Advances and Applications in Statistics

http://www.pphmj.com
http://dx.doi.org/10.17654/AS051060427
Volume 51, Number 6, 2017, Pages 427-444 **ISSN: 0972-3617**

A STATISTICAL STUDY ON THE VARIABILITY OF AGGREGATE CHARACTERISTICS OF ASPHALT MIXTURES

Saeed Maghsoodloo, Randy West and Dilcu Barnes

ISE Department
Auburn University
Shelby Center, AL 36849-5346, U. S. A.
e-mail: maghsood@eng.auburn.edu

National Center for Asphalt Technology
277 Technology Parkway
Auburn, AL 36830, U. S. A.
e-mail: westran@auburn.edu

College of Business
The University of Alabama in Huntsville
301 Sparkman Drive
Huntsville, AL 35899, U. S. A.
e-mail: dh0066@uah.edu

Abstract

In the design of asphalt paving mixtures, aggregates from the stockpiles are blended at selected proportions to satisfy specified gradation ranges. In a laboratory environment, the mix designer controls the proportion from each stockpile so that the proportions (or weights) can be assumed fixed. This is unlike in a plant operation where aggregates are fed from bins into a mixing-drum so that the

Received: September 14, 2017; Revised: November 8, 2017; Accepted: December 1, 2017
2010 Mathematics Subject Classification: 62P30.
Keywords and phrases: statistical properties, asphalt mixtures, expected value, variance.

proportion of each aggregate, from each bin is a random variable. Statistical analyses of the former case are well-known and are first repeated in Sections 2 and 3. The objective of this article is to provide statistical information for the case when aggregate proportions are treated as random variables. The formulas in Section 4 provide the reader with a method of calculating approximate first two moments of asphalt paving mixture characteristic treated as random variables witnessed in plant operations.

1. Introduction

Historically, asphalt is a derivative of petroleum and was first used more than 5000 years ago [6]. Typically, asphalt is produced from a petroleum residuum. A residuum or resid (pl. residua, resids) is the non-distillable fraction of petroleum [12].

The aggregate is the hard-inert material, such as sand, gravel, crushed stone, slag or rock dust that are mixed with the asphalt (binder) for the construction of roadways. However, the choice of aggregate is not an easy “pick-any material” choice and is far from being a simple procedure. The aggregate must be selected according to the properties of the asphalt binder as well as the conditions that will exist when the roadway is completed [6]. Aggregate gradation plays an important role in the behaviors of asphalt mixtures [4].

Many works are published regarding the aggregate gradation [14, 9], effect of aggregate properties, size and type on asphalt mixtures [2, 11, 7] and measurement of the variability of material properties of asphalt. Valle and Thom [10] present the results of a review on variability of key pavement design input variables and assess effects on pavement performance. They address the statistical characterization of layer thickness variation, asphalt stiffness and subgrade stiffness.

There are various properties of asphalt mixes that can be considered. However, this paper analyzes the variability related with aggregate proportions of asphalt mixtures specifically focusing on analyzing aggregate proportions as random variables.

In the design of asphalt paving mixtures, aggregates from n stockpiles (generally $n = 2, 3, 4, \ldots, 10$) are blended at selected proportions to satisfy specified gradation ranges. In a laboratory environment, the proportion w_i $(i = 2, 3, \ldots, 10)$ from each stockpile is controlled by the mix designer so that the n proportions (or weights) can be assumed fixed. This is unlike a plant operation where aggregates are fed from n bins into a mixing-drum so that the proportion of each aggregate, W_i, from each bin is a random variable.

Let $X_1, X_2, \ldots, X_n$ be aggregate characteristics from n stockpiles. Suppose $X_1, X_2, \ldots, X_n$ are random variables with known process means $\mu_1, \mu_2, \ldots, \mu_n$, known process variances $\sigma_{11}, \sigma_{22}, \ldots, \sigma_{nn}$, respectively, and known covariances $COV(X_i, X_j) = \sigma_{ij}(i \neq j) = \rho_{ij}\sigma_i\sigma_j$, where ρ_{ij} is the correlation coefficient between the inputs X_i and X_j. In the field of statistics, μ_i's are also referred to as the population first origin moments, and σ_{ii}'s are called the *population second central moments*. Let the characteristic of a mixture, such as %-passing through a sieve, having n aggregates be denoted by $Y_n = \sum_{i=1}^{n} W_i X_i$, where the proportions W_i's are random variables with also known first two moments. This paper obtains the first two moments of the output Y_n under all different scenarios based on the nature of W_i's and their relationships to X_i's. The developments are presented in the order of the simplest to the most complicated, where W_i's and X_i's are correlated variates and pair-wise correlated together.

