Increasing the Confidence in Student's t Interval

BU-1099-M

December 1990

Constantinos Goutis
University College London

George Casella[1]
Cornell University

Keywords and phrases: Conditional confidence, Relevant sets.

[1]Research supported by National Science Foundation Grant No. DMS89-0039.

Summary

The usual confidence interval, based on Student's t distribution, has conditional confidence that is larger than the nominal confidence level. Although this fact is known, along with the fact that increased conditional confidence can be used to improve a confidence assertion, the confidence assertion of Student's t interval has never been critically examined. We do so here, and construct a confidence estimator that allows uniformly higher confidence in the interval, and is closer (than $1-\alpha$) to the indicator of coverage.

1. Introduction and Summary

The usual confidence interval for a normal mean, when the population variance is unknown, is based on Student's t distribution. More precisely, let $x_1,\cdots,x_n$ be the realized value of $X_1,\cdots,X_n$, iid random variables from a normal distribution with unknown mean μ and unknown variance σ^2. The set

$$C(\bar{x}, s) = \left\{\mu : \left|\frac{\bar{x}-\mu}{s}\right| \leq k\right\} \tag{1.1}$$

is a $1-\alpha$ confidence interval for μ, where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$, and k is a constant (based on the t distribution with n−1 degrees of freedom) that gives $C(\bar{x}, s)$ coverage probability $1-\alpha$.

The interval (1.1) has constant coverage probability that is equal to the confidence coefficient, $1-\alpha$. In cases where the coverage probability is nonconstant, the confidence coefficient is the infimum of the coverage probabilities, but for (1.1) we have

$$\inf_{\mu,\sigma^2} P_{\mu,\sigma^2}\left[\mu \in C(\bar{X}, S)\right] = 1-\alpha = P_{\mu,\sigma^2}\left[\mu \in C(\bar{X}, S)\right], \quad \forall \mu,\sigma^2 . \tag{1.2}$$

A frequentist must usually be content with reporting the confidence coefficient, but because of the equalities in (1.2), the coverage probability can be reported. This makes the value $1-\alpha$ particularly meaningful as a report of confidence that $C(\bar{x}, s)$ covers μ.

The value $1-\alpha$, however, is independent of the data. If we observe $\bar{X} = \bar{x}$ and $S^2 = s^2$, might this new knowledge alter our assessment of confidence in $C(\bar{x}, s)$? This question originated with Fisher (1956), who was concerned with the behavior of confidence intervals on *recognizable* subsets (subsets of the sample space). If there are recognizable subsets on which confidence can be altered, then a confidence report of $1-\alpha$ may not be appropriate.

Brown (1967), building on the work of Buehler and Fedderson (1963), proved that there exists $\epsilon > 0$ and a constant c such that

$$P_{\mu,\sigma^2}\left[\mu \in C(\bar{X}, S) \,\middle|\, |\bar{X}/S| \leq c\right] \geq 1-\alpha+\epsilon \qquad \forall \mu,\sigma^2 . \tag{1.3}$$

The set $\left\{(\bar{x}, s) : |\bar{x}/s| \leq c\right\}$ is now called a *relevant* set (Robinson, 1979a). The existence of a set satisfying the inequality in (1.3) immediately implies that the confidence report $1-\alpha$ can be improved upon, using a data-dependent confidence report $\gamma(\bar{x}/s)$, with $\gamma(\bar{x}/s) \geq 1-\alpha$.

The applicability of Brown's result is limited by the fact that ϵ is only shown to exist, and thus an improved confidence report cannot be constructed. A main goal of this paper is to find a computable value ϵ_0 (along with a constant c) that satisfies (1.3). Once that is done, an improved confidence statement for the interval (1.1) can be constructed.

In Section 2 we provide some background on conditional confidence, and formalize what is meant by an "improved confidence statement." Section 3 contains the key result of the paper, the lemma giving an explicit value of ϵ that satisfies (1.3). The proof of the lemma is extremely lengthy, and occupies the remainder of Section 3. (A detailed reading of the proof is not necessary to follow the main ideas of the paper.) In Section 4 we show how to construct improved confidence statements using the lemma of Section 3. We also apply a Brewster-Zidek-type construction (Brewster and Zidek, 1974) to exhibit smoother confidence estimators. The size of the gains is also investigated numerically. Section 5 contains a concluding discussion.

2. Conditional Confidence

Suppose X is a random variable with density $f(\cdot|\theta)$ and, after observing $X = \mathbf{x}$, a confidence interval $C(\mathbf{x})$ for (possibly vector valued) θ is constructed, with the property that $P_\theta[\theta \in C(X)] \geq 1-\alpha$. After the interval is obtained, we would like to report a number that represents our confidence that the interval covers the true parameter. The conventional frequentist report is $1-\alpha$, the infimum of the pre-data coverage probability. Although such a report is probably the best possible before the data are obtained, it may be inadequate as a post-data confidence report. This is especially true if $P_\theta[\theta \in C(X)]$ is not constant or if an ancillary statistic exists. In the first case the coverage probability is under-reported, and in the second case the infimum can be quite misleading. Indeed, the coverage probability can be misleading even if it is constant, as is the case here. (These points are also illustrated by Berger and Wolpert, 1989 and Robert and Casella, 1990.)

Many researchers have been concerned with these problems, starting with Fisher (1956, 1959). Work by Basu (1964, 1981), Buehler (1959), Robinson (1976, 1977, 1979a, 1979b) all address these points. (A review of the development of conditional confidence ideas is given by Casella, 1990). A possible solution to these conditional problems was given by Kiefer (1977), who advocated using estimated (data-dependent) confidence statements.

