Interim Research Report
Research Project Agreement T9903.2, Task A5
Asphalt Concrete Specifications

A Quantification and Evaluation of WSDOT's Hot Mix Asphalt Concrete Statistical Acceptance Specification

by

Stephen T. Muench
Graduate Student
Department of Civil and Environmental Engineering
University of Washington

Joe P. Mahoney
Professor
Department of Civil and Environmental Engineering
University of Washington

Washington State Transportation Center (TRAC)
University of Washington, Box 354802
University District Building
1107 NE 45th Street, Suite 535
Seattle, Washington 98105-4631

Washington State Department of Transportation
Technical Monitor
Jim Walter
Construction Materials Engineer

Prepared for

Washington State Transportation Commission
Department of Transportation
and in conjunction with
U.S. Department of Transportation
Federal Highway Administration

October 2001

TECHNICAL REPORT STANDARD TITLE PAGE

1. REPORT NO.	2. GOVERNMENT ACCESSION NO.	3. RECIPIENT'S CATALOG NO.
WA-RD 517.1		

4. TITLE AND SUBTITLE	5. REPORT DATE
A Quantification and Evaluation of WSDOT's Hot-Mix Asphalt Concrete Statistical Acceptance Specification	October 2001
	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S)	8. PERFORMING ORGANIZATION REPORT NO.
Stephen T. Muench, Joe P. Mahoney	
9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. WORK UNIT NO.
Washington State Transportation Center (TRAC) University of Washington, Box 354802 University District Building; 1107 NE 45th Street, Suite 535 Seattle, Washington 98105-4631	11. CONTRACT OR GRANT NO. Agreement T9903, Task A5
12. SPONSORING AGENCY NAME AND ADDRESS	13. TYPE OF REPORT AND PERIOD COVERED
Research Office Washington State Department of Transportation Transportation Building, MS 47370 Olympia, Washington 98504-7370 Keith Anderson, Project Manager, 360-709-5405	Interim research report 14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

This study was conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration.

16. ABSTRACT

This report clarifies the Washington State Department of Transportation (WSDOT) hot-mix asphalt (HMA) statistical acceptance specification's statistical basis and how its key components influence overall specification performance and contractual pay.

The WSDOT specification uses a variable sampling plan to measure in-place density, asphalt content, and aggregate gradation. Specification bands on these three quality characteristics are reasonable and consistent with typical material, sampling, testing, and construction variability. Pay is determined by calculating a percentage within limits (PWL), then applying a series of parabolic pay equations, depending on sample size.

The methods used by WSDOT balance risk between the contractor and WSDOT well but result in two issues that differ from similar specifications: (1) Expected pay for material produced at acceptable quality level (AQL) is greater than 1.0, and (2) WSDOT's AQL is 95 PWL, but contractors seem to consistently produce material near 90 PWL The issues are not critical, and therefore, the specification should not be changed. However, if the third and final report in this series, which studies quality characteristics to be measured for Superpave design mixes, results in major recommended changes, then we recommend reviewing the AQL and pay factors to bring the specification into line with actual practice.

17. KEY WORDS	18. DISTRIBUTION STATEMENT
Specification, statistics, HMA, hot mix asphalt, asphalt concrete, pay factor, statistical acceptance specification, Washington State	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22616

19. SECURITY CLASSIF. (of this report)	20. SECURITY CLASSIF. (of this page)	21. NO. OF PAGES	22. PRICE
None	None		

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Washington State Transportation Commission, Department of Transportation, or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

Since 1989 the Washington State Department of Transportation (WSDOT) has used a statistical acceptance specification to pay contractors for hot mix asphalt concrete (HMA). Statistical acceptance specifications provide a better method of specifying HMA construction than previously used specifications often referred to as "method" specifications. Where method specifications subjectively inspect the process that produces HMA pavement, statistical acceptance specifications monitor the contractor's process and accept or reject the end product based on objective specified measurements (Bowery and Hudson, 1976). This creates several key advantages of statistical acceptance specifications over method specifications (Bowery and Hudson, 1976):

- Responsibility for material and construction quality resides with the party that can best control these factors: the contractor.
- The contractor is allowed greater latitude in choice of materials, equipment, and method, which allows more control over material and construction quality as well as contract profitability.
- Acceptance / rejection decisions are objective, consistent and statistically defensible.
- Quick inspection and pay calculations on a relatively small subsections of material / construction give contractors the opportunity to take corrective action before large quantities of out-of-specification material or construction are produced.

In order to fully exploit these advantages while avoiding potential pitfalls, the statistical acceptance specification in use needs to be well-understood by all parties involved. This report clarifies the WSDOT HMA statistical acceptance specification's statistical basis and how its key components influence overall specification performance and contractual pay. Specifically, this clarification of the WSDOT HMA statistical acceptance specification will:

1. *Provide a general background on statistical acceptance specifications*. This will define statistical acceptance specifications, their applications and their limitations.

2. *Quantify the WSDOT statistical acceptance plan's key components*. This will provide a common knowledge base as well as provide the background for determining the implications of each component.

3. *Evaluate the WSDOT statistical acceptance plan's key components and compare them to key components of the Federal Highway Administration (FHWA) and the Federal Aviation Administration (FAA)*. This evaluation and comparison will give a sense which WSDOT statistical acceptance specification components are most critical and how they compare to national specifications of the same type.

2 BACKGROUND

This section provides a practical background for statistical acceptance specifications by establishing what statistical acceptance specifications are, why they are used, and what their applications and limitations are. A statistical acceptance specification is only one way of accepting or rejecting work; other methods such as accepting with no inspection and full inspection are also commonly used, however each method has its specific areas of applicability. Regardless of the details of accept / reject methods, they are all only designed to monitor or audit a process. They do not provide any means of process or product improvement and should not be used to achieve such goals.

2.1 Acceptance Plan Basics

In general, a statistical acceptance specification is simply an acceptance procedure. An acceptance procedure is a formal procedure used to decide whether work should be accepted, rejected, or accepted at a reduced payment (Freeman and Grogan, 1998). This makes acceptance procedures a form of quality assurance. Specifically, they are *monitoring* methods used to determine whether or not a particular process is meeting quality standards. Acceptance procedures are not, however, a form of quality control. "Quality control" refers to a system employed to ensure the maintenance of proper quality standards within a project. Acceptance procedures simply accept or reject lots based on their quality; they do not ensure proper quality standards. Acceptance procedures should never be used as a method to control or improve quality; process controls are used to control and systematically improve quality (Montgomery, 1997).

Acceptance procedures can take one of the following three broad forms: (Montgomery, 1997):

1. *Accept with no inspection* is generally used when there is no economic justification to look for defective units or material.

2. *100 percent inspection* is generally used where components or material are extremely critical and passing any defective components or material would result in an unacceptably high failure cost.

3. *Acceptance sampling* is generally used when there is some economic justification to look for defective material and either (1) some small finite percentage of defective material is acceptable or (2) it is not economical or practical to use 100 percent inspection. Acceptance sampling uses statistics to estimate information about an entire lot from a small random sample.

Of these three approaches, HMA construction typically uses acceptance sampling because excessive out-of-specification (defective) material will substantially affect long-term pavement performance but it is neither practical nor economical to inspect everything. Basically, acceptance sampling uses random sampling to make quality and material property estimates about a large amount of material. This highlights two key concepts involved in the effective use of acceptance sampling: (1) acceptance sampling only estimates material properties, and (2) acceptance sampling depends on random sampling.

Acceptance sampling uses a small number of random samples to draw conclusions about a large amount of material (called a "lot"). Since the entire lot is not inspected, these conclusions are only estimates of actual lot properties and will therefore involve some amount of uncertainty as to their accuracy. The only way to determine lot properties with certainty is to test the entire lot (100 percent inspection).

Acceptance samples must be random. If samples are not random then the statistical basis for evaluating them and drawing conclusions about an entire lot is invalid. Thus, any exercise in judgment as to whether or not a sample will produce a good, failing, or average test result nullifies the random sampling assumption and therefore the assumptions on which a statistically oriented specification is based (Bowery and Hudson, 1976).

HMA construction acceptance sampling uses a modified version of random sampling that satisfies the random sampling assumption. In true random sampling any location or item within a lot must have an equal probability of being sampled. In rare instances, this results in all samples being clustered together through random chance (Freeman and Grogan, 1998). Although this sample clustering is statistically valid, HMA specifications usually strive to ensure samples are spread more evenly throughout the lot. Therefore, stratified random sampling, which involves dividing lots into several equal-sized sublots, is generally used (Weed, 1982 as cited in Freeman and Grogan, 1998). Each individual sublot is still randomly sampled, but stratification ensures that samples are more evenly spread throughout the lot. Stratified random sampling conforms to the requirements of random sampling as long as three rules are obeyed (Weed, 1982 as cited in Freeman and Grogan, 1998):

1. The same number of samples are taken from each sublot.
2. Sublots are of equal size.
3. Samples are selected randomly from within sublots.

For example, WSDOT divides its in-place density lots into five equal-sized sublots and takes one random sample from each sublot (WSDOT, 2000b).

In summary, acceptance sampling is only one of several acceptance procedure options. As such, it does not provide any direct form of quality control; it simply accepts or rejects lots. Acceptance sampling only makes estimates of actual lot properties and is dependant on random sampling to make these estimates. In order to ensure a uniform distribution of samples, the HMA construction industry (both in Washington and nationwide) uses stratified random sampling.

2.2 Acceptance Plan Components

Acceptance sampling has been in general use for well over 60 years (Montgomery, 1997). Therefore, the theoretical underpinnings behind acceptance sampling are well proven. The key is then to appropriately apply acceptance sampling and its associated statistics to the HMA construction industry to create a viable overall plan. Correct application involves proper implementation of the following acceptance sampling components: (1) acceptance sampling type, (2) quality characteristics, (3) specification limits, (4) statistical model, (5) quality level goals, (6) risk and (7) pay factors. Decisions regarding these components will significantly impact final acceptance plan performance.

2.2.1 Types of Acceptance Sampling

There are two basic types of acceptance sampling: (1) attribute sampling, and (2) variable sampling. Both attribute and variable sampling are used in HMA construction; however variable sampling is more prevalent (Bowery and Hudson, 1976; Schmitt, et al., 1998).

In attribute sampling, each sample is inspected for the presence or absence of one or several attributes (often called quality characteristics). Measurements used to detect these quality characteristics are not retained. Rather, they are compared to a standard then recorded as either passing or failing. WSDOT's asphalt concrete aggregate fracture test is an example of attribute sampling. Aggregate is accepted or rejected based on a minimum quality characteristic of one fractured face on a specified percentage of the material. The actual percentage of fractured face is not recorded; instead, a simple pass-fail record is used.

In variable sampling, measured quality characteristics are used as continuous variables, which means that, unlike attribute sampling, measurement values are retained. Because these values are retained rather than converted into a discrete pass-fail criterion, variable sampling plans retain more information per sample than do attribute sampling plans (Freeman and Grogan, 1998). This means that compared to attribute sampling, it takes fewer variable samples to get the same information. Because of this, most HMA statistical acceptance plans use variable sampling.

However, variable sampling does have disadvantages. Foremost, variable sampling plans are predicated on a known distribution of the measured property. Therefore, most variable sample plans assume a normal distribution of the measured property. For instance, WSDOT acceptance testing for asphalt concrete compaction assumes that in-place asphalt concrete densities (the measured property) are normally distributed. If this normal distribution assumption is not true then the resulting estimates of lot quality will be wrong. Fortunately, construction-related lot characteristics are usually normally distributed (Markey, Mahoney, and Gietz, 1994; Aurilio and Raymond, 1995; Cadicamo, 1999).

Therefore, although both attribute and variable sampling are used in HMA construction, variable sampling is more prevalent because it provides more information per sample and its necessary assumption of a normal distribution of the quality characteristic is usually satisfied.

2.2.2 Quality Characteristic Selection

Quality characteristics are those material characteristics or properties that a particular acceptance plan measures to determine quality. Quality characteristics can be any measurable material or construction property but they must be carefully selected for two reasons: (1) their quality should accurately reflect overall project quality and (2) they should be relatively independent of one another.

Construction contracts, including HMA contracts, generally require full payment at substantial completion. However, since the constructed HMA pavement performs for many years after construction, contracting agencies usually use some predictive method to relate construction quality to long-term pavement performance. Statistical acceptance plans typically accomplish this by choosing construction quality characteristics that are most predictive of pavement performance. These quality characteristics typically include mix properties (such as aggregate gradation, asphalt content, and mix volumetrics), in-place density, and pavement smoothness (Schmitt, et al., 1998).

Quality characteristics must also be chosen to avoid correlation with one another. If not carefully selected, a change in one quality characteristic (such as aggregate gradation) could result in a change in another quality characteristic (such as voids in the mineral aggregate - VMA). Lin, Solaimanian, and Kennedy (2001) point out that this correlation will always cause biases in pay factor determination. In the gradation-VMA instance mentioned previously, the bias occurs because a poorly graded aggregate would be penalized not only by lower pay for poor gradation but also by lower pay for the correlated poor VMA. Bias in the opposite direction (higher pay for well-graded aggregate) is equally likely. Therefore, biased pay factors will unfairly penalize either WSDOT or the contractor.

Acceptance sampling determines overall construction quality by measuring quality characteristics. Proper selection of these characteristics ensures that (1) their quality accurately reflects construction quality, which should in turn reflect long-term pavement performance and (2) they are relatively independent of one another so that final pay is not biased in either direction.

2.2.3 Specification Limits

Specification limits establish limits for quality characteristic measurements to differentiate between adequate material, and inadequate or defective material. For instance, the WSDOT lower specification limit for in-place asphalt concrete pavement density is typically 91.0 percent of the reference theoretical maximum density (TMD) (WSDOT, 2000b). Therefore, a measurement of 91.0 percent of TMD or higher represents adequate density while a measurement below 91.0 percent of TMD represents inadequate density.

