

# UNIFIED VARIANCE BOUNDS AND A STEIN–TYPE IDENTITY\*

N. Papadatos<sup>(a)</sup> and V. Papathanasiou<sup>(b),\*\*</sup>

<sup>(a)</sup> University of Cyprus, Department of Mathematics and Statistics,  
P.O. Box 20537, 1678 Nicosia, Cyprus

<sup>(b)</sup> University of Athens, Department of Mathematics,  
Panepistemiopolis, 157 84 Athens, Greece

## Abstract

Let  $X$  be an absolutely continuous (a.c.) random variable (r.v.) with finite variance  $\sigma^2$ . Then, there exists a new r.v.  $X^*$  (which can be viewed as a transform on  $X$ ) with a unimodal density, satisfying the extended Stein-type covariance identity

$$\text{Cov}[X, g(X)] = \sigma^2 \mathbb{E}[g'(X^*)]$$

for any a.c. function  $g$  with derivative  $g'$ , provided that  $\mathbb{E}|g'(X^*)| < \infty$ . Properties of  $X^*$  are discussed and, also, the corresponding unified upper and lower bounds for the variance of  $g(X)$  are derived.

*AMS 1991 subject classifications.* Primary 60E15; secondary 60E05.

*Key words and phrases:* Stein–type identity, variance bounds, transformation of random variables.

## 1 Introduction

The well-known Stein’s identity, [16], [17], for the standard normal r.v.  $Z$  is formulated as follows. For every absolutely continuous (a.c.) function  $g$  with derivative  $g'$  such that  $\mathbb{E}|g'(Z)| < \infty$ ,

$$\mathbb{E}[Zg(Z)] = \mathbb{E}[g'(Z)] \tag{1.1}$$

(throughout this paper, the term ‘a.c.’=‘absolutely continuous’ will be used either to describe an r.v. having a density with respect to Lebesgue measure on  $\mathbb{R}$ , or to denote an ordinary a.c. function; in any case, the meaning will be clear from the context). This identity has had many important applications in several areas of Probability and Statistics; see for example [16], [12] and [13]. Several generalizations to other r.v.’s can be found in [4], [5] and [12].

---

\*Research partially supported by the Ministry of Industry, Energy and Technology of Greece under grant 1369.

\*\*Part of this work was done when the author was visiting the University of Bristol.

Another interesting result from the point of view of upper variance bounds is the inequality of Chernoff, [10],

$$\text{Var}[g(Z)] \leq \mathbb{E}[(g'(Z))^2] \quad (1.2)$$

with equality iff (if and only if)  $g$  is linear. This and its multivariate normal analogue, [8], motivated the generalizations to arbitrary discrete and continuous r.v.'s as well as corresponding lower variance bounds, [1], [3].

An important role of the general covariance identity given in [4] (see (2.4), below) is in the derivation of simple and elementary proofs of the Central Limit Theorem (see [6], [7]). Other applications concerning the rate of convergence in the Local Limit Theorem for sums of independent r.v.'s are given in [2].

However, all the preceding results (and, in particular, the validity of the generalized covariance identity given in [4] and the variance bounds given in [3] and [4]) require an interval support of the basic r.v.  $X$ , which is rather restrictive.

In the present paper we avoid the restriction of an interval support, introducing an appropriate (smooth) transformation  $X^*$ , which is in fact a new r.v. corresponding to  $X$ . The r.v.  $X^*$  is always uniquely defined (provided that  $X$  is a.c. with finite variance), having itself an absolutely continuous unimodal density. Thus, for any a.c. r.v.  $X$  with finite second moment, there exists a new smooth r.v.  $X^*$  with unimodal a.c. density satisfying the generalized Stein covariance identity. This transformation behaves well to convolutions of independent r.v.'s. Moreover, it appears in the upper and lower bounds for the variance of any a.c. function  $g$  of  $X$  (and it is, in fact, the only r.v. with this property; Theorem 3.1).

It should be noted that in a recent paper, independently of our results, Goldstein and Reinert [11] used a similar approach (the so-called *zero bias transformation*), which, in fact, is based on the same covariance identity when  $\mathbb{E}[X] = 0$ . They also fruitfully applied this identity to estimate the rate of convergence in the CLT, obtaining an  $O(n^{-1})$  bound for smooth functions (see Corollary 3.1 in [11]). Furthermore, they presented a nice application for dependent samples. However, except of the definition, the results of the present paper are completely different; our main interest is on unified variance bounds and their connection with Stein's identity. We are also interested on the behavior of the inverse transform  $X^* \rightarrow X$ ; in fact we show that, under general conditions, the distribution of  $X^*$  uniquely determines that of  $X$  (Theorems 2.1 and 3.2). Finally, we also include some illustrative examples.

