The Annals of Statistics
1999, Vol. 27, No. 2, 610–626

BROWN'S PARADOX IN THE ESTIMATED CONFIDENCE APPROACH

BY HSIUYING WANG

Academia Sinica

A widely held notion of classical conditional theory is that statistical inference in the presence of ancillary statistics should be independent of the distribution of those ancillary statistcs. In this paper, ancillary paradoxes which contradict this notion are presented for two scenarios involving confidence estimation. These results are related to Brown's ancillary paradox in point estimation. Moreover, the confidence coefficient, the usual constant coverage probability estimator, is shown to be inadmissible for confidence estimation in the multiple regression model with random predictor variables if the dimension of the slope parameters is greater than five. Some estimators better than the confidence coefficient are provided in this paper. These new estimators are constructed based on empirical Bayes estimators.

1. Introduction. Consider the canonical problem in which X is a p-dimensional multivariate normal random variable with mean θ and covariance matrix $I_{p\times p}$, the $p \times p$ identity matrix, and $W = (W_1, \dots, W_p)'$ is an independent observation. Define

$$\eta = \sum_{i=1}^{p} W_i \theta_i = W'\theta. \tag{1}$$

A customary estimator of η is $W'X$. For a confidence interval

$$C_{X,W} = \left\{\eta\colon \frac{|W'X - \eta|}{|W|} \le c_\gamma\right\}$$

of η, the reported confidence statement is usually $1-\gamma$, where c_γ satisfies $P(\eta \in C_{X,W}) = 1-\gamma$. Kiefer (1977) pointed out that, in place of the constant value $1-\gamma$, a better approach is to provide a data dependent estimate $r(X, W)$ of the value of the coverage function $I(\eta \in C_{X,W})$, where

$$I(\eta \in C_{X,W}) = \begin{cases} 1, & \text{if } \eta \in C_{X,W}, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Therefore, we consider a confidence procedure $\langle I(\eta \in C_{X,W}), r(X, W)\rangle$. The interpretation of $\langle I(\eta \in C_{X,W}), r(X, W)\rangle$ is that $r(X, W)$ states a degree of belief or level of confidence in the proposition that $\eta \in C_{X,W}$ after $(X, W) = (x, w)$ has been observed. Related literature is Robinson (1979a, b) and Lu

Received September 1996; revised September 1998.
AMS 1991 *subject classifications*. Primary 62C15; secondary 62C10.
Key words and phrases. Confidence interval, admissibility, coverage function, the usual constant coverage probability estimator, ancillary statistic.

and Berger (1989). For a conditional report $r(X, W)$, we consider squared error loss

$$L[\eta, r(X, W)] = \left[r(X, W) - I\left(\frac{|W'X - \eta|}{|W|} \le c_\gamma\right)\right]^2. \tag{3}$$

We shall show that the constant report $1 - \gamma$ is an inadmissible estimator of $I_{C_{X,W}}(\eta)$ under squared error loss (3) if $p \ge 5$ when W is random, whereas it is admissible when W is fixed. Note that the classical approach suggests that one should condition on W when making inference, since W here is an ancillary statistic. However, this leads to a one-dimensional problem and therefore the natural estimator for the confidence is $1 - \gamma$, which is admissible in that one-dimensional problem. What seems surprising is that $1 - \gamma$ is actually inadmissible in the unconditional framework. A similar phenomenon was demonstrated in Brown (1990) regarding the point estimation problem for η.

Sandved (1968) has some results in conditional inference. He shows that confidence statements, conditional on an ancillary statistic, dominate unconditional confidence statements under squared error loss. Indeed, since

$$E[1 - \gamma - I(\xi \in C_X)]^2 = E[1 - \gamma - r(A)]^2 + E[r(A) - I(\xi \in C_X)]^2,$$

where C_X is a confidence interval for θ, A is an ancillary statistic, $r(A) = P(\theta \in C_X | A)$ and $1 - \gamma = P(\theta \in C_X)$, it is clear that

$$E[r(A) - I(\xi \in C_X)]^2 \le E[1 - \gamma - I(\xi \in C_X)]^2.$$

In this paper, since a confidence statement conditional on ancillary statistic W is the same as an unconditional confidence statement, Sandved's result does not provide any information regarding the paradoxical phenomenon.

In the proof of inadmissibility, we will construct an estimator which is better than $1 - \gamma$. Robinson (1979b) and Robert and Casella (1994) have results related to this issue. Let $X \sim N(\theta, I_{p\times p})$ and $C_X = \{\theta\colon |X - \theta| \le c_1\}$, where c_1 satisfies $P(\theta \in C_X) = 1 - \gamma$. Then Robinson (1979b) and Robert and Casella (1994) show that $\delta(X) = 1 - \gamma + \varepsilon/(1 + |X|^2)$ dominates $1 - \gamma$ for estimating $I(\theta \in C_X)$ under squared error loss if $p = 5$ and $p \ge 5$, respectively, where ε is a sufficiently small positive number. Here $\delta(X)$ is closely related to the James–Stein estimator in point estimation since they are both derived from the empirical Bayes approach. In Theorem 2.2.6, we will show that an estimator, which is also constructed based on the empirical Bayes estimator, dominates $1 - \gamma$ if $p \ge 6$ when W is random. The better estimator is

$$1 - \gamma + \frac{a(\mathbf{1}'\Omega^{-1/2}X)^2}{2p(b + X'\Omega^{-1}X)^2} + \frac{a}{(b + X'\Omega^{-1}X)}, \tag{4}$$

where a and b are certain positive constants, $\mathbf{1}' = (1, \ldots, 1)_{1\times p}$ and Ω is a $p \times p$ matrix defined in Lemma 2.2.4 below.

It is worth noting that the above paradoxical phenomenon in confidence procedures also happens in the multiple regression model. Consider a multiple linear regression with random predictor variables and assume that the dimension of the slope parameters is greater than 5. In Section 3, we shall prove that the usual constant coverage probability estimator of the coverage function of the confidence interval for the unknown coefficient parameter is not admissible. Note that, if the predictor variables are fixed, the same estimator is admissible. Thus, this is another important example of the phenomenon. A better estimator, similar to (4), is also constructed in this model.

2. The paradox for the confidence procedure.

2.1. *Admissibility when W is fixed.* In this section, we will prove that the constant coverage probability estimator is admissible for (2) when the ancillary statistic has a degenerate distribution. This is a natural result since the problem then reduces to a one-dimensional problem in which the ancillary statistic can be taken to be any prespecified constant value. For convience, $I(|\cdot| < c)$ is denoted by $I_c(|\cdot|)$ throughout the paper.

