Some Applications of Covariance Identities and Inequalities to Functions of Multivariate Normal Variables

Christian HOUDRÉ*

We apply general covariance identities and inequalities to some functions of multivariate normal variables. We recover, in particular, a recent covariance identity due to Siegel and provide simple estimates on the variance of order statistics. We also present some computations.

KEY WORDS: Financial futures; Order statistics; Stein's identity.

1. INTRODUCTION

The main purpose of this article is to point out some consequences and particular cases of the covariance identities and inequalities obtained by Houdré and Pérez-Abreu (1995) and Houdré and Kagan (1995). The framework of these papers was rather general and abstract, and it is felt that isolating some important particular cases is worthwhile. This is particularly true in view of potential consequences in statistical theory, as illustrated by results of Siegel (1993). Motivated by the analysis of futures contracts, Siegel found a simple expression for $\text{cov}[X_1, \min(X_1, \ldots, X_n)]$, where $(X_1, \ldots, X_n)$ is multivariate normal. This covariance is the numerator of the hedge ratio $R_{\min} = \text{cov}[X_1, \min(X_1, \ldots, X_n)]/\text{var}[\min(X_1, \ldots, X_n)]$, which quantifies the amount to trade to minimize the financial risk. Aside from recovering Siegel's result, our approach also provides a simple upper bound for $\text{var}[\min(X_1, \ldots, X_n)]$, and hence a lower bound for the absolute value of $R_{\min}$. Although an exact expression for $\text{var}[\min(X_1, \ldots, X_n)]$ is known (see Afonja 1972), the complexity of this expression makes it difficult to handle, and our simple upper bound will provide a useful estimate. Another advantage of our approach comes from its generality. It applies not only to the minimum but also to any order statistic, thus providing, for example, identities involving the median price in futures contracts or estimates on $R_{\text{med}} = \text{cov}[X_1, \text{med}(X_1, \ldots, X_n)]/\text{var}[\text{med}(X_1, \ldots, X_n)]$. Moreover, we also give upper estimates for the variance of order statistics of a multivariate normal vector $(X_1, \ldots, X_n)$ which should prove useful in other statistical contexts.

To focus on the most important distribution, we deal here only with normal variables. Similar results hold for the Poisson distribution or the uniform distribution on the circle. In fact, properly modified, these results hold for any infinitely divisible distribution, because the main identity (eq. (1) in Sec. 2) on which they are based has a version for such distributions (see Houdré, Pérez-Abreu, and Surgailis 1994).

After submission of this article, we received a technical report (Liu 1994) in which Siegel's result was obtained and extended via Stein's identity. This technical report and the present article have some overlap. Also, Anderson (1993) presented a multivariate version of Stein's identity from which Siegel's formula follows. The identity presented by Anderson is a particular case of Equation (1) when Φ (defined in Sec. 2) is a linear function. Also, Rinott and Samuel-Cahn (1994), following Siegel's method, established the covariance result where the minimum is replaced by any order statistic. As already indicated, our method also gives these results for order statistics.