2. W_i's = w_i's are Known Constants

In this case, the output Y_n reduces to $\sum_{i=1}^{n} w_i X_i$ and is referred to as a linear combination (LC). Thus, we have complete information about the first two moments of the n inputs X_i's, and the objective is to use them to compute the first two moments of the linear output Y_n. Such LCs occur frequently in industrial applications and in the field of statistics (the simplest of all examples is the case of sample mean $Y_n = \bar{\mathbf{x}}$, which is a LC with each $w_i = 1/n$),

$$\mu(Y_n) = E(Y_n) = \sum_{i=1}^{n} w_i \mu_i, \tag{1a}$$

where E represents the linear expected-value operator throughout this paper. Equation (1a) shows that the mean of the mixture $E(Y_n)$ is the same LC of μ_i's as Y_n is of X_i's. The variance $V(Y_n) = \sigma^2(Y_n)$, whose expression is also given in numerous sources, can be computed by applying the nonlinear variance-operator V and is provided below,

$$V(Y_n) = \sigma^2(Y_n) = \sum_{i=1}^{n} w_i^2 \sigma_i^2 + 2\sum_{i=1}^{n-1}\sum_{j>i}^{n} w_i w_j \sigma_{ij}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_i \sigma_j \rho_{ij}, \tag{1b}$$

where $\sigma_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)]$, $\rho_{ij} = \sigma_{ij}/(\sigma_i \sigma_j)$, and variances $V(X_i) = \sigma_{ii} = \sigma_i^2$, $i = 1, 2, ..., n$ are all known parameters. If $X_1, X_2, ..., X_n$ are stochastically independent, then σ_{ij} in equation (1b) for all $i \neq j$ is identically zero, and as a result the $V(Y_n)$ reduces to $\sum_{i=1}^{n} w_i^2 \sigma_{ii} = \sum_{i=1}^{n} w_i^2 \sigma_i^2$. For example, if Y_n is the mean of a random sample from an infinite population, then the previous formula yields the very well-known expression for the variance of the mean as $V(\bar{x}) = \sigma^2/n$ [20-22], where σ^2 is the variance of individuals in the target population. Further, if X_i's are also normally distributed (besides being jointly independent), then Y_n is also normally distributed and expressed as $N\left(\sum_{i=1}^{n} w_i \mu_i, \sum_{i=1}^{n} w_i^2 \sigma_i^2\right)$. However, if X_i's are correlated (i.e., $\sigma_{ij} \neq 0$ for $i \neq j$) and are also normally distributed, then from statistical theory the linear combination $Y_n = \sum_{i=1}^{n} w_i X_i$ is still normally distributed (or Laplace-Gaussian) with $E(Y_n) = \sum_{i=1}^{n} w_i \mu_i$ and $V(Y_n)$ $= \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{ij}$.

The simple LC, $Y_n = \sum_{i=1}^{n} w_i X_i$, occurs frequently in many industrial applications, and specifically, in the design and production of asphalt paving mixtures, where a laboratory combines aggregates from *n* stockpile samples in such a manner that the blend of aggregates meets certain design specifications. The aggregate characteristics, X_i's, from each stockpile have known mean and variance based on quality control tests on samples obtained from the stockpiles. An example is provided below.

Example 1. Suppose an asphalt mixture (or a job-mix formula) is designed to contain aggregates from $n = 3$ stockpiles, having means $E(X_1)$ $= \mu_1 = 35.1$, $\mu_2 = E(X_2) = 46.4$, and $\mu_3 = 62.8\%$ passing through the 2.36mm-sieve, and variances $V(X_1) = \sigma_{11} = 8.60$, $V(X_2) = \sigma_{22} = 16.40$, and $V(X_3) = \sigma_{33} = 12.96\%^2$. The percentages passing certain sieve sizes are key characteristics used in the design and control of asphalt paving mixtures. For the design of asphalt paving mixtures in a laboratory environment, the proportion w_i $(i = 1, 2, 3)$ from each of the 3 stockpiles can be controlled precisely by the mix designer. Thus, for a Laboratory Job Mix Formula (JMF), unlike a plant operation, we can assume that the proportions from each stockpile can be controlled and are not random variables. If $w_1 = 0.40$,

$w_2 = 0.35$, and $w_3 = 0.25$, the characteristic of interest for the combined blend is obtained from the LC: $Y_3 = 0.40X_1 + 0.35X_2 + 0.25X_3$. Then, for the combined % passing the 2.36mm-sieve, equations (1a) and (1b) yield the mean $E(Y_3) = 0.40 \times 35.1 + 0.35 \times 46.4 + 0.25 \times 62.8 = 45.98\%$, and variance $V(Y_3) = 0.40^2 \times 8.60 + 0.35^2 \times 16.40 + 0.25^2 \times 12.96 = 4.195 \rightarrow \sigma(Y_3) =$ 2.05%. The coefficient of variation (or variation coefficient) of Y_3 is given by $CV(Y_3) = 2.05/45.98 = 4.45\%$. Thus, assuming the X_i's are normally distributed, then the overall % passing the 2.36mm-sieve has a sampling distribution which is normal with process mean 45.98% and process variance $4.195\%^2$, designated as $N(45.98, 4.195)$.

