A direct calculation of moments of the sample variance

Gonzalo Vegas-Sánchez-Ferrero a,*, Santiago Aja-Fernández a, Marcos Martín-Fernández a, César Palencia b

a Laboratorio de Procesado de Imagen – Universidad de Valladolid, Spain
b Departamento de Matemática Aplicada y Computación – Universidad de Valladolid, Spain

Received 28 November 2010; received in revised form 5 October 2011; accepted 3 November 2011
Available online 22 November 2011

Abstract

A systematic method to deal with the interrelations of systems with multi-index quantities (random variables) is proposed. The method differs of the well-known Polykays. An application of the theoretical results here presented is the calculation of the moments of the sample variance for general populations in a direct way. The main advantage of the proposed methodology is that no conversion formulae and other complicated Polykays rules are needed. However, the proposed method is compatible with Polykays philosophy and conversion formulae and multiplication rules can be derived by using the theoretical results of this work. For practical purposes, two algorithms for the calculation of the moments of the sample variance are proposed.

© 2011 IMACS. Published by Elsevier B.V. All rights reserved.

Keywords: Sample variance; Moments; Polykays; Sample moments

1. Introduction

The approximation of the probability density function (PDF) of sample variances from nonnormal universes becomes a classical and important problem for studying the deviation from normality on the analysis of variance and covariance [18]. One practical example can be found in [2], where authors use the sample (local) variance distribution of Rician and noncentral Chi data to model the background/signal areas or magnetic resonance images. Other examples comprise gene classification and DNA analysis [10] and studies in nonlinear matter power spectrum [14].

Some methods have been reported in literature to approximate the PDF of sample variance of general populations, see for instance [4,5,13,19,18]. In order to estimate the PDF, these methods need the computation of higher order moments of the sample variance. This becomes a difficult task, since there is not a closed form to calculate them for general populations. The usual way to do it is by using the Polykays method by Tukey [21]. However, this is a complex methodology which needs conversion formulae and rules for the product of Polykays that have to be calculated in a non-direct way.

Concretely, Tukey uses the augmented symmetric functions for a set of \( x_1, x_2, \ldots, x_n \) defined as:

\[
[p_1^{p_1}, p_2^{p_2}, \ldots, p_n^{p_n}] = \sum x_1^{p_1} x_2^{p_2} \ldots x_q^{p_q} x_{q+1}^{p_{q+1}} \ldots x_{q+n-1}^{p_{q+n-1}}
\]  

* Corresponding author. Tel.: +34 677432975.
E-mail addresses: gvegsan@lpi.tel.uva.es (G. Vegas-Sánchez-Ferrero), sanaja@tel.uva.es (S. Aja-Fernández), marcma@tel.uva.es (M. Martín-Fernández), cesarpal@mac.uva.es (C. Palencia).
where there are $\pi_1$ powers $p_1$, $\pi_2$ powers $p_2$, and so on. All the suffixes are different and the summation takes place over all values of $x$. Thus, the expression has $n(n-1)(n-2)\ldots$ terms in the summation. For example:

$$[1^223] = \sum x_i^2 x_j x_k; \quad [2^3] = \sum x_i^2 x_j^2$$

Additionally, Tukey also uses the monomial symmetric functions defined as:

$$(p_{\pi_1}^{\pi_1}, p_{\pi_2}^{\pi_2}, \ldots, p_{\pi_s}^{\pi_s}) = \frac{[p_{\pi_1}^{\pi_1}, p_{\pi_2}^{\pi_2}, \ldots, p_{\pi_s}^{\pi_s}]}{\pi_1!\pi_2!\cdots\pi_s!}$$

From these definitions, one can state the following fundamental result:

$$E([p_{\pi_1}^{\pi_1}, p_{\pi_2}^{\pi_2}, \ldots, p_{\pi_s}^{\pi_s}]) = n(n-1)\cdot(n-\rho+1)\mu_{\pi_1}\mu_{\pi_2}\cdots\mu_{\pi_s}$$

where $\rho = \sum_{i=1}^{s} \pi_i$ and $\mu_i$ is the $i$th raw moment.

The calculation of the moments of sample variance can be defined by expressions including terms with monomial symmetric functions, so the problem is reduced to express the sample variance as sums of augmented symmetric functions. Some tables exist giving these functions in terms of one another. However, the derivation of these tables is not systematic and needs the calculation of power-sums of lower weights. The original tables were heroically (manually) calculated up to 10th order by M.G. Kendall in two different ways and were independently checked by F.N. David in [6].

In this work we propose a systematic method to deal with the interrelations of systems with multi-index quantities (random variables) in a different way as it is done with Polykays. Two theorems are proposed that will allow a direct computation of moments of sample variance for general populations. The main advantage of this new methodology is that there is no need of conversion formulae, multiplication tables and other complicated Polykays rules [21,20,6]. So, the moments can be calculated in a closed form for any order.

As a result, the calculation of the higher order moments will be straightforward, and the algorithm here proposed will show a computational gain up to order 15.

The paper is structured as follows: Section 2 establishes the problem of the moments of the sample variance and presents two theorems as the main results of this work. Additionally, the case of the variance of the sample variance is explained as an example following the theorems presented (Section 3). Section 4 is devoted to the implementation details, where two different implementations of the method are described: a direct one and a more refined method in order to reduce the computations. In Section 5, we conclude.

2. Theory

2.1. The problem of moments of sample variance

Let $V$ be the (unbiased) sample variance of $X_1, X_2, \ldots, X_N$ Independent and Identically Distributed (IID) random variables, defined as:

$$V = \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \bar{X})^2 = \frac{1}{N(N-1)} \left( N \sum_{i=1}^{N} X_i^2 - \left( \sum_{i=1}^{N} X_i \right)^2 \right)$$
Fig. 1. This graph represents all possible combinations of \( M = m + n \) groups of \( N \) indices per group. Sums of \( N \) different indices are depicted in columns, whereas rows show the product combination. Each path from left to right gives a different combination of products of random variables.

