C. R. Math. Rep. Acad. Sci. Canada Vol. **42** (2) 2020, pp. 25–28

On an Extension of Stein's Lemma

Christian Genest, FRSC

Abstract. An extension of Stein's lemma to arbitrary random pairs was recently proposed in [2]. It is shown that this generalization is valid only for linear functions and hence trivial.

Résumé. Une généralisation du lemme de Stein à toute paire d'aléas a récemment été proposée en [2]. On montre que cette extension n'est valable que pour les fonctions affines et qu'elle est donc triviale.

1. Introduction In its simplest form, Stein's lemma [7] states that if a pair (X, Y) of random variables is jointly normal and $f : \mathbb{R} \to \mathbb{R}$ is any differentiable function with derivative f' such that $\mathrm{E}\,|f'(X)| < \infty$, then

$$\operatorname{cov}\{f(X), Y\} = \mathrm{E}\{f'(X)\} \operatorname{cov}(X, Y),$$

provided that all necessary moments exist.

This result is useful for deriving covariances between functions of component random variables in various contexts and has implications, e.g., in portfolio analysis. It has been extended in various ways, e.g., to random vectors of arbitrary dimension whose joint distribution is elliptical [5] and even skew-elliptical [1].

In a recent paper published in this journal [2], an extension of Stein's lemma was proposed which is claimed to be valid for any differentiable function f satisfying the above condition and any random pair (X, Y), irrespective of the relation between its components. The author claims that there exists a constant $\gamma \in \mathbb{R}$ depending only on f and the distribution of X such that

$$\operatorname{cov}\{f(X), Y\} = \gamma \operatorname{cov}(X, Y). \tag{1}$$

The author adds that one can take $\gamma \neq 0$ without loss of generality.

The purpose of this note is to point out that this result is incorrect and to emphasize that an identity of the form (1) cannot possibly hold unless the nature of the dependence between the variables X and Y is taken into account. These two objectives are achieved in Sections 2 and 3, respectively.

Received by the editors on June 15, 2020.
AMS Subject Classification: Primary: 62E10; secondary: 62H05.
Keywords: Covariance, Stein's lemma.

2. A contradiction Suppose that identity (1) holds for a given constant γ which depends only on a given random variable X and an arbitrary differentiable function f. If $\mathrm{var}(X) = 0$, both X and $f(X)$ are almost surely constant and relation (1) is true but trivial because one then has

$$\mathrm{cov}\{f(X), Y\} = \mathrm{cov}(X, Y) = 0.$$

Therefore, assume henceforth that $\mathrm{var}(X) > 0$. Upon taking $Y = X$ almost surely as a special case, one deduces that

$$\mathrm{cov}\{f(X), X\} = \gamma\,\mathrm{cov}(X, X) = \gamma\,\mathrm{var}(X) \tag{2}$$

while by taking $Y = f(X)$ almost surely, one finds

$$\mathrm{cov}\{f(X), f(X)\} = \gamma\,\mathrm{cov}\{X, f(X)\}$$

and hence also, as a consequence of (2),

$$\mathrm{var}\{f(X)\} = \gamma^2\,\mathrm{var}(X). \tag{3}$$

Relations (2) and (3) are both mentioned in [2]. Observe that if $\gamma = 0$, then $f(X)$ is almost surely constant even though X isn't, and hence relation (1) is again true but trivial.

When $\mathrm{var}(X) > 0$ and $\gamma \neq 0$, then $\mathrm{var}\{f(X)\} > 0$ and it immediately follows from relations (2) and (3) that

$$\mathrm{corr}\{f(X), X\} = \frac{\mathrm{cov}\{f(X), X\}}{\sqrt{\mathrm{var}(X)\,\mathrm{var}\{f(X)\}}} = \frac{\gamma\,\mathrm{var}(X)}{|\gamma|\,\mathrm{var}(X)} = \pm 1.$$

As it well known, the absolute value of the correlation between two random variables can only be 1 if these random variables are almost surely linear functions of one another. In the present case, this means that there exist reals $a \neq 0$ and $b \in \mathbb{R}$ such that $\Pr\{f(X) = aX + b\} = 1$; see, e.g., Theorem 4.5.7 in [3].

Therefore, it transpires that the relaxation of the distributional assumption on the pair (X, Y) given in [2] has merely been obtained at the cost of a reduction of the applicability of the result to linear functions. As Stein's identity is trivial for such functions, this is hardly an achievement.

3. Discussion Instead of pointing out what went wrong in not just one but two different mathematical derivations of identity (1) contained in [2], it is perhaps more useful to illustrate through simple examples why an identity of the form (1) cannot possibly hold universally and why, when it does, the form of the constant γ does not depend only on f and the distribution of X, but also critically on the nature of the relationship between the variables X and Y. For this reason, any extension of Stein's lemma is necessarily limited to classes of distributions; the larger the class, of course, the more useful the result.

The first example illustrates the fact that when a relation of the form (1) exists, the constant γ does not depend only on f and X but also on the nature of the dependence between X and Y within a specified class of distributions.

