ISSN 2090-3359 (Print)
ISSN 2090-3367 (Online)

# Advances in Decision Sciences

*Volume 24*
*Issue 4*
*December 2020*

Michael McAleer

Editor-in-Chief
University Chair Professor
Asia University, Taiwan

Published by Asia University, Taiwan

ADS@ASIAUNIVERSITY

# Extension of Stein's Lemmas to General Functions and Distributions*

**Moawia Alghalith**

University of the West Indies

**Wing-Keung Wong****

Department of Finance, Fintech Center, and Big Data Research Center
Asia University
and
Department of Medical Research
China Medical University Hospital
and
Department of Economics and Finance
The Hang Seng University of Hong Kong

Revised: November 2020

***** The authors are grateful to a reviewer for very helpful comments and suggestions. The second author would like to thank Robert B. Miller and Howard E. Thompson for their continuous guidance and encouragement. This research has been partially supported by grants from Asia University, China Medical University Hospital, The Hang Seng University of Hong Kong, the Research Grants Council of Hong Kong (Project Number 12500915), and Ministry of Science and Technology (MOST, Project Numbers 106-2410-H-468-002 and 107-2410-H-468-002-MY3), Taiwan.

******Corresponding author: wong@asia.edu.tw

# Abstract

In this paper, we extend the lemmas in Stein (1973, 1981) and others to include situations in which the variables are dependent and non-normally distributed. There is no restriction on the form of the function, which could be linear or nonlinear, provided that the function is differentiable and the expectation of the derivative of the function exists. Thereafter, we give some examples of non-normal distributions and nonlinear functions to illustrate the theorems developed in the paper to hold, and show that the assertion of Genest (2020) is incorrect. In addition, we discuss applications of using the theorems in decision sciences.



**JEL:** C0, G0.

**Keywords:** Stein's Lemma, dependence, non-normality, differentiability, expectations.

## 1. Introduction

In this paper, we extend Stein (1973, 1981), Adcock (2007), Zhang, Li, and Meng (2008), and others to include the situations in which the variables are dependent and non-normally distributed and there is no restriction on the form of the function that it could be linear or nonlinear. We first extend Stein (1973) and others to introduce the covariance of X and a function, say, f(X), not only by relaxing the normal assumption, but also by dropping the requirement of the form of distribution for X, so that X can follow any distribution, under the assumption that f(X) is differentiable.

We then extend Stein (1981) to develop the covariance of one variable X and a function f(Y) by relaxing the joint bivariate normal assumption, but also by dropping the requirement of the form of joint distribution for X and Y.

Then, we give some examples of non-normal distributions and nonlinear functions to illustrate the theorems developed in our paper, and show that the assertion of Genest (2020) is incorrect and give some examples to illustrate the assertion of Genest (2020) is incorrect. Finally, we discuss the applications of the theorems developed in our paper.

## 2. Theory

We first modify Stein's Lemma developed by Stein (1973) as stated in the following lemma:

**Lemma 1.** *Let X be a normally-distributed random variable with mean,* $\mu_X$ *, f(x) is a differentiable function , and Ef'(X) exists. Then:*

$$Ef(X)(X-\mu_X) = Var(X)Ef'(X)$$

In addition, we modify another version of Stein's Lemma developed by Stein (1981) as stated in the following lemma:

**Lemma 2.** *Let X and Y be a pair of random variables following a bivariate normal distribution with covariance Cov(X,Y), f(x) is differentiable and Ef'(X) exists. Then:*

*Cov(f(X),Y) = Ef'(X)Cov(X,Y).*

Our results are more flexible than those in Stein (1973, 1981) because the constant is not necessarily equal to *Ef'(X)*.

There are some extensions of the lemmas to include other distributions including the exponential and the elliptical distributions, see, for example, Landsman, *et al.* (2015) and Shushi (2018). We note that the extensions of Lemma1 and Lemma 2 in the literature are based on extensions to specific distributions. In this paper, we extend the theory further by removing the distribution assumption to get the distribution-free results as shown in the following theorems:

**Theorem 1.** *Let X be a random variable with finite mean,* $\mu_X$, *and finite standard deviation,* $\sigma_X$, *f(x) is differentiable function, and a is a non-arbitrary constant. Then:*

$Ef(X)(X-\mu_X) = aVar(X)$

**Proof.** Using Taylor's extension, we have:

$f(x) = f(c) + f'(x^{\wedge})(x-c)$

In addition, by applying the mean-value theorem, we obtain:

$\Phi = \Sigma f''(x_i)/n$

and

$\beta = \lim \Sigma f''(x_i)/n \quad as\ n \rightarrow \infty$

It is intuitive that as $n \to \infty$ (that is, for a sufficiently large interval), the convergence will hold. The intuition is that the larger the sample/interval, the smaller the impact of a change in x on the average of f'(x) until it becomes negligible. Thus:

$$(d\,\Phi/dx) = f''(x)/n$$

so that (for a sufficiently large interval) as $n \to \infty$, $\Phi \to \beta$, where $\beta$ is a constant. Therefore, $f(x)$ can be expressed as:

$$f(x) = c + \beta x \quad \text{as } n \to \infty.$$

We note that $\beta$ is not necessarily equal to $Ef'(X)$ as shown in Lemma 1, and thus, Theorem 2 is an extension of Lemma 2 introduced by Stein (1973). We extend Lemma 2 in the following theorem:

**Theorem 2.** *Let X and Y be a pair of random variables with covariance Cov(X,Y), f(x) is differentiable, and b is a non-arbitrary constant. Then:*

$$Cov(f(X),Y) = bCov(X,Y).$$

Using the proof of Theorem 1, one could obtain the proof of Theorem 2 by using a similar argument. We also recognize the crucial difference between a linearized function and a linear function, and a non-arbitrary constant and an arbitrary constant. Note that Zhang, Li, and Meng (2008) extend Stein's Lemma to include the case when the variables are dependent and normally distributed and independent and non-normally distributed.

In Theorem 2, we extend Stein (1973, 1981), Zhang, Li, and Meng (2008), and others to include the situations in which the variables could be dependent and non-normally distributed. To be more specific, Theorem 2 is an extension of Lemma 2 introduced by Stein (1981) and others to establish a property for the covariance of a random variable Y and a function of another random variable X such that it is equal to the product of the covariance of X and Y and a function of another random variable X and a non-arbitrary constant in which there is no restriction on the form of the function, and the variables could be dependent and non-normally distributed.

In Theorems 1 and 2, instead of using a particular distribution as in most, if not all, of the extensions of Stein's lemma, we will shift the focus to the functional form f(x). If this function can be transformed to a tractable form, under certain conditions, both Theorems1 and 2 can be derived regardless of the distribution. For example, it could be arcsine, exponential, exponential-logarithmic, elliptical, gamma, (generalized) extreme value distributions, Gompertz, hyperbolic secant, Levy, logistic, log-logistic, Laplace, Maxwell, normal, Pareto, Rayleigh, semicircle, Student's t, triangle, U-power, uniform, Weibull, Wald distribution, and actually any distribution as long as the moments and the function are finite.

Ideally, if the function can be linearized in some sense (such as some form convergence); that is, if under certain conditions, if:

$$f(x) \rightarrow c + \beta x$$

clearly the constant will not be arbitrary. In general, it will depend on the functional form, the distribution, and the interval of the variable.

Clearly, one tool to achieve this is by applying Taylor expansions; that is, if we can find conditions under which the sum remainders of Taylor expansions are minimized.

Consequently, the Stein-like results follow directly from this convergence. We also recognize the crucial difference between a linearized function and a linear function; and a non-arbitrary constant and an arbitrary constant.

Genest (2020) argued that such identities (equations in both (Stein 1973) and (Stein 1981)) hold only for a linear function and he made the following assertion[1]:

**Assertion:** It transpires that the relaxation of the distributional assumption on the pair (X,Y) given in both (Stein 1973) and (Stein 1981) has merely been obtained at the cost of a reduction of the applicability of the result to linear functions.

The proofs of Theorems 1 and 2 show that Assertion statement1 is not correct and, in the contract, both (Stein 1973) and (Stein 1981) hold for so many nonlinear functions. In addition, we construct some examples to show that Statement 1 is not correct.

## 3. Illustrations

In this section, we will construct some examples to show that both (Stein 1973) and (Stein 1981) hold not only for linear functions, but also for nonlinear functions. We also construct an example to show that both Theorems 1 and 2 do not hold. We first construct a simple nonlinear function as shown in Example 1:

**Example 1.** *Clearly, both Theorems 1 and 2 hold for the following function:*

$f(X)=aX^2 +bX^3$

*for a specific value of a/b.*

One could easily show the function constructed in Example 1 hold for Theorems 1 and 2.
We turn to construct a more complicated example shown as follows:

**Example 2.** *Clearly, both Theorems 1 and 2 hold for the following function:*

---

[1]We have rewritten his statement to fit the contents of our paper.

$$f(X)=aX+v_1 X^3 +v_2 X^4 +v_3e^{bX}$$

*then, for specific values of* $v_i$*.*

**Example 3.** *Theorems 1 and 2 hold for the following function:*

$f(X) = aX+g(X),$

*where g is uncorrelated with Y.*

In Example 3, the dependence structure between X and Y is automatically captured by the non-arbitrary constant b. To show this, if X and Y are dependent, then $X = h(Y)$ and thus $f'(X)$ in the paper (he is referring to) is expressed a $f'(h(Y)$ and $b$ depends on $f'(x)$. Thus, in general, b is affected by the dependence structure. Also the choice of the constant $c$ is in accordance with the functional forms and the distributions, and $b$ depends on $f'(c)$. Furthermore, as is the case with the classical Stein's lemma, the constant does not have to always exist.