The \( j \)th raw moment of the sample variance is

\[
E \{ V^j \} = E \left\{ \frac{1}{N^j (N - 1)^j} \left( N \sum_{i=1}^{N} X_i^2 - \left( \sum_{i=1}^{N} X_i \right)^2 \right)^j \right\} 
\]

\[
= E \left\{ \frac{1}{N^j (N - 1)^j} \sum_{k=0}^{j} \binom{j}{k} N^k (-1)^{j-k} \left( \sum_{i=1}^{N} X_i^2 \right)^k \left( \sum_{i=1}^{N} X_i \right)^{2(j-k)} \right\} 
\]

\[
= \frac{1}{N^j (N - 1)^j} \sum_{k=0}^{j} \binom{j}{k} N^k (-1)^{j-k} E \left\{ \left( \sum_{i=1}^{N} X_i^2 \right)^k \left( \sum_{i=1}^{N} X_i \right)^{2(j-k)} \right\} 
\]

(5)

So, according to Eq. (5), the problem of the calculation of the raw moments of the sample variance is reduced to the calculus of the expected value of all the combinations of products of sums of two different powers of random variables.

2.2. Sketch of the method

For the sake of simplicity, we will refer to this problem in its general form as:

\[
E \left\{ \left( \sum_{i=1}^{N} X_i^2 \right)^m \left( \sum_{i=1}^{N} X_i \right)^n \right\} 
\]

(6)

The expansion of the product of two sums will give sums of different arrangements of random variables of different indices. Since \( X_i \) are IID random variables, the expectation of those arrangements are the product of the raw moments of the same order as the exponent of the random variable. Hence, all arrangements with the same number of different indices are equivalent in terms of its expectations and they can only differ in the power of each random variable, i.e. the order of each moment.

This problem can be seen as a graph (see Fig. 1), where each vertex is an index of each sum which can be combined with the indices of the following sums. Thus, this graph represents all possible combinations of \( M = m + n \) groups of \( N \) indices per group. Obviously the number of combinations is \( N^M \).

In order to construct the minimum set of combinations, the method will focus on those arrangements of different number of indices. This will avoid calculating redundant arrangements of indices. In the next section we give some theoretical results that provide a methodology to calculate the solution of Eq. (6) in a suitable way.

Concretely, two theorems will be demonstrated; the first one allows constructing all the combinations as the disjoint union of sets of \( s^r \) different indices with a certain multiplicity, \( c_s \), for each index. Note that this theorem performs a suitable way to describe the whole set of combinations in order to make an easier way to treat them, but no distinction is taken into account concerning the exponent of each random variable. The second theorem makes use of the result of the first, and allows us to distinguish between variables with different exponents.
2.3. Main results

**Theorem 1.** Let $\mathcal{A}$ be the set of all possible combinations of $M$ groups of indices with $N$ indices per group, and let $\mathcal{A}_{s^*}$ be a set of all possible combinations of $s^*$ different indices where $s^* = \{1, \ldots, \min(M, N)\}$. Then, the sets $\mathcal{A}_{s^*}$ are mutually disjoint sets of the form:

$$\mathcal{A}_{s^*} = \bigcup_{(s, c) \in S_{s^*} \times C_{s^*}} B(s, c)$$  \hspace{1cm} (7)

and $\mathcal{A}$ is the union of all these sets:

$$\mathcal{A} = \bigcup_{s^*=1}^{\min(M,N)} \mathcal{A}_{s^*} = \bigcup_{s^*=1}^{\min(M,N)} \bigcup_{(s, c) \in S_{s^*} \times C_{s^*}} B(s, c)$$  \hspace{1cm} (8)

where $S_{s^*}$ is the set of the combinations of $s^*$ indices out of $N$, $C_{s^*}$ is the set of all possible solutions of the composition of $M$ as the sum of $s^*$ strictly positive integers, and $B(s, c)$ is the set of all permutations without repetition of $s \in S_{s^*}$ with multiplicity $c \in C_{s^*}$.

**Proof.** The set $S_{s^*}$ of $s^*$ different indices out of $N$ possibilities is constructed from the combinations of $N$ elements in groups of $s^*$ elements, thus, its cardinal is $\binom{N}{s^*}$. From this set, we can construct the set $\mathcal{A}_{s^*}$, of all possible combinations of $M$ elements with $s^*$ different indices. This is done just by knowing the multiplicity of indices that must hold the following Diophantine equation:

$$c_1 + c_2 + \cdots + c_{s^*} = M$$  \hspace{1cm} (9)

where $c_i$ with $i = 1, \ldots, s^*$ is the non-zero multiplicity of each index.

Note that the solutions of Eq. (9) are the same as the problem in Number Theory of the composition of a number, $M$, as a sum of $s^*$ strictly positive integers. It also can be seen as a partition problem when order is taken into consideration. We call $C_{s^*}$ to the set of solutions of the form $[c_1, c_2, \ldots, c_{s^*}]$ of Eq. (9) for $s^*$ different indices.

Every element of the product set $(s, c) \in S_{s^*} \times C_{s^*}$ defines a unique set of combinations, $B(s, c)$ of $\binom{M}{c}$ elements which is constructed as the permutations without repetition of indices $s$ with multiplicity $c$. Note that, for a fix $s \in S_{s^*}$, $B(s, c)$ is unique because, by construction, each pair $(s, c)$ defines a different set $B(s, c)$.