2. BACKGROUND

Let $\mathbf{X} = (X_1, \ldots, X_n) \sim N(\mu, \Sigma)$ be an n-dimensional real normal vector with covariance matrix $\Sigma = (\sigma_{i,j})_{i,j=1,\ldots,n}$. Let $\Phi, \Psi : \mathbf{R}^n \to \mathbf{R}^p$ have a sufficient number of square-integrable (with respect to the Gaussian measure of covariance Σ) derivatives and let $\text{COV}(\Phi(\mathbf{X}), \Psi(\mathbf{X})) = (\text{cov}(\varphi_i(\mathbf{X}), \psi_j(\mathbf{X})))_{i,j=1,\ldots,p}$, denote the covariance matrix of $\Phi(\mathbf{X}) = (\varphi_1(\mathbf{X}), \ldots, \varphi_p(\mathbf{X}))^t$ and of $\Psi(\mathbf{X}) = (\psi_1(\mathbf{X}), \ldots, \psi_p(\mathbf{X}))^t$, where t denotes transpose. (Throughout, COV and VAR are used for matrix arguments and cov and var are reserved for scalar ones.) Let ∇ be the gradient operator; that is, if $\varphi : \mathbf{R}^n \to \mathbf{R}$, then $\nabla\varphi = (\partial\varphi/\partial x_1, \ldots, \partial\varphi/\partial x_n)$, whereas for $\Phi = (\varphi_1, \ldots, \varphi_p)^t : \mathbf{R}^n \to \mathbf{R}^p$, $\nabla\Phi = (\nabla\varphi_1, \ldots, \nabla\varphi_p)^t$. Finally, for $k \geq 2$, let ∇^k be the iterated gradient operator; that is, for $\varphi : \mathbf{R}^n \to \mathbf{R}$, $\nabla^k\varphi$ is the n^k row vector $\nabla^k\varphi = (\partial\nabla^{k-1}\varphi/\partial x_1, \ldots, \partial\nabla^{k-1}\varphi/\partial x_n)$, whereas for $\Phi = (\varphi_1, \ldots, \varphi_p)^t : \mathbf{R}^n \to \mathbf{R}^p$, $\nabla^k\Phi$ is the $p \times n^k$ matrix $\nabla^k\Phi = (\nabla^k\varphi_1, \ldots, \nabla^k\varphi_p)^t$. Then Houdré and Pérez-Abreu (1995, sec. 3) showed that

$$\text{COV}(\Phi(\mathbf{X}), \Psi(\mathbf{X})) = \sum_{k=1}^{N} \frac{(-1)^{k+1}}{k!} E\{(\nabla^k\Phi(\mathbf{X}))\Sigma^{\otimes k}(\nabla^k\Psi(\mathbf{X}))^t\} + R_N, \quad N \geq 1, \quad (1)$$

where E is expectation and $\Sigma^{\otimes k}$ is the kth Kronecker product of Σ with itself. The form of the remainder term R_N in (1) is not important for our present considerations; let us just

* Christian Houdré is Assistant Professor, School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332. This research was supported in part by a National Science Foundation (NSF) Mathematical Sciences Postdoctoral Fellowship while at Department of Statistics, Stanford University, Stanford, CA 94305 and by an NSF-NATO Postdoctoral Fellowship while at CEREMADE, Université Paris-Dauphine, 75775 Paris, France, and at CERMA, Ecole Nationale des Ponts et Chaussées, Noisy Le Grand, France. The author thanks an associate editor and the editor, who greatly helped to make the exposition friendlier to the general JASA readership, and Y. L. Tong for helpful editorial comments and a reference.

Journal of the American Statistical Association
September 1995, Vol. 90, No. 431, Theory and Methods

say that it is an expression involving the gradients of order $N+1$.

In particular, from (1) and for $\Phi = \Psi$, we get the following inequalities (we are comparing $p \times p$ matrices, and for such matrices, $\mathbf{A} \le \mathbf{B}$ will mean that the difference $\mathbf{B} - \mathbf{A}$ is a positive semidefinite matrix):

$$\sum_{k=1}^{2N} \frac{(-1)^{k+1}}{k!} E\{(\nabla^k \Phi(\mathbf{X}))\mathbf{\Sigma}^{\otimes k}(\nabla^k \Phi(\mathbf{X}))^t\}$$

$$\le \mathrm{VAR}\{\Phi(\mathbf{X})\}$$

$$\le \sum_{k=1}^{2N-1} \frac{(-1)^{k+1}}{k!} E\{(\nabla^k \Phi(\mathbf{X}))\mathbf{\Sigma}^{\otimes k}(\nabla^k \Phi(\mathbf{X}))^t\}. \quad (2)$$

The inequalities (2) were already obtained in the univariate case by Houdré and Kagan (1995), whereas the right side of (2), for $N = 1$ and for univariate variables, can be found, for example, in the work of Chernoff (1981).

3. DEVELOPMENT

Inequality 1 of this section is a consequence of (2) when $N = 1$. Before proving it, we make a remark and set a convention. In what follows, the maximum is only a.s. differentiable, since Lipschitz, but an approximation argument will give the right side of (2), for $N = 1$ and for Lipschitz functions. Furthermore, we also assume that the X_i, $i = 1, \ldots, n$, are distinct. If such is not the case, then $P\{\max \mathbf{X} = X_i\}$ should be divided by the number of elements in the various distinct classes. A similar remark and convention (properly modified) apply throughout the article, when needed.