## 2 Properties of the transformation

Let  $X$  be an a.c. r.v. with density  $f$ , mean  $\mu$ , variance  $\sigma^2$  and support  $S(X)$  (for the sake of simplicity, we will always mean the support of an a.c. r.v.  $X$  with density  $f$  to be the set  $S(X) = \{x : f(x) > 0\}$ ). We simply define  $X^*$  to be a random variable with density  $f^*$ , given by the relation

$$f^*(x) = \frac{1}{\sigma^2} \int_{-\infty}^x (\mu - t)f(t) dt = \frac{1}{\sigma^2} \int_x^{\infty} (t - \mu)f(t) dt. \quad (2.1)$$

Obviously the right hand sides of (2.1) are equal and thus,  $f^*$  is nonnegative. An application of Tonelli's Theorem shows that  $f^*$  integrates to 1, and therefore it is indeed a probability density (c.f. [4], [6]). The following Lemma summarizes the properties of  $X^*$  and shows the generalized Stein's identity (2.2).

**Lemma 2.1** (i)  $f^*$  is a unimodal a.c. density with mode  $\mu$  and maximal value

$$f^*(\mu) = \frac{\mathbb{E}|X - \mu|}{2\sigma^2}.$$

Furthermore, the function  $(f^*(x))'/(\mu - x)$  (defined almost everywhere) is non-negative and integrable.

(ii) For each a.c. function  $g$  with  $\mathbb{E}|g'(X^*)| < \infty$ ,

$$\text{Cov}[X, g(X)] = \sigma^2 \mathbb{E}[g'(X^*)]. \quad (2.2)$$

(iii) If the r.v.  $Y$  satisfies the identity

$$\text{Cov}[X, g(X)] = \sigma^2 \mathbb{E}[g'(Y)]$$

for every a.c.  $g$  with  $\mathbb{E}|g'(Y)| < \infty$ , then  $Y \stackrel{d}{=} X^*$  (in the sense that  $Y$  and  $X^*$  have the same distribution).

(iv)  $S(X^*) = (\text{ess inf } S(X), \text{ess sup } S(X))$ .

(v) For arbitrary scalars  $a \neq 0$  and  $b$ ,

$$(aX + b)^* \stackrel{d}{=} aX^* + b.$$

(vi) For independent a.c. r.v.'s  $X_1, X_2$  with means  $\mu_1, \mu_2$ , variances  $\sigma_1^2, \sigma_2^2$  and arbitrary scalars  $a_1$  and  $a_2$  with  $a_1 a_2 \neq 0$ ,

$$(a_1 X_1 + a_2 X_2)^* \stackrel{d}{=} B(a_1 X_1^* + a_2 X_2) + (1 - B)(a_1 X_1 + a_2 X_2^*), \quad (2.3)$$

where  $X_1, X_2, X_1^*, X_2^*$  and  $B$  are mutually independent with

$$\mathbb{P}[B = 1] = \frac{a_1^2 \sigma_1^2}{a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2} = 1 - \mathbb{P}[B = 0].$$

**Proof:** (i), (iv) and (v) are obvious. (ii) follows from the definition of  $X^*$  using Fubini's theorem, since by the assumption that  $\mathbb{E}|g'(X^*)| < \infty$ , the nonnegative functions  $g_1(x, t) = |g'(x)|(\mu - t)f(t)I(t < x)$  and  $g_2(x, t) = |g'(x)|(t - \mu)f(t)I(t > x)$  are integrable over  $(-\infty, \mu] \times (-\infty, \mu]$  and  $[\mu, \infty) \times [\mu, \infty)$ , respectively (see also Lemma 3.1 in [4]). (iii) follows from (ii) and the fact that

$$\mathbb{E}[G(X^*)] = \mathbb{E}[G(Y)]$$

for every bounded (measurable) function  $G$ . Regarding (vi), observe that for  $S = a_1X_1 + a_2X_2$  and  $g'$  bounded, we have from (ii):

$$\text{Cov}[S, g(S)] = \sigma^2 \mathbb{E}[g'(S^*)],$$

where  $\sigma^2 = a_1^2\sigma_1^2 + a_2^2\sigma_2^2$ . On the other hand,

$$\begin{aligned} \text{Cov}[S, g(S)] &= a_1 \text{Cov}[X_1, g(S)] + a_2 \text{Cov}[X_2, g(S)] \\ &= a_1^2\sigma_1^2 \mathbb{E}[g'(a_1X_1^* + a_2X_2)] + a_2^2\sigma_2^2 \mathbb{E}[g'(a_1X_1 + a_2X_2^*)]. \end{aligned}$$