Specification limits must be based on sound engineering judgment and sound statistical analysis. Specifically, engineering judgment is used to establish a target value for each quality characteristic and statistical analysis is used to establish an acceptable range around the target value. This range is used to account for the various sources of variability inherent in producing and testing HMA. Specifically, there are four types of variability to consider: (Hughes, 1996):

1. *The material's inherent variability* is the true random variation of the material and is a function of material characteristics alone. A contractor's manufacturing and construction process cannot control this variability.

2. *Sampling variability* is the variation in sample characteristics from sample-to-sample that is attributable to variations in sampling technique. A contractor's manufacturing and construction process cannot control this variability.

3. *Testing variability* is the lack of repeatability of test results. Operators, equipment condition, calibration, and test procedure all contribute to testing variability. A contractor's manufacturing and construction process cannot control this variability.

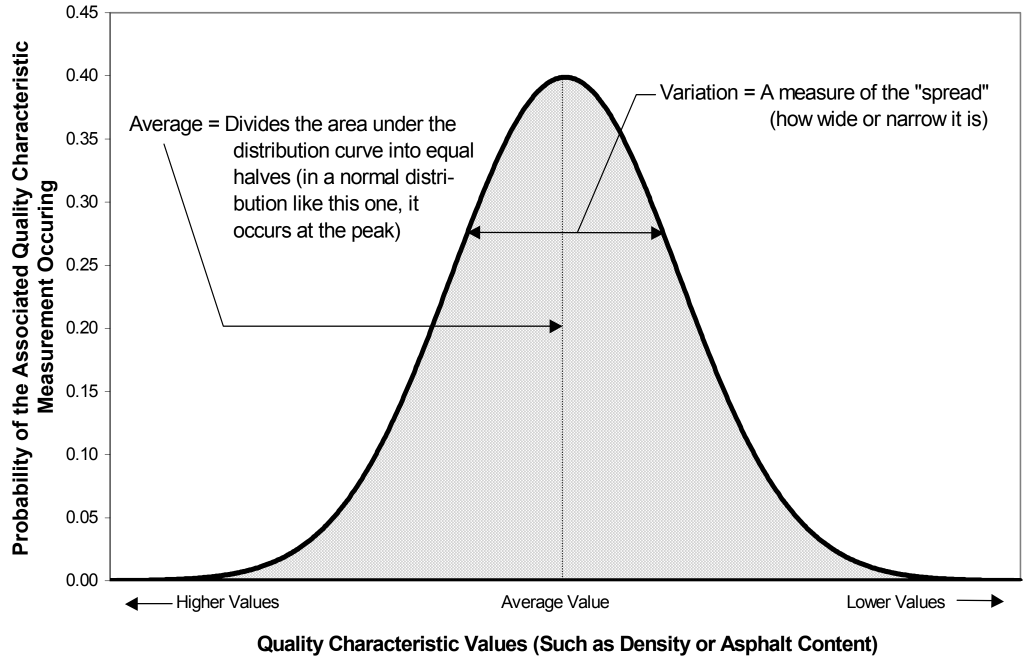
4. *Manufacturing and construction variability* is the variation in material caused by the manufacturing and construction process. These variations can be extremely localized within a lot and therefore difficult to detect by random sampling (like density differentials and pavement thickness variations) or they can be more global (e.g. between lots or days) and therefore more easily detected by random sampling (like changes in asphalt content or aggregate gradation between lots). Contractor quality control can minimize these types of variability.

Since contractors can only control manufacturing and construction variability, if the sum of inherent material, sampling and testing variability is greater than the allowable specification band, a potentially large amount of material will be judged out-of-specification for no contractor-correctible reason. For instance, an asphalt content specification of the JMF ± 0.1 percent does not make statistical sense because the combination of inherent asphalt content variability, sampling variability, and testing variability will typically cause test results to vary by more than ±0.1 percent on either side of the JMF (Hughes, 1996). A more practical approach (such as WSDOT's), which adequately accounts for material, sampling, and testing variability might specify the JMF asphalt content ±0.5 percent. In sum, specification limits should be tight enough to detect manufacturing and construction variability, but loose enough to allow a reasonable amount of testing, sampling, and inherent material variability.

2.2.4 Statistical Model

The statistical model used by an acceptance plan determines how the plan relates actual random sample test results to the distribution of the quality characteristic within the lot. This distribution is then used to determine lot quality.

Statistical models all rely on random samples, which provide two pieces of data: (1) the average of the sample measurements and (2) the variation in sample measurements. Both pieces of data are needed to estimate the distribution of the measured quality characteristic within a lot (see Figure 1).



Note: This distribution represents hypothetical quality characteristic measurement results if an entire lot were broken down into infinitesimally small sections and the quality characteristic associated with each section was measured. As stated earlier, this distribution can never be known for certain unless a 100 percent inspection method is used.

Figure 1: A Generic Example of a Quality Characteristic Distribution

There are typically three different ways of using sample data:

1. *Use the average of sample measurements only*. This method calculates the sample average and uses this to estimate lot average. It does not calculate sample variation, thus it is unable to estimate the overall distribution of the quality characteristic within the lot.

2. *Use the average of sample measurements and assume typical lot variation*. This method estimates lot average as the calculated sample average and assumes a typical lot variation based on historical data of the measured quality characteristic. By assuming a typical lot variation, this method can use the standard normal distribution (a relatively well-understood distribution) to estimate the overall distribution of the quality characteristic within the lot. This estimate is only accurate if the actual variation of the quality characteristic within the lot is close to the assumed variation (Freeman and Grogan, 1998).

3. *Use the average of sample measurements and variation in sample measurements*. This method estimates lot average as the calculated sample average and estimates lot variation as the calculated sample variation. It estimates the overall

distribution of the quality characteristic within the lot by applying the *non-central t distribution* (Johnson and Welch, 1940).

Methods like #3 are typically preferable because they fully describe the distribution of the quality characteristic within a lot and make the fewest assumptions. However, methods such as #1 and #2 are still often used.

"Quality" is then defined as the fraction of the overall quality characteristic distribution that falls within specification limits. It is usually expressed as either (TRB, 1999):

- *Percent defective* (PD) – also called *percent nonconforming*. The percentage of the lot falling outside specification limits.

- *Percent within limits* (PWL)—also called *percent conforming.* The percentage of the lot falling above a lower specification limit, below an upper specification limit, or between upper and lower specification limits. PWL is related to PD by the following: PWL = 100% - PD.

To summarize, the statistical model determines how and to what extent the overall quality characteristic distribution is estimated. Some models are quite simple and only estimate an average quality characteristic value while other models are more complete and estimate both average and variation, which then provides the ability to estimate lot quality. Lot quality, expressed as PD or PWL, is simply the fraction of the lot that falls within specifications.

2.2.5 Quality Level Goals

Quality level goals consist of an acceptable quality limit (AQL) and a rejectable quality limit (RQL). AQL is the minimum level of actual quality at which the material or construction can be considered fully acceptable (TRB, 1999). RQL is the maximum level of actual quality at which a material or construction can be considered unacceptable and thus, rejectable (TRB, 1999).

The appropriate levels of AQL and RQL are matters of judgment. It would be nice but unrealistic to expect all material within a lot to meet specifications (PWL = 100). However, some small fraction of defective material must be permitted due to the unavoidable variability that accompanies any material or production process (Comisky, 1974 as cited in Freeman and Grogan, 1998). To account for this, AQL should be some value less than 100 PWL. Additionally, AQL should also be set at a value equal to the maximum amount of defective material present within the pavement that will not substantially degrade overall road quality (Freeman and Grogan, 1998). These considerations result in typical AQL values of 90 or 95 PWL.

RQL is generally set much lower than AQL because it should represent a PWL below which the HMA pavement is essentially worthless to the contracting agency. Typical values of RQL range from 60 PWL down to 30 PWL and often depend upon sample size. If the actual material quality level is between AQL and RQL then it is often accepted at reduced pay because although defects in the material will degrade overall road performance they will not degrade it to a point that makes the pavement worthless.

AQL and RQL are difficult to accurately set. Typically there is not enough data to accurately relate material quality to final pavement worth. Although current research is addressing this issue (Weed, 1998; Deacon et al, 2001), most AQL and RQL values seem to be set using a combination of historical data, experience, and statistical tradition.

2.2.6 Risk

Using samples to make estimates about the quality of a large amount of construction material involves risk; there is some probability that a random sample will not be representative of the material as a whole, and will thus be an incorrect estimate of material quality. Therefore, risk is an inherent part of statistical acceptance plans. An incorrect estimate, or error, and its associated risk can be either of two types:

- *Type I error (α risk)*. Acceptable construction quality will be rejected as unsatisfactory. This is the contractor's (seller's) risk and can result in unnecessary removal and reconstruction of large pavement sections. There are two types:
    - *Primary type I error (primary α risk)*. The contractor's risk that material produced at AQL will be either rejected or subject to reduced pay.
    - *Secondary type I error (secondary α risk)*. The contractor's risk that material produced at AQL will be rejected.

- *Type II error (β risk)*. Unacceptable construction quality will be accepted as satisfactory. This is the contracting agency's (buyer's) risk and can result in additional maintenance costs, and premature pavement failure. There are two types:
    - *Primary type II error (primary β risk)*. The contracting agency's risk that material produced at RQL will be accepted at bonus pay.
    - *Secondary type II error (secondary β risk)*. The contracting agency's risk that material produced at RQL will be accepted.

These risks can be calculated (see Appendix B) and must be balanced. For a given sample size, reducing the likelihood of accepting poor material usually means increasing the likelihood of rejecting good material and vice versa (Freeman and Grogan, 1998). To simultaneously reduce both of these risks, the sample plan must make more accurate estimates. This usually means increasing the sample size, which means higher inspection and testing costs to the contracting agency. Therefore, the contracting agency will try and achieve an acceptable balance between sample size (accuracy) and inspection and testing costs.

Selecting the appropriate contractor risk and contracting agency risk is a matter of judgment. However, these risks should be related to the criticality of the quality characteristic as well as economic considerations (Freeman and Grogan, 1998). If the failure of a certain material characteristic will render an entire project useless, then it is a critical material characteristic. Therefore, the probability of accepting poor material (β risk) should be set quite small. Conversely, if a material characteristic is not critical, then the probability of accepting poor material (β risk) can be set higher (Freeman and Grogan, 1998). For road construction, the primary α risk is often set near 5 percent and the primary β risk is often set near 10 percent (Cominsky, 1974 as cited in Freeman and Grogan, 1998). As long as these risks are quantified and known in advance, both parties can account for them in their respective budgets and bids.

The risks involved in a particular acceptance plan are often expressed using an operating characteristic (OC) curve. An OC curve describes the relationship between a lot's quality and its probability of acceptance for a given sample size. Each sample size has a different OC curve. Figure 2 shows a WSDOT OC curve for a sample size of five (n = 5). The better the sampling plan is at estimating actual lot quality, the steeper the OC Curve. Figure 3 shows a much steeper OC curve for a sample size of 50. OCPLOT (FHWA, 1994; Weed, 1995) is a readily available software product that can produce these OC curves specifically for HMA pavement acceptance plans.