Inequality 1. Let $\mathbf{X} \sim N(\mu, \mathbf{\Sigma})$ with $\mathbf{\Sigma} = (\sigma_{i,j})_{i,j=1,\ldots,n}$. Then

$$\mathrm{var}\{\max_{1\le i\le n} X_i\} \le \sum_{i=1}^{n} \mathrm{var}\{X_i\} P\{\max \mathbf{X} = X_i\}$$

$$\le \max_{1\le i\le n} \mathrm{var}\{X_i\}. \quad (3)$$

Proof. The right inequality is clear. To prove the left one, we apply (2) to the function $\Phi : \mathbf{R}^n \to \mathbf{R}$ defined via $\Phi(x_1, \ldots, x_n) = \max(x_1, \ldots, x_n)$. Because $\max(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i I_{A_i}(x)$, where $A_i = \{\max(x_1, \ldots, x_n) = x_i\}$ and $x = (x_1, \ldots, x_n)$, then $\nabla \Phi(\mathbf{X}) = (I_{A_1}(\mathbf{X}), \ldots, I_{A_n}(\mathbf{X}))$. Thus by the right side of (2) with $N = 1$,

$$\mathrm{var}\{\max_{1\le i\le n} X_i\} \le E\{(I_{A_1}(\mathbf{X}), \ldots, I_{A_n}(\mathbf{X})) \times \mathbf{\Sigma}(I_{A_1}(\mathbf{X}), \ldots, I_{A_n}(\mathbf{X}))^t\}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_{i,j} E\{I_{A_i}(\mathbf{X}) I_{A_j}(\mathbf{X})\}$$

$$= \sum_{i=1}^{n} \sigma_{i,i} P\{\max \mathbf{X} = X_i\}.$$

The inequality $\mathrm{var}\{\max_{1\le i\le n} X_i\} \le \max_{1\le i\le n}\mathrm{var}\{X_i\}$ is known (see, for example, Cirel'son, Ibragimov, and Sudakov 1976). Only the first inequality in (3) appears to be new. The estimates obtained in (3) are of course mainly interesting for dependent variables, and they extend (via an approximation argument) to Gaussian processes $\{X_t\}_{t\in T}$, where T is a metric space with a dense countable subset. Estimates of the type (3) continue to hold for any multivariate vector $\mathbf{X}$ for which $\mathrm{var}\{\Phi(\mathbf{X})\} \le CE\{(\nabla\Phi(\mathbf{X}))\mathbf{\Sigma}(\nabla\Phi(\mathbf{X}))^t\}$, for some constant C independent of $\Phi : \mathbf{R}^n \to \mathbf{R}$.

The inequalities just obtained also provide a "law of large numbers" for the maximum of normal variables. If $\{X_n\}_{n\in\mathbf{N}}$ is a sequence of univariate normal variables such that $\lim_{n\to+\infty} \sum_{i=1}^{n} \mathrm{var}\{X_i\} P\{\max_{1\le j\le n} X_j = X_i\} = 0$, then $\lim_{n\to+\infty}(\max_{1\le i\le n} X_i - E\max_{1\le i\le n} X_i) = 0$ in probability. This is in particular the case if $\lim_{n\to+\infty}\max_{1\le i\le n}\mathrm{var}\{X_i\} = 0$.

If the X_i are replaced by $|X_i|$, which by our convention are assumed distinct, then the following inequality is obtained.