We consider a *confidence procedure* to be a pair $<C(\mathbf{x}), \gamma(\mathbf{x})>$ where $\gamma(\mathbf{x})$ is the reported confidence in the set $C(\mathbf{x})$. We think of $C(\mathbf{x})$ being constructed in a predetermined way and our goal is to determine a reasonable $\gamma(\mathbf{x})$. We treat the choice of a confidence report as an estimation problem and we are concerned whether, for a given $X = \mathbf{x}$, the true parameter is covered by $C(\mathbf{x})$. This suggests that $\gamma(\mathbf{x})$ could be treated as an estimate of the indicator function of coverage, that is,

$$I_{C(\mathbf{x})}(\theta) = \begin{cases} 1 & \text{if} \quad \theta \in C(\mathbf{x}) \\ 0 & \text{if} \quad \theta \notin C(\mathbf{x}) \end{cases}. \tag{2.1}$$

This approach was discussed by Berger (1985a, 1985b, 1988), and taken in Lu and Berger (1989), George and Casella (1990), and Robert and Casella (1990). Our approach is closely related to the theory of conditional inference as formalized by Robinson (1979 a,b).

For a given $C(\mathbf{x})$ the confidence statements will be compared according to the squared error loss

$$\text{(2.2)} \qquad L(\gamma, \theta, \mathbf{x}) = \Big[\gamma(\mathbf{x}) - I_{C(\mathbf{x})}(\theta)\Big]^2 .$$

If we are concerned with the performance of $\gamma(\mathbf{x})$ from a frequentist view, we must evaluate the performance of $\gamma(\mathbf{x})$ by considering the risk function

$$\text{(2.3)} \qquad R(\gamma, \theta) = E\Big\{\Big(\gamma(X) - I_{C(X)}(\theta)\Big)^2\Big\} .$$

A reported confidence $\gamma_1(\mathbf{x})$ will be inadmissible if there is a $\gamma_2(\mathbf{x})$ such that

$$\text{(2.4)} \qquad R(\gamma_1, \theta) \geq R(\gamma_2, \theta)$$

for all θ, with strict inequality for some θ.

If a relevant subset exists for $<C(\mathbf{x}), \gamma(\mathbf{x})>$ (or, in the more general terminology of Robinson (1979a), a *relevant betting procedure*), then $\gamma(\mathbf{x})$ is an inadmissible confidence report for $C(\mathbf{x})$, using (2.4). A *positively biased* relevant subset M satisfies, for some $\epsilon > 0$,

$$\text{(2.5)} \qquad E_\theta\Big\{\Big[I_{C(X)}(\theta) - \gamma(X)\Big] I_M(X)\Big\} \geq \epsilon\, E_\theta I_M(X) .$$

Define $\gamma'(\mathbf{x}) = \gamma(\mathbf{x}) + \epsilon I_M(\mathbf{x})$. Then, using (2.5), it is immediate that $R\big[\gamma'(X), \theta\big] \leq R\big[\gamma(X), \theta\big]$, with strict inequality for some θ as long as M has positive probability. (Note that if $\gamma(\mathbf{x}) = 1-\alpha$, then (2.5) can be written as the inequality $P\big[\theta \in C(X) \mid X \in M\big] \geq 1-\alpha+\epsilon$, implying that $1-\alpha$ reports too little conditional confidence.)

A *negatively-biased* relevant subset can similarly be defined, as satisfying

$$\text{(2.6)} \qquad -E_\theta\Big\{\Big[I_{C(X)}(\theta) - \gamma(X)\Big] I_M(X)\Big\} \geq \epsilon\, E_\theta I_M(X) ,$$

and, here, $\gamma(\mathbf{x})$ is dominated by $\gamma'(\mathbf{x}) = \gamma(\mathbf{x}) - \epsilon I_M(\mathbf{x})$. This can be interpreted as saying $\gamma(\mathbf{x})$ reports too much confidence in $C(\mathbf{x})$. Brown (1967) identified a positively biased relevant subset for the t procedure $<C(\bar{x}, s), 1-\alpha>$, and Robinson (1976) showed that the procedure has no negatively biased relevant subsets.

Estimation of confidence has been studied, though not extensively. Lu and Berger (1989) consider the case of a multivariate normal mean and concentrate on improved confidence statements for sets recentered at positive-part James-Stein estimator. As shown by Hwang and Casella (1982), such sets have uniformly higher coverage probability than the usual confidence region. Lu and Berger (1989) established that, if the dimension is more than 5, the constant confidence is inadmissible for the class of estimators exhibited in Hwang and Casella (1982). They also developed classes of alternative confidence coefficients. George and Casella (1990) developed empirical Bayes confidence estimators dominating the usual one.

Hwang and Brown (1990) established admissibility of $1-\alpha$ as a confidence report for the usual confidence set for a multivariate normal mean of dimension less than or equal to 4. Robert and Casella (1990) showed inadmissibility for dimensions more than 5, and also developed improved confidence statements for the usual multivariate normal confidence set.

For the Student's t procedure $\langle C(\bar{x}, s), 1-\alpha \rangle$, Brown (1967) showed that the set

$$C = \{(\bar{x}, s) : |\bar{x}/s| < c\} \tag{2.7}$$

is a positively biased relevant subset. Thus, a confidence estimator of the form $\gamma(\bar{x}/s) = 1-\alpha + \epsilon I_C(\bar{x}/s)$ will dominate $1-\alpha$ in risk, using (2.3). If values of ϵ and c can be computed, $\gamma(\cdot)$ can be constructed. The next section shows how to compute these values.

