ECONTECHMOD. AN INTERNATIONAL QUARTERLY JOURNAL – 2017, Vol. 06, No. 4, 93–96

Note on the moments of random variables product

A. Kornacki

Department of Applied Mathematics and Computer Science
University of Life Sciences in Lublin
Akademicka 12, 20-950 Lublin Poland, e-mail:Andrzej.Kornacki@up.lublin.pl

Received: May 09. 2017; Accepted: November 15. 2017

Abstract. In this paper the formulae for central moments of independent random variables product are introduced. Generally, the formulae are very complex and not very readable, therefore the article focuses on the most important – terms of applications – moments of $r = 2, 3, 4$ orders. These moments occur in the determination of such characteristics as variance, skewness or kurtosis. In cases $r = 3$ and 4 only two random variables are considered. Apart from exact formulae in the considered situations the approximate formulae were also presented. For the variance the approximation effectiveness was also assessed.

Key words: central moment of random variable, coefficient of skewness, kurtosis, variation coefficient, approximate formula.

INTRODUCTION

In a variety of natural sciences the quantities are encountered which are products of other quantities (e.g. physics, economics, finances). In statistics [1], the equivalent values are estimators that are products of other estimators. Thus a question arises how to express the moments of product using the moments of factors. In this paper such formulae (theorem 4) are presented. In terms of practical applications of particular importance are the moments of orders $r = 2, 3, 4$. They are the starting point while calculating the variance, skewness or kurtosis. For such situations both exact and approximate formulae were provided in the article. The results for $r = 3$ and 4 are generalizations for variance formulae provided by Goodman [2, 3].

MATERIAL AND METHODS

I Notations and basic lemmas

The following notations are used in the study: $X_1, X_2, ..., X_K$ denote independent random variables with non-zero expected values, α_r^i - r-th ordinary moment of random variable X_i, μ_r^i - r-th central moment of random variable X_i, $H_r^i = \frac{\mu_r^i}{[\alpha_1^i]^r}$, $Z_i = X_i - \alpha_1^i$, $z_i = \frac{Z_i}{\alpha_1^i}$.

Thus, $\frac{X_i}{\alpha_1^i} z_i + 1$, and R denote variation coefficient. Moreover, the operators of taking of the he expected value, variance, variation coefficient and of the r-th central moment of the random variable are denoted as $E()$, $V()$, $R()$ and $\mu_r()$, respectively.

The calculation of the r-th central moment of random variables is based on the expression:

$$\prod_{i=1}^{K} X_i - E(\prod_{i=1}^{K} X_i).$$

Its form is described by lemma 1.

Lemma 1. If X_i are independent random variables ($i = 1, 2, \ldots, K$), then:

$$\prod_{i=1}^{K} X_i - E(\prod_{i=1}^{K} X_i) = \prod_{i=1}^{K} \alpha_1^i [\sum_{i=1}^{K} z_i + \sum_{i_1<i_2} z_{i_1} z_{i_2} + \ldots + z_1 z_2 z \ldots z_n]. \quad (1)$$

Proof: Due to the independence of random variables X_i and adopted notations, we have, respectively:

$$\prod_{i=1}^{K} X_i - E(\prod_{i=1}^{K} X_i) = \prod_{i=1}^{K} X_i - \prod_{i=1}^{K} E(X_i) = \prod_{i=1}^{K} E(X_i) \left[\frac{\prod_{i=1}^{K} X_i}{\prod_{i=1}^{K} E(X_i)} - 1 \right] =$$

$$= \prod_{i=1}^{K} E(X_i) \left[\prod_{i=1}^{K} \frac{X_i}{E(X_i)} - 1 \right] = \prod_{i=1}^{K} \alpha_1^i \left[\prod_{i=1}^{K} (z_i + 1) - 1 \right] =$$

$$= \prod_{i=1}^{K} \alpha_1^i [\sum_{i=1}^{K} z_i + \sum_{i_1<i_2} z_{i_1} z_{i_2} + \ldots + z_1 z_2 z \ldots z_n].$$

This ends the proof of the lemma 1.

While formulating the basic result we will use the polynomial formula, which will be repeated now in the form of next lemma.