Finally, the set $\mathcal{A}_{s^*}$ is constructed as the union of all the disjoint sets $B(s, c)$

$$\mathcal{A}_{s^*} = \bigcup_{(s, c) \in S_{s^*} \times C_{s^*}} B(s, c)$$  \hspace{1cm} (10)

So the cardinal of $\mathcal{A}_{s^*}$ can be calculated as follows:

$$|\mathcal{A}_{s^*}| = \sum_{(s, c) \in S_{s^*} \times C_{s^*}} |B(s, c)| = \sum_{(s, c) \in S_{s^*} \times C_{s^*}} \binom{M}{c} = \binom{N}{s^*} \sum_{c \in C_{s^*}} \binom{M}{c} = \binom{N}{s^*} \sum_{k=0}^{s^*} \binom{s^*}{k} (s^* - k)^M (-1)^k$$  \hspace{1cm} (11)

The last equality is obtained by using the multinomial theorem:

$$\sum_{c_1 \geq 0} c_1 + \cdots + c_{s^*} = M \binom{M}{c_1, \ldots, c_{s^*}} = (s^*)^M$$  \hspace{1cm} (12)
in combination with the inclusion–exclusion principle of the sets of non-negative solutions of Eq. (9):

\[
\sum_{c_i \geq 0} \binom{M}{c_1, \ldots, c_s} = \sum_{c_i \geq 0} \binom{M}{c_1, \ldots, c_s} = \binom{M}{s^*} - \binom{s^*}{1} \sum_{c_i \geq 0} \binom{M}{c_1, \ldots, c_{(s^*-1)}} + \cdots + \binom{s^*}{s^* - 1} \sum_{c_i \geq 0} \binom{M}{c_1} \tag{13}
\]

Hence, for each combination of \(s^* = 1, \ldots, \min(M, N)\) different indices, the set of all possible combinations, \(\mathcal{A}\), is:

\[
\mathcal{A} = \bigcup_{s^* = 1}^{\min(M, N)} \mathcal{A}_{s^*} = \bigcup_{s^* = 1}^{\min(M, N)} \bigcup_{(s, e) \in S_{s^*} \times C_{s^*}} \mathcal{B}(s, e) \tag{14}
\]

and its cardinal is, from Eqs. (14) and (11):

\[
|\mathcal{A}| = \sum_{s^* = 1}^{\min(M, N)} \sum_{(s, e) \in S_{s^*} \times C_{s^*}} \binom{M}{c} = \sum_{s^* = 1}^{\min(M, N)} \binom{N}{s^*} \sum_{k=0}^{s^*} \binom{s^*}{k} (s^* - k)^M (-1)^k = N^M \tag{15}
\]

This theorem allows us to calculate all the combinations of indices in the following way: since \(X_i\) are IID random variables, the index of each variable does not matter but the number of different indices per combination. So, for each number of different indices, \(s^*\), we have \(\binom{N}{s^*}\) equivalent combinations.

The solutions of the combination problem provide different arrangements of the combinations of indices. However, it is important to note that many of them are equivalent and there is no need to calculate all of them.

For instance, in Table 1 an example of the construction of the set \(\mathcal{A}\) is presented for \(M = 4\) and \(N = 4\). Note that when \(X_i\) are IID random variables, all the compositions that just differ in order are equivalent. Thus, the set of solutions of the composition problem can be dramatically reduced into a smaller set of solutions with multiplicities for each one. The sets of solutions can be calculated as an integer partition problem where order does not matter and the multiplicities \(\sigma(\cdot)\) can be calculated as a multinomial coefficient, we denote the set of solutions of the integer partition problem as \(\mathcal{P}_{s^* = 1, \ldots, 4}\):

\[
p_1 + \cdots + p_{s^*} = 4
\]

\(\mathcal{P}_{s^* = 1, \ldots, 4} = \{\{4\}, \{2, 2\}, \{3, 1\}, \{2, 1, 1\}, \{1, 1, 1, 1\}\}
\]

\[
\sigma(\mathcal{P}_{s^* = 1, \ldots, 4}) = \left\{ \binom{1}{1} \cdot \binom{2}{2} \cdot \binom{1}{1} \cdot \binom{3}{1} \cdot \binom{4}{4} \right\}
\]

Table 2 shows a comparison between the number of solutions of the composition problem and the partition problem. The number of solutions of the composition problem, \(|\mathcal{C}_{s^* = 1, \ldots, M}|\), can be calculated easily by placing either a plus sign or a comma in each of the \(M - 1\) boxes of the array:

\[
\begin{array}{c}
\hline
1 \quad 1 \quad \cdots \quad 1 \\
\hline
\end{array}
\]

\(M\)

(16)

Since every composition of \(M\) can be determined by an assignment of pluses and commas, the number of compositions is given by the binomial coefficient \(\binom{M - 1}{k - 1}\) for \(k\) parts, where \(k = 1, \ldots, M\):

\[
|\mathcal{C}_{s^* = 1, \ldots, M}| = \sum_{k=1}^{M} \binom{M - 1}{k - 1} = 2^{M-1} \tag{17}
\]

In the case of the partition problem, \(|\mathcal{P}_{s^* = 1, \ldots, M}|\), one can compute the number of solutions by means of the Hardy–Ramanujan–Rademacher formula [3].
Table 1
Example of the construction of the set \( A \) for \( N = 3 \) and \( M = 4 \) for each value of \( s^* = \{1, 2, 3\} \) different indices for \( (\sum_{i=1}^{3} X_i^2)^2(\sum_{i=1}^{3} X_i)^2 \).

| \( s^* \) | \( S_{s^*}, C_{s^*} \) | \( |S_{s^*}| = \binom{N}{i^*}, |C_{s^*}| \) | \( |\mathcal{B}_{(s,c)}| = \binom{M}{c} \) | \( |A_{s^*}| \) |
|---------|-----------------|--------------------------|-----------------|----------------|
| 1       | \( S_1 = \{(1), (2), (3)\} \) | \( |S_1| = 3, |C_1| = 1 \) | \( \mathcal{B}_{(i,(4))} = \{(i, i, i, i)\} \) | \( |A_1| = 3 \) |
|         | \( C_1 = \{(4)\} \) | | | |
| 2       | \( S_2 = \{(1, 2)\}, \{(1, 3)\}, \{(2, 3)\}, \{(1, 3)\}, \{(2, 2)\}, \{(3, 1)\} \) | \( |S_2| = 3, |C_2| = 3 \) | \( \mathcal{B}_{(i,(1,3))} = \{(i, j, j, j), (j, i, j, j), (j, j, i, j), (j, j, j, i)\} \) | \( |A_2| = 42 \) |
| 3       | \( S_3 = \{(1, 2, 3)\} \) | \( |S_3| = 1, |C_3| = 3 \) | \( \mathcal{B}_{(i,(k,(2,1,1)))} = \{(i, i, j, k), \ldots\} \) | \( |A_3| = 36 \) |

\( A = \bigcup_{s^*=1}^{\min(M,N)} A_{s^*} = \{(1, 1, 1, 1), (2, 2, 2, 2), (3, 3, 3, 3), (1, 1, 1, 2), \ldots\} \) | \( |A| = N^M = 81 \)
Comparison between the number of solutions of Eq. (9) when it is considered a composition problem versus when it is considered as an integer partition problem.

