The Annals of Mathematical Statistics
1972, Vol. 43, No. 5, 1541-1548

SHARP ONE-SIDED CONFIDENCE BOUNDS OVER POSITIVE REGIONS[1]

BY ROBERT BOHRER AND GEORGE K. FRANCIS

University of Illinois at Urbana-Champaign

The paper develops one-sided analogs to Scheffé's two-sided confidence bounds for a function $f(\mathbf{x}), \mathbf{x} \in R^n$. If the domain X^* of f is a subset of $R_+^n = \{\mathbf{x} : x_i \geqq 0, \forall i\}$, then the upper Scheffé bounds are conservative upper confidence bounds, which can be sharpened, often to great practical advantage. This sharpening, accomplished by a non-trivial extension of Scheffé's method, is developed by the geometry-probability argument of Section 2. Section 3 derives coverage probabilities for general 2- and 3-parameter functions and illustrates savings by the sharp bounds in two examples.

1. Upper confidence bounds. We consider confidence bounds for a function $f(\mathbf{x}; \boldsymbol{\beta}) = \mathbf{x}'\boldsymbol{\beta}$ based on statistics $(\hat{\boldsymbol{\beta}}, S)$, where $\hat{\boldsymbol{\beta}}$ is normal $N(\boldsymbol{\beta}, B\sigma^2)$ and where $\nu S^2/\sigma^2$ is $\chi^2(\nu)$ independent of $\hat{\boldsymbol{\beta}}$. The parameters of the n by 1 vector $\boldsymbol{\beta}$ and σ^2 are not known, but the elements of the symmetric, positive definite n by n matrix B are known.

For example, this is the case in the general analysis of variance ([7] Chapter 2), where $\hat{\boldsymbol{\beta}}$ is the vector of least squares estimators with variance $B\sigma^2$ and S^2 is the usual unbiased estimator of σ^2.

A *coefficient-α upper confidence bound* for the function $f(\mathbf{x}; \boldsymbol{\beta})$, $\mathbf{x} \in X^*$, based on $(\hat{\boldsymbol{\beta}}, S)$, is a random function $U(\mathbf{x}; \hat{\boldsymbol{\beta}}, S)$, $\mathbf{x} \in X^*$, such that

$$\text{(1.1)} \qquad \Pr_{(\boldsymbol{\beta}, \sigma^2)} \{f(\mathbf{x}; \boldsymbol{\beta}) \leqq U(\mathbf{x}; \hat{\boldsymbol{\beta}}, S), \ \ \forall \mathbf{x} \in X\} \geqq 1 - \alpha$$

holds uniformly over all $(\boldsymbol{\beta}, B\sigma^2)$. The bound is said to be *sharp* if equality holds for the second inequality sign in (1.1) for all $(\boldsymbol{\beta}, B\sigma^2)$; if for some $(\boldsymbol{\beta}, B\sigma^2)$ the inequality is strict, the bound is said to be *conservative*. Coefficient-α lower confidence bounds $L(\mathbf{x}; \hat{\boldsymbol{\beta}}, S)$ are defined by reversing the interior inequality of (1.1). Their analysis reverts to that of upper bounds if f is replaced by its negative. Evidently, if

$$\text{(1.2)} \qquad \Pr_{(\boldsymbol{\beta}, \sigma^2)} \{L(\mathbf{x}; \hat{\boldsymbol{\beta}}, S) \leqq f(\mathbf{x}; \boldsymbol{\beta}) \leqq U(\mathbf{x}; \hat{\boldsymbol{\beta}}, S), \ \ \forall \mathbf{x} \in X^{**}\} \geqq 1 - \alpha ,$$

then every function U_1 exceeding U on a subset X_1^* of X^* is also a coefficient-α upper confidence bound on X_1^*.

Scheffé ([7] Section 3.5) considers two-sided bounds for f over all of n-space R^n, of the form $U, L = \mathbf{x}'\hat{\boldsymbol{\beta}} \pm cSS_{\mathbf{x}}$, where $S_{\mathbf{x}}^2 = \operatorname{Var} f(\mathbf{x}; \hat{\boldsymbol{\beta}})/\sigma^2 = \mathbf{x}'B\mathbf{x}$, which is proportional to the variance $\sigma_{\mathbf{x}}^2$ of the unbiased estimator $f(\mathbf{x}; \hat{\boldsymbol{\beta}})$. Hence the

Received March 26, 1971; revised January 20, 1972.
[1] Supported partially by NSF grants 23835 and 13145.
AMS subject classifications. Primary 6240, 6256; Secondary 1570, 5040.

expected excess of an upper bound of this type, $E\{U(\mathbf{x}; \hat{\boldsymbol{\beta}}, S) - f(\mathbf{x}; \boldsymbol{\beta})\}$, is proportional to the precision of $f(\mathbf{x}; \hat{\boldsymbol{\beta}})$, as measured by $\sigma_{\mathbf{x}}$. Scheffé found that the value $c^* = (nF_\alpha(n, \nu))^{\frac{1}{2}}$ yields sharp two-sided bounds over R_n:

$$\text{(1.3)} \qquad \Pr_{(\beta,\sigma^2)}\{|\mathbf{x}'(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})| \leqq c^*S(\mathbf{x}'\beta\mathbf{x})^{\frac{1}{2}}, \ \forall \mathbf{x} \in R^n\} = 1 - \alpha ,$$

where $F_\alpha(n, \nu)$ is the $100(1 - \alpha)$ percentile of the $F(n, \nu)$ distribution.

