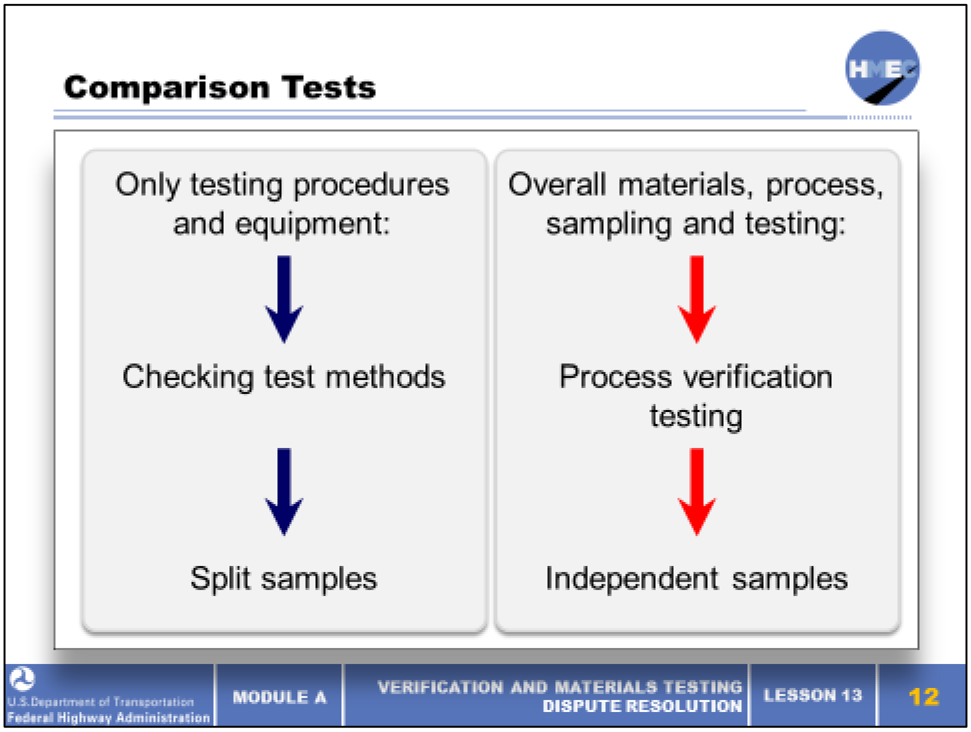
The topic of hypothesis tests was discussed in Lesson 11 as they relate to risks; however, another use of them can be for comparing tests between two labs. During the initial production (start-up) for any process, it is important to obtain an early estimate of the product quality. When there is no prior knowledge of the particular process, it may be necessary to adjust testing frequencies to reduce the time needed to produce a reliable early estimate of the degree of contract compliance. This will accomplish not only the desirable early estimate of contract compliance in a statistically valid manner, but will also produce confidence in the sampling and testing program (if statistical analysis indicates that the agency's tests and contractor's tests represent the same population). If the analysis indicates that the results likely came from different populations, the sources of difference in the test results should be determined and corrected as soon as possible.

Slide 11



The greater the number of tests, the greater the ability of the procedure to identify statistically valid differences. This is the reason for using a larger amount of agency testing early in a project. In this way, more data will be available to make a determination regarding whether the contractor testing is consistent with the agency tests.

Slide 12



The type of sample taken for comparison depends upon the agency's objective. The difference in information provided between split and independent samples is very important here. Each testing procedure has a specific use that doesn't extend to the other.

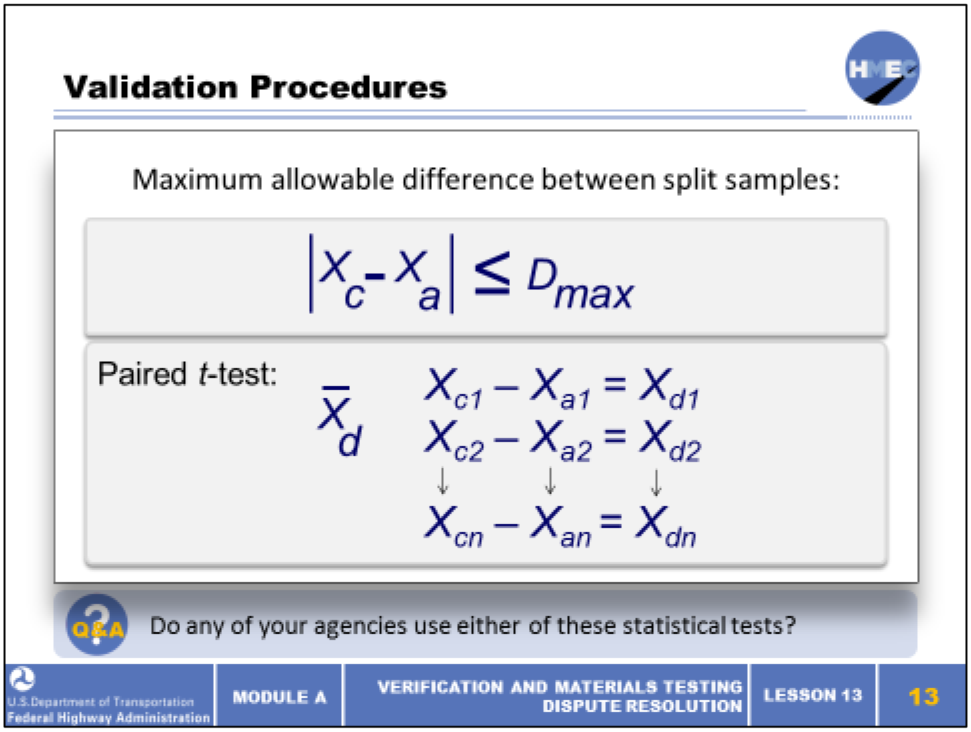
If the agency's objective for this comparison is simply to identify discrepancies in the testing procedures or equipment, the statistical testing should be done on split samples. There are two ways of doing this with one being much more powerful than the other.

If the agency's objective for this comparison is to identify discrepancies in the overall material, process, sampling, and testing processes, then validation testing must be done on independent samples, not split samples. Split samples do not meet the regulatory requirements for verification, but are used in the independent assurance program.

Whether split or independent samples are used, until initial verification or validation testing has been done, it is recommended that only agency test results be used for contract compliance determination. Once the validation is successfully accomplished, the agency can then reduce its testing frequency and, if appropriate, use contractor tests in the acceptance decision.

____
____
____
____
____
____
____
____
____
____

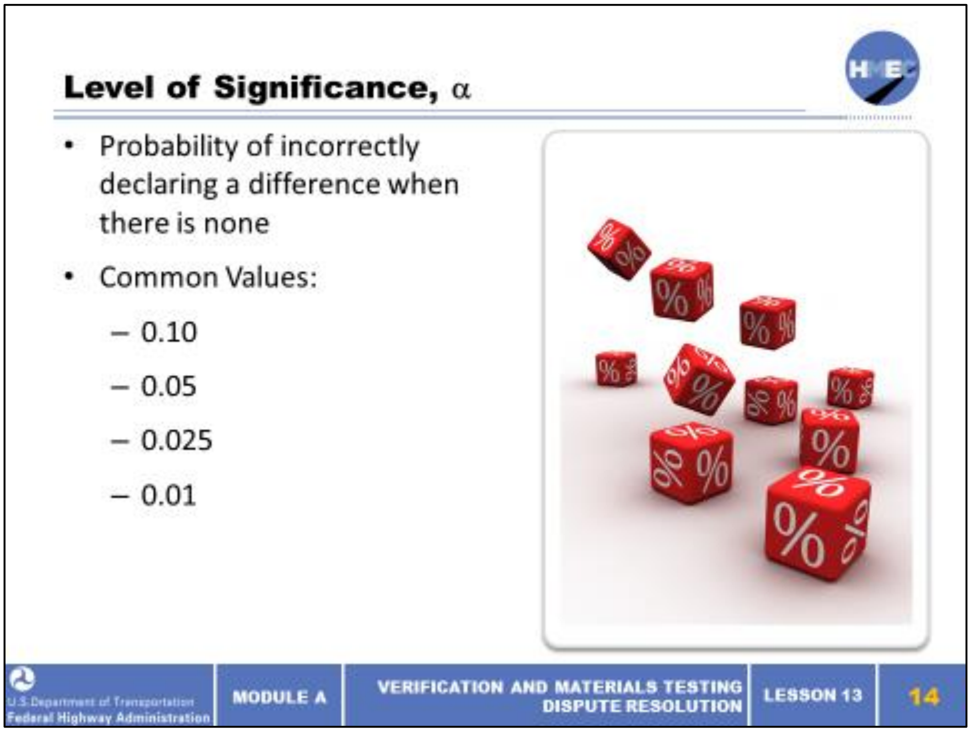
Slide 13



One procedure uses the absolute difference between two tests from split samples (one tested by the contractor and one by the agency) and compares this difference to some allowable difference established for the specific test method. These allowable differences might be taken from an ASTM or AASHTO precision statement for the test procedure, or they might be developed based on split sample data collected by the agency. A system such as this has been employed when comparing independent assurance test results.

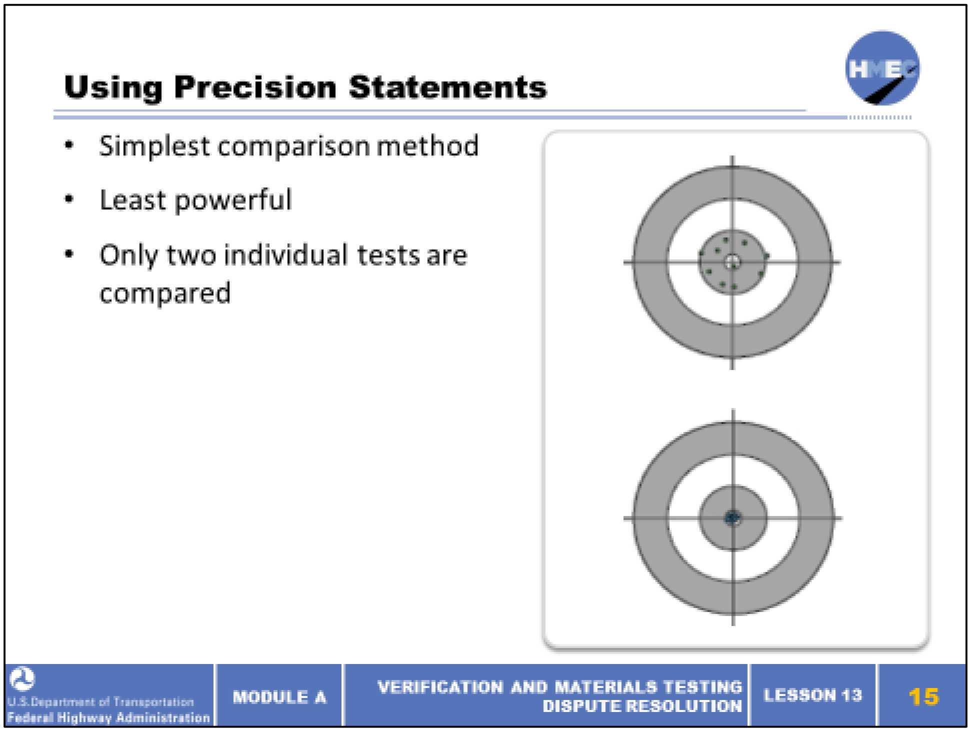
However, not all test procedures have precision statements. In this case (and in the case where it is desirable to compare more than one contractor test and agency test from split samples), a second statistical process, the *t*-test for paired measurements (referred to as a paired *t*-test), can be used. Both of these procedures are based on the same underlying principle. However, since more data are employed, the paired *t*-test is more powerful at detecting differences when they exist. Each of these procedures is discussed on the slide.

Slide 14



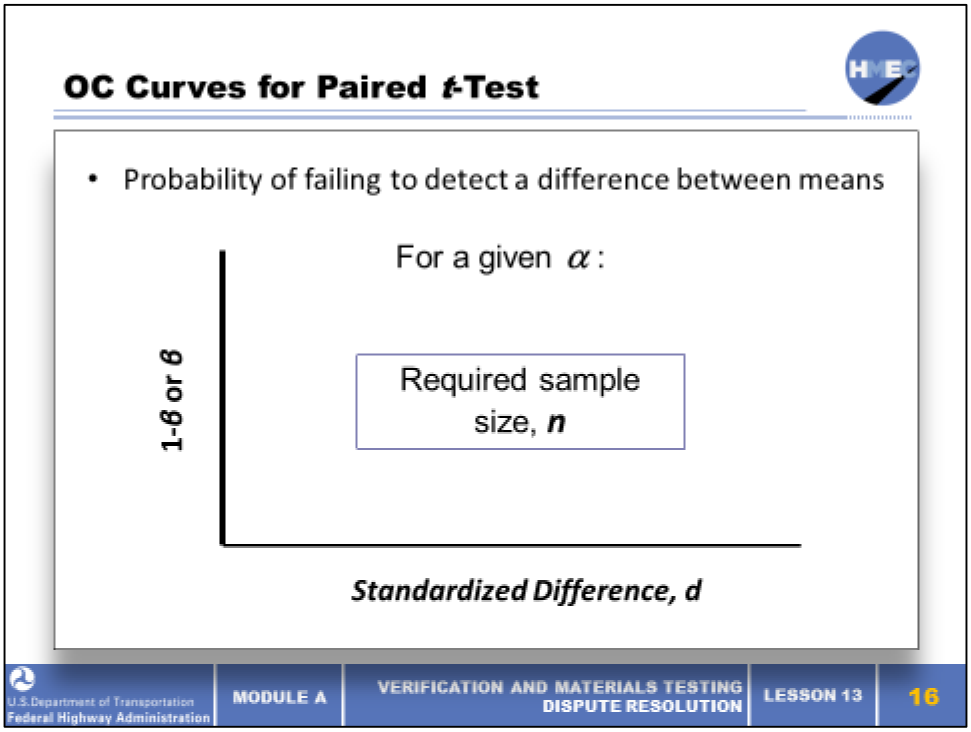
A level of significance, α, must be selected before comparing contractor and agency samples. While α values of 0.10, 0.05, 0.025, and 0.01 are common, many agencies select a value of 0.01 to minimize the likelihood of incorrectly concluding that the results are different when they actually came from the same population. However, it should be recognized that selecting a low α value reduces the chance of detecting a difference when one actually exists. In an effort to compromise these risks, FHWA recommends an α of 0.025.

Slide 15



It's the least powerful because it includes a comparison between only two test values and is best suited when starting the checking process.

Slide 16



The beneficial use of operating characteristic (OC) curves was discussed in Lesson 11 when we wanted to analyze the influence of sample size on acceptance plans and its use is appropriate here. But the question arises, how many pairs of test results should be used? This is where an OC curve is helpful. This form of OC curve, for a given level of α, plots on the vertical axis the probability of either not detecting β, or detecting $1 - \beta$, a difference between two populations. The standardized difference, *d*, between the two population means is plotted on the horizontal axis.

____

____

____

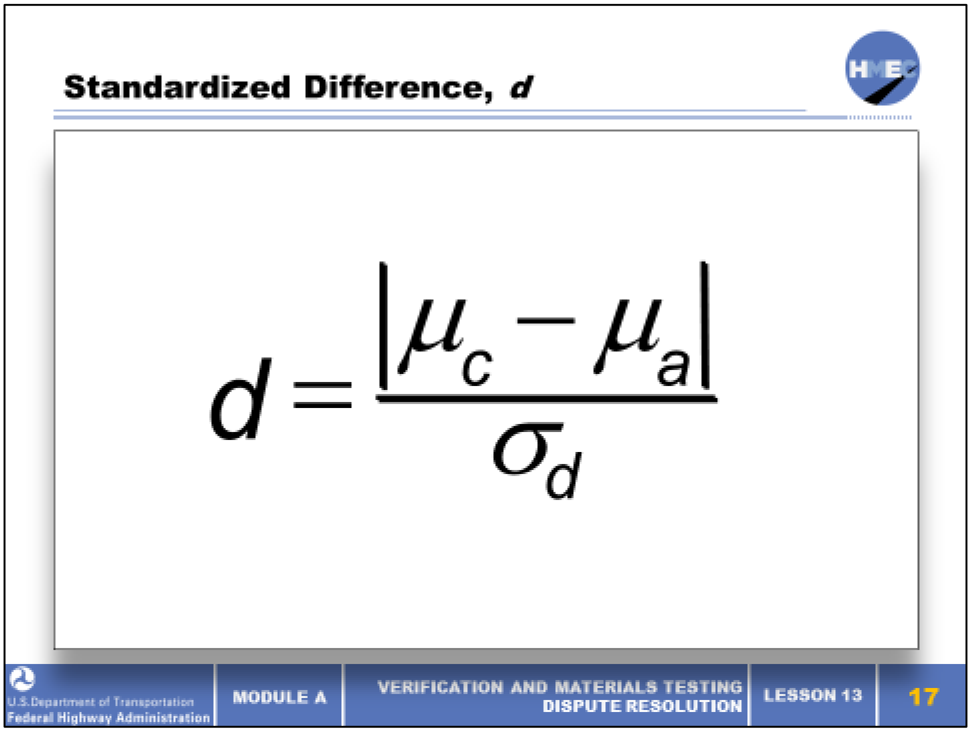
____

____

____

____

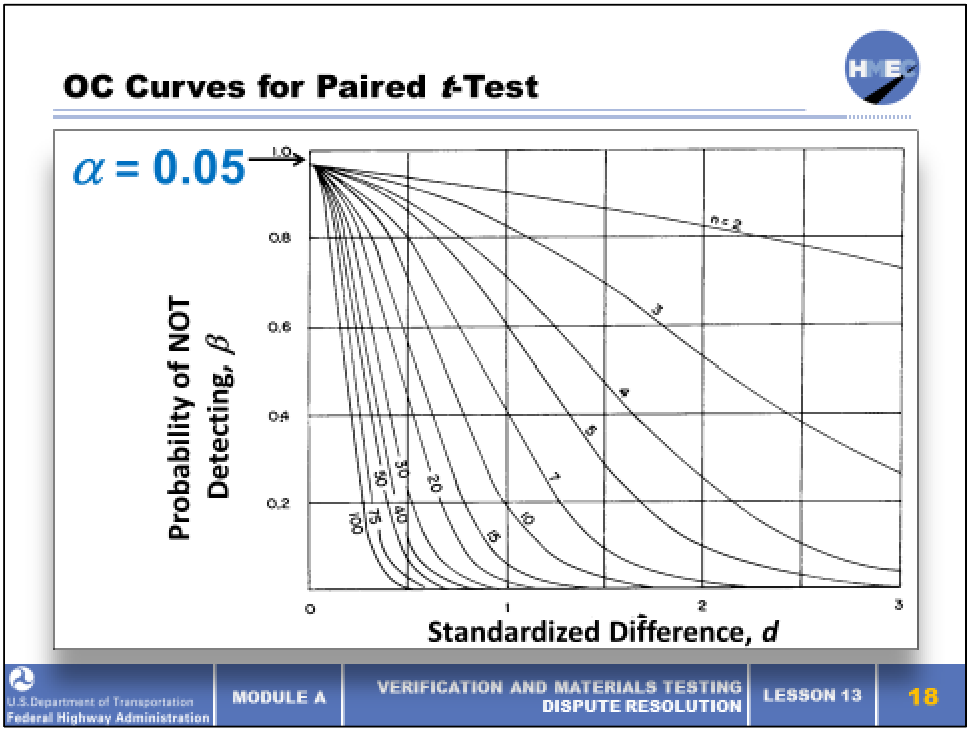
Slide 17



The standardized difference, *d*, is measured as: $d = |\mu_c - \mu_a| / \sigma_d$ where:

- $|\mu_c - \mu_a|$ = the true absolute difference between the mean, μ_c, of the contractor's test result population (which is unknown) and the mean, μ_a, of the agency's test result population (which is unknown); and
- σ_d = the standard deviation of the true population of signed differences between the paired tests (which is unknown).

Slide 18



It is evident from the OC curves that for any probability of not detecting a difference, β, (value on the vertical axis), the required *n* will increase as the difference, *d*, decreases (value on the horizontal axis). In some cases, the desired β or difference may require prohibitively large sample sizes, which we'll discuss in a moment.

In that case, a compromise must be made between the discriminating power desired, the cost of the amount of testing required, and the risk of claiming a difference when none exists. To use these OC curves, the true standard deviation of the signed differences, σ_d , is assumed known, (or approximated based on published literature). After experience is gained with the process, σ_d can be more accurately defined and a better idea of the required number of tests determined.

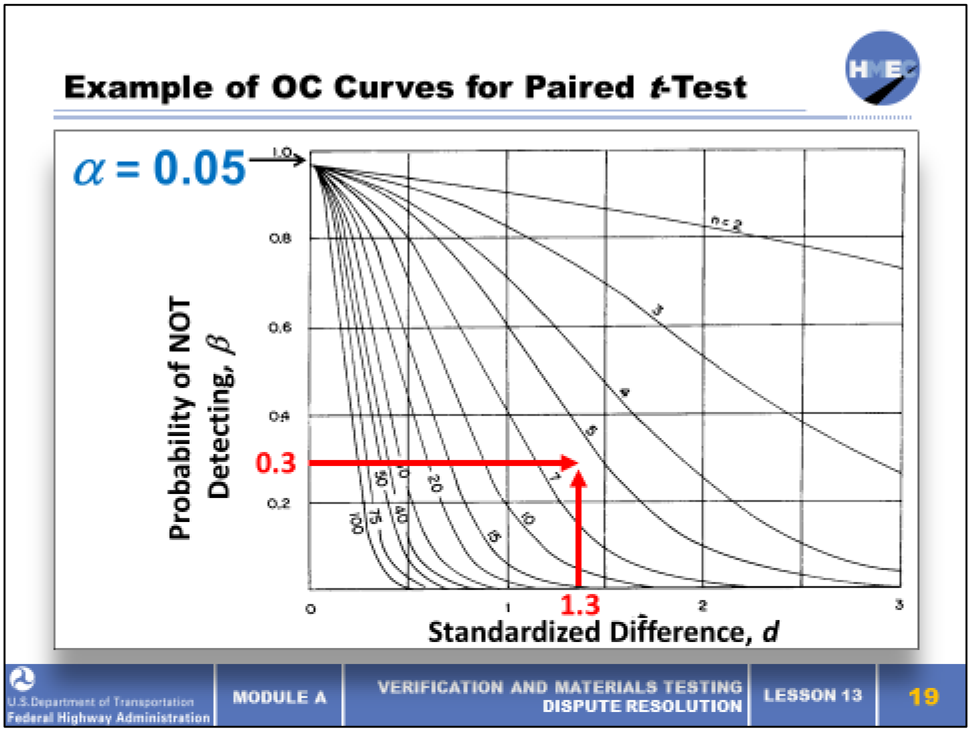
Larger versions of the OC curves for paired *t*–tests for α values of 0.05 and 0.01 appear in Figures 10–1 and 10–2, respectively, in the Resources folder on your tablet.

____

____

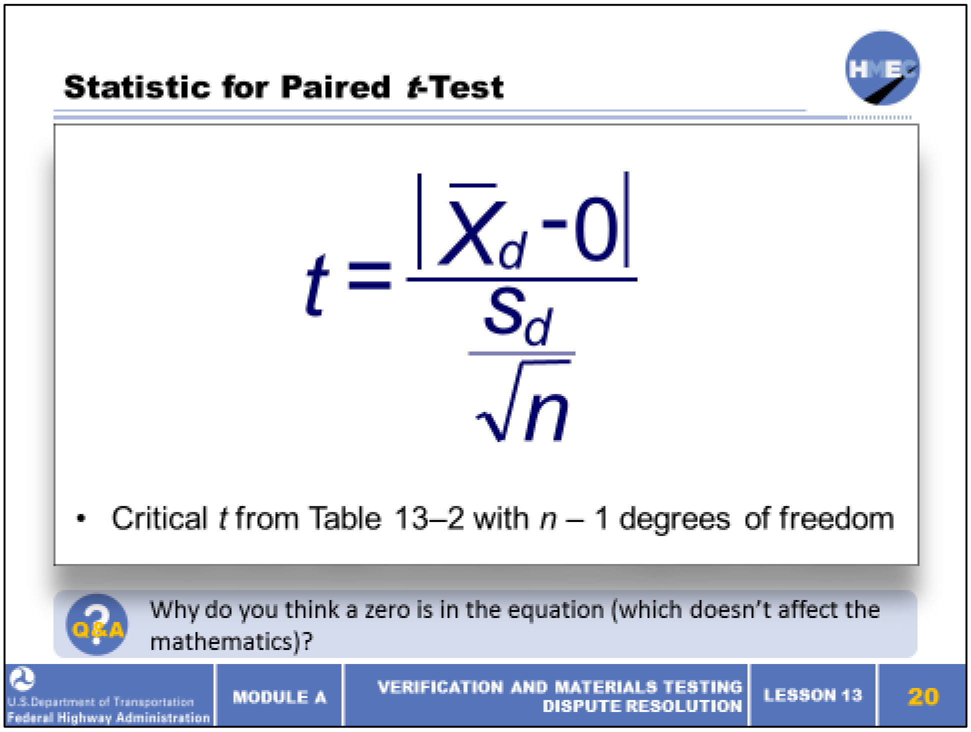
____

Slide 19



The number of pairs of split samples for verification of laboratory-compacted air voids using the Superpave Gyratory Compactor (SGC) is desired. The probability of not detecting a difference, β, is chosen as 30% or 0.30. (Some OC curves use 1-b, known as the "power of the test") on the vertical axis, but the only difference is the scale change, with 1-β in this case being 70%). Assume the absolute difference between μ_c and μ_a should not be greater than 2.0% and that the standard deviation of the differences, σ_d, using the SGC is 1.5%. Thus, $d = \left| \mu_c - \mu_a \right| / \sigma_d$ = 2.0/1.5 = 1.3. Reading these values on the horizontal scale and a β of 0.3 on the vertical scale shows that about six paired split *t*-tests are necessary for comparison.