Inequality 2. Let $\mathbf{X} \sim N(\mu, \mathbf{\Sigma})$. Then

$$\mathrm{var}\{\max_{1\le i\le n} |X_i|\} \le \sum_{i=1}^{n} \mathrm{var}\{X_i\} P\{\max|\mathbf{X}| = |X_i|\}$$

$$\le \max_{1\le i\le n} \mathrm{var}\{X_i\}. \quad (4)$$

Proof. Because $\max|\mathbf{X}| = \max_{1\le i\le n}|X_i| = \sum_{i=1}^{n} X_i I_{B_i}(\mathbf{X}) - \sum_{i=1}^{n} X_i I_{C_i}(\mathbf{X})$, where B_i and C_i are the events $\{X_i \ge 0, \max|\mathbf{X}| = |X_i|\}$ and $\{X_i < 0, \max|\mathbf{X}| = |X_i|\}$, $i = 1, \ldots, n$, we have $\nabla \max_{1\le i\le n}|X_i| = (I_{B_1}(\mathbf{X}) - I_{C_1}(\mathbf{X}), \ldots, I_{B_n}(\mathbf{X}) - I_{C_n}(\mathbf{X}))$. By (2), it follows that

$$\mathrm{var}\{\max_{1\le i\le n} |X_i|\}$$

$$\le \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_{i,j} E\{(I_{B_i}(\mathbf{X}) - I_{C_i}(\mathbf{X})) \times (I_{B_j}(\mathbf{X}) - I_{C_j}(\mathbf{X}))\}$$

$$= \sum_{i=1}^{n} \sigma_{i,i} E\{(I_{B_i}(\mathbf{X}) - I_{C_i}(\mathbf{X}))^2\}$$

$$= \sum_{i=1}^{n} \sigma_{i,i} P\{\max|\mathbf{X}| = |X_i|\}$$

$$\le \max_{1\le i\le n} \mathrm{var}\{X_i\}.$$

Of course, if $\mathbf{L} : \mathbf{R}^n \to \mathbf{R}^n$ is a nonsingular linear transformation, then $\mathbf{LX} \sim N(\mathbf{L}\mu, \mathbf{L\Sigma L}^t)$ and

$$\mathrm{var}\{\max_{1\le i\le n} (\mathbf{LX})_i\}$$

$$\le \sum_{i=1}^{n} \mathrm{var}\{(\mathbf{LX})_i\} P\{\max \mathbf{LX} = (\mathbf{LX})_i\}$$

$$\le \max_{1\le i\le n} \mathrm{var}\{(\mathbf{LX})_i\}. \quad (5)$$

In particular, if $S_i = \sum_{k=1}^{i} X_k$ and if $\mathbf{S} = (S_1, \ldots, S_n)$, then $\mathrm{var}\{\max_{1\le i\le n} S_i\} \le \sum_{i=1}^{n} \mathrm{var}\{S_i\} P\{\max \mathbf{S} = S_i\} \le \max_{1\le i\le n}\mathrm{var}\{S_i\}$.

Contrary to the normal distribution, the Poisson distribution is not preserved under linear transformations. But inequalities similar to (5) remain true and can be proved directly by the methods of proof of Inequalities 1 and 2, combined with the version of (2) for Poisson variables obtained by Houdré and Pérez-Abreu (1995).

Another instance where (2) is potentially useful is to give variance estimates on nonlinear transformations of multivariate normal variables (e.g., order statistics); for example, for the median. Indeed, let $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$, where $X_{(i)}$ is the ith order statistic of the random sample $X_1, \ldots, X_n$. Assume that n is odd and let med $\mathbf{X} = X_{[(n+1)/2]}$. Then the following holds: $\operatorname{var}\{\text{med } \mathbf{X}\} \le \sum_{i=1}^n \operatorname{var}\{X_i\} P\{\text{med } \mathbf{X} = X_i\}$. (To prove this last inequality, it is enough to write med $\mathbf{X}$ in the form med $\mathbf{X} = \sum_{i=1}^n X_i I_{\{\text{med}\mathbf{X} = X_i\}}$ and to proceed as in the proofs of Inequalities 1 and 2.) In that context, we can talk about the "universal domination" of the maximal variance in that $\operatorname{var}\{X_{(i)}\} \le \max_{1\le i\le n} \operatorname{var}\{X_i\}$.

More refined inequalities can be obtained. For instance, and still assuming that the $|X_i|$ are distinct, we have the following.

