A CONSTRUCTION OF INCOMPLETE SUFFICIENT UNBIASED ESTIMATORS OF THE NORMAL CORRELATION COEFFICIENT

Kōsei Iwase* and Noriaki Setô*

A class of unbiased estimators of the correlation coefficient of a bivariate normal distribution with known means and variances is constructed in terms of the incomplete sufficient statistic, and the variances of these estimators are obtained analytically for any sample size. Moreover, some properties of the estimators in the case that the sample size is equal to unity are investigated both analytically and numerically, and then an unbiased estimator of the correlogram of a stationary Gaussian process with known mean and variance is proposed.

1. Introduction

This paper deals with the unbiased estimation, based on the incomplete sufficient statistic, of the correlation coefficient of the two variates having a bivariate normal distribution with known means and variances.

Olkin & Pratt [12] constructed the uniformly minimum variance unbiased estimators of the correlation coefficient in the cases that all of the five population parameters were unknown and that only the population variances were known to be equal. In both cases, as the sufficient statistic is complete, the unbiased estimator are uniquely determined. If the population variances are known, however, the statistic is not complete, so that there exist infinitely many unbiased estimators of the correlation coefficient based on the sufficient statistic.

Recently, a class of incomplete sufficient unbiased estimators of the correlation coefficient was obtained and some statistical properties were derived by one of the authors [11]. The class contains "the simplified estimator of \( \rho \)" which was suggested by Takahasi & Husimi [15] and discussed by Huzii [4, 5, 6], Sibuya [14], Iwase [9, 10] and Inagaki & Kondo [8]. This also contains other estimators due to Olkin & Pratt [12] and Pradhan & Sathe [13]. In this paper a new type of estimators which includes the above class is presented by using a hypergeometric function.

In the next section, the construction of the new type estimators is performed and, in Section 3, the variances of these estimators are derived analytically for any sample size. An outline of the derivation is described in Appendix. In Section 4, it is shown numerically that the subset of the class has uniformly smaller variances with respect to the population correlation coefficient than those of Iwase's class. Some properties of the estimators in the case that the sample size equals unity are considered both analytically and numerically, and then an unbiased estimator of the correlogram of a stationary Gaus-
sian process with known mean and variance is proposed. The final section is applied to some discussions.

2. Construction of unbiased estimators

We shall consider \( n \) pairs of mutually independent random variables \((X_1, Y_1), \ldots, (X_n, Y_n)\), and assume that each pair obeys the bivariate normal distribution \( N(\mu, \Sigma) \) with \( \mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \) and \( \Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \). It is known that the sufficient statistic is

\[
\left( \sum_{i=1}^{n} (X_i + Y_i), \sum_{i=1}^{n} X_i Y_i \right)
\]

for the construction of an unbiased estimator of the correlation parameter \( \rho \). It is also known that the family of the distributions specified by the above statistic is incomplete, so that the unbiased estimator cannot be uniquely determined. An alternative form of the sufficient statistic is \((S, T)\), where \( S \) and \( T \) are defined by

\[
S = \sum_{i=1}^{n} \left[ \frac{X_i + Y_i}{2} \right]^2, \quad T = \sum_{i=1}^{n} \left[ \frac{X_i Y_i}{2} \right]^2.
\]

The random variables \( S \) and \( T \) are independent, and \( 2S/(1+\rho) \) and \( 2T/(1-\rho) \) are distributed according to \( \chi^2 \)-distribution with \( n \) degrees of freedom.

Here we shall momentarily modify the above setting, and investigate the estimation problem in which \( S \) and \( T \) are independent and \( 2aS \) and \( 2bT \) obey the \( \chi^2 \)-distribution. For the original \( S \) and \( T \) in (2.1), \( a \) and \( b \) are equal to \( 1/(1+\rho) \) and \( 1/(1-\rho) \), respectively. We shall assume, however, for a while that \( a \) and \( b \) are two arbitrary (functionally independent) positive numbers. In this case the statistic \((S, T)\) is complete, and the expectation of \( F(S, T) \) is calculated to be

\[
E[F(S, T)] = \frac{1}{2\pi} \int_0^\infty \int_0^\infty F\left( \frac{s}{2a \cdot 2b} \right) \frac{1}{\Gamma\left( \frac{n}{2} \right)} (ab)^{n/2} t^{-n/2-1} e^{-s^2/4t - t^2} ds dt
\]

\[
= \frac{1}{\Gamma\left( \frac{n}{2} \right)} (ab)^{n/2} \int_0^\infty F(s, t) dt^{-n/2-1} e^{-as -bt} ds dt
\]

\[
= \phi(a, b).
\]

From this we have the following

**Lemma 1.** The uniformly minimum variance unbiased (UMVU) estimator \( F(S, T) \) of \( f(a, b) \) is given by

\[
F(S, T) = F^3\left( \frac{n}{2} \right) \langle ST \rangle^{-n/2+1} \mathcal{L}^{-1}\{ (ab)^{-n/2} f(a, b) \}(S, T),
\]

where \( \mathcal{L}^{-1} \) denotes the double inverse Laplace transform with respect to the variables \( a \) and \( b \).