Slide 20



In the example in the last slide, it may be feasible to take six paired samples the first day of testing, thereby providing sufficient data to perform the *t*-test on the second day. After the split samples are collected and tested, the *t*-test for paired measurements can be performed. This test uses the difference between each pair of tests and determines whether the difference is statistically different from 0. Thus, it is the difference *within* pairs, not between pairs, that is being tested. The form of the *t*–test for paired measurements is: $t = |X\text{-}bar_d - 0| / s_d$ /square root *n*, where:

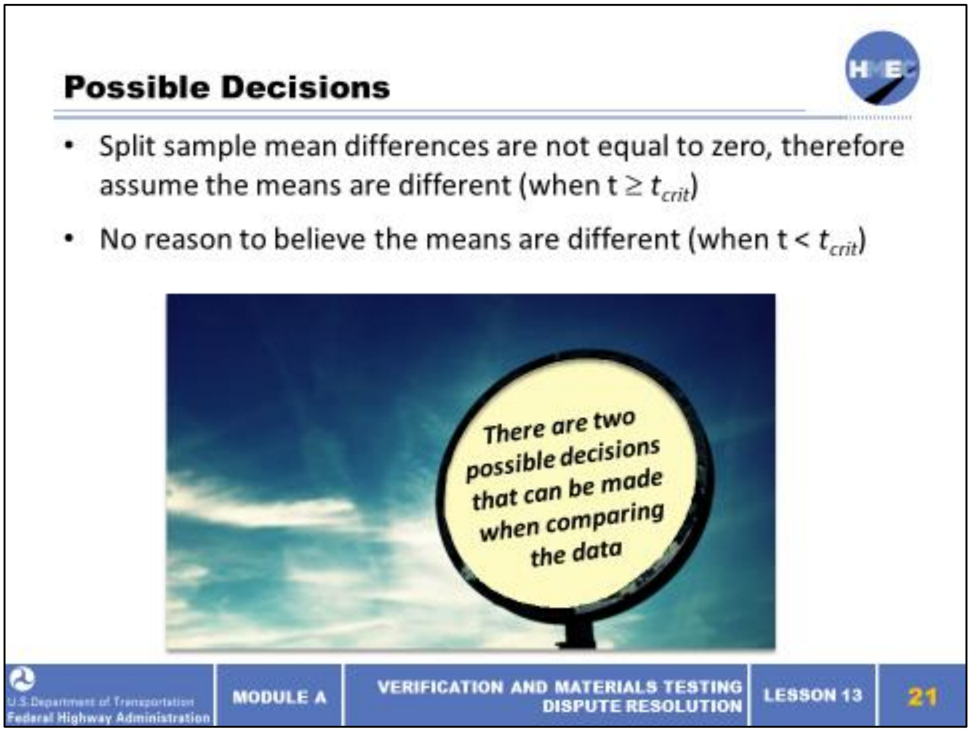
- $X\text{-}bar_d$ = mean of the differences between the split sample test results;
- s_d = standard deviation of the differences between the split sample test results; and
- *n* = number of split samples.

The calculated *t* value is then compared to the critical value, t_{crit}, obtained from Table 13–2 with *n* – 1 degrees of freedom.

____

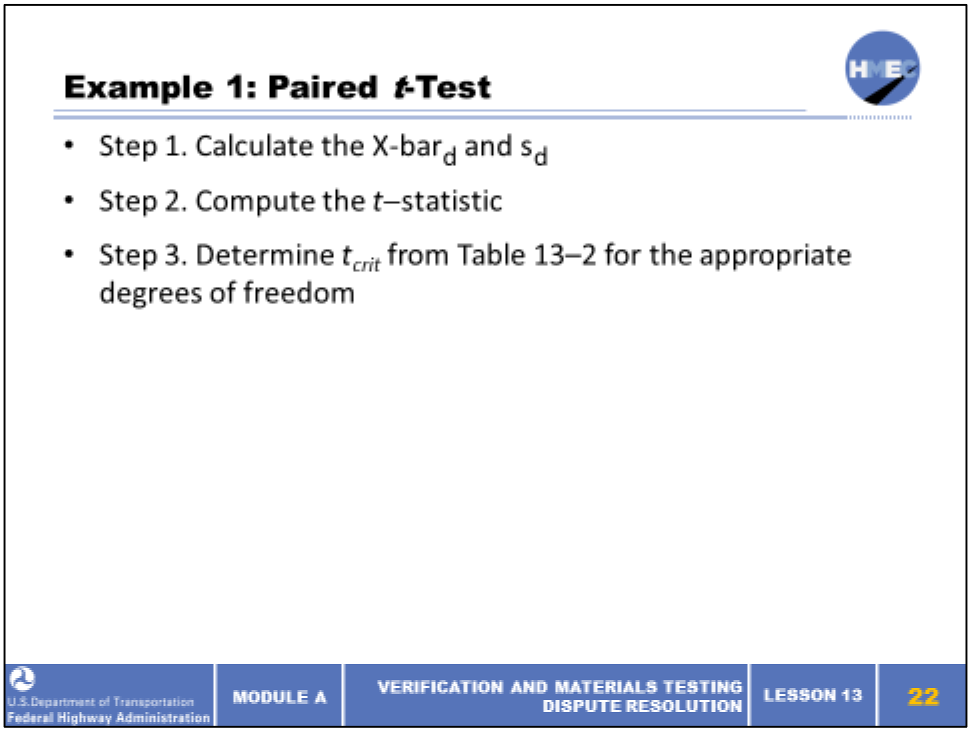
____

Slide 21



The difference between the paired test results is greater than likely to occur from chance if they are actually equal (when $t \geq t_{crit}$). There is no reason to believe that the paired test results are different because the average difference is not so different from 0 as to be unlikely to have occurred from chance if they are actually equal (when $t < t_{crit}$).

Slide 22



The table on the next slide presents the split sample test results for laboratory-compacted air voids. Use a paired *t*-test (a more common way of saying, "*t*-test for paired measurements") to determine whether a difference exists between the contractor and agency results. Before we use the paired t-test equation, we have to calculate the X-bar$_d$ and s$_d$.

Slide 23

Step 1: Calculate X-bar$_d$ and s$_d$

Sample	Contr.	Agency	Diff.
1	4.6	3.8	0.8
2	4.4	3.4	1.0
3	3.7	3.1	0.6
4	4.5	4.1	0.4
5	4.8	3.9	0.9
6	4.0	3.7	0.3

U.S. Department of Transportation Federal Highway Administration | MODULE A | VERIFICATION AND MATERIALS TESTING DISPUTE RESOLUTION | LESSON 13 | 23

The table on the slide presents the split sample test results for laboratory-compacted air voids. Use a paired *t*-test (a more common way of saying, "*t*-test for paired measurements") to determine whether a difference exists between the contractor and agency results. Before we use the paired t-test equation, we have to calculate the X-bar$_d$ and s$_d$.

____

____

____

____

____

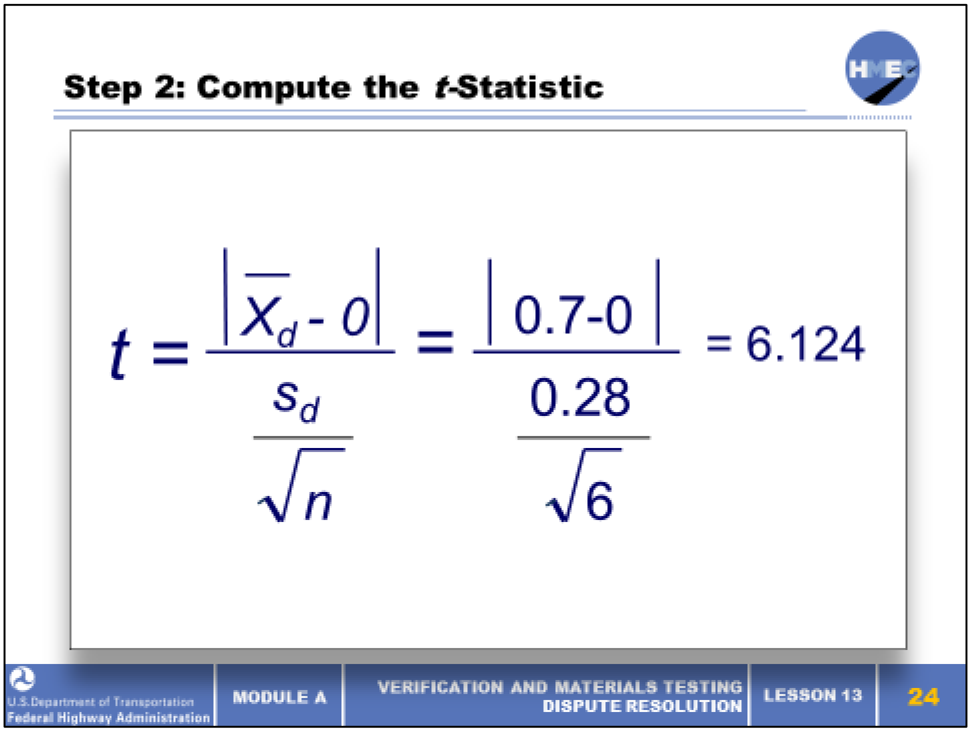
____

____

____

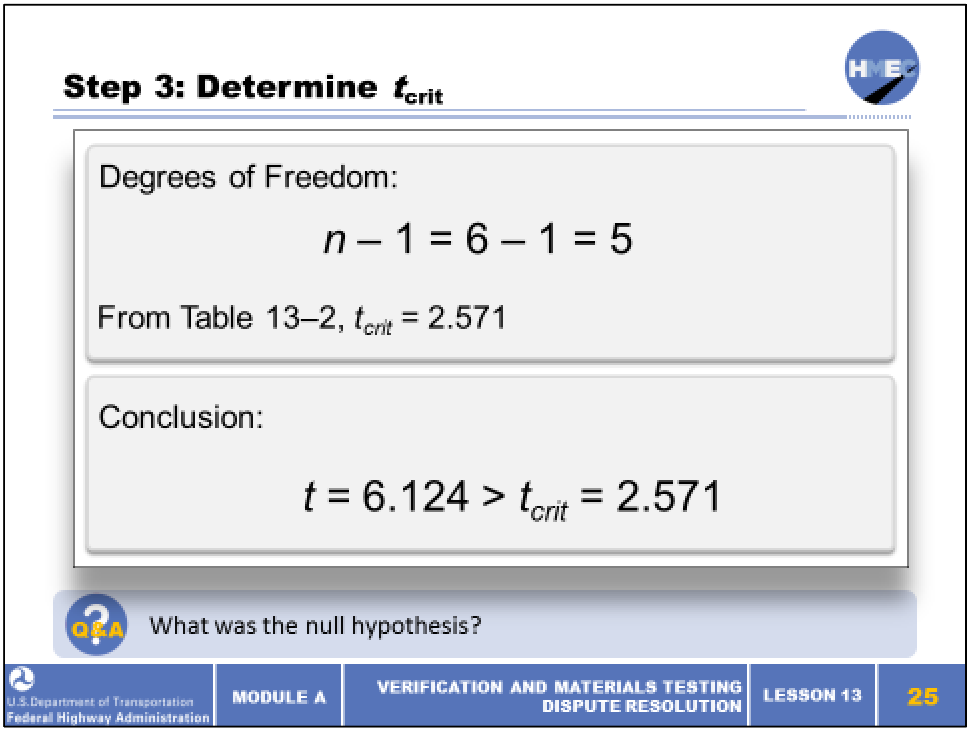
____

Slide 24



First, compute the *t*-statistic for the paired measurements (i.e., the differences between the split sample test results). Next we'll compare the *t*-statistic with t_{crit}.

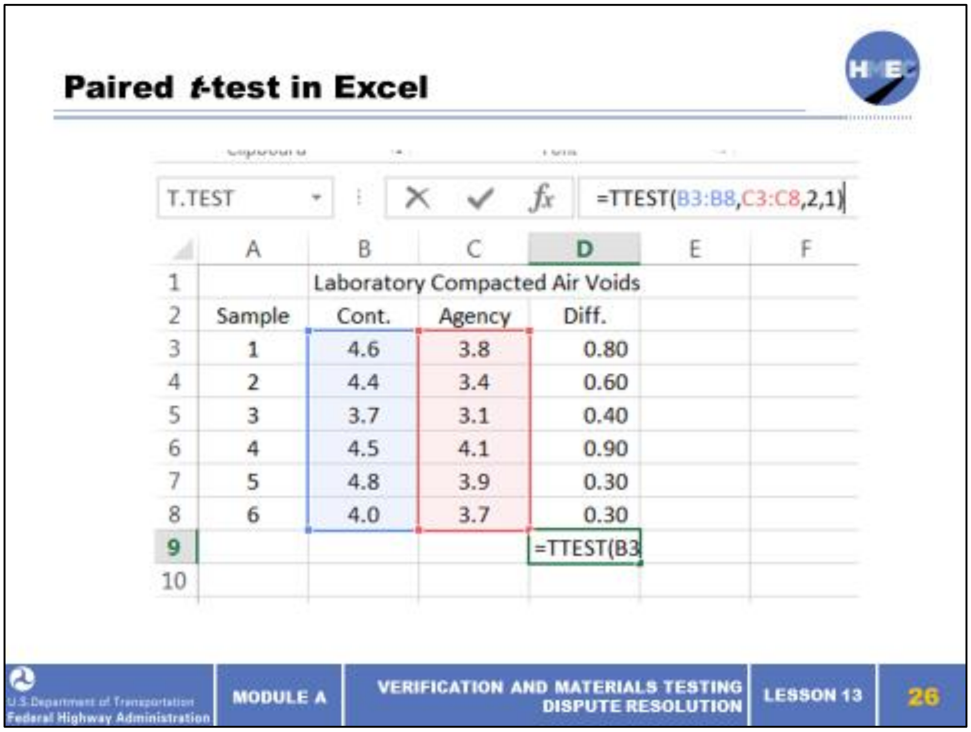
Slide 25



For $n – 1$ (6 – 1) degrees of freedom and $\alpha = 0.05$, (advance slide) the t_{crit} value can be determined from Table 13–2: $t_{crit} = 2.571$.

Conclusion: Since 6.124 > 2.571, we reject the null hypothesis, and assume that the paired tests are different. We therefore assume that the contractor and agency test results from paired measurements indicate that the test method, technicians, and/or test equipment are not providing similar results. Keep in mind that we can conclude nothing about the material or production variation.

Slide 26



Enter the contractor and agency data in the appropriate columns (B, and C, respectively). Although it is not necessary to determine the *t*-value, it is a good idea to complete the Difference column by subtracting the value in column C from the value in column B.

Then, go to cell D-9, click on the formula (f_x) sign and choose TTEST. This will bring up the next screen shot.

____

____

____

____

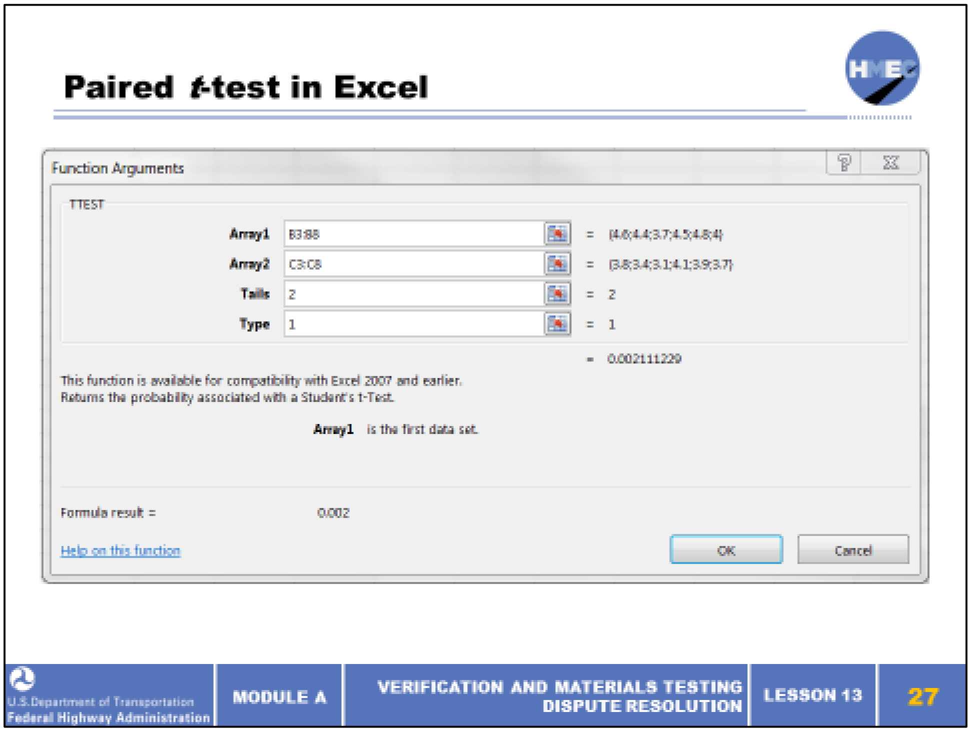
____

____

____

____

Slide 27



Click on Array 1 and (shown on the previous slide) copy the values from the cells for the contractor data to this row. Click on Array 2 and do the same thing for the agency data. Next, click on Tails; the two choices are one-tail distribution, 1, or two tailed distribution, 2. Choose 2. Then, click on Type; the three choices are paired-*t* test, 1, two-sample equal variance, 2, or two-sample unequal variance, 3. Chose 1. This will then enter the result, 0.002, in cell D-9, which is shown on the next screen.

____

____

____

____

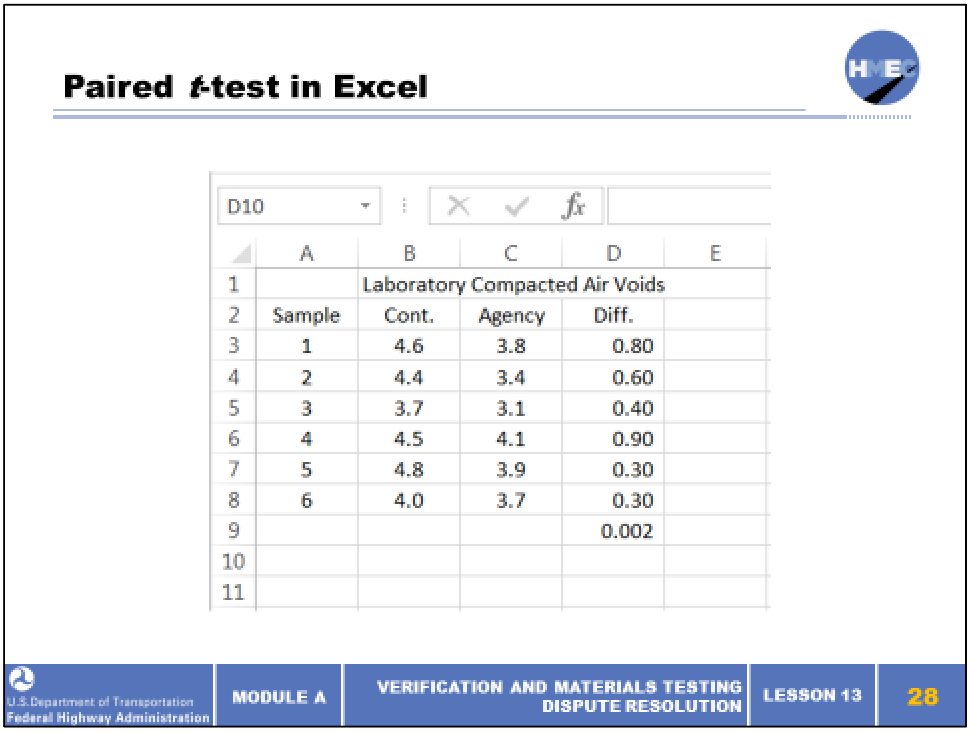
____

____

____

____

Slide 28



	A	B	C	D	E
1		Laboratory Compacted Air Voids			
2	Sample	Cont.	Agency	Diff.	
3	1	4.6	3.8	0.80	
4	2	4.4	3.4	0.60	
5	3	3.7	3.1	0.40	
6	4	4.5	4.1	0.90	
7	5	4.8	3.9	0.30	
8	6	4.0	3.7	0.30	
9				0.002	
10					
11					

The result in cell D-9 = 0.002 is less than 0.05 and we reject the null hypothesis, and assume that the paired tests are different.

Slide 29

Exercise 1: Paired *t*-Test

Sample	Contr.	Agency	Diff.
1	5.2	5.4	
2	5.7	6.0	
3	5.1	6.2	
4	6.0	6.5	
5	5.8	5.7	
6	5.3	5.8	

Step 1	Calculate the X-bar$_d$ and s_d
Step 2	Compute the t–statistic
Step 3	Determine t_{crit} from Table 13–2 for the appropriate degrees of freedom for an α=0.01

Let's break into groups of 2-3 for this exercise. You have 10 minutes to determine if the results are statistically different or not.

U.S. Department of Transportation Federal Highway Administration | MODULE A | VERIFICATION AND MATERIALS TESTING DISPUTE RESOLUTION | LESSON 13 | 29

For the percent passing the 0.075 mm (#200) sieve, it is desirable to verify the contractor's tests. Using Figure 13-2, the OC curve for α = 0.01, it has been determined that six paired measurements must be made for a β, the probability of not detecting a difference, of 30% or 0.30, with a difference between means of 0.75% and an assumed standard deviation of 0.4%. The paired *t*-test results that were obtained are shown in the table on the slide. Are the results statistically different or not?

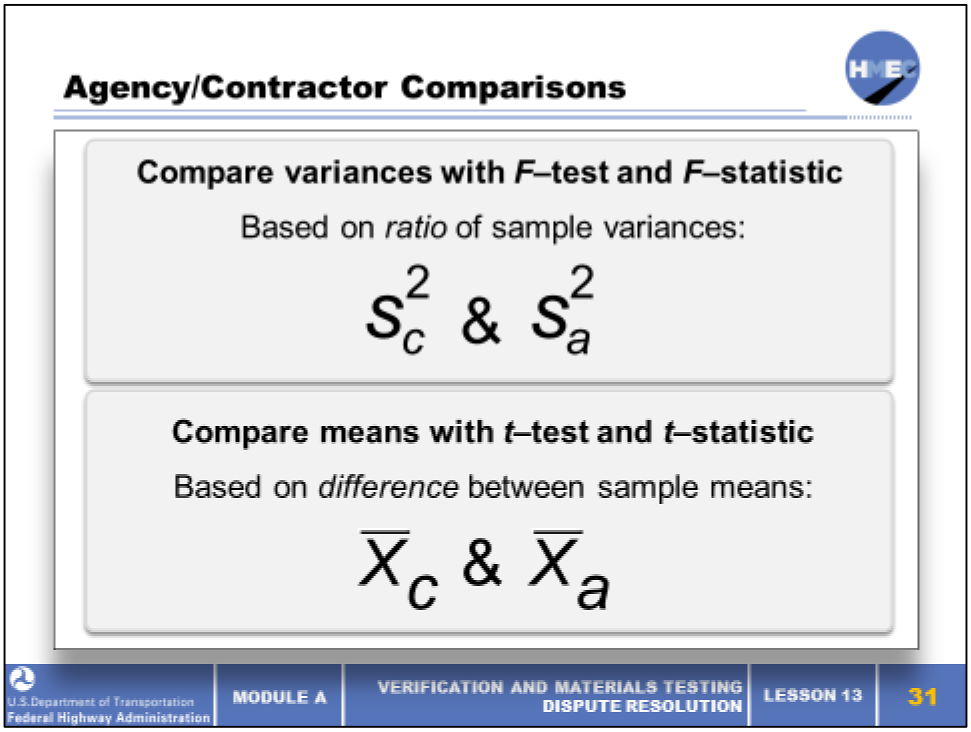
Refer to Table 13-2 *t*-table located in the Resources folder on your tablet (electronic document).