3. The Construction Lemma

In this section we state and prove the lemma needed to construct an improved estimator of confidence for the interval (1.1), giving an explicit lower bound on the conditional confidence of $C(\bar{x}, s)$.

Lemma: Suppose $n > 2$. Define the constant c^* by

$$c^* = \max\left\{k, \frac{k}{\sqrt{1+k^2}-1}, c_0\right\} \tag{3.1}$$

with k given in (1.1), and c_0 satisfying

$$c_0 = \begin{cases} k + \sqrt{k^2+1} & \text{if } 3^{n-2}\frac{\alpha}{1-\alpha} < 1 \\ \cot\omega_0 & \text{otherwise} \end{cases}$$

where ω_0 is the solution to

$$P\left(F_{1,n-1} < \frac{n-1}{\cot^2\omega_0}\right) = (1-\alpha)\left[1 + \left(\frac{\sin\omega_0}{\sin(3\omega_0)}\right)^{n-2}\right], \tag{3.2}$$

and $F_{1,n-1}$ is an F random variable with 1 and n−1 degrees of freedom. Then for all $c > c^*$,

$$P_{\mu,\sigma^2}\left(\frac{|\bar{X}-\mu|}{S} \le k \,\middle|\, \frac{|\bar{X}|}{S} \le c\right) \ \ge\ P_{0,\sigma^2}\left(\frac{|\bar{X}|}{S} \le k \,\middle|\, \frac{|\bar{X}|}{S} \le c\right). \tag{3.3}$$

Remark: For $\mu = 0$, the conditional probability on the right-hand side of (3.3) does not depend on σ^2, so the bound is independent of all parameters.

Proof: Define the two sets

$$K = \{(\bar{x}, s): |\bar{x}-\mu|/s \le k\}, \quad C = \{(\bar{x}, s): |\bar{x}|/s \le c\}. \tag{3.4}$$

Following Brown (1967), we will constantly refer to Figure 1 (which is a reproduction of his Figure 1). The area K is contained in $A_1A_0A_2$ and C is contained in B_1OB_2. Since K and C depend only on the ratios $\bar{x}/s$, μ/s, and μ/σ, all the probabilities we consider are only functions of μ/σ. Thus, without loss of generality we can assume $\sigma = 1$, and because of symmetry, we can take $\mu \ge 0$.

Let $\varphi = \cot^{-1}(k)$ and $\omega = \cot^{-1}(c)$. Since $c > k$, the lines A_0A_2 and OB_2 intersect at Q_2, while

OB_1 and A_0A_1 do not intersect. Note that $c > k$ is equivalent to $\varphi > \omega$. Also, we have taken $A_0P \perp OB_2$, and $c > k/\left(\sqrt{1+k^2}-1\right)$ implies $\varphi < \pi/2 - 2\omega$, or $A_0\widehat{C}_1P < A_0\widehat{C}_2P$. (The "$\widehat{\ }$" denotes angle.)

Consider the system of polar coordinates with A_0 as its center. Let $r^2 = (\bar{x}-\mu)^2 + s^2$ and $\theta = \arctan[s/(\bar{x}-\mu)]$. The values r and θ can be considered values of random variables, R and Θ with probability density

$$f(r, \theta) \propto r^{n-1} e^{\frac{-nr^2}{2}} (\sin\theta)^{n-2} \qquad r \geq 0 \qquad 0 \leq \theta \leq \pi .$$

It can be easily seen that R and Θ are independent. Define

$$p_1(r) = P\left\{K \cap C \cap \left\{\theta : \theta \leq \tfrac{\pi}{2}\right\} \middle| R = r\right\}$$

$$p_2(r) = P\left\{C \cap \left\{\theta : \theta \leq \tfrac{\pi}{2}\right\} \middle| R = r\right\}$$

$$p_3(r) = P\left\{K \cap C \cap \left\{\theta : \theta > \tfrac{\pi}{2}\right\} \middle| R = r\right\}$$

$$p_4(r) = P\left\{C \cap \left\{\theta : \theta > \tfrac{\pi}{2}\right\} \middle| R = r\right\}$$

$$\rho(r) = P_\mu\left\{\frac{|\bar{X}-\mu|}{S} < k \,\middle|\, \frac{|\bar{X}|}{S} < c,\ R = r\right\} = \frac{p_1(r) + p_3(r)}{p_2(r) + p_4(r)} \qquad \text{for} \qquad r > A_0P .$$

We will omit the dependence on r for notational convenience, whenever no confusion arises. Obviously

$$P_\mu\left\{\frac{|\bar{X}-\mu|}{S} < k \,\middle|\, \frac{|\bar{X}|}{S} < c\right\} = E_\mu(\rho) .$$

If $\mu = 0$, $\rho(r)$ is independent of r, $1-\alpha < \rho(r) < 1$, and can be written

(3.5) $$\rho(r) = \frac{\int_{\varphi}^{\frac{\pi}{2}} (\sin\theta)^{n-2} d\theta}{\int_{\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2} d\theta} \equiv 1-\alpha+\epsilon_0 ,$$

which defines ϵ_0. Observe that for every μ,

(3.6) $$\lim_{r \to \infty} \rho(r) = 1-\alpha+\epsilon_0 .$$

To show that this limit is a lower bound, we must consider a number of cases. We will consider two cases depending on the range of c and, for each of these cases there will be four subcases, depending on the range of r.

Case I: $c > k + \sqrt{1+k^2}$, or equivalently, $\varphi > 2\omega$ or $A_0Q_2 < A_0O$.