In particular, using the formula 5.5 (40) in [2], we immediately find

**Lemma 2.** The UMVU estimators \( F_n(S, T; u) \) and \( G_n(S, T; u) \) for

\[
f(a, b; u) = \frac{1}{2} \left( \frac{1}{a} - \frac{1}{b} \right) \exp \left\{ -\left( \frac{1}{a} + \frac{1}{b} \right) \frac{u}{2} \right\}, \quad (u \geq 0)
\]
and
\[(2.5) \quad g(a, b; u) = \exp \left\{ -\frac{1}{a + b} \frac{u}{2} \right\}, \quad (u \geq 0)\]
respectively, are
\[(2.6) \quad F_n(S, T; u) = 2^{\left(n - 3/2\right)} \Gamma_{1/2}^2 \left( \frac{n}{2} \right) \left( ST \right)^{1-n/4} \exp \left( -\frac{u}{2} \right) \]
\[\times \left\{ \frac{1}{\sqrt{T}} J_{n/2}(\sqrt{2uS}) J_{n/2-1}(\sqrt{2uT}) - \frac{1}{\sqrt{S}} J_{n/2-1}(\sqrt{2uS}) J_{n/2}(\sqrt{2uT}) \right\}\]
and
\[(2.7) \quad G_n(S, T; u) = 2^{\left(n - 5/2\right)} \Gamma_{3/2}^2 \left( \frac{n}{2} \right) \left( ST \right)^{-\left(n - 3/2\right)} \exp \left( -\frac{u}{2} \right) \]
\[\times J_{n/2-1}(\sqrt{2uS}) J_{n/2-1}(\sqrt{2uT})\],
where \( J \) is a Bessel function of the first kind.

Now we return to our original situation, that is, \( S \) and \( T \) are given by (2.1). Because of the equalities \( a = 1/(1 + \rho) \) and \( b = 1/(1 - \rho) \), Lemma 2 implies that
\[(2.8) \quad E[F_n(S, T; u)] = \rho e^{-u}, \quad E[G_n(S, T; u)] = e^{-u}.
\]
This leads to

**Theorem 1.** Let \( \tilde{\rho}_n(q, r) \) be defined by
\[(2.9) \quad \tilde{\rho}_n(q, r) = \int_0^\infty F_n(S, T; u) q(u) \, du + \int_0^\infty G_n(S, T; u) r(u) \, du,
\]
then this is an unbiased estimator of \( \rho \), provided that \( q \) and \( r \) satisfy
\[(2.10) \quad \int_0^\infty q(u) e^{-u} \, du = 1, \quad \int_0^\infty r(u) e^{-u} \, du = 0
\]
and also provided that the order of the integration of taking the expectation and that appearing in (2.9) can be exchanged.

**Proof.** From the assumption in Theorem 1 and from (2.8), it follows that
\[E[\tilde{\rho}_n(q, r)] = \rho (\mathcal{L}q)(1) + (\mathcal{L}r)(1),\]
where \( \mathcal{L} \) denotes the (single) Laplace transformation. The condition (2.10) leads \( E[\tilde{\rho}_n(q, r)] = \rho \).

We have thus obtained a class of unbiased estimators of \( \rho \) depending on two arbitrary functions. To restrict this class, we hereafter set \( r \equiv 0 \). This amounts to requiring that the estimator should be an odd function with respect to the interchange of \( S \) and \( T \). By specializing the functional form of \( q \) to
\[(2.11) \quad q(u) = h(u; c, \kappa) = \frac{(1+c)^r}{\Gamma(\kappa)} u^{r-1} e^{-\kappa u}, \quad \kappa > 0, \quad c \geq 0,
\]
we obtain

**Theorem 2.** A random variable
\[(2.12) \quad \tilde{\rho}_n(c) = \frac{(1+c)^r}{\Gamma(\kappa)} \int_0^\infty F_n(S, T; u) u^{r-1} e^{-\kappa u} \, du\]
is an unbiased estimator of $\rho$ under the conditions

\[(2.13) \quad \kappa > 0 \quad \text{for } c > 0,\]

or

\[(2.14) \quad n + 1 > 2\kappa > 0 \quad \text{for } c = 0.\]

**Proof.** The conditions (2.13) and (2.14) assure the exchangeability of the integration in (2.12) and the $E$-operation, so that Theorem 1 together with the obvious equality $\langle \mathcal{L}k(\cdot ; c, \kappa) \rangle (1) = 1$ implies Theorem 2.

For a special case $c = 0$, the integration in (2.12) can be performed explicitly with the help of the formula 6.574.1 in [3]. The result is summarized in the following

**Corollary 1.** Under the condition (2.14), the estimator in (2.12) is expressed as

\[(2.15) \quad \tilde{\rho}_{n, k} := \tilde{\rho}_{n, k}(0) = \frac{2^{-\kappa} \Gamma(n/2)}{\Gamma(n/2 + 1 - \kappa)} \frac{(S - T) \cdot \kappa + 1}{(S + T)^{\kappa/2}} \cdot \text{$_2F_1$} \left( \frac{\kappa}{2}, \frac{\kappa + 1}{2}; \frac{n}{2}; \frac{4ST}{(S + T)^2} \right),\]

where $\text{$_2F_1$}$ is a hypergeometric function.