<table>
<thead>
<tr>
<th>M</th>
<th>Composition solutions</th>
<th>Partition solutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>512</td>
<td>42</td>
</tr>
<tr>
<td>20</td>
<td>524288</td>
<td>627</td>
</tr>
<tr>
<td>30</td>
<td>536870912</td>
<td>5604</td>
</tr>
<tr>
<td>40</td>
<td>549755813888</td>
<td>37338</td>
</tr>
<tr>
<td>50</td>
<td>562949953421312</td>
<td>204226</td>
</tr>
<tr>
<td>60</td>
<td>57646075230342348</td>
<td>966467</td>
</tr>
<tr>
<td>70</td>
<td>59029581035870561712</td>
<td>4087968</td>
</tr>
<tr>
<td>80</td>
<td>604462909807314587353088</td>
<td>15796476</td>
</tr>
<tr>
<td>90</td>
<td>618970019642690137449562112</td>
<td>56634173</td>
</tr>
<tr>
<td>100</td>
<td>633825300114117048351602688</td>
<td>190569292</td>
</tr>
</tbody>
</table>

Obviously, this way for calculating the solutions of Eq. (9) is highly more efficient as M increases.

In the following theorem we link the arrangements of indices with the exponents of each random variable.

**Theorem 2.** Let $X_i$ be IID random variables with $2m+n$ finite raw moments $\mu_1, \ldots, \mu_{2m+n}$. Then, the following equality holds:

$$E\left\{ \left( \sum_{i=1}^{N} X_i^2 \right)^m \left( \sum_{i=1}^{N} X_i \right)^n \right\} = \min(M, N) \sum_{s^*=1}^{M} \sum_{p \in P_s} \sum_{j=1}^{s^*} \prod_{k=1}^{n} \mu E_{j,k}$$

where $M=m+n$, $P_s$ is the set of solutions of the integer partition problem for $M$ into $s^*$ terms. $\sigma(p)$ is a multinomial coefficient which indicates the number of permutations without repetition of $p \in P_s$, $E_{j,k} = \sum_{i=1}^{M} e(i) \delta(b_j(i) - k)$ with $e = [2, \ldots, 2, 1, \ldots, 1]$, $\delta$ is the Kronecker delta and $b_j \in B_{s^*,p}$ is a permutation without repetition of $s^*$ elements with multiplicities the elements of $p$.

**Proof.** First we prove that there is no need to construct the set $B_{s,c}$ for each $s \in S_{s^*}$ when dealing with IID random variables. This is an important issue for its numerical implementation.

Let $b \in B_{s,c}$ be one of the vector of indices $s$ with multiplicities $c$. Products of random variables with indices $s$ can be defined as:

$$\lambda(b, s) = \prod_{k=1}^{s^*} X_{c_k}^{E_k(s)}$$

where

$$E_k(s) = \sum_{i=1}^{M} e(i) \delta(b(i) - s(k))$$

and $e = [2, \ldots, 2, 1, \ldots, 1]$. Let $B_{s^*,c}$ be the set of combinations of indices $[1, 2, \ldots, s^*]$ with multiplicities $c$. Note that this set of combinations does not depend on the elements of $S_{s^*}$ but on the number of different indices $s^*$. So, for a fixed $s^*$ and $c$, a bijection can be established between elements of $B_{s^*,c}$ and $B_{s,c}$ in such a way that, for every $b \in B_{s,c}$, there exists one $b^* \in B_{s^*,c}$ such that:

$$E_k(s) = \sum_{i=1}^{M} e(i) \delta(b^*(i) - k) = E_k$$
And, $|B_{k,e}| = |B_{s^*, e}|$. Hence,

$$E\{\lambda(b, s)\} = E\left\{ \prod_{k=1}^{s^*} X_{E_k}^{E_k(i)} \right\} = \prod_{k=1}^{s^*} E\{X_k^{E_k}\} = E\{\lambda(b^*)\}$$  \(\text{(21)}\)

which demonstrates the equality:

$$E\left\{ \sum_{a \in A_{s^*}} \lambda(a) \right\} = E\left\{ \sum_{(s,e) \in S_s \times C_e} \sum_{b_j \in B_{s,e}} \lambda(b) \right\} = E\left\{ \left( \sum_{s^*} \sum_{e \in C_e} \sum_{b_j \in B_{s^*, e}} \lambda(b^*) \right) \right\}$$  \(\text{(22)}\)

Now, the construction of $B_{s^*, e} = \{b_1, b_2, \ldots, b \binom{M}{c}\}$ becomes direct just by creating all the permutations without repetition of a vector of indices with multiplicity the elements of $e$.