Slide 30

Exercise 1 Solution: Paired *t*-Test

Sample	Contr.	Agency	Diff.
1	5.2	5.4	0.2
2	5.7	6.0	0.3
3	5.1	6.2	1.1
4	6.0	6.5	0.5
5	5.8	5.7	0.1
6	5.3	5.8	0.5

Slide 31



When comparing the two data sets, it is important to compare both the means and the variances. A different test is used for each of these properties. The *F*-test provides a method for comparing the variances (standard deviation squared) of two sets of data. Differences in means are assessed by the *t*-test. Previously, it was noted that construction processes and material properties usually follow a normal distribution. The ratios of variances follow an *F*-distribution, while the means of relatively small samples follow a *t*-distribution. Using these distributions, hypothesis tests can be conducted using the same concepts that were used in the prior examples using pavement thickness and the normal distribution.

For samples from the same normal population, the statistic *F*, which is the ratio of the two sample variances, has a sampling distribution called the *F*-distribution. Tables are available for the *F*-distribution just as they are for the normal distribution. For validation and verification testing, the *F*-test is based on the ratio of the sample variance of the contractor's test results, and the sample variance of the agency's test results.

Similarly, the *t*-statistic and the *t*-test can be used to test whether the sample mean of the contractor's

____

____

____

____

____

____

____

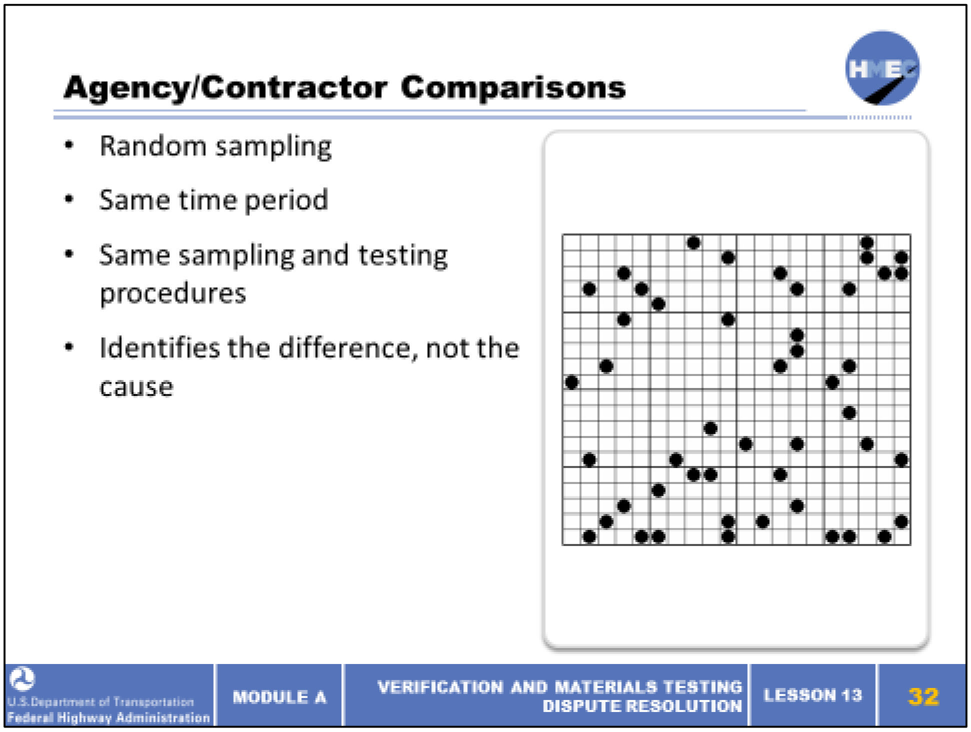
____

____

____

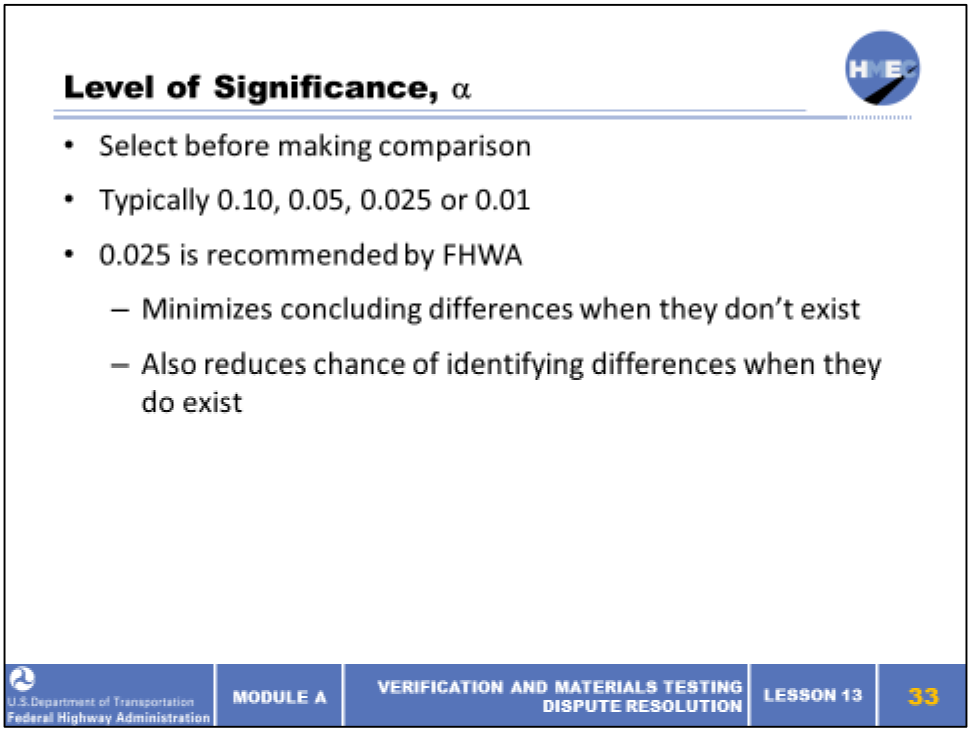
____

Slide 32



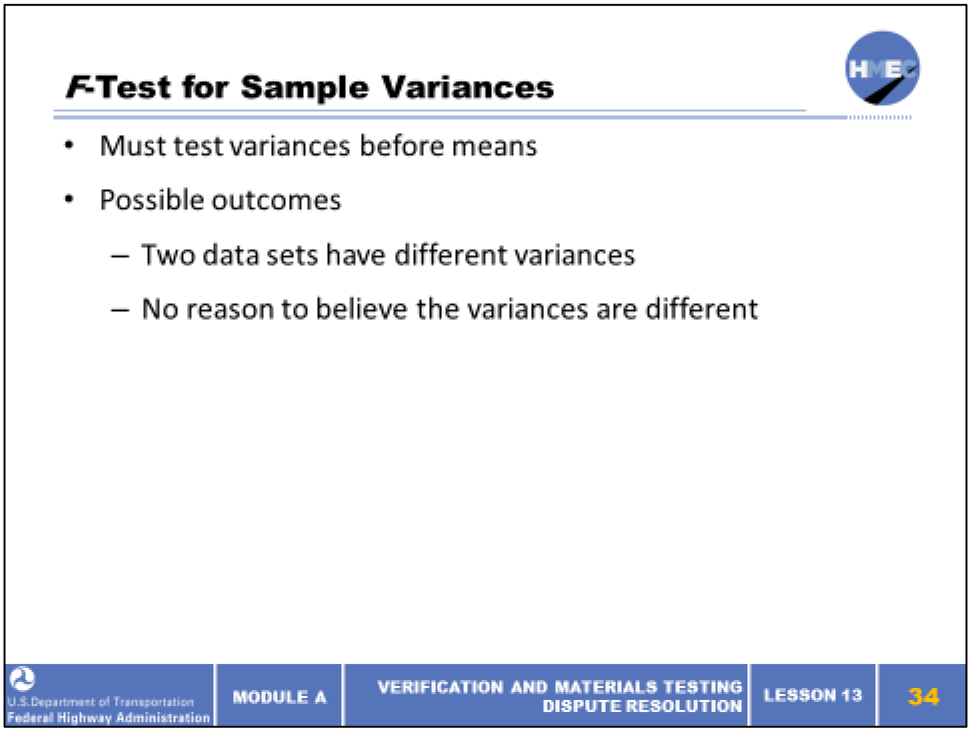
It is important that random sampling was used when obtaining the samples. Also, because sources of variability influence the population parameters, the two sets of test results must have been sampled over the same time period, and the same sampling and testing procedures must have been used. If it is determined that a significant difference is likely between either the variance or the means, the source of the difference should be identified. The identification of a difference is just that, i.e., it identifies that a difference exists. The reason for the difference must still be determined.

Slide 33



As previously mentioned, the same discussions, including the choice of α values of 0.10, 0.05, 0.025 and 0.01, apply here. Many agencies select a value of 0.01 to minimize the likelihood of incorrectly concluding that the results are different when they actually came from the same population. However, this also means that selecting a low α value reduces the chance of detecting a difference when one actually exists. FHWA recommends 0.025 as a compromise.

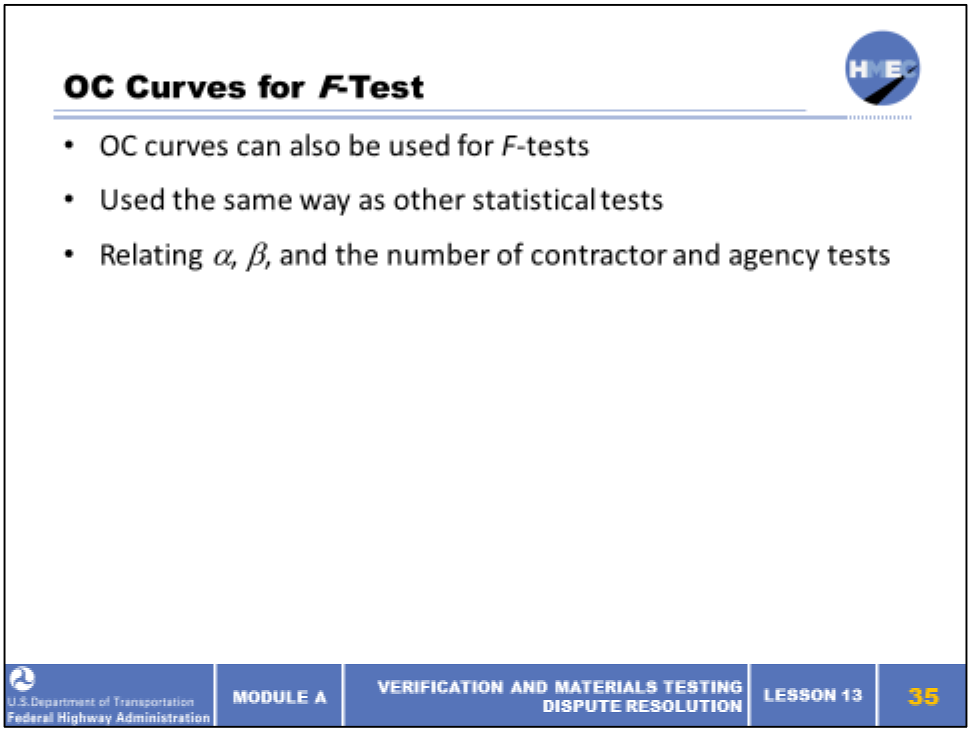
Slide 34



The intent is to determine whether the difference in the variability of the contractor's tests and the agency's tests is larger than what might be expected by chance if they came from the same population. It does not matter which variance is larger. After comparing the *F*-test results, one of the following will be concluded:

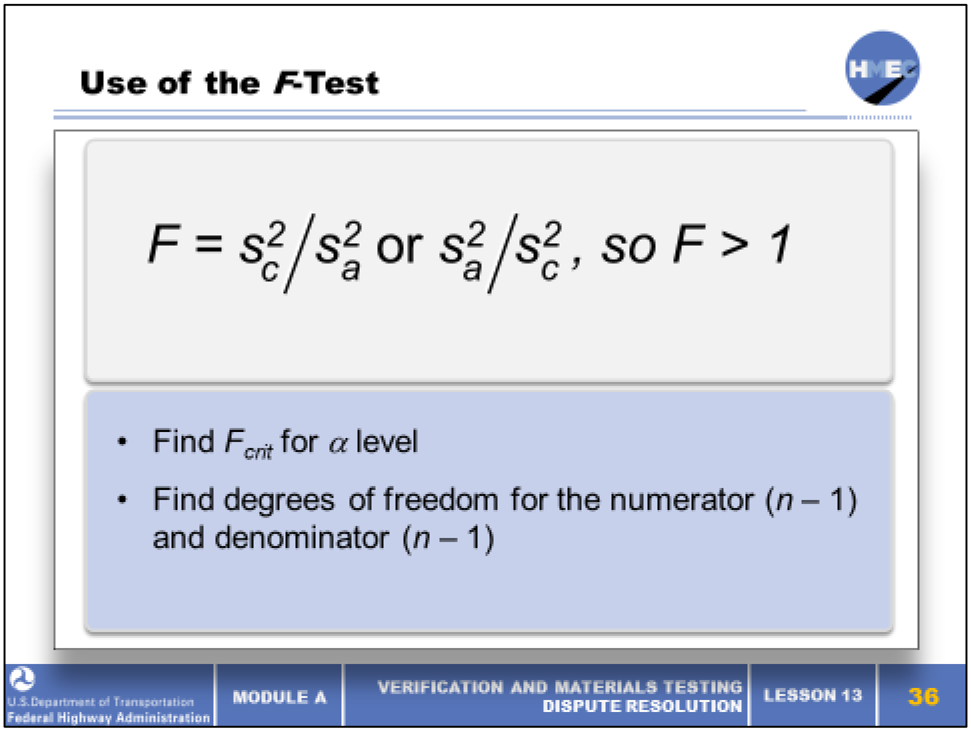
• The two sets of data have different variances because the difference between the two sets of test results is greater than is likely to occur from chance if their variances are actually equal.
• There is no reason to believe the variances are different because the difference is not so great as to be unlikely to have occurred from chance if the variances are actually equal.

Slide 35



OC curves for *F*-tests are not covered here, but you should be aware that they exist.

Slide 36



The first step is to compute the variance for the contractor's tests, s_c^2, and the agency's tests, s_a^2. Next, use the simple ratio equation to compute *F*, where s_c^2/s_a^2 *or* s_a^2/s_c^2. *Always use the larger of the variances in the numerator so that the ratio will be greater than 1.* The next step is to choose α, the level of significance for the test. Here $\alpha = 0.01$ is used. The third step is to determine the critical *F* value, F_{crit}, from the *F*-table (see Table 13–1 at the end of this lesson) for the α level of significance chosen, and using the degrees of freedom (n – 1) associated with each set of test results. Thus, the degrees of freedom associated with the contractor's variance, s^2_c, is $(n_c - 1)$ and the degrees of freedom associated with the agency's variance, s^2_a, is $(n_a - 1)$.

The values in this *F*-table are tabulated to test if there is a difference (either larger or smaller) between the two variance estimates. This is known as a two-sided or two-tailed test. Care must be taken when using other tables of the *F*-distribution, since they are usually based on a one-tailed test, i.e., testing whether one variance is larger than another. This means that the F_{crit} values in Table 13–1 are the same values that would be listed at the 99.5 percentile (even though the 99.0 percentile would normally be associated with $\alpha = 0.01$) for a one-sided test. Once the value for F_{crit} is determined from the table (making sure the appropriate degrees of freedom for the numerator and denominator are used), if $F \geq F_{crit}$ then decide that the two sets of tests have significantly different variabilities. If $F < F_{crit}$ then decide that there is no reason to believe that the variabilities are significantly different. Variances are assumed different if $F \geq F_{crit}$.

____

____

____

____

____

____

____

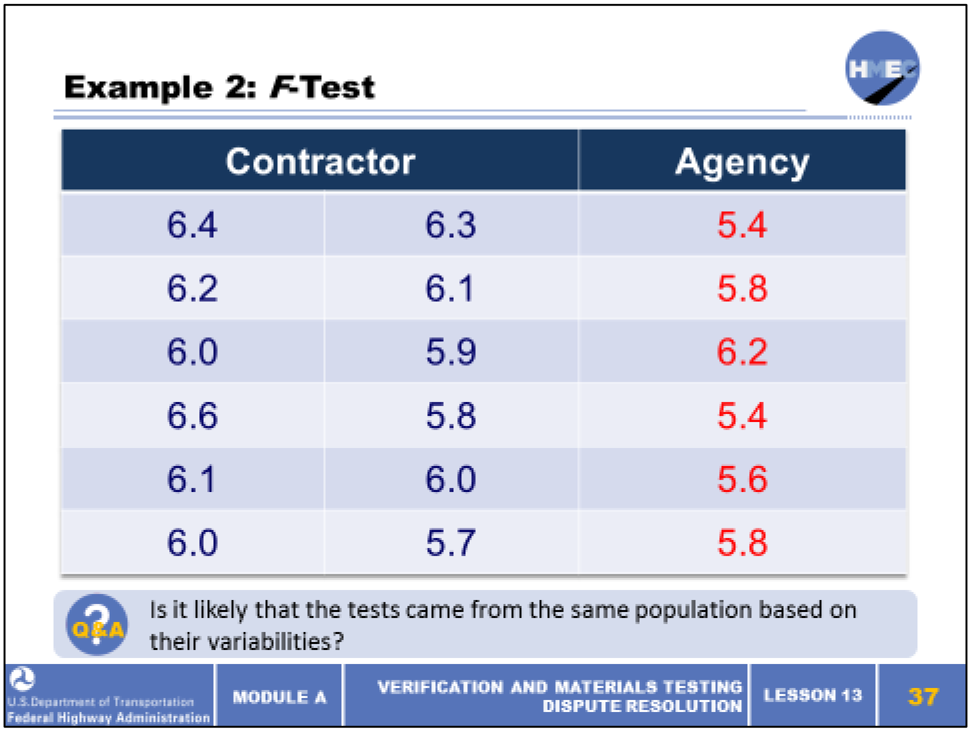
____

____

____

____

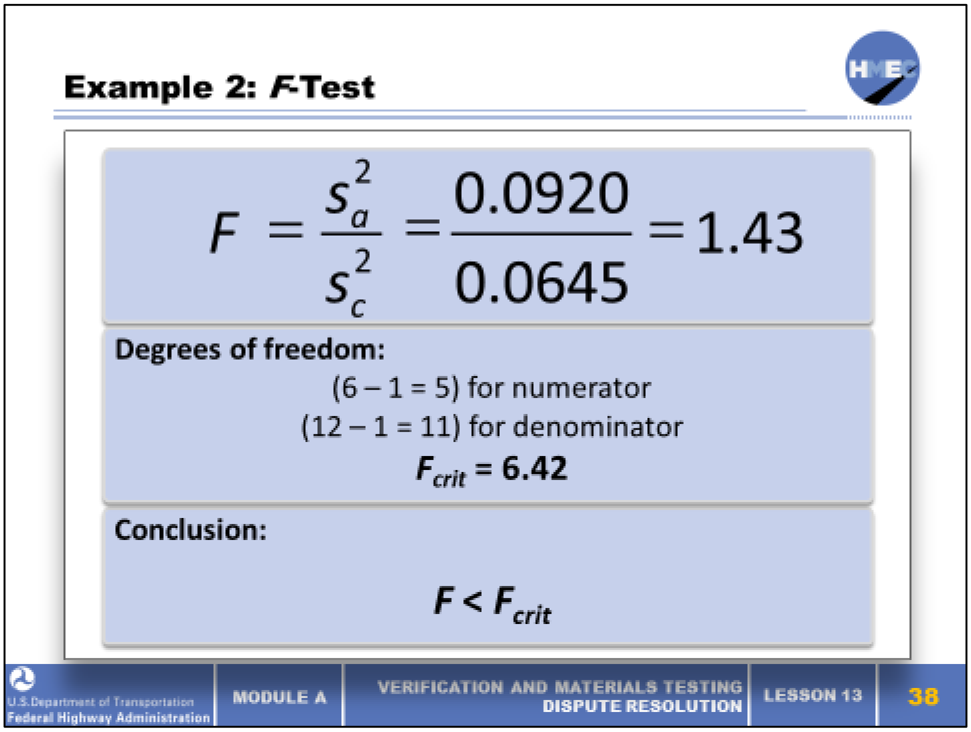
Slide 37



Contractor		Agency
6.4	6.3	5.4
6.2	6.1	5.8
6.0	5.9	6.2
6.6	5.8	5.4
6.1	6.0	5.6
6.0	5.7	5.8

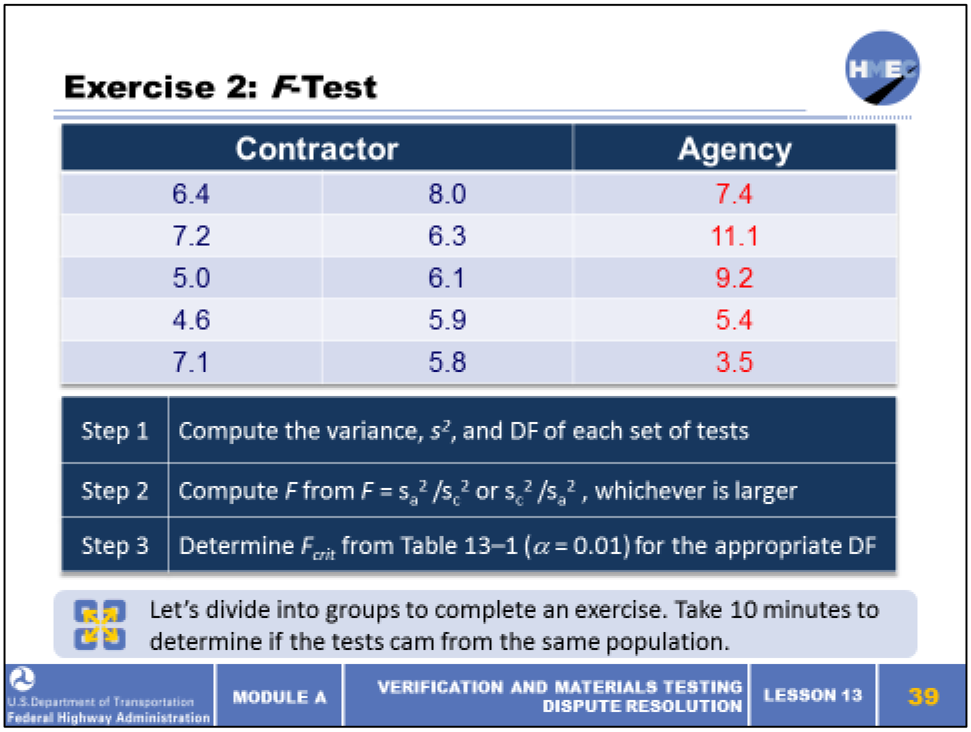
A contractor has run 12 asphalt content tests (advance slide) and the agency has run six tests over the same period of time using the same sampling and testing procedures.

Slide 38



• Step 1: Compute the variance, s^2, for each set of tests. $s_c^2 = 0.0645$ $s_a^2 = 0.0920$. Since s_a^2 is the larger of the two, use this in the numerator.
• Step 2: Compute F : $F = 0.0920/0.0645 = 1.43$.
• Step 3: Determine F_{crit} from the F-distribution table (13-1) making sure to use the correct degrees of freedom for the numerator ($n_a - 1 = 6 - 1 = 5$) and the denominator ($n_c - 1 = 12 - 1 = 11$). From Table 13–1, $F_{crit} = 6.42$.
• *Conclusion:* Since $F < F_{crit}$ (i.e., 1.43 < 6.42), there is no reason to believe that the two sets of data have different variabilities. That is, they could have come from the same population.

Slide 39



Exercise 2: *F*-Test

Contractor		Agency
6.4	8.0	7.4
7.2	6.3	11.1
5.0	6.1	9.2
4.6	5.9	5.4
7.1	5.8	3.5

Step 1	Compute the variance, s^2, and DF of each set of tests
Step 2	Compute F from $F = s_a^2 / s_c^2$ or s_c^2 / s_a^2, whichever is larger
Step 3	Determine F_{crit} from Table 13–1 ($\alpha = 0.01$) for the appropriate DF

Let's divide into groups to complete an exercise. Take 10 minutes to determine if the tests cam from the same population.

A contractor has run 10 air void tests from cores and the agency has run five air void tests over the same period of time using the same sampling and testing procedures. The results are shown on the slide.

Based on their variabilities, is it likely that the tests came from the same population?