It follows that for any bounded function  $G$ ,

$$\begin{aligned} \mathbb{E}[G(S^*)] &= \frac{a_1^2\sigma_1^2}{\sigma^2} \mathbb{E}[G(a_1X_1^* + a_2X_2)] + \frac{a_2^2\sigma_2^2}{\sigma^2} \mathbb{E}[G(a_1X_1 + a_2X_2^*)] \\ &= \mathbb{E}[G(B(a_1X_1^* + a_2X_2) + (1 - B)(a_1X_1 + a_2X_2^*))], \end{aligned}$$

which completes the proof.

It should be noted that in the previous Lemma, and elsewhere in this paper, the term ‘unimodal’ is reserved to denote a function  $h : \mathbb{R} \rightarrow \mathbb{R}$  with the property that there exists some  $m \in \mathbb{R}$  such that  $h(x)$  is nondecreasing for  $x \leq m$  and is nonincreasing for  $x \geq m$ ; each  $m$  with this property is called a ‘mode’ of  $h$ . Of course, the assertion that  $h$  is a unimodal and a.c. function implies that the mode(s) of  $h$  form a compact interval  $[a, b]$  with  $-\infty < a \leq b < \infty$  (which, in the case  $a = b$ , reduces to single point).

The known identity of [4], (2.4) below, requiring an interval support of  $X$ , follows immediately from (2.2). Indeed, when  $S(X)$  is a (finite or infinite) interval,  $S(X)$  can always be taken to be open, and then  $S(X) = S(X^*)$ . Thus, for the nonnegative function  $w(x) = f^*(x)/f(x)$  (defined on  $S(X)$ ), (2.2) becomes

$$\text{Cov}[X, g(X)] = \sigma^2 \mathbb{E}[w(X)g'(X)]. \quad (2.4)$$

The identity (2.2) remains valid for any non-decreasing (or non-increasing) a.c. function  $g$ , even in the case where  $\mathbb{E}[g'(X^*)] = \infty$  (or  $-\infty$  if  $g$  is non-increasing), as it follows by an application of Tonelli’s (instead of Fubini’s) theorem. In this case,  $\mathbb{E}[(X - \mu)(g(X) - g(\mu))] = \infty$  (or  $-\infty$ ). If, however,  $\mathbb{E}|g'(X^*)| = \infty$  and  $g$  is arbitrary, it may happen that  $\mathbb{E}|(X - \mu)(g(X) - g(\mu))| < \infty$ , as the following example shows.

**Example 1** Assume that  $X$  has density  $f(x) = (3/8) \min\{1, x^{-4}\}$ ,  $-\infty < x < \infty$  and

$$\begin{aligned} g(x) &= \sum_{n=1}^{\infty} (2n - 1)(|x| - a_{2n-2}) I(a_{2n-2} \leq |x| < a_{2n-1}) \\ &\quad + \sum_{n=1}^{\infty} 2n(a_{2n} - |x|) I(a_{2n-1} \leq |x| < a_{2n}), \end{aligned}$$

where  $a_0 = 0$  and  $a_n = 1 + 1/2 + \dots + 1/n$  for  $n \geq 1$ . It follows that  $0 \leq g(x) \leq 1$  and  $g(-x) = g(x)$  for all  $x$ . Moreover,  $g$  is a.c. with derivative  $g'(x)$  (outside the

set  $\{0, \pm a_1, \pm a_2, \dots\}$  satisfying  $|g'(x)| \geq |x| \geq |g(x)| = g(x)$  for almost all  $x$ . Since  $\mathbb{E}[X] = 0$ ,  $\mathbb{E}|X| = 3/4$ ,  $\text{Var}[X] = 1$  and  $S(X) = (-\infty, \infty)$ , we conclude that  $X^*$  has the density  $f^*(x) = (3/16)[(2 - x^2)I(|x| \leq 1) + x^{-2}I(|x| > 1)]$  supported by the entire real line, and  $\mathbb{E}|g'(X^*)| \geq \mathbb{E}|X^*| = \infty$ , while  $\mathbb{E}|Xg(X)| \leq \mathbb{E}|X| = 3/4$ .