It should be highlighted that the output, Y_n, is not the same as in the classical mixture experiments, where X_i's themselves are proportion of n ingredients that constitute a mixture so that $\sum_{i=1}^{n} X_i$ is constrained to equal 1 (or 100%). The mere objective in statistical mixture designs is to identify $(n-1)$ of the X_i's in such a manner that some characteristic of the final mixture is optimized. Cornell [3] provides an example of a mixture experiment where $n = 3$ ingredients, $X_1 =$ proportion of polyethylene, $X_2 =$ proportion of polystyrene, and $X_3 =$ proportion of propylene, were blended to form fiber that would be spun into yarn for draperies. The objective of this mixture experiment was to determine the approximate values of X_1, X_2, (and by necessity X_3) such that the resulting yarn elongation, measured in kilograms of force applied, was maximized, bearing in mind that the constraint $X_1 + X_2 + X_3 \equiv 1$ among the 3 ingredients must be satisfied. While in this paper, as illustrated in Example 1 above, the output Y_n represents the overall characteristic of a mixture, comprised of aggregates with differing known proportions from n stockpiles with constraint $\sum_{i=1}^{n} w_i \equiv 1$.

3. W_i's Are Correlated Random Variables but Independent of Correlated X_i's

Before we formulate results for the sum of products of n random variables given by $Y_n = \sum_{i=1}^{n} W_i X_i$, we first obtain the mean and variance of the product of two independent random variables, which has been known in statistical literature for well over 60 years (e.g., see [18, p. 99]) but repeated below for completeness.

Let W_1 and X_1 be two independent random variables with known process means ξ_1, μ_1, and known process variances $\omega_{11} = \omega_1^2$ and $\sigma_{11} = \sigma_1^2$, respectively. Let $Y_1 = W_1 X_1$; our objective is to obtain the expected-value (or mean) and the variance of the random variable Y_1, assuming W_1 and X_1 are independent,

$$E(Y_1) = E(W_1 X_1) = E(W_1)E(X_1) = \xi_1 \mu_1, \tag{2a}$$

$$\begin{aligned}
V(Y_1) &= E[(W_1 X_1 - \xi_1 \mu_1)^2] = E[(W_1 X_1 - W_1 \mu_1 + W_1 \mu_1 - \xi_1 \mu_1)^2] \\
&= E[W_1(X_1 - \mu_1) + \mu_1(W_1 - \xi_1)]^2 \\
&= E[W_1^2(X_1 - \mu_1)^2] + \mu_1^2 E(W_1 - \xi_1)^2 + 2\mu_1 E[W_1(X_1 - \mu_1)(W_1 - \xi_1)] \\
&= E(W_1^2)\sigma_1^2 + \mu_1^2 \omega_1^2 + 2\mu_1 E(X_1 - \mu_1) E[W_1(W_1 - \xi_1)] \\
&= (\omega_1^2 + \xi_1^2)\sigma_1^2 + \mu_1^2 \omega_1^2 + 0 = \omega_1^2 \sigma_1^2 + \xi_1^2 \sigma_1^2 + \mu_1^2 \omega_1^2.
\end{aligned} \tag{2b}$$

In the above developments leading to equations (2a) and (2b), we have used the well-known fact that the expected-value of a product of two independent random variables is equal to the product of their expectations, i.e., $E(W_1 X_1) = E(W_1)E(X_1)$ [1, pp. 213-220], [8, pp. 259-260]. However, the converse of this last statement is not necessarily true; in other words, the equality $E(W_1 X_1) = E(W_1)E(X_1)$ may hold, but still the two random variables may not be stochastically independent. In the case of a bivariate normal random vector $[W_1 \quad X_1]'$, the equality $E(W_1 X_1) = E(W_1)E(X_1)$ does guarantee that the random components W_1 and X_1 are independent [5, pp. 181-183], [13, p. 315]. The probability density function (pdf) of $Y_1 = W_1 X_1$, when the two variates are independent and normal, is provided by Springer [19] and has been studied by other authors such as [15, 17].

Now consider the general nonlinear combination, NLC, $Y_n = \sum_{i=1}^{n} W_i X_i$, where W_i's have known covariances ω_{ij} $(i \neq j)$ but are stochastically independent of all X_i's, and X_i's have also known covariances σ_{ij} $(i \neq j)$. Further, the 1st two moments are also known (or can be estimated accurately) and given by $E(W_i) = \xi_i$, $E(X_i) = \mu_i$, $V(W_i) = \omega_{ii} = \omega_i^2$ and $V(X_i) = \sigma_{ii} = \sigma_i^2$, $i = 1, 2, ..., n.$ As before, our objective is to obtain the first two moments of the output Y_n using the known first origin moments of W_i's, X_i's, and known covariance structures ω_{ij} and σ_{ij}.