Slide 40

Exercise 2 Solution: *F*-Test

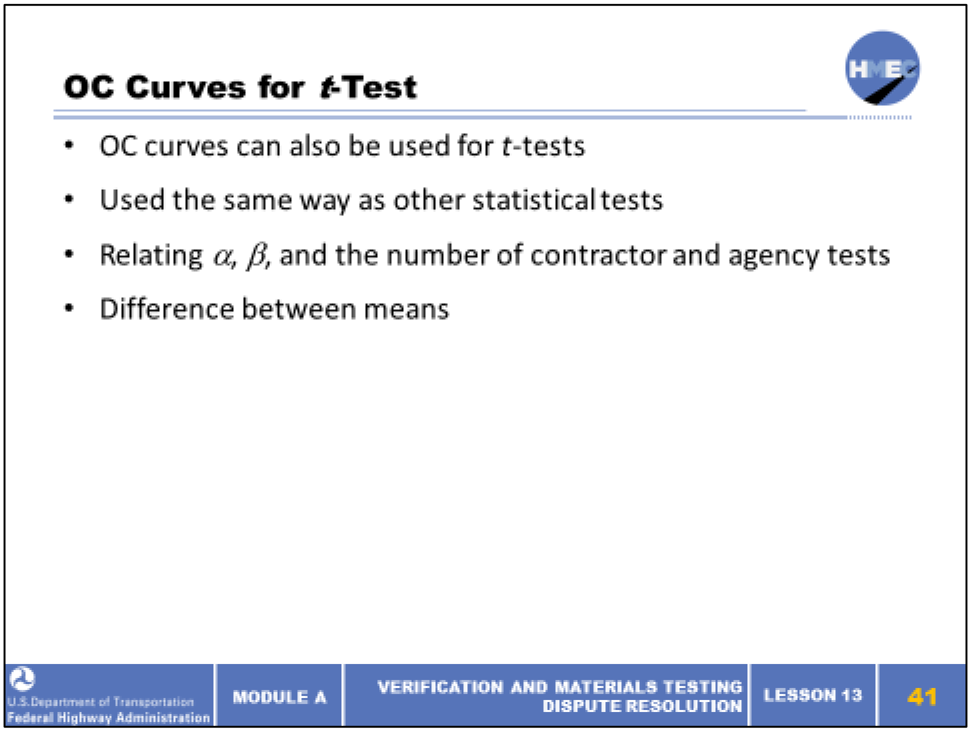
Contractor		Agency
6.4	8.0	7.4
7.2	6.3	11.1
5.0	6.1	9.2
4.6	5.9	5.4
7.1	5.8	3.5

Step 1	$s^2_c = 1.0382$; $s^2_a = 9.0270$; $DF_c = (10-1) = 9$; $DF_a = (5-1) = 4$
Step 2	$F = s^2_a / s^2_c = 9.0270/1.0382 = 8.695$
Step 3	$F_{crit} = 7.96$ (Conclude: Since $F > F_{crit}$, assume that the variances are different)

U.S. Department of Transportation Federal Highway Administration | MODULE A | VERIFICATION AND MATERIALS TESTING DISPUTE RESOLUTION | LESSON 13 | 40

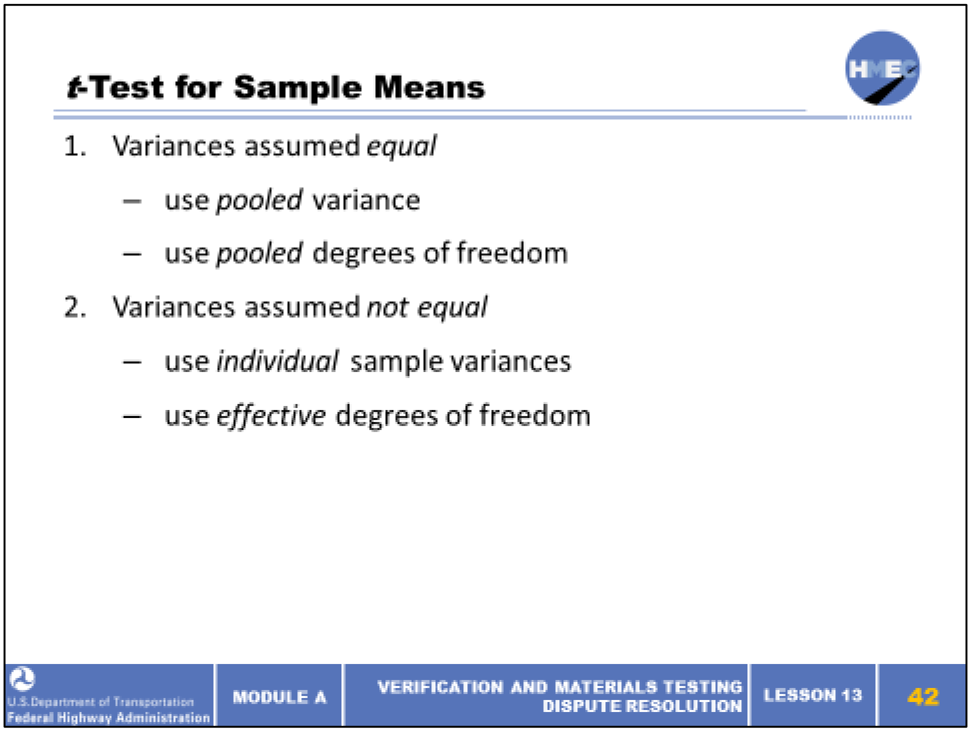
So, based on their variabilities, is it likely that the tests came from the same population?

Slide 41



Note that the procedure for using the OC curves for the paired *t*-test that was presented previously cannot be used for the two-sample *t*-test with different sample sizes. Either different OC curves or a different method for calculating the standardized difference, *d** as opposed to *d*, must be used. These other procedures are not covered here due to the limited amount of time. The equations for the *t*-test will be covered conceptually but it is recommended that a computer program be used in practice to perform the calculations. Spreadsheet programs such as Microsoft Excel have both *F*- and *t*-tests.

Slide 42



The intent is to determine whether it is reasonable to assume that the contractor's test came from the same population as the agency's test. A *t*-test is used to compare the sample means. Two approaches for the *t*-test are necessary. If the sample variances are assumed equal from the *F*-test, then the *t*-test is conducted based on the two sets of samples using a pooled estimate for the variance and the pooled degrees of freedom, which will be explained shortly (Example 2 above). This approach is Case 1 described later in this lesson. If the sample variances are assumed to be different, then the *t*-test is conducted using the individual sample variances, the individual sample sizes, and the effective degrees of freedom, which will also be explained shortly (estimated from the sample variances and sample sizes). This approach is Case 2, which we'll look at soon.

____

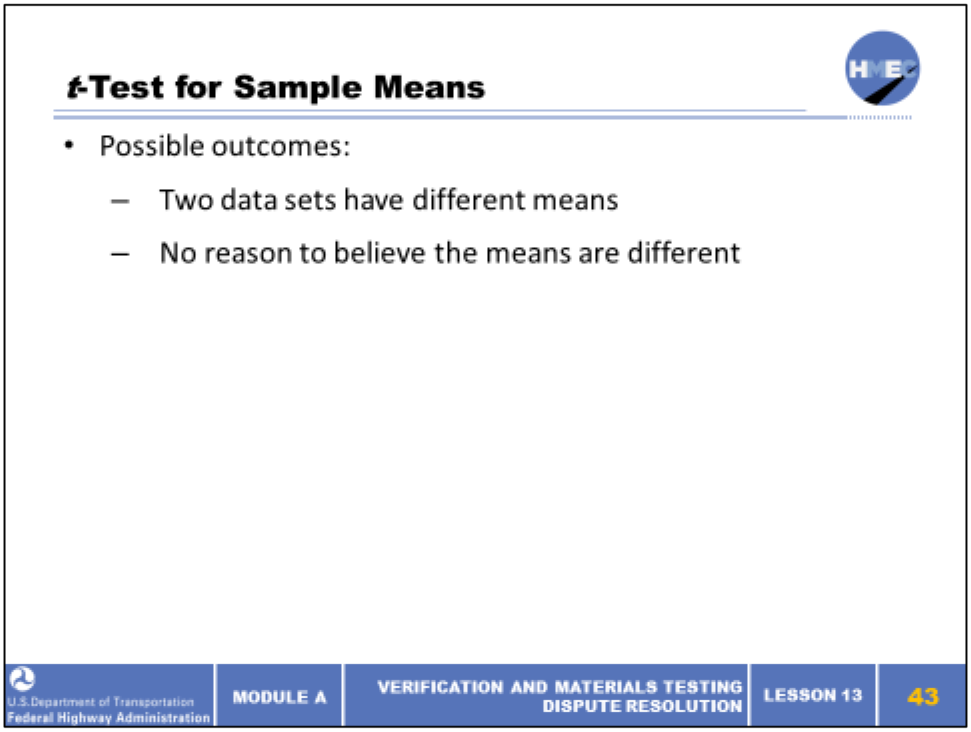
____

____

____

____

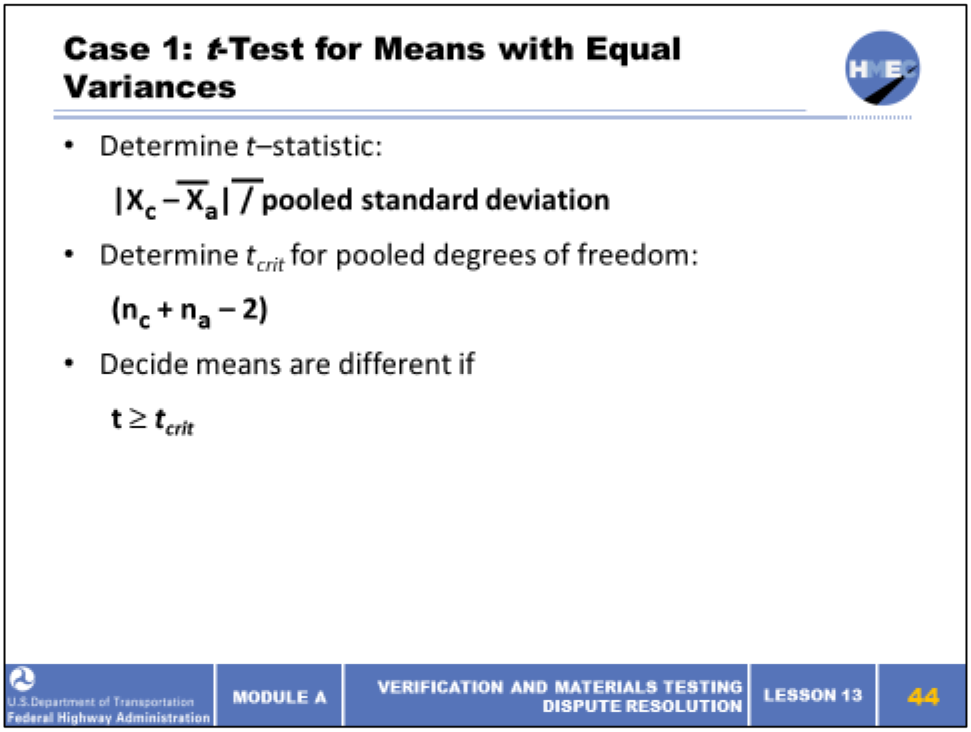
Slide 43



In either of the two cases discussed in the previous slide, one of following decisions is made:

- The two sets of data have different means because the difference in the sample means is greater than is likely to occur from chance if their means are actually equal.
- There is no reason to believe that the means are different because the difference in the sample means is not so great as to be unlikely to have occurred from chance if the means are actually equal.

Slide 44



Conceptually, for the *t*-test where the sample variances are equal, the equation used to calculate the *t* value divides the difference between the two means by the pooled standard deviation. The pooled standard deviation is the square root of the pooled variance, which is the weighted average of the two variances using the degrees of freedom for each sample as the weighting factor. (Conceptually, this is similar to the Z equation in which the difference between the mean and a point of interest is expressed in terms of the number of standard deviations. But, because small sample sizes are used, the *t*-distribution is used.)

To determine the critical *t* value, t_{crit}, against which the computed *t* value is compared, it is necessary to select the level of significance, α. Again, a value of $\alpha = 0.01$ is recommended. Next, the critical *t* value, t_{crit}, is obtained from the *t*-table for the pooled degrees of freedom. The value for pooled degrees of freedom for the case where the sample variances are assumed equal is $(n_c + n_a - 2)$. If $t \geq t_{crit}$, then decide that the two sets of tests have significantly different means. If $t < t_{crit}$, then decide that there is no reason to believe the means are significantly different.

____

____

____

____

____

____

____

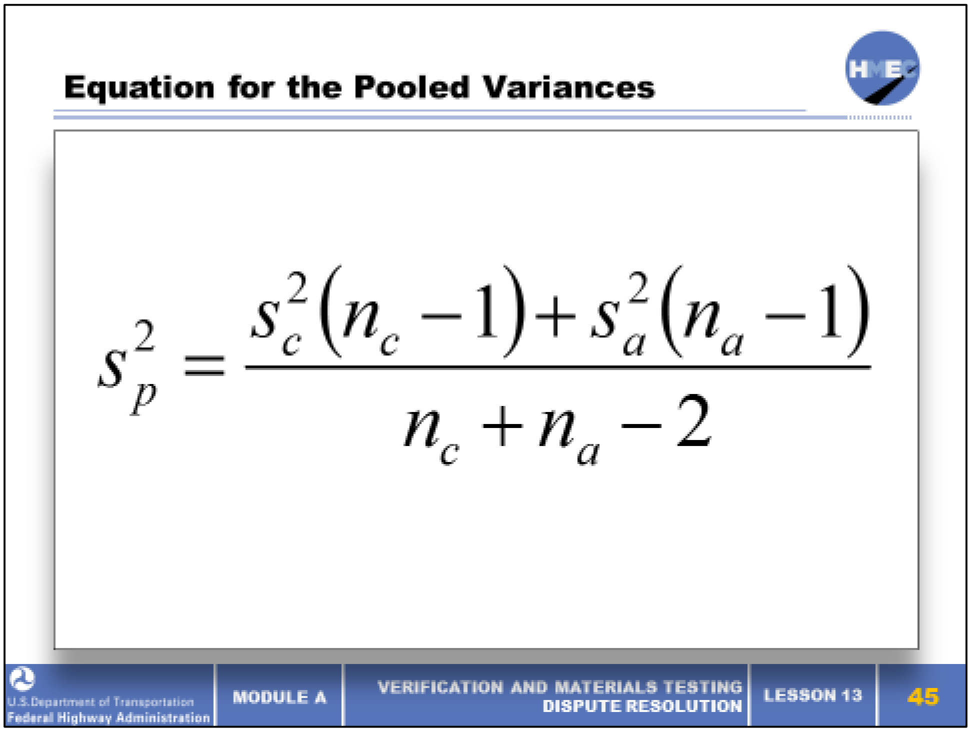
____

____

____

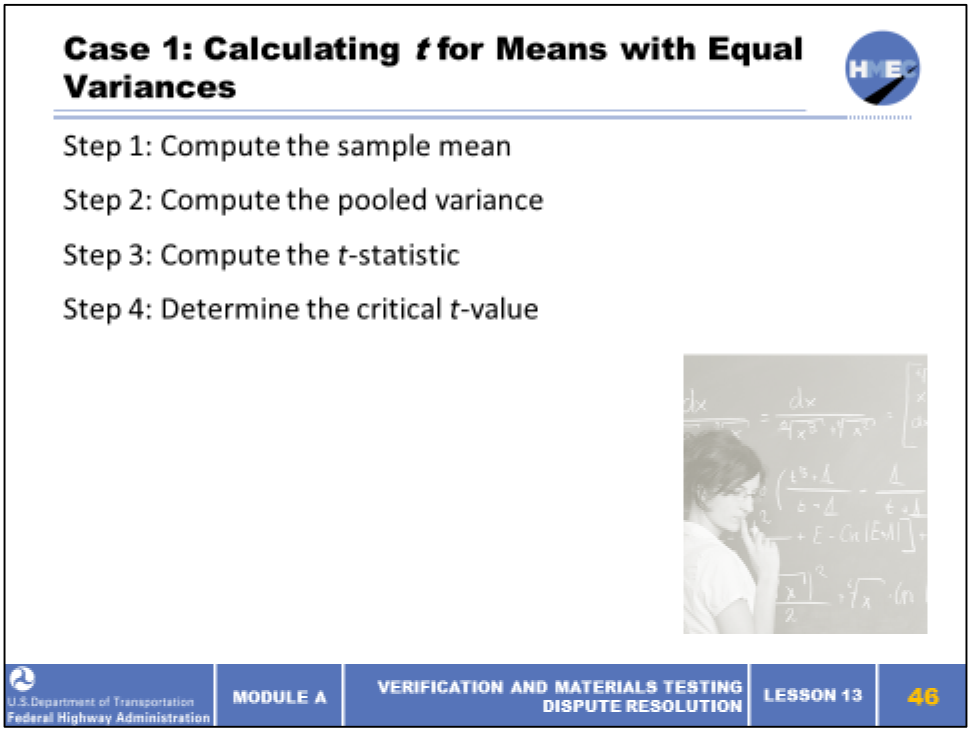
____

Slide 45



The pooled variances is the weighted average of the two variances using the degrees of freedom for each sample as the weighting factor. From this formula, we will take the square root to use in the *t*-test.

Slide 46



These three steps show the procedure for calculating the *t*-value for means with equal variances using the equation on the next slide.

Slide 47

Case 1: t-Test for Means with Equal Variances

Contractor		Agency
6.4	6.3	5.4
6.2	6.1	5.8
6.0	5.9	6.2
6.6	5.8	5.4
6.1	6.0	5.6
6.0	5.7	5.8

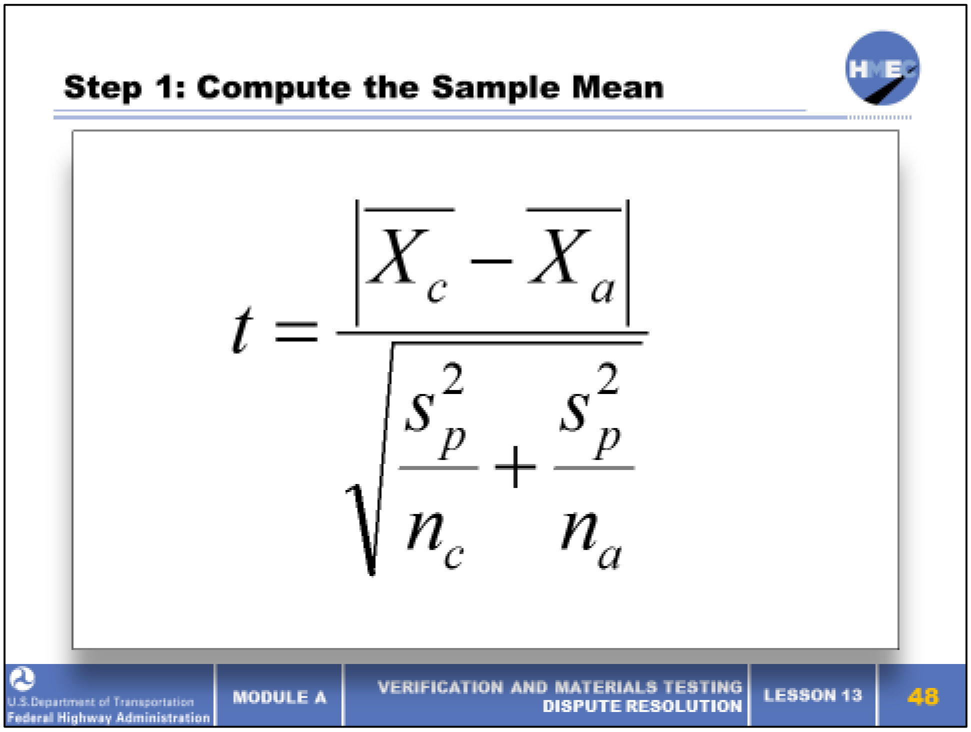
Q&A: Based on their means, is it likely that the tests came from the same population?

U.S. Department of Transportation Federal Highway Administration | MODULE A | VERIFICATION AND MATERIALS TESTING DISPUTE RESOLUTION | LESSON 13 | 47

This example problem is the one in which a contractor has run 12 asphalt content tests and the agency has run six tests over the same period of time using the same sampling and testing procedures. The results are shown on the slide. Remember, the variances were found to be equal for these data.

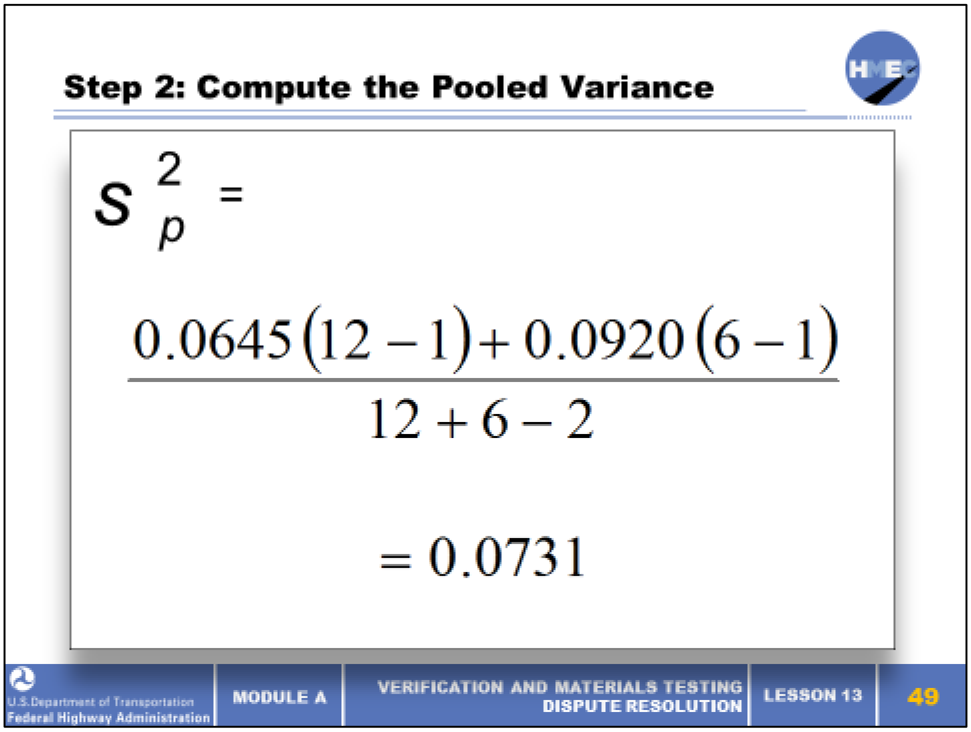
In this case, $\alpha = 0.01$.

Slide 48



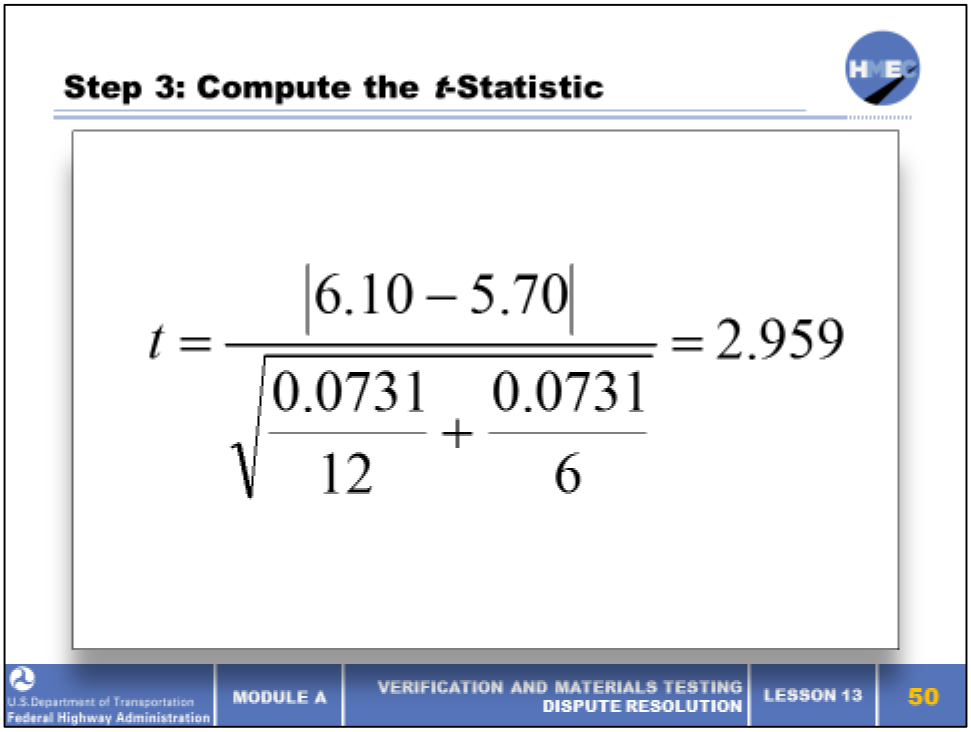
The first step is to compute the sample mean. This equation is for the *t*-test where the sample variances are equal. The *t*-value is determined by dividing the absolute difference between the two means by the pooled standard deviation shown previously. (Conceptually, this test is similar to the Z equation in which the difference between the mean and a point of interest is expressed in terms of the number of standard deviations. But, because small sample sizes are used, the *t*-distribution is used instead of the normal distribution.)