Case Ia: $r > A_0O$

Let $\psi = D_4\hat{A}_0O$ and $\xi = D_2\hat{A}_0x$. Note that D_4 lies on the line OB_1. Let D_0' and D_4' be the reflections of D_0 and D_4 with respect to x axis, respectively. Clearly $D_4'\hat{A}_0O = \psi$. By simple geometry we know that the sum of the radian measure of the arcs $D_4'D_0'$ and D_2D_0 equals twice the measure of $O\hat{C}_2A_0$, hence $\psi + \xi = 2\omega$.

Taking the derivative of $p_2 + p_4$ with respect to ψ, we have

$$\text{(3.7)} \qquad \frac{d(p_2+p_4)}{d\psi} \propto \frac{d}{d\psi}\left[\int_{\psi}^{\frac{\pi}{2}}(\sin\theta)^{n-2}d\theta + \int_{2\omega-\psi}^{\frac{\pi}{2}}(\sin\theta)^{n-2}d\theta\right] = -(\sin\psi)^{n-2} + \left[\sin(2\omega-\psi)\right]^{n-2}$$

which is positive since $\xi = 2\omega-\psi > \psi$ and $n > 2$. Since ψ is increasing in r, $\frac{d(p_2+p_4)}{dr} > 0$. On the other hand, for $r > A_0Q_2 > A_0C_1$, p_1 and p_3 are constant, hence ρ is decreasing in r. Using (3.6) we conclude that $\rho > 1-\alpha+\epsilon_0$.

Case Ib: $A_0O > r > A_0Q_2$

Similar to the previous case, p_1 and p_3 are constant, whereas now both p_2 and p_4 are increasing in r. Therefore ρ decreases to $\rho(A_0O)$, which is greater than $1-\alpha+\epsilon_0$.

Case Ic: $A_0Q_2 > r > A_0C_1$

Let $\psi = D_4\hat{A}_0O$ as above, and observe that ψ is decreasing in r. If $\psi > \omega$ then $\rho > 1-\alpha+\epsilon_0$ and $r = A_0Q_2$ corresponds to $\psi = \varphi-2\omega$. Hence we are interested in ψ such that $\varphi-2\omega < \psi < \omega$. (If $\varphi > 3\omega$ then $\rho > 1-\alpha+\epsilon_0$ trivially.) Since p_4 is decreasing in ψ,

$$\rho = \frac{\int_{2\omega+\psi}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta + \int_{\varphi}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta}{\int_{2\omega+\psi}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta + \int_{\psi}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta} \geq \frac{\int_{2\omega+\psi}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta + \int_{\varphi}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta}{\int_{2\omega+\psi}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta + \int_{\varphi-2\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta}$$

(3.8)

$$\geq \frac{\int_{3\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta + \int_{\varphi}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta}{\int_{3\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta + \int_{\varphi-2\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta}.$$

Comparing the lower bound in (3.8) to (3.5), to show $\rho \geq 1-\alpha + \epsilon_0$, it suffices to show

$$\int_{3\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta \int_{\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta + \int_{\varphi}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta \int_{\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta$$

(3.9)

$$-\int_{\varphi}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta \int_{3\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta - \int_{\varphi}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta \int_{\varphi-2\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta \geq 0,$$

for $2\omega < \varphi < 3\omega$ and $0 < \omega < \pi/8$. This last restriction is a consequence of the fact that $2\omega < \varphi < \pi/2-2\omega$.

Differentiating the LHS of (3.9) with respect to φ we obtain

$$-(\sin\varphi)^{n-2} \int_{\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta + (\sin\varphi)^{n-2} \int_{3\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta$$

$$+ (\sin\varphi)^{n-2} \int_{\varphi-2\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta + [\sin(\varphi-2\omega)]^{n-2} \int_{\varphi}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta$$

which is positive since

$$\int_{\varphi-2\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2} d\theta > \int_{\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2} d\theta .$$

Hence (3.9) is true if and only if, for $0 < \omega < \pi/8$,

$$\frac{\int_{3\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2} d\theta + \int_{2\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2} d\theta}{\int_{3\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2} d\theta + \int_{0}^{\frac{\pi}{2}} (\sin\theta)^{n-2} d\theta} \geq \frac{\int_{2\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2} d\theta}{\int_{\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2} d\theta}, \tag{3.10}$$

or equivalently

$$\frac{\int_{3\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2} d\theta}{\int_{3\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2} d\theta + \int_{0}^{\omega} (\sin\theta)^{n-2} d\theta} \geq \frac{\int_{2\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2} d\theta}{\int_{\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2} d\theta}. \tag{3.11}$$

Inverting numerators and denominators in (3.11), after some algebra, we obtain the equivalent expression

$$1 + \frac{\int_{2\omega}^{3\omega} (\sin\theta)^{n-2} d\theta}{\int_{3\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2} d\theta} \leq \frac{\int_{\omega}^{2\omega} (\sin\theta)^{n-2} d\theta}{\int_{0}^{\omega} (\sin\theta)^{n-2} d\theta}. \tag{3.12}$$

The LHS of (3.12) is less than 2 since the interval of integration of the denominator is wider and lies on the right of the interval of integration of the numerator, since $\omega < \pi/8$ and $(\sin\theta)^{n-2}$ is increasing. Hence a sufficient condition for (3.12) to hold is

$$(3.13) \qquad \frac{\int_0^{2\omega} (\sin\theta)^{n-2} d\theta}{\int_0^{\omega} (\sin\theta)^{n-2} d\theta} \geq 3$$

or, for $\omega > 0$,

$$(3.14) \qquad \int_0^{2\omega} (\sin\theta)^{n-2} d\theta - 3 \int_0^{\omega} (\sin\theta)^{n-2} d\theta \geq 0 .$$

Since (3.14) [and (3.13)] are true for $\omega = 0$, it suffices to show that the LHS of (3.14) is nondecreasing in ω. Differentiating with respect to ω, we need to establish

$$2(\sin 2\omega)^{n-2} - 3(\sin\omega)^{n-2} \geq 0$$

or, equivalently

$$(3.15) \qquad (2\cos\omega)^{n-2} \geq \tfrac{3}{2} .$$

Inequality (3.15) is true for $0 < \omega < \pi/8$, $n > 2$ since $\cos\omega \geq \cos \pi/8 = 0.9239$, thus establishing (3.10).