THEOREM 2.1.1. *Let $X \sim N(\theta, I_{p\times p})$ and $W = (W_1, \dots, W_p)$ be fixed at $w = (w_1, \dots, w_p)$. The estimator $1 - \gamma$ of* (2) *is admissible under the usual squared error loss* (3).

2.2. *Inadmissibility when W is random.* Now assume that the values of $(w_1, \dots, w_p)$ in the previous section are coordinates of a random variable $W \in \Re^p$. We also assume that the distribution of W is known and W is independent of X. Since the distribution of W is independent of θ, W is an ancillary statistic.

From now on, we assume that $X \sim N(\theta, I_{p\times p})$ and $W = (W_1, \dots, W_p)$, where W_i are i.i.d. with $EW_i = \mu$, and $\operatorname{Var}(W_i) = \sigma^2 > 0$, the distribution being unimodal and symmetric about μ. Note that μ and σ^2 are assumed to be known.

In proving the main results of this section, we need the following lemmas. Here E_0 denotes expectation with respect to X and W when $\theta = 0$.

LEMMA 2.2.1.

$$E_0\left\{X_1^2\left[1 - \gamma - I_c\left(\frac{|W'X|}{|W|}\right)\right]\right\} > 0.$$

LEMMA 2.2.2. *If $\mu \neq 0$, then*

$$E_0\left\{X_1X_2\left[1 - \gamma - I_c\left(\frac{|W'X|}{|W|}\right)\right]\right\} > 0.$$

LEMMA 2.2.3.

$$E_0\left\{X_1X_2\left(1 - \gamma - I_c\left(\frac{|W'X|}{|W|}\right)\right]\right\} < E_0\left\{X_1^2\left[1 - \gamma - I_c\left(\frac{|W'X|}{|W|}\right)\right]\right\}. \tag{5}$$

LEMMA 2.2.4. *Let $U' = (\mu, \dots, \mu)_{1\times p}$ and $\Omega = I_{p\times p} + kUU'$, where k is a positive constant. Then there exists a constant $h_k = \{-1 \pm [1 - kp\mu^2/(1 + kp\mu^2)]^{1/2}\}/(p\mu^2)$ such that $\Omega^{-1/2} = I_{p\times p} + h_kUU'$.*

LEMMA 2.2.5. *If $\mu \neq 0$, let*

$$g = \frac{E_0\{X_1 X_2[1-\gamma - I_c(|W'X|/|W|)]\}}{E_0\{X_1^2[1-\gamma - I_c(|W'X|/|W|)]\}} \tag{6}$$

and

$$k = \left[4 - \left(\frac{2\varepsilon}{p+\varepsilon}\right)^2\right] \bigg/ \left[p\mu^2\left(\frac{2\varepsilon}{p+\varepsilon}\right)^2\right], \tag{7}$$

where ε is a positive constant such that $\varepsilon < 2/g - 2$. Then

$$E_0\left\{Y_1 Y_2\left[1-\gamma - I_c\left(\frac{|W'X|}{|W|}\right)\right]\right\} < 0,$$

where $Y = \Omega^{-1/2}X$ and $\Omega^{-1/2}$ is given in Lemma 2.2.4.

The proofs of Lemmas 2.2.1–2.2.5 are given in the Appendix. We now state and prove the main theorem in this section.

THEOREM 2.2.6. *Assume that $X \sim N(\theta, I_{p\times p})$ and $W = (W_1, \dots, W_p)$ are independent and that the W_i are i.i.d. with $E(W_i) = \mu$, $\mathrm{Var}(W_i) = \sigma^2 > 0$, the distributions being unimodal and symmetric about μ. Then, for $p \geq 6$, the estimator $r_0 = 1 - \gamma$ of (2) is inadmissible for loss (3). A better estimator is given by*

$$r_1(X) = 1 - \gamma + \frac{a(\mathbf{1}'\Omega^{-1/2}X)^2}{2p(b + X'\Omega^{-1}X)^2} + \frac{a}{(b + X'\Omega^{-1}X)}, \tag{8}$$

where $a > 0$ and $b > 0$ are sufficiently small and sufficiently large constants, respectively.

PROOF. First, we consider the case $\mu \neq 0$. Note that

$$\begin{aligned} R(r_0, \theta) - R(r_1, \theta) = {}& -2aE_0\left\{\left(\frac{A}{2p} + B\right)\left[1-\gamma - I_c\left(\frac{|W'X|}{|W|}\right)\right]\right\} \\ & - a^2 E_0\left[\left(\frac{A}{2p} + B\right)^2\right], \end{aligned} \tag{9}$$

where

$$A = [\mathbf{1}'\Omega^{-1/2}(X+\theta)]^2[b + (X+\theta)'\Omega^{-1}(X+\theta)]^{-2},$$
$$B = [b + (X+\theta)'\Omega^{-1}(X+\theta)]^{-1} \quad \text{and} \quad \mathbf{1}' = (1, \dots, 1)_{1\times p}.$$

Define

$$Y = \Omega^{-1/2}X \quad \text{and} \quad \phi = \Omega^{-1/2}\theta. \tag{10}$$

Below we use k_1, k_2 and k_3 to denote $\mathbf{1}'\phi$, $b+\phi'\phi$ and $\phi'\phi$, respectively. By substituting (10) into A and then using a Taylor expansion, we have

$$
\begin{aligned}
A = {} & k_1^2 k_2^{-2} + \sum_{i=1}^{p} Y_i A_{1,i} + \tfrac{1}{2}\sum_{i=1}^{p} Y_i^2 A_{2,i} \\
& + \tfrac{1}{2}\sum_{i,j=1}^{p} Y_i Y_j A_{3,i,j} + e_1(b,\phi,Y),
\end{aligned}
\tag{11}
$$

where

$$
\begin{aligned}
A_{1,i} &= 2k_1k_2^{-2} - 4k_1^2k_2^{-3}\phi_i,\\
A_{2,i} &= 2k_2^{-2} - 8k_1k_2^{-3}\phi_i - 8k_1k_2^{-3}\phi_i + 24k_1^2k_2^{-4}\phi_i^2 - 4k_1^2k_2^{-3},\\
A_{3,i,j} &= 2k_2^{-2} - 8k_1k_2^{-3}\phi_j - 8k_1k_2^{-3}\phi_i + 24k_1^2k_2^{-4}\phi_i\phi_j
\end{aligned}
$$

and

$$
\begin{aligned}
B = {} & k_2^{-1} + \sum Y_i[-k_2^{-2}2\phi_i] + \tfrac{1}{2}\sum Y_i^2[2k_2^{-3}(2\phi_i)^2 - 2k_2^{-2}] \\
& + \tfrac{1}{2}\sum Y_iY_j[2k_2^{-3}(2\phi_i)(2\phi_j)] + e_2(b,\phi,Y).
\end{aligned}
\tag{12}
$$