Slide 49



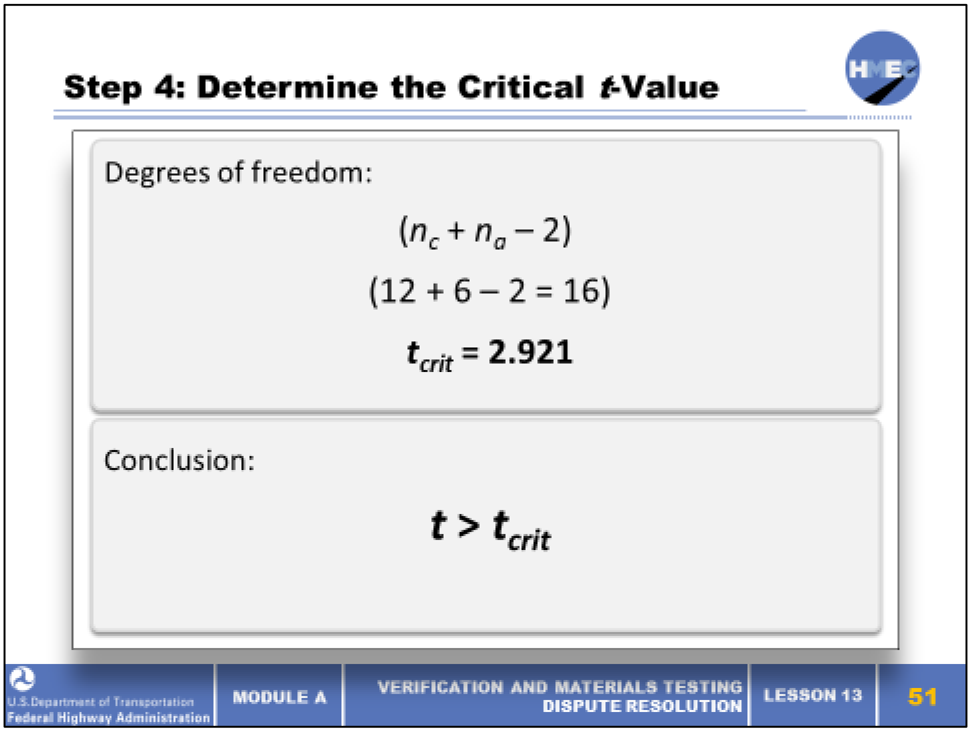
Step 2: From the s_c^2 = 0.0645 and s_a^2 = 0.0920, calculate the s_p^2 . s_p^2 =[0.0645 (12-1) + 0.0920 (6-1)]/ (12 + 6 + 2) = 0.0731

Slide 50

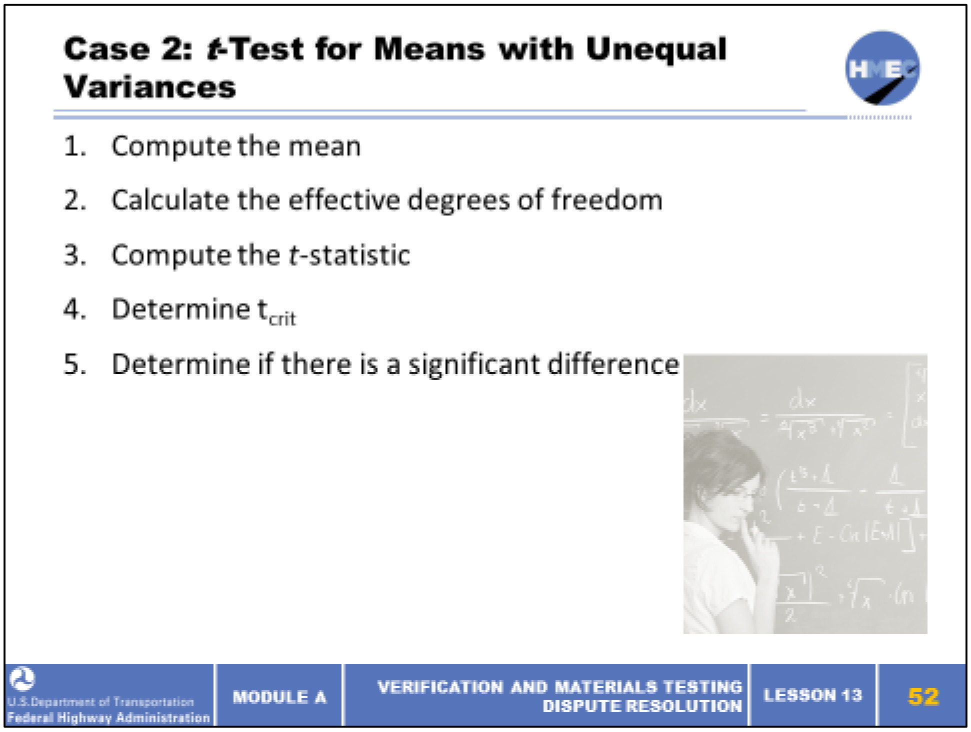


Substituting the means, pooled variances, and DF provides a *t* equal to 2.959.

Slide 51

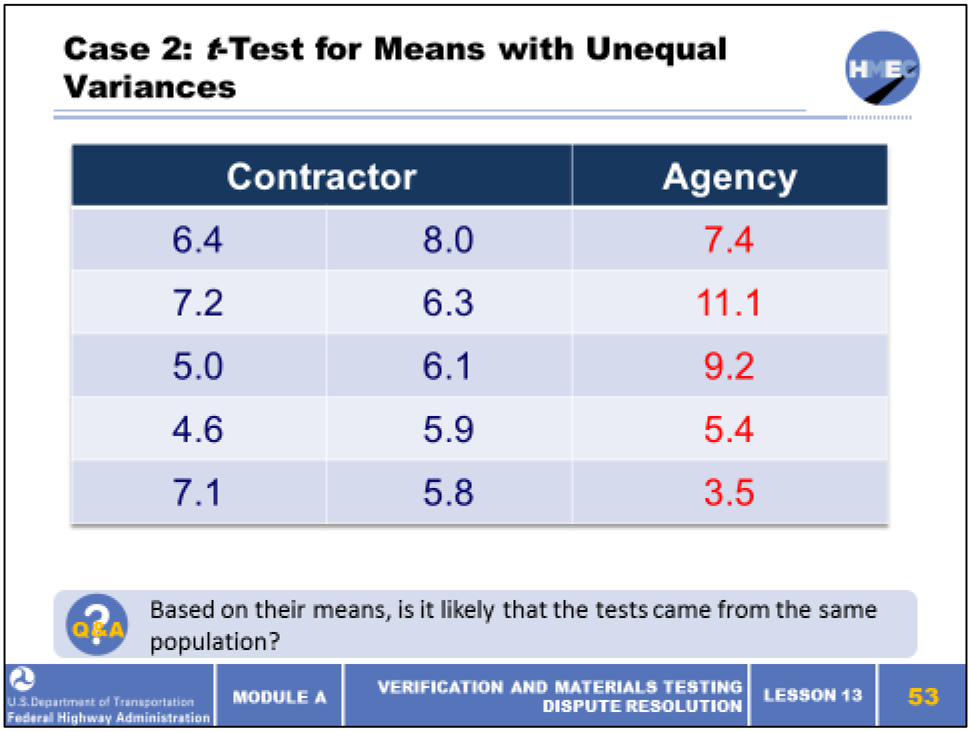

Step 4: Determine the Critical t-Value
HMEC
Degrees of freedom:
$(n_c + n_a - 2)$
$(12 + 6 - 2 = 16)$
$t_{crit} = 2.921$
Conclusion:
$t > t_{crit}$
U.S. Department of Transportation
Federal Highway Administration
MODULE A
VERIFICATION AND MATERIALS TESTING
DISPUTE RESOLUTION
LESSON 13
51

Slide 52



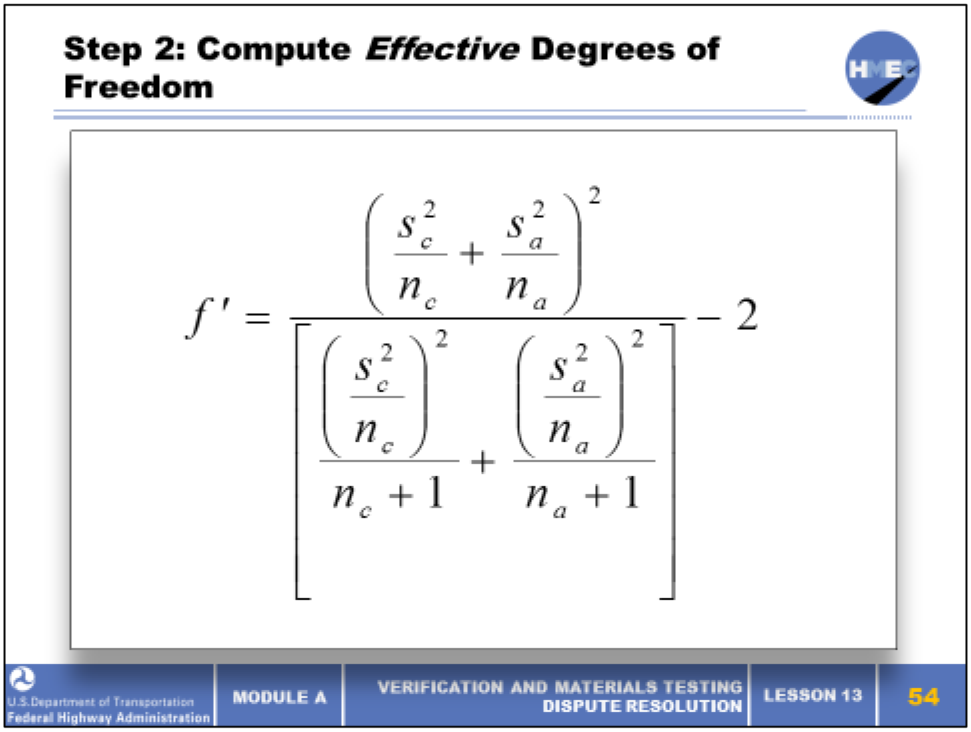
The equation is similar to the t-statistic with equal variances. One difference is that the individual variances must be used (they can’t be pooled because they were found to be significantly different). As we shall see, the degrees of freedom will be different than in Case 1.

Slide 53



The first step is to compute the mean. These are the same results as we saw previously where the variances were found to be unequal.

Slide 54



It essentially uses the variances to weight the degrees of freedom. This is Step 3 from slide 39, to determine the effective degrees of freedom.

____

____

____

____

____

____

____

____

____

____

Slide 55

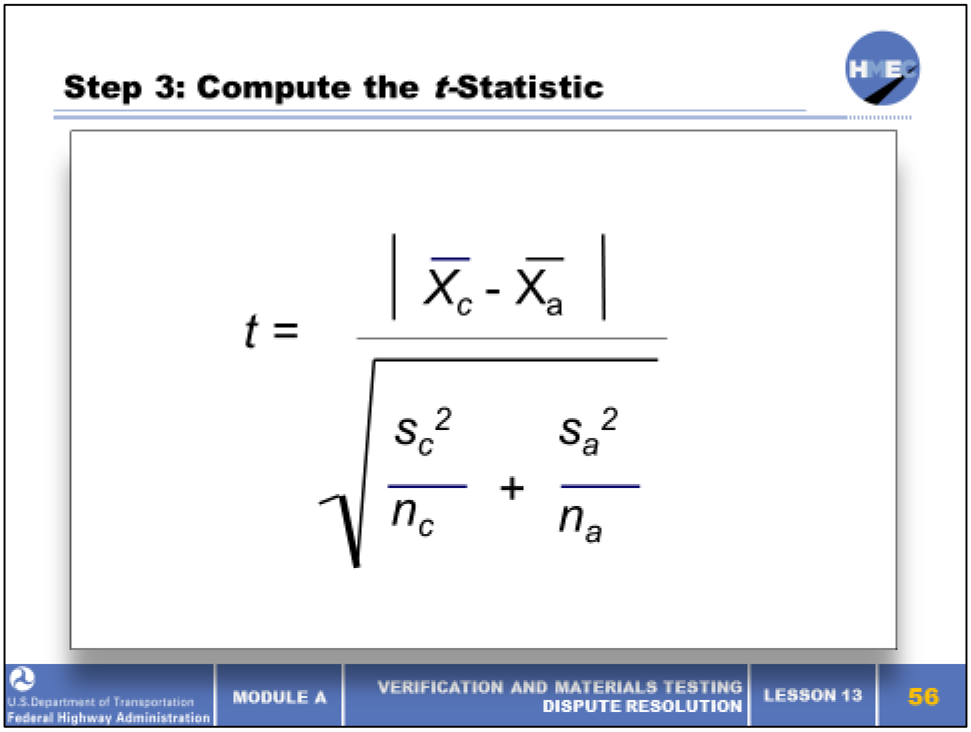
Step 2: Compute *Effective* Degrees of Freedom

$$f' = \frac{\left(\frac{1.038}{10} + \frac{9.027}{5}\right)^2}{\left[\frac{\left(\frac{1.038}{10}\right)^2}{10+1} + \frac{\left(\frac{9.027}{5}\right)^2}{5+1}\right]} - 2 = 4.70$$

U.S. Department of Transportation Federal Highway Administration | MODULE A | VERIFICATION AND MATERIALS TESTING DISPUTE RESOLUTION | LESSON 13 | 55

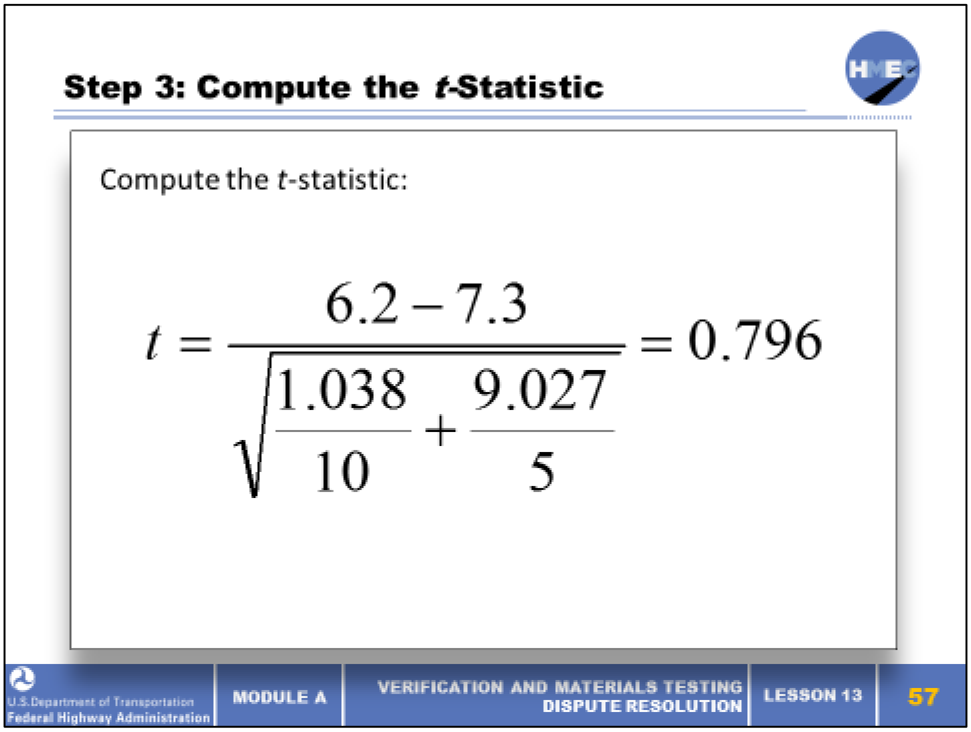
Since effective degrees of freedom must be a whole number, we round 4.7 to the closest integer, 5.

Slide 56



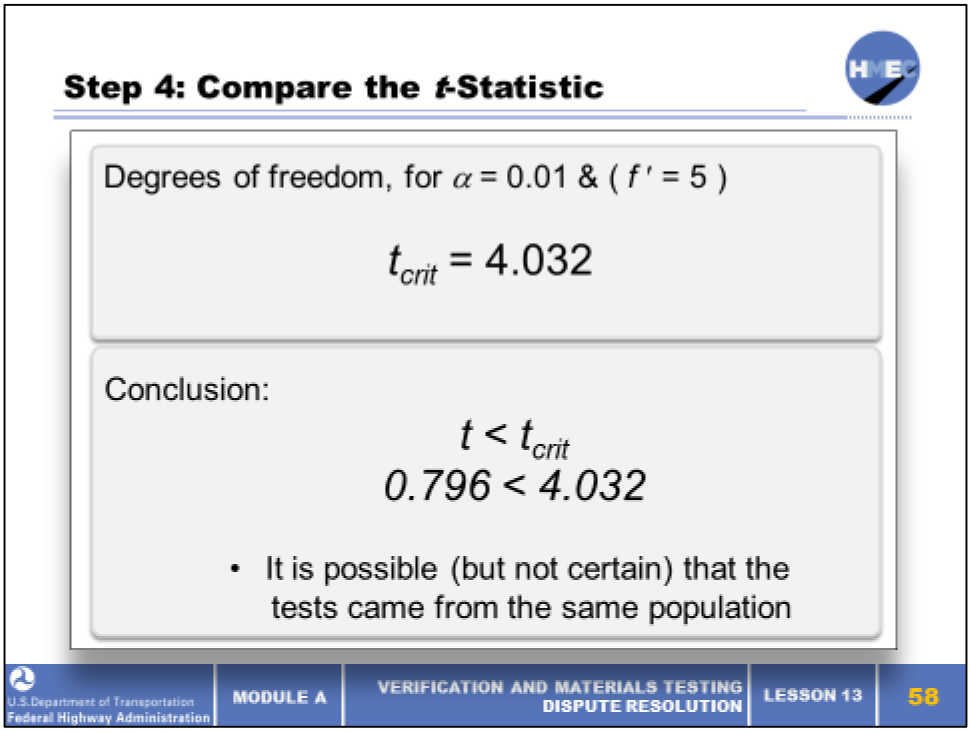
The *t*-statistic is the ratio of the absolute difference between the means divided by the square root of the addition of each variance divided by the sample size.

Slide 57



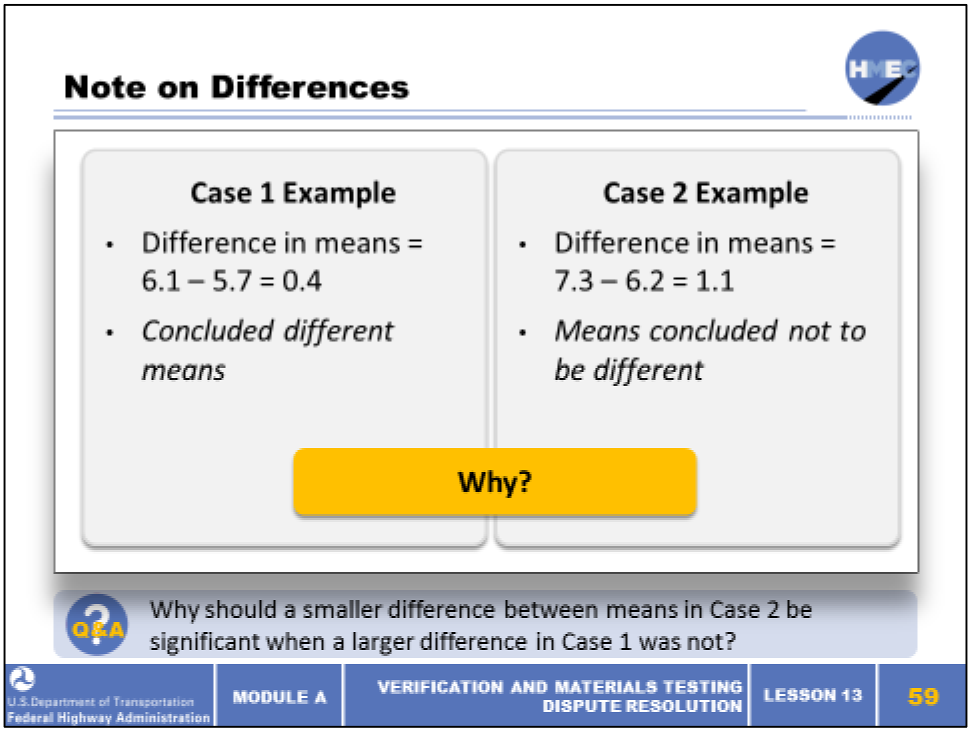
The *t*-statistic is the ratio of the absolute difference between the means divided by the square root of the addition of each variance divided by the sample size.

Slide 58

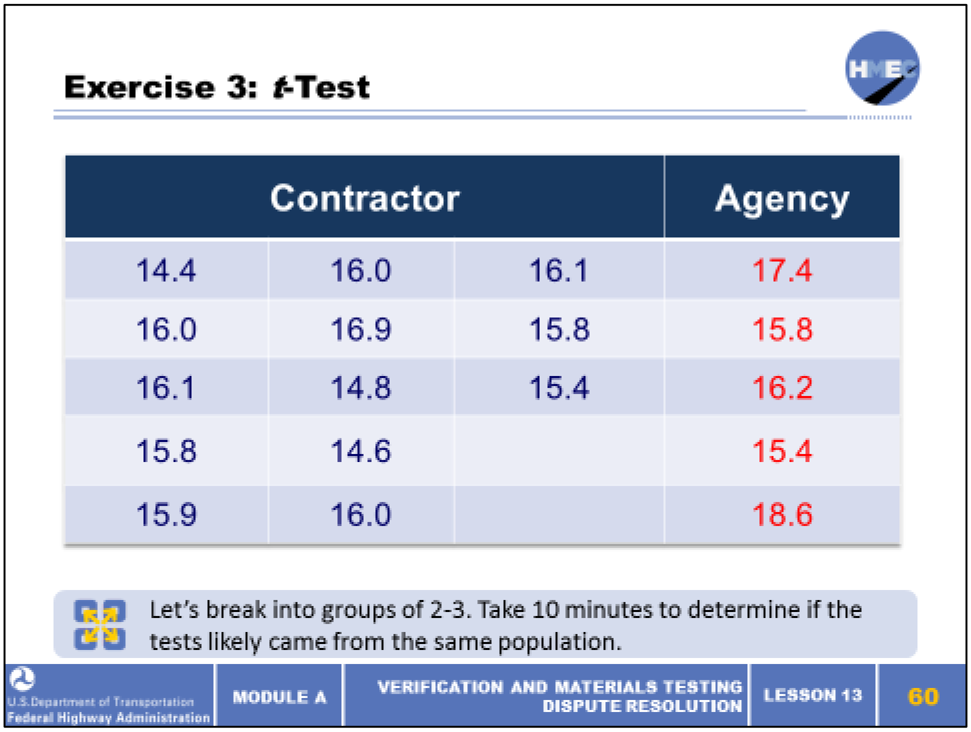


The value for effective degrees of freedom is less than would have been used if the variances had been assumed to be equal (13).