This is nothing but the estimator obtained in [11], although in that case the parameter $\kappa (= k$ according to the notation there) is allowed to assume also negative values. For the case $c > 0$ and $\kappa = n/2 + 1$, by using a modified formula of 6.633.2 in [3], we can obtain an explicit expression for $\tilde{\rho}_{n, k}(c)$:

**Corollary 2.**

\[(2.16) \quad \tilde{\rho}_{n, k/2 + 1}(c) = 2^{n/2 - 1} \frac{\Gamma(n/2)}{n} \frac{(1 + c)^{n/2 + 1}}{c^2} (S - T)(ST)^{(n - 3)/4} \times I_{n/2 - 1} \left( \sqrt{ST} \right) \exp \left( - \frac{S + T}{2c} \right),\]

where $I$ denotes a modified Bessel function of the first kind.

Also for the case $c > 0$ and $n = 1$, the integration in (2.12) can be carried out by means of the formulas 6.633.1 in [3] and (A.6) in Appendix;

**Corollary 3.**

\[(2.17) \quad \tilde{\rho}_{1, k}(c) = \frac{1}{2} \left( \frac{1 + c}{c} \right)^{\kappa} (S - T) \{ _1F_1 \left( \kappa; \frac{3}{2}; - \frac{(\sqrt{S} - \sqrt{T})^2}{2c} \right) + _1F_1 \left( \kappa; \frac{3}{2}; - \frac{(\sqrt{S} + \sqrt{T})^2}{2c} \right) \} = \frac{1}{2} \left( \frac{1 + c}{c} \right)^{\kappa} \sqrt{Y} \left\{ _1F_1 \left( \kappa; \frac{3}{2}; - \frac{X^2}{2c} \right) + _1F_1 \left( \kappa; \frac{3}{2}; - \frac{Y^2}{2c} \right) \right\}.\]

For the general case $n > 1$ and $c > 0$, by employing a similar method to that in Appendix, we have

**Corollary 4.**

\[(2.18) \quad \tilde{\rho}_{n, k}(c) = \frac{1}{n} \left( \frac{1 + c}{c} \right)^{\kappa} (S - T) \int_0^1 (v(1 - v))^{(n - 3)/2} \times _2F_1 \left( \kappa, n - 1; \frac{n}{2} + 1; - \frac{1}{2c} \left( (\sqrt{S} + \sqrt{T})^2 - 4\sqrt{ST}v \right) \right) dv, \quad n > 1.\]
This formula enables us to express $\hat{\rho}_{n,e}(c)$ for $n$ odd in terms of a finite series of hypergeometric functions; for example

$$
\hat{\rho}_{n,e}(c) = \frac{c}{4(\kappa-1)} \left( \frac{1+c}{c} \right)^{\frac{S-T}{ST}} \left\{ \text{Hermite } F_1 \left( \kappa-1, 1; \frac{3}{2}; \frac{1}{2c} \left( \sqrt{S} - \sqrt{T} \right)^2 \right) - F_1 \left( \kappa-1, 1; \frac{3}{2}; -\frac{1}{2c} \left( \sqrt{S} + \sqrt{T} \right)^2 \right) \right\}.
$$

The formula for $\hat{\rho}_{n,e}(c)$, (2.18), seems to be complicated compared with that for $\hat{\rho}_{n,e}$ in (2.15). In spite of this, the variance of $\hat{\rho}_{n,e}(c)$ can be expressed in a form as simple as that of $\hat{\rho}_{n,e}$. This is the theme to be investigated in the following section.

3. The variances of unbiased estimators

The second moment of the estimator $\hat{\rho}_n(q) := \hat{\rho}_n(q, r=0)$ in Theorem 1 is expressed as

$$
E[(\hat{\rho}_n(q))^2] = \int_0^{\infty} \int_0^{\infty} q(u) \cdot E[F_n(S, T; u)F_n(S, T; v)] \cdot q(v) \, du \, dv,
$$

provided, of course, that the $E$-operation and the integration with respect to $u$ and $v$ can be commuted. The expectation in the right hand side of (3.1) can be evaluated with the help of the formula 6.633.2 in [3]:

**Lemma 3.**

$$
E[(\hat{\rho}_n(q))^2] = 2^{n-3} \Gamma^2 \left( \frac{n}{2} \right) (ab)^{n-1}(uv)^{-(n-1)/2} \exp \left[ -\frac{1}{2} \left( \frac{1}{a} + \frac{1}{b} \right) (u+v) \right] \times \left\{ \frac{\sqrt{uv}}{a} \left( \frac{b}{a} + \frac{a}{b} - 1 \right) I_{n/2-1} \left( \frac{\sqrt{uv}}{a} \right) I_{n/2-1} \left( \frac{\sqrt{uv}}{b} \right) - \frac{\sqrt{uv}}{uv} I_{n/2} \left( \frac{\sqrt{uv}}{a} \right) I_{n/2} \left( \frac{\sqrt{uv}}{b} \right) \right\} \times \left\{ a \left( 1 - \frac{n}{2} \frac{u+v}{2b} \right) + \frac{u+v}{2} \right\} I_{n/2} \left( \frac{\sqrt{uv}}{a} \right) I_{n/2} \left( \frac{\sqrt{uv}}{b} \right) + \left( b \left( 1 - \frac{n}{2} \frac{u+v}{2a} \right) + \frac{u+v}{2} \right) I_{n/2} \left( \frac{\sqrt{uv}}{a} \right) I_{n/2} \left( \frac{\sqrt{uv}}{b} \right),
$$

where $a=1/(1+\rho)$ and $b=1/(1-\rho)$.