Now we are ready to take the expectation. Since $(X_1,\dots,X_p)$ have the same distribution as $(-X_1,-X_2,\dots,-X_p)$, it follows that the expectation of $X_1[1-\gamma-I_c(|W'X|/|W|)]$ is zero. This implies that

$$
E_0\left\{Y_1\left[(1-\gamma) - I_c\left(\frac{|W'X|}{|W|}\right)\right]\right\} = 0,
$$

since Y_1 is a linear combination of $X_1,\dots,X_p$. Also by (10) and the expression of $\Omega^{-1/2}$ in Lemma 2.2.4, it can be shown that $(Y_1,\dots,Y_p)$ have a permutation invariant distribution. Using this, (10), (11) and (12), $-E_0\{(A/2p+B)[1-\gamma-I_c(|W'X|/|W|)]\}$ in (9) equals

$$
\begin{aligned}
\frac{1}{2k_2^4} m\Biggl\{ & b^2\left[\frac{-2p+2p(p-1)h}{2p} + 2p\right] + k_3^2\left[\frac{-2p+2p(p-1)h}{2p} - 8 + 2p - 8h\right] \\
& + k_3k_1^2\left[\frac{-8+4p-8h(p-1)-8h(p-1)-24h}{2p} + 8h\right] \\
& + k_1^4\frac{24h}{2p} + bk_3\left[\frac{-4p+4p(p-1)h}{2p} - 8 + 4p - 8h\right] \\
& + bk_1^2\left[\frac{16+4p-8(p-1)h-8(p-1)h}{2p} + 8h\right]\Biggr\} + e(b,\phi),
\end{aligned}
$$

where the error term is shown in Wang (1998b) to be

$$
\begin{aligned}
e(b,\phi) &= o\left[\frac{1}{k_2^2}\right],\\
m &= E_0\left\{Y_1^2\left[1-\gamma-I_c\left(\frac{|W'X|}{|W|}\right)\right]\right\}
\end{aligned}
\tag{13}
$$

and

$$h = -\frac{E_0\{Y_1 Y_2[1-\gamma - I_c(|W'X|/|W|)]\}}{E_0\{Y_1^2[1-\gamma - I_c(|W'X|/|W|)]\}}.$$

Using this and a direct computation of $E_0[(A/2p+B)^2]$, we have

$$R(r_0,\theta) - R(r_1,\theta) = a\frac{1}{(b+\phi'\phi)^4 p} m\rho - a^2\tau + o\left[\frac{1}{k_2^2}\right], \tag{14}$$

where

$$\begin{aligned}\rho = {} & b^2[p^2(4+2h) + p(-2-2h)] + k_3^2[p^2(4+2h) + p(-18-18h)] \\ & + k_3 k_1^2(4p - 8 - 8h) + k_1^4 24h \\ & + bk_3[p^2(4h+8) + p(-20-20h)] + bk_1^2(4p+16+16h)\end{aligned}$$

and

$$\tau = \left[\frac{k_1^2 k_2^{-2}}{2\rho} + k_2^{-1}\right]^2.$$

We shall show that m and ρ are greater than zero when $p \geq 6$. Then, by the relation that $(\mathbf{1}\phi)^2 < p(\phi'\phi)$, we conclude that (14) is greater than zero for all θ when $p \geq 6$ if b goes to infinity and a goes to zero. Thus, the proof will be completed.

Now consider m of (14) which equals $E_0\{(eX_1 + dX_2 + \cdots + dX_p)^2[1-\gamma - I_c(|W'X|/|W|)]\}$ where $e = 1 + h_k\mu^2$, $d = h_k\mu^2$, k is given in (7) and h_k is defined in Lemma 2.2.4. Expanding the term $(eX_1 + \cdots + dX_p)^2$, then replacing X_i^2, $i = 1,\dots,p$, by X_1X_2 in the above expansion and using Lemma 2.2.3, gives the lower bound of m,

$$E_0\left\{[e+(p-1)d]^2 X_1 X_2\left[1-\gamma - I_c\left(\frac{|W'X|}{W}\right)\right]\right\},$$

which is positive by Lemma 2.2.2 and the fact that $e+(p-1)d \neq 0$. This shows that m is positive. To show $\rho > 0$, it suffices to show that the coefficients of b^2, $(\phi'\phi)^2$, $(\phi'\phi)(\mathbf{1}'\phi)^2$, $(\mathbf{1}'\phi)^4$, $b\phi'\phi$ and $b(\mathbf{1}'\phi)^2$ of ρ are all greater than 0 when $p \geq 6$. We only give the details for the more involved coefficient of $(\phi'\phi)^2$. The others can be dealt with similarly. First note that

$$0 < h < 1, \tag{15}$$

which is established in Wang (1998b). By a straightforward caculation, the coefficient of $(\phi'\phi)^2$ (i.e., k_3^2) in ρ can be written as

$$\left[(4+2h)\left(p + \frac{-9-9h}{4+2h}\right)^2 - \frac{(-9-9h)^2}{4+2h}\right], \tag{16}$$

which is greater than 0 if and only if $p > 9 - 9/(h+2)$. Note that $3 < 9/(h+2) < 5$, because $0 < h < 1$. Hence $p \geq 6$ obviously implies that (16) > 0.

Since the coefficient of a in (14) is greater than 0, we can choose $a > 0$ small enough and $b > 0$ large enough such that $R(r_0, \theta) - R(r_1, \theta) > 0\ \forall\ \theta$. Therefore, the case that $\mu \neq 0$ is proved.

When $\mu = 0$, Ω equals $I_{p\times p}$. Note that the inequality in Lemma 2.2.2 then becomes an equality. Thus, by an argument similar to that in the first case, $R(r_0, \theta) - R(r_1, \theta)$ can be shown to be positive for all θ if a is small enough and b is large enough. □

In Theorem 2.2.7, if $\mu = 0$, then there exists another better estimator when $p \geq 5$.

THEOREM 2.2.7. *Under the assumptions of Theorem* 2.2.6 *and, in addition, assuming* $\mu = 0$ *and* $p \geq 5$, *a better estimator than* $1 - \gamma$ *for estimating* (2) *is*

$$1 - \gamma + \frac{a}{b + |X|^2}, \tag{17}$$

where a and b are some positive constants.

The proof of Theorem 2.2.7 can be established by arguments similar to Theorem 2.2.6. This is because the mean of W is 0 and the order of the error term is the same as in the situation without an ancillary statistic W. Note that (17) is independent of W and (8) depends on W only through the mean of W.