Inequality 3. Let $\mathbf{X} \sim N(\mathbf{0}, \mathbf{\Sigma})$. Then

$$\operatorname{var} \sum_{i=1}^n |X_i| \le \frac{2}{\pi} \sum_{i=1}^n \sum_{j=1}^n \sigma_{i,j} \arcsin\left\{\frac{\sigma_{i,j}}{\sqrt{\sigma_{i,i}\sigma_{j,j}}}\right\}. \quad (6)$$

Proof. $\sum_{i=1}^n |X_i| = \sum_{i=1}^n X_i I_{\{X_i \ge 0\}} - \sum_{i=1}^n X_i I_{\{X_i<0\}}$. Hence

$$\nabla \Phi(\mathbf{X}) = (I_{\{X_1\ge 0\}} - I_{\{X_1<0\}}, \ldots, I_{\{X_n \ge 0\}} - I_{\{X_n<0\}})$$

and

$$\begin{aligned}
&\operatorname{var} \sum_{i=1}^n |X_i| \\
&\le \sum_{i=1}^n \sum_{j=1}^n \sigma_{i,j} E\{(I_{\{X_i\ge 0\}} - I_{\{X_i<0\}})(i_{\{X_j \ge 0\}} - I_{\{X_j \le 0\}})\} \\
&= \sum_{i=1}^n \sum_{j=1}^n \sigma_{i,j}(2E\{I_{\{X_i\ge 0\}} I_{\{X_j \ge 0\}}\} \\
&\qquad + 2E\{I_{\{X_i \le 0\}} I_{\{X_j \le 0\}}\} - 1) \\
&= \sum_{i=1}^n \sum_{j=1}^n \sigma_{i,j}\left(\frac{1}{2} + \frac{1}{\pi} \arcsin\left\{\frac{\sigma_{i,j}}{\sqrt{\sigma_{i,i}\sigma_{j,j}}}\right\}\right. \\
&\qquad \left. + \frac{1}{2} + \frac{1}{\pi} \arcsin\left\{\frac{\sigma_{i,j}}{\sqrt{\sigma_{i,i}\sigma_{j,j}}}\right\} - 1\right) \\
&= \frac{2}{\pi} \sum_{i=1}^n \sum_{j=1}^n \sigma_{i,j} \arcsin\left\{\frac{\sigma_{i,j}}{\sqrt{\sigma_{i,i}\sigma_{j,j}}}\right\}, \qquad (7)
\end{aligned}$$

where in obtaining the second equality of (7) we used a so-called orthant probability formula (see, for example, Johnson and Kotz 1972, pp. 93–95).

To this point, we have dealt only with the simplest case of (2); that is, with one derivative and the upper bound estimate. We now present an identity where a bit more of the complexity of (1) is used. In what follows, we again assume that the X_i, $i = 1, \ldots, n$ are distinct and that n is odd, and we note that the median is a Lipschitz function.

Identity 4. Let $\mathbf{X} \sim N(\boldsymbol{\mu}, \mathbf{\Sigma})$. Then,

$$\operatorname{cov}\left(\sum_{i=1}^n X_i, \text{med } \mathbf{X}\right) = \sum_{i=1}^n \sum_{j=1}^n \sigma_{i,j} P\{\text{med } \mathbf{X} = X_j\}. \quad (8)$$

Proof. Because $\Phi(\mathbf{X}) = \sum_{i=1}^n X_i$, and because $\Psi(\mathbf{X}) = \text{med } \mathbf{X} = \sum_{i=1}^n X_i I_{\{\text{med}\mathbf{X} = X_i\}}$ is a Lipschitz function, we have

$$\nabla \Phi(\mathbf{X}) = (1, \ldots, 1)$$

and

$$\nabla \Psi(\mathbf{X}) = (I_{\{\text{med}\mathbf{X} = X_1\}}, \ldots, I_{\{\text{med}\mathbf{X} = X_n\}}).$$