Surprisingly ([3] Theorem 1), the upper bound $U^* = \mathbf{x}'\hat{\boldsymbol{\beta}} + c^*SS_{\mathbf{x}}$ is also sharp as a one-sided bound over R^n. This is in striking contrast to the case of bounding f at a single point $\mathbf{x}_0$. Then coefficient-α two-sided bounds are $\mathbf{x}_0'\hat{\boldsymbol{\beta}} \pm SS_{\mathbf{x}_0}(F_\alpha(1, \nu))^{\frac{1}{2}}$, whereas, for $\alpha \leqq \frac{1}{2}$, the one-sided bounds $\mathbf{x}_0'\hat{\boldsymbol{\beta}} + SS_{\mathbf{x}_0}(F_{2\alpha}(1, \nu))^{\frac{1}{2}}$ are shorter by a factor of $(F_{2\alpha}(1, \nu)/F_\alpha(1, \nu))^{\frac{1}{2}} < 1$.

However, if, as in the case of the two examples we present in Section 3, the domain X^* is a subset of the nonnegative orthant $R_+^{\ n} = \{\mathbf{x} : x_i \geqq 0, \forall i\}$, then the value $c^\sharp$ which gives sharp bounds can be considerably smaller than c^*. Indeed, for the most tractable case of B diagonal and σ^2 known, up to a 30 percent saving was noted in [3]. As we shall see, even greater improvement can obtain in the more general case considered here.

Note throughout that the case of variance known is obtained as the limit as $\nu \to \infty$ in the present case.

Sharp two-sided bounds over $R_+^{\ n}$ for the case that $B = I$, the identity matrix, have been treated in [2]. Sharp one-sided bounds for linear regression over an interval are treated in [4]. There we also compare the average width of these bounds, which, being proportional to $\sigma_{\mathbf{x}}$, yield hyperbolic confidence bands about the estimated regression line, with sharp one-sided bounds of constant width. The question of the optimum shape for this criterion in general regression will be taken up in a subsequent paper by the first author.

2. Sharp bounds on $R_+^{\ n}$. We define the *coverage probability* as

$$\text{(2.1)} \qquad \mathscr{P}(c) = \Pr_{(\beta,\sigma^2)}\{\mathbf{x}'\boldsymbol{\beta} \leqq \mathbf{x}'\hat{\boldsymbol{\beta}} + cS(\mathbf{x}'\mathbf{B}\mathbf{x})^{\frac{1}{2}}, \ \forall \mathbf{x} \in R_+^{\ n}\} ,$$

where S is a statistic independent of the $N(\boldsymbol{\beta}, B\sigma^2)$ statistic $\hat{\boldsymbol{\beta}}$ such that $\nu S^2/\sigma^2$ is a random variable with a $\chi^2(\nu)$ distribution, B is a positive definite symmetric n by n matrix, and $R_+^{\ n} = \{\mathbf{x} : x_i \geqq 0, i = 1, \cdots, n\}$ is the nonnegative orthant in n-space. Let $S^{n-1} = \{\mathbf{x} : ||\mathbf{x}|| = 1\}$ be the unit hypersphere. Let $V = [\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n]$ be a square root of B ([1] page 277) with column vectors $\mathbf{v}_i$, i.e. $B = V'V$. Let $W = V'^{-1} = [\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_n]$ be the transpose inverse of V. For a subset of indices, $P \subset \{1, 2, \cdots, n\}$, let P' denote the complement of P and card $(P) = p$, the number of indices in P. The independent column vectors of V and W determine the following spherical simplices and associated ratios:

$$\text{(2.2)} \qquad \begin{aligned} \Delta_P &= \{\mathbf{y} = \textstyle\sum_{i \in P} \alpha_i \mathbf{v}_i;\ \alpha_i > 0 \ \text{and} \ ||\mathbf{y}|| = 1\} , \\ \Delta'_{P'} &= \{\mathbf{y} = -\textstyle\sum_{i \in P'} \alpha_i \mathbf{w}_i;\ \alpha_i \geqq 0 \ \text{and} \ ||\mathbf{y}|| = 1\} , \\ \rho_P &= \text{cont}\,(\Delta_P)/\text{cont}\,(S^{p-1}) , \qquad \text{and} \\ \rho'_{P'} &= \text{cont}\,(\Delta'_{P'})/\text{cont}\,(S^{n-p-1}) , \end{aligned}$$

where the content is, respectively, unity, cardinality, length, area, volume, $\cdots$, of a subset of R^n that is respectively empty, finite, one-, two-, three-, $\cdots$, -dimensional.