REMARK 1. If the distribution of W is not degenerate, $r_1(x)$ is shown to be better than $1 - \gamma$ for estimating (2) if $p \geq 6$. If the mean of W is zero, (17) is also shown to be better than $1 - \gamma$ if $p \geq 5$. According to the latter result, we conjecture that when $p = 5$, $1 - \gamma$ is inadmissible if the mean of W is not equal to zero.

In addition to Theorem 2.2.6, which shows that $1 - \gamma$ is inadmissible if $p \geq 6$, Wang (1998a) has proved that $1 - \gamma$ is admissible if $p \leq 4$.

Thus far we have only considered the confidence interval $C_{X,W}$. Now we turn to consider the confidence interval $C^*_{X,W} = \{\eta\colon |W'X - \eta| \leq c_2\}$, where c_2 is any positive constant. Let $\lambda(W) = P(|W'X - \eta| \leq c_2 | W)$.

THEOREM 2.2.8. *Under the assumptions of Theorem* 2.2.6, *the estimator* $\lambda(W)$ *is not admissible for estimating* $I(\eta \in C^*_{X,W})$ *if* $p \geq 6$ *under squared error loss. A better estimator is given by*

$$\lambda(X, W) = \lambda(W) + \frac{a^*(\mathbf{1}'\Omega^{-1/2}X)^2}{2p(b^* + X'\Omega^{-1}X)^2} + \frac{a^*}{(b^* + X'\Omega^{-1}X)},$$

where a^* and b^* are some positive constants.

The proof is similar to that of Theorem 2.2.6 and is omitted.

2.3. *Empirical Bayes considerations.* In this section, an empirical Bayes argument leading to (8) is presented. The empirical Bayes estimator of (2) with respect to the prior $\theta|\tau \sim N(0, \tau^2 I_p)$ is

$$F_\lambda\left(c\left(1+\frac{1}{\tau^2}\right)\right) - F_\lambda\left(-c\left(1+\frac{1}{\tau^2}\right)\right),$$

where $\lambda = w'x/(\tau^2|w|)$ and $F_\lambda(\cdot)$ is the cdf of the normal distribution with mean λ and unit variance. Using $|x|^2$ to estimate $(1+\tau^2)$ and applying a Taylor expansion yields

$$\begin{aligned}
&P\left(\frac{|w'\theta - w'x|}{|w|} \le c|x\right)\\
&\quad\simeq F_\lambda(c) + \frac{c}{\tau^2} f_\lambda(c) - F_\lambda(-c) + \frac{c}{\tau^2} f_\lambda(-c)\\
&\quad\simeq [F_0(c) - F_0(-c)] + \lambda[F_0'(c) - F_0'(-c)]\\
&\qquad + \frac{c}{\tau^2}[f_0(c) + f_0(-c)] + \frac{c}{\tau^2}\lambda\frac{\partial}{\partial\lambda}(f_\lambda(c) + f_\lambda(-c))|_{\lambda=0}\\
&\quad\simeq 1 - \gamma + \frac{w'x}{(|x|^2-1)|w|}(F_0'(c) - F_0'(-c)) + \frac{c}{|x|^2-1}(f_0(c) + f_0(-c))\\
&\qquad + \frac{w'x}{(|x|^2-1)^2|w|}\frac{\partial}{\partial\lambda}(f_\lambda(c) + f_\lambda(-c))|_{\lambda=0},
\end{aligned}$$

where $f_\lambda(c) = (\partial/\partial x)F_\lambda(x)|_{x=c}$. The last expression and the form of the estimator (2.1.4) in Brown (1990) yield (8).

3. The multiple regression problem.

3.1. *Admissibility in a regression model.* In Section 2, we considered the case where the ancillary statistic W is independent of the major random variable X. In this section, consider the usual normal multiple linear regression. Denote the $(p+1)$ unknown parameters by $\alpha \in R$, $\beta = (\beta_1, \dots, \beta_p)' \in R^p$. Let $V_i = (V_{i1}, \dots, V_{ip})'$, $i = 1, \dots, n$, denote the observed (i.e., known) regression constants. Let $Y_1, \dots, Y_n$ be independent normal random variables with

$$E(Y_i) = \alpha + V_i'\beta, \qquad \mathrm{Var}(Y_i) = \sigma^2,\ i = 1, \dots, n,$$

where σ^2 is known. The usual estimators of α and β are

$$\hat\alpha = \bar Y - \bar V\hat\beta \quad \text{and} \quad \hat\beta = S^{-1}V'(Y - \bar Y\mathbf{1}),$$

where $V' = (V_1, \dots, V_n)$, $\bar Y = n^{-1}\mathbf{1}'Y$, $\bar V = n^{-1}\mathbf{1}'V$, $S = (V - \mathbf{1}\bar V)'(V - \mathbf{1}\bar V)$ and $\mathbf{1}' = (1, \dots, 1)_{1\times n}$. Note that

$$\begin{pmatrix}\hat\alpha\\ \hat\beta\end{pmatrix} \sim N\left(\begin{pmatrix}\alpha\\ \beta\end{pmatrix}, \sum(V)\right),$$

$$\sum(V) = \begin{pmatrix} n^{-1} + \bar V S^{-1}\bar V' & -\bar V S^{-1}\\ -S^{-1}\bar V' & S^{-1}\end{pmatrix}.$$

Consequently, a $1-\gamma$ confidence interval for α is

$$C_{\hat{\alpha}, V} = \left\{\alpha \colon \frac{|\hat{\alpha}-\alpha|}{\sqrt{n^{-1}+\bar{V}S^{-1}\bar{V}'}} \le c\right\},$$

where c is the upper $1-\gamma/2$ cutoff point of $N(0,1)$. Here we are interested in estimating $I(\alpha \in C_{\hat{\alpha}, V})$. For an estimator $r(\hat{\alpha}, \hat{\beta})$, consider the squared error loss

$$[r(\hat{\alpha}, \hat{\beta}) - I(\alpha \in C_{\hat{\alpha}, V})]^2. \tag{18}$$

THEOREM 3.1. *In the preceding problem*, $1-\gamma$ *is an admissible estimator of* $I(\alpha \in C_{\hat{\alpha},V})$.

PROOF. If V is fixed, then $\hat{\alpha} \sim N(\alpha, n^{-1}+\bar{V}S^{-1}\bar{V}')$. It reduces to the one-dimensional normal problem. Thus, $1-\gamma$ is an admissible estimator of $I(\alpha \in C_{\hat{\alpha}, V})$. □

3.2. *Inadmissibility result.* In Section 3.1, the design variables $V_{i,j}$ were assumed to be fixed, known constants. This assumption is realistic in many applications where the $V_{i,j}$ are preassigned by the experimenters. However, in many other situations, the V_i are independent vector random variables. In this section, we assume that the $V_{i,j}$ are i.i.d $N(\mu, 1)$ with μ known. In the following, we shall first prove an inadmissibility result for the usual estimator, $1-\gamma$, of $I(\alpha \in C_{\hat{\alpha}, V})$.