Case Id: $A_0C_1 > r > A_0P$

Clearly $\rho = 1 > 1-\alpha+\epsilon_0$ for this range of r.

Case II: $c < k + \sqrt{k^2+1}$ or, equivalently, $\varphi < 2\omega$ or $A_0O < A_0Q_2$. We partition the space of possible values of r in a similar way.

Case IIa: $r > A_0Q_2$ Here ρ decreases to $1-\alpha+\epsilon_0$, as before.

Case IIb: $A_0Q_2 > r > A_0O$ The functions p_1 and p_2 are equal and increasing in r whereas p_4 is decreasing and p_3 is constant. Hence

$$(3.16) \qquad \frac{dp_1}{dr} = \frac{dp_2}{dr} > \frac{d(p_2+p_4)}{dr} .$$

The derivative of ρ has the same sign as

$$(3.17) \qquad \frac{dp_2}{dr}(p_2+p_4) - \frac{d(p_2+p_4)}{dr}(p_1+p_3)$$

which is positive since $p_2 + p_4 > p_1 + p_3$ and (3.16). Hence ρ is increasing in r and it is greater than $1-\alpha+\epsilon_0$ if and only if $\rho(A_0O) > 1-\alpha+\epsilon_0$.

Case IIc: $A_0O > r > A_0C_1$ Define $\psi = D_4\hat{A}_0O$. If $\psi \geq \omega$ we have $\rho > 1-\alpha+\epsilon_0$, so $\rho > 1-\alpha+\epsilon_0$ if and only if

$$(3.18) \quad \int_{2\omega+\psi}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta + \int_{\varphi}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta - (1-\alpha+\epsilon_0)\left[\int_{2\omega+\psi}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta + \int_{\psi}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta\right] \geq 0$$

for all ψ such that $0 < \psi < \omega$.

Since (3.18) is true for $\psi = \omega$, it is sufficient (but not necessary) that the LHS be nonincreasing in ψ. Differentiating with respect to ψ, the derivative is nonpositive if and only if

$$(3.19) \qquad \frac{1}{1-\alpha+\epsilon_0} - 1 \geq \left(\frac{\sin\psi}{\sin(2\omega+\psi)}\right)^{n-2}.$$

Since the RHS of (3.19) is increasing in ψ, using the definition of $1-\alpha+\epsilon_0$ (from (3.5)), an equivalent condition is

$$(3.20) \qquad \frac{\int_{\omega}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta}{\int_{\varphi}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta} - 1 \geq \left(\frac{\sin\omega}{\sin 3\omega}\right)^{n-2}$$

for $0 < \omega < \varphi$. Now observe that the LHS, as a function of ω, decreases to zero whereas the RHS is increasing. Hence if

$$(3.21) \qquad \frac{\int_{0}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta}{\int_{\varphi}^{\frac{\pi}{2}} (\sin\theta)^{n-2}d\theta} -1 > \left(\frac{1}{3}\right)^{n-2},$$

inequality (3.20) holds for $\omega < \omega_0$ where ω_0 is the value of ω that makes (3.20) an equality.

Note that integrals of the form $\int (\sin\theta)^{n-2}\, d\theta$ can be expressed as beta integrals, using the transformation $u = \sin^2\theta$. Then using the relationship between the beta and F distribution, equality in (3.20) is exactly condition (3.2). Moreover, condition (3.21) reduces to the requirement $3^{n-2}\, \alpha/(1-\alpha) > 1$. Therefore, for $c > c_0 = \cot \omega_0$, $\rho > 1-\alpha+\epsilon_0$.

Case IId: $A_0C_1 > r > A_0P$. For the range of r it is clear that $\rho = 1$, hence $\rho > 1-\alpha+\epsilon_0$.

Since $\rho(r) \geq 1-\alpha+\epsilon_0$ for every μ, with equality for $\mu = 0$, the lemma is proved by taking expectations over r. □

The assumptions of the lemma are clearly sufficient but not necessary. Numerical evidence shows that the result holds for c smaller than the bounds given in the statement of the lemma. However if $c < k$ the RHS of (3.3) equals 1 and, as Brown (1967) points out, if $c < 1/k$ the conditional probability tends to 0 as μ tends to infinity.

4. Increased Confidence for the t Interval

The lemma of Section 3 gives an easily computable bound on the conditional coverage probability. We now use this bound to construct a post-data confidence estimator $\gamma(\bar{x}/s)$ which improves on $1-\alpha$.