We have shown in Lemma 2.1(iii) that  $X^*$  is the only r.v. satisfying the identity (2.2), and thus, an equivalent definition of  $X^*$  could be given via the covariance identity; the latter approach is due to Goldstein and Reinert ([11], Definition 1.1), who proved that such a transformation is uniquely defined by this identity. Moreover, their approach extends to r.v.'s that are not necessarily a.c., e.g., for the symmetric Bernoulli r.v.  $X$  taking the values  $\pm 1$  with probability  $1/2$ ,  $X^*$  is uniformly distributed over the interval  $(-1, 1)$ . Our approach, however, is restricted to a.c. r.v.'s; for this reason, the analytic definition (2.1) is possible and, moreover, the density  $f^*$  itself turns out to be an a.c. unimodal function.

Our results also go to the opposite direction; the following Theorem shows that, in general, the distribution of  $X^*$  uniquely determines that of  $X$ .

**Theorem 2.1** *Assume that the r.v.  $Y$  has a unimodal a.c. density  $h$ .*

- (i) *If the mode  $m$  of  $h$  is unique (i.e.,  $h(x) < h(m)$  for all  $x \neq m$ ), there exists an r.v.  $X_m$  such that  $X_m^* \stackrel{d}{=} Y$  iff the function  $h'(x)/(m - x)$  is integrable. Moreover,  $X_1^* \stackrel{d}{=} Y$  and  $X_2^* \stackrel{d}{=} Y$  implies  $X_1 \stackrel{d}{=} X_2$  (and thus,  $X_m$  is unique).*
- (ii) *If  $\{x : x \text{ is a mode of } h\} = [a, b]$  with  $a < b$ , then for each  $\mu \in (a, b)$ , there always exists a unique r.v.  $X_\mu$  such that  $\mathbb{E}[X_\mu] = \mu$  and  $X_\mu^* \stackrel{d}{=} Y$ . Moreover, for  $\mu = a$  or  $\mu = b$ , there exists an r.v.  $X_\mu$  such that  $\mathbb{E}[X_\mu] = \mu$  and  $X_\mu^* \stackrel{d}{=} Y$  iff the function  $h'(x)/(\mu - x)$  is integrable. Finally, if  $X_1^* \stackrel{d}{=} Y$ ,  $X_2^* \stackrel{d}{=} Y$  and  $\mathbb{E}[X_1] = \mathbb{E}[X_2]$ , then  $X_1 \stackrel{d}{=} X_2$ .*

**Proof:** (i) If  $X$  is an r.v. with mean  $\mu$  variance  $\sigma^2$  and density  $f$  such that  $X^* \stackrel{d}{=} Y$ , it follows from Lemma 2.1(i) that  $\mu$  must be a mode of  $f^* = h$  and that  $h'(x)/(\mu - x)$  is integrable. Thus,  $\mu = m$  and  $h'(x)/(m - x)$  is integrable. Assume now that the function  $h'(x)/(m - x)$  (defined almost everywhere) is integrable. Observe that it is also nonnegative (because  $m$  is the mode of  $h$ ) and define the r.v.  $X_m$  with density

$$f_m(x) = \frac{h'(x)}{c(m - x)}, \text{ where } c = \int_{-\infty}^{\infty} \frac{h'(x)}{m - x} dx > 0.$$

Since  $\lim_{\pm\infty} h(x) = 0$  (because  $h$  is a unimodal density), we have  $\mathbb{E}[m - X_m] = 0$ . Applying Tonelli's Theorem we have

$$\int_{-\infty}^m (m - x)h'(x) dx = \int_{-\infty}^m \int_x^m h'(x) du dx = \int_{-\infty}^m h(u) du,$$

and similarly,

$$\int_m^\infty (x-m)(-h'(x)) dx = \int_m^\infty h(u) du,$$

yielding  $\text{Var}[X_m] = 1/c$ . Therefore,  $X_m^* \stackrel{d}{=} Y$ . We now show that  $X_m$  is unique. Indeed, if  $X^* \stackrel{d}{=} Y$  for an r.v.  $X$  with mean  $\mu$ , variance  $\sigma^2$  and density  $f$ , it follows from Lemma 2.1(i) that  $\mu$  must be a mode of  $f^* = h$  and therefore  $\mu = m$ . Hence,

$$f(x) = \frac{\sigma^2 h'(x)}{m-x} = c\sigma^2 f_m(x)$$

for almost all  $x$ , and thus  $X \stackrel{d}{=} X_m$ .