THEOREM.

$$\mathscr{P}(c) = \sum_P \rho_P \rho'_{P'} \Pr\{F(p, \nu) \leqq c^2/p\} . \tag{2.3}$$

PROOF. Let us denote the probability density function for the $\chi^2(\nu)$ distribution by $f_\nu(s)$, for the $N(\boldsymbol{\beta}, B\sigma^2)$ distribution by $g_{\beta, B\sigma^2}(\mathbf{b})$ and for the distribution of the statistic S by $h_\sigma(s)$. Recall that

$$\begin{aligned} f_\nu(s) &= 2^{-\frac{1}{2}\nu}\Gamma(\tfrac{1}{2}\nu)^{-1} \exp(-\tfrac{1}{2}s)s^{\frac{1}{2}\nu-1} , \qquad s \geqq 0 , \\ g_{\beta, B\sigma^2}(\mathbf{b}) &= (2\pi)^{-\frac{1}{2}n}(\det B)^{-\frac{1}{2}} \exp(-\tfrac{1}{2}(\boldsymbol{\beta} - \mathbf{b})'B^{-1}(\boldsymbol{\beta} - \mathbf{b})/\sigma^2)\sigma^{-n} . \end{aligned} \tag{2.4}$$

Let $D(s) = \{\mathbf{b} : \mathbf{x}'\boldsymbol{\beta} \leqq \mathbf{x}'\mathbf{b} + cs(\mathbf{x}'B\mathbf{x})^{\frac{1}{2}}, \forall \mathbf{x} \in R_+{}^n\}$. Then, the coverage probability is given analytically by

$$\mathscr{P}(c) = \int_0^\infty \int_{D(s)} g_{\beta, B\sigma^2}(\mathbf{b})\, d\mathbf{b} h_\sigma(s)\, ds . \tag{2.5}$$

A change of variable, $\boldsymbol{\theta} = W(\boldsymbol{\beta} - \mathbf{b})/\sigma$, on the inside integral of (2.5) leads to the identity

$$\int_{D(s)} g_{\beta, B\sigma^2}(\mathbf{b})\, d\mathbf{b} = \int_{R(\lambda)} g_{\mathbf{0}, I}(\boldsymbol{\theta})\, d\boldsymbol{\theta} , \tag{2.6}$$

where $\lambda = cs/\sigma$ and $R(\lambda) = \{\boldsymbol{\theta} : \mathbf{y}'\boldsymbol{\theta} \leqq \lambda||\mathbf{y}||, \forall \mathbf{y} \in VR_+{}^n\}$. The transform of the integrand follows from (2.4). The transform of the domain of integration follows from the fact that setting $\mathbf{y} = V\mathbf{x}$ leads to $\mathbf{x}'(\boldsymbol{\beta} - \mathbf{b})/\sigma = \mathbf{y}'\boldsymbol{\theta}$ and $cs(\mathbf{x}'B\mathbf{x})^{\frac{1}{2}}/\sigma = \lambda||\mathbf{y}||$. Hence $\mathbf{y}'\boldsymbol{\theta} \leqq \lambda||\mathbf{y}||$ if and only if $\mathbf{x}'(\boldsymbol{\beta} - \mathbf{b}) \leqq cs(\mathbf{x}'B\mathbf{x})^{\frac{1}{2}}$.

To calculate the right side of (2.6) we decompose $B(\lambda)$ into a union of subregions with mutually disjoint interiors. For a subset $X \subset R^n$ and $0 \leqq \lambda \leqq \infty$, denote by $X\langle\lambda\rangle$ the *cone* of radius λ on X, that is

$$X\langle\lambda\rangle = \{\mathbf{y} : \mathbf{y} = t\mathbf{x}, \mathbf{x} \in X, t \geqq 0 \text{ and } ||\mathbf{y}|| \leqq \lambda\} .$$

Given two subsets $X, Y \subset R^n$, such that $\mathbf{x}'\mathbf{y} = 0$ for all $\mathbf{x} \in X$ and $\mathbf{y} \in Y$, we shall set $X \oplus Y = \{\mathbf{x} + \mathbf{y} : \mathbf{x} \in X, \mathbf{y} \in Y\}$.