Theorem 4.1 Let $c > c^*$ and k be fixed constants, where c^* and k satisfy the assumptions of the lemma of Section 3. Define $\gamma_c(\bar{x}/s)$ as follows:

$$(4.1) \qquad \gamma_c(\bar{x}/s) = \begin{cases} \dfrac{1-\alpha}{P\left(|t_{n-1}| < c\sqrt{n-1}\right)} & \text{if} \quad \dfrac{|\bar{x}|}{s} < c \\ 1-\alpha & \text{if} \quad \dfrac{|\bar{x}|}{s} \geq c \end{cases}$$

where t_{n-1} denotes a t random variable with n−1 degrees of freedom. Then

$$(4.2) \qquad E_{\mu,\sigma^2}\left[\left(\gamma_c(\bar{X}/S) - I_{C(\bar{X},S)}(\mu)\right)^2\right] < E_{\mu,\sigma^2}\left[\left(1-\alpha - I_{C(\bar{X},S)}(\mu)\right)^2\right]$$

for all μ and σ^2, where $C(\bar{x}, s)$ is the $1-\alpha$ Student's t interval of (1.1).

Proof Since $\gamma_c(\bar{x}/s) = 1-\alpha$ for $|\bar{x}|/s \geq c$, it suffices to show

$$(4.3) \qquad E_{\mu,\sigma^2}\left[\left(\gamma_c(\bar{X}/S) - I_{C(\bar{X},S)}(\mu)\right)^2 \,\middle|\, \frac{|\bar{X}|}{S} < c\right] < E_{\mu,\sigma^2}\left[\left(1-\alpha - I_{C(\bar{X},S)}(\mu)\right)^2 \,\middle|\, \frac{|\bar{X}|}{S} < c\right]$$

for all μ. The last inequality can be seen to be true by observing that, for $|\bar{X}|/S < c$,

$$(4.4) \qquad \gamma_c(\bar{X}/S) = P_{0,\sigma^2}\left(\frac{|\bar{X}|}{S} < k \,\middle|\, \frac{|\bar{X}|}{S} < c\right) > 1-\alpha$$

and using inequality (3.3) (as in Robinson, 1979a). □

Since a confidence report based on a partition of possible values of $|\bar{x}|/s$ improves upon $1-\alpha$, it seems plausible that, by taking a finer partition, we could construct a report that is even better. The technique of further partitioning has been introduced by Brewster and Zidek (1974) and has been applied in similar problems in Goutis and Casella (1990) and Shorrock (1990). We have the following theorem:

Theorem 4.2 Let $\underset{\sim}{c}_2 = (c_1, c_2)$, $c_1 > c_2 > c^*$ and k be constants, where c^* and k satisfy the assumptions of the lemma of Section 3. Define $\gamma_{\underset{\sim}{c}_2}(\bar{x}/s)$ as follows:

$$\gamma_{\underset{\sim}{c}_2}(\bar{x}/s) = \begin{cases} \dfrac{1-\alpha}{P\left(|t_{n-1}| < c_2\sqrt{n-1}\right)} & \text{if} \quad \dfrac{|\bar{x}|}{s} \leq c_2 \\ \dfrac{1-\alpha}{P\left(|t_{n-1}| < c_1\sqrt{n-1}\right)} & \text{if} \quad c_2 < \dfrac{|\bar{x}|}{s} < c_1 \\ 1-\alpha & \text{if} \quad c_1 \leq \dfrac{|\bar{x}|}{s}\,. \end{cases}$$

Then

$$E_{\mu,\sigma^2}\left[\gamma_{\underset{\sim}{c}_2}\left((\bar{X}/S) - I_{C(\bar{X},S)}(\mu)\right)^2\right] < E_{\mu,\sigma^2}\left[\gamma_{c_1}\left((\bar{X}/S) - I_{C(\bar{X},S)}(\mu)\right)^2\right]$$

for all μ and σ^2, where γ_{c_1} is given in (4.1).

Proof The result follows immediately after observing that $P_{0,\sigma^2}\left\{\frac{|\bar{X}|}{S} < k \,\middle|\, \frac{|\bar{X}|}{S} < c\right\}$ is decreasing in c. □

An immediate consequence of the above theorem is that $\gamma_{\underset{\sim}{c}_2}(\bar{x}/s)$ improves upon $1-\alpha$. We can easily generalize and take more cutoff points. We create an array $\underset{\sim}{c}_m = (c_{m,1}, c_{m,2} \cdots c_{m,m})$ such that $c^* < c_{m,1} < c_{m,2} \cdots < c_{m,m} < +\infty$, $\lim_{m\to\infty} c_{m,m} = +\infty$ and $\lim_{m\to\infty} \max_i (c_{m,i} - c_{m,i-1}) = 0$. As $m \to +\infty$ the reported confidence will tend to $\gamma(\bar{x}/s)$ defined by

$$\gamma(\bar{x}/s) = \begin{cases} \dfrac{1-\alpha}{P\left(|t_{n-1}| < (|\bar{x}|/s)\sqrt{n-1}\right)} & \text{if} \quad |\bar{x}|/s > c^* \\ \dfrac{1-\alpha}{P\left(|t_{n-1}| < c^*\sqrt{n-1}\right)} & \text{if} \quad |\bar{x}|/s < c^* \end{cases}. \tag{4.5}$$

Note that $P\left(|t_{n-1}| < (|\bar{x}|/s)\sqrt{n-1}\right) = 1-p(\mathbf{x})$ where $p(\mathbf{x})$ is the p-value associated with the hypothesis $\mu = 0$. By the Dominated Convergence Theorem, $\gamma(\bar{x}/s)$ dominates $1-\alpha$ in terms of risk.