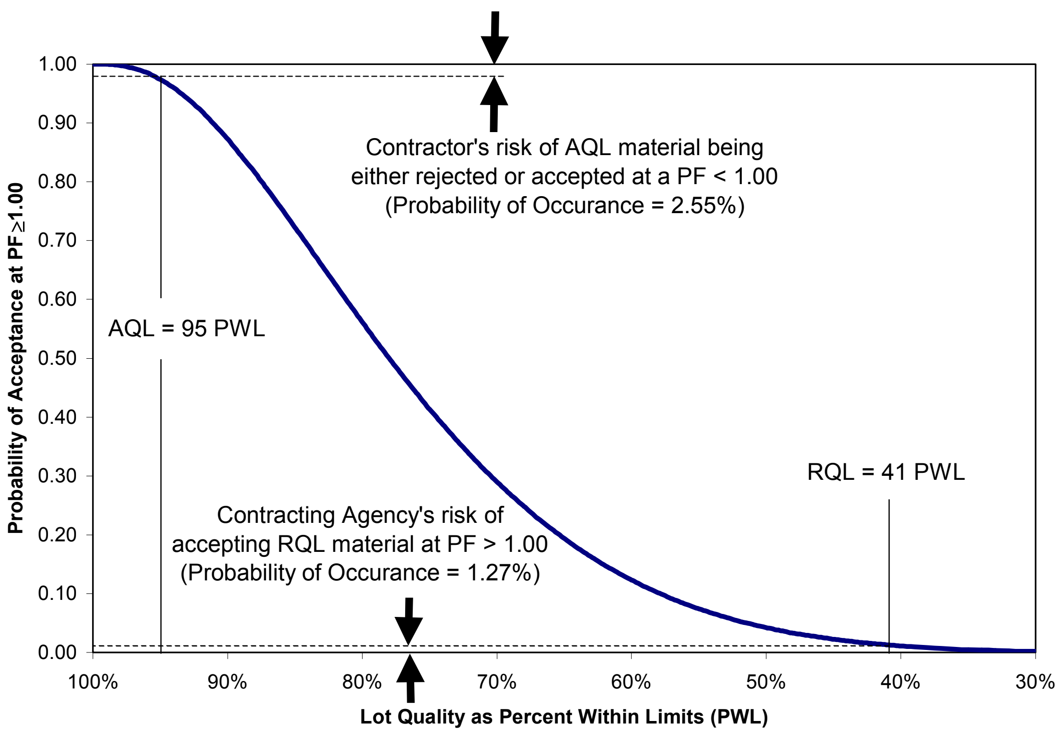


Figure 2: WSDOT Operating Characteristic (OC) Curve for a Sample Size of 5

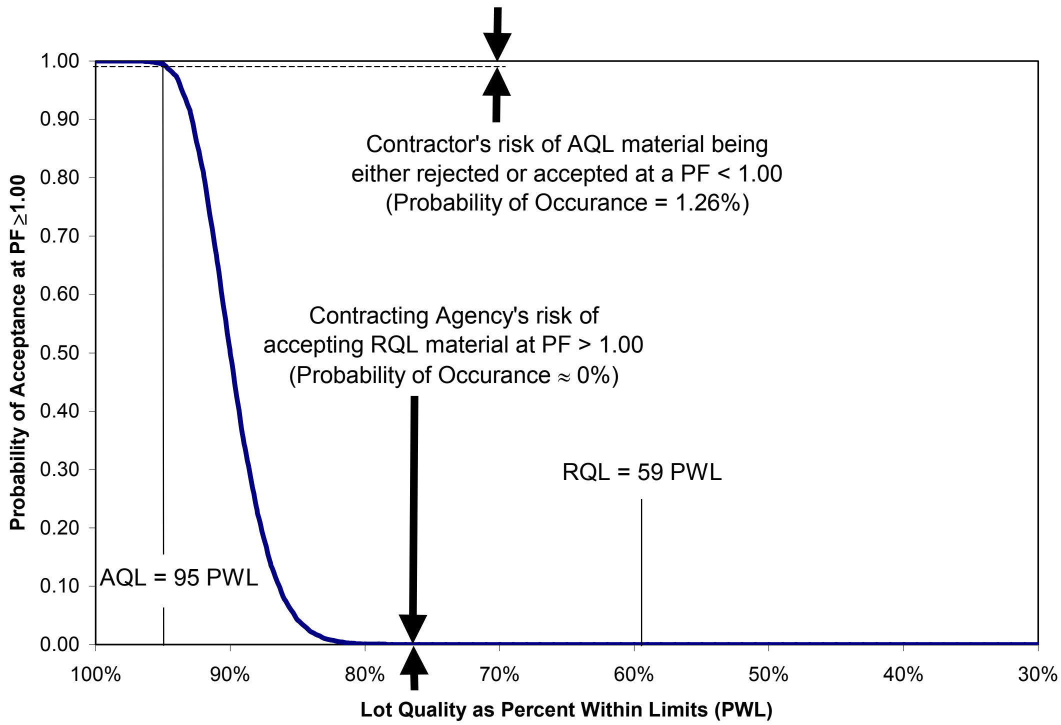


Figure 3: WSDOT Operating Characteristic (OC) Curve for a Sample Size of 50

2.2.7 Pay Factors

Pay factors relate quality to actual pay. In broad terms, a pay factor (PF) is a multiple applied to the contract price of a particular item. Most acceptance plans apply a pay factor to the contract price based on the calculated quality (expressed as PD or PWL) of a particular quality characteristic. Pay factors usually range from a high between 1.00 and 1.12 down to a low between 0.50 and 0.75 (Mahoney and Backus, 2000). Ideally, material produced at AQL receives a pay factor of 1.00, material produced at RQL is rejected, material produced between AQL and RQL receives a pay factor less than 1.00 depending on quality and material produced in excess of AQL receives a pay factor greater than 1.00. Pay factors are not, however, as simple as they seem for two reasons: (1) expected pay is different than contractual pay and (2) material produced at AQL may not receive a 1.00 pay factor.

First, the pay a contractor can expect for consistently producing material at a particular quality level is not necessarily the same as the pay factor shown in the specification for that quality level (referred to as the contractual pay factor). For instance, the WSDOT *Standard Specifications* (2000a) show that for five samples (n = 5), material estimated at AQL (95 PWL) shall receive a 1.04 pay factor. However, a contractor consistently producing AQL material should expect to receive, over time, an average pay factor near

1.03[1]. Figure 4 shows this difference between the specified, or contractual, pay factor and the expected pay factor for the WSDOT specification.

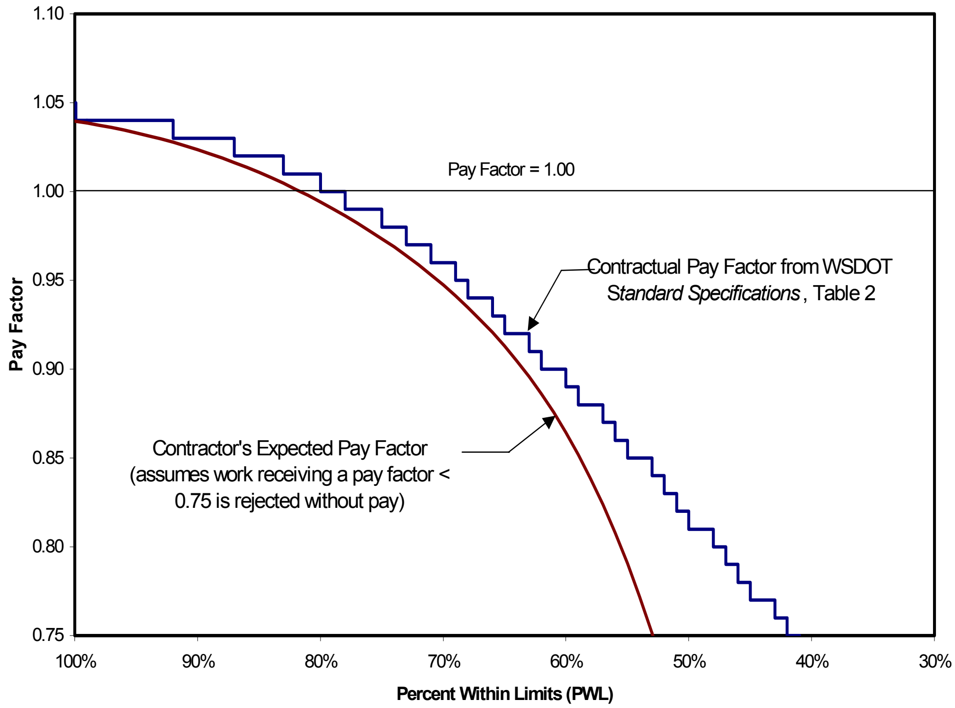


Figure 4: Expected vs. Contractual Pay Factor for WSDOT (n = 5)

This difference occurs because sampling only estimates actual material quality. Therefore, material produced at AQL may be estimated by sampling to be either above or below AQL. Over time, sample estimates of quality will be normally distributed about a mean equal to the actual material quality. Figure 5 shows how this looks for material produced at AQL under WSDOT's acceptance plan using an ideal normal distribution of samples (the large number of lots with estimated quality at 100 PWL occur because 100 PWL is the maximum achievable quality, therefore the entire portion of the normal distribution that falls above 100 PWL is represented by the 100 PWL value). Since each lot receives a contractual pay factor, Figure 6 shows the resulting pay factors associated with Figure 5. Figure 6 shows that material consistently produced at AQL (95 PWL) will not receive the contractual pay factor (1.04) associated with AQL but rather a lesser pay factor (1.0349 in this example). Simulations run by the FAA (FAA, 1999b) and Weed (1995, 1998) have also shown this type of behavior, which is a characteristic of almost all statistical acceptance plans that use pay factors.

[1] For this report expected pay wis calculated assuming a normal distribution of quality (as measured by PWL) about the quality level in question with a standard deviation of about 19 percent PWL. Results using this model compare almost identically with results from the OCPLOT simulation software contained in the FHWA's demonstration Project 89, *Quality Assurance Software for the Personal Computer* (1996b).

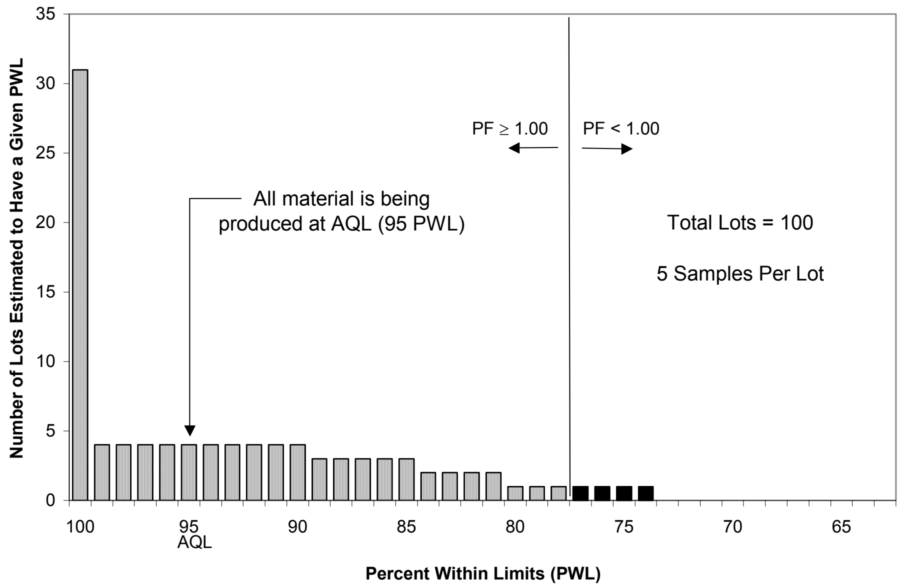


Figure 5: Typical Sample Distribution for Material Produced at AQL for a Hypothetical Project Consisting of 100 Lots

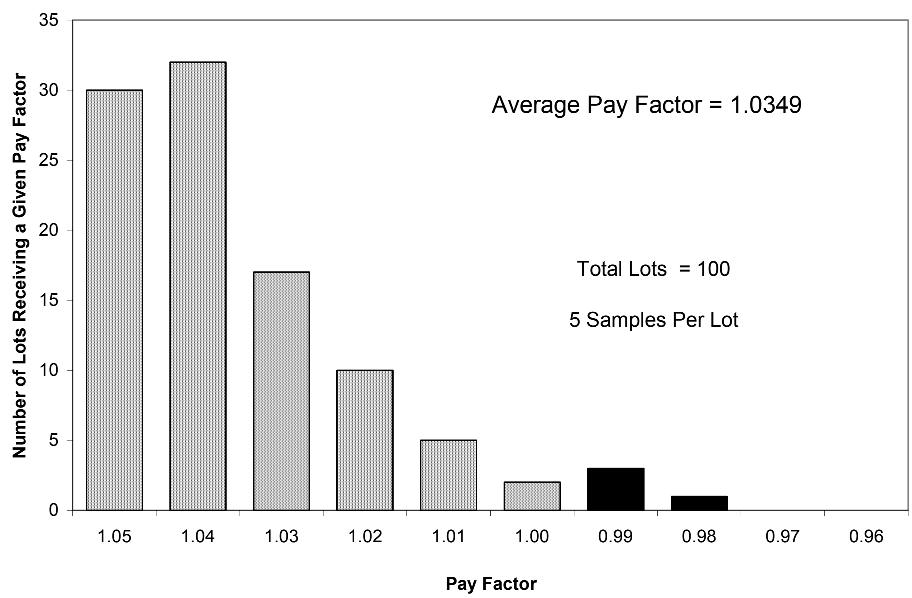


Figure 6: Typical Pay Factor Distribution for Material Produced at AQL for a Hypothetical Project Consisting of 100 Lots

Second, material produced at AQL does not always receive a 1.00 pay factor. In the example shown in Figures 2 and 3, AQL material produced a 1.0349 pay factor. Therefore, material produced at the contractually specified quality is paid at a higher rate than the contractually specified price. Conversely, in acceptance plans that do not include pay factors above 1.00 AQL material could receive a pay factor significantly less than 1.00. In these cases, material produced at the contractually specified quality is paid at a lower rate than the contractually specified price.

Pay factors relate material quality to actual pay. An ideal pay factor system typically allows bonus pay for material produced in excess of AQL, pays the contractual price for AQL material, applies a deduction for material produced between AQL and RQL and rejects material produced at or below RQL. Meeting all four of these goals is quite difficult because expected pay is often different than contractual pay and providing bonus pay for material produced in excess of AQL may lead to expected pay above the contractual price for AQL material.

2.3 Acceptance Plan Categories

Statistical acceptance plans can be categorized according to their specification limits and decision criteria structure. Depending upon the category, different components of the plan will carry different levels of importance. These categories are (Freeman and Grogan, 1998):

1. *Single specification limit, single decision criterion*. Single specification limits are used when a material must be controlled above a minimum or below a maximum. An AQL is set and material is either accepted or rejected based on it. There is no pay factor provision.

2. *Single specification limit, dual decision criteria*. An AQL and RQL are set. Material at or above AQL is accepted at full or bonus pay while material below RQL is rejected. Material with an estimated quality level between AQL and RQL is usually accepted at reduced pay according to a pay scale. WSDOT's asphalt concrete compaction specification uses a single specification limit, dual decision criteria acceptance plan.

3. *Dual specification limits, single decision criterion*. Dual specification limits are used when a material must be controlled within a range of values. The percent of material between these values is calculated as the PWL and compared to the AQL. Material is then either accepted or rejected. There is no pay factor provision.

4. *Dual specification limits, dual decision criteria*. An AQL and RQL are set. Material at or above AQL is accepted at full or bonus pay while material below RQL is rejected. Material with an estimated quality level between AQL and RQL is usually accepted at reduced pay according to a pay scale. WSDOT's asphalt content and aggregate gradation specifications use dual specification limit, dual decision criteria acceptance plans.

2.4 Summary

HMA statistical acceptance specifications use acceptance sampling to audit construction quality. Acceptance sampling is a powerful audit tool because it allows reasonably accurate estimates of lot quality to be made based on test results from a relatively small number of random samples within the lot.