From the positivity of the second moment (3.1), it follows that the kernel function $K_n$ defined in Lemma 3 satisfies the (Schwarz) inequality

$$
|K_n(u, v; \rho)| \leq \sqrt{K_n(u, u; \rho)K_n(v, v; \rho)}.
$$

The asymptotic behaviours of $K_n(u, u; \rho)$ as $u \to 0$ and $\sim \infty$ are

$$
K_n(u, u; \rho) \sim \begin{cases} \rho^2 + \frac{1}{n} (1 + \rho^2), & u \sim 0, \\ \frac{2^{n-3}}{\pi} \Gamma^2 \left( \frac{n}{2} \right) (1-\rho^2)^{(1-n)/2} u^{-n}, & u \sim \infty \end{cases} \quad (|\rho| < 1)
$$

and

$$
K_n(u, u; \pm 1) \sim \begin{cases} 1 + \frac{2}{n}, & u \sim 0, \\ \frac{\Gamma(n/2)}{4\sqrt{n}} u^{-(n+1)/2}, & u \sim \infty \end{cases} \quad (|\rho| = 1).
$$
From these we can obtain a sufficient condition for the existence of the second moment (or the variance) of \( \tilde{\rho}_{n,s}(c) \). In this case, the function \( f \) in (3.1) is given by (2.11), so that the inequality (3.3) together with (3.4) or (3.5) implies that the second moment exists if the following conditions are satisfied:

\[
\kappa > 0 \quad \text{for } c > 0 \quad \text{or} \quad n > 2 \kappa > 0 \quad \text{for } c = 0, \quad (|\rho| < 1) , \]

\[
\kappa > 0 \quad \text{for } c > 0 \quad \text{or} \quad (n+1)/2 > 2 \kappa > 0 \quad \text{for } c = 0. \quad (|\rho| = 1)
\]

We shall see later that for the case \( c = 0 \) the variance exists in a range of parameter \( \kappa \) a little wider than that specified above.

After some calculations, a part of which will be reproduced in Appendix, we have

**Theorem 3.** The variance of \( \tilde{\rho}_{n,s}(c) \) is given by

\[
\text{Var} [\tilde{\rho}_{n,s}(c)] = \frac{I(n-1)}{I(n-1)} \int_0^1 (t(1-t))^{(n-3)/2} \left\{ \left( \frac{2c}{1+c} - \rho^2 + (1-\rho^2)t \right) \right. \\
\times \frac{1}{2} F_1 \left( \kappa, \kappa; \frac{n}{2}; \frac{1}{1+c} \right) - \left[ 1 - \frac{2c}{n} - \frac{2\rho^2}{1+c} \right] \left( \frac{2\kappa}{n} \right) \\
\left. - (1-\rho^2) \left( 1 - \frac{2}{n} \right) t \right\} 2 F_1 \left( \kappa, \kappa; \frac{n}{2} + 1; \frac{1}{1+c} \right) \right\} dt - \rho^2,
\]

for \( n > 1 \) and

\[
\begin{align*}
\text{Var} [\tilde{\rho}_{1,s}(c)] &= \left( \frac{1}{2} + \rho^2 \right) \frac{\kappa}{1+c} + \frac{\rho^2}{1+c} + \frac{1}{2} + \frac{1}{2} \left( 1 - \frac{2\kappa}{n} \right) \\
&= \left( \frac{1}{2} + \frac{1-2\kappa}{1+c} \rho^2 \right) 2 F_1 \left( \kappa, \kappa; \frac{3}{2}; \frac{1}{1+c} \right) - \rho^2.
\end{align*}
\]

**Proof.** See Appendix.

The variance (3.7) for the case \( n = 1 \), can be derived formally from (3.6) if we put

\[
\frac{I(n-1)}{I(n-1)} \left( (1-t)(1-t) \right)^{-(n-3)/2} = \frac{1}{2} (\delta(t) + \delta(1-t)) \quad \text{for } n = 1 ,
\]

where \( \delta(\cdot) \) is so called Dirac’s delta function. The integral in (3.6) has the same structure as that in (2.18), so that for \( n \) odd we can express \( \text{Var} [\tilde{\rho}_{n,s}(c)] \) through the integration by parts as a finite sum of the hypergeometric functions.

For the case

\[
n + \frac{1}{2} > 2 \kappa (\geq 0) \quad \text{for} \quad |\rho| < 1 \quad \text{or} \quad \frac{n}{2} + 1 > 2 \kappa (\geq 0) \quad \text{for} \quad |\rho| = 1 ,
\]

the limit of (3.6) as \( c \to 0 \) exists and the variance of \( \tilde{\rho}_{n,s} \) for \( n > 1 \) is given by the right hand side of (3.6) with \( c = 0 \). So we have