The first origin moment (or the mean) of Y_n is easily obtained by applying the linear expected-value operator to Y_n,

$$E(Y_n) = E\left(\sum_{i=1}^{n} W_i X_i\right) = \sum_{i=1}^{n} E(W_i X_i) = \sum_{i=1}^{n} E(W_i)E(X_i) = \sum_{i=1}^{n} (\xi_i \times \mu_i). \quad (3)$$

The second central moment of Y_n can be obtained by applying the variance-operator V to Y_n,

$$V(Y_n) = V\left(\sum_{i=1}^{n} W_i X_i\right) = \sum_{i=1}^{n} V(W_i X_i) + 2\sum_{i=1}^{n-1}\sum_{j>i}^{n} COV(W_i X_i,\, W_j X_j). \quad (4)$$

However, by definition

$$\begin{aligned} COV(W_i X_i,\, W_j X_j) &= E(W_i X_i,\, W_j X_j) - \xi_i \mu_i \xi_j \mu_j \\ &= E(W_i W_j) \times E(X_i X_j) - \xi_i \xi_j \mu_i \mu_j \end{aligned}$$

$$= (\omega_{ij} + \xi_i\xi_j) \times (\sigma_{ij} + \mu_i\mu_j) - \xi_i\xi_j\mu_i\mu_j$$

$$= \omega_{ij}\sigma_{ij} + \mu_i\mu_j\omega_{ij} + \xi_i\xi_j\sigma_{ij}. \quad (5)$$

Substituting equation (5) into (4) using results of (2b),

$$V(Y_n) = \sum_{i=1}^{n}(\omega_{ii}\sigma_{ii} + \xi_i^2\sigma_{ii} + \mu_i^2\omega_{ii}) + 2\sum_{i=1}^{n-1}\sum_{j>i}^{n}(\omega_{ij}\sigma_{ij} + \mu_i\mu_j\omega_{ij} + \xi_i\xi_j\sigma_{ij}). \quad (6)$$

One special case of equation (6), that occurs frequently in the production control of asphalt paving mixtures is when aggregate characteristics from *n* stockpiles (i.e., X_i's) are independent (that is $\sigma_{ij} = 0$ for all $i \neq j$), but the feed rates from each stockpile into a mixing-drum are constrained such that the variable proportions from the *n* stockpiles that enter the mixing-drum add to 1, i.e., the constraint $\sum_{i=1}^{n} W_i \equiv 1$ must be satisfied. Thus, in this special case, equation (3) still holds true except for the fact that $\xi_n = E(W_n) = 1 - \xi_1 - \xi_2 - \cdots - \xi_{n-1}$, while equation (6) reduces to

$$V(Y_n) = \sum_{i=1}^{n}(\omega_{ii}\sigma_{ii} + \xi_i^2\sigma_{ii} + \mu_i^2\omega_{ii}) + 2\sum_{i=1}^{n-1}\sum_{j>i}^{n}\mu_i\mu_j\omega_{ij}. \quad (7)$$

Due to the constraint $\sum_{i=1}^{n} W_i \equiv 1$, equation (7) further reduces to a special form, which is proven below by starting with $Y_2 = W_1X_1 + W_2X_2$, with $W_1 + W_2 \equiv 1$.

For $n = 2$ stockpiles, $\omega_{22} = V(W_2) = V(1 - W_1) = V(W_1) = \omega_{11}$, and

$$\omega_{12} = COV(W_1, W_2) = E[(W_1 - \xi_1) \times (W_2 - \xi_2)]$$

$$= E[(W_1 - \xi_1) \times (1 - W_1 - 1 + \xi_1)]$$

$$= -E[(W_1 - \xi_1) \times (W_1 - \xi_1)] = -\omega_{11}.$$

Thus, when $n = 2$, equation (7) reduces to

$$V(Y_2) = \sum_{i=1}^{2} (\omega_{ii}\sigma_{ii} + \xi_i^2\sigma_{ii} + \mu_i^2\omega_{ii}) + 2\mu_1\mu_2\omega_{12}$$

$$= \sum_{i=1}^{2} (\omega_{ii}\sigma_{ii} + \xi_i^2\sigma_{ii} + \mu_i^2\omega_{ii}) - 2\mu_1\mu_2\omega_{11}.$$

Substituting $\omega_{22} = V(W_2) = \omega_{11}$ and $\xi_2 = E(W_2) = 1 - \xi_1$ into this last expression and combining common terms yield

$$V(Y_2) = \sum_{i=1}^{2} \xi_i^2\sigma_{ii} + [(\mu_1 - \mu_2)^2 + (\sigma_{11} + \sigma_{22})] \times \omega_{11}. \tag{8}$$

Equation (8) shows that if an asphalt mixture in a plant-operation is blended from two stockpiles with variable feed rates such that $W_1 + W_2 \equiv 1$, then the exact variance of any characteristic of the mixture is given by equation (8). Bonaquist and Christensen [16] report the following equation for the variance of a two-stockpile mixture characteristic, $m = \alpha a + (1 - \alpha)b$, as

$$V(m) = \alpha^2\sigma_a^2 + (1 - \alpha)^2\sigma_b^2 + (\overline{X}_a - \overline{X}_b)^2\sigma_\alpha^2.$$