For an observed value of $\bar{x}/s$, the confidence report attached to the t interval will be $\gamma(\bar{x}/s)$ of (4.5). To give some idea of the shape of this function, Figure 2 shows values of $\gamma(\bar{x}/s)$ for n=5 and n=9. It can be seen that a confidence report of almost 94% confidence is possible when using a nominal 90% interval. Figure 3 shows the difference in the risks plotted against μ/σ. Although the improvement is not large, $\gamma(\bar{x}/s)$ does dominate 1-α. Note that the size of the risk improvement is really only of secondary concern, as the experimenter may be able to assert a sizeable increase in confidence. Indeed, the experimenter really does not care about the magnitude of the risk improvement as long as the confidence is maximally increased. As the degrees of freedom get large, for the usual confidence coefficients, the constant k is rather small. Since $c^* \geq \frac{k}{\sqrt{1+k^2}-1} > \frac{1}{k}$, the bound c^* is large, so $\gamma(\bar{x}/s)$ is quite close to $1-\alpha$ and we cannot expect a large improvement.

5. Discussion

For the t interval, the existence of a relevant subset implies that the unconditional confidence statement may not be the most appropriate assertion of confidence. Since the conditional confidence level can be bounded above the nominal $1-\alpha$ for all parameter values, an increased confidence assertion is appropriate.

Although it is known, in theory, how to use a relevant set to improve confidence, the implementation of the theory may be difficult. This was the case for Student's t, as the verification of the lower bound on the conditional probabilities was quite involved. The end result, however, is the construction of a confidence estimator that allows a confidence assertion uniformly higher than $1-\alpha$, and is closer to the true indicator of coverage.

Many of the arguments used here took advantage of properties of the normal and t distribution, so the extent of generalizability is not clear. It is probably the case, however, that our results can be extended to the F distribution in the analysis of variance, allowing for increased confidence in Scheffé's simultaneous intervals. (Olshen, 1973, was able to extend some of Brown's 1967 arguments to the analysis of variance. Some of those arguments may prove useful.)

For the case of known variance, the usual interval does not allow relevant sets so the construction used in this paper will not result in improved confidence. (Actually, Robinson, 1979b established a stronger result, the absence of *semi*relevant sets.) Also, recall that Hwang and Brown (1990) have shown that $1-\alpha$ is an admissible confidence report in four or fewer dimensions. For five or more dimension, even though there still are no relevant sets, the confidence in the usual set can be improved (Robert and Casella, 1990). Thus, although the existence of a relevant set results in an improved confidence statement, it is a sufficient but not necessary condition.

Perhaps the key idea is Fisher's, that an interval's confidence should be evaluated on recognizable subsets. Then, if the confidence, conditional on that subset, is different from the nominal value, that information should be used to improve the confidence assertion. Fisher concentrated on sets defined by ancillary statistics, while here the conditioning set is the acceptance region of a hypothesis test, and is not based on an ancillary statistic. The moral may be to examine recognizable sets of interest, perhaps dictated by the problem at hand, and see if confidence assertions are affected by conditioning. If so, the nominal confidence coefficient, $1-\alpha$, may not be an adequate confidence report.

References

Basu, D. (1964). Recovery of Ancillary Information. Contribution to *Statistics*, Pergammon Press, Oxford, 7-20. Republished in *Sankhyā* 26 (1964), 3-16. Also in *Statistical Information and Likelihood. A Collection of Critical Essays by Dr. D. Basu*, 1988, J.K. Ghosh (ed.). Lecture Notes in Statistics. New York: Springer-Verlag.

Basu, D. (1981). On Ancillary Statistics, Pivotal Quantities, and Confidence Statements. In *Topics in Applied Statistics*, Y.P. Chaubey and T.D. Dwivedi (eds.). Concordia University, Montreal, 1-29. Also in *Statistical Information and Likelihood. A Collection of Critical Essays by Dr. D. Basu*, 1988, J.K. Ghosh (ed.). Lecture Notes in Statistics. New York: Springer-Verlag.

Berger, J. O. (1985a). *Statistical Decision Theory and Bayesian Analysis, Second Edition.* New York: Springer-Verlag.

Berger, J. O. (1985b). The frequentist viewpoint and conditioning. In *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer.* (L. LeCam and R. Olshen, eds.) Monterey: Wadsworth.

Berger, J. O. (1988). An Alternative: The Estimated Confidence Approach, in *Statistical Decision Theory and Related Topics IV*, Volume I, S.S. Gupta and J.O. Berger (eds.)., 85-90.

Berger, J. O. and Wolpert, R. (1989). *The Likelihood Principle, Second Edition.* Institute of Mathematical Statistics Monograph. Haywood, California.

Brewster, J. F. and Zidek, J. V. (1974). Improving on equivariant estimators. *Ann. Statist.* **2**, 21-38.

Brown, L. D. (1967). The Conditional Level of Student's t-test. *Ann. Math. Statist.* 38, 1068-1071.

Buehler, R. J. (1959). Some Validity Criteria for Statistical Inference. *Ann. Math. Statist.* 30, 845-863.

Buehler, R. J. and Fedderson, A. P. (1963). Note on a Conditional Property of Student's t. *Ann. Math. Statist.* 34, 1098-1100.

Casella, G. (1990). Conditional Inference from Confidence Sets. Technical Report BU-1001-M, Biometrics Unit, Cornell University. To appear in the volume to honor D. Basu.

Fisher, R. A. (1956). On a Test of Significance in Pearson's Biometrika Tables (No. 11). *J. Roy. Statist. Soc. Ser. B* 18, 56-60.

Fisher, R. A. (1959). *Statistical Methods and Scientific Inference*, second edition. New York: Hafner.

George, E. I. and Casella, G. (1990). An empirical Bayes confidence report. Technical Report BU-1062-M, Biometrics Unit, Cornell University.