LEMMA. *The region $R(\lambda)$ is the union of 2^n regions with mutually disjoint interiors given by*

$$R_P(\lambda) = \Delta_P\langle\lambda\rangle \oplus \Delta'_{P'}\langle\infty\rangle , \qquad P \subset \{1, 2, \cdots, n\} .$$

PROOF. The matrix $W^{-1}V$, being positive definite, has positive principal minors. Hence ([6] page 807) the column vectors of V and $-W$ comprise a special case of a "partition" ([6] page 805) of R^n. That is, the 2^n cones $\Delta_P\langle\infty\rangle \oplus \Delta'_{P'}\langle\infty\rangle$, $P \subset \{1, 2, \cdots, n\}$, have mutually disjoint interiors and their union fills out R^n. Thus, for a vector $\boldsymbol{\theta}$ there is a $P \supset \{1, 2, \cdots, n\}$, with $\boldsymbol{\theta} = \boldsymbol{\theta}_P + \boldsymbol{\theta}'_{P'} \in \Delta_P\langle\infty\rangle \oplus \Delta'_{P'}\langle\infty\rangle$. If $\boldsymbol{\theta} \in R_P(\lambda)$ then $||\boldsymbol{\theta}_P|| \leqq \lambda$. So for every $\mathbf{y} \in VR_+{}^n$,

$$\mathbf{y}'\boldsymbol{\theta} = \mathbf{y}'\boldsymbol{\theta}_P \leqq ||\mathbf{y}||\,||\boldsymbol{\theta}_P|| \leqq \lambda||\mathbf{y}||$$

and $\boldsymbol{\theta} \in R(\lambda)$. Conversely, if $\boldsymbol{\theta} \in R(\lambda)$, choose $\mathbf{y} = \boldsymbol{\theta}_P \in VR_+{}^n$. Then $\|\boldsymbol{\theta}_P\|^2 = \mathbf{y}'\boldsymbol{\theta} \leqq \lambda\|\mathbf{y}\| = \lambda\|\boldsymbol{\theta}_P\|$. So $\boldsymbol{\theta}_P \in \Delta_P\langle\lambda\rangle$ and $\boldsymbol{\theta} \in R_P(\lambda)$. The lemma is proved. □

As a consequence of this lemma and (2.4), we have

$$\int_{R(\lambda)} g_{0,I}(\boldsymbol{\theta})\, d\boldsymbol{\theta} = \sum_P \int_{R_P(\lambda)} g_{0,I}(\boldsymbol{\theta})\, d\boldsymbol{\theta}\,, \quad \text{where} \tag{2.7}$$

$$\int_{R_P(\lambda)} g_{0,I}(\boldsymbol{\theta})\, d\boldsymbol{\theta} = \int_{\Delta_P\langle\lambda\rangle} g_{0,I_P}(\boldsymbol{\theta}_P)\, d\boldsymbol{\theta}_P \int_{\Delta'_{P'}\langle\infty\rangle} g_{0,I_{P'}}(\boldsymbol{\theta}'_{P'})\, d\boldsymbol{\theta}'_{P'}\,.$$

To evaluate the first factor integral in (2.7) we change to p-dimensional polar coordinates ([8] page 53f.), setting $r = \|\boldsymbol{\theta}_P\|$ and $d\Omega =$ the area density on S^{p-1}. Thus we have

$$\int_{\Delta_P\langle\lambda\rangle} g_{0,I_P}(\boldsymbol{\theta}_P)\, d\boldsymbol{\theta}_P = (2\pi)^{-\frac{1}{2}p} \int_0^\lambda e^{-\frac{1}{2}r^2} r^{p-1}\, dr \int_{\Delta_P} d\Omega\,. \tag{2.8}$$

Since $\int_{\Delta_P} d\Omega = \text{cont}\,(\Delta_P) = \rho_P \,\text{cont}\,(S^{p-1}) = \rho_P 2\Gamma(\frac{1}{2}p)^{-1}\pi^{\frac{1}{2}p}$, changing variable to $\mu = r^{\frac{1}{2}}$, we obtain

$$(2.8) = 2^{-\frac{1}{2}p}\Gamma(\tfrac{1}{2}p)^{-1} \int_0^{\lambda^2} e^{-\frac{1}{2}\mu} \mu^{\frac{1}{2}p-1}\, d\mu \rho_P = \rho_P \int_0^{\lambda^2} f_p(r)\, dr\,.$$

A similar argument for the second factor integral in (2.7) leads to its value of $\rho'_{P'}$. Collecting, we have that (2.5) is

$$\mathscr{P}(c) = \sum_P \rho_P \rho'_{P'} \int_0^\infty \int_0^{c^2 s^2/\sigma^2} f_p(r)\, dr h_\sigma(s)\, ds\,. \tag{2.9}$$

We complete the argument as follows. Let X^2 denote a $\chi^2(p)$ random variable that is independent of S^2. The iterated integral in (2.9), is, as a function of p,

$$\Pr\{X^2 \leqq c^2 S^2/\sigma^2\} =$$

$$\Pr\left\{\frac{X^2/p}{(\nu S^2/\sigma^2)/\nu} \leqq \frac{c^2}{p}\right\} = \Pr\{F(\rho, \nu) \leqq c^2/p\}\,. \qquad □$$

Having completed the proof of the theorem, we next consider the limiting cases $c = 0$; $c = \infty$; $B = I$; $\nu = \infty$; $B = I$ and $\nu = \infty$. Note that the ratios ρ_P and $\rho'_{P'}$ are functions of B only. Let $1 - \alpha_0 = \rho'_{\phi'}$ be the ratio corresponding to $\Delta'_{\phi'}$ where ϕ is the empty subset of $\{1, 2, \cdots, n\}$. (Geometrically, this hyperspherical $(n-1)$-simplex is the reflection in the origin of the polar simplex to the fundamental simplex $\text{Clos}\, \Delta_{\{1,2,\cdots,n\}} = S^{n-1} \cap VR_+{}^n$.)