In general, there are seven components of a statistical acceptance specification that define its performance: (1) the sampling type, (2) the quality characteristics, (3) the specification limits, (4) the statistical model, (5) the quality level goals, (6) the risk inherent in the plan, and (7) the pay factor. All of these components interact with one another to determine how the specification pays for HMA construction.

3 QUANTIFICATION AND EVALUATION OF THE WSDOT STATISTICAL ACCEPTANCE SPECIFICATION

WSDOT's statistical acceptance specification has been in use since 1989 and is almost identical to the Federal Highway Administration's (FHWA) 1985 *Standard Specifications for Construction of Roads and Bridges on Federal Highway Projects* (FP-85). Although the FHWA has since revised their specification, WSDOT has not. This section discusses and evaluates the seven basic components of WSDOT's statistical acceptance specification: (1) sampling type, (2) quality characteristics, (3) specification limits, (4) statistical model, (5) quality level goals, (6) risk and (7) pay factors. It also makes comparisons to the current FHWA and FAA statistical acceptance plans where applicable because they are similar plans with national application. Table 1 provides a basic summary of the WSDOT statistical acceptance plan and Table 2 summarizes the comparisons and also includes the 1996 AASHTO *Guide Specification* values. For a comparison of State DOT practices, see Mahoney and Backus (2000) and Schmitt, et al. (1998). All references to the FHWA and FAA in this section are from their respective current specifications (FHWA, 1996a; FAA, 1991, 1999a) unless otherwise stated.

A Note on Comparisons

Comparisons between different statistical acceptance specifications should be made with caution. Specifically, there are two distinct components to each specification: the statistical plan, and the application of this plan to the material in question. Statistical plans can be compared directly to one another. Risk, quality limits, and pay factor comparisons can be made between plans without regard to the specific quality characteristic being measured or how often it is measured. For instance, assessing α and β risks at a specific sample size across several different plans is completely valid.

However, comparing the application of these statistical plans to specific material may not always be valid. Sampling frequency can vary widely between plans. This has a profound effect on the resulting estimate of actual quality level. For instance, WSDOT typically takes five in-place density samples for each 400-ton lot, but the FHWA only takes one in-place density sample every 500 tons. Therefore, on a hypothetical 10,000-ton job, WSDOT will typically use 125 samples while the FHWA will typically use only 20 samples. Depending upon how these samples are divided up into lots and sublots, this may result in higher or lower pay, depending upon actual material quality. Therefore, the sampling frequency must be carefully considered when comparing statistical acceptance specifications.

Table 1: WSDOT Statistical Acceptance Plan Summary[1]

	Sample Size[2]														
	n =3	n = 4	n = 5	n = 6	n = 7	n = 8	n = 9	n = 10-11	n = 12-14	n = 15-18	n = 19-25	n = 26-37	n = 38-69	n = 70-200	n = 201+
Primary α Risk[3]	2.07	2.26	2.55	2.46	2.12	1.95	1.91	1.99	1.75	1.44	1.19	1.65	1.26	0.55	0.83
Secondary α Risk[4]	0.02	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Primary β Risk[5]	5.80	2.89	1.27	0.75	0.48	0.25	0.17	0.08	0.02	0	0	0	0	0	0
Secondary β Risk[6]	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50
AQL[7] (in PWL[8])	95	95	95	95	95	95	95	95	95	95	95	95	95	95	95
c[9]	68	74	78	80	81	82	83	84	85	86	87	89	90	91	93
RQL[10] (in PWL[8])	33	38	41	44	46	47	49	50	51	53	55	57	59	62	65

Notes:

1. Values shown as zero (0) are not truly zero but rather some small fraction, which can be effectively rounded to zero. See Appendix B for Calculation Formulas.
2. All table values calculated for sample size ranges (e.g. n = 15 – 18) are calculated for a sample size corresponding to the smallest sample size within the range (e.g. 15 for the n = 15 – 18 range).
3. Primary α Risk = The contractor's risk that material produced at AQL will be either rejected or subject to reduced pay.
4. Secondary α risk = The contractor's risk that material produced at AQL will be rejected.
5. Primary β risk = WSDOT's risk that material produced at RQL will be accepted at bonus pay.
6. Secondary β risk = WSDOT's risk that material produced at RQL will be accepted.
7. AQL (Acceptable Quality Limit) = The minimum level of actual quality at which the material or construction can be considered fully acceptable.
8. PWL (Percent Within Limits) = The percentage of the lot falling above a lower specification limit, beneath an upper specification limit, or between upper and lower specification limits
9. c = The PWL that receives a PF = 1.00.
10. RQL (Rejectable Quality Limit) = The maximum level of actual quality at which a material or construction can be considered unacceptable and thus, rejectable.

Table 2: Comparison of WSDOT, FHWA, FAA, and AASHTO Guide Statistical Acceptance Plans

Agency	Sample Size Range	A Typical Sample Size	Typical Sublot Size	Typical Lot Size	AQL (PWL)	RQL (PWL)	Expected Pay Factor at AQL[a]	Expected Pay Factor at RQL	Primary α Risk[b]	Primary β Risk[c]
WSDOT	3 - ∞	5	80 tons (compaction) ≤ 800 tons (AC, gradation)	400 tons (compaction) 1 lot per JMF (AC, gradation)	95	41[d]	1.033	0.406[e]	2.55 %	1.27 %
FHWA	5 - ∞	5	500 tons	≥ 2500 tons	95	41[d]	1.007	0.437[e]	2.55 %	1.27 %
FAA[f]	3 - 8	4	≤ 500 tons	≤ 2000 tons	90	55	0.998	0.679[g]	37.30%[h]	1.04 %[h]
AASHTO Guide	3 - 50	i	i	i	90	i	0.992	i	i	i

Notes:

a) Based on a model that assumes an ideal normal distribution of the quality characteristic.
b) Primary α risk corresponds to AQL and the typical sample size in column 3.
c) Primary β risk corresponds to AQL and the typical sample size in column 3.
d) RQL varies with sample size. This RQL corresponds to the typical sample size in column 3.
e) Assumes below RQL work as indicated by sampling is rejected at no pay.
f) The FAA uses a typical sample size of 4; therefore it is not directly comparable to WSDOT and FHWA numbers.
g) Assumes below RQL work is accepted at 50 percent pay.
h) From FAA, 1999b, Table 3.1, p. 15 (see References Section).
i) AASHTO leaves this value to the discretion of the individual agency.

3.1 Sampling Type

Although WSDOT uses both attribute and variable sampling in its specification, those quality characteristics subject to statistical acceptance (in-place density, asphalt content and aggregate gradation) use variable sampling. Attribute sampling is used in simple pass-fail test such as aggregate fractured-face percentage.

3.2 Quality Characteristics

WSDOT uses three main quality characteristics: (1) asphalt content, (2) aggregate gradation, and (3) in-place density. This combination of characteristics has both advantages and disadvantages.

First, these characteristics are few in number and relatively simple to measure. This makes them easy to comprehend and relatively inexpensive to measure. Certainly there is considerable debate over which quality characteristics ought to be measured and when (Chamberlain, 1995; Lin, Solaimanian and Kennedy, 2001), however, this debate concerns engineering judgment and economics rather than statistical acceptance plans, therefore this paper will not address it.

Second, these characteristics are not entirely independent of one another. Specifically, in-place density will change based on a change in asphalt content or gradation. However, these changes are related in such a way that as pay for one quality characteristic goes up, pay for the correlated quality characteristic goes down. For example, excess asphalt content will reduce air voids and increase in-place density, which will cause asphalt content quality (and thus pay factors) to decrease and in-place density quality (and thus pay factors) to increase.

Finally, these correlations are small in magnitude and subordinate to other construction parameters that typically have a much greater effect on in-place density such as compactive effort, and mat temperature during compaction. Therefore, although there is some bias built into WSDOT's specification based on the choice of quality characteristics, this bias is generally small in magnitude, tolerable and represents a good compromise between simplicity and correlation.

3.3 Specification Limits

WSDOT specification limits and bands vary depending upon the class of HMA used. Typical Class A or Class B dense-graded HMA specification limits are (WSDOT, 2000a):

- In-place density ≥ 91.0 percent TMD
- Asphalt content = JMF ± 0.5 percent (< 20 percent recycled asphalt pavement)
- Gradation: Passing the ¼-inch sieve = ± 6 percent
  Passing the No. 10 sieve = ± 5 percent
  Passing the No. 40 sieve = ± 4 percent
  Passing the No. 200 sieve = ± 2.0 percent

This paper does not evaluate the engineering judgment and historical data used to develop these specific specification limits, however (1) they are fairly close to FHWA and FAA specifications and (2) comparing the specification bands to data compiled by Hughes (1996) shows that they are loose enough to account for material, sampling, and testing variation but still tight enough to identify manufacturing and construction variability. Although WSDOT measures other construction characteristics they are not evaluated using statistical acceptance.

3.4 Statistical Model

The WSDOT acceptance specification uses sample measurement average and standard deviation (a measure of variation) to estimate overall lot quality. The FHWA and FAA also use this method. Basically, sample measurement average and standard deviation are used to compute a quality index (Q). This quality index is then compared with a statistical probability distribution used for estimating percent defective called the *non-central t distribution* (Johnson and Welch, 1940) to obtain a quality level (expressed as PWL).

This method of calculating the PWL is the most descriptive of several common practices because by using the *non-central t distribution* it (1) calculates a sample standard deviation and (2) based on the sample standard deviation it actually estimates (rather than assumes) a lot variation. Other less descriptive plans, such as those used by Puerto Rico, Wisconsin, Georgia, and Ohio use sample average only. By not computing sample variation, these plans do not estimate overall lot variation and therefore cannot estimate a PWL. Therefore, they simply pay based on sample average, which does not fully describe the lot.

3.5 Quality Level Goals

As discussed previously, quality level goals consist of an AQL and a RQL. WSDOT uses a set AQL of 95 PWL regardless of sample size while it varies RQL depending on sample size.

By setting AQL at 95 PWL, WSDOT implies that it believes five percent defective material is the maximum amount of defective material that will not substantially degrade overall road quality. However, although an AQL of 95 PWL is fairly typical throughout the industry (most are set at either 90 or 95 PWL), this belief is not based on any documented engineering evidence. Most likely, WSDOT's selection of 95 PWL is based on the typical 95 percent confidence level that statistical disciplines conventionally accept as adequate certainty.

Additionally, there is substantial evidence that, on the average, Washington State contractors are producing material below the established AQL. Table 3, based on values obtained from the WSDOT Materials Laboratory, shows pay factors and their back-calculated theoretical PWL. Based on Table 3, contractors are producing 90 PWL material on average. This implies that, on average, contractors are producing material that does not meet WSDOT specifications for quality. Actual material contains about twice as much (10 percent vice five percent) defective material as WSDOT is willing to accept. However, if WSDOT is generally satisfied with current HMA construction quality (at the 90 PWL level) then the current AQL is too high. If this is the case, then AQL should be set at 90 PWL. For comparison, the FHWA uses an AQL of 95 PWL while the FAA uses 90 PWL. Although the FAA's AQL is more lenient than either WSDOT or the FHWA, their stricter pay factors compensate for this to produce comparable expected pay.

Table 3: WSDOT Pay Factors and Corresponding Theoretical PWL[a,b]

Year	Overall Pay Factor	Overall PWL[c]	Compaction Pay Factor	Compaction PWL[c]	Gradation and Asphalt Content Pay Factor	Gradation and Asphalt Content PWL[c]
2000	1.0235	90	1.0235	90	1.0235	90
1994 - 2000	1.0231	90	1.0208	89	1.0246	91

Notes:

a) This table excludes all contracts receiving pay factors less than 0.7500. This amounts to an exclusion of 3.2 % of the contracts for 2000 and 4.1 % of the contracts from 1994-2000.

b) Data provided by WSDOT Materials Laboratory and Kim Willoughby

c) The corresponding PWL is the nearest whole number PWL corresponding to an overall pay factor based on an ideal expected pay factor model. True PWL values may be slightly different.

Unlike AQL, WSDOT's RQL changes with sample size. It can vary from 33 PWL for small samples (n = 3) to 65 PWL for large samples (n ≥ 201). In general, the larger the sample size, the more accurate its estimate of actual quality, and therefore the higher RQL can be without rejecting adequate lots that are merely sampled inaccurately. For a given sample size, WSDOT RQL is the same as that used by the FHWA but substantially lower than that used by the FAA (see Table 2). For small sample sizes it is also substantially lower than 60 PWL – the level at which the AASHTO *Quality Assurance Guide Specification* (1996) recommends that "the Agency…make a special evaluation of the material and determine the appropriate action". However, WSDOT augments its RQL by requiring the contractor to "…shut down operations and…not resume asphalt concrete placement until such time as the engineer is satisfied that specification material can be produced whenever the Composite Pay Factor (CPF) for a lot in progress…drops below 1.00 and the Contractor is taking no corrective action, or…is less than 0.75." Based on this clause and FHWA RQL values, WSDOT's RQL values are reasonable.

3.6 Risk

Risk is inherent in all statistical acceptance plans and interacts with AQL and RQL. Risk interacts with AQL through an acceptance value (c), which is the lot quality associated with a 1.00 pay factor (Appendix C illustrates how an acceptance value is calculated). Most statistical acceptance plans choose to either establish constant risks and vary acceptance values and RQL with sample size, or establish constant acceptance values and RQL and vary the risks with sample size. Since the WSDOT plan is designed to maintain a relatively constant primary α risk, its acceptance values and RQL vary with sample size.

Risk associated with WSDOT's statistical acceptance specification is summarized in Table 1, however these risks represent the statistical risks associated with WSDOT's *Standard Specifications* (2000a) Table 2. They are not necessarily the true risks involved with pavement construction. As with most specifications, WSDOT's *Standard Specifications* (2000a) contain several contractual clauses that can make the actual risk different from the statistically calculated risk. These clauses, found in Section 5-04.3(8), are:

1. "The Engineer may, without sampling, reject any batch, load, or section of roadway that appears defective in gradation or asphalt cement content."

