



**FACULTAD DE CIENCIAS EMPRESARIALES Y ECONOMIA**

**Serie de documentos de trabajo del Departamento de Economía /  
Department of Economics Working Papers Series**

## **The variance upper bound for a mixed random variable**

Martín Egozcue Universidad de Montevideo, Uruguay

Luis Fuentes García  
Universidad de Coruña, Spain.

The working papers of the Department of Economics, Universidad de Montevideo are circulated for discussion and comment purposes. They have not been peer reviewed nor been subject to the review by the University's staff. © 2017 by Martín Egozcue and Luis Fuentes García.

All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

# The variance upper bound for a mixed random variable

**Martín Egozcue**

*Universidad de Montevideo, Montevideo 11600, Uruguay*

**Luis Fuentes García**

*Departamento de Métodos Matemáticos e de Representación, Escola Técnica Superior de Enxeñeiros de Camiños, Canais e Portos, Universidade da Coruña, 15071 A Coruña, Spain.*

**Abstract** In this note, we derive upper bounds on the variance of a mixed random variable. Our results are an extension of previous results for unimodal and symmetric random variables. The novelty of our findings is that this mixed random variable does not necessary need to be symmetric and is multimodal. We also characterize the cases when these bounds are optimal.

*Keywords and phrases:* Gruss' inequality; Popoviciu's inequality; unimodal; symmetry

# 1 Introduction

It is well known that the sharp upper bound on the variance of any random variable  $X$  with support on  $[a, b]$  is equal to

$$\mathbf{V} [X] \leq \frac{(b - a)^2}{4}. \quad (1)$$

This inequality is usually known as Popoviciu's inequality (Popoviciu, 1935), but is also a special case of Gruss' inequality (Gruss, 1935) that states that if two random variables  $X$  and  $Y$  have finite support on  $[a, b]$  and  $[c, d]$  respectively, then

$$\mathbf{Cov} [X, Y] \leq \frac{(b - a)(d - c)}{4}.$$

Hereafter, we refer inequality (1) as Popoviciu-Gruss' inequality.

If we have additional information of the random variable such as its mean  $\mu$ , then the upper bound of the variance is  $(b - \mu)(\mu - a)$ , (cf., Muilwijk, 1966). See also Bathia and Davis (2000) for further improvements of this type of bound.

One can improve these bounds with further information of  $X$ . For instance, Gray and Odell (1967) show that if  $X$  is symmetric and unimodal with support on  $[a, b]$  then

$$\mathbf{V} [X] \leq \frac{(b - a)^2}{12}. \quad (2)$$

Jacobson (1969) shows that bound (2) increases to  $(b - a)^2/9$  if the symmetry assumption is relaxed. Different variations of these bounds have been explored by Seaman, Young and Turner (1987), Abouammoh et al. (1994), Ageel (2000), Liu and Li (2014), among others.

In addition, Dharmadhikari and Joag-Dev (1989) establish the following bounds when the continuous random variable is unimodal with mode  $m$ , mean  $\mu$  and support on  $[0, 1]$ :

$$\begin{aligned} \text{i) } \mathbf{V} [X] &\leq \frac{2\mu(1+m)-3\mu^2-m}{3}, \text{ and} \\ \text{ii) } \mathbf{V} [X] &\leq \frac{1-m(1-m)}{9}. \end{aligned} \quad (3)$$

Also, Sharma and Bhandari (2014) find the following upper bound of the variance for continuous random variables with unique mode  $m$ , mean  $\mu$  and support on  $[a, b]$

$$\mathbf{V} [X] \leq (\beta - \mu)(\mu - \alpha), \quad (4)$$

where

$$\alpha = \frac{(a + b + m) - \sqrt{a^2 + b^2 + m^2 - (am + bm + ab)}}{3},$$

and

$$\beta = \frac{(a + b + m) + \sqrt{a^2 + b^2 + m^2 - (am + bm + ab)}}{3}.$$

In this note, we derive new upper bounds of the variance considering a random variable of a mixture of  $n$  random variables. This type of distribution is commonly used in applied mathematics (cf. Kaas et al., 2008). To clarify the motivation of this work we present the next example.