**Corollary 5.** Under the condition (3.9), the variance of \( \tilde{\rho}_{n,s} \) for \( n > 1 \) is expressed as

\[
\text{Var} [\tilde{\rho}_{n,s}] = \frac{I(n/2)}{\sqrt{n} I(n-1)} \int_0^1 (1-x^2)^{(n-3)/2} \left\{ \frac{1}{2} F_1 \left( \kappa-1, \kappa-1; \frac{n}{2}; \frac{1}{2}(1+\rho^2-(1-\rho^2)x) \right) \\
- \left( \frac{2\kappa}{n} \right) (1-\rho^2)(1-x) \cdot \frac{1}{2} F_1 \left( \kappa, \frac{n}{2} + 1; \frac{1}{2}(1+\rho^2-(1-\rho^2)x) \right) \right\} dx - \rho^2.
\]
The variance of $\hat{\rho}_{n_1}$ is equal to the right hand side of (3.7) with $c=0$ for $3/4 > \kappa (>0)$. This was already obtained by one of the authors [11]. Also for the case $n=3$, we can easily recognize that the expression for $Var[\hat{\rho}_{n_1}]$ given in [11] coincides with (3.11). From (3.6) the asymptotic behaviour of $Var[\hat{\rho}_{n_1}(c)]$ for large $n$ can be readily derived:

**Corollary 6.** The variance of $\hat{\rho}_{n_1}(c)$ for fixed $\kappa$ and $c$ and large $n$ behaves as

$$n \cdot Var[\hat{\rho}_{n_1}(c)] = \rho^2(1+\rho^2)\left(\frac{\kappa}{1+c}\right)^2 - 4\rho^2\frac{\kappa}{1+c} + \rho^2 + 1$$

$$+ \frac{1}{n}\left(\frac{\kappa}{1+c}\right)^2 \left[ \frac{1}{2} \rho^2(1+\rho^2)^2 \left(\frac{\kappa+1}{1+c}\right)^2 - 4\rho^2(1+\rho^2)\frac{\kappa+1}{1+c} \right] + O(n^{-1}).$$

As was mentioned in the final remark in Section 2, the expression for $Var[\hat{\rho}_{n_1}(c)]$, (3.6), is not so simplified even if we put $c=0$. The merit of introducing an extra parameter $c$ besides $\kappa$ will be clarified in the next section.

4. Some properties of $\hat{\rho}_{n_1}(c)$ for $n=1$

For the case that the sample size equals unity, Sibuya [14] considered the “simplified estimator of $\rho$”:

$$\hat{\rho}^* = \sqrt{\frac{2n}{4}} \left( [X \sgn(Y) + Y \sgn(X)] \right)$$

where $\sgn(x) = 1$, 0 or $-1$ according to $x>0$, $x=0$ or $x<0$. The variance of this estimator is

$$Var[\hat{\rho}^*] = \frac{n}{4} + \frac{\rho^2}{2} \text{Arcsin} \rho + \frac{\sqrt{1-\rho^2}}{2} - \rho^2.$$

If we restrict to $n=1$ in (2.12), we get

$$\hat{\rho}_{1_1}(c) = \frac{1}{2} \left( 1 + c \right)^{-1} \lambda Y \left\{ \exp \left[ - \frac{X^2}{2c} \right] F_1 \left( \frac{3}{2} - \kappa ; \frac{3}{2} ; \frac{Y^2}{2c} \right) + \exp \left[ - \frac{Y^2}{2c} \right] F_1 \left( \frac{3}{2} - \kappa ; \frac{3}{2} ; \frac{X^2}{2c} \right) \right\}.$$

From Table 13.6 in [1], it can be shown that $\hat{\rho}_{1_1}(c)$ is represented as

$$\hat{\rho}_{1_1}(c) = (1+c)^{-1} \sqrt{\frac{2\pi}{4}} \left\{ X \sgn(Y) \text{erf} \left( \frac{|Y|}{\sqrt{2c}} \right) + Y \sgn(X) \text{erf} \left( \frac{|X|}{\sqrt{2c}} \right) \right\},$$

where

$$\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \exp(-t^2) \, dt.$$

It is evident that $\hat{\rho}_{1_1}(c)$ approaches the conventional estimator $\hat{\rho}^*$ as $c \to 0$. In this sense we can identify $\hat{\rho}_{1_1}(0)$ with $\hat{\rho}^*$.

Now we try to find estimators which have uniformly smaller variances than those of $\hat{\rho}^*$. We shall determine the set of $(\kappa, c)$, if any, which satisfy the inequality $Var[\hat{\rho}_{1_1}(c)] \leq Var[\hat{\rho}^*]$ for any real number $-1 \leq \rho \leq 1$. For this purpose, let us first introduce the set of $(\kappa, c)$, denoted by $\Theta$, such that
The values in the right hand side of the above inequalities are the variances of \( \hat{\beta}^* \) at \(|\rho|=0 \) and 1 respectively. For the numerical determination of the set \( \Theta \), we use the following corollary which can be obtained as the special case of Theorem 3:

**Corollary 7.** For \( n \geq 1 \) it holds that

\[
(4.6) \quad \text{Var} \left[ \hat{\beta}_{n,c}(c) \right] \bigg|_{\rho=0} = \frac{1}{2} \cdot \frac{1}{c} F_1 \left( \frac{n-1}{2}, \kappa, \kappa; n, \frac{n}{2}, \frac{1}{(1+c)^2} \right) - \frac{1}{2} \left( 1 - \frac{2}{n} \right) \frac{1}{F_1 \left( \frac{n+1}{2}, \kappa, \kappa; n, \frac{n}{2} + 1, \frac{1}{(1+c)^2} \right)} ,
\]

where \( \kappa > 0 \) for \( c > 0 \) or \( n + \frac{1}{2} > 2\kappa > 0 \) for \( c = 0 \), and

\[
(4.7) \quad \text{Var} \left[ \hat{\beta}_{n,c}(c) \right] \bigg|_{\rho=1} = \frac{2c}{1+c} \cdot \frac{1}{c} F_1 \left( \kappa, \kappa; \frac{n}{2}, \frac{1}{(1+c)^2} \right) - \left[ 1 - \frac{2}{n} - \frac{2}{1+c} \left( 1 - \frac{2\kappa}{n} \right) \right] F_1 \left( \kappa, \kappa; \frac{n}{2} + 1, \frac{1}{(1+c)^2} \right) - 1 ,
\]

where \( \kappa > 0 \) for \( c > 0 \) or \( \frac{n}{2} + 1 > 2\kappa > 0 \) for \( c = 0 \).

The set \( \Theta \) can be found numerically from (4.6) and (4.7) for the case \( n = 1 \). The result is shown in Fig. 1, in which \( \Theta \) is the inside of the solid curves. If we take a point \((\kappa^*, c^*)\) belonging to \( \Theta \), then \( \text{Var} \left[ \hat{\beta}_{1,\rho}(c^*) \right] \) is smaller than \( \text{Var} \left[ \hat{\beta}^* \right] \) at least at the three points \( \rho = -1, 0 \) and 1. We have further to investigate \( \text{Var} \left[ \hat{\beta}_{1,\rho}(c^*) \right] \) for other values of \( \rho \). By using the formula (3.7) in Theorem 3, we perform the numerical computations of \( \text{Var} \left[ \hat{\beta}_{1,\rho}(c^*) \right] \), assuming

\[
\kappa^* = c^* + \frac{1}{2} , \quad c^* = \frac{i}{4} , \quad i = 0, 1, 2, 3, 4 \text{ and } 5.
\]
The results are listed in Table 1. We may conclude from this table that the variances of \( \tilde{\rho}_{n,c}(c^*) \) corresponding to the above cases are smaller than those of \( \tilde{\rho}^* \) for all values of \( \rho \).

### Table 1. Numerical values of \( \text{Var} [\tilde{\rho}_{n,c}(c)] \).

<table>
<thead>
<tr>
<th>( \kappa )</th>
<th>( c )</th>
<th>( \rho = 0 )</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/2</td>
<td>0</td>
<td>1.285</td>
<td>1.277</td>
<td>1.255</td>
<td>1.218</td>
<td>1.166</td>
<td>1.099</td>
<td>1.018</td>
<td>0.923</td>
<td>0.816</td>
<td>0.697</td>
<td>0.570</td>
</tr>
<tr>
<td>3/4</td>
<td>1/4</td>
<td>1.212</td>
<td>1.204</td>
<td>1.180</td>
<td>1.140</td>
<td>1.085</td>
<td>1.014</td>
<td>0.930</td>
<td>0.832</td>
<td>0.724</td>
<td>0.610</td>
<td>0.500</td>
</tr>
<tr>
<td>1</td>
<td>1/2</td>
<td>1.234</td>
<td>1.225</td>
<td>1.199</td>
<td>1.156</td>
<td>1.095</td>
<td>1.019</td>
<td>0.929</td>
<td>0.826</td>
<td>0.714</td>
<td>0.599</td>
<td>0.493</td>
</tr>
<tr>
<td>5/4</td>
<td>3/4</td>
<td>1.253</td>
<td>1.244</td>
<td>1.217</td>
<td>1.172</td>
<td>1.109</td>
<td>1.031</td>
<td>0.938</td>
<td>0.833</td>
<td>0.721</td>
<td>0.609</td>
<td>0.512</td>
</tr>
<tr>
<td>3/2</td>
<td>1</td>
<td>1.269</td>
<td>1.260</td>
<td>1.232</td>
<td>1.186</td>
<td>1.122</td>
<td>1.043</td>
<td>0.949</td>
<td>0.844</td>
<td>0.734</td>
<td>0.628</td>
<td>0.539</td>
</tr>
<tr>
<td>7/4</td>
<td>5/4</td>
<td>1.283</td>
<td>1.273</td>
<td>1.245</td>
<td>1.198</td>
<td>1.134</td>
<td>1.054</td>
<td>0.961</td>
<td>0.857</td>
<td>0.750</td>
<td>0.648</td>
<td>0.569</td>
</tr>
</tbody>
</table>

**lower bound** | 1.000 | 0.970 | 0.886 | 0.759 | 0.608 | 0.450 | 0.301 | 0.174 | 0.079 | 0.019 | 0.000 |

---

5. **Discussion**

We pointed out in this paper that, for \( N(0, 0; 1, 1; \rho) \), there exists infinitely many incomplete sufficient unbiased estimators of \( \rho \) having uniformly smaller variance than that of the simplified estimator \( \tilde{\rho}^* \) when, at least, the sample size is equal to unity. The estimator \( \tilde{\rho}_{n,c}(c) \) is expected to have this good property also when \( n \) is larger than 1. If \( n \) is sufficiently large, however, the formula of the asymptotic variances of \( \tilde{\rho}_{n,c}(c) \) depends only on \( \kappa/(1+c) \) (cf. (3.11)). This is essentially a one-parameter function, so that the merit of introducing an extra parameter disappears. If we could somehow manage \( c \) to be a random variable depending suitably upon \( S \) and \( T \), the variance of the resulting estimator
would attain the Cramer-Rao lower bound. This deserves further investigation.