In our notation, $n = 2$, $m = Y_2$, $\alpha = W_1$, $1 - \alpha = W_2$, $a = X_1$, $b = X_2$, $\overline{X}_a = E(X_1) = \mu_1$, and $\overline{X}_b = E(X_2) = \mu_2$. Note that the Bonaquist and Christensen's [16] formula for $V(m)$ is an approximation to our equation (8) because the last two terms, $(\sigma_{11} + \sigma_{22}) \times \omega_{11}$, are left out of $V(m)$. However, the last two terms of equation (8), $(\sigma_{11} + \sigma_{22}) \times \omega_{11}$, are small relative to the other 3 terms unless the $CV(X_1) = \sigma_1/\mu_1$, $CV(X_2) = \sigma_2/\mu_2$, and $CV(W_1) = \omega_1/\xi_1$ all exceed 30%. It can be shown, see Appendix, that the $V(Y_2)$ given by equation (8), where $n = 2$ stockpiles, generalizes to our main result

$$V(Y_n) = \sum_{i=1}^{n} (\xi_i^2 + \omega_{ii})\sigma_{ii} + \sum_{i=1}^{n-1} (\mu_i - \mu_n)^2\omega_{ii}, \tag{9a}$$

for $n > 2$ stockpiles. Further, if we apply the approximation recommended by Bonaquist and Christensen [16] for $n = 2$ stockpiles, then equation (9a) further reduces to

$$V(Y_n) \cong \sum_{i=1}^{n} \xi_i^2 \sigma_{ii} + \sum_{i=1}^{n-1} (\mu_i - \mu_n)^2 \omega_{ii}. \tag{9b}$$

Equations (9a) and (9b) assume that only $(n - 1)$ out of the random proportions $W_1, W_2, \ldots, W_n$ are independent due to the constraint $\sum_{i=1}^{n} W_i \equiv 1$, i.e., there are only $(n - 1)$ degrees of freedom among the variates $W_1, W_2, \ldots, W_n$. Clearly, the stockpile designated as n impacts the variance given in equations (9a) and (9b). If the user denotes either the stockpile with maximum (or minimum) characteristic as n, then $V(Y_n)$ of equations (9) attain its near maximum (or conservative) value. Example 2 provides an application of the special case of $V(Y_n)$ as provided by equations (9a) and (9b).

Example 2. Suppose an asphalt plant produces a paving mixture from $n = 3$ stockpiles with the same parameter values as in Example 1, i.e., $E(X_1) = \mu_1 = 35.1$, $\mu_2 = E(X_2) = 46.4$, $\mu_3 = 62.8\%$, $V(X_1) = \sigma_{11} = 8.60$, $V(X_2) = \sigma_{22} = 16.40$, and $V(X_3) = \sigma_{33} = 12.96\%^2$. However, the feed rates cannot be exactly controlled such that $CV(W_i) = \omega_i / \xi_i = 15\%$ for $i = 1, 2, 3$, but $W_1 + W_2 + W_3 \equiv 1$ at any point in time during the process with $\xi_1 = E(W_1) = 0.40$, $\xi_2 = 0.35$ and $\xi_3 = E(W_3) = 0.25$. As noted above, $E(Y_3)$ remains unaffected and remains as $E(Y_3) = 45.98\%$ passing the 2.36mm-sieve, but the variance now must be computed from equation (9a). Because the coefficient of variation of each W_i is assumed to be 15%, then $\xi_1 = 0.40$, $\xi_2 = 0.35$, $\xi_3 = E(W_3) = 0.25$ imply that $\omega_{11} = V(W_1) = (0.15 \times 0.40)^2 = 0.0036$, $\omega_{22} = (0.15 \times 0.35)^2 = 0.0028$, and $\omega_{33} = (0.15 \times 0.25)^2 = 0.0014$. Substituting ω_{ii}'s $(i = 1, 2, 3)$ and $\mu_1 = 35.1$,

$\mu_2 = 46.4$, $\mu_n = 62.8\%$, $\sigma_{11} = 8.60$, $\sigma_{22} = 16.40$, $\sigma_{33} = 12.96$ into equation (9a) results in $V(Y_3) = V(W_1X_1 + W_2X_2 + W_3X_3) = 7.805356$, $\sigma(Y_3) = 2.79381$, and $CV(Y_3) = 6.05\%$. As expected, if W_i's are random variables, then $\sigma(Y_3)$ of Example 1 increases from 2.050% to 2.794%. This 36.283% increase clearly depends on the $CV(W_i)$; e.g., at $CV(W_i) = 10\%$, $\sigma(Y_3) = 2.4071\%$, which is an increase of 17.42%. Further, if we use the approximation of equation (9b), used by [16], then at $CV(W_i) = 15\%$, $\sigma(Y_3) \cong 2.77675\%$; compared to the exact 2.794%, demonstrating a close approximation to our exact value. It should be highlighted that in general the in-plant variability for X_i's are larger than those in the lab. In Example 2, we assumed the same variability for X_i's as those of Example 1 for accessing the increase in the $V(Y_3)$ when W_i's are random variables.