Lemma 2 Rao [4]:

$$(x_1+x_2+\dots+x_s)^r = \sum_{\substack{k_1,k_2,\dots k_s \\ \sum_{i=1}^{s} k_i = r}} \binom{r}{k_1,k_2,\dots,k_s} x_1^{k_1} x_2^{k_2} \cdots x_s^{k_s}, \qquad (2)$$

where polynomial factors have a form:

$$\binom{r}{k_1,k_2,\dots,k_s} = \frac{r!}{k_1!k_2!\cdots k_s!}\,. \qquad (3)$$

Moreover, we will use the properties of variables z_i, Z_i, and H_i. They are formulated by lemma 3.

Lemma 3. Z_i are independent for i=1,2,…,K; z_i are independent for i=1,2,…,K; H_r^i are independent for i=1,2,…,K. Moreover, we have:

$$E(z_i)=E(Z_i)=0, \quad E(z_i^2)=H_2^i, \quad E(z_i^3)=H_3^i. \qquad (4)$$

Proof**:** The truth of the lemma results from the independence of variables X_i and definitions variables z_i, Z_i and H_r^i.

RESULTS

I Main result

Now we will present the basic result of central moments of the product of independent random variables.

Theorem 1. The central moments of the product of independent random variables take the following form:

$$\mu_r(\prod_{i=1}^{K} X_i) = \prod_{i=1}^{K} (\overset{i}{\alpha}_1)^r E[\sum_{i=1}^{K} z_i + \sum_{i_1<i_2} z_{i_1} z_{i_2} + \dots + z_1 z_2 z \dots z_n]^r =$$

$$=[\prod_{i=1}^{K} (\overset{i}{\alpha}_1) \sum_{\substack{k_1,k_2,\dots k_s \\ \sum_{i=1}^{s} k_i = r}} \binom{r}{k_1,k_2,\dots,k_s} (\sum_{i=1}^{K} z_i)^{k_1} (\sum_{i_1<i_2} z_{i_1} z_{i_2})^{k_2} \dots (z_1 z_2 \dots z_K)^{k_s}]. \qquad (5)$$

In our case we have $s = 2^K$-1.

Proof**:** Formula (5) is obtained by direct application to the definition of the central moment of the r order of lemmas 1 and 2.

Due to the fact that in the general case theorem 1 has a complex and not very readable form, in the further part of the paper we will focus on special cases r = 2, 3, 4 that are important in practice.

II Case r = 2 (Variance)

For $r = 2$ the central moment of the random variable becomes its variance. In such a case, the following result is true [2, 3].

Theorem 2. The variance of the product of independent random variables X_i, (i = 1, …, K), equals:

$$V(\prod_{i=1}^{K} X_i) = \prod_{i=1}^{K} (\overset{i}{\alpha}_1)^2 [\sum_{i=1}^{K} H_2^i + \sum_{i_1<i_2} H_2^{i_1} H_2^{i_2} + \dots + H_2^1 H_2^2 \cdots H_2^K]. \qquad (6)$$

Proof: Due to the independence of random variables X_i we have:

$$E(\prod_{i=1}^{K} X_i^2) = \prod_{i=1}^{K} E(X_i^2)$$

and from formula $V(X)=EX^2$-(EX) we get:

$$V(\prod_{i=1}^{K} X_i) = E(\prod_{i=1}^{K} X_i)^2 - [E(\prod_{i=1}^{K} X_i)]^2 = \prod_{i=1}^{K} E(X_i^2) - [E(\prod_{i=1}^{K} X_i)]^2 =$$

$$= \prod_{i=1}^{K} [V(X_i) + (E(X_i))^2] - [E(\prod_{i=1}^{K} (X_i)]^2 =$$

$$= \prod_{i=1}^{K} \{(E(X_i)^2 [\frac{V(X_i)}{(E(X_i)^2} + 1]\} - [\prod_{i=1}^{K} E(X_i)]^2 =$$

$$= \prod_{i=1}^{K} \{(E(X_i)^2 [H_2^i + 1]\} - \prod_{i=1}^{K} (E(X_i))^2 = \prod_{i=1}^{K} \{(E(X_i)^2 \{\prod_{i=1}^{K} (H_2^i + 1]\} - 1\} =$$

$$= \prod_{i=1}^{K} [\overset{i}{\alpha}_1]^2 [\sum_{i=1}^{K} H_2^i + \sum_{i_1<i_2} H_2^{i_1} H_2^{i_2} + \dots + H_2^1 H_2^2 \cdots H_2^K].$$

This ends the proof of the theorem 2.