Throughout this section, the notation E_0 is used to denote the expectation with respect to $\hat{\alpha}$ and $\hat{\beta}$ when $\alpha = 0$ and $\beta = 0_{p\times 1}$ and with respect to V as well.

THEOREM 3.2.1. *Let* Y *be defined as in Section* 3.1 *and let* $V_{i,j}$, $i = 1, \dots, n$, $j = 1, \dots, p$, *be i.i.d. with* $E(V_{i,j}) = \mu$ *and* $\mathrm{Var}(V_{i,j}) = 1$. *Define* $\Omega = I_{p\times p} + kUU'$, *where* $U' = (\mu, \dots, \mu)$ *and* k *will be specified in Lemma* 3.2.4 *below. Then, for* $p \ge 6$,

$$r_0 = 1-\gamma \tag{19}$$

is not admissible under loss (18). *A better estimator is given by*

$$r(\hat{\beta}) = \begin{cases} 1-\gamma + \dfrac{a_1(\mathbf{1}'\Omega^{-1/2}\hat{\beta})^2}{2p(b_1+\hat{\beta}'\Omega^{-1}\hat{\beta})^2} + \dfrac{a_1}{(b_1+\hat{\beta}'\Omega^{-1}\hat{\beta})}, \\ \qquad \text{if } E_0\left\{\hat{\beta}_1\hat{\beta}_2\left[1-\gamma - I_c\left(\dfrac{|\hat{\alpha}-\alpha|}{\sqrt{n^{-1}+\bar{V}S^{-1}\bar{V}'}}\right)\right]\right\} > 0, \\ 1-\gamma + \dfrac{a_2}{b_2+|\hat{\beta}|^2}, \\ \qquad \text{if } E_0\left\{\hat{\beta}_1\hat{\beta}_2\left[1-\gamma - I_c\left(\dfrac{|\hat{\alpha}-\alpha|}{\sqrt{n^{-1}+\bar{V}S^{-1}\bar{V}'}}\right)\right]\right\} \le 0, \end{cases} \tag{20}$$

where a_1, b_1, a_2 *and* b_2 *are some positive constants.*

Although it is formidable to show that

$$E_0\left\{\hat{\beta}_1\hat{\beta}_2\left[1-\gamma-I_c\left(\frac{|\hat{\alpha}-\alpha|}{\sqrt{n^{-1}+\bar{V}S^{-1}\bar{V}'}}\right)\right]\right\}\geq 0,$$

statistical simulation indicates that it may hold for all μ. Based on this and the similarity of the first part of $r(\hat{\beta})$ and $r_1(X)$ of (8) in Section 2.2, we conjecture that the first part of $r(\hat{\beta})$ in (20) dominates $1-\gamma$ for all α and β.

In Lemmas 3.2.2–3.2.4, all results are established under the same assumptions as in Theorem 3.2.1.

LEMMA 3.2.2.

$$E_0\left\{\hat{\beta}_1^2\left[1-\gamma-I_c\left(\frac{|\hat{\alpha}-\alpha|}{\sqrt{n^{-1}+\bar{V}S^{-1}\bar{V}'}}\right)\right]\right\}>0. \tag{21}$$

LEMMA 3.2.3.

$$\left|E_0\left\{\hat{\beta}_1\hat{\beta}_2\left[1-\gamma-I_c\left(\frac{|\hat{\alpha}-\alpha|}{\sqrt{n^{-1}+\bar{V}S^{-1}\bar{V}'}}\right)\right]\right\}\right|$$
$$< E_0\left\{\hat{\beta}_1^2\left[1-\gamma-I_c\left(\frac{|\hat{\alpha}-\alpha|}{\sqrt{n^{-1}+\bar{V}S^{-1}\bar{V}'}}\right)\right]\right\}.$$

LEMMA 3.2.4. *Assume that* $E_0\{\hat{\beta}_1\hat{\beta}_2[1-\gamma-I_c(|\hat{\alpha}-\alpha|/\sqrt{n^{-1}+\bar{V}S^{-1}\bar{V}'})]\}>0$. *Let*

$$g=\frac{E_0\{\hat{\beta}_1\hat{\beta}_2[1-\gamma-I_c(|\hat{\alpha}-\alpha|/\sqrt{n^{-1}+\bar{V}S^{-1}\bar{V}'})]\}}{E_0\{\hat{\beta}_1^2[1-\gamma-I_c(|\hat{\alpha}-\alpha|/\sqrt{n^{-1}+\bar{V}S^{-1}\bar{V}'})]\}}$$

and let $k=\{4-[2\varepsilon/(p+\varepsilon)]^2\}/\{p\mu^2[2\varepsilon/(p+\varepsilon)]^2\}$, *where* ε *is a positive constant such that* $\varepsilon<2/g-2$. *Then*

$$E_0\left\{\xi_1\xi_2\left[1-\gamma-I_c\left(\frac{|\hat{\alpha}-\alpha|}{\sqrt{n^{-1}+\bar{V}S^{-1}\bar{V}'}}\right)\right]\right\}<0,$$

where $\xi=\Omega^{-1/2}\hat{\beta}$ *and* $\Omega^{-1/2}$ *is specified in Lemma* 2.2.4.

The proof of Lemma 3.2.2 is given in Wang (1998b), by a similar argument to that in Lemma 2.2.2. The proofs of Lemmas 3.2.3–3.2.4 are similar to those of Lemmas 2.2.3 and 2.2.5 and are omitted.

PROOF OF THEOREM 3.2.1. If $E_0\{\hat{\beta}_1\hat{\beta}_2[1-\gamma-I_c(|\hat{\alpha}-\alpha|/\sqrt{n^{-1}+\bar{V}S^{-1}\bar{V}'})]\}\geq 0$, then by using the lemmas in Section 3, the proof of the first part is established by the same argument as in Theorem 2.2.6. If $E_0\{\hat{\beta}_1\hat{\beta}_2[1-\gamma-I_c(|\hat{\alpha}-\alpha|/\sqrt{n^{-1}+\bar{V}S^{-1}\bar{V}'})]\}\leq 0$, the proof is also by the similar argument, see Wang (1998b). □

Now, we consider another confidence interval

$$C^*_{\hat\alpha} = \{\alpha\colon |\hat\alpha - \alpha| \le c_1\},$$

where c_1 is any positive constant. Let $r(V) = P(|\hat\alpha - \alpha| \le c_1 | V)$.