All element $b_j \in B_{s^*, e}$ has a direct relationship with the exponents of the random variables, i.e. the first $m$ indices are those of random variables with power 2, whereas the last $n$ indices are those of power 1. So, the order of the moment of the random variable is the sum of all the powers of the same index:

$$\lambda(B_{s^*, e}) = \left\{ \begin{array}{c}
X_1^{E_1.1}X_2^{E_1.2}\ldots X_{s^*}^{E_1.s^*} \\
X_1^{E_2.1}X_2^{E_2.2}\ldots X_{s^*}^{E_2.s^*} \\
\vdots \\
X_1^{E} \binom{M}{c.1} X_2^{E} \binom{M}{c.2} \ldots X_{s^*}^{E} \binom{M}{c.s^*} \\
\end{array} \right\}$$  \(\text{(23)}\)

where

$$E_{j,k} = \sum_{i=1}^{M} e(i) \delta(b_j(i) - k)$$
$$e = [2, \ldots, 2, 1, \ldots, 1]$$

All in all, from Eqs. (8) and (22), we have:

$$E \left\{ \left( \sum_{i=1}^{N} X_i^2 \right)^m \left( \sum_{i=1}^{N} X_i \right)^n \right\} = E \left\{ \sum_{a \in A} \lambda(a) \right\} = E \left\{ \sum_{s^*=1}^{\min(M,N)} \sum_{(s,e) \in S_s \times C_e} \sum_{b_j \in B_{s^*, e}} \lambda(b) \right\}$$
$$= \sum_{s^*=1}^{\min(M,N)} \binom{M}{s^*} \sum_{p \in F_{s^*}} \sigma(p) \sum_{b_j \in B_{s^*, p}} E\{\lambda(b)\}$$  \(\text{(24)}\)

where the last equality is obtained due to the equivalence of the compositions that just differ in order.
So, finally, Eq. (23) leads to

\[
\min(M,N) \sum_{s^*=1} \left( \begin{array}{c} N \\ s^* \end{array} \right) \sum_{p \in \mathcal{P}_{s^*}} \sigma(p) \sum_{j=1} E[\lambda(b_j)] = \sum_{s^*=1} \left( \begin{array}{c} M \\ p \end{array} \right) \sum_{j=1} E\{X_{1}^{E_{j,1}}X_{2}^{E_{j,2}} \ldots X_{s^*}^{E_{j,s^*}}\} = \sum_{s^*=1} \left( \begin{array}{c} N \\ s^* \end{array} \right) \sum_{p \in \mathcal{P}_{s^*}} \sigma(p) \sum_{j=1} \prod_{k=1} \mu_{E_{j,k}}
\]

which completes the proof. □

For the sake of clarity, let us see the example of Table 1. For \( s^* = 2 \) it is clear that:

\[
\sum_{b \in \mathcal{B}_{(0,1,2),1}} E[\lambda(b)] = \sum_{b \in \mathcal{B}_{(0,1,2),1}} E[\lambda(b)]
\]

and the same happens for \( s^* = 3 \):

\[
\sum_{b \in \mathcal{B}_{(0,1,2,3),1}} E[\lambda(b)] = \sum_{b \in \mathcal{B}_{(0,1,2,3),1}} E[\lambda(b)] = \sum_{b \in \mathcal{B}_{(0,1,2,3),1}} E[\lambda(b)]
\]

One of the potentials of Theorem 2 is that it can be easily generalized to more complex products of sums of random variables, since the multiplicity of each sum is defined by \( e \). Then, the following corollary of Theorem 2 can be established.

**Corollary 1.** Let \( X_i \) be IID random variables with \( t_1 m_1 + t_2 m_T \) finite raw moments. Then, the following equality holds:

\[
E\left\{ \left( \sum_{i=1}^{N} X_i^1 \right)^{m_1} \ldots \left( \sum_{i=1}^{N} X_i^T \right)^{m_T} \right\} = \min(M,N) \sum_{p \in \mathcal{P}_{s^*}} \sum_{s^*=1} \left( \begin{array}{c} N \\ s^* \end{array} \right) \sum_{j=1} \prod_{k=1} \mu_{E_{j,k}}
\]

where \( M = \sum_{i=1}^{T} m_i \), \( \mathcal{P}_{s^*} \) is the set of solutions of the integer partition problem for \( M \) into \( s^* \in \{1, \ldots, M\} \) terms. \( \sigma(p) \) is a multinomial coefficient which indicates the number of permutations without repetition of \( p \in \mathcal{P}_{s^*} \), \( E_{j,k} = \sum_{i=1}^{M} e(i) \delta(b_j(i) - k) \) with \( e = [t_1, \ldots, t_1, \ldots, t_T, \ldots, t_T] \), \( \delta \) is the Kronecker delta and \( b_j \in \mathcal{B}_{s^*,p} \) is a permutation without repetition of \( s^* \) elements with multiplicities the elements of \( p \).

The proof can be derived in the same way as was done for Theorem 2.

3. An example: variance of sample variance

To illustrate the results of the previous section with a practical example, we will calculate the variance of the unbiased sample variance:

\[
\text{Var} \{ V \} = E\{ V^2 \} - E\{ V \}^2
\]
with \( V = \frac{1}{N(N-1)} \left(N\sum_{i=1}^{N} X_i^2 - \left(\sum_{i=1}^{N} X_i\right)^2\right) \) and \( N \geq 4 \). In this case we need to calculate just two raw moments, i.e. \( E[V] \) and \( E[V^2] \):

\[
E[V] = \mu_2 - \mu_1^2
\]

\[
E[V^2] = \frac{1}{N^2(N-1)^2} \sum_{k=0}^{2} \binom{2}{k} N^k (-1)^{2-k} E\left\{ \left(\sum_{i=1}^{N} X_i\right)^k \left(\sum_{i=1}^{N} X_i\right)^{2(2-k)} \right\}
\]

\[
= \frac{1}{N^2(N-1)^2} \left[ E\left\{ \left(\sum_{i=1}^{N} X_i\right)^4 \right\} - 2NE\left\{ \left(\sum_{i=1}^{N} X_i\right)^2 \left(\sum_{i=1}^{N} X_i\right) \right\} + N^2 E\left\{ \left(\sum_{i=1}^{N} X_i^2\right)^2 \right\} \right]
\]