2. "…the Engineer may also isolate from a normal sublot any material that is suspected of being defective in gradation or asphalt cement content."

3. "If an entire sublot is rejected in accordance with Section 1-06.2, four additional random samples from this sublot will be obtained and the sublot evaluated as an independent lot with the original test result included as a fifth test with the new independent lot instead of with the original lot."

4. "The Contractor shall shut down operations and shall not resume asphalt concrete placement until such time as the engineer is satisfied that specification material can be produced whenever the Composite Pay Factor (CPF) for a lot in progress:

    a. Drops below 1.00 and the Contractor is taking no corrective action, or
    b. Is less than 0.75."

There is no way of statistically quantifying these clauses; therefore they are not accounted for in this risk quantification and analysis. The following two subsections discuss risk as related to (1) the contractor and (2) WSDOT.

3.6.1 Contractor Risk

The WSDOT acceptance plan is designed to maintain a relatively constant primary α risk between 0.55 and 2.55 percent depending on sample size. To create this low primary α risk, the acceptance value is set to a value lower than AQL. For instance, at a sample size of five (n = 5), AQL is 95 PWL (as it always is in the WSDOT plan) while the acceptance value is 78 PWL. This means that for a sample size of five (n = 5), any lot estimated at 78 PWL or better receives at least a 1.00 pay factor. Figure 7 shows the relation between AQL, acceptance value and primary α risk. Figure 8 shows a similar plot for material produced at RQL.

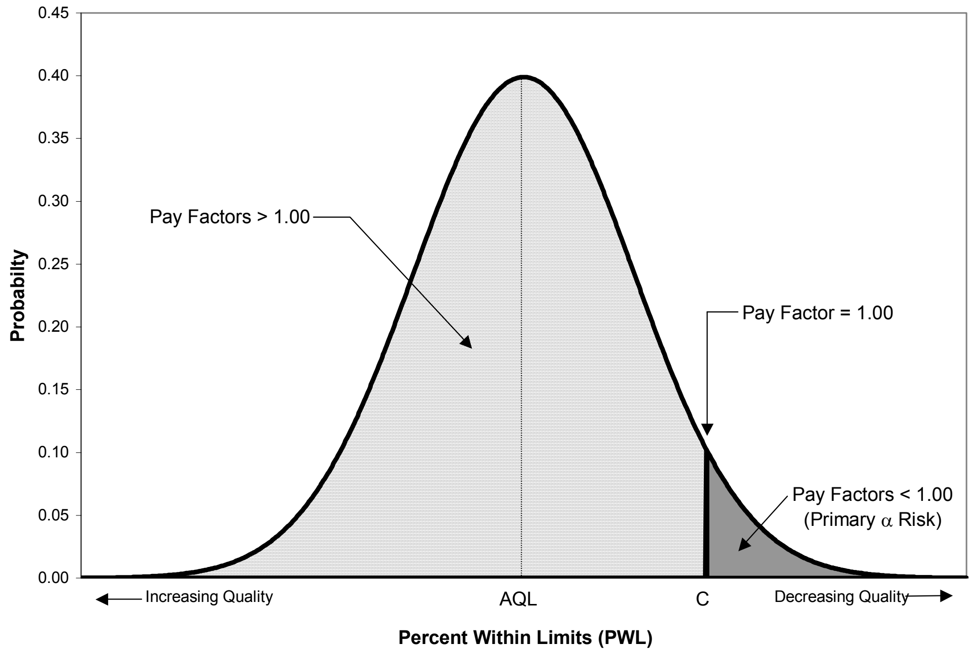


Figure 7: Sample Probability Distribution for Material Produced at AQL

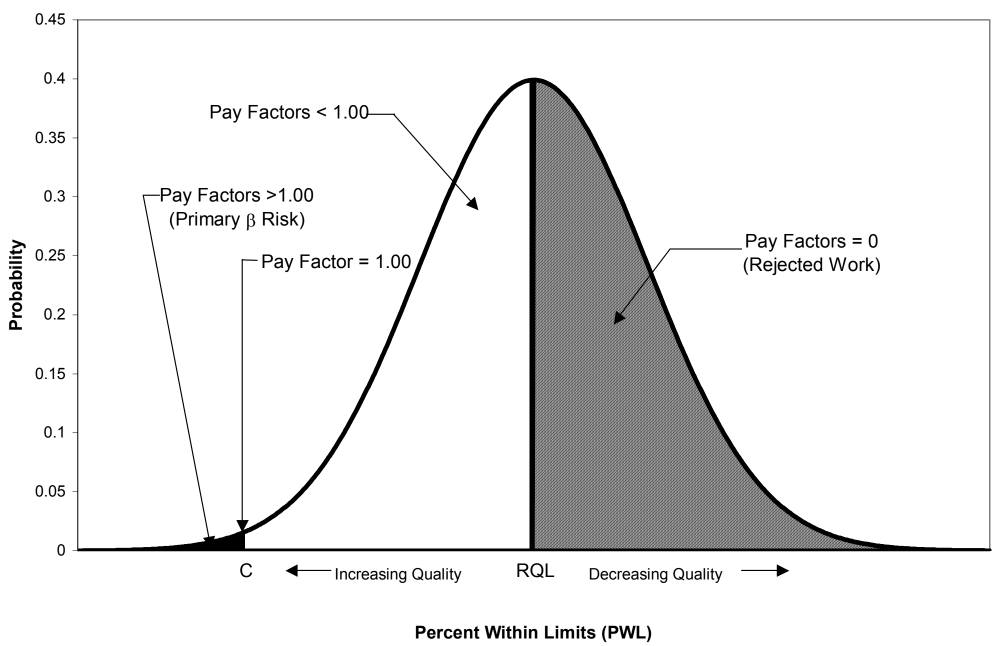


Figure 8: Sample Probability Distribution for Material Produced at RQL

Much has been written on the importance of the primary α risk in acceptance plans. However, this literature is usually in reference to single decision criterion acceptance plans, which do not use pay factors. In a dual decision criteria acceptance plan like WSDOT's (one that uses AQL, RQL, and pay factors), the expected pay as well as the primary α risk should be considered. Single decision criterion acceptance plans are designed to reject an entire lot without pay if samples estimate lot quality to be below a certain minimum. These plans use α risk as the primary method of quantifying how often a contractor will receive zero pay instead of full pay for acceptable material. Therefore, the α risk completely, although indirectly, describes how a single decision criterion acceptance plan pays contractors.

However, in dual decision criteria acceptance plans (like the WSDOT, FHWA, or FAA plans) the primary α risk does not fully describe how contractors are paid; it only indicates the risk that pay will be below 100 percent for acceptable material. The pay factor quantifies this pay reduction and therefore these two statistics, the primary α risk and the pay factor, should be used to fully describe a dual decision criteria acceptance plan.

For comparison, WSDOT's primary α risk is almost identical to the FHWA's but is substantially lower than the FAA's primary α risk of 37.30 percent (for $n = 4$). Although the difference between the WSDOT / FHWA primary α risk and the FAA primary α risk is large, the difference in expected pay is much less due to pay factors (see Table 2).

The secondary α risk for WSDOT, the FHWA and the FAA is essentially zero no matter the sample. This value should be extremely small since it represents the risk that a contractor will receive zero pay for acceptable material.

To summarize, the WSDOT statistical acceptance specification purposefully maintains a small, relatively constant contractor risk, which is typical of dual decision criteria acceptance plans (one that uses AQL, RQL, and pay factors). Although this risk is considered quite important in single decision criterion acceptance plans, in dual decision criteria acceptance plans like WSDOT's expected pay, which directly addresses compensation, is most important to the contractor.

3.6.2 WSDOT Risk

WSDOT's primary β risk varies with sample size and is relatively small. Because of the relatively constant primary α risk and the mathematical interaction between α and β risks, WSDOT assumes more risk with small sample sizes while the contractor risk remains relatively constant (see Table 1). This is a common feature of constant primary α risk plans like WSDOT's. The FHWA, which maintains a relatively constant primary α risk, limits its small sample primary β risk by not using sample sizes less than five. This limits their maximum primary β risk to 1.27 percent. The FAA sets its primary β risk at 1.04 percent for its most common sample size ($n = 4$), which is comparable to both WSDOT and the FHWA.

WSDOT's secondary β risk is constant at 50 percent no matter the sample size. This is consistent with the FHWA and FAA. A secondary β risk of 50 percent is acceptable in dual decision criteria plans for two reasons:

1. The expected pay factor for RQL material is quite small (0.406 at $n = 5$). Therefore, even if RQL material is erroneously accepted, payment will be less than half the contract price.

2. These plans make it difficult to continually produce RQL material. Specifically, WSDOT (2000a) requires the contractor to "…shut down operations and…not resume asphalt concrete placement until such time as the engineer is satisfied that specification material can be produced whenever the Composite Pay Factor (CPF) for a lot in progress…drops below 1.00 and the Contractor is taking no corrective action, or…is less than 0.75."

To summarize, WSDOT's β risks vary with sample size and are generally higher for small sample sizes, which is consistent with the FHWA. The FAA, which uses a set sample size of four ($n = 4$) also has comparable β risks. All three plans use a secondary β risk of 50 percent, which may seem quite high but is acceptable because (1) expected pay for RQL material is quite small and therefore even if RQL material is accepted it is accepted at a greatly reduced price and (2) all three plans make it difficult to produce RQL material without taking corrective action.

3.7 Pay Factors

WSDOT's pay factors are based on a series of parabolic equations; one for each sample size group (Phillips, 1995). There are two main issues with WSDOT pay factors: (1) their basis is undocumented and (2) expected pay for AQL material is greater than 1.00.

First, although WSDOT worked to match its statistical acceptance specification with historical data, the engineering basis, if any, was not documented. Other authors have attempted to establish an engineering basis for pay factors by developing ways to relate pay factors directly to materials, projected life, and projected rehabilitation costs (Weed, 1982; Aurilio and Raymond, 1995; Weed, 1998; Deacon, et al., 2001). However, these methods are typically complex or involve many broad assumptions and therefore do not necessarily offer substantial improvement.

Second, the WSDOT pay factor system will produce expected pay factors greater than 1.00 for AQL material. This is because WSDOT's pay factors are most likely based on a pay factor table contained in the original 1985 FHWA *Standard Specifications*, which when followed resulted in pay factors greater than 1.00 for AQL material. This is especially pronounced for small samples sizes (say, $n \leq 10$) because for small sample sizes the acceptance value is significantly lower than AQL (for $n = 5$, AQL = 95 PWL while $c = 78$ PWL). As previously discussed, the acceptance value is that PWL that receives a pay factor of 1.00. Therefore, material produced at AQL (which is substantially above the acceptance value for small sample sizes) receives pay factors that are also substantially above 1.00. The FHWA believes this leads to "a significant overpayment of money to the contractor" and has since corrected for it by limiting the maximum allowed pay factor for small sample sizes (Wasill, 2001). The FAA achieves an expected pay factor of near 1.00 for AQL material through an alternative method by using a relatively high primary α risk of 37.30 percent.

Market forces in Washington State may have already provided a solution to pay in excess of 1.00 for AQL material. Although contractors receive bonus pay (on average) for hot mix work (Table 3 shows the average pay factor to be near 1.02), anecdotal evidence suggests that contractors who know this may frequently bid slightly lower on hot mix jobs with the assumption that they will receive bonus pay. This method of competitive bidding has tended to eliminate a substantial portion of the overpayment by WSDOT for AQL material.

Figure 9 compares FHWA, FAA, and WSDOT expected pay factors. The FHWA and WSDOT curves are for a sample size of five (n = 5) while the FAA curve is for a sample size of four (n = 4) and is presented for comparison (Figure 9 assumes that work receiving a pay factor less than 0.75 under the WSDOT and FHWA plans will be rejected without pay). Even though all three acceptance plans are shown on the same graph, comparisons can be difficult because of differing lot and sublot sizes (see Table 3). Appendix A applies all three acceptance plans to the same actual WSDOT data to make a case-specific example comparison.

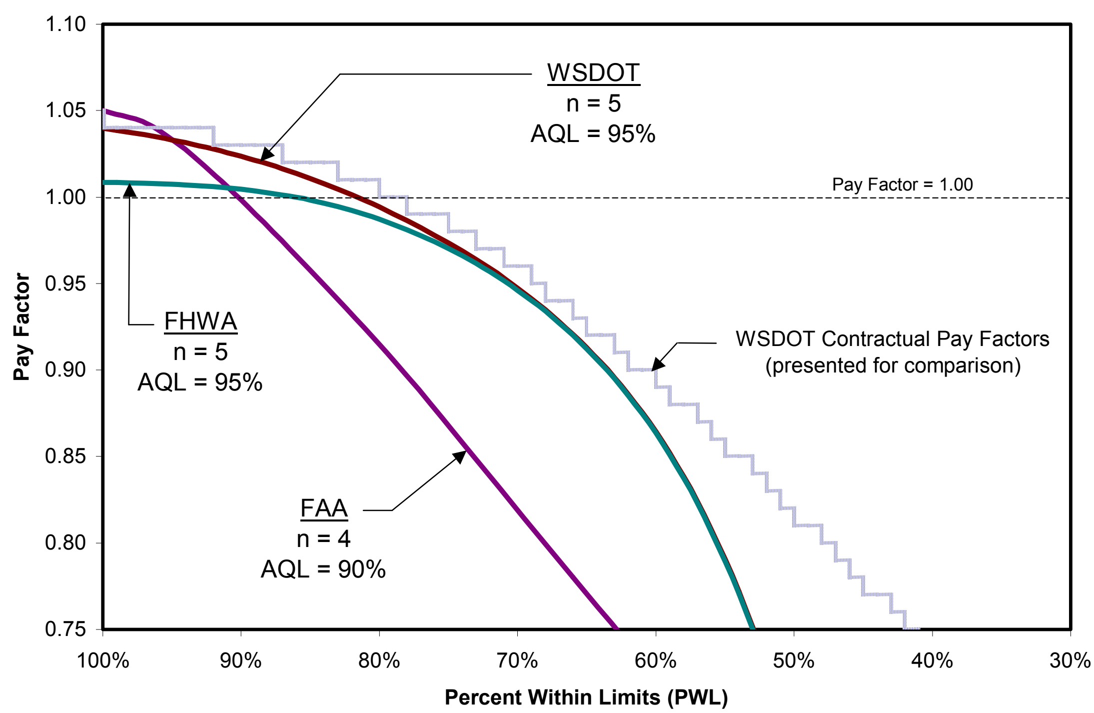


Figure 9: Expected Pay Factors for WSDOT, FHWA, and FAA

Expected pay factors provide the most direct and accurate way of describing how WSDOT's specification pays for HMA construction. Furthermore, contractors can use expected pay along with their own internal quality control data to accomplish any of the following three items:

1. Calculate their expected pay for HMA construction and confidently incorporate this value into their contract bid.

2. Accurately set production goals to achieve a specific pay factor.

3. Determine whether a specific pay factor is realistically achievable given their current production capabilities.

Cumulatively, these items can eliminate a majority of the guesswork involved in HMA construction bidding.