THEOREM 3.2.5. *Under the assumptions of Theorem* 3.2.1, $r(V)$ *is an inadmissible estimator of* $I(\alpha \in C^*_{\hat\alpha})$ *under squared error loss if* $p \ge 6$. *A better estimator is given by*

$$r_1(\hat\beta) = \begin{cases} r(V) + \dfrac{a_1^*(\mathbf{1}'\Omega^{-1/2}\hat\beta)^2}{2p(b_1^* + \hat\beta'\Omega^{-1}\hat\beta)^2} + \dfrac{a_1^*}{(b_1^* + \hat\beta'\Omega^{-1}\hat\beta)}, \\ \qquad\qquad \text{if } E_0\{\hat\beta_1\hat\beta_2[r(V) - I_c(|\hat\alpha - \alpha|)]\} > 0, \\ r(V) + \dfrac{a_2^*}{b_2^* + |\hat\beta|^2}, \\ \qquad\qquad \text{if } E_0\{\hat\beta_1\hat\beta_2[r(V) - I_c(|\hat\alpha - \alpha|)]\} \le 0, \end{cases}$$

where a_1^*, b_1^*, a_2^*, *and* b_2^* *are some positive constants.*

The previous discussion focuses on the intercept parameter. It is, of course, important to have a similar result for the slope parameter. Let $q = l_1\alpha + l_2'\beta$, where $l_1 \in R$ and $l_2 \in R^p$ are known constants.

THEOREM 3.2.6. *Under the assumptions of Theorem* 3.2.1, *let* $C_{\hat q} = \{q\colon |l_1\hat\alpha + l_2'\hat\beta - q|/\sqrt{\mathrm{Var}(l_1\hat\alpha + l_2'\hat\beta)} \le c_2\}$, *where* c_2 *is the* $1-\gamma/2$ *cutoff point of* $N(0,1)$. *Then, for* $p \ge 6$, $1-\gamma$ *is inadmissible for estimating* $I(q \in C_{\hat q})$ *under squared error loss. A better estimator is given by*

$$r_2(\hat\beta) = \begin{cases} 1 - \gamma + \dfrac{a_3(\mathbf{1}'\Omega^{-1/2}\hat\beta)^2}{2p(b_3 + \hat\beta'\Omega^{-1}\hat\beta)^2} + \dfrac{a_3}{(b_3 + \hat\beta'\Omega^{-1}\hat\beta)}, \\ \qquad\qquad \text{if } E_0\left(\hat\beta_1\hat\beta_2\left(1 - \gamma - I_c\left(\dfrac{|l_1\hat\alpha + l_2'\hat\beta - k|}{\sqrt{\mathrm{Var}(l_1\hat\alpha + l_2'\hat\beta)}}\right)\right)\right) \ge 0, \\ 1 - \gamma + \dfrac{a_4}{b_4 + |\hat\beta|^2}, \\ \qquad\qquad \text{if } E_0\left(\hat\beta_1\hat\beta_2\left(1 - \gamma - I_c\left(\dfrac{|l_1\hat\alpha + l_2'\hat\beta - k|}{\sqrt{\mathrm{Var}(l_1\hat\alpha + l_2'\hat\beta)}}\right)\right)\right) \le 0, \end{cases}$$

where a_3, b_3, a_4 *and* b_4 *are some positive constants.*

Theorems 3.2.5 and 3.2.6 can be established by arguments similar to Theorem 3.2.1.

REMARK 2. For the usual confidence interval of a linear combination of the intercept and slope parameters, the usual constant coverage probability estimator is shown to be inadmissible for estimating the coverage function of the confidence interval if the dimension of the slope parameters is greater than 6.

4. Simulation results. In Sections 2.2 and 3.2, $r_1(X)$ and $r(\hat{\beta})$ are shown to dominate the usual constant coverage probability estimator. In this section, simulation results are presented to show the practical gains that are potentially available. Table 1 compares the risks $R(r_1(X), \theta)$ and $R(1-\gamma, \theta)$, with $p = 8$, $\mu = 10$, $c = 1.96$, $\theta = (l, 2, 4, 6, 8, 10, 12, 14)$, $a = 0.5$, $b = 10$ and ε in Lemma 2.2.5 chosen to be $0.8 \times (2/g - 2)$.

Table 2 gives the results under the same conditions as Table 1, except that $\mu = 50$ and $\theta = (l, 1, -2, 3, 2, -3, -1, -5)$.

Table 3 provides the ratio of the risks in the regression model in Section 3.2 when $p = 6$, $\mu = 1$, $c = 1.96$, $\beta = (l, 3, 7, 4, 9, 20)$, $a = 0.5$, $b = 10$ and ε in Lemma 3.2.4 is chosen to be $0.8 \times (2/g - 2)$.

Table 4 provides the results under the same conditions as Table 3, except that $\mu = -20$ and $\beta = (l, -10, 1, 30, 4, 1)$.

All simulations were based on 10,000 replications, so that the simulation error is about 0.01 (2 standard deviations). From the above tables, it is clear that substantial improvements in risk are sometimes, but not always, available.

TABLE 1

l	**1**	**5**	**10**	**50**	**100**
$\dfrac{R(r_1(x), \theta)}{R(1-\gamma, \theta)}$	0.94	0.93	0.95	0.98	0.99

TABLE 2

l	**1**	**5**	**10**	**50**	**100**
$\dfrac{R(r_1(x), \theta)}{R(1-\gamma, \theta)}$	0.7	0.77	0.89	0.991	0.998

TABLE 3

l	**1**	**5**	**10**	**50**	**100**
$\dfrac{R(r(\hat{\beta}), \beta)}{R(1-\gamma, \beta)}$	0.96	0.95	0.95	0.992	0.998

TABLE 4

l	**1**	**5**	**10**	**50**	**100**
$\dfrac{R(r(\hat{\beta}), \beta)}{R(1-\gamma, \beta)}$	0.98	0.98	0.98	0.994	0.998

Simulation results also show that, if a is between 0 and 1 and $b \geq 10$, then $R(r_1(x), \theta)$ and $R(r(\hat{\beta}), \beta)$ are better than $R(1-\gamma, \theta)$ and $R(1-\gamma, \beta)$, respectively. Therefore, a suggested choice of a and b is $a = 0.5$ and $b = 10$.