(ii) Since  $h$  is constant in  $[a, b]$ ,  $h' \equiv 0$  in  $(a, b)$ . Hence, for all  $\mu \in (a, b)$ ,

$$\int_{-\infty}^\infty \frac{h'(x)}{\mu-x} dx = \int_{-\infty}^a \frac{h'(x)}{\mu-x} dx + \int_b^\infty \frac{-h'(x)}{x-\mu} dx \leq \frac{h(a)}{\mu-a} + \frac{h(b)}{b-\mu} < \infty.$$

Consider the r.v.  $X_\mu$  with density

$$f_\mu(x) = \frac{h'(x)}{c_\mu(\mu-x)}, \text{ where } c_\mu = \int_{-\infty}^\infty \frac{h'(x)}{\mu-x} dx.$$

Then  $\mathbb{E}[X_\mu] = \mu$ ,  $\text{Var}[X_\mu] = 1/c_\mu$ , and therefore  $X_\mu^* \stackrel{d}{=} Y$ . It is easy to see that if  $X^* \stackrel{d}{=} Y$  and  $\mathbb{E}[X] = \mu$  for some r.v.  $X$  with density  $f$  and variance  $\sigma^2$ , then  $X \stackrel{d}{=} X_\mu$ . Indeed, it follows from (2.1) and Lemma 2.1(i) that

$$f(x) = \frac{\sigma^2 h'(x)}{\mu-x} = c_\mu \sigma^2 f_\mu(x)$$

for almost all  $x$ . Therefore, since the mean of any r.v.  $X$  satisfying  $X^* \stackrel{d}{=} Y$  must be a mode of  $h$ , either  $\mathbb{E}[X] = \mu \in (a, b)$  (and thus  $X \stackrel{d}{=} X_\mu$ ) or  $\mathbb{E}[X] = a$  (and  $h'(x)/(a-x)$  is integrable) or  $\mathbb{E}[X] = b$  (and  $h'(x)/(b-x)$  is integrable).

The following example shows that all the cases described by Theorem 2.1 are possible.

**Example 2** (i) Assume that  $Y$  has the unimodal a.c. density

$$h(x) = \frac{1}{3} (\min\{2 - |x|, 1\})^+$$

with derivative  $h'(x) = -(1/3)\text{sign}(x)I(1 < |x| < 2)$  for almost all  $x$ . Then, for all  $\mu \in (-1, 1)$ , the r.v.  $X_\mu$  with density

$$f_\mu(x) = (3c_\mu|x-\mu|)^{-1} I(1 < |x| < 2), \text{ where } c_\mu = \frac{1}{3} \log \left[ \frac{4-\mu^2}{1-\mu^2} \right],$$

satisfies  $\mathbb{E}[X_\mu] = \mu$  and  $X_\mu^* \stackrel{d}{=} Y$ . Moreover,  $h'(x)/(\mu-x)$  is not integrable for  $\mu = \pm 1$ .

(ii) Assume that  $Y$  has the unimodal a.c. density

$$h(x) = \frac{6}{19} (\min \{x + 2, 1\} I(-2 < x < 1) + x(2 - x)I(1 \leq x < 2)).$$

Then, for any  $\mu \in (-1, 1]$ , the r.v.  $X_\mu$  with density

$$f_\mu(x) = \frac{6}{19c_\mu|x - \mu|} (I(-2 < x < -1) + 2(x - 1)I(1 < x < 2)),$$

where  $c_\mu = (6/19) (2 + \log [(\mu + 2)/(\mu + 1)] - 2(1 - \mu) \log [(2 - \mu)/(1 - \mu)])$ , satisfies  $\mathbb{E}[X_\mu] = \mu$  and  $X_\mu^* \stackrel{d}{=} Y$ , while  $h'(x)/(-1 - x)$  is not integrable. Similarly, for the r.v.  $W = -Y$  with density  $h(-x)$ , there exists an r.v.  $R_\mu$  such that  $\mathbb{E}[R_\mu] = \mu$  and  $R_\mu^* \stackrel{d}{=} W$  iff  $\mu \in [-1, 1)$ .

(iii) For the r.v.  $Y$  with density

$$h(x) = \frac{3}{10} (|x|(2 - |x|)I(1 < |x| < 2) + I(|x| \leq 1)),$$

there exists an r.v.  $X_\mu$  such that  $\mathbb{E}[X_\mu] = \mu$  and  $X_\mu^* \stackrel{d}{=} Y$  for any mode  $\mu \in [-1, 1]$ .