4. The Mean and Variance of Product of Two Correlated Random Variables

Consider the product $Y_1 = W_1X_1$, where $\upsilon_{11} = COV(W_1, X_1) = \omega_1\sigma_1\rho_{W_1, X_1} \neq 0$ is known and the objective is to compute the mean and variance of the product $Y_1 = W_1X_1$. By definition, $COV(W_1, X_1) = E(W_1X_1) - \xi_1\mu_1$, and thus

$$E(Y_1) = E(W_1X_1) = \xi_1\mu_1 + \upsilon_{11}. \tag{10}$$

Because W_1 and X_1 are not independent, then the $V(Y_1)$ is no longer given by equation (2a) as illustrated below:

$$V(W_1X_1) = E[(W_1X_1)^2] - E[(W_1X_1)]^2 = E(W_1^2X_1^2) - (\xi_1\mu_1 + \upsilon_{11})^2. \tag{11}$$

The 1st term on the RHS of equation (11) cannot be exactly computed unless the $COV(W_1^2, X_1^2)$ is known. Therefore, we resort to a Taylor's expansion of any function $f(W_1, X_1)$ about ξ_1 and μ_1:

$$f(W_1, X_1) = f(\xi_1, \mu_1) + \frac{\partial f}{\partial W_1}\bigg|_{(\xi_1, \mu_1)}(W_1 - \xi_1) + \frac{\partial f}{\partial X_1}\bigg|_{(\xi_1, \mu_1)}(X_1 - \mu_1)$$

$$+ \frac{1}{2}\frac{\partial^2 f}{\partial W_1^2}\bigg|_{(\xi_1, \mu_1)}(W_1 - \xi_1)^2 + \frac{1}{2}\frac{\partial^2 f}{\partial X_1^2}\bigg|_{(\xi_1, \mu_1)}(X_1 - \mu_1)^2$$

$$+ \frac{\partial^2 f}{\partial W_1 \partial X_1}\bigg|_{(\xi_1, \mu_1)}(W_1 - \xi_1)(X_1 - \mu_1) + R(W_1, X_1), \qquad (12)$$

where $R(W_1, X_1)$ is of order 3 or higher. Because in our special case $f(W_1, X_1) = W_1 X_1,$ then its Taylor's expansion from equation (12) reduces to

$$Y_1 = W_1 X_1 = \xi_1\mu_1 + \mu_1(W_1 - \xi_1) + \xi_1(X_1 - \mu_1) + (W_1 - \xi_1)(X_1 - \mu_1). \quad (13)$$

Note that in the special case of $f(W_1, X_1) = W_1 X_1,$ the Taylor expansion in (12) is an exact identity. To obtain the mean of $W_1 X_1,$ we apply the expected-value operator to both sides of equation (13):

$$E(W_1 X_1) \cong \xi_1\mu_1 + 0 + 0 + E[(W_1 - \xi_1)(X_1 - \mu_1)]$$

$$= \xi_1\mu_1 + Cov(W_1, X_1) = \xi_1\mu_1 + \upsilon_{11}. \qquad (14)$$

The mean of $W_1 X_1$ given in equation (14) is identical to that of equation (10), as expected.

To approximate the variance of $Y_1 = W_1 X_1,$ we apply the variance operator to equation (13) and ignore the last order-2 term. Thus,

$$V(W_1 X_1) \cong V[\mu_1(W_1 - \xi_1) + \xi_1(X_1 - \mu_1)]$$

$$\cong \mu_1^2 V(W_1 - \xi_1) + \xi_1^2 V(X_1 - \mu_1) + 2COV[\mu_1(W_1 - \xi_1), \xi_1(X_1 - \mu_1)]$$

$$\cong \mu_1^2\omega_{11} + \xi_1^2\sigma_{11} + 2\xi_1\mu_1 COV(W_1, X_1)$$

$$= \mu_1^2\omega_{11} + \xi_1^2\sigma_{11} + 2\xi_1\mu_1\upsilon_{11}. \qquad (15)$$

The approximate $V(W_1X_1)$ in equation (15) is fairly close to the exact $V(Y_1) = \xi_1^2\sigma_1^2 + \mu_1^2\omega_1^2 + \omega_1^2\sigma_1^2$ given in equation (2b) for the case when W_1 and X_1 are independent. Unfortunately, the approximation in equation (15) does not reduce to the exact result of $\xi_1^2\sigma_1^2 + \mu_1^2\omega_1^2 + \omega_1^2\sigma_1^2$ when W_1 and X_1 are independent for which $\upsilon_{11} = 0$ because the Taylor expansion was truncated. However, when both ω_1/ξ_1 and σ_1/μ_1 are less than 30%, the product $\omega_1^2\sigma_1^2 < (0.30\xi_1)^2(0.30\mu_1)^2 = 0.0081\xi_1^2\mu_1^2$ so that $\omega_1^2\sigma_1^2$ is much smaller than either $\mu_1^2\omega_1^2$ or $\xi_1^2\sigma_1^2$, and thus, the approximation in (15) is fair agreement with equation (2b). For the worst-case scenario of $CV \geq 30\%$, equation (15) further shows that the $V(Y_1) = V(W_1X_1)$ is an increasing function of the process correlation coefficient, ρ_{W_1, X_1}, between W_1 and X_1.