APPENDIX

PROOF OF LEMMA 2.2.1. Since

$$\frac{W_2X_2 + \cdots + W_pX_p}{|W|}\Big| W \sim N(0, \sigma_W^2),$$

where $\sigma_W^2 = (W_2^2 + \cdots + W_p^2)/|W|^2$,

$$E_0\left[I_c\left(\frac{|W'X|}{|W|}\right)\Big| X_1, W\right]$$

is a decreasing function with respect to $|X_1|$ for every W. Thus

$$\begin{aligned} &E_0\left\{X_1^2\left[1-\gamma - I_c\left(\frac{|W'X|}{|W|}\right)\right]\right\} \\ &\qquad = E_0\left\{X_1^2 E_0\left[1-\gamma - I_c\left(\frac{|W'X|}{|W|}\right)\Big| X_1, W\right]\right\} > 0. \end{aligned}$$ □

PROOF OF LEMMA 2.2.2. Since

$$E_0\left\{X_1X_2\left[1-\gamma - I_c\left(\frac{|W'X|}{|W|}\right)\right]\right\} = -E_0\left[X_1X_2I_c\left(\frac{|W'X|}{|W|}\right)\right],$$

it suffices to show that

$$E_0\left[X_1X_2I_c\left(\frac{|W'X|}{|W|}\right)\right] < 0. \tag{A.1}$$

Note that the left-hand side of inequality (A.1) equals

$$\begin{aligned} &\int x_1x_2\left[\int I_c\left(\frac{|w'x|}{|w|}\right)\phi_{p-2}f_2(w_1, w_2)dx_3\cdots dx_p dw_1 dw_2\right] \\ &\qquad \times \phi_2(x_1, x_2)f_{p-2}(w_3, \ldots, w_p)dx_1dx_2dw_3\cdots dw_p, \end{aligned}$$

where $\phi_p(x_1, \ldots, x_p)$, $\phi_2(x_1, x_2)$ and $\phi_{p-2}(x_3, \ldots, x_p)$, respectively, denote the joint p.d.f.'s of $N(0, I_{p\times p})$, $N(0, I_{2\times 2})$ and $N(0, I_{(p-2)\times(p-2)})$, and that $f_p(w_1, \ldots, w_p)$, $f_2(w_1, w_2)$ and $f_{p-2}(w_3, \ldots, w_p)$, respectively, denote the joint p.d.f.'s of $(W_1, \ldots, W_p)$, (W_1, W_2) and $(W_3, \ldots, W_p)$. Let

$$\begin{aligned} &p(x_1, x_2, w_1, \ldots, w_p) \\ &\qquad = E_0\left[I_c\left(\frac{|W'X|}{|W|}\right)\Big| X_1 = x_1,\ X_2 = x_2,\ W_i = w_i,\ i = 1, \ldots, p\right]. \end{aligned}$$

It follows that

$$E_0\left[X_1X_2I_c\left(\frac{|W'X|}{|W|}\right)\right] = \int x_1x_2q(x_1, x_2, |w_1|, |w_2|, w_3, \dots, w_p)\phi_2(x_1, x_2) \times f_{p-2}(w_3, \dots, w_p)d|w_1|d|w_2|dx_1dx_2dw_3\cdots dw_p,$$

where

$$\begin{aligned} q(x_1&, x_2, |w_1|, |w_2|, w_3, \dots, w_p)\\ &= p(x_1, x_2, |w_1|, |w_2|, w_3, \dots, w_p)f_2(|w_1|, |w_2|)\\ &\quad + p(x_1, x_2, -|w_1|, -|w_2|, w_3, \dots, w_p)f_2(-|w_1|, -|w_2|)\\ &\quad + p(x_1, x_2, -|w_1|, |w_2|, w_3, \dots, w_p)f_2(-|w_1|, |w_2|)\\ &\quad + p(x_1, x_2, |w_1|, -|w_2|, w_3, \dots, w_p)f_2(|w_1|, -|w_2|). \end{aligned}$$

For convenience, we use $p(x_1, x_2, w_1, w_2)$ and $q(x_1, x_2, w_1, w_2)$ to denote $p(x_1, x_2, w_1, w_2, \dots, w_p)$ and $q(x_1, x_2, w_1, w_2, \dots, w_p)$ in the following. It is then obvious that (A.1) can be established if we could show for any fixed $x_1, x_2, w_1, \dots, w_p$ that

$$q(|x_1|, |x_2|, |w_1|, |w_2|) - q(-|x_1|, |x_2|, |w_1|, |w_2|) < 0$$

and

$$q(-|x_1|, -|x_2|, |w_1|, |w_2|) - q(|x_1|, -|x_2|, |w_1|, |w_2|) < 0.$$

Note that, given $W_1, W_2, \dots, W_p, X_1$ and X_2, the conditional distribution of $|W'X|$ under the condition $\theta = 0$ depends on W_1, W_2, X_1 and X_2 only through $W_1X_1 + W_2X_2$. Therefore,

$$\begin{aligned} p(|x_1|, |x_2|, |w_1|, |w_2|) &= p(|x_1|, |x_2|, -|w_1|, -|w_2|)\\ &= p(-|x_1|, |x_2|, -|w_1|, |w_2|)\\ &= p(-|x_1|, |x_2|, |w_1|, -|w_2|) \end{aligned} \tag{A.2}$$

and

$$\begin{aligned} p(|x_1|, |x_2|, -|w_1|, |w_2|) &= p(|x_1|, |x_2|, |w_1|, -|w_2|)\\ &= p(-|x_1|, |x_2|, |w_1|, |w_2|)\\ &= p(-|x_1|, |x_2|, -|w_1|, -|w_2|). \end{aligned} \tag{A.3}$$

Also, we have

$$f_2(|w_1|, |w_2|) + f_2(-|w_1|, -|w_2|) > f_2(-|w_1|, |w_2|) + f_2(|w_1|, -|w_2|), \tag{A.4}$$

since W_1 and W_2 are independent and

$$\begin{aligned} f_1(|w_1|)&f_1(|w_2|) + f_1(-|w_1|)f_1(-|w_2|)\\ &- f_1(-|w_1|)f_1(|w_2|) - f_1(|w_1|)f_1(-|w_2|)\\ &\quad = [f_1(|w_1|) - f_1(-|w_1|)][f_1(|w_2|) - f_1(-|w_2|)] > 0, \end{aligned}$$

where f_1 is the density of W_1. The last strict inequality holds because W_1 is unimodal and symmetric about μ. Thus, by (A.2) and (A.3) we have

$$\begin{aligned}
&q(|x_1|, |x_2|, |w_1|, |w_2|) - q(-|x_1|, |x_2|, |w_1|, |w_2|)\\
&\quad = [p(|x_1|, |x_2|, |w_1|, |w_2|) - p(-|x_1|, |x_2|, |w_1|, |w_2|)]\\
&\qquad \times [f_2(|w_1|, |w_2|) + f_2(-|w_1|, -|w_2|)\\
&\qquad\quad - f_2(-|w_1|, |w_2|) - f_2(|w_1|, -|w_2|)],
\end{aligned}$$