Example 1. Fix $\theta \in [-1, 1]$ and consider a random pair (X, Y) whose cumulative distribution function is given, for all $x, y \in [0, 1]$, by

$$\Pr(X \leq x, Y \leq y) = xy + \theta xy(1-x)(1-y).$$

The variables X and Y are then uniformly distributed on the interval $[0, 1]$ and their joint distribution is called a Farlie–Gumbel–Morgenstern (FGM) copula with parameter θ; see, e.g., [6]. It is easily checked that

$$\operatorname{cov}\{f(X), Y\} = \frac{\theta}{6} \int_0^1 f(x)(2x-1)dx$$

and in particular, $\operatorname{cov}(X, Y) = \theta/36$. Therefore, relation (1) holds with

$$\gamma = 6 \int_0^1 f(x)(2x-1)dx, \tag{4}$$

irrespective of the value of $\theta \in [-1, 1]$. In other words, a version of Stein's lemma holds for the entire FGM copula family with γ of the form (4).

However, if it is supposed instead that the random pair (X, Y) is such that, for some $\theta \in [0, 1]$ and all $x, y \in [0, 1]$,

$$\Pr(X \leq x, Y \leq y) = xy + \theta xy^2(1-x)^2(1-y),$$

then the variables X and Y are still uniformly distributed on $[0, 1]$ but

$$\operatorname{cov}\{f(X), Y\} = \frac{\theta}{12} \int_0^1 f(x)(1-x)(3x-1)dx$$

and in particular, $\operatorname{cov}(X, Y) = \theta/144$. Therefore, relation (1) holds with

$$\gamma = 12 \int_0^1 f(x)(1-x)(3x-1)dx. \tag{5}$$

Therefore, while X is uniform on $[0, 1]$ in both models, the expression for γ is not the same in equations (4) and (5). In other words, there is a version of Stein's lemma for these two classes of distributions but the multiplicative constant in terms of f is specific to each class. It is also interesting to note that in both cases, $\gamma \neq \mathrm{E}\{f'(X)\}$, i.e., it is of a different form than in Stein's original lemma.

The second example shows that for some classes of distributions, a relation of the form (1) does not exist.

EXAMPLE 2. Fix $\theta \in [-1, 1]$ and $\varphi \in [\theta - 1, (3 - \theta + \sqrt{9 - 6\theta - 3\theta^2})/2]$. Consider a random pair (X, Y) such that, for all $x, y \in [0, 1]$,

$$\Pr(X \le x, Y \le y) = xy + \theta xy(1 - x)(1 - y) + \varphi x^2 y^2 (1 - x)(1 - y).$$

Once again, the variables X and Y are uniformly distributed on the interval $[0, 1]$ and their joint distribution is an iterated FGM copula with parameters θ and φ; see, e.g., [4]. One can easily verify that

$$\mathrm{cov}\{f(X), Y\} = \frac{\theta}{6} \int_0^1 f(x)(2x - 1)dx + \frac{\varphi}{12} \int_0^1 f(x)(3x^2 - 2x)dx$$

and that in particular, $\mathrm{cov}(X, Y) = \theta/36 + \varphi/144$. Therefore, if relation (1) were valid for this entire two-parameter family of copulas, it would hold in particular when $\varphi = 0$, implying that γ is as given in Equation (4) because when $\varphi = 0$, the distribution reduces to a classical bivariate FGM copula with parameter θ. At the same, though, the relation (1) would need to hold when $\theta = 0$, and hence

$$\gamma = 12 \int_0^1 f(x)(3x^2 - 2x)dx = 6 \int_0^1 f(x)(2x - 1)dx$$

identically for all possible choices of f. However, this cannot be. For example if $f(x) = x^n$ for all $x \in [0, 1]$ and some non-negative integer n, the above identity is seen to fail except if $n = 1$, i.e., if f is linear.

In the last paragraph of [2], the author states that his results "are intuitive and they will be extremely useful in many disciplines, since they express an arbitrary function in a linear form." It should be clear from the above discussion that such a sweeping conclusion is unwarranted. There is, in effect, no truth or substance to the extension of Stein's lemma described in that article.

Acknowledgment Research funded in part by the Canada Research Chairs Program and the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

1. C. Adcock, Z. Landsman and T. Shushi, *Stein's lemma for generalized skew-elliptical random vectors*, Comm. Statist. Theory Methods (2020), in press.
2. M. Alghalith, *Generalizing Stein's Lemma*, C. R. Math. Rep. Acad. Sci. Canada **42** (2020), pp. 21–24.
3. G. Casella and R. L. Berger, *Statistical Inference*, 2nd edition, Duxbury Press, Pacific Grove, CA, 2002.
4. J. S. Huang and S. Kotz, *Correlation structure in iterated Farlie–Gumbel–Morgenstern distributions*, Biometrika **71** (1984), pp. 633–636.
5. Z. Landsman and J. Nešlehová, *Stein's lemma for elliptical random vectors*, J. Multivariate Anal. **99** (2008), pp. 912–927.
6. R. B. Nelsen, *An Introduction to Copulas*, 2nd edition, Springer, New York, 2006.
7. C. M. Stein, *Estimation of the mean of a multivariate normal distribution*, Ann. Statist. **9** (1981), pp. 1135–1151.

Department of Mathematics and Statistics, McGill University, 805, rue Sherbrooke ouest, Montréal (Québec) Canada H3A 0B9
e-mail: christian.genest@mcgill.ca