Now consider the most general output $Y_n = \sum_{i=1}^{n} W_i X_i$, where all the $2 \times n$ random variables are correlated with covariance structure

$$\Sigma_W = \begin{bmatrix} \omega_{11} & \omega_{12} & \omega_{13} & \cdots & \omega_{1n} \\ \omega_{21} & \omega_{22} & \omega_{23} & \cdots & \omega_{1n} \\ \omega_{31} & \omega_{32} & \omega_{33} & \cdots & \omega_{1n} \\ \vdots & \vdots & \vdots & & \vdots \\ \omega_{n1} & \omega_{n2} & \omega_{n3} & \cdots & \omega_{nn} \end{bmatrix},$$

where $\omega_{ij} = E[(W_i - \xi_i) \times (W_j - \xi_j)]$ represents the covariance between W_i and W_j. Similarly, X_i's are correlated random variables with means $E(X_i) = \mu_i$ and covariance structure

$$\Sigma_X = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \cdots & \sigma_{1n} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & \cdots & \sigma_{1n} \\ \vdots & \vdots & \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} & \sigma_{n3} & \cdots & \sigma_{nn} \end{bmatrix},$$

and W_i and X_i also are correlated with covariance structure

$$\Sigma_{WX} = \begin{bmatrix} \upsilon_{11} & \upsilon_{12} & \cdots & \upsilon_{1n} \\ \upsilon_{21} & \upsilon_{22} & \cdots & \upsilon_{2n} \\ . & . & \cdots & . \\ \upsilon_{n1} & \upsilon_{n2} & \cdots & \upsilon_{nn} \end{bmatrix},$$

then we have the following 2nd-order approximation for the mean of Y_n:

$$E(Y_n) \approx \sum_{i=1}^{n} \xi_i \mu_i + \sum_{i=1}^{n} \upsilon_{ii} \tag{16a}$$

and rough 1st-order approximation for the variance of Y_n is given by

$$V(Y_n) \cong \sum_{i=1}^{n} (\mu_i^2 \omega_{ii} + \xi_i^2 \sigma_{ii}) + 2\sum_{i=1}^{n-1} \sum_{j>i}^{n} (\mu_i \mu_j \omega_{ij} + \xi_i \xi_j \sigma_{ij})$$

$$+ 2\sum_{i=1}^{n} \sum_{j=i}^{n} \mu_i \xi_j \upsilon_{ij}. \tag{16b}$$

5. Conclusions

This article first generalized the known approximate result for the variance of a mixture characteristic having two ingredients to the case of more than $n = 2$ ingredients. Equation (9a) is an exact formula and (9b) is the corresponding approximation for the practitioner. Secondly, equations (16a) and (16b) give the approximate formulas for the mean and variance, respectively, of an output $Y_n = \sum_{i=1}^{n} W_i X_i$ under the most general case that the $2n$ random variables W_i and X_i are correlated. These formulas provide the reader with a method of calculating approximate mean and variance of an asphalt paving mixture treated as random variables witnessed in plant operations. We have prepared an Excel® spreadsheet to assist with calculations. The spreadsheet is available upon request.

References

[1] J. L. Devore, Probability and Statistics for Engineers and Sciences, 9th ed., Brooks Cole, Boston, USA, 2015.

[2] S. Abo-Qudais and H. Al-Shweily, Effect of aggregate properties on asphalt mixtures stripping and creep behavior, Construction and Building Materials 21 (2007), 1886-1898.

[3] J. A. Cornell, Experiments with Mixtures: Designs, Models, and the Analysis of Mixture Data, 2nd ed., John Wiley and Sons, New York, USA, 2002.

[4] H. Al-Mosawe, N. Thom and G. Airey, Effect of aggregate gradation on the stiffness of asphalt mixtures, Inter. J. Pavement Engin. Asphalt Tech. 16(2) (2015), 39-49.

[5] D. C. Montgomery and G. C. Runger, Applied Statistics and Probability for Engineers, 6th ed., John Wiley and Sons, New York, USA, 2014.

[6] J. G. Speight, Asphalt Materials Science and Technology, Elsevier Inc., MA, USA, 2015.

[7] A. Kuity and A. Das, Study on aggregate size distribution in asphalt mix using images obtained by different imaging techniques, Transport. Research Procedia 17 (2016), 340-348.

[8] D. D. Wackerly, W. Mendenhall and R. L. Scheaffer, Mathematical Statistics with Applications, 7th ed., Thomson Learning, Inc., CA, USA, 2008.

[9] Y. J Jiang and L. F. Fan, An experimental investigation of optimal asphalt-aggregate ratio for different compaction methods, Construction and Building Materials 91 (2015), 111-115.

[10] P. D. Valle and N. Thom, Variability in pavement design, Inter. J. Pavement Engin. Asphalt Tech. 16(2) (2015), 50-67.

[11] G. Valdes-Vidal, A. Calabi-Floody, R. Miro-Recasens and J. Norambuena-Contreras, Mechanical behaviors of asphalt mixtures with different aggregate type, Construction and Building Materials 101 (2015), 474-481.