**Example 1** *Willassen (1981) derives best bounds on a expected utility functional. Consider  $u : [a, b] \rightarrow R$  to be an increasing concave function on  $[0, \infty)$  with  $u(0) = 0$  and  $u''' > 0$  and  $X$  a random variable with mean and variance  $\mu$  and  $\sigma^2$ , respectively. Then the greatest lower bound of  $\mathbf{E}[u(X)]$  is:*

$$\frac{\mu^2}{\mu^2 + \sigma^2} u\left(\mu + \frac{\sigma^2}{\mu}\right). \quad (5)$$

*Suppose that  $X$  is a mixture of two random variables  $X_1$  and  $X_2$  with support on  $[a_i, b_i]$  for  $i = 1, 2$  and we know some characteristics of each  $X_i$  such as their symmetry, modality or even the mean, but not, for instance, their variance. To use bound (5) we need to estimate the variance of  $X$  and our findings would help to estimate an upper bound of such variance.*

The novelty of our results is that this mixed random variable is multimodal and does not necessary has to be symmetric. The optimality of these bounds is also discussed.

## 2 Main results

Let  $X_i$  be a family of random variables with  $i = 1, 2, \dots, n$ . Consider a mixture variable  $Z$ , such that  $\mathbf{P}[Z = X_i] = p_i$  with  $p_i \geq 0$  and  $\sum_{i=1}^n p_i = 1$ . It is well known that:

$$\mathbf{E}[Z^k] = \sum_{i=1}^n p_i \mathbf{E}[X_i^k].$$

Therefore,

$$\begin{aligned}
\mathbf{V}[Z] &= \sum_{i=1}^n p_i \mathbf{E}[X_i^2] - \left( \sum_{i=1}^n p_i \mathbf{E}[X_i] \right)^2, \\
&= \sum_{i=1}^n p_i (\mathbf{V}[X_i] + \mathbf{E}[X_i]^2) - \left( \sum_{i=1}^n p_i \mathbf{E}[X_i] \right)^2, \\
&= \sum_{i=1}^n p_i \mathbf{V}[X_i] + \sum_{i=1}^n p_i \mathbf{E}[X_i]^2 - \left( \sum_{i=1}^n p_i \mathbf{E}[X_i] \right)^2.
\end{aligned} \tag{6}$$

The next lemma plays a fundamental role in our main findings.

**Lemma 1** *Under the previous assumptions:*

$$\sum_{i=1}^n p_i \mathbf{E}[X_i]^2 - \left( \sum_{i=1}^n p_i \mathbf{E}[X_i] \right)^2 \leq \frac{1}{4} (\max \mathbf{E}[X_i] - \min \mathbf{E}[X_j])^2.$$

Moreover, the bound is optimal.

**Proof.** Consider a discrete random variable  $Y$  with  $\mathbf{P}[Y = \mathbf{E}[X_i]] = p_i$  for  $i = 1, 2, \dots, n$ . Then,

$$\mathbf{V}[Y] = \sum_{i=1}^n p_i \mathbf{E}[X_i]^2 - \left( \sum_{i=1}^n p_i \mathbf{E}[X_i] \right)^2.$$

This random variable has support on  $[\min \mathbf{E}[X_j], \max \mathbf{E}[X_i]]$ . Now, applying Popoviciu-Gruss' inequality we conclude that:

$$\mathbf{V}[Y] \leq \frac{1}{4} (\max \mathbf{E}[X_i] - \min \mathbf{E}[X_j])^2. \tag{7}$$

The optimality of Popoviciu-Gruss bound implies the optimality of this bound. In particular, the bound is sharp for the special weights  $p_{i_0} = p_{j_0} = \frac{1}{2}$ , where  $\mathbf{E}[X_{i_0}] = \max \mathbf{E}[X_i]$  and  $\mathbf{E}[X_{j_0}] = \min \mathbf{E}[X_j]$ . ■

The next result works for a mixture of random variables all with the same variance.