Hence, applying (1) and taking into account the fact that the higher derivatives of Φ are zero, we get

$$\begin{aligned}
\operatorname{cov}(\Phi(\mathbf{X}), \Psi(\mathbf{X})) &= E\{(\nabla\Phi(\mathbf{X}))\mathbf{\Sigma}(\nabla\Psi(\mathbf{X}))^t\} \\
&= E\left\{\sum_{i=1}^n \sum_{j=1}^n \sigma_{i,j} I_{\{\text{med}\mathbf{X} = X_j\}}\right\} \\
&= \sum_{i=1}^n \sum_{j=1}^n \sigma_{i,j} P\{\text{med } \mathbf{X} = X_j\}. \qquad (9)
\end{aligned}$$

As already noted by Houdré and Pérez-Abreu (1995), the form of the identity used in (9) is Stein's identity. On the other hand, applying this method to $\Phi(\mathbf{X}) = X_1$ and $\Psi(\mathbf{X}) = \min_{1\le i\le n} X_i$ does yields the results of Siegel (1993)—namely, $\operatorname{cov}(X_1, \min \mathbf{X}) = \sum_{i=1}^n \sigma_{1,i} P\{\min \mathbf{X} = X_i\}$. Moreover, in view of the method of proof of inequality (1), we get

$$\begin{aligned}
|R_{\min}| &= \frac{|\operatorname{cov}(X_1, \min \mathbf{X})|}{\operatorname{var}\{\min \mathbf{X}\}} \\
&\ge \frac{|\sum_{i=1}^n \sigma_{1,i} P\{\min \mathbf{X} = X_i\}|}{\sum_{i=1}^n \sigma_{i,i} P\{\min \mathbf{X} = X_i\}}. \qquad (10)
\end{aligned}$$

By now, it is clear that estimates similar to (10) also hold when the minimum is replaced by any order statistic or by certain functions of order statistics. For example, for n odd and if the median is defined as previously, we have

$$\begin{aligned}
|R_{\text{med}}| &= \frac{|\operatorname{cov}(X_1, \text{med } \mathbf{X})|}{\operatorname{var}\{\text{med } \mathbf{X}\}} \\
&\ge \frac{|\sum_{i=1}^n \sigma_{1,i} P\{\text{med } \mathbf{X} = X_i\}|}{\sum_{i=1}^n \sigma_{i,i} P\{\text{med } \mathbf{X} = X_i\}}. \qquad (11)
\end{aligned}$$

On the other hand, if n is even, let med $\mathbf{X} = \{X_{(n/2)} + X_{[(n/2)+1]}\}/2$. Then med $\mathbf{X}$ is no longer an order statistic but rather a function of order statistics. However, still writing $\Psi(\mathbf{X}) = \text{med } \mathbf{X} = \frac{1}{2}\sum_{i=1}^n X_i\{I_{\{X_{(n/2)} = X_i\}} + I_{\{X_{[(n/2)+1]} = X_i\}}\}$, we get $\nabla\Psi(\mathbf{X}) = \frac{1}{2}(I_{\{X_{(n/2)} = X_1\}} + I_{\{X_{[(n/2)+1]} = X_1\}}, \ldots, I_{\{X_{(n/2)} = X_n\}} + I_{\{X_{[(n/2)+1]} = X_n\}})$. Thus via Equation (1), $\operatorname{cov}(X_1, \text{med } \mathbf{X}) = \frac{1}{2}\sum_{i=1}^n \sigma_{1,i}\{P\{X_{(n/2)} = X_i\} + P\{X_{[(n/2)+1]} = X_i\}\}$, and via (2),

Table 1. Comparisons of the Right Side Inequalities in (2) for N = 1 and N > 1 and Some Specific Functions

$f(x)$	$E\|f'(X)\|^2$	$\sum_{k=1}^{2N-1} \frac{(-1)^{k+1}E\|f^{(k)}(X)\|^2}{k!}$
x^3	27	15
x^5	2,625	945
ax^r, r odd	$a \cdot r^2 \cdot 1 \cdot 3 \cdot 5 \cdots (2r-3)$	$a \cdot \sum_{k=1}^{r} \frac{(-1)^{k+1}[P(N,k)]^2(1 \cdot 3 \cdot 5 \cdots (2(N-k)-1))}{k!}$
$\Phi(x)$	$\frac{\sqrt{3}}{6\pi}$	$\approx \frac{\sqrt{3}}{6\pi}(.8915)$
$\sin x$	$\frac{1+e^{-2}}{2} \approx .5676$	$\frac{1}{2}\sum_{k=1}^{\infty} \frac{e^{-2}+(-1)^{k+1}}{k!} \approx .4323$