COROLLARY 1.

(a) $\mathscr{P}(0) = 1 - \alpha_0(B) > 0$;

(b) $\mathscr{P}(\infty) = \sum_P \rho_P \rho'_{P'} = 1$;

(c) $\mathscr{P}(c) = \sum_{p=0}^n 2^{-n} \binom{n}{p} \Pr\{F(p, \nu) \leqq c^2/p\}$, *if* $B = I$;

(d) $\mathscr{P}(c) = \sum_P \rho_P \rho'_{P'} \Pr\{\chi^2(p) \leqq c^2\}$, *if* $\nu = \infty$ *(variance known)*;

(e) $\mathscr{P}(c) = 2^{-n} \sum_{p=0}^n \binom{n}{p} \Pr\{X^2(p) \leqq c^2\}$, *if* $B = I$ *and* $\nu = \infty$.

PROOF. Setting $c = 0$ in (2.1) leads to $\lambda = 0$ in (2.6). Hence $R(0) = \{\boldsymbol{\theta} : \mathbf{y}'\boldsymbol{\theta} \leqq 0, \mathbf{y} \in VR_+{}^n\} = \Delta'_{\phi'}\langle\infty\rangle$. Thus we have directly that

$$\mathscr{P}(0) = \int_0^\infty \int_{R(0)} g_{0,I}(\boldsymbol{\theta})\, d\boldsymbol{\theta} h_\sigma(s)\, ds$$

$$= \int_0^\infty \int_{\Delta'_{\phi'}\langle\infty\rangle} g_{0,I}(\boldsymbol{\theta})\, d\boldsymbol{\theta} h_\sigma(s)\, ds = \rho'_{\phi'} \int_0^\infty h_\sigma(s)\, ds = 1 - \alpha_0\,.$$

So (a) holds. For $c \to \infty$ in (2.3) we have the rest of the identity (b). If $B = I$ then $\Delta_P = S_+^{p-1}$ and $\Delta'_{P'} = -S_+^{n-p-1}$, where $S_+^{p-1} = S^{p-1} \cap R_+^n$. Since $\text{cont}\,(S_+^{p-1})/\text{cont}\,(S^{p-1}) = 2^{-p}$, and we may collect together those partitions of equal cardinality to arrive at (c). Since $pF(p, \infty)$ has the distribution of $\chi^2(p)$, we have (d). As a consequence of (c) and (d) we have (e), which was previously obtained in [3]. □

The $F(p, \nu)$ densities involved in (2.3) are continuous and positive for c positive; their coefficients are also positive. Hence $\mathscr{P}(c)$ is a continuously increasing function of c from $1 - \alpha_0$ to 1. Consequently, for every sufficiently large coverage probability $1 - \alpha$, a unique value $c^\#$ of c will yield sharp bounds (2.1). By formula (a) in the preceding corollary, there is a nonzero lower limit $1 - \alpha_0$ for the coverage probability, which depends only on the covariance matrix B.

COROLLARY 2. *For each* α *and* ν, $1 - \alpha_0(B) \leqq 1 - \alpha \leqq 1$ *and* $1 \leqq \nu \leqq \infty$, *there exists a unique* $c^\# = c^\#(\alpha, B, \nu)$ *such that* $\mathscr{P}(c^\#) = 1 - \alpha$.

We next investigate the range of the coverage probability over all covariance matrices.

COROLLARY 3.

$$\sup_B \mathscr{P}(c) = \tfrac{1}{2} + \tfrac{1}{2} \Pr\{F(1, \nu) \leqq c^2\}\,. \tag{2.10}$$

PROOF. Note that for any $\mathbf{x}^0 \in VR_+^n$

$$\begin{aligned}\mathscr{P}(c) &= \Pr\{\mathbf{x}'\boldsymbol{\theta} \leqq cS\|\mathbf{x}\|/\sigma,\ \forall \mathbf{x} \in VR_+^n\} \\ &\leqq \Pr\{\mathbf{x}_0'\boldsymbol{\theta} \leqq cS\|\mathbf{x}_0\|/\sigma\} = (2.10)\,;\end{aligned}$$

the first line follows by using (2.6) in (2.5). This limiting case can in fact be approximated arbitrarily closely by those cases in which the covariance matrix is $B_\varepsilon = J + \varepsilon I$, where J is the matrix of all unit entries, $\varepsilon > 0$ and $\varepsilon \to 0$. Note that $J\mathbf{x} = (\sum x_j)\mathbf{1}$, where $\mathbf{1}$ is the vector all of whose components are equal to one. Hence $J^2 = nJ$. It is easy to check that $V_\varepsilon = [(\varepsilon + n)^{\frac{1}{2}} + \varepsilon^{\frac{1}{2}}]^{-1}J + \varepsilon^{\frac{1}{2}}I$ is a square root of B_ε and that $V_\varepsilon R_+^n$ tends to the ray through $\mathbf{1}$ as $\varepsilon \to 0$. Setting $\mathbf{x}_0 = \mathbf{1}$ above, we have (2.10). □