The former is straightforward, but note that for the latter moments of products of sums must be calculated. For \( E\left\{ \left(\sum_{i=1}^{N} X_i\right)^4 \right\} \), applying Theorem 2, we have:

| \( s^* \) | \( \mathcal{P}_{s^*} \) | \( B_{(s^*, p)} \), \( |B_{(s^*, p)}| = \binom{N}{p} \) | \( \mathcal{X}(B_{(s^*, p)}) \) | \( \sum_{\alpha \in A_{s^*}} E[\mathcal{X}(\alpha)] \)
|-----------------|-----------------|-----------------|-----------------|-----------------|
| \( s^* = 1 \) | \( \mathcal{P}_1 = \{(4)\} \) | \( B_{(1,(4))} = \{(1,1,1,1)\} \) | \( \mathcal{X}(B_{(1,(4))}) = \{X_4^4\} \) | \( \binom{N}{1} \mu_4 \)
| \( \sigma(P_1) = \{1\} \) | | | | |
| \( s^* = 2 \) | \( \mathcal{P}_2 = \{(3,1)\} \) \((2,2)\) | \( B_{(2,(3,1))} = \{(1,1,1,2)\} \) \( B_{(2,(2,2))} = \{(1,1,2,2)\} \) | \( \mathcal{X}(B_{(2,(3,1))}) = \{X_1^2X_2\} \) \( \mathcal{X}(B_{(2,(2,2))}) = \{X_1^2X_2^2\} \) | \( \binom{N}{2}(8\mu_1\mu_3 + 6\mu_2^2) \)
| \( \sigma(P_2) = \{2,1\} \) | \( |B_{(2,(3,1))}| = 4 \) \( |B_{(2,(2,2))}| = 6 \) | | | |
| \( s^* = 3 \) | \( \mathcal{P}_3 = \{(2,1,1)\} \) | \( B_{(3,(2,1,1))} = \{(1,1,2,3)\} \) | \( \mathcal{X}(B_{(3,(2,1,1))}) = \{X_1^2X_2X_3\} \) | \( \binom{N}{3}(36\mu_1^2\mu_2) \)
| \( \sigma(P_3) = \{3\} \) | | \( |B_{(3,(2,1,1))}| = 12 \) | | |
| \( s^* = 4 \) | \( \mathcal{P}_4 = \{(1,1,1,1)\} \) | \( B_{(4,(1,1,1,1))} = \{(1,2,3,4)\} \) | \( \mathcal{X}(B_{(4,(1,1,1,1))}) = \{X_1X_2X_3X_4\} \) | \( \binom{N}{4}(24\mu_4^3) \)
| \( \sigma(P_4) = \{24\} \) | \( |B_{(4,(1,1,1,1))}| = 24 \) | | | |

\[
E\left\{ \left(\sum_{i=1}^{N} X_i\right)^4 \right\} = \binom{N}{4}(\mu_4^4) + \binom{N}{2}(8\mu_1\mu_3 + 6\mu_2^2) + \binom{N}{3}(36\mu_1^2\mu_2) + \binom{N}{4}(24\mu_4^3)
\]
Now for $E \left\{ \left( \sum_{i=1}^{N} X_i^2 \right) \left( \sum_{i=1}^{N} X_i \right)^2 \right\}$ we have:

| $s^*$ | $\mathcal{P}_{s^*}$ | $B_{(s^*,p_1)}$, $|B_{(s^*,p_1)}| = \binom{M}{p_1}$ | $X(\mathcal{B}_{(s^*,p_1)}) = \{X_1^2 X_i X_1\}$ | $(N) \sum_{\alpha \in A_{s^*}} E(\chi(\alpha))$ |
|-------|------------------|--------------------------------|--------------------------------|--------------------------------|
| $s^* = 1$ | $\{ (3) \}$ | $B_{(1, (3))} = \{(1, 1, 1)\}$ | $X(\mathcal{B}_{(1, (3))}) = \{X_1^2 X_i X_1\}$ | $(N) \mu_4$ |
| $s^* = 2$ | $\{ (2, 1) \}$ | $B_{(2, (2, 1))} = \{(1, 1, 2), (1, 2, 1), (2, 1, 1)\}$ | $X(\mathcal{B}_{(2, (2, 1))}) = \{X_1^2 X_2 X_1, X_1^2 X_2 X_1, X_1^2 X_1 X_1\}$ | $(N) (4\mu_1 \mu_3 + 2\mu_2^2)$ |
| $s^* = 3$ | $\{ (1, 1, 1) \}$ | $B_{(2, (1, 1, 1))} = \{(1, 2, 3), (1, 3, 2), \ldots\}$ | $X(\mathcal{B}_{(2, (1, 1, 1))}) = \{X_1^2 X_2 X_3, X_1^2 X_3 X_2, \ldots\}$ | $(N) (6\mu_1^2 \mu_2)$ |

Then, the variance of the sample variance can be written as

$$\text{Var} \{V\} = \frac{1}{N} \mu_4 - \frac{4}{N} \mu_1 \mu_3 - \frac{(N - 3)}{N(N - 1)} \mu_2^2 + \frac{4(2N - 3)}{N(N - 1)} \mu_1^2 \mu_2 - \frac{2(2N - 3)}{N(N - 1)} \mu_4$$

This result, first derived by Gauss, has been derived recently by quite heuristic algebraic methods \[4,5\] in terms of zero-mean random variables which was also obtained in \[21\]. Quoting \[16\], deriving this result “is already messy enough to warrant thinking very carefully about the algebraic formulation one adopts, and any desire to obtain more general expressions of the same kind focuses the mind greatly on the same issue”. The contribution of Theorem 2 is that it offers a simple methodology for calculating the sample variance moments which is easy to implement even for higher orders avoiding, as much as possible, the combinatorial explosion. Additionally, Theorem 2 corollary offers the desired general expressions.
4. Numerical implementation

There are two important problems that must be solved in order to calculate arbitrary moments of the sample variance. The first one is the partition problem. This is a very common problem in combinatorics, algebra and many algorithms. For computational purposes one is interested in generating the set of all the partitions of an integer or those that satisfy some conditions. In our case, we are interested in generating all the set of partitions, \( \mathcal{P}_{s^*} \). Many algorithms to generate the set of solutions have appeared in the literature \([3,9,8,11,17,22]\), so we can use any of them.