[12] J. G. Speight, The Chemistry and Technology of Petroleum, 5th ed., Taylor and Francis Group, FL, USA, 2014.

[13] M. H. DeGroot and M. J. Schervish, Probability and Statistics, 3rd ed., Addison-Wesley, Boston, USA, 2002.

[14] L. Bruno, G. Parla and C. Celauro, Image analysis for detecting aggregate

gradation in asphalt mixture from planar images, Construction and Building Materials 28 (2012), 21-30.

[15] L. A. Aroian, The probability function of a product of two normally distributed variables, Ann. Math. Stat. 18 (1947), 265-271.

[16] R. Bonaquist and D. Christensen, A method to estimate the effect of RAP variability, NCHRP Project 9-33, 2008.

[17] C. C. Craig, On the frequency function of *xy*, Ann. Math. Stat. 7 (1936), 1-15.

[18] K. C. Kapur and L. R. Lamberson, Reliability in Engineering Design, 1st ed., John Wiley and Sons, New York, USA, 1977.

[19] M. D. Springer, The Algebra of Random Variables, 1st ed., John Wiley and Sons, New York, USA, 1979.

[20] W. Navidi, Statistics for Engineers and Scientists, 2nd ed., McGraw-Hill Companies, New York, NY, USA, 2008.

[21] R. A. Johnson, Miller and Freund's Probability and Statistics for Engineers, 7th ed., Pearson Education, NJ, USA, 2005.

[22] G. Vining and S. M. Kowalski, Statistical Methods for Engineers, 2nd ed., Thomson Brooks Cole, CA, USA, 2006.

Appendix

The derivation of $V\left(\sum_{i=1}^{n} W_i X_i\right)$, for $n > 2$, under the constraint $\sum_{i=1}^{n} W_i \equiv 1$ and the assumption of independent X_i's

The constraint $\sum_{i=1}^{n} W_i \equiv 1$ implies that $E(W_n) = \xi_n = 1 - \sum_{i=1}^{n-1} \xi_i$, and $V(W_n) = \omega_{nn} = V(1 - W_1 - W_2 - W_3 - \cdots - W_{n-1}) = \sum_{i=1}^{n-1} \omega_{ii}$. Further, W_1, W_2, ..., W_{n-1} are jointly independent but $W_1, W_2, ..., W_{n-1}$ are correlated with W_n, i.e., $\omega_{ij} = COV(W_i, W_j) = 0$ for all i and $j \neq n$, but

$$\omega_{1n} = COV(W_1, W_n) = COV(W_1, 1 - W_1 - W_2 - \cdots - W_{n-1})$$

$$= -COV(W_1, W_1 + W_2 + \cdots + W_{n-1}) = -V(W_1) = -\omega_{11}.$$

Similarly, $\omega_{in} = -\omega_{ii}$ for all $i = 1, 2, \ldots, n-1$. The use of equation (9) leads to

$$V(Y_n) = V\left(\sum_{i=1}^{n} W_i X_i\right) = \sum_{i=1}^{n} (\omega_{ii}\sigma_{ii} + \xi_i^2\sigma_{ii} + \mu_i^2\omega_{ii}) + 2\sum_{i=1}^{n-1}\sum_{j>i}^{n} \mu_i\mu_j\omega_{ij}$$

$$= \sum_{i=1}^{n} (\omega_{ii}\sigma_{ii} + \xi_i^2\sigma_{ii}) + \sum_{i=1}^{n} \mu_i^2\omega_{ii} + 2\sum_{i=1}^{n-1}\sum_{j>i}^{n} \mu_i\mu_j\omega_{ij}.$$

Substituting $\omega_{nn} = \sum_{i=1}^{n-1}\omega_{ii}$ and $\omega_{ij} = COV(W_i, W_j) = 0$ for all i and $j \neq n$ into the last formula, we obtain

$$V(Y_n) = \sum_{i=1}^{n} (\omega_{ii}\sigma_{ii} + \xi_i^2\sigma_{ii}) + \sum_{i=1}^{n-1} \mu_i^2\omega_{ii} + \mu_n^2\omega_{nn} + 2\sum_{i=1}^{n-1} \mu_i\mu_n\omega_{in}$$

$$= \sum_{i=1}^{n} (\omega_{ii}\sigma_{ii} + \xi_i^2\sigma_{ii}) + \sum_{i=1}^{n-1} \mu_i^2\omega_{ii} + \mu_n^2\sum_{i=1}^{n-1} \omega_{ii} - 2\sum_{i=1}^{n-1} \mu_i\mu_n\omega_{ii}$$

$$= \sum_{i=1}^{n} (\omega_{ii}\sigma_{ii} + \xi_i^2\sigma_{ii}) + \sum_{i=1}^{n-1} (\mu_i^2 + \mu_n^2 - 2\mu_i\mu_n)\omega_{ii}$$

$$= \sum_{i=1}^{n} (\omega_{ii} + \xi_i^2)\sigma_{ii} + \sum_{i=1}^{n-1} (\mu_i - \mu_n)^2\,\omega_{ii},$$

completing the proof.