### 3 Application to variance bounds

Upper bounds for the variance of a function  $g(X)$  of a normal r.v.  $X$  in terms of  $g'$  are known as *the inequality of Chernoff* [10] (see also [8] and [18]). Upper and lower variance bounds of  $g(X)$  for an arbitrary r.v.  $X$  were considered in [1] and [3] (see also [4]–[7] and references therein). Both upper and lower variance bounds may be obtained as by-products of the Cauchy-Schwarz inequality. The following Lemma summarizes and unifies these bounds in terms of the r.v.  $X^*$ ; in effect, (3.1) is a Chernoff-type, [8], upper bound; (3.2) is a Cacoullos-type, [14], [15], lower bound as obtained in [1] and [3], in terms of a function  $w$  (see also (3.4) below).

**Lemma 3.1** *Let  $X$  be an a.c. r.v. with mean  $\mu$  and variance  $\sigma^2$ . Then, for every a.c. function  $g$  with derivative  $g'$ , we have the following bounds.*

$$(i) \quad \mathbb{E}[(g(X) - g(\mu))^2] \leq \sigma^2 \mathbb{E}[(g'(X^*))^2], \quad (3.1)$$

with equality iff either  $\mathbb{E}[g^2(X)] = \infty$  or

$$g(x) - g(\mu) = \begin{cases} a_1(x - \mu) & \text{if } x \leq \mu, \\ a_2(x - \mu) & \text{if } x \geq \mu, \end{cases}$$

for some constants  $a_1, a_2$  and for all  $x \in S(X^*)$ .

(ii) If  $\mathbb{E}|g'(X^*)| < \infty$ ,

$$\text{Var}[g(X)] \geq \sigma^2 \mathbb{E}^2[g'(X^*)], \quad (3.2)$$

with equality iff  $\mathbb{P}[g(X) = aX + b] = 1$  for some constants  $a$  and  $b$ .

**Proof:** (i) Let  $f$  be a density of  $X$ . We then have

$$\begin{aligned} & \sigma^2 \mathbb{E}[(g'(X^*))^2] \\ &= \int_{-\infty}^{\mu} (g'(x))^2 \int_{-\infty}^x (\mu - t)f(t) dt dx + \int_{\mu}^{\infty} (g'(x))^2 \int_x^{\infty} (t - \mu)f(t) dt dx \\ &= \int_{-\infty}^{\mu} f(t)(\mu - t) \int_t^{\mu} (g'(x))^2 dx dt + \int_{\mu}^{\infty} f(t)(t - \mu) \int_{\mu}^t (g'(x))^2 dx dt \\ &\geq \int_{-\infty}^{\mu} f(t) (g(\mu) - g(t))^2 dt + \int_{\mu}^{\infty} f(t) (g(t) - g(\mu))^2 dt \\ &= \mathbb{E}[(g(X) - g(\mu))^2], \end{aligned}$$

from Tonelli's theorem and the Cauchy-Schwarz inequality for integrals. Observe that if  $\mathbb{E}[g^2(X)] = \infty$ , the equality holds in a trivial way ( $\infty = \infty$ ); otherwise, the equality holds iff there exist constants  $a_1$  and  $a_2$  such that  $g'(x) = a_1 + (a_2 - a_1)I(x \geq \mu)$  for almost all  $x \in S(X^*)$ , which completes the proof.

(ii) We have from (2.2),

$$\sigma^4 \mathbb{E}^2[g'(X^*)] = \text{Cov}^2[X, g(X)] \leq \sigma^2 \text{Var}[g(X)]$$

by the Cauchy-Schwarz inequality for r.v.'s, and the proof is complete.

**Corollary 3.1** For every a.c. function  $g$ ,

$$\text{Var}[g(X)] \leq \sigma^2 \mathbb{E}[(g'(X^*))^2], \quad (3.3)$$

with equality iff either  $\mathbb{E}[g^2(X)] = \infty$  or  $g$  is linear on  $S(X^*)$ .

Note that (3.2) continues to hold for any non-decreasing (non-increasing) a.c. function  $g$ , even in the case where  $\mathbb{E}[g'(X^*)] = \pm\infty$  (in this case,  $\mathbb{E}[g^2(X)] = \infty$ ). Moreover, since  $\mathbb{P}[X \in S(X^*)] = 1$  (because the measure produced by  $X$  is absolutely continuous with respect to that produced by  $X^*$ ), equality in (3.3) implies the equality in (3.2). The converse is not always true, as the following example shows.