The estimators $\hat{\rho}_{1,3/4}(c)$ and $\hat{\rho}_{1,3/2}(c)$ are expressed in terms of a modified Bessel and an exponential functions respectively. For practical applications, we concern with the case ($c^* = 3/2$, $c^* = 1$);

$$\hat{\rho}_{1,3/4}(1) = \sqrt{2} XY \left\{ \exp \left( -\frac{1}{2} X^2 \right) + \exp \left( -\frac{1}{2} Y^2 \right) \right\} .$$

This estimator is an improvement on the simplified estimator $\hat{\rho}^*$ with regard to variances. For this case, moreover, we have

$$\text{Var} [\hat{\rho}_{1,3/4}(1)] = \frac{4(4-2\rho^2 + \rho')}{(4-\rho^2)^{3/2}} + \frac{4}{3\sqrt{2}} - \rho^2$$

for $-1 \leq \rho \leq 1$.

The estimator $\hat{\rho}^*$ has been considered by Sibuya as the special case of the simplified estimator of correlogram $\hat{\rho}(h)$;

$$\hat{\rho}(h) = \frac{\sqrt{2\pi}}{4N} \sum_{i=1}^{N} [X_i \text{sgn}(X_{t+h}) + X_{t+h} \text{sgn}(X_i)]$$

for a stationary Gaussian process $X(t)$ with $E[X(t)] = 0$ and $\text{Cov}[X(t), X(t+h)] = \rho(h)$. If we take it the other way round, we may suggest a new estimator of correlogram $\rho(h)$, that is,

$$\hat{\rho}^*(h) = \frac{\sqrt{2}}{N} \sum_{i=1}^{N} X_i X_{t+h} \left\{ \exp \left( -\frac{1}{2} X_i^2 \right) + \exp \left( -\frac{1}{2} X_{t+h}^2 \right) \right\} .$$

The estimator $\hat{\rho}^*(h)$ is expected to have a better property than $\hat{\rho}(h)$ at least with regard to variances. The analytical and numerical calculations of the variances of $\rho^*(h)$ will be given in a supplementary article.

**Appendix**

We shall outline the derivation of (3.10), or its generalized form

$$\text{E}[\hat{\rho}_{n,1}(c)\hat{\rho}_{n,1}(d)] = \frac{I(n-1)}{I^2 \left( \frac{n-1}{2} \right)} \int_0^1 (1-t)^{(n-3)/2}
\times \left[ \left( \frac{c}{1+c} + \frac{d}{1+d} \right) \rho^2 + (1-\rho^2) t \right] _2 F_1 \left( \kappa, \lambda; \frac{n}{2}; \frac{1-(1-\rho^2) t}{(1+c)(1+d)} \right)
- \left[ \frac{2}{n} (1-(1-\rho^2) t) - \frac{1}{1+c} \left( 1-\frac{2\kappa}{n} \right) + \frac{1}{1+d} \left( 1-\frac{2\lambda}{n} \right) \right] \rho^2
\times _2 F_1 \left( \kappa, \lambda; \frac{n}{2} + 1; \frac{1-(1-\rho^2) t}{(1+c)(1+d)} \right) dt .$$

The integral exists under the condition (3.9) with $2\kappa$ replaced by $\kappa + \lambda$ for the case $c = d = 0$, otherwise it always exists since the argument in the hypergeometric functions is less than unity. In the above formula we understand that for the case $n = 1$ the replacement (3.8) should be made.

First we shall evaluate the integral

$$\int_0^\infty K_n(u, v; \rho) h(u; c, \kappa) dv .$$
The key formula is

\[(A.3) \int_0^\infty v^{\mu}e^{-(v+\rho)^2} I_v\left(\frac{\sqrt{uv}}{a}\right) I_v\left(\frac{\sqrt{uv}}{b}\right) dv = 2^{-\rho^2/2} \Gamma(v+1)(1+c)^{-\alpha-n+1/2} \cdot a^{-\rho^2} \cdot b^{n/2} \cdot u^{(n+1)/2} \]

\[\times \sum_{m=0}^\infty \frac{\Gamma(m+\mu+\lambda+2)}{m! \Gamma(m+\mu+1)} \left(\frac{c}{4a^2(1+c)}\right)^m \times \left(-m, -\mu - m; v+1; \left(\frac{a}{b}\right)^2\right), \quad (\mu + \nu + \lambda + 2 > 0)\]

which is obtained from 6.633.1 in [3] by setting \(x=\sqrt{v}, \beta=\sqrt{u}/a, \gamma=\sqrt{u}/b\) and \(\alpha=1+c\).