In summary, the two issues with WSDOT pay factors are not as critical as they may seem. First, although the basis for WSDOT pay factors is undocumented and it is therefore difficult to determine why they were chosen, there is currently no consensus best practice for establishing pay factors. Second, although expected pay for AQL is greater than 100 percent, market forces appear to correct for this. Finally, expected pay, which can be calculated from pay factors, is the most direct and accurate way of describing how WSDOT's specification pays for HMA construction.

4 CONCLUSIONS

This report has set out to provide: (1) an increased knowledge of its statistical basis and what this basis implies about its capabilities, (2) a detailed quantification of the specification's components and their implications and (3) an evaluation of the appropriateness and basis for these components. These items are needed to properly use the WSDOT HMA statistical acceptance specification and take advantage of the many improvements it offers over traditional method specifications. Table 4 summarizes the basic conclusions of this report.

The issues raised in the current specification's evaluation are not critical and therefore the specification should not be changed. Since its implementation 12 years ago, WSDOT's statistical acceptance specification seems to have attained an acceptable equilibrium; both WSDOT and contractors seem reasonably content with it. This contentment as well as the familiarity associated with an established specification is a valuable asset that should not be disrupted for non-critical changes. However, if the third and final report in this series, which studies quality characteristics to be measured for Superpave design mixes, results in major recommended changes, then we recommend considering the following:

1. *Establish a new AQL*. Develop a sound engineering basis for AQL. Based on the 2000 average pay factor, contractors are, on the average, not producing what WSDOT has defined as AQL material. If WSDOT believes current average contractor output is acceptable, then the current AQL is set too high. Based on Table 3, a more appropriate value may be near 90 PWL. Further research based on engineering principles and long-term pavement performance could determine the most appropriate AQL. Regardless of its value, any adjustment to AQL will require new pay factor tables in order to maintain current α and β risks.

2. *Examine more appropriate pay factors*. Other researchers are investigating more appropriate pay factors (Deacon, et al., 2001). Although their methods are often complex and involve broad assumptions of the most critical variables they may prove useful. Soundly reasoned, empirically-based pay factors should result in an expected pay factor near 1.00 for AQL work and pay adjustments based on actual increased costs incurred by WSDOT for work accepted at reduced quality levels. Presently, pay factors are not definitively linked to future agency costs.

Finally, an improved understanding of WSDOT's HMA statistical acceptance specification is desirable for everyone involved. Therefore, continued examination, education and discussion regarding the specification are in the best interest of all parties involved.

Table 4: Report Summary

Component	General Background	WSDOT Statistical Acceptance Plan	Evaluation and Comparison
Sampling Type	Two types: attribute and variable	Uses variable sampling for those quality characteristics evaluated by statistical acceptance.	Typical of almost all acceptance plans and consistent with FHWA and FAA.
Quality Characteristics	Should be selected such that: 1. Their quality accurately reflects overall project quality. 2. They are independent of one another.	Uses in-place density, asphalt content and aggregate gradation.	Relatively independent and similar to the FHWA. FAA uses more quality characteristics, some of which are highly correlated.
Specification Limits	Should be tight enough to detect manufacturing and construction variability, but loose enough to allow a reasonable amount of testing, sampling, and inherent material variability	Typical Class A/B specifications are: • In-place density ≥ 91.0 percent TMD • Asphalt content = JMF ± 0.5 % • Gradation: Passing the ¼-inch sieve = ± 6 % Passing the No. 10 sieve = ± 5 % Passing the No. 40 sieve = ± 4 % Passing the No. 200 sieve = ± 2.0 %	WSDOT specification bands are be tight enough to detect manufacturing and construction variability, but loose enough to allow a reasonable amount of testing, sampling, and inherent material variability.
Statistical Model	Several statistical models can be used.	Estimates lot average and variation then uses the *non-central t distribution* to calculate lot quality (expressed as PWL).	Is the most descriptive and makes the fewest assumptions of several common practices. FHWA and FAA use the same model.
Quality Level Goals	AQL and RQL relate the fraction of acceptable material within a lot to whether or not it will be accepted at full pay (AQL) or rejected at zero pay (RQL).	AQL = 95 PWL regardless of sample size c = varies with sample size (from 68 PWL up to 93 PWL) RQL = varies with samples size (from 33 PWL up to 65 PWL)	• AQL is higher than average contractor quality • RQL is low but adequate considering additional specification clauses • FHWA AQL = 95 PWL • FAA AQL =90 PWL
Risk	All statistical acceptance plans involve risk to both the contractor and the contracting agency. This risk can be quantified	Primary α risk = 0.55 to 2.55 % Secondary α risk ≈ 0 Primary β risk = 5.80 to 0 % Secondary β risk = 50 %	• Small α risk, which is typical and similar to FHWA • Secondary β risk of 50 % is typical • Expected pay better describes the plan than risk
Pay Factors	Pay factors relate lot quality to actual pay. Expected pay is different from contractual pay and should be near 1.00 for AQL material.	Uses a set of roughly parabolic equations. Maximum PF = 1.05 Minimum PF = 0.75 Expected pay at AQL ≈ 1.03	• Undocumented basis • PF > 1.00 for AQL material but this is largely corrected by market forces • Expected pay best describes the plan

REFERENCES

American Association of Highway and Transportation Officials (AASHTO). (1996). *Quality Assurance Guide Specification*. AASHTO Highway Subcommittee on Construction. Washington, D.C.

Aurilio, V., and Raymond, C. (1995). Development of End Result Specification for Pavement Compaction. *Transportation Research Record 1491*. pp. 11-17.

Bowery, F.J., and Hudson, S.B. (1976). *National Cooperative Highway Research Program Synthesis of Highway Practice 38: Statistically Oriented End-Result Specifications*. Transportation Research Board, National Research Council. Washington, D.C.

Cadicamo, M.P. (1999). *A Partial Analysis of the Washington State Department of Transportation Quality Assurance Specification*. Master's Thesis. University of Washington. Seattle, WA.

Chamberlain, W.P. (1995). *National Cooperative Highway Research Program Synthesis of Highway Practice 212: Performance-Related Specifications for Highway Construction and Rehabilitation*. Transportation Research Board, National Research Council. Washington, D.C.

Cominsky, R.J. (1974). Session 18: Development of Acceptance Plans. *Statistical Quality Control of Highway Construction*. The Pennsylvania State University. University Park, PA. pp. 18.1-18.37.

Deacon, J.A., et al. (2001). *Pay Factors for Asphalt-Concrete Construction: Effect of Construction Quality on Agency Costs DRAFT*. Caltrans TM-UCB-PRC-2001-1. Sacramento, CA.

FAA. (1991). *AC No:150/537-10A Change 1-3: Standards for Specifying Construction of Airports*. FAA. Washington, D.C. http://www.faa.gov/arp/5370-10a.htm

FAA. (1999a). *AC 150/5370-10A: Item P-401 Plant Mix Bituminous Pavements thru CHG 12*. FAA. Washington, D.C. http://www.faa.gov/arp/5370p401.doc

FAA. (1999b). *Engineering Brief No. 56: Development of Revised Acceptance Criteria for Item P-401 and Item P-501*. FAA. Washington, D.C. http://www.faa.gov/arp/pdf/EB56.PDF

FHWA. (1985). *Standard Specifications for Construction of Roads and Bridges on Federal Highway Projects*. FP-85. FHWA, Washington, D.C.

FHWA. (1996a). *Standard Specifications for Construction of Roads and Bridges on Federal Highway Projects*. FP-96. FHWA, Washington, D.C.

FHWA. (1996b). *Quality Assurance Software for the Personal Computer*. Demonstration Project 89. Publication No. FHWA-SA-96-026. FHWA. Washington, D.C.

Freeman, R.B. and Grogan, W.P. (May 1998) *Statistical Acceptance Plan for Asphalt Pavement Construction*. U.S. Army Corps of Engineers. Washington, D.C.

Halstead, W.J. (1979). *National Cooperative Highway Research Program Synthesis of Highway Practice 65: Quality Assurance*. Transportation Research Board, National Research Council. Washington, D.C.

Hancher, D.E. (1994). *National Cooperative Highway Research Program Synthesis of Highway Practice 195: Use of Warranties in Road Construction*. Transportation Research Board, National Research Council. Washington, D.C.

Hughes, C.S. (1996). *National Cooperative Highway Research Program Synthesis of Highway Practice 232: Variability in Highway Pavement Construction*. Transportation Research Board, National Research Council. Washington, D.C.

Johnson, N.L. and Welch, B.L. (March 1940). Applications of the Non-Central t-Distribution. *Biometrika*, v.31, no. 3/4. pp. 362-389.

Lin, H.H., Solaimanian, M., and Kennedy, T. (2001). General Approach to Payment Adjustments for Flexible Pavements. *Journal of Transportation Engineering* (Jan/Feb 2001). pp. 39-46.

Mahoney, J.P., and Backus, A.W. (2000). *QA Specification Practices*. Washington State Transportation Center (TRAC). Seattle, WA.

Montgomery, D.C. (1997). *Introduction to Statistical Quality Control, 3rd Ed*. John Wiley & Sons, Inc. New York, NY.

Phillips, H.C. (1995). *A Risk and Pay Factor Analysis of Washington State's Department of Transportation 1994 Standard Specification*. Master's Thesis. University of Washington. Seattle, WA.

Rapol, Jeff. (2001). Memorandum from Jeff Rapol to Jack Scott regarding an initial review of this report (Subject: Washington State DOT HMA Statistical Acceptance Specifications) dated 6/26/01. Received by facsimile from Jack Scott on 29 June 2001.

Schmitt, R.L., et al. (1998). Summary of Current Quality Control / Quality Assurance Practices for Hot-Mix Asphalt Construction. *Transportation Research Record 1632*. pp. 22-31.

TRB. (1999). *Glossary of Highway Quality Assurance Terms*. Transportation Research Circular, No. E-C010. Transportation Research Board, National Research Council. Washington, D.C.

WSDOT. (2000a). *Standard Specifications for Road, Bridge, and Municipal Construction*. Publication M 41-10. Washington State Department of Transportation, Olympia, WA.

WSDOT. (2000b). *Materials Manual*. Publication M 46-01. Washington State Department of Transportation, Olympia, WA.

Wasill, B. (2001). E-mail message from Bruce Wasill (FHWA Vancouver, WA office) to Cathy Nicholas (FHWA, Olympia, WA office) dated 18 March 2001. Subject: Flawed PWL Acceptance Plan. FHWA.

Weed, R.M. (1982). *Statistical Specification Development*. New Jersey Department of Transportation, Report No. FHWA/NJ-83/007. FHWA. Washington, D.C.

Weed, R.M. (1995). OCPLOT: PC Program to Generate Operation Characteristic Curves for Statistical Construction Specifications. *Transportation Research Record 1491*. pp. 18-26.

Weed, R.M. (1998). A Rational Method for Relating As-Built Quality to Pavement Performance and Value. *Transportation Research Record 1632*. pp. 32-39.

APPENDIX A: EXAMPLE COMPARISON

This Appendix contains an example comparison of the in-place density portions of the WSDOT, FHWA, and FAA statistical acceptance plans. Data from an actual WSDOT contract is used (contract information is shown in Table A1). This example contains the assumptions and limitations listed below:

- The example only compares in-place density. Other statistically specified material characteristics are not addressed.

- The actual data comes from a specific WSDOT project. Because lot size and sample frequencies are different, randomly selected samples were selected from the WSDOT data to simulate hypothetical FHWA and FAA samples. Specifically:

    - The FHWA samples approximately once every 500 tons. Therefore, one WSDOT sample per 400-ton WSDOT lot was selected at random to represent a simulated FHWA sample. To get the proper number of simulated FHWA samples, 11 WSDOT lots were excluded at random from this process. Therefore, there were 32 simulated FHWA samples for 15,987 tons of HMA placed.

    - The FAA samples approximately four times in a 2000-ton lot. Lots are also broken up by day. Therefore, the WSDOT data was broken up into approximately 2000-ton lots (usually five 400-ton WSDOT lots were used) or one day's production (this restriction was predominant). This resulted in 12 simulated FAA lots with four simulated samples each.