Note that every region $R(\lambda)$ in the proof of the Theorem properly contains the radius-λ ball $S(\lambda) = \{\boldsymbol{\theta} : \|\boldsymbol{\theta}\| \leqq \lambda\}$. The coverage probability $\mathscr{P}(c)$ therefore exceeds the coverage probability of the λ-ball, which is $\Pr\{F(n, \nu) \leqq c^2/n\}$. In other words, the Scheffé upper confidence bound $c^* = (nF_\alpha(n, \nu))^{\frac{1}{2}}$ discussed in Section 1, exceeds the sharp bound $c^\#$ over R_+^n for all covariances B. We therefore express the relative savings of sharp bounds over Scheffé bounds by

COROLLARY 4.

$$\begin{aligned}&\sup_B (c^\#/c^*) \leqq 1\,, \quad \textit{and for} \quad \alpha \leqq \tfrac{1}{2}\,, \\ &\inf_B (c^\#/c^*) = (F_{2\alpha}(1, \nu)/nF_\alpha(n, \nu))^{\frac{1}{2}}\,, \qquad \textit{and} \\ &\lim_{n\to\infty} \lim_{\nu\to\infty} \inf_B (n^{\frac{1}{2}}c^\#/c^*) = (\chi^2_{2\alpha}(1))^{\frac{1}{2}}\,.\end{aligned}$$

Thus, unlike the case considered in [3], the length of the $c^\ddagger$ bounds relative to the Scheffé bounds can be arbitrarily close to zero for large n.

We close with a general observation on the sharpness of our bounds over a proper subset $X^* \subset R_+{}^n$. The inequality in (2.1) is homogeneous in the vector $\mathbf{x}$ and persists in the limit for any convergent sequence of rays through points $\mathbf{x}_n$ in X^*. Hence, we have

COROLLARY 5. *$U(\mathbf{x}; \hat{\boldsymbol{\beta}}; S) = \mathbf{x}'\hat{\boldsymbol{\beta}} + c^\ddagger S(\mathbf{x}'B\mathbf{x})^{\frac{1}{2}}$ is a sharp coefficient-α upper confidence bound for $f(\mathbf{x}; \boldsymbol{\beta}) = \mathbf{x}'\boldsymbol{\beta}$ on $X^* \subset R_+{}^n$ if and only if the cone $X^*\langle\infty\rangle$ is dense in $R_+{}^n$.*

3. Use and efficiency of sharp bounds. In the case $n = 2$ and 3, we compute the ratios ρ_P and $\rho'_{P'}$ to obtain explicit formulas (3.1) and (3.2) for (2.3). With $n = 2$, the region $R(1)$ defined after (2.6) decomposes as illustrated in Fig. 1. For $P = \{1, 2\}$, P' is empty, so $\rho'_{P'} = 1$. $\Delta_{\{1,2\}}$ is the arc between $\mathbf{v}_1$ and $\mathbf{v}_2$ on the circle, with content equal to the angle in radian measure. We have that arc length $(\Delta_P) = \operatorname{arc}\cos(\mathbf{v}_1'\mathbf{v}_2/||\mathbf{v}_1||\,||\mathbf{v}_2||) = \operatorname{arc}\cos(B_{12}/(B_{11}B_{22})^{\frac{1}{2}})$, since $B = V'V = [\mathbf{v}_i'\mathbf{v}_j]$. So $\rho_{\{1,2\}} = \operatorname{arc}\cos(B_{12}/(B_{11}B_{22})^{\frac{1}{2}})/2\pi$. The case $P = \phi$, hence $P' = \{1, 2\}$, is analogous, and yields $\rho_\phi = 1$ and $\rho'_{\phi'} = \operatorname{arc}\cos(B^{12}/(B^{11}B^{22})^{\frac{1}{2}})/2\pi$, where $B^{-1} = [B^{ij}]$. For $P = \{i\}$, $i = 1, 2$, $\rho_P = \rho'_{P'} = \frac{1}{2}$, since $\Delta_{\{i\}} = \{\mathbf{v}_i/||\mathbf{v}_i||\}$, whereas $S^0 = \{-1, +1\}$. Collecting, the coverage probability is