Putting (3.2) and (2.11) into (A.2) and applying the formula (A.3) to the resulting expression, we obtain several infinite series of type (A.3) with \(\mu, \nu\) and \(\lambda\) suitably chosen. We then transform the hypergeometric functions appearing in the infinite series by using the formula 9.134.3 in [3] and the recursion relations in the following way \((a=1/(1+\rho)\) and \(b=1/(1-\rho)\):

\[zF_1\left(-m, -\frac{n}{2}+1-m; \frac{n}{2}; \left(\frac{a}{b}\right)^2\right) = (4a^2)^m F_1\left(-m, \frac{n-1}{2}; n-1; 1-\rho^2\right),\]

\[zF_1\left(-m, -\frac{n}{2}+1-m; \frac{n}{2}+1; \left(\frac{a}{b}\right)^2\right) = \frac{\Gamma(n/2)}{n\Gamma(n/2+m)} (4a^2)^m F_1\left(-m, \frac{n-1}{2}; n-1; 1-\rho^2\right)
\]

\[+ \frac{m}{n} (4a^2)^{m-1} F_1\left(-m+1, \frac{n+1}{2}; n+1; 1-\rho^2\right)\]

and

\[zF_1\left(-m, -\frac{n}{2}-m; \frac{n}{2}; \left(\frac{a}{b}\right)^2\right) = (4a^2)^m F_1\left(-m, \frac{n-1}{2}; n-1; 1-\rho^2\right)
\]

\[+ 2m \left(\frac{a}{b}\right)^2 (4a^2)^{m-1} F_1\left(-m+1, \frac{n+1}{2}; n+1; 1-\rho^2\right) .\]

After these transformations, (A.2) turns out to be

\[\left(\frac{\Gamma(n/2)}{\Gamma(n)}\right) e^{-u} \left(\sum_{m=0}^\infty \frac{\Gamma(k+m)}{m! \Gamma(n/2+m)} \left[\frac{1}{2}(1+3\rho^2) + \left(\frac{1}{2}(1+\rho^2)\right) \left(1 - \frac{n}{2}\right) - \rho^2 u\right] \frac{1}{n/2+m} \right.
\]

\[- \rho^2 \cdot \frac{k+m}{n/2+m} \left[\frac{u}{1+c}\right]
\]

\[\times \left[\frac{1}{1+c}\left(\frac{1}{4n} (1-\rho^2)^3 \right) \frac{\Gamma(k+1+m)}{n/2+1+m} \left(\frac{1-n/2}{1+n/2+1+m}\right)\right.
\]

\[\times \left(\frac{u}{1+c}\right)^m F_1\left(-m, \frac{n+1}{2}; n+1; 1-\rho^2\right) .\]

By multiplying (A.4) by \(h(u; d, \lambda)\), integrating over \(u\) and making use of the integral representation

\[zF_1\left(-m, \frac{n-1}{2}; n-1; 1-\rho^2\right) = \frac{\Gamma(n-1)}{\Gamma\left(\frac{n-1}{2}\right)} \int_0^1 t^{(1-\rho)^2} (1-t)^{\alpha-1/2} \left(1 - (1-\rho^2)t\right)^m dt , \quad (n>1)\]

and further summing over \(m\), we find that the correlation function in the left hand side of (A.1) is reduced to
Integrating by parts in the second integral of the above expression and employing suitable recursion formulas for the hypergeometric functions, we finally arrive at (A.1). The excluded case, \( n=1 \), can be easily recovered if we use the relations.

\[
\begin{align*}
&\int_0^1 \left( (1-t)^{(n-1)/2} \left[ \frac{1}{2} \left(1+3\rho^2\right)_{\nu} F_1 \left( \kappa, \lambda; \frac{n}{2}+1; \frac{1-(1-\rho^2)t}{(1+c)(1+d)} \right) \\
&+ \frac{1}{2} \left(1+\rho^2\right) \left( \frac{2}{n} - 1 \right)_{\nu} F_1 \left( \kappa+1, \lambda; \frac{n}{2}+1; \frac{1-(1-\rho^2)t}{(1+c)(1+d)} \right) \right] dt + \frac{\Gamma(n+1)}{\Gamma(\frac{n+1}{2})} \frac{1}{4n} \frac{(1-\rho^2)^2}{(1+c)(1+d)} \\
&\times \left[ \left( \frac{2}{n} - 1 \right) \left( \frac{2\lambda}{n+2} \right)_{\nu} F_1 \left( \kappa+1, \lambda+1; \frac{n}{2}+2; \frac{1-(1-\rho^2)t}{(1+c)(1+d)} \right) \right] \\
&- \frac{2\lambda}{n} \left( \frac{2\lambda}{n+2} \right)_{\nu} F_1 \left( \kappa+1, \lambda+1; \frac{n}{2}+1; \frac{1-(1-\rho^2)t}{(1+c)(1+d)} \right) \right) dt.
\end{align*}
\]

Integrating by parts in the second integral of the above expression and employing suitable recursion formulas for the hypergeometric functions, we finally arrive at (A.1). The excluded case, \( n=1 \), can be easily recovered if we use the relations.

\[
\lim_{n \to 1} F_1 \left( -m, \frac{n-1}{2}; n-1; z \right) = \frac{1}{2} (1+(1-z)^m),
\]

and perform the summation over \( m \) in (A.4). By the way the upper formula in (A.6) is consistent with (A.5) under the replacement \((3.8)\).


Note added in proof: Professor A. R. Sampson kindly informed the authors of his paper “Simple BAN estimators of correlations for certain multivariate normal models with known variances” (*Journal of the American Statistical Association* **70** 859–862 (1978)), in which he obtained the maximum likelihood estimator in the situation more general than that in the present paper.