- WSDOT and the FHWA use specification limits based on percent of TMD. The FAA uses a specification limit based on percent compaction as related to a laboratory-compacted specimen (compacted using the Marshall method). In order to make all three density specifications comparable, I assumed laboratory compaction to result in 4.2-percent air voids. This is the upper limit that the FAA specifies in item *P-401 Plant Mix Bituminous Pavements,* Section 401-3.2 (the specified range is 2.8-percent to 4.2-percent). The FAA specification is then translated to percent of TMD by the following formula:

$$Simulated\ \ FAA\ Specification\ \ Limit = \left(100\% - 4.2\%\right) \times (96.3\%) = 92.26\%\ \ of\ TMD$$

- This example comparison represents a single data point and therefore does not represent a trend or imply any mathematical relationship between WSDOT, FHWA, and FAA pay factors. It does, however, show a qualitative relationship between the three specifications.

Table A1: WSDOT Contract Information for Data Used in this Example

Contract Number:	5848
Location:	SR 395, Mile Post 36.1 – 45.36
Description:	East Elm Road to SR 17
Paving Dates:	6/21/2000 – 7/10/2000
Total Tonnage:	15,987 tons
Mix Class:	19 mm Superpave
Binder Type:	PG 70-28
Mix Design:	G9478

Table A2: Example Comparison Summary

	WSDOT	FHWA	FAA
Total Tonnage Placed	15,987	15,987	15,987
Number of Samples per Lot	5	32	4
Total Number of Samples	215	32	48
Specification Limit	91	90	92.26
Average Lot Compaction	92.0377	91.9625	92.2993
Average Lot Standard Deviation of Compaction	0.7753	1.0877	0.8719
Average Quality Level (Q_L)	1.4243	1.8043	-0.3459
Average PD	10.8%	3.3%	57.1%
Average PWL	89.2%	96.7%	42.9%
Overall Pay Factor	**1.0251**	**1.0400**	**0.6012**

Conclusions

1. **WSDOT and FHWA pay factors are similar** ***for the densities in this example***. Data for this example come from a WSDOT-specified contract. Therefore, the contractor's efforts were designed to meet WSDOT specifications. Since FHWA specifications are similar to WSDOT specifications it is expected that the hypothetical FHWA pay factor would be close to the actual WSDOT pay factor.

2. **The FAA pay factors are substantially less than either WSDOT or FHWA pay factors** ***for the densities in this example***. In other words, FAA compaction requirements are typically more stringent than WSDOT or FHWA requirements. Therefore, to receive an FAA pay factor equal to WSDOT or FHWA pay factors a contractor would have to achieve greater compaction. However, while HMA paving by WSDOT or the FHWA is generally intended for vehicular use, HMA paving by the FAA is generally intended for airport pavements, which experience less traffic and have different operating requirements (e.g. up to 60,000 lb. wheel loads, grooving and rubber removal) (Rapol, 2001). Generally the FAA believes its aircraft pavements need 1 to 1.5 percent more initial density than vehicular pavements to meet their functional requirements (Rapol, 2001).

3. **For a given amount of material, WSDOT takes more samples than the FHWA or FAA**. In this example WSDOT took 215 samples compared to a hypothetical 32 for the FHWA and 48 for the FAA. This does not imply that WSDOT specifications provide a better estimate of actual quality. Lot quality estimates become more accurate as the number of samples *per lot* is increased. Table A3 compares results from the WSDOT specification and results from a hypothetical specification that uses only one lot per JMF with all 215 samples coming from that lot. The difference in PWL and pay factor is a direct result of the larger sample standard deviation of the hypothetical specification. The WSDOT standard deviation is smaller because the WSDOT pay factor system, like most others, does not directly account for lot-to-lot variability.

Table A3: Comparison Between WSDOT Specification and a Hypothetical Specification that Uses the Same Number of Samples but Only One Lot Per JMF

	Density	Standard Deviation	Quality Level	PWL	Pay Factor
Results Using WSDOT Specifications (43 Lots, 5 Samples Each)	92.0377[a]	0.7753[b]	1.4243	89 %	1.0251
Hypothetical Results Assuming Only One Lot (215 Samples)	92.0377[c]	1.0467[d]	0.9914	84 %	0.9300

Notes:

a. Calculated as an average of the 43 five-sample averages.
b. Calculated as an average of the 43 five-sample standard deviations.
c. Calculated as an average of all 215 samples.
d. Calculated as the standard deviation of all 215 samples.

WSDOT Actual Data

Summary Data

Total Tonnage Placed =	15987 tons	Average Quality Level (Q_L) =	1.4243
Number of samples per lot =	5 samples/lot	Average PD =	11%
Total Number of Samples =	215 samples	Average PWL =	89%
Specification Limit =	91 % TMD	Weighted Average Pay Factor =	**1.0251**
Average Lot Compaction =	92.0377 % TMD		
Average Lot Standard Deviation of Compaction =	0.7753 % TMD		

Lot Number	Lot_Date	Lot Size (tons)	Measured Density (% Rice) Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Average Density (% Rice)	Standard Dev. (% Rice)	Q_L	x	PD	PWL	PF	weighted PF
1	20-Jun-00	433	96.2	91.2	96.2	96.8	94.8	95.0400	2.2700	1.7797	0.0025	0%	100%	1.05	454.65
2	22-Jun-00	400	91.0	91.2	91.1	91.3	94.9	91.9000	1.6800	0.5357	0.3503	31%	69%	0.96	384.00
3	22-Jun-00	400	92.5	91.2	93.1	91.7	92.0	92.1000	0.7300	1.5068	0.0788	4%	96%	1.04	416.00
4	22-Jun-00	400	91.1	94.2	91.5	91.7	92.8	92.2600	1.2500	1.0080	0.2183	16%	84%	1.02	408.00
5	22-Jun-00	400	92.3	91.8	92.2	91.7	92.7	92.1400	0.4000	2.8500	-0.2966	0%	100%	1.05	420.00
6	22-Jun-00	315	92.2	91.1	93.0	93.2	91.9	92.2800	0.8500	1.5059	0.0791	4%	96%	1.04	327.60
7	23-Jun-00	400	91.3	91.4	92.2	90.9	91.6	91.4800	0.4800	1.0000	0.2205	16%	84%	1.02	408.00
8	23-Jun-00	400	93.1	92.3	92.0	91.0	90.9	91.8600	0.9200	0.9348	0.2387	18%	82%	1.01	404.00
9	23-Jun-00	448	90.8	91.2	91.5	91.3	90.9	91.1400	0.2900	0.4828	0.3651	33%	67%	0.94	421.12
10	23-Jun-00	450	91.1	91.5	90.8	92.6	91.4	91.4800	0.6800	0.7059	0.3027	26%	74%	0.98	441.00
11	26-Jun-00	400	91.5	91.4	91.1	91.0	90.9	91.1800	0.2600	0.6923	0.3065	26%	74%	0.98	392.00
12	26-Jun-00	400	92.8	92.2	93.8	91.7	92.6	92.6200	0.7800	2.0769	-0.0805	0%	100%	1.05	420.00
13	26-Jun-00	443	92.7	93.3	91.8	91.7	91.3	92.1600	0.8200	1.4146	0.1046	6%	94%	1.04	460.72
14	27-Jun-00	400	93.4	91.4	93.8	93.6	93.5	93.1400	0.9800	2.1837	-0.1104	0%	100%	1.05	420.00
15	27-Jun-00	400	92.5	94.9	92.0	91.0	93.1	92.7000	1.4500	1.1724	0.1723	11%	89%	1.03	412.00
16	27-Jun-00	400	92.1	92.3	92.7	92.6	91.1	92.1600	0.6400	1.8125	-0.0066	0%	100%	1.05	420.00
17	27-Jun-00	257	91.1	91.8	91.6	91.7	92.0	91.6400	0.3400	1.8824	-0.0261	0%	100%	1.05	269.85
18	28-Jun-00	400	93.3	93.6	93.1	91.8	92.4	92.8400	0.7300	2.5205	-0.2045	0%	100%	1.05	420.00
19	28-Jun-00	400	91.5	91.0	92.4	91.7	91.8	91.6800	0.5100	1.3333	0.1273	7%	93%	1.04	416.00
20	28-Jun-00	400	92.2	93.3	91.5	90.8	92.7	92.1000	0.9800	1.1224	0.1863	13%	87%	1.03	412.00
21	28-Jun-00	160	92.2	91.9	91.5	92.0	91.9	91.9000	0.2500	3.6000	-0.5062	0%	100%	1.05	168.00
22	29-Jun-00	400	91.6	91.2	91.7	91.9	91.7	91.6200	0.2600	2.3846	-0.1665	0%	100%	1.05	420.00
23	29-Jun-00	400	91.0	91.0	92.1	92.2	91.5	91.5600	0.5800	0.9655	0.2301	17%	83%	1.02	408.00
24	29-Jun-00	400	91.5	91.0	91.6	91.8	91.3	91.4400	0.3000	1.4667	0.0901	4%	96%	1.04	416.00
25	29-Jun-00	400	91.5	91.7	91.0	91.2	91.7	91.4200	0.3100	1.3548	0.1213	7%	93%	1.04	416.00
26	29-Jun-00	413	92.7	91.2	92.3	91.0	91.7	91.7800	0.7200	1.0833	0.1972	14%	86%	1.02	421.26
27	30-Jun-00	400	91.5	92.6	91.2	93.6	91.9	92.1600	0.9600	1.2083	0.1623	11%	89%	1.03	412.00
28	30-Jun-00	400	91.7	95.8	91.6	91.6	91.3	92.4000	1.9100	0.7330	0.2951	25%	75%	0.99	396.00
29	30-Jun-00	400	92.6	91.9	91.8	93.3	93.5	92.6200	0.7800	2.0769	-0.0805	0%	100%	1.05	420.00
30	30-Jun-00	125	92.7	92.4	93.4	91.6	91.1	92.2400	0.9100	1.3626	0.1191	7%	93%	1.04	130.00
31	05-Jul-00	400	92.6	92.6	91.5	92.1	91.4	92.0400	0.5800	1.7931	-0.0012	0%	100%	1.05	420.00
32	05-Jul-00	400	91.8	92.5	93.9	93.0	91.6	92.5600	0.9300	1.6774	0.0311	1%	99%	1.04	416.00
33	05-Jul-00	400	92.3	92.6	91.6	92.3	91.4	92.0400	0.5100	2.0392	-0.0700	0%	100%	1.05	420.00
34	05-Jul-00	400	92.6	91.0	92.8	91.7	91.7	91.9600	0.7400	1.2973	0.1374	8%	92%	1.04	416.00
35	05-Jul-00	400	92.4	91.6	92.9	92.5	91.8	92.2400	0.5300	2.3396	-0.1539	0%	100%	1.05	420.00
36	05-Jul-00	305	92.6	92.6	91.0	93.4	91.5	92.2200	0.9600	1.2708	0.1448	9%	91%	1.03	314.15
37	07-Jul-00	400	92.6	91.1	94.5	91.6	91.4	92.2400	1.3800	0.8985	0.2488	19%	81%	1.01	404.00
38	07-Jul-00	400	91.1	91.7	91.6	92.6	91.9	91.7800	0.5400	1.4444	0.0963	5%	95%	1.04	416.00
39	07-Jul-00	400	91.3	91.1	92.5	91.9	91.0	91.5600	0.6300	0.8889	0.2515	20%	80%	1.01	404.00
40	07-Jul-00	400	91.2	91.0	91.0	91.0	91.2	91.0800	0.1100	0.7273	0.2967	25%	75%	0.99	396.00
41	07-Jul-00	85	91.3	91.9	90.7	91.3	90.9	91.2200	0.4600	0.4783	0.3663	33%	67%	0.94	79.90
42	10-Jul-00	400	91.7	92.0	94.1	91.8	92.9	92.5000	1.0100	1.4851	0.0849	4%	96%	1.04	416.00
43	10-Jul-00	153	92.3	92.0	90.7	90.5	90.2	91.1400	0.9400	0.1489	0.4584	45%	55%	0.86	131.58

FHWA Hypothetical Data

Total Tonnage Placed =	15987 tons	Quality Level (Q_L) =	84.5486
Number of samples per lot =	32 samples/lot	Average PD =	3%
Total Number of Samples =	32 samples	Average PWL =	97%
Specification Limit =	90 % TMD	Pay Factor =	**1.0400**
Lot Compaction =	91.9625 % TMD		
Lot Standard Deviation of Compaction =	1.0877 % TMD		

highlighted area indicates a value selected at random from the WSDOT samples to represent a hypothetical FHWA sample