**Theorem 1** *Let  $Z$  be a random variable that consist of a mixture of the random variables  $X_1, \dots, X_n$  all with the same variance equal to  $\sigma^2$ . Then*

$$\mathbf{V}[Z] \leq \sigma^2 + \frac{1}{4} (\max \mathbf{E}[X_i] - \min \mathbf{E}[X_j])^2,$$

and the bound is optimal.

**Proof.** Combining the assumption that  $\sum_{i=1}^n p_i = 1$  and  $\mathbf{V}[X_i] = \sigma^2$  for each  $i = 1, 2, \dots, n$  then the expression (6) becomes:

$$\mathbf{V}[Z] = \sigma^2 + \sum_{i=1}^n p_i \mathbf{E}[X_i]^2 - \left( \sum_{i=1}^n p_i \mathbf{E}[X_i] \right)^2.$$

The final assertion follows immediately after applying Lemma 1. ■

As a corollary of the previous theorem, we consider a mixture of symmetric unimodal random variables with the same variance.

**Corollary 1** *Let  $X_i$  be unimodal random variables with the same variance  $\sigma^2$  and symmetric on  $[a_i, b_i]$  for  $i = 1, 2, \dots, n$ . Let  $Z$  be a random variable mixture of  $X_1, \dots, X_n$ . Then*

$$\mathbf{V}[Z] \leq \frac{1}{12} \min \{(b_i - a_i)^2\} + \frac{1}{16} (\max \{a_i + b_i\} - \min \{a_j + b_j\})^2,$$

and the bound is optimal.

**Proof.** Since, each  $X_i$  is symmetric on  $[a_i, b_i]$  then

$$\mathbf{E}[X_i] = \frac{b_i + a_i}{2}. \quad (8)$$

Using inequality (2) we have an optimal upper bound for each variance

$$\mathbf{V}[X_i] = \sigma^2 \leq \frac{(b_i - a_i)^2}{12}. \quad (9)$$

Now, the result follows immediately from Theorem 1. ■

In the next theorem, we consider a special case of Corollary 1 relaxing the equal variance assumption, but assuming that the length of the supports of each  $X_i$  are equal.

**Theorem 2** *Let  $X_i$  be unimodal random variables symmetric on  $[a_i, b_i]$  for  $i = 1, 2, \dots, n$ . Let  $Z$  be a random variable mixture of  $X_1, \dots, X_n$ . Then*

$$\mathbf{V}[Z] \leq \frac{1}{12} \max \{(b_i - a_i)^2\} + \frac{1}{16} (\max \{a_i + b_i\} - \min \{a_j + b_j\})^2. \quad (10)$$

If  $b_i - a_i = k$  for all  $i = 1, 2, \dots, n$  then the bound is optimal.

**Proof.** It follows immediately applying Lemma 1 and inequality (2), but now observing that since the variance of  $X_i$  are not necessary equal, the bound on the first term must be replaced by the maximum. ■

One can derive new results for mixed random variables with different bounds on  $\mathbf{V}[X_i]$ . For instance, exploiting the bounds (3) we establish the following results:

**Theorem 3** Let  $X_i$  be unimodal random variables with mode  $m_i$ , mean  $\mu_i$  and support on  $[0, 1]$  for  $i = 1, 2, \dots, n$ . Let  $Z$  be a random variable mixture of  $X_1, \dots, X_n$ . Then:

1.  $\mathbf{V}[Z] \leq \max\{(2\mu_i(1+m_i) - 3\mu_i^2 - m_i)/3\} + \frac{1}{4}(\max \mathbf{E}[X_i] - \min \mathbf{E}[X_j])^2$ ,
2.  $\mathbf{V}[Z] \leq \max\{(1 - m_i(1 - m_i))/9\} \sum p_i + \frac{1}{4}(\max \mathbf{E}[X_i] - \min \mathbf{E}[X_j])^2$ .

**Proof.** The result follows by combining equation (6) and bound (3). ■

Finally, one can use the results in Sharma and Bhandari (2014) to find an upper bound for a mixed random variable as we do next.