**Example 3** Let  $X$  be uniformly distributed over  $(-2, -1) \cup (1, 2)$  and

$$g(x) = xI(|x| > 1) + x^3I(|x| \leq 1).$$

Then  $\mu = 0$ ,  $\sigma^2 = 7/3$  and  $X^*$  has the density  $f^*(x) = (3/28)(\min\{4 - x^2, 3\})^+$ . It follows that  $\text{Var}[g(X)] = \mathbb{E}[(g(X) - g(\mu))^2] = \sigma^2$  and  $\sigma^2 \mathbb{E}^2[g'(X^*)] = \text{Var}[g(X)] < \sigma^2 \mathbb{E}[(g'(X^*))^2] = 53/15$ . This shows that  $g$  is 'linear' with respect to the measure produced by  $X$ , and it is 'non-linear' with respect to the measure produced by  $X^*$ .



This example hinges on the fact that  $S(X)$  fails to be an interval. If, however,  $S(X)$  is a (finite or infinite) interval, the known upper and lower bounds for the variance of  $g(X)$  take the form (see [3], [4])

$$\sigma^2 \mathbb{E}^2[w(X)g'(X)] \leq \text{Var}[g(X)] \leq \sigma^2 \mathbb{E}[w(X)(g'(X))^2], \quad (3.4)$$

for some nonnegative function  $w$  defined on  $S(X)$  (in fact,  $w = f^*/f$ ), where both equalities hold iff  $g$  is linear on  $S(X)$ , provided that  $\mathbb{E}[w(X)(g'(X))^2] < \infty$ .

The following result shows the equivalence between the variance bounds and the covariance identity.

**Theorem 3.1** *Assume that for the a.c. r.v.  $X$  with finite variance  $\sigma^2$ , the r.v.  $Y$  satisfies one of the following.*

(i) *For every a.c. function  $g$  with derivative  $g'$ ,*

$$\text{Var}[g(X)] \leq \sigma^2 \mathbb{E}[(g'(Y))^2].$$

(ii) *For every a.c. function  $g$  with derivative  $g'$  such that  $\mathbb{E}|g'(Y)| < \infty$ ,*

$$\text{Var}[g(X)] \geq \sigma^2 \mathbb{E}^2[g'(Y)].$$

Then  $Y \stackrel{d}{=} X^*$ .

**Proof:** Assume that (i) holds. Let  $h'$  be any (measurable) bounded function and consider the a.c. function  $g(x) = x + \lambda h(x)$ , where  $\lambda$  is an arbitrary constant and  $h$  an indefinite integral of  $h'$ . It follows that  $g$  is a.c. with bounded derivative  $g' = 1 + \lambda h'$ . Thus,  $\text{Var}[g(X)] < \infty$ ,  $\text{Var}[h(X)] < \infty$  and

$$\begin{aligned} \text{Var}[g(X)] &= \sigma^2 + \lambda^2 \text{Var}[h(X)] + 2\lambda \text{Cov}[X, h(X)] \\ &\leq \sigma^2 \left( 1 + \lambda^2 \mathbb{E}[(h'(Y))^2] + 2\lambda \mathbb{E}[h'(Y)] \right). \end{aligned}$$

Therefore, by using standard arguments (see [4]), the quadratic  $\theta\lambda^2 + 2\delta\lambda$  (where  $\theta = \sigma^2 \mathbb{E}[(h'(Y))^2] - \text{Var}[h(X)] \geq 0$ ,  $\delta = \sigma^2 \mathbb{E}[h'(Y)] - \text{Cov}[X, h(X)]$ ) is nonnegative for all  $\lambda$ , and thus  $\delta = 0$ . Hence, taking into account (2.2) we conclude that for any bounded function  $H$ ,

$$\mathbb{E}[H(Y)] = \mathbb{E}[H(X^*)]$$

and the result follows. The same arguments apply to (ii).

Finally, by using similar arguments and the results of Theorem 2.1, a converse of Theorem 3.1 and Lemma 2.1(iii) can be easily established:

**Theorem 3.2** *Let  $Y$  be an arbitrary r.v. Then,  $Y$  has a unimodal a.c. density  $h$  such that the function  $h'(x)/(m-x)$  is integrable for some mode  $m$  of  $h$  iff there exists some a.c. r.v.  $X$  with finite second moment such that any one of the following holds.*

(i) For every a.c. function  $g$  with derivative  $g'$  such that  $\mathbb{E}|g'(Y)| < \infty$ ,

$$\text{Cov}[X, g(X)] = \text{Var}[X] \mathbb{E}[g'(Y)].$$

(ii) For every a.c. function  $g$  with derivative  $g'$ ,

$$\text{Var}[g(X)] \leq \text{Var}[X] \mathbb{E}[(g'(Y))^2].$$

(iii) For every a.c. function  $g$  with derivative  $g'$  such that  $\mathbb{E}|g'(Y)| < \infty$ ,

$$\text{Var}[g(X)] \geq \text{Var}[X] \mathbb{E}^2[g'(Y)].$$

(iv)  $X^* \stackrel{d}{=} Y$ .