Lot Number	Lot_Date	Lot Size (tons)	Measured Density (% Rice) Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Simulated FHWA Density (% Rice)
1	20-Jun-00	433	96.2	91.2	96.2	96.8	94.8	96.2
2	22-Jun-00	400	91.0	91.2	91.1	91.3	94.9	91.0
3	22-Jun-00	400	92.5	91.2	93.1	91.7	92.0	91.7
4	22-Jun-00	400	91.1	94.2	91.5	91.7	92.8	91.7
5	22-Jun-00	400	92.3	91.8	92.2	91.7	92.7	
6	22-Jun-00	315	92.2	91.1	93.0	93.2	91.9	93.0
7	23-Jun-00	400	91.3	91.4	92.2	90.9	91.6	92.2
8	23-Jun-00	400	93.1	92.3	92.0	91.0	90.9	
9	23-Jun-00	448	90.8	91.2	91.5	91.3	90.9	91.2
10	23-Jun-00	450	91.1	91.5	90.8	92.6	91.4	
11	26-Jun-00	400	91.5	91.4	91.1	91.0	90.9	90.9
12	26-Jun-00	400	92.8	92.2	93.8	91.7	92.6	92.6
13	26-Jun-00	443	92.7	93.3	91.8	91.7	91.3	91.7
14	27-Jun-00	400	93.4	91.4	93.8	93.6	93.5	
15	27-Jun-00	400	92.5	94.9	92.0	91.0	93.1	92.0
16	27-Jun-00	400	92.1	92.3	92.7	92.6	91.1	91.1
17	27-Jun-00	257	91.1	91.8	91.6	91.7	92.0	
18	28-Jun-00	400	93.3	93.6	93.1	91.8	92.4	93.6
19	28-Jun-00	400	91.5	91.0	92.4	91.7	91.8	91.5
20	28-Jun-00	400	92.2	93.3	91.5	90.8	92.7	
21	28-Jun-00	160	92.2	91.9	91.5	92.0	91.9	
22	29-Jun-00	400	91.6	91.2	91.7	91.9	91.7	91.6
23	29-Jun-00	400	91.0	91.0	92.1	92.2	91.5	91.0
24	29-Jun-00	400	91.5	91.0	91.6	91.8	91.3	91.5
25	29-Jun-00	400	91.5	91.7	91.0	91.2	91.7	91.5
26	29-Jun-00	413	92.7	91.2	92.3	91.0	91.7	92.3
27	30-Jun-00	400	91.5	92.6	91.2	93.6	91.9	91.5
28	30-Jun-00	400	91.7	95.8	91.6	91.6	91.3	91.3
29	30-Jun-00	400	92.6	91.9	91.8	93.3	93.5	91.9
30	30-Jun-00	125	92.7	92.4	93.4	91.6	91.1	
31	05-Jul-00	400	92.6	92.6	91.5	92.1	91.4	92.6
32	05-Jul-00	400	91.8	92.5	93.9	93.0	91.6	
33	05-Jul-00	400	92.3	92.6	91.6	92.3	91.4	91.4
34	05-Jul-00	400	92.6	91.0	92.8	91.7	91.7	92.8
35	05-Jul-00	400	92.4	91.6	92.9	92.5	91.8	92.9
36	05-Jul-00	305	92.6	92.6	91.0	93.4	91.5	93.4
37	07-Jul-00	400	92.6	91.1	94.5	91.6	91.4	91.6
38	07-Jul-00	400	91.1	91.7	91.6	92.6	91.9	
39	07-Jul-00	400	91.3	91.1	92.5	91.9	91.0	92.5
40	07-Jul-00	400	91.2	91.0	91.0	91.0	91.2	91.0
41	07-Jul-00	85	91.3	91.9	90.7	91.3	90.9	90.9
42	10-Jul-00	400	91.7	92.0	94.1	91.8	92.9	
43	10-Jul-00	153	92.3	92.0	90.7	90.5	90.2	90.7

FAA Hypothetical Data

Summary Data

Total Tonnage Placed =	15987 tons	Average Quality Level (Q_L) =	-0.3459
Number of samples per lot =	4 samples/lot	Average PD =	57%
Total Number of Samples =	48 samples	Average PWL =	43%
Specification Limit* =	92.26 % TMD	Weighted Average Pay Factor =	**0.6012**
Average Lot Compaction =	92.2993 % TMD		
Average Lot Standard Deviation of Compaction =	0.8719 % TMD		

***Specification limit is approximated as(100% - 4.2%) x 96.3%**

highlighted area indicates a value selected at random from the WSDOT samples to represent a hypothetical FAA sample

Lot Number	Lot_Date	Lot Size (tons)	Measured Density (% Rice) Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sim. Avg. FAA Density (% Rice)	Sim. FAA Std. Dev. (% Rice)	Q_L	x	PD	PWL	PF	weighted PF
1	20-Jun-00	433	96.2	91.2	96.2	96.8	94.8	94.6000	2.3608	0.9912	0.1696	17%	83%	0.97	417.93
2	22-Jun-00	400	91.0	91.2	91.1	91.3	94.9	93.0250	0.9416	0.8124	0.2292	23%	77%	0.94	1791.30
3	22-Jun-00	400	92.5	91.2	93.1	91.7	92.0								
4	22-Jun-00	400	91.1	94.2	91.5	91.7	92.8								
5	22-Jun-00	400	92.3	91.8	92.2	91.7	92.7								
6	22-Jun-00	315	92.2	91.1	93.0	93.2	91.9								
7	23-Jun-00	400	91.3	91.4	92.2	90.9	91.6	91.3000	0.3096	-3.1011	1.5337	100%	0%	0.50	849.00
8	23-Jun-00	400	93.1	92.3	92.0	91.0	90.9								
9	23-Jun-00	448	90.8	91.2	91.5	91.3	90.9								
10	23-Jun-00	450	91.1	91.5	90.8	92.6	91.4								
11	26-Jun-00	400	91.5	91.4	91.1	91.0	90.9	92.1250	1.0079	-0.1339	0.5446	54%	46%	0.50	621.50
12	26-Jun-00	400	92.8	92.2	93.8	91.7	92.6								
13	26-Jun-00	443	92.7	93.3	91.8	91.7	91.3								
14	27-Jun-00	400	93.4	91.4	93.8	93.6	93.5	92.5500	0.9000	0.3222	0.3926	39%	61%	0.73	1064.15
15	27-Jun-00	400	92.5	94.9	92.0	91.0	93.1								
16	27-Jun-00	400	92.1	92.3	92.7	92.6	91.1								
17	27-Jun-00	257	91.1	91.8	91.6	91.7	92.0								
18	28-Jun-00	400	93.3	93.6	93.1	91.8	92.4	92.0000	0.9967	-0.2609	0.5870	59%	41%	0.50	680.00
19	28-Jun-00	400	91.5	91.0	92.4	91.7	91.8								
20	28-Jun-00	400	92.2	93.3	91.5	90.8	92.7								
21	28-Jun-00	160	92.2	91.9	91.5	92.0	91.9								
22	29-Jun-00	400	91.6	91.2	91.7	91.9	91.7	91.8333	0.2986	-1.4289	0.9763	98%	2%	0.50	1006.50
23	29-Jun-00	400	91.0	91.0	92.1	92.2	91.5								
24	29-Jun-00	400	91.5	91.0	91.6	91.8	91.3								
25	29-Jun-00	400	91.5	91.7	91.0	91.2	91.7								
26	29-Jun-00	413	92.7	91.2	92.3	91.0	91.7								
27	30-Jun-00	400	91.5	92.6	91.2	93.6	91.9	92.2333	0.3697	-0.0721	0.5240	52%	48%	0.50	662.50
28	30-Jun-00	400	91.7	95.8	91.6	91.6	91.3								
29	30-Jun-00	400	92.6	91.9	91.8	93.3	93.5								
30	30-Jun-00	125	92.7	92.4	93.4	91.6	91.1								
31	05-Jul-00	400	92.6	92.6	91.5	92.1	91.4	91.9500	1.1413	-0.2716	0.5905	59%	41%	0.50	600.00
32	05-Jul-00	400	91.8	92.5	93.9	93.0	91.6								
33	05-Jul-00	400	92.3	92.6	91.6	92.3	91.4								
34	05-Jul-00	400	92.6	91.0	92.8	91.7	91.7	92.4500	0.6164	0.3082	0.3973	40%	60%	0.72	799.84
35	05-Jul-00	400	92.4	91.6	92.9	92.5	91.8								
36	05-Jul-00	305	92.6	92.6	91.0	93.4	91.5								
37	07-Jul-00	400	92.6	91.1	94.5	91.6	91.4	91.7750	0.8057	-0.6020	0.7007	70%	30%	0.50	842.50
38	07-Jul-00	400	91.1	91.7	91.6	92.6	91.9								
39	07-Jul-00	400	91.3	91.1	92.5	91.9	91.0								
40	07-Jul-00	400	91.2	91.0	91.0	91.0	91.2								
41	07-Jul-00	85	91.3	91.9	90.7	91.3	90.9								
42	10-Jul-00	400	91.7	92.0	94.1	91.8	92.9	91.7500	0.7141	-0.7141	0.7380	74%	26%	0.50	276.50
43	10-Jul-00	153	92.3	92.0	90.7	90.5	90.2								

APPENDIX B: CALCULATING RISK FOR A GIVEN SPECIFICATION

This appendix is based on a more thorough treatment of the topic in Freeman and Grogan's *Statistical Acceptance Plan for Asphalt Pavement Construction* (1998) Appendix C.

1. Select the sample size for which risk is to be calculated (n). Each sample size typically has a unique combination of α and β risks, acceptance value, and rejection value.

2. Determine AQL, RQL, and the acceptance value (c, the PD that exactly receives a pay factor of 1.00). Determine the rejection value (r, the PD at which material is considered rejectable) if it is different than RQL. In WSDOT's specification the rejection value is the same as RQL.

3. Determine the standard normal values associated with AQL, RQL, c, and r (usually designated z_{AQL}, z_{RQL}, z_c, and z_r).

4. Determine the standard normal values associated with α and β risks.

Primary α risk

$$z(\alpha_c) = \frac{z_{AQL} - z_c}{1/\sqrt{n}}$$

Secondary α risk

$$z(\alpha_r) = \frac{z_{AQL} - z_r}{1/\sqrt{n}}$$

Primary β risk

$$z(\beta_c) = \frac{z_c - z_{RQL}}{1/\sqrt{n}}$$

Secondary β risk

$$z(\beta_r) = \frac{z_r - z_{RQL}}{1/\sqrt{n}}$$

5. Determine the probabilities associated with the standard normal values calculated for the α and β risks.

$1 - P(Z > z(\alpha_c))$ = Primary α Risk $\quad$ $1 - P(Z > z(\alpha_r))$ = Secondary α Risk

$1 - P(Z > z(\beta_c))$ = Primary β Risk $\quad$ $1 - P(Z > z(\beta_r))$ = Secondary β Risk

<u>Example</u>: Calculation of α and β risks of the WSDOT statistical acceptance specification at a sample size of five (n = 5).

1. Select the sample size for which risk is to be calculated.

   Sample Size = 5 (given)

2. Determine AQL, RQL, and the acceptance value (c, the PD that exactly receives a pay factor of 1.00). Determine the rejection value (r, the PD at which material is considered rejectable) if it is different than RQL. *PWL and PD are expressed as a percent*.

   AQL = 95 PWL = 0.95 (WSDOT's specification uses an AQL of 95 PWL)

   RQL = 41 PWL = 0.41 (From Section 1-06.2, Table 2 using the n-5 column at a pay factor of 0.75)

   c = 78 PWL = 0.78 (From Section 1-06.2, Table 2 using the n-5 column at a pay factor of 1.00)

   r = 41 PWL = 0.41 (r = RQL in the WSDOT specification)

3. Determine the standard normal values associated with AQL, RQL, c, and r (z_{AQL}, z_{RQL}, z_c, and z_r). This can be done on Microsoft Excel using the NORMSINV function.

   $z_{AQL} = z_{0.95} = 1.645$

   $z_{RQL} = z_{0.41} = -0.228$

   $z_c = z_{0.78} = 0.772$

   $z_r = z_{0.41} = -0.228$ *(same as z_{RQL} since r = RQL in the WSDOT specification)*

4. Determine the standard normal values associated with α and β risks.

   <u>Primary α risk</u>: $$z(\alpha_c) = \frac{z_{AQL} - z_c}{1/\sqrt{n}} = \frac{1.645 - 0.772}{1/\sqrt{5}} = 1.951$$

   <u>Secondary α risk</u>: $$z(\alpha_r) = \frac{z_{AQL} - z_r}{1/\sqrt{n}} = \frac{1.645 - (-0.228)}{1/\sqrt{5}} = 4.187$$

Primary β risk: $z(\beta_c) = \dfrac{z_c - z_{RQL}}{1/\sqrt{n}} = \dfrac{0.772 - (-0.228)}{1/\sqrt{5}} = 2.235$

Secondary β risk: $z(\beta_r) = \dfrac{z_r - z_{RQL}}{1/\sqrt{n}} = \dfrac{-0.228 - (-0.228)}{1/\sqrt{5}} = 0$

5. Determine the probabilities associated with the standard normal values calculated for the α and β risks. This can be done on Microsoft Excel using the NORMSDIST function.

   Primary α risk: $1 - P(Z > z(\alpha_c)) = 1 - 0.9745 = 0.0255$

   Secondary α risk: $1 - P(Z > z(\alpha_r)) = 1 - 0.99999 = 0.00001$

   Primary β risk: $1 - P(Z > z(\beta_c)) = 1 - 0.9873 = 0.0127$

   Secondary β risk: $1 - P(Z > z(\beta_r)) = 1 - 0.5000 = 0.5000$

APPENDIX C: CALCULATION OF THE ACCEPTANCE VALUE (C)

This appendix provides a step-by-step method for calculating the acceptance value (c) for an acceptance plan with a set primary α risk. This appendix is based on a more thorough treatment of the topic in Freeman and Grogan's *Statistical Acceptance Plan for Asphalt Pavement Construction* (1998) Appendix C.

1. Determine the acceptable quality limit (AQL) in percent defective (PD).

$$PD = 100 - PWL$$

2. Set the primary α risk (the contractor's risk that material produced at AQL will be either rejected or subject to reduced pay).

For WSDOT's statistical acceptance plan, $\alpha = 0.05$

3. Determine the sample size to be used (n).

4. Determine the z-statistic associated with the primary α risk, $z(\alpha_c)$. This is just the cumulative normal probability value associated with the primary α risk and can be obtained with Microsoft Excel (NORMSDIST function) or standard statistical tables.

5. Use the basic equation below to solve for z_c.

basic equation	rearranged to solve for z_c
$z(\alpha_c) = \dfrac{z_{AQL} - z_c}{1/\sqrt{n}}$	$z_c = z_{AQL} - \dfrac{z(\alpha_c)}{\sqrt{n}}$

where: $z(\alpha_c)$ = z-statistic associated with the primary α risk
z_{AQL} = z-statistic associated with AQL
z_c = z-statistic associated with the acceptance value (c)
n = sample size

6. Determine the acceptance value (c) from z_c. This can be done with Microsoft Excel (NORMSINV function) or standard statistical tables.