It should be noted that H_2^i is equal to the square of the coefficient of variation of random variable X_i. From theorem 2 we get the approximation formula for the variance of the product of random variables in the following form:

$$\tilde{V}(\prod_{i=1}^{K} X_i) = \prod_{i=1}^{K} [\overset{i}{\alpha}_1]^2 (\sum_{i=1}^{K} H_2^i). \qquad (7)$$

Formula (7) is a generalisation of the approximation formula for the variance of product of independent random variables from two for K variables. As one can see the variance of the product calculated from the approximation formula (7) is smaller than the variance for the exact formula (6). The approximation will be the better the greater degree of negligence of the summand applied:

$$\sum_{i_1<i_2} H_2^{i_1} H_2^{i_2} + \dots + H_2^1 H_2^2 \cdots H_2^K.$$

Let us denote the effectiveness factor of the "approximation" by:

$$q_1 = \frac{\tilde{V}(\prod_{i=1}^{K} X_i)}{V(\prod_{i=1}^{K} X_i)}. \qquad (8)$$

The following theorem allows us to express the effectiveness of approximation of the exact formula (6) by approximation (7) by coefficients of variation of variables X_i.

Theorem 3. If $H_2^i = t$, $(i = 1, 2, \ldots, K)$, then:

$$q_1 = \frac{K}{K + \binom{K}{2} t + \cdots + \binom{K}{K} t^{K-1}}. \tag{9}$$

In the light of formulae (6), (7), (8) and assumptions we have:

$$q_1 = \frac{\tilde{V}(\prod_{i=1}^{K} X_i)}{V(\prod_{i=1}^{K} X_i)} = \frac{\prod_{i=1}^{K} [\alpha_1^i]^2 (\sum_{i=1}^{K} H_2^i)}{\prod_{i=1}^{K} (\alpha_1^i)^2 [\sum_{i=1}^{K} H_2^i + \sum_{i_1<i_2} H_2^{i_1} H_2^{i_2} + \ldots + H_2^1 H_2^2 \cdots H_2^K]} =$$

$$= \frac{Kt}{Kt + \binom{K}{2} t^2 + \cdots + \binom{K}{K} t^K} = \frac{K}{K + \binom{K}{2} t + \cdots + \binom{K}{K} t^{K-1}}.$$

This ends the proof of the theorem 3.

Putting K=3, we obtain efficiency coefficient of the form:

$$q_1 = \frac{3}{3 + 3t + t^2}. \tag{10}$$

Example. Putting q_1=0,95 we have $\frac{3}{3+3t+t^2} = 0{,}95$ due to (10). Thus, $t = 0{,}05$. Due to the equality $H_2^i = R_i^2$ we obtain: $R_i = 0{,}23$. Therefore, if the coefficients of variation of variables X_i (for 3 variables) are equal to 0,23 then the effectiveness of "approximation" of exact variance (6) using the approximation formula (7) amounts to 95%.

III Case. *r* = 3 (*K*=2)

We are going to present now the formula for the third order central moment of the product of independent random variables. For the sake of simplicity, we will only consider the case of two random variables. Theorem 4 presents the basic result.

Theorem 4. The third central moment of the product of two independent random variables is:

$$\mu_3(X_1X_2) = (\alpha_1^1)^3(\alpha_1^2)^3[H_3^1 + H_3^2 + H_3^1H_3^2 + 3H_3^1H_2^2 + 3H_3^2H_2^1 + 6H_2^1H_2^2]. \tag{11}$$

Proof: Using lemmas 1 and 2, we get:

$$\mu_3(X_1X_2) = (\alpha_1^1)^3(\alpha_1^2)^3[Z_1^3 + Z_2^3 + Z_1^3Z_2^3 + 3Z_1^2 z_2 + Z_1^3 z_2 + 3Z_2^2 z_1 + 3z_1 Z_2^3 + 3Z_1^3 Z_2^2 + 3Z_2^3 Z_1^2 + 6Z_1^2 Z_2^2].$$

After the application of lemma 3, formula takes the form (11). This ends the proof of the theorem 4.