Furthermore,  $X$  is unique iff the mode of  $h$  is unique.

Theorems 3.1 and 3.2 characterize the r.v.'s  $X$  and  $Y$  which admit *Poincaré type inequalities* or *differential inequalities* (c.f. [9], [18]). It should be noted that in the preceding inequalities the constant equals  $\text{Var}[X]$ , and this implies that the equality is attained for ‘linear’  $g$ . There are, however, other kinds of Poincaré type inequalities where the constant does not equal  $\text{Var}[X]$ . The following example is relative to this subject.

**Example 4** Let  $X$  be uniformly distributed over  $(0, 1)$  and  $Y$  have the density  $h(x) = (3/4)(1 - x^2)I(|x| < 1)$ . It follows that for every a.c. function  $g$ ,

$$\text{Var}[g(X)] \leq \frac{1}{12} \mathbb{E}[(g'(X^*))^2] = \frac{1}{2} \int_0^1 x(1-x)(g'(x))^2 dx \leq c \mathbb{E}[(g'(Y))^2],$$

for some constant  $c \leq 1/3$ , since  $(1/2)x(1-x)I(0 < x < 1) \leq (1/3)h(x)$  for  $x \in (-1, 1)$ . On the other hand, an application to the function  $g(x) = x^2$  shows that  $c \geq 1/9 > 1/12 = \text{Var}[X]$ .

## References

- [1] Cacoullos, T. (1982). On upper and lower bounds for the variance of a function of a random variable. *Ann. Probab.* **10** 799–809.
- [2] Cacoullos, T., Papadatos, N. and Papathanasiou, V. (1997). Variance inequalities for covariance kernels and applications to central limit theorems. *Theory Probab. Appl.* **42** 195–201.
- [3] Cacoullos, T. and Papathanasiou, V. (1985). On upper bounds for the variance of functions of random variables. *Statist. Probab. Lett.* **3** 175–184.
- [4] Cacoullos, T. and Papathanasiou, V. (1989). Characterizations of distributions by variance bounds. *Statist. Probab. Lett.* **7** 351–356.

- [5] Cacoullos, T. and Papathanasiou, V. (1995). A generalization of covariance identity and related characterizations. *Math. Methods Statist.* **4** 106–113.
- [6] Cacoullos, T., Papathanasiou, V. and Utev, S. (1992). Another characterization of the normal law and a proof of the central limit theorem connected with it. *Theory Probab. Appl.* **37** 648–657 (in Russian).
- [7] Cacoullos, T., Papathanasiou, V. and Utev, S. (1994). Variational inequalities with examples and an application to the central limit theorem. *Ann. Probab.* **22** 1607–1618.
- [8] Chen, L.H.Y. (1982). An inequality for the multivariate normal distribution. *J. Multivariate Anal.* **12** 306–315.
- [9] Chen, L.H.Y. (1988). The central limit theorem and Poincaré type inequalities. *Ann. Probab.* **16** 300–304.
- [10] Chernoff, H. (1981). A note on an inequality involving the normal distribution. *Ann. Probab.* **9** 533–535.
- [11] Goldstein, L. and Reinert, G. (1997). Stein’s method and the Zero Bias Transformation with application to simple random sampling. *Ann. Appl. Probab.* **7** 935–952.
- [12] Hudson, H.M. (1978). A natural identity for exponential families with applications to multiparameter estimation. *Ann. Statist.* **6** 473–484.
- [13] Liu, J.S. (1994). Siegel’s formula via Stein’s identities. *Statist. Probab. Lett.* **21** 247–251.
- [14] Papathanasiou, V. (1990). Characterizations of multidimensional exponential families by Cacoullos-type inequalities. *J. Multivariate Anal.* **35** 102–107.
- [15] Papathanasiou, V. (1993). Some characteristic properties of the Fisher information matrix via Cacoullos-type inequalities. *J. Multivariate Anal.* **44** 256–265.
- [16] Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **2** 583–602. Univ. California Press, Berkeley.
- [17] Stein, C. (1981). Estimation of the mean of a multivariate normal distribution. *Ann. Statist.* **9** 1135–1151.
- [18] Vitale, R.A. (1989). A differential version of the Efron-Stein inequality: Bounding the variance of an infinitely divisible variable. *Statist. Probab. Lett.* **7** 105–112.