**Theorem 4** Let  $X_i$  be continuous random variables with mode  $m_i$ , mean  $\mu_i$  and support on  $[a_i, b_i]$  for  $i = 1, 2, \dots, n$ . Let  $Z$  be a random variable mixture of  $X_1, \dots, X_n$ . Then:

$$\mathbf{V}[Z] \leq \frac{1}{12} \max\{(\mu_i - \alpha_i)(\beta_i - \mu_i)\} + \frac{1}{4}(\max(\mu_i) - \min(\mu_j))^2,$$

where

$$\alpha_i = \frac{(a_i + b_i + m_i) - \sqrt{a_i^2 + b_i^2 + m_i^2 - (a_i m_i + b_i m_i + a_i b_i)}}{3}$$

and

$$\beta_i = \frac{(a_i + b_i + m_i) + \sqrt{a_i^2 + b_i^2 + m_i^2 - (a_i m_i + b_i m_i + a_i b_i)}}{3}.$$

**Proof.** The proof follows from applying equation (6) and bounds (4). ■

### 3 Concluding remarks

In this work, we establish upper bounds on the variance of a mixture of random variables. We extend previous findings relaxing the assumptions of unimodality and symmetry of the underlying random variable. We finally show the conditions under which these bounds are optimal.

**Acknowledgments.** This research has been partly supported by grants from the Agencia Nacional de Innovación e Investigación (ANII), Uruguay. Also, it has been partially supported by the Ministerio de Economía y Competitividad (grant DPI2015-68431-R) of the Spanish Government and by the Consellería de Cultura, Educación e Ordenación Universitaria of the Xunta de Galicia (grant GRC2014/039 ) cofinanced with FEDER funds, and the Universidade da Coruña.

## References

- Abouammoh, A. M and A.F. Mashhour (1994). Variance upper bounds and convolutions of  $\alpha$ -unimodal distributions. *Statistics & Probability Letters*, 21(4), 281-289.
- Ageel, M. (2000). Variance upper bounds and a probability inequality for discrete  $\alpha$ -unimodality. *Applicationes Mathematicae* 27.4 (2000): 403-410
- Bhatia, R., and C. Davis (2000). A Better Bound on the Variance. *The American Mathematical Monthly*, 107(4), 353-357.
- Dharmadhikari, S. W. and K. Joag-Dev (1989). Upper bounds for the variances of certain random variables. *Communications in Statistics-Theory and Methods*, 18(9), 3235-3247.
- Gray, H. L. and P.L. Odell (1967). On least favorable density functions. *SIAM Review*, 9(4), 715-720.
- Grüss, G. (1935). Über das Maximum des absoluten Betrages von. *Mathematische Zeitschrift*, 39(1), 215-226.
- Jacobson, H. I. (1969). The maximum variance of restricted unimodal distributions. *The Annals of Mathematical Statistics*, 1746-1752.
- Kaas, R., M. Goovaerts, J. Dhaene and M. Denuit (2008). *Modern Actuarial Risk Theory: Using R (Vol. 128)*. Springer Science & Business Media.
- Liu, G. and W.V. Li (2014). Variance Bounds for Functions of Unimodal Random Variable. *Communications in Statistics-Theory and Methods*.<http://dx.doi.org/10.1080/03610926.2012.718842>
- Muilwijk, J. (1966). Note on a theorem of MN Murthy and VK Sethi. *Sankhya Ser. B*, 28(1), 2.
- Popoviciu, T. (1935). Sur les équations algébriques ayant toutes leurs racines réelles. *Mathematica*, 9, 129-145.
- Seaman, J. W., D.M. Young and D.W. Turner (1987). On the variance of certain bounded



random variables. *Math. Sci*, 12, 109-116.

Sharma, R. and R. Bhandari (2014). On Variance Upper Bounds for Unimodal Distributions. *Communications in Statistics-Theory and Methods*, 43(21), 4503-4513.

Willassen, Y. (1981). Expected Utility, Chebichev Bounds, Mean-Variance Analysis. *The Scandinavian Journal of Economics*, 419-428.