which is less than zero by (A.4) and the fact that

$$p(|x_1|, |x_2|, |w_1|, |w_2|) - p(-|x_1|, |x_2|, |w_1|, |w_2|) < 0.$$

Similarly,

$$\begin{aligned}
&q(-|x_1|, -|x_2|, |w_1|, |w_2|, w_3, \dots, w_p)\\
&\quad - q(|x_1|, -|x_2|, |w_1|, |w_2|, w_3, \dots, w_p) < 0,
\end{aligned}$$

establishing the lemma. □

PROOF OF LEMMA 2.2.3. Since X_1 and X_2 are identically distributed, (5) is equivalent to

$$\begin{aligned}
&E_0\left\{X_1^2\left[1-\gamma-I_c\left(\frac{|W'X|}{|W|}\right)\right]\right\} + E_0\left\{X_2^2\left[1-\gamma-I_c\left(\frac{|W'X|}{|W|}\right)\right]\right\}\\
&\quad - 2E_0\left\{X_1X_2\left[1-\gamma-I_c\left(\frac{|W'X|}{|W|}\right)\right]\right\} > 0
\end{aligned}$$

which is equivalent to

$$E_0\left\{(X_1-X_2)^2\left[1-\gamma-I_c\left(\frac{|W'X|}{|W|}\right)\right]\right\} > 0. \tag{A.5}$$

Let $K_1 = (X_1 - X_2)/\sqrt{2}$, $K_2 = (X_1 + X_2)/\sqrt{2}$ and $K_i = X_i\ \forall\ i = 3 \cdots p$. It is easy to show that the K_i are i.i.d. $N(0,1)$ under the assumption that $\theta = 0$. Consequently,

$$\begin{aligned}
&E_0\left\{(X_1-X_2)^2\left[1-\gamma-I_c\left(\frac{|W'X|}{|W|}\right)\right]\right\}\\
&\quad = 2E_0\left\{K_1^2\left[1-\gamma-I_c\left(\left|K_1\left(\frac{W_1-W_2}{\sqrt{2}}\right)+K_2\left(\frac{W_1+W_2}{\sqrt{2}}\right)\right.\right.\right.\right.\\
&\qquad\qquad\qquad\qquad \left.\left.\left.\left.+ W_3K_3 + \cdots + W_pK_p\right| \cdot |W|^{-1}\right)\right]\right\} > 0.
\end{aligned}$$

The last inequality follows from an argument similar to that for Lemma 2.2.1 and hence the proof is complete. □

PROOF OF LEMMA 2.2.4. The proof is by directly solving

$$\Omega^{-1/2}\Omega^{-1/2} = \Omega^{-1} = I - \frac{k}{1+kp\mu^2}UU'.$$

See Wang (1998b).

PROOF OF LEMMA 2.2.5. By definition,

$$\begin{aligned}
&E_0\left\{Y_1Y_2\left[1-\gamma-I_c\left(\frac{|W'X|}{|W|}\right)\right]\right\}\\
&\quad= E_0\biggl\{(eX_1+dX_2+\cdots+dX_p)\\
&\qquad\qquad \times (dX_1+eX_2+dX_3+\cdots+dX_p)\\
&\qquad\qquad\qquad \times \left[1-\gamma-I_c\left(\frac{|W'X|}{|W|}\right)\right]\biggr\}\\
&\quad= E_0\left\{X_1^2\left[1-\gamma-I_c\left(\frac{|W'X|}{|W|}\right)\right]\right\}\\
&\qquad \times \bigl\{e^2g+2ed[g(p-2)+1]+d^2(p-2)\\
&\qquad\qquad + gd^2[p-1+(p-2)^2]\bigr\},
\end{aligned}\tag{A.6}$$

where $e = 1 + h_k\mu^2$, $d = h_k\mu^2$ and h_k is given in Lemma 2.2.4. For the value of k given in (7), straightforward calculation yields

$$e = -(p-1+\varepsilon)d.$$

Substituting this into (A.6), we have

$$E_0\left\{X_1^2\left[1-\gamma-I_c\left(\frac{|W'X|}{|W|}\right)\right]\right\}d^2[p(g-1)+\varepsilon(g\varepsilon+2g-2)] < 0,$$

where the last inequality follows from Lemma 2.2.1 and the fact that $0 < g < 1$ (Lemma 2.2.3) and $\varepsilon < 2/g - 2$. The proof is complete. □

Acknowledgments. This paper is part of the author's Ph.D. thesis submitted to Tsing Hua University. The author thanks Professor Gene Hwang at Cornell University for his supervision. I also thank Professor Yuang-Chin Chiang at Tsing Hua University, Taiwan, and other Ph.D. committee members for their kind help and useful comments. Also, I am grateful to the referee, the Associate Editor and the Editor for their many valuable comments and suggestions, which substantially improved this paper.

REFERENCES

BROWN, L. D. (1990). An ancillarity paradox which appears in multiple linear regression (with discussion). *Ann. Statist.* **18** 471–538.

KIEFER, J. (1977). Conditional confidence statements and confidence estimators. *J. Amer. Statist. Assoc.* **72** 789–827.

LU, K. L. and BERGER, J. O. (1989). Estimated confidence procedures for multivariate normal means. *J. Statist. Plann. Inference* **23** 1–19.

ROBERT, C. and CASELLA, G. (1994). Improved confidence statements for the usual multivariate normal confidence set. In *Statistical Decision Theory and Related Topics 5* (S. S. Gupta and J. O. Berger, eds.) 351–368. Springer, New York.

ROBINSON, G. K. (1979a). Conditional properties of statistical procedures. *Ann. Statist.* **7** 742–755.

ROBINSON, G. K. (1979b). Conditional properties of statistical procedures for location and scale parameters. *Ann. Statist.* **7** 756–771.

SANDVED, E. (1968). Ancillary statistics and prediction of the loss in estimation problems. *Ann. Math. Statist.* **39** 1756–1758.

WANG, H. (1998a). Admissibility of the constant coverage estimator for estimating the coverage function. *Statist. Probab. Lett.* **36** 365–372.

WANG, H. (1998b). Brown's paradox in the estimated confidence approach. Technical Report, Institute of Statistical Science, Academia Sinica, Taiwan.

INSTITUTE OF STATISTICAL SCIENCE
ACADEMIA SINICA
TAIPEI 11529
TAIWAN
E-MAIL: hywang@stat.sinica.edu.tw