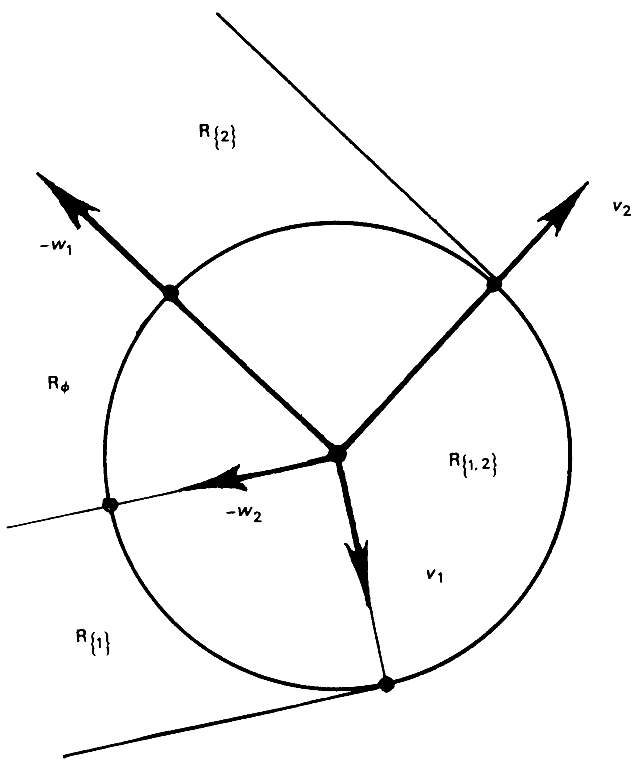

FIG. 1.

$$\mathscr{P}(c) = \arc\cos(B^{12}/(B^{11}B^{22})^{\frac{1}{2}})/2\pi + \tfrac{1}{2}\{\Pr F(1, \nu) \leqq c^2\} + \arc\cos(B_{12}/(B_{11}B_{22})^{\frac{1}{2}}) \Pr\{F(2, \nu) \leqq c^2/2\}/2\pi . \tag{3.1}$$

Observe that the leading term is the residual $1 - \alpha_0(B)$.

For $n = 3$ and $P = \{1, 2, 3\}$, $P' = \phi$, hence $\rho'_{\{1,2,3\}'} = 1$. To evaluate $\rho_{\{1,2,3\}}$ we use Euler's theorem in spherical trigonometry ([8] page 66). $\Delta_{\{1,2,3\}}$ is the spherical triangle with vertices $\mathbf{v}_i/||\mathbf{v}_i||$. Hence Area $(\Delta_{\{1,2,3\}}) = \sum_{i=1}^{3} \varphi_i - \pi$, where the φ_i are the angles at the vertices of the triangle. Equivalently, $\varphi_i = \pi - \varphi_i{}^*$, where $\varphi_i{}^*$ is the arc length of the corresponding side of the polar triangle with vertices $\mathbf{w}_i/||\mathbf{w}_i||$. Since Area $(S^2) = 4\pi$, it follows that $\rho_{\{1,2,3\}} = (4\pi)^{-1}[\sum_{i<j} (\pi - \arc\cos(\mathbf{w}_i'\mathbf{w}_j/||\mathbf{w}_i||\,||\mathbf{w}_j||)) - \pi] = \frac{1}{2} - (4\pi)^{-1} \sum_{i<j} \arc\cos(B^{ij}/(B^{ii}B^{jj})^{\frac{1}{2}})$. Similarly, for $P = \phi$, $\rho_\phi = 1$, and $\rho'_{\phi'} = \frac{1}{2} - (4\pi)^{-1} \arc\cos(B_{ij}/(B_{ii}B_{jj})^{\frac{1}{2}})$. Partitions $P = \{i, j\}$ and $P = \{k\}$ are evaluated as in the case $n = 2$, leading to the final formula

$$\begin{aligned}\mathscr{P}(c) = \tfrac{1}{2} &- (4\pi)^{-1} \textstyle\sum_{i<j} \arc\cos(B_{ij}/(B_{ii}B_{jj})^{\frac{1}{2}}) \\ &+ (4\pi)^{-1} \textstyle\sum_{i<j} \arc\cos(B^{ij}/(B^{ii}B^{jj})^{\frac{1}{2}}) \Pr\{F(1, \nu) \leqq c^2\} \\ &+ (4\pi)^{-1} \textstyle\sum_{i<j} \arc\cos(B_{ij}/(B_{ii}B_{jj})^{\frac{1}{2}}) \Pr\{F(2, \nu) \leqq c^2/2\} \\ &+ [\tfrac{1}{2} - (4\pi)^{-1} \textstyle\sum_{i<j} \arc\cos(B^{ij}/(B^{ii}B^{jj})^{\frac{1}{2}})] \Pr\{F(3, \nu) \leqq c^2/3\} .\end{aligned} \tag{3.2}$$

Evaluation of $c^\# = c^\#(\alpha, B, \nu)$ requires in practice evaluation of (3.1) or (3.2) for various values of c to find that for which $\mathscr{P}(c) = 1 - \alpha$. Tabulation of $c^\#$ is impractical, because of the large number of parameters on which it depends. The search for $c^\#$ may, however, be relegated to a digital computer. A program usable on any computer that accepts BASIC FORTRAN is available from the authors.