Slide 59


Note on Differences
HMEC
Case 1 Example
Difference in means = 6.1 – 5.7 = 0.4
Concluded different means
Case 2 Example
Difference in means = 7.3 – 6.2 = 1.1
Means concluded not to be different
Why?
Q&A
Why should a smaller difference between means in Case 2 be significant when a larger difference in Case 1 was not?
U.S. Department of Transportation
Federal Highway Administration
MODULE A
VERIFICATION AND MATERIALS TESTING DISPUTE RESOLUTION
LESSON 13
59

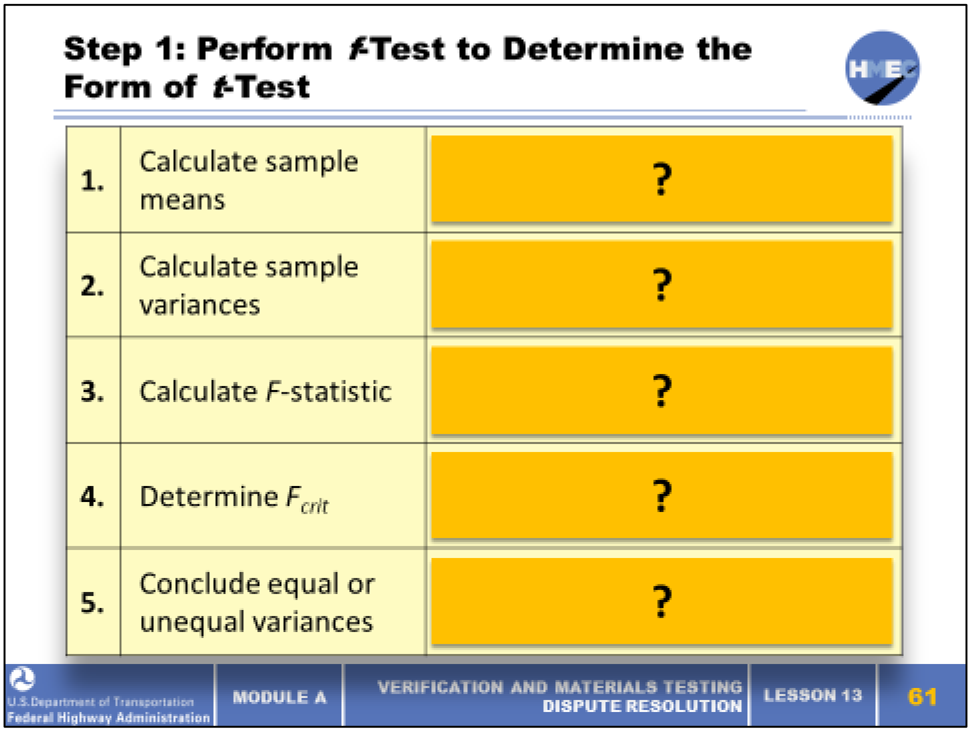
Slide 60



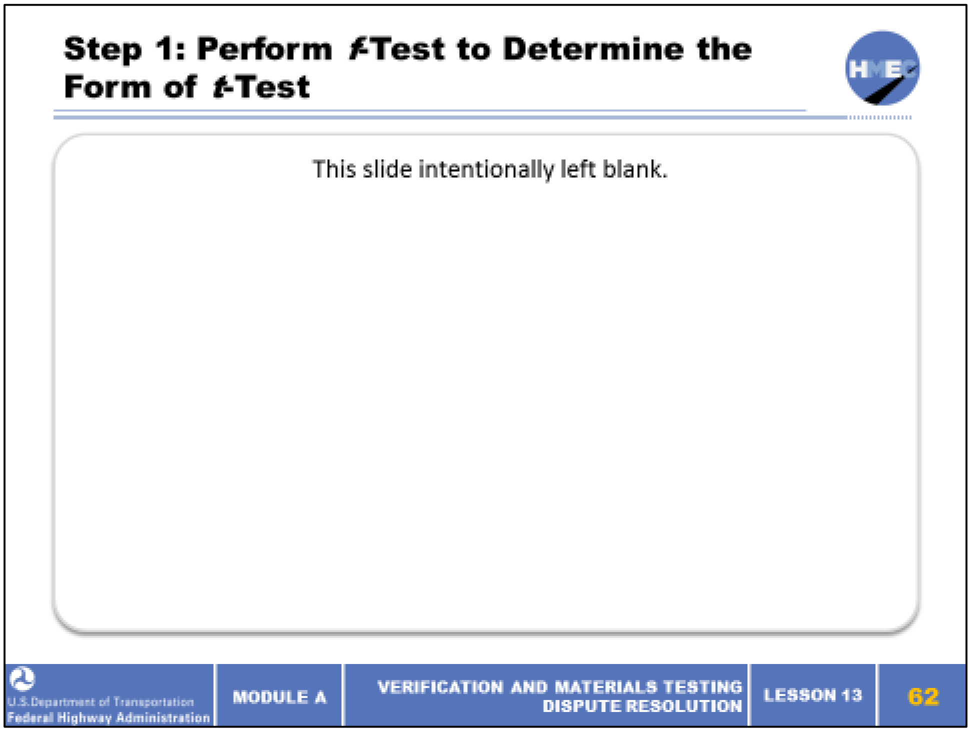
Contractor			Agency
14.4	16.0	16.1	17.4
16.0	16.9	15.8	15.8
16.1	14.8	15.4	16.2
15.8	14.6		15.4
15.9	16.0		18.6

A contractor has run 13 tests on the voids in the mineral aggregate (VMA) and the agency has run 5 tests over the same period of time using the same sampling and testing procedures. The results are shown on the slide. Is it likely that the tests came from the same population?

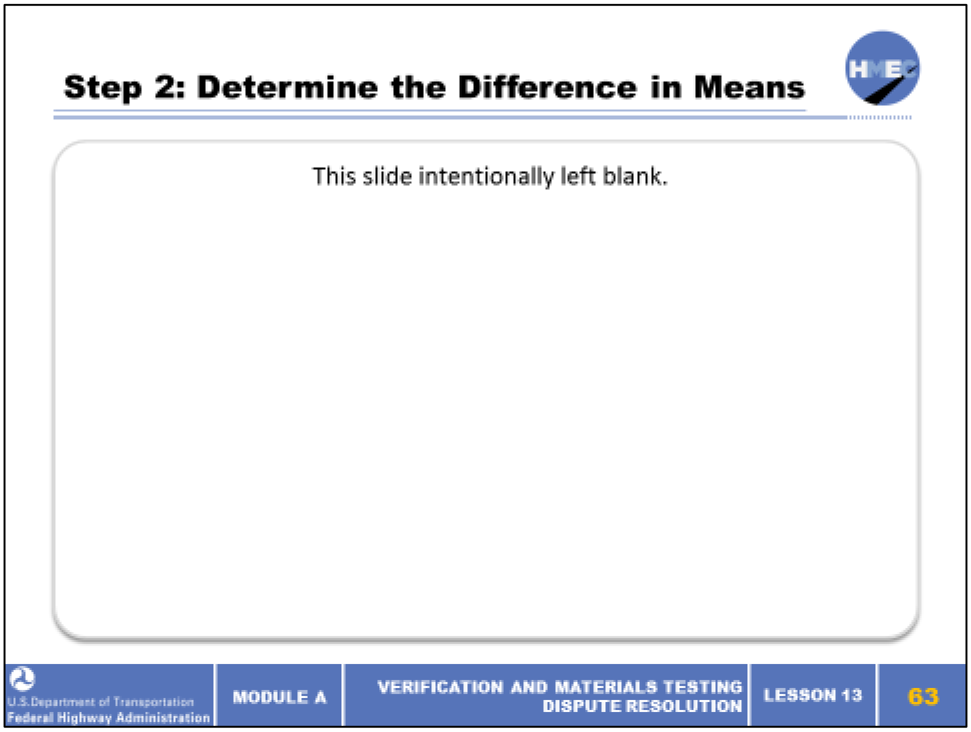
Slide 61


Step 1: Perform F-Test to Determine the Form of t-Test
HMEC
1. Calculate sample means
?
2. Calculate sample variances
?
3. Calculate F-statistic
?
4. Determine F_{crit}
?
5. Conclude equal or unequal variances
?
U.S. Department of Transportation
Federal Highway Administration
MODULE A
VERIFICATION AND MATERIALS TESTING
DISPUTE RESOLUTION
LESSON 13
61

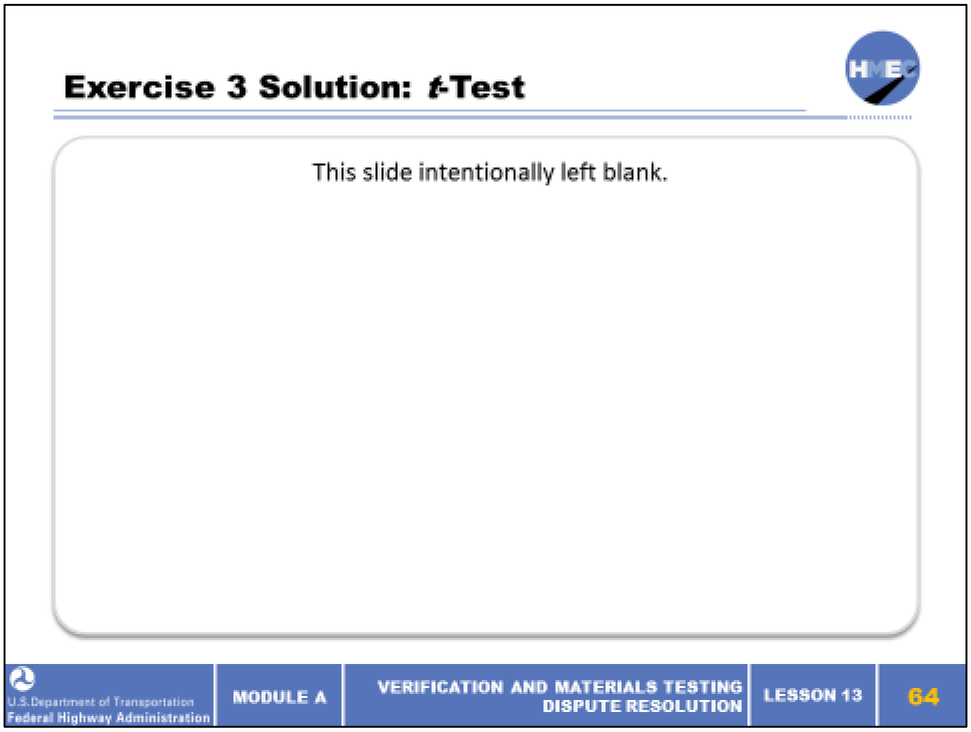
Slide 62


Step 1: Perform F-Test to Determine the Form of t-Test
This slide intentionally left blank.
U.S. Department of Transportation
Federal Highway Administration
MODULE A
VERIFICATION AND MATERIALS TESTING DISPUTE RESOLUTION
LESSON 13
62

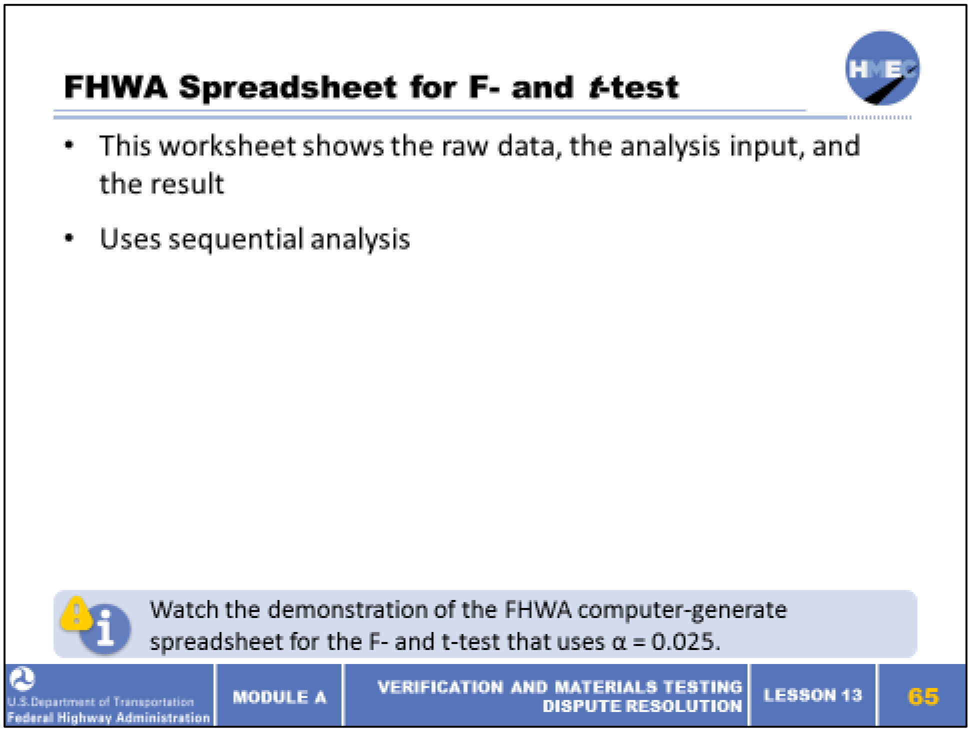
Slide 63


Step 2: Determine the Difference in Means
HMEC
This slide intentionally left blank.
U.S. Department of Transportation
Federal Highway Administration
MODULE A
VERIFICATION AND MATERIALS TESTING
DISPUTE RESOLUTION
LESSON 13
63

Slide 64


Exercise 3 Solution: *t*-Test
HMEC
This slide intentionally left blank.
U.S. Department of Transportation
Federal Highway Administration
MODULE A
VERIFICATION AND MATERIALS TESTING
DISPUTE RESOLUTION
LESSON 13
64

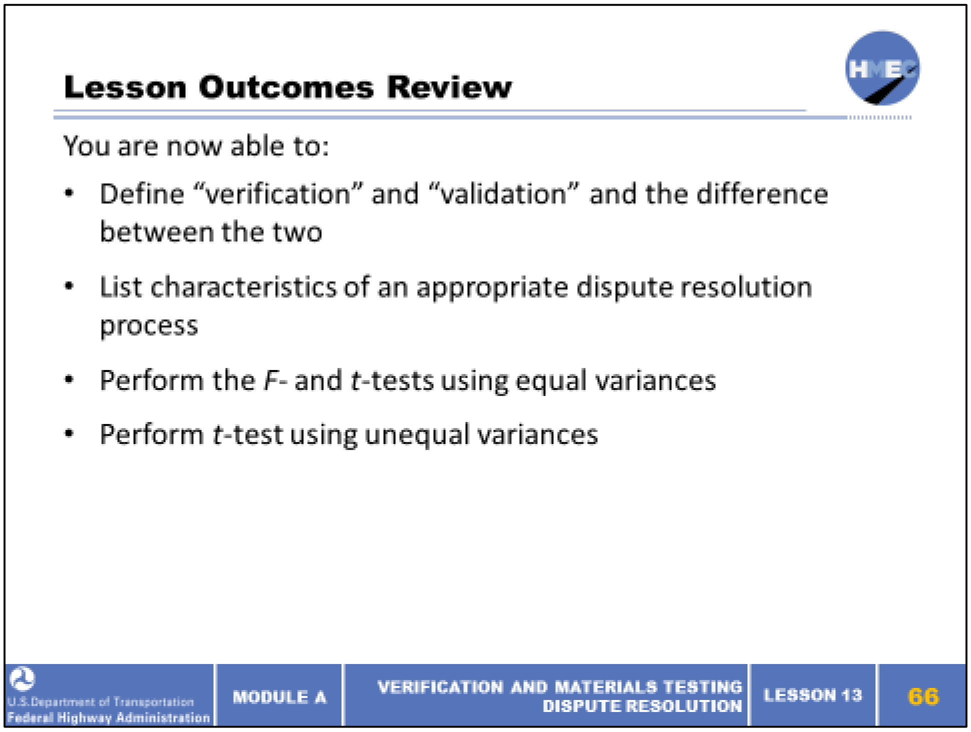
Slide 65



This worksheet shows the raw data, the analysis input, and the result. Another interesting aspect of this analysis, it that it uses a sequential analysis. It first compares contractor's tests 2-17 with the agency's 4 tests over the same period, then compares contractor's test 5-21 with the agency's 4 tests over the same period, and so forth.

F and t Worksheet for HIPT 8-2009 in the Resources folder on your tablet. This is an Excel spreadsheet (.xls). This is for future reference.

Slide 66


Lesson Outcomes Review
HMEC
You are now able to:
• Define "verification" and "validation" and the difference between the two
• List characteristics of an appropriate dispute resolution process
• Perform the F- and t-tests using equal variances
• Perform t-test using unequal variances
U.S. Department of Transportation
Federal Highway Administration
MODULE A
VERIFICATION AND MATERIALS TESTING DISPUTE RESOLUTION
LESSON 13
66

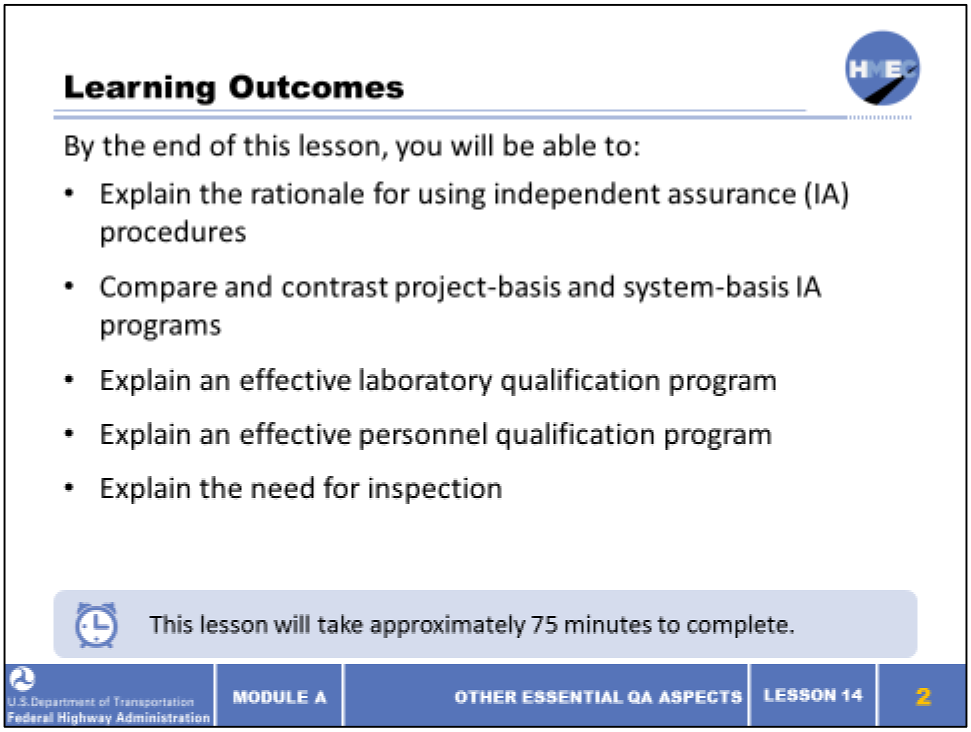
This page is intentionally blank.

Slide 1


HMEC
Highway Materials Engineering Course
Lesson 14: Other Essential QA Aspects
Quality Assurance
U.S.Department of Transportation
Federal Highway Administration
MODULE
A

____
____
____
____
____
____
____
____
____
____
____
____

Slide 2


Learning Outcomes
HMEC
By the end of this lesson, you will be able to:
• Explain the rationale for using independent assurance (IA) procedures
• Compare and contrast project-basis and system-basis IA programs
• Explain an effective laboratory qualification program
• Explain an effective personnel qualification program
• Explain the need for inspection
This lesson will take approximately 75 minutes to complete.
U.S. Department of Transportation
Federal Highway Administration
MODULE A
OTHER ESSENTIAL QA ASPECTS
LESSON 14
2

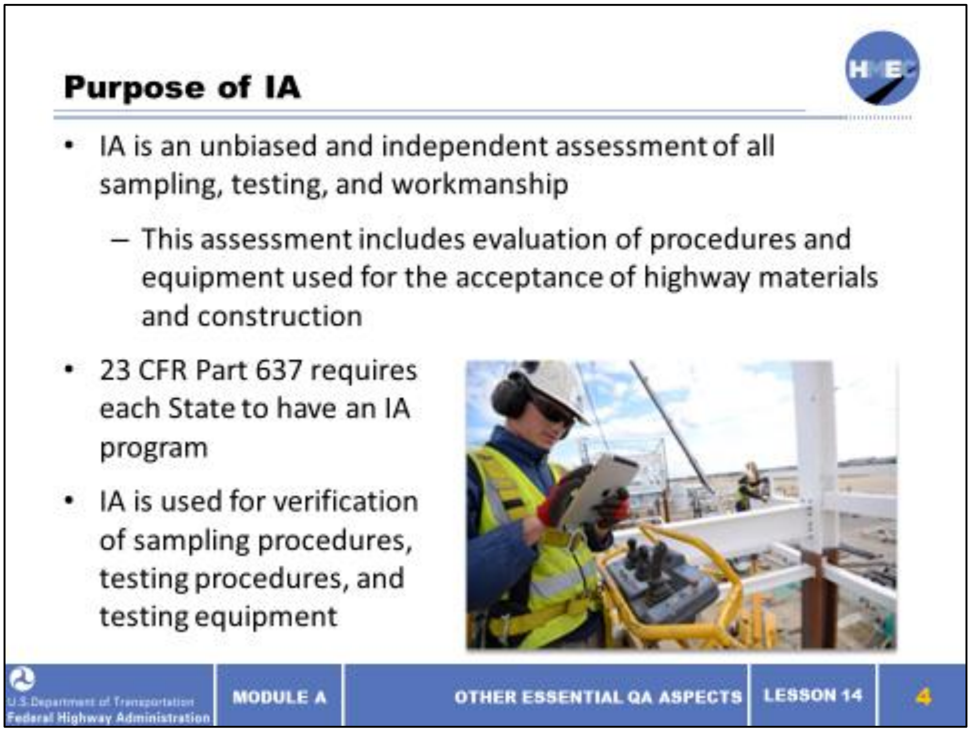
Slide 3

Definition of Independent Assurance (IA)

- Independent Assurance (IA)
  - Activities that are an unbiased and independent evaluation of all the sampling and testing (or inspection) procedures used in the quality assurance (QA) program
- IA provides an independent verification of the reliability of the acceptance (or verification) data obtained by the agency and the contractor
  - The results of IA testing or inspection are not to be used as a basis of acceptance
  - IA provides information for quality system management

U.S. Department of Transportation Federal Highway Administration | MODULE A | OTHER ESSENTIAL QA ASPECTS | LESSON 14 | 3

Slide 4



This assessment includes evaluation of procedures and equipment used for the acceptance of highway materials and construction. FHWA 23 CFR Part 637 requires each State to have an IA program. IA is used for verification of sampling procedures, testing procedures, and testing equipment.

Slide 5

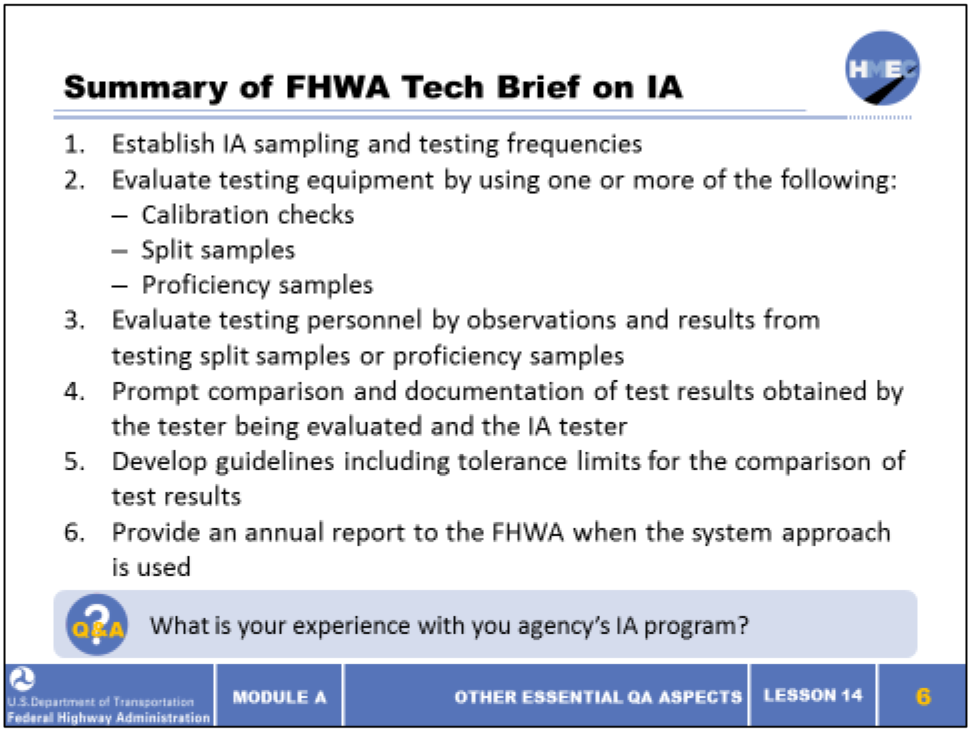
The Function of the IA Program

- IA is distinct from, and not intended as, an acceptance process or for use in verification of contractor sampling and testing results
  - Also distinct from, and not intended for, quality control (QC) purposes
- If IA results indicate a potential problem with quality, the results may be used to initiate additional testing

U.S. Department of Transportation Federal Highway Administration | MODULE A | OTHER ESSENTIAL QA ASPECTS | LESSON 14 | 5

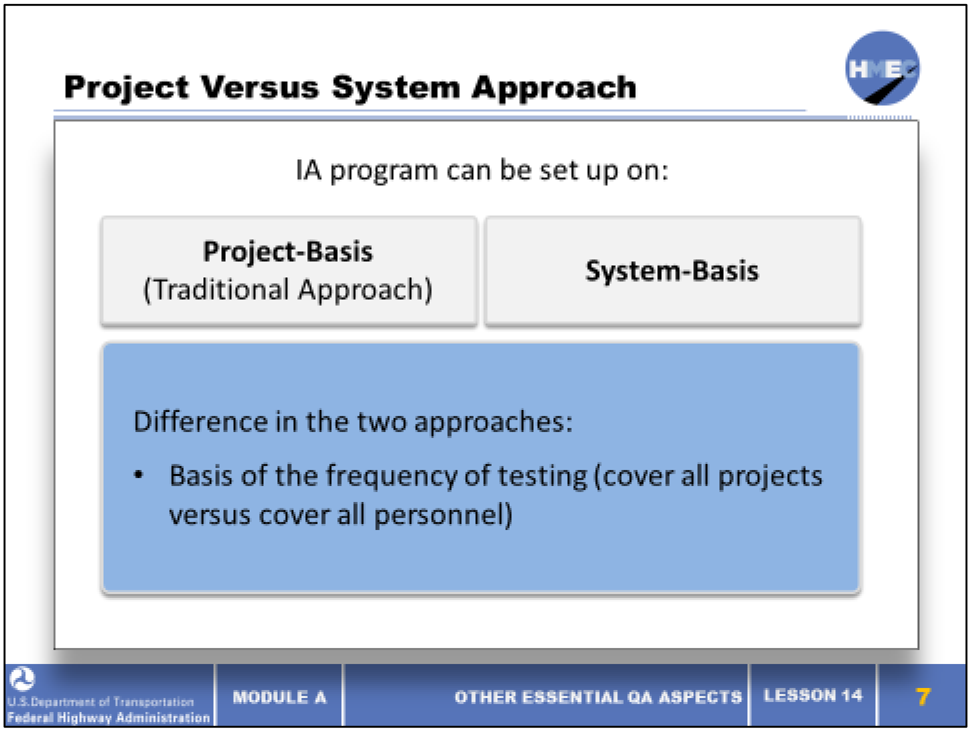
IA is distinct and not intended as an acceptance process or for use in verification of contractor sampling and testing results. It is also distinct and not intended for quality control (QC) purposes. If IA results indicate a potential problem with quality, the results may be used to initiate additional testing.