$$|R_{\text{med}}| = \frac{|\text{cov}(X_1, \text{med } \mathbf{X})|}{\text{var}\{\text{med } \mathbf{X}\}} \geq \frac{|\sum_{i=1}^{n} \sigma_{1,i}\{P\{X_{(n/2)} = X_i\} + P\{X_{[(n/2)+1]} = X_i\}\}|}{D}, \quad (12)$$

where the denominator D is equal to

$$\sum_{i=1}^{n} \sigma_{i,i}\{P\{X_{(n/2)} = X_i\} + P\{X_{[(n/2)+1]} = X_i\}\} + 2\sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n} \sigma_{i,j} P\{X_{(n/2)} = X_i, X_{[(n/2)+1]} = X_j\}.$$

It is also clear that expressions similar to (12) continue to hold for, say, the range $\{\max \mathbf{X} - \min \mathbf{X}\}$ or the midrange $\{\max \mathbf{X} + \min \mathbf{X}\}/2$ of the random sample.

We next present some computations comparing the simplest cases of the inequality (2). Table 1 (kindly provided by Michael Hernández) lists some functions $f\colon \mathbf{R} \to \mathbf{R}$ and the upper bounds on the variances of these functions when X is a univariate standard normal variable. The table also compares the estimates $E|f'(X)|^2$ and

$$\sum_{k=1}^{2N-1} \frac{(-1)^{k+1}E|f^{(k)}(X)|^2}{k!}.$$

In general, the second estimate gives a sharper upper bound. Of course, for polynomials, and N large enough, the variance is completely recovered. In Table 1, $\Phi(x)$ is the standard normal distribution.

4. CONCLUSION

We have presented a method to estimate the variance or the covariance of functions of multivariate normal variables. In particular, our method gives rather simple estimates for the variances of order statistics. As an example, we also showed how our method yields the covariance formula of Siegel (1993).

[Received November 1993. Revised December 1994.]

REFERENCES

Afonja, B. (1972), "The Moments of the Maximum of Correlated Normal and *t* Variates," *Journal of the Royal Statistical Society,* Ser. B, 34, 251–262.

Anderson, C. L. (1993), "Extension of Surprising Covariances," *Journal of the American Statistical Association,* 88, 1478.

Chernoff, H. (1981), "A Note on an Inequality Involving the Normal Distribution," *The Annals of Probability,* 9, 533–535.

Cirel'son B. S., Ibragimov, I. A., and Sudakov, V. N. (1976), "Norms of Gaussian Sample Functions," in *Proceedings of the 3rd Japan-USSR Symposium on Probability Theory,* Lecture Notes in Mathematics 550, New York: Springer-Verlag, pp. 20–41.

Houdré, C., Pérez-Abreu, V., and Surgailis, D. (1994), "An Interpolation Principle for Functions of Infinitely Divisible Variables," preprint.

Houdré, C., and Kagan, A., (1995), "Variance Inequalities for Functions of Gaussian Variables," *The Journal of Theoretical Probability,* 8, 23–30.

Houdré, C., and Pérez-Abreu, V. (1995), "Covariance Identities and Inequalities for Functionals on Wiener and Poisson Spaces," *The Annals of Probability,* Vol. 23.

Johnson, N., and Kotz, S. (1972), *Distributions in Statistics: Continuous Multivariate Distributions 1,* New York: John Wiley.

Liu, J. S. (1994), "Siegel's Formula via Stein's Identities," preprint.

Siegel, A. (1993), "A Surprising Covariance Involving the Minimum of Multivariate Normal Variables," *Journal of the American Statistical Association,* 88, 77–80.

Rinott, Y., and Samuel-Cahn, E. (1994), "Covariance Between Variables and Their Order Statistics for Multivariate Normal Variables," *Statistics and Probability Letters,* 21, 153–155.