On the other hand, the problem of generating the set \( \mathcal{B}_{s^*,p} \) is a permutation problem with possible repeated terms. There are also many algorithms to solve this problem \([1,7,12,15]\) so we can also choose any of them.

The direct implementation of \( E \left\{ \left( \sum_{i=1}^N X_i^2 \right)^m \left( \sum_{i=1}^N X_i \right)^n \right\} \) is presented in Algorithm 1. Note that following the results presented in Theorems 1 and 2 this implementation is straightforward. The heaviest computational part of the algorithm is the construction of \( \mathcal{B}_{s^*,p} \) since it needs to build all unique combinations of \( s^* \) indices with a particular arrangement of multiplicities \( p \). As an example, for the 5th raw moment of the sample variance, the biggest set \( \mathcal{B}_{s^*,p} \) will have 3, 628, 800 elements and it will cost a considerably amount of time and memory.

Algorithm 1. Algorithm to calculate \( E \left\{ \left( \sum_{i=1}^N X_i^2 \right)^m \left( \sum_{i=1}^N X_i \right)^n \right\} \)

\[
\begin{align*}
m &\leftarrow \text{Power of } \left( \sum_{i=1}^N X_i^2 \right)^m \\
n &\leftarrow \text{Power of } \left( \sum_{i=1}^N X_i \right)^n \\
M &\leftarrow m + n \\
v &\leftarrow \text{Vector of } M \text{ raw moments of } X_i \\
e &\leftarrow [2, \ldots, 2, 1, \ldots, 1] \\
\mathcal{P}_{s^*} &\leftarrow \text{Sets of all the partitions of } M \\
value &\leftarrow 0 \\
\text{for } s^* &\leftarrow 1 \text{ to } \min(M, N) \text{ do} \\
\text{for } p &\in \mathcal{P}_{s^*} \text{ do} \\
\mathcal{B}_{s^*,p} &\leftarrow \text{Sets of permutations without repetition} \\
E_{j,k} &\leftarrow \sum_{i=1}^M e(i) b(j, i) - k) \\
\sigma(p) &\leftarrow \text{Number of unique permutations of } p \\
\text{value} &\leftarrow \text{value} + \binom{N}{s^*} \sigma(p) \sum_{j=1}^M \prod_{k=1}^{s^*} v(E(j, k)) \\
\text{end for} \\
\text{end for} \\
\text{return value}
\end{align*}
\]

In order to reduce the number of elements to calculate, one can realize that the set of all combinations of \( \mathcal{B}_{s^*,p} \) can be decomposed into two different groups of \( m \) and \( n \) elements. The fist one refers to those indices of random variables of power 2: \( X_i^2 \). The second group to those of power 1: \( X_i \). This way, it is clear that every permutation inside these two groups is equivalent in terms of the expectation so, one just need to know the cardinal of all equivalent permutations, which is easily calculated by means of a multinomial coefficient with multiplicities those of indices in each group. So, the following corollary can be established.

Corollary 2. Let \( X_i \) be IID random variables with \( 2m + n \) finite raw moments \( \mu_1, \ldots, \mu_{2m+n} \). Then, the following equality holds:

\[
E \left\{ \left( \sum_{i=1}^N X_i^2 \right)^m \left( \sum_{i=1}^N X_i \right)^n \right\} = \sum_{s^*=1}^{\min(M, N)} \binom{N}{s^*} \sum_{p \in \mathcal{P}_{s^*}} \sigma(p) \sum_{j=1}^m \prod_{k=1}^{s^*} \mu_{E_{j,k}}
\]

where \( M = m + n \), \( \mathcal{P}_{s^*} \) is the set of solutions of the integer partition problem for \( M \) into \( s^* \) terms. \( \sigma(p) \) is a multinomial coefficient which indicates the number of permutations without repetition of \( p \in \mathcal{P}_{s^*} \). \( D_{(s^*,p,m)} \) is the set of
Proof. For each $s^*$ and each $p$ let $D(s^*, p, m)$ be the set of unique combinations of $M$ elements ($s^*$ indices with multiplicities those of $p$) taken $m$ at time with multiplicities $p = (p_1, p_2, \ldots, p_{s^*})$. $\hat{E}_{j,k}$ is defined as:

$$
\hat{E}_{j,k} = 2 \times \pi(d_j)(k) + \pi^*(d_j)(k)
$$

where $\pi(b_j) = (m_1, \ldots, m_{s^*})$ are the multiplicities of each index of $b_j \in D(s^*, p, m)$ and $\pi^*(b_j) = (p_1 - m_1, \ldots, p_{s^*} - m_{s^*})$.

We avoid intentionally the dependence of $\pi$ with respect to $s^*$ and $p$ in order to make notation simpler.

Now, let $d_j^* = D(s^*, p, m)$ be the complement of each element $d_j \in D(s^*, p, m)$, i.e., the indices and the multiplicities of the indices that are not in $d_j \in D(s^*, p, m)$. The multiplicities of $d_j^*$ can be easily calculated by $\pi(d_j^*) = \pi^*(d_j) = (p_1 - m_1, p_2 - m_2, \ldots, p_{s^*} - m_{s^*})$.

$$
\pi^* : D(s^*, p, m) \rightarrow N_{s^*} \quad \begin{array}{c} \mathbf{d} \rightarrow (m_1, \ldots, m_{s^*}) \end{array}
$$

(27)

where

$$
m_i = \sum_{k=1}^{s^*} \delta(d(k) - i), \quad \text{with } i = 1, \ldots, s^*
$$

(28)

From these sets it is easy to calculate the number of equivalent arrangements since all unique permutations of each group belong to the same sort of random variable ($X^2_i$ or $X_i$) and the order of indices does not matter.