Slide 6



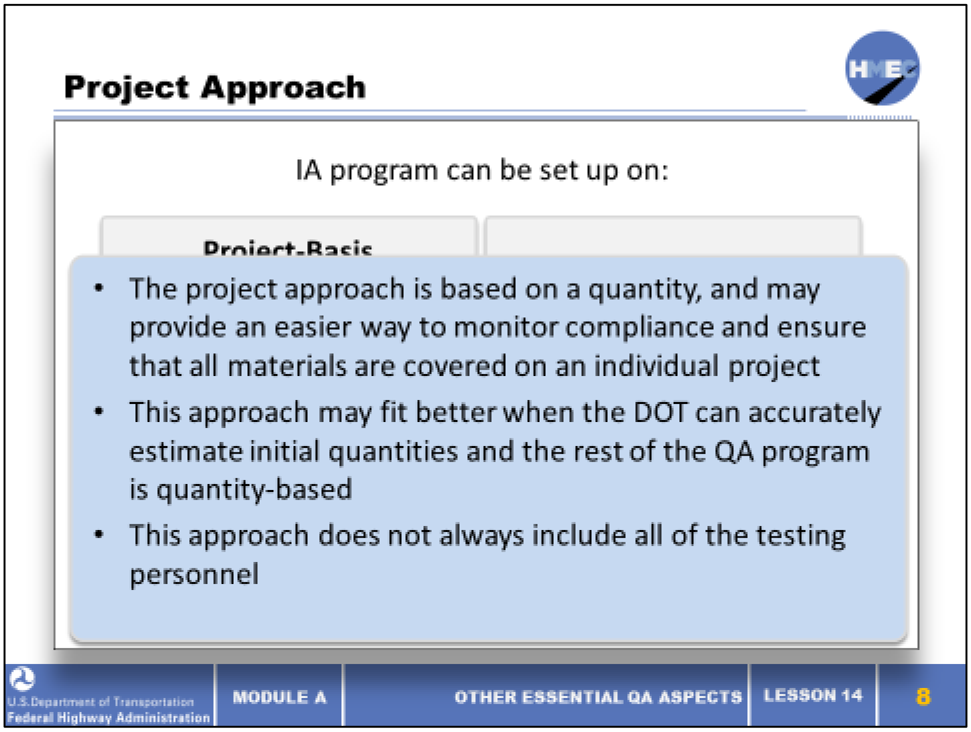
- Technical Brief on Independent Assurance: http://www.fhwa.dot.gov/pavement/materials/hif12001.pdf
- 23 CFR Part 637: http://www.access.gpo.gov/nara/cfr/waisidx_03/23cfr637_03.html

Slide 7



The difference in the two approaches is the basis of the frequency of testing (cover all projects versus cover all personnel).

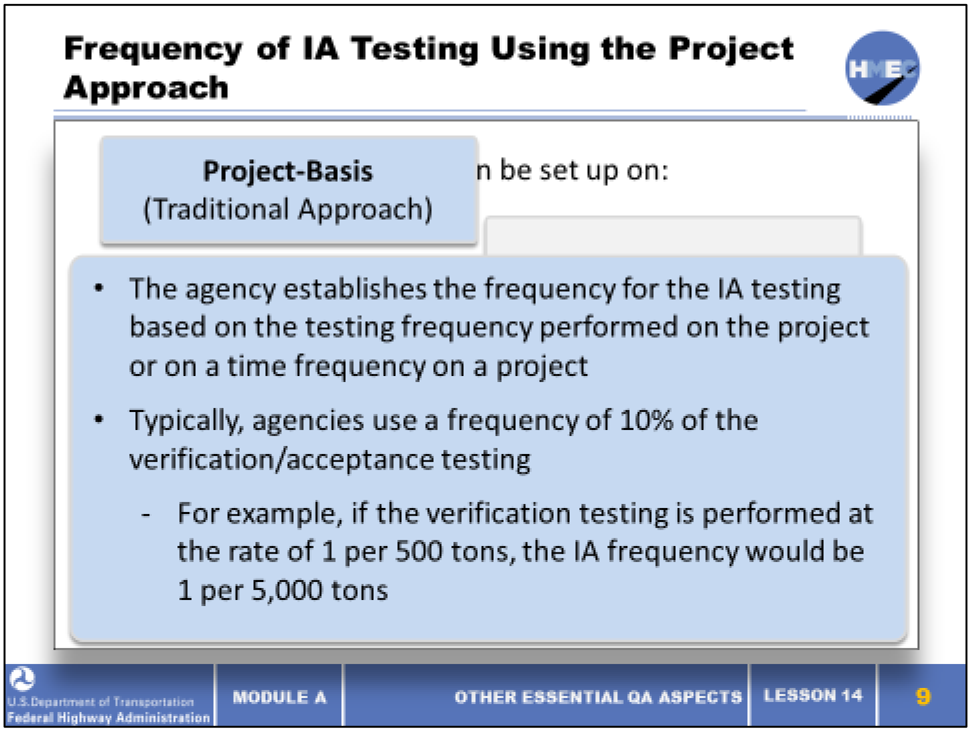
Slide 8



This approach may fit better when the DOT can accurately estimate initial quantities and the rest of the QA program is quantity-based.

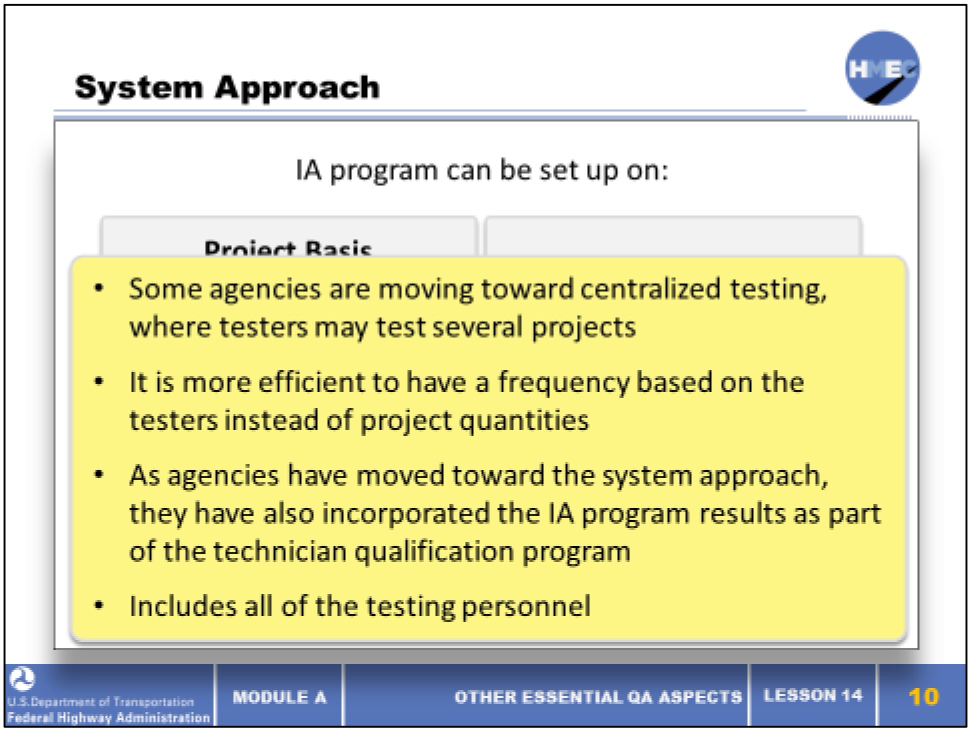
____
____
____
____
____
____
____
____
____
____
____

Slide 9



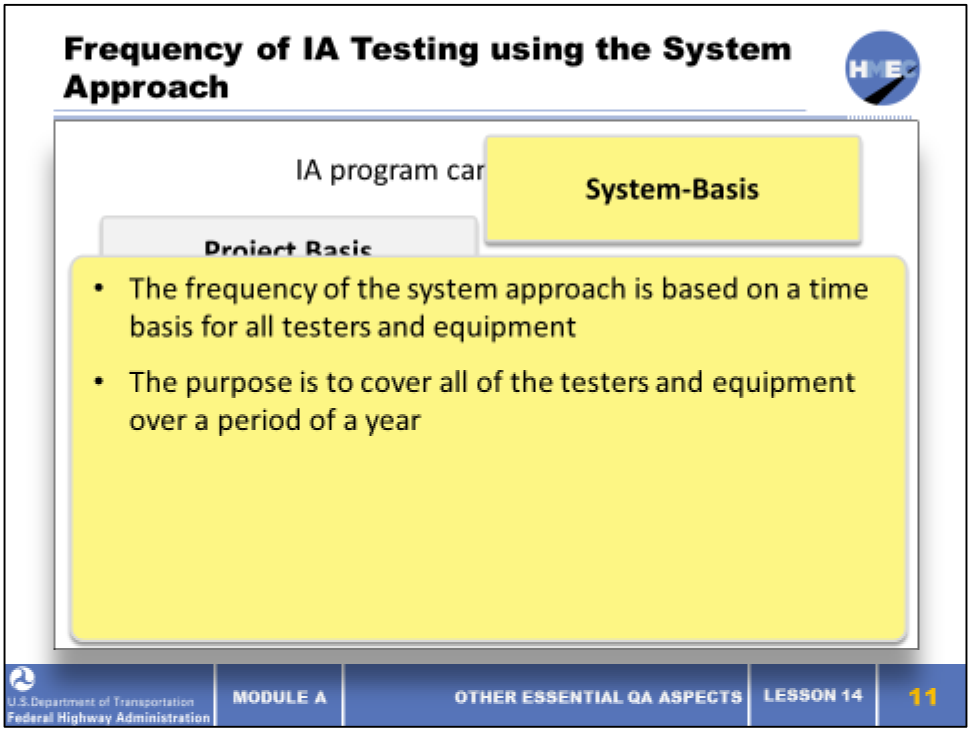
The agency establishes the frequency for the IA testing based on the testing frequency performed on the project or on a time frequency on a project. Typically, the agencies use a frequency of 10% of the verification/acceptance testing. For example, if the verification testing is performed at the rate of 1 per 500 tons, the IA frequency would be 1 per 5,000 tons.

Slide 10



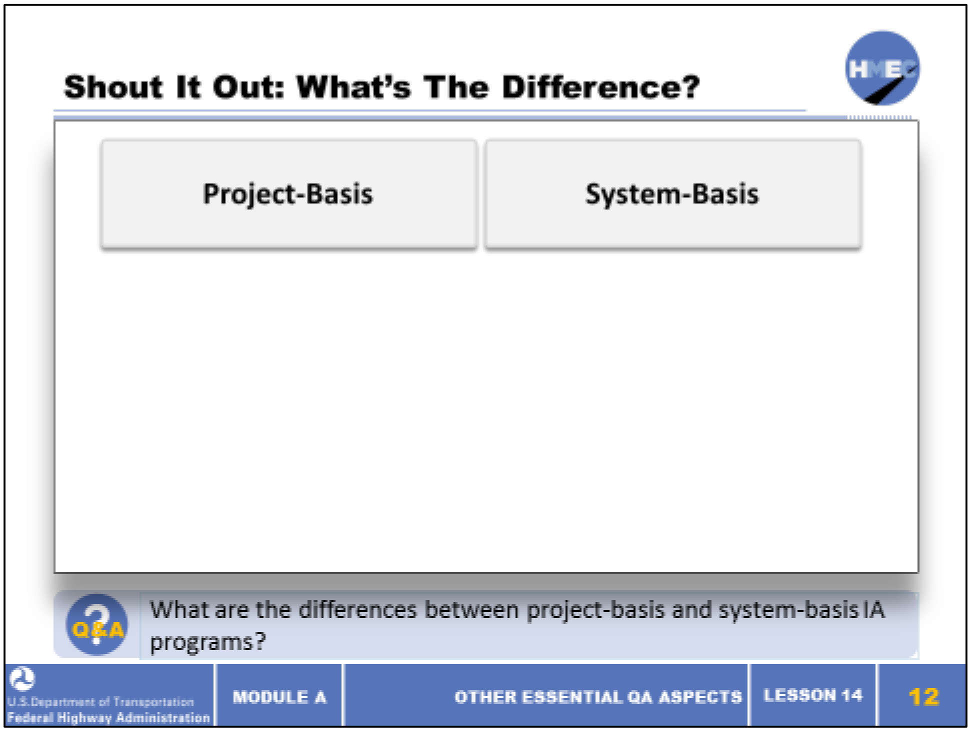
They are moving toward centralizing testing away from the project level. As this occurs, testers may perform testing on several projects and it becomes more efficient to have a frequency based on the testers instead of projects quantities. As agencies have moved toward the system approach, they have also incorporated the IA program results as part of the technician qualification program.

Slide 11

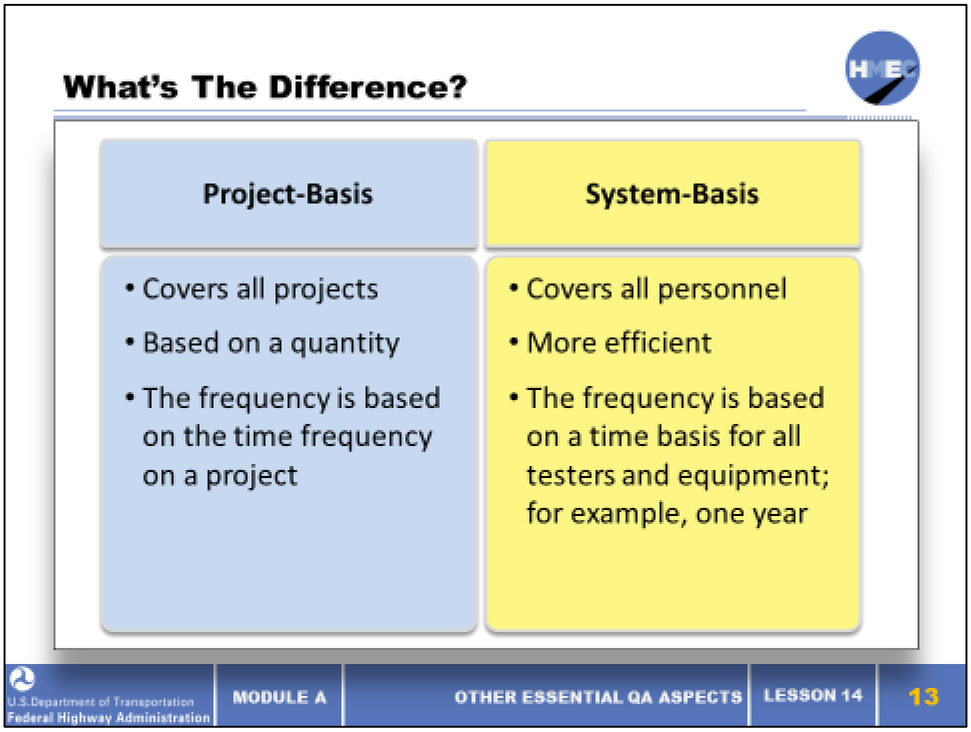


In this case, the personnel and equipment would be verified on a system basis. The purpose is to cover all of the testers and equipment over a period of a year. While States strive to reach all testers, it is not always possible. States typically set a goal of reaching 90% of the active testers. Active testers are defined as those testers that are performing testing in a given year. In most States, this is a subset that is smaller than all "qualified" testers since some qualified personnel may have retired, move to other jobs, or resigned. The system approach can be a more effective means of performing IA since it ensures that most testers are reviewed and that the same testers are not continually reviewed.

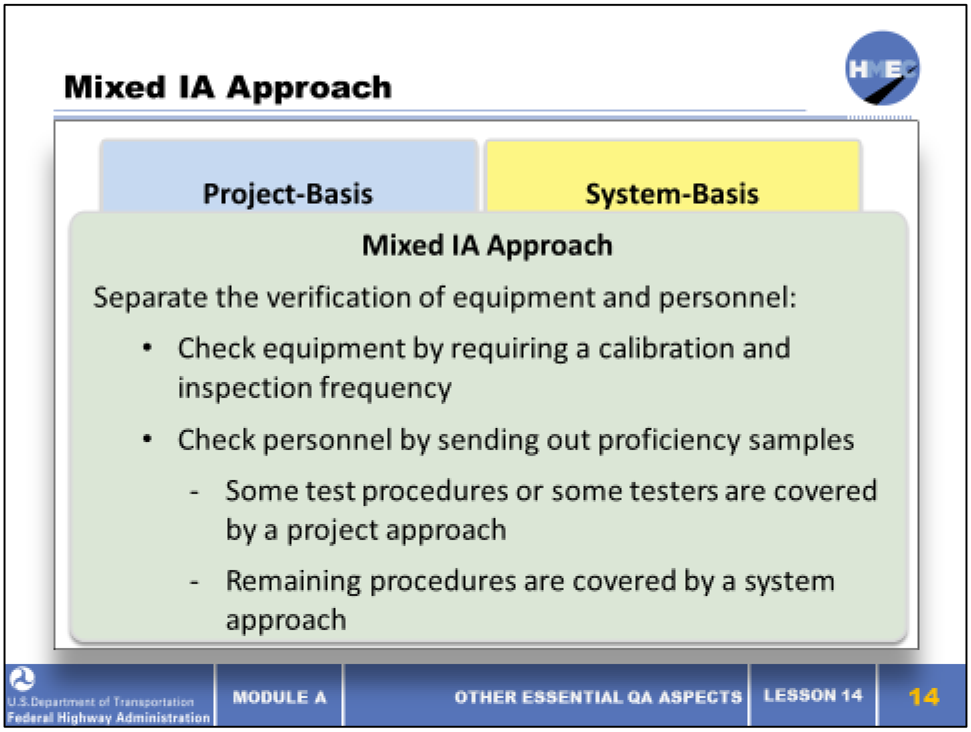
Slide 12


Shout It Out: What's The Difference?
HMEC
Project-Basis
System-Basis
Q&A
What are the differences between project-basis and system-basis IA programs?
U.S. Department of Transportation
Federal Highway Administration
MODULE A
OTHER ESSENTIAL QA ASPECTS
LESSON 14
12

Slide 13


What's The Difference?
Project-Basis
• Covers all projects
• Based on a quantity
• The frequency is based on the time frequency on a project
System-Basis
• Covers all personnel
• More efficient
• The frequency is based on a time basis for all testers and equipment; for example, one year
U.S. Department of Transportation
Federal Highway Administration
MODULE A
OTHER ESSENTIAL QA ASPECTS
LESSON 14
13

Slide 14



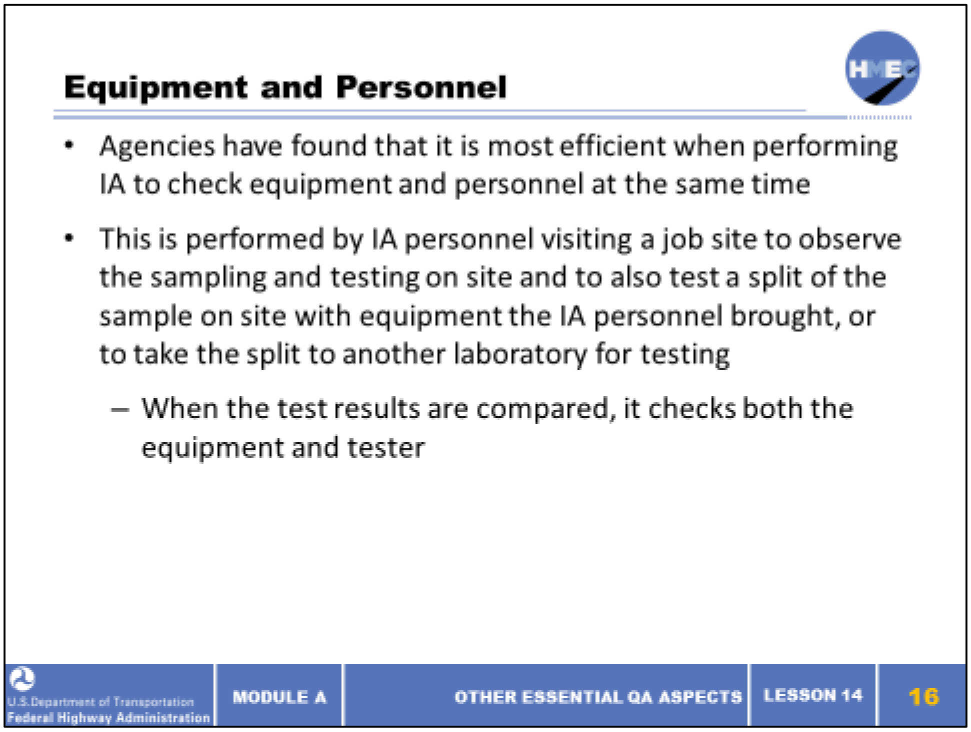
This is called the mixed approach. It is permissible to separate the verification of equipment and personnel, i.e., one method to check equipment is to require a calibration and inspection frequency. Personnel can be checked by sending out proficiency samples. It is permissible to use a mixed approach, i.e., where some test procedures and or some testers are covered by a project approach and the remaining procedures are covered by a system approach.

Slide 15

Three types of laboratories need to be accredited; state central labs, third party labs performing IA; and third party labs performing dispute resolution. Labs performing acceptance testing only need to be qualified.

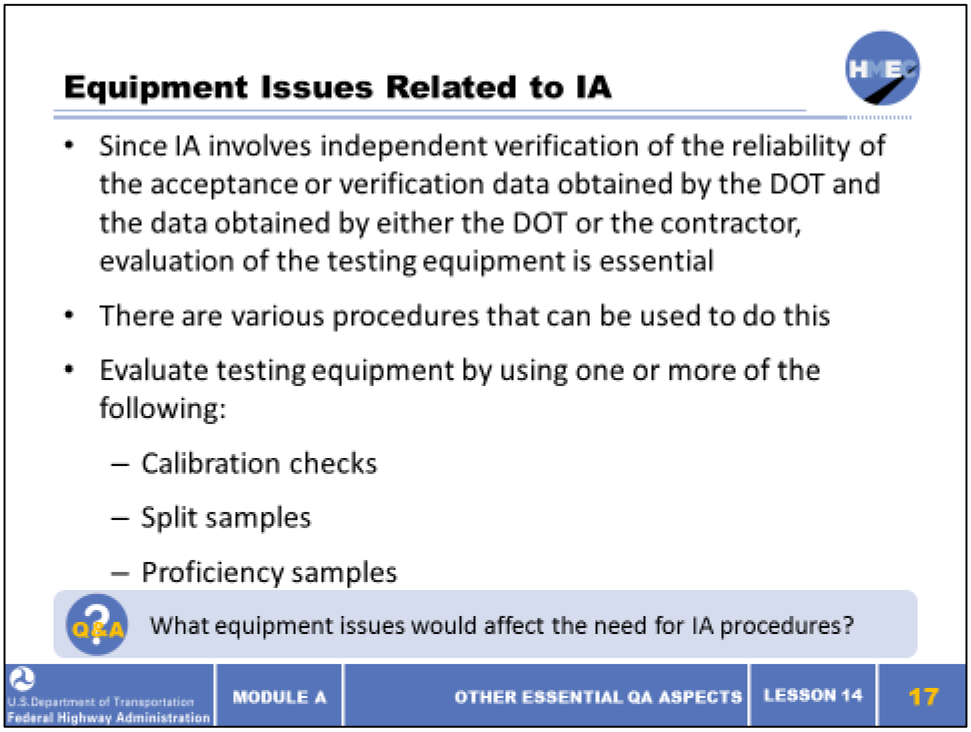
In Lesson 9, you learned about these issues as related to QC. But in the case of IA and dispute resolution, they must also be accredited.