From theorem 4 obtain the approximation formula of the following form:

$$\tilde{\mu}_3(X_1X_2) = (\alpha_1^1)^3(\alpha_1^2)[H_3^1 + H_3^2 + H_3^1H_3^2]. \tag{12}$$

By analogy to formula (9), let us determine the effectiveness factor of "approximation" for the third central moment by:

$$q_2 = \frac{\tilde{\mu}_3(X_1X_2)}{\mu_3(X_1X_2)}. \tag{13}$$

This factor has a property described by theorem 5.

Theorem 5.

If $H_3^1 = H_3^2 = H_2^1 = H_2^1 = t$, then:

$$q_2 = \frac{2+t}{2+13t}. \tag{14}$$

Proof: Due to formulae (11), (12) and (13) as well as the assumption we have subsequently:

$$q_2 = \frac{(\alpha_1^1)^3(\alpha_1^2)[H_3^1 + H_3^2 + H_3^1H_3^2]}{(\alpha_1^1)^3(\alpha_1^2)^3[H_3^1 + H_3^2 + H_3^1H_3^2 + 3H_3^1H_2^2 + 3H_3^2H_2^1 + 6H_2^1H_2^2]} =$$

$$= \frac{2t + t^2}{2t + 13t^2} = \frac{2+t}{2+13t}.$$

This ends the proof of the theorem 5.

IV Case. *r* = 4 (*K*=2)

The fourth central moment of two independent random variables is described by theorem 6.

Theorem 6.

The fourth central moment of two independent random variables is equal:

$$\begin{cases} \mu_4(X_1X_2) = (\alpha_1^1)^4(\alpha_1^2)^4[H_4^1 + H_4^2 + 7H_4^1H_4^2 + 4H_4^1H_3^2 + 4H_3^1H_4^2 + 6H_8^1H_4^2 + \\ + 6H_8^2H_4^1 + 12H_3^1H_2^2 + 12H_3^2H_2^1 + 12H_3^1H_3^2] \end{cases} \tag{15}$$

We obtain formula (15) using the definition of the central moment and using lemmas 1, 2 and 3.

Due to a very complex form of formula (15), it seems natural that only the approximate formula should be used which in this case takes the form:

$$\tilde{\mu}_4 = (\alpha_1^1)^4(\alpha_1^2)^4[H_4^1 + H_4^2 + 7H_4^1H_4^2] \cdot \tag{16}$$

CONCLUSIONS

1. In many areas of science there are quantities that are products of other quantities. (e.g. physics, economics, engineering, finance, etc.). In this paper, central moments corresponding to these magnitudes of random variables are investigated. Central moments of the order $r = 2, 3, 4$ are the basis for the definition of statistically significant characteristics such as variance, asymmetry coefficient, and kurtosis.

2. In this paper, graceful formulas are given for the variance of the product of independent random variables. They are expressed through the squares of the coefficients of variation of the individual factors. This allows us to evaluate accuracy when approximating exact formulas with the approximation formula.

3. Due to the complex and unreadable form, it seems that for moments of order $r = 3, 4$ makes sense to use only approximation formulas.

REFERENCES

1. **Fisz M. 1967.** Probability and mathematical statistics. Warszawa PWN. (in Polish).
2. **Goodman L.A. 1960.** On the exact variance of products. Journal of American Statistical Association, 55. 708-713.
3. **Goodman L.A. 1962.** The variance of product of random variables. Journal of American Statistical Association, 57. 54-60.
4. **Rao C.R. 2002.** Linear Statistical Inference and Its Applications. John Wiley & Sons New York.
5. **Yates F. 1953.** Sampling methods for censuses and surveys. Second edition. London Charles Griffin and Co .Ltd.