There are $\left(\frac{m}{\pi(d_j)}\right) \left(\frac{n}{\pi(d_j^*)}\right)$ equivalent elements for each $p$. And now we can avoid using $e$ since the order of the moments per index becomes

$$
\hat{E}_{j,k} = 2 \times \pi(d_j)(k) + \pi^*(d_j)(k)
$$

(30)

So, finally, for each $s^*$ and $p$ we have:

$$
\sum_{j=1}^{M} \prod_{k=1}^{s^*} \mu E_{j,k} = \sum_{j=1}^{\left|D(s^*, p, m)\right|} \left(\frac{m}{\pi(d_j)}\right) \left(\frac{n}{\pi(d_j^*)}\right) \prod_{k=1}^{s^*} \mu E_{j,k} \quad \square
$$

(31)

The method to implement this reduction of combinations is presented in Algorithm 2.
Algorithm 2. Alternative method for \( E \left\{ (\sum_{i=1}^{N} X_i^2)^m (\sum_{i=1}^{N} X_i)^n \right\} \)

\[
m \leftarrow \text{Power of } (\sum_{i=1}^{N} X_i^2)^m \\
n \leftarrow \text{Power of } (\sum_{i=1}^{N} X_i)^n \\
M \leftarrow m + n \\
v \leftarrow \text{Vector of } M \text{ raw moments of } X_i \\
P_{s=1,\ldots,M} \leftarrow \text{Sets of all the partitions of } M \\
\text{value} \leftarrow 0 \\
\text{for } s^* = 1 \text{ to } \min(M,N) \text{ do} \\
\quad \text{for } p \in P_{s^*} \text{ do} \\
\quad \quad D_{(s^*,p,m)} \leftarrow \text{set of unique combinations of } M \text{ elements taken } m \text{ at time } (s^* \text{ indices with multiplicities those of } p) \\
\quad \quad \sigma(p) \leftarrow \text{Number of unique permutations of } p \\
\quad \quad \text{value} \leftarrow \text{value} + \left( \binom{N}{s^*} \sigma(p) \sum_{j=1}^{\left| D_{(s^*,p,m)} \right|} \left( \frac{m}{\pi(d_j)} \right) \left( \frac{n}{\pi^*(d_j)} \right) \prod_{k=1}^{s^*} v(\hat{E}(j,k)) \right) \\
\quad \text{end for} \\
\text{end for} \\
\text{return value} \\
\]

5. Discussion

In this paper the problem of the calculation of the moments of the sample variance has been studied. Two theorems are proposed that allow calculating in a direct way the moments of the products of sums of different powers of random variables. The main contribution of these theorems is that there is no need of conversion formulae of multiplication tables as the Polykay philosophy. Additionally, if one is interested in conversion formulae or multiplication tables, they can be derived with the methods here proposed in a direct way.

From a practical point of view, two algorithms for numerical implementation are proposed. The first one is a direct implementation of the main result of this work. This algorithm can be easily generalized for more complex products. The second one is a refined method for the case of two different powers of random variables.

The complexity of both algorithms is related to the number of elements in each of the sets \( B_{s^*,p} \) and \( D_{(s^*,p,m)} \). These sets can be obtained from algorithms with constant delay (see Ref. [22]), this means that the time for obtaining each element of the sets from the previous one can be assumed to be constant and included in the time of the operations inside each loop.

So, considering \( N \geq M \), the instructions inside the loops consume a time \( T_{(1,s^*)} \) for each element of the sets \( B_{s^*,p} \) and \( D_{(s^*,p,m)} \). This time can be considered fix since the operations to be done are products in the interval \( (1,s^*) \) which,

| \( j \) | Algorithm 1. max \( |B_{s^*,p}| \) | Algorithm 2. max \( |D_{(s^*,p,m)}| \) |
|---|---|---|
| 1 | 2 | 1 |
| 2 | 24 | 3 |
| 3 | 720 | 6 |
| 4 | 40320 | 15 |
| 5 | 362880 | 35 |
| 6 | 479001600 | 84 |
| 7 | 87178291200 | 210 |
| 8 | 20922789888000 | 495 |
| 9 | 6402373705728000 | 1287 |
| 10 | 2432902008176640000 | 3003 |
| 11 | 1124000727777607680000 | 8008 |
| 12 | 620448401733239439360000 | 19448 |
| 13 | 403291461126605655384000000 | 50388 |
| 14 | 30488344611713680501504000000 | 125970 |
| 15 | 265252859812191058636308480000000 | 319770 |
in the worst case, is \( T_{1(M)} \). So, the time consumed in the loops, \( T_1 \) and \( T_2 \), for the first and the second algorithm respectively are:

\[
T_1 = \sum_{s^*} \sum_{p \in P_{s^*}} \sum_{j=1}^{M} T_{1(s^*)} \leq T_{1(M)} \sum_{s^*} \max_p |B_{s^*,p}| |P_{s^*}|
\]

\[
T_2 = \sum_{s^*} \sum_{p \in P_{s^*}} \sum_{j=1}^{M} T_{1(s^*)} \leq T_{1(M)} \sum_{s^*} \max_p |D_{(s^*,p,m)}| |P_{s^*}|
\]

(32)

(33)

Table 3 shows the maximum number of combinations for each algorithm when \( N \geq M \), to calculate the \( j \)th raw moment of the sample variance, \( \langle E\{V^j\} \rangle \), up to order 15. The higher performance of this implementation becomes evident.

Acknowledgements

The authors acknowledge: Gerencia Regional de Salud, Consejería de Sanidad, Junta de Castilla y Leon for grants GRS 474/A/10; Consejería de Educación, Junta de Castilla y Leon for grant VA376A11-2; Ministerio de Ciencia e Innovacion for grants CEN-20091044, TEC2010-17982 and MTM2007-63257; Proyectos de investigacion en salud, Accion Estrategica en Salud, Instituto de Salud Carlos III for grant PI11-01492 and Centro para el Desarrollo Tecnico Industrial, Ministerio de Ciencia e Innovacion, CEN-20091044 (cvREMOD).

References