We turn to construct the following examples to show that both Theorems 1 and 2 do not hold. We first construct an example of distribution that both Theorems 1 and 2 do not hold.

**Example 4.** *A simple example that both Theorems 1 and 2 do not hold is when X follows Cauchy distribution. When X follows Cauchy distribution, the moments do not exist, and thus, Theorem 1 does not hold.*

We now construct an example of a function that both Theorems 1 and 2 do not hold.

**Example 4.** *One could simply obtain a function that the function is not differentiable, then both Theorems 1 and 2 do not hold. It is because both Theorems 1 and 2 require the assumption that the function is differentiable. A simple function that is non-differentiable everywhere is the*

*Weierstrass function* $f(x) = \Sigma a^n \cos(b^n \Pi x)$ *with restricted parameters. For this function, Theorems 1 and 2 do not hold.*

## 4. Applications in Decision Sciences

There are many studies using Theorems 1 and 2 in developing their models in many areas of Applications in Decision Sciences, including mathematics, statistics, economics, finance, risk management, and many other areas. Here we discuss a few in this section.

There are many studies that use Theorems 1 and 2 in developing their models in mathematics. For example, Egozcue, *et al.* (2009) and others use Theorems 1 and 2 to develop some properties for the covariance inequalities and copulas. There are many studies that use the theorems in statistics. For instance, Wong and Miller (1990), and others apply Theorems 1 and 2 to develop robust estimation for regression and time series models. Bai et al. (2015) apply the theorems in developing causality tests.

There are studies that used the theorems in economic modelling. For example, Alghalith and Wong (2020) applied them to examine the properties of welfare gains from macro-hedging. Others applied them in financial modelling. Lam, *et al.* (2012) and others used them to develop theories of behaviors for different investors. Several studies used them to develop models in risk management. For example, Leung and Wong (2008) and others apply Theorems 1 and 2 to introduce different risk measures or different.

## 5. Discussions and Concluding Remarks

In this paper, we extended Stein (1973, 1981), Adcock (2007), Zhang, Li, and Meng (2008) and others to include the cases in which the variables are dependent, non-normally distributed. In doing so, we used a novel approach in extending Stein's lemma. Instead of focusing on Ef'(X), we emphasize the role of a constant expressed as a limit of a simple average of f'.

We then give some examples of non-normal distributions and nonlinear functions that Theorems 1 and 2 will hold. We also give examples of a distribution and a function that Theorems 1 and 2 do not hold. We also show that the assertion of Genest is incorrect. Future studies can extend our results to a multivariate framework.

## References


Adcock, C.J. (2007), Extensions of Stein's Lemma for the Skew-Normal Distribution, Communications in Statistics - Theory and Methods, 36(9), 1661-1671.

Alghalith, M., Wong, W.K. (2020), Welfare Gains from Macro-Hedging, Annals of Financial Economics, 15(2), 2050009.

Bai, Z.D., Li, H., McAleer, M., Wong, W.K. (2015), Stochastic Dominance Statistics for Risk Averters and Risk Seekers: An Analysis of Stock Preferences for USA and China, Quantitative Finance, 15(5), 889-900.

Egozcue, M., Fuentes Garcia, F., Wong, W.K. (2009), On Some Covariance Inequalities for Monotonic and Non-monotonic Functions, Journal of Inequalities in Pure and Applied Mathematics, 10(3), Article 75, 1-7.

Genest, C. (2020), On an Extension of Stein's Lemma, C. R. Math. Rep. Acad. Sci. Canada 42, 25-28.

Lam, K., Liu, T.S., Wong, W.K. (2012), A New Pseudo Bayesian Model with Implications to Financial Anomalies and Investors' Behaviors, Journal of Behavioral Finance, 13(2), 93-107.

Landsman, Z, Vanduffel, S., Yao, J. (2015), Some Stein-type Inequalities for Multivariate Elliptical Distributions and Applications, Statistics & Probability Letters, 97, 54-62.

Leung, P.L., Wong, W.K. (2008), On Testing the Equality of the Multiple Sharpe Ratios, with Application on the Evaluation of Ishares, Journal of Risk, 10(3), 1-16.

Shushi, T. (2018), Stein's Lemma for Truncated Elliptical Random Vectors, Statistics & Probability Letters, 137, 297-303.

Stein, C.M (1973), Estimation of the Mean of a Multivariate Normal Distribution, Proceedings of Prague Symposium on Asymptotic Statistics, 345-381.

Stein, C.M. (1981), Estimation of the Mean of a Multivariate Normal Distribution, Annals of Statistics, 9, 1135-1151.

Wong, W.K., R.B. Miller (1990), Analysis of ARIMA-Noise Models with Repeated Time Series, Journal of Business and Economic Statistics, 8(2), 243-250.

Zhang, C., Li, J., Meng, J. (2008), On Stein's Lemma, Dependent Covariates and Functional Monotonicity in Multi-dimensional Modeling, Journal of Multivariate Analysis, 99(10), 2285-2303.