Slide 16



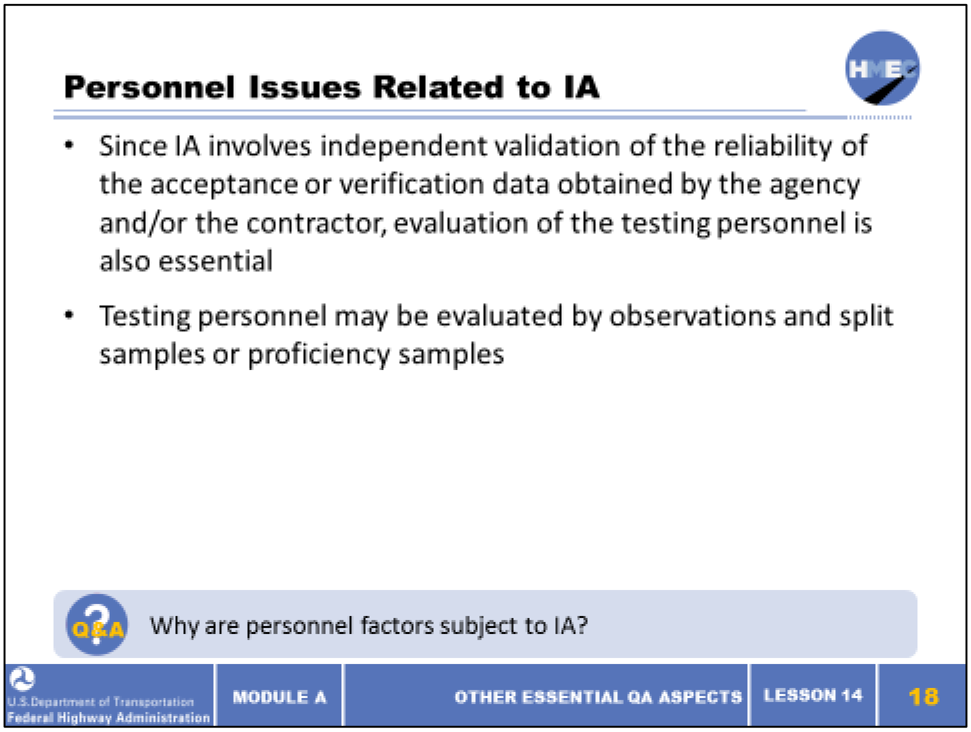
This is performed by IA personnel visiting a job site to observe the sampling and testing on site and to also test a split of the sample on site with equipment the IA personnel brought or to take the split to another laboratory for testing. When the test results are compared, it checks both the equipment and tester. If a set of samples do not compare, further analysis is required to determine if the source of the error is in procedure or the equipment.

Slide 17



Evaluate testing equipment by using one or more of the following: calibration checks, split samples, or proficiency samples.

Slide 18

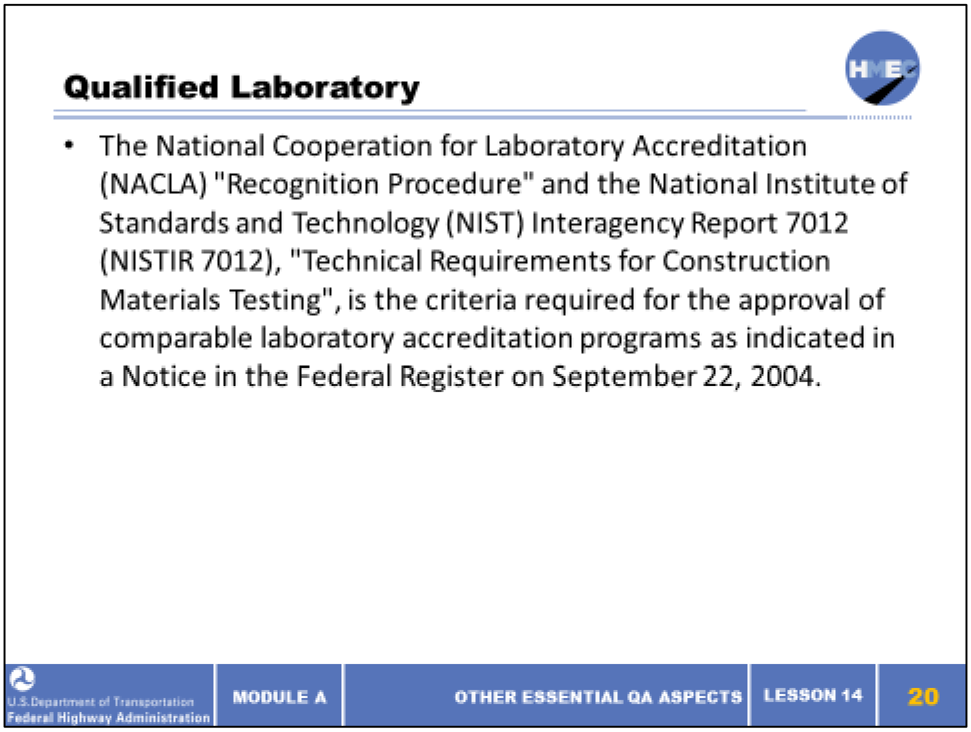


Testing personnel may be evaluated by observations and split samples or proficiency samples.

Slide 19

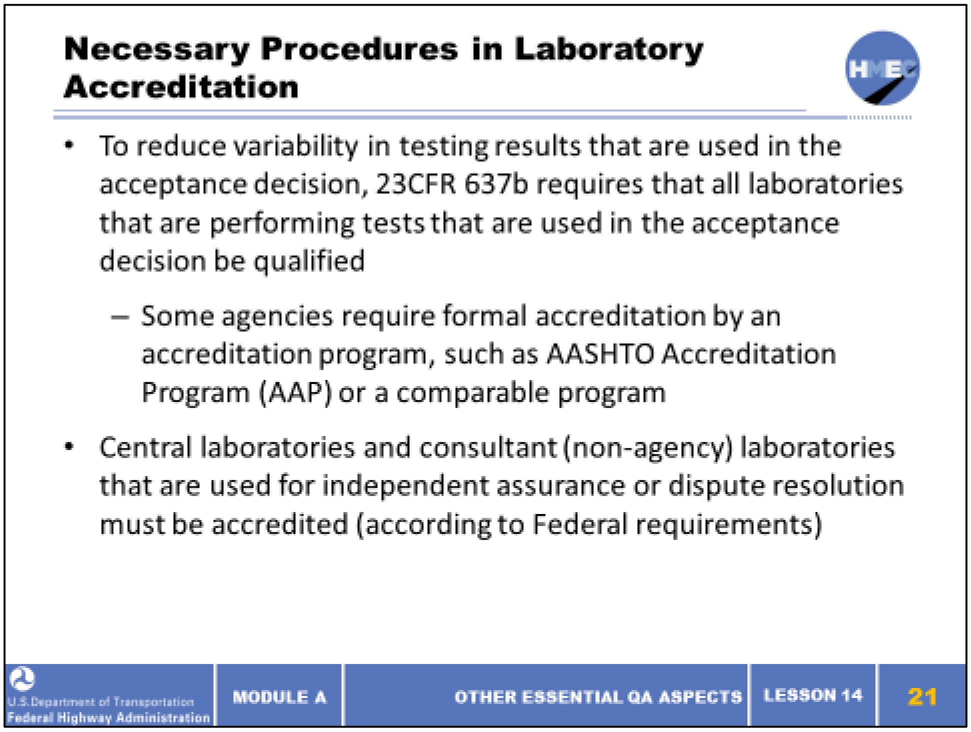
The requirements for qualified labs are defined by the state. In some cases, that may mean accreditation. By definition, accredited laboratories are laboratories that are recognized by a formal accrediting body as meeting quality system requirements including demonstrated competence to perform standard test procedures. The type of testing done in the lab is important. For example, a lab used for only QC must be qualified but does not have to be accredited.

Slide 20



The accreditation bodies will be evaluated against the NACLA Recognition Procedure and the Technical Requirements for Construction Materials Testing, and they must be recognized by NACLA with the Technical Requirements for Construction Materials Testing listed within its scope before the accreditation bodies will be approved by the Federal Highway Administration (FHWA). To meet the quality assurance requirements in 23 CFR 637.209(a)(2), (3), and (4), the laboratories' scope of accreditation must indicate that the laboratory was assessed according to the requirements in NISTIR 7012.

Slide 21



To reduce variability in testing results that are used in the acceptance decision, 23CFR 637b requires that all laboratories that are performing tests that are used in the acceptance decision be qualified; some agencies require formal accreditation by an accreditation program, such as the AASHTO Accreditation Program (AAP) or a comparable program. Central laboratories and consultant (non-agency) laboratories that are used for independent assurance or dispute resolution must be accredited (according to Federal requirements).

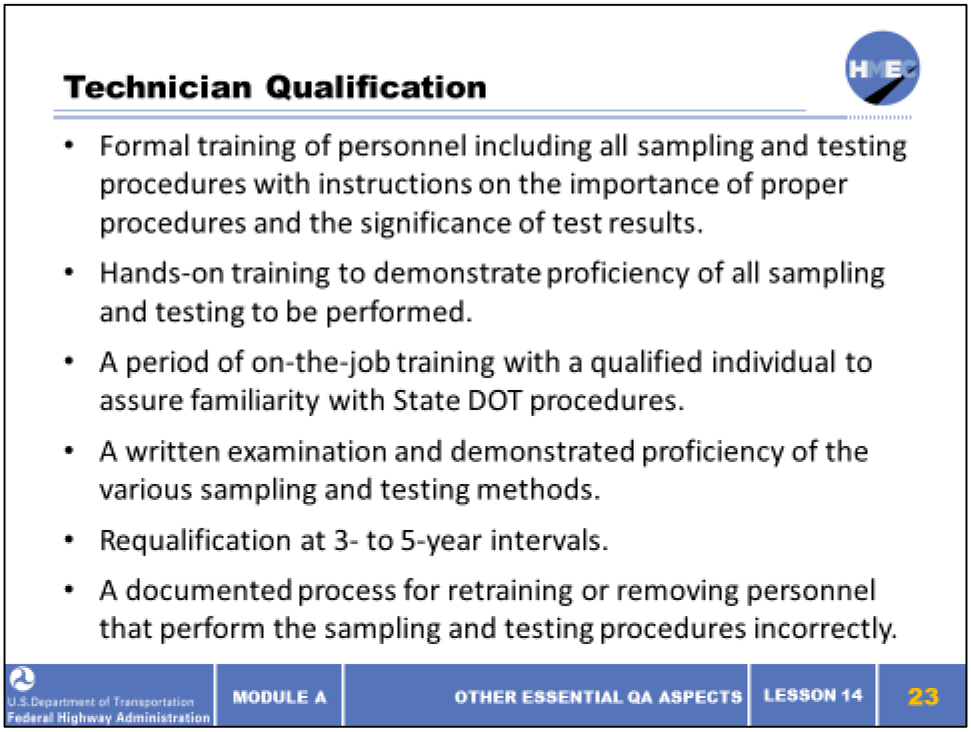
FHWA has also approved 2 other accreditation bodies to be comparable with the AAP. The Laboratory Accreditation Bureau (LAB) and the Construction Materials Engineering Council (CMEC).
http://www.fhwa.dot.gov/pavement/materials/120604.cfm
http://www.fhwa.dot.gov/pavement/materials/111007.cfm

Slide 22

Federal regulations require qualification and some agencies require certification of personnel as part of the qualification program.

FHWA has also approved 2 other accreditation bodies to be comparable with the AAP. The Laboratory Accreditation Bureau (LAB) and the Construction Materials Engineering Council (CMEC):
http://www.fhwa.dot.gov/pavement/materials/120604.cfm
http://www.fhwa.dot.gov/pavement/materials/111007.cfm

Slide 23



Here is some guidance for reviewing and revising a technician qualification program.

- Formal training of personnel including all sampling and testing procedures with instructions on the importance of proper procedures and the significance of test results.
- Hands-on training to demonstrate proficiency of all sampling and testing to be performed.
- A period of on-the-job training with a qualified individual to assure familiarity with State DOT procedures.
- A written examination and demonstrated proficiency of the various sampling and testing methods.
- Requalification at 3- to 5-year intervals (data from the IA program can be used as one element of requalification).
- A documented process for retraining or removing personnel that perform the sampling and testing procedures incorrectly.

The following are not appropriate criteria for achieving or maintaining qualification status: Grandfathering, the acceptance of a Professional Engineer or Engineer-in-Training certificate, or lifetime qualification.

___

___

___

___

___

___

___

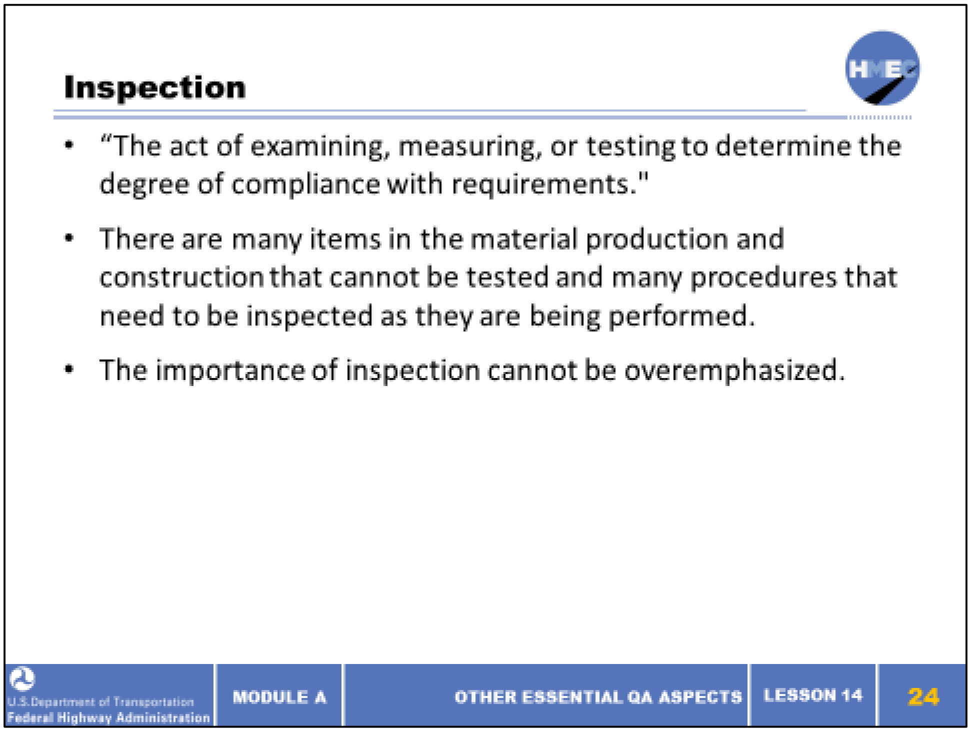
___

___

___

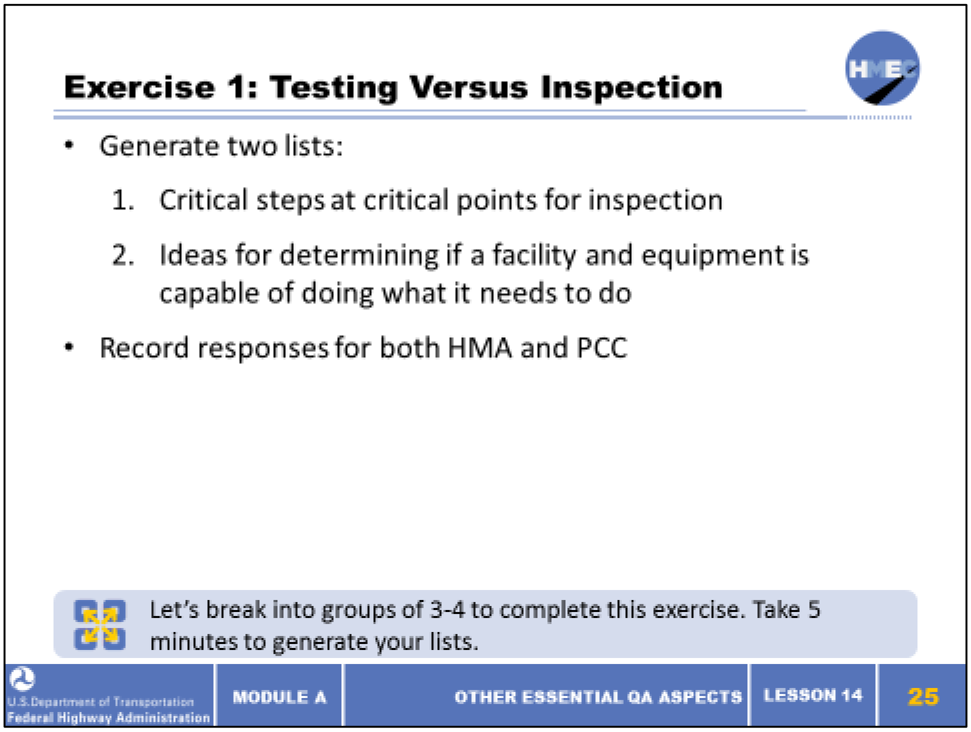
___

Slide 24



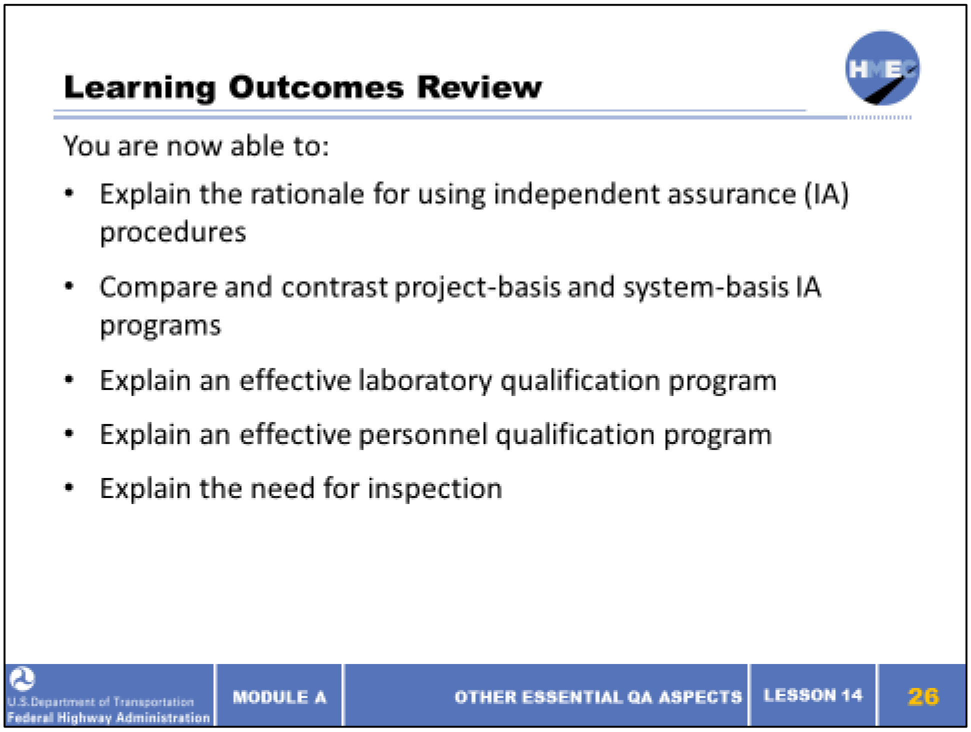
There are many items in the material production and construction that cannot be tested and many procedures that need to be inspected as they are being performed. The importance of inspection cannot be overemphasized.

Slide 25



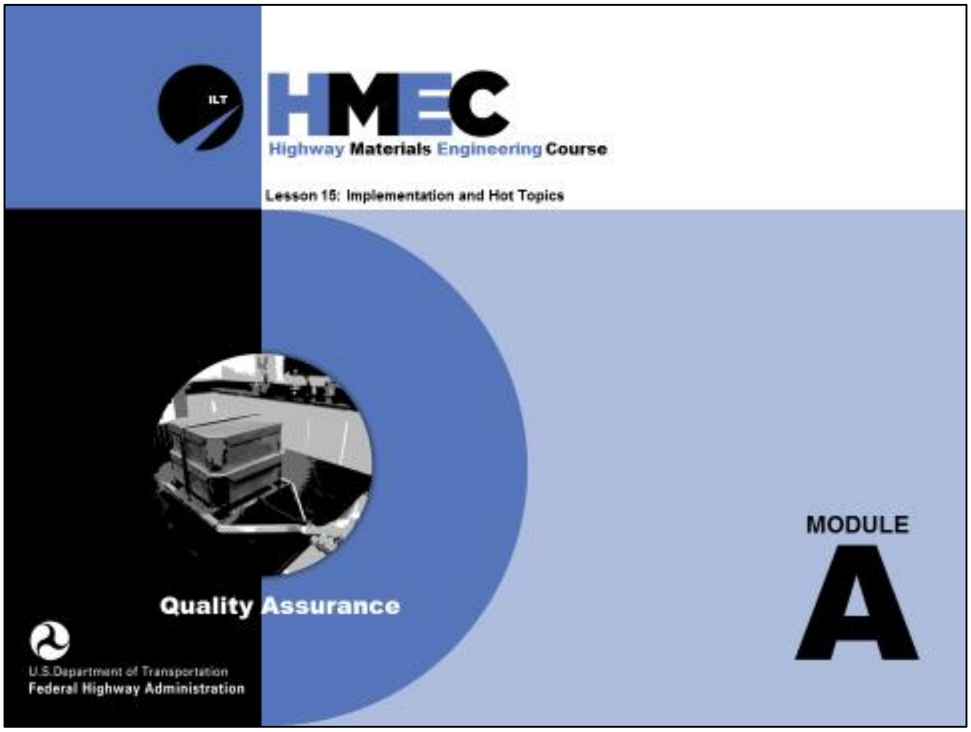
In plant operations and construction, you can't test everything. Also some items need to be inspected before they are made unavailable for testing. Production facilities, testing facilities, plant, and construction equipment all need inspection periodically to ensure they continue to produce a quality product.

Slide 26


Learning Outcomes Review
You are now able to:
• Explain the rationale for using independent assurance (IA) procedures
• Compare and contrast project-basis and system-basis IA programs
• Explain an effective laboratory qualification program
• Explain an effective personnel qualification program
• Explain the need for inspection
U.S. Department of Transportation
Federal Highway Administration
MODULE A
OTHER ESSENTIAL QA ASPECTS
LESSON 14
26

This page is intentionally blank.

Slide 1


HMEC
Highway Materials Engineering Course
Lesson 15: Implementation and Hot Topics
Quality Assurance
U.S.Department of Transportation
Federal Highway Administration
MODULE
A

This page is intentionally blank.

Appendix A: Acronyms

The following are acronyms referenced throughout the course that are important agencies or organizations:

Acronym	Proper Name
AASHTO	American Association of State Highway and Transportation Officials
ACAA	American Coal Ash Association
ACI	American Concrete Institute
ACPA	American Concrete Paving Association
AI	Asphalt Institute
ASTM	American Society for Testing and Materials
AWS	American Welding Society
CFR	Code of Federal Regulations
DOT	U.S. Department of Transportation
EPA	Environmental Protection Agency
FHWA	Federal Highway Administration
NACE	National Association of Corrosion Engineers
NAPA	National Asphalt Pavement Association
NCAT	National Center for Asphalt Technology
NCHRP	National Cooperative Highway Research Program
NEPCOAT	North East Protective Coating
NHI	National Highway Institute
NRC	National Recycling Coalition

Acronym	Proper Name
NRMCA	National Ready Mixed Concrete Association
NSA	National Slag Association
NSBA	National Steel Bridge Alliance
NTPEP	National Transportation Product Evaluation Program
OSHA	Occupational Safety and Health Administration
RCSC	Research Council on Structural Connections
SSPC	Society for Protective Coatings
TRB	Transportation Research Board
USGS	U.S. Geological Survey

Appendix B: Resources

Additional information regarding Module A can be found in the following sources.

Primary Resources

- AASHTO Quality Assurance Guide Specifications
- ASTM Manual on Presentation of Data and Control Chart Analysis
- NHI-131023 Module A, Quality Assurance (updated July 2009)

Additional Resources

- Basic Math Symbols
http://www.rapidtables.com/math/symbols/Basic_Math_Symbols.htm
- FHWA Code of Federal Regulations 23 Part 637b (pgs. 210-213)
http://www.gpo.gov/fdsys/pkg/CFR-2003-title23-vol1/pdf/CFR-2003-title23-vol1.pdf
- FHWA Percent Within Limits Workshop
https://www.fhwa.dot.gov/pavement/pwl/basic_pwl.cfm
- NCHRP Project – Research Results Digest 371
http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rrd_371.pdf
- Random Number Table
https://harpersenterprise.adobeconnect.com/_a1005638364/randomnumbertable/
- Technical Advisory 6120.3 https://www.fhwa.dot.gov/construction/t61203.cfm
- Technical Brief on Independent Assurance:
http://www.fhwa.dot.gov/pavement/materials/hif12001.pdf
- TRB Glossary http://onlinepubs.trb.org/onlinepubs/circulars/ec173.pdf

AASHTO/ASTM Standards

- AASHTO R-9
- AASHTO R-10
- AASHTO R-18
- AASHTO R-25
- AASHTO R-38
- AASHTO R-44
- ASTM D 2726
- ASTM E 178