For $n > 3$, the general problem of expressing the ratios ρ_P and $\rho'_{P'}$ of (2.3) as explicit functions of the entries of B is beyond practical scope. According to [5], the computation of the content of hyperspherical simplices of dimension exceeding 2 involves the evaluation of a sequence of recursively defined integrals. If, however, the ρ_P and $\rho'_{P'}$ are known, as for example in cases (c) and (e) of Corollary 1, the solution to the equation $\mathscr{P}(c^\#) = 1 - \alpha$ is easily programmable.

We next describe the improvement of the sharp upper bounds over $R_+{}^3$ as compared to Scheffé upper bounds in the following two examples:

EXAMPLE 1. On the basis of past observations $\{Y(t): t = -1, \cdots, -T\}$ of a time series of economic gains $Y(t)$, with expectations

$$E\{Y(t)\} = \beta_1 + \beta_2 t + \beta_3 t^2 , \tag{3.3}$$

we seek a lower bound on the expected future gain. Here $\mathbf{x}' = (1, t, t^2)$, $f(\mathbf{x}; \boldsymbol{\beta}) = \mathbf{x}'\boldsymbol{\beta} = E\{Y(t)\}$ and $X^* = \{\mathbf{x}(t): t \geqq 0\}$ is a curve in $R_+{}^3$.

Sharpening, as measured by $c^\#/c^*$, is greatest when σ^2 is unknown and must be estimated from few observations. Here, savings reach 55% for $T = 5$; they decrease to the case for known variance, where savings are still about 25—40% for reasonable α values. Since the corresponding maximal savings are 13—20% for diagonal B and known variance, as shown in [3], this example suggests that

TABLE I
$c^{\#}/c^{*}$

$1-\alpha$ / T	.90	.95	.99	.90	.95	.99
5	.442	.453	.469	.578	.653	.742
∞	.515	.677	.760	.615	.677	.760
	variance unknown			variance known		

there are cases of even more practical interest to which the present work gives even better savings than were obtained in [3]. See Table I.

EXAMPLE 2. Let $Y(l_1, l_2)$ denote the response of an individual to l_i units of medication i, $i = 1, 2$, with expectation

$$E\{Y(l_1, l_2)\} = \beta_1 + \beta_2 l_1 + \beta_3 l_2 \,. \tag{3.4}$$

We seek an upper bound on the dose-response function $f(\mathbf{x}; \boldsymbol{\beta}) = \mathbf{x}'\boldsymbol{\beta} = E\{Y(l_1, l_2)\}$ where $\mathbf{x}' = (1, l_1, l_2)$. Since dosage is nonnegative, $X^* = \{\mathbf{x} : x_1 = 1, x_2, x_3 \geqq 0\} \subset R_+^{\,3}$. Note that in this case, X_1^* is dense in $S_+^{\,2}$ and so, by Corollary 5, the bounds $c^{\#}$ are sharp here too. In Table II we present an abstract of the computations for this example based on the usual analysis of variance applied to $\{Y(l_1, l_2) : l_1, l_2 = 0, 1, \cdots, D\}$, $D = 2, \cdots, 10$ and $D = \infty$. See Table II.

TABLE II
$c^{\#}/c^{*}$

$1-\alpha$ / D	.90	.95	.99	.90	.95	.99
3	.895	.904	.915	.914	.928	.947
∞	.904	.920	.940	.904	.920	.940
	variance unknown			variance known		

REFERENCES

[1] BIRKHOFF, G. and MACLANE, S. (1953). *A Survey of Modern Algebra.* MacMillan, New York.

[2] BOHRER, R. (1967). On sharpening Scheffé bounds. *J. Roy. Statist. Soc. B* **29** 110–114.

[3] BOHRER, R. (1969). On one-sided confidence bounds for response surfaces. *Bull. Internat. Statist. Inst.* **43** 255–257.

[4] BOHRER, R. and FRANCIS, G. (1972). Sharp one-sided confidence bounds for linear regression over an interval. *Biometrika* **59** No. 1.

[5] MÖLLER, H. (1967). *Beiträge zur Integration der Schläflischen Differentialform für Simplexinhalte in nichteuklidischen Räumen höherer Dimension.* Bonner Mathematische Schriften, No. 29.

[6] SAMELSON, H., THRALL, R. M. and WESLER, O. (1958). A partition theorem for Euclidean n-space. *Proc. Amer. Math. Soc.* **9** 805–807.

[7] SCHEFFÉ, H. (1959). *The Analysis of Variance.* Wiley, New York.

[8] VAN DER WAERDEN, B. L. (1969). *Mathematical Statistics.* Springer, New York.

MATH INSTITUTE
UNIVERSITY OF KENT
CANTERBURY, ENGLAND