Goutis, C. and Casella, G. (1990). Improved Invariant Confidence Intervals for a Normal Variance. Biometrics Unit Technical Report BU-1038-M, Cornell University. To appear in the *Annals of Statistics.*

Hwang, J.T. and Brown, L. D. (1990). Estimated Confidence Approach and The Validity Admissibility Criterion. Technical Report, Statistics Center, Cornell University. To appear in the *Annals of Statistics.*

Hwang, J.T. and Casella, G. (1982). Minimax confidence sets for the mean of a multivariate normal distribution. *Ann. Statist.* **10**, 868-881.

Kiefer, J.C. (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. Inf.* **23**, 1-19.

Olshen, R. A. (1973). The Conditional Level of the F-Test. *J. Amer. Statist. Assoc.* 68, 692-698.

Robert, C. and Casella, G. (1990). Improved Confidence Statements for the Usual Multivariate Normal Confidence Set. Technical Report BU-1041-M, Biometrics Unit, Cornell University.

Robinson, G. K. (1976). Properties of Student's t and the Behrens-Fisher solution to the two means problem. *Ann. Statist.* 4, 963-971.

Robinson, G. K. (1977). Conservative statistical inference. *J. Royal Statist. Soc. Ser. B* **39**, 381-386.

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.

Shorrock, G. (1990). Improved confidence intervals for a normal variance. *Ann. Statist.* 18, 972-980.

Figure 1. Regions used for calculation of conditional coverage probabilities.

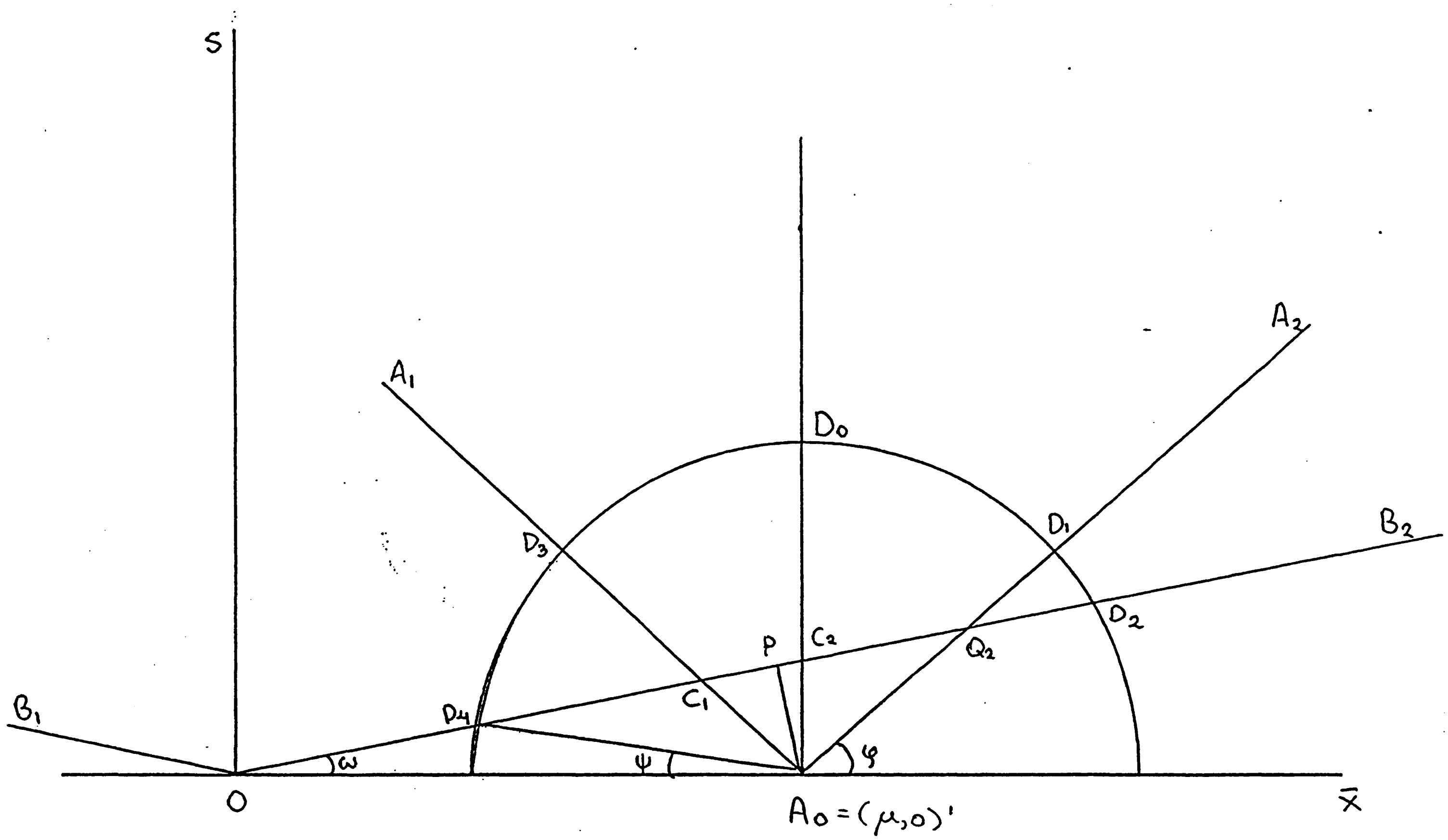

Figure 2. Values of $\gamma(\bar{x}/s)$ of (4.5) for n=5 (solid line), n=9 (dotted line) and $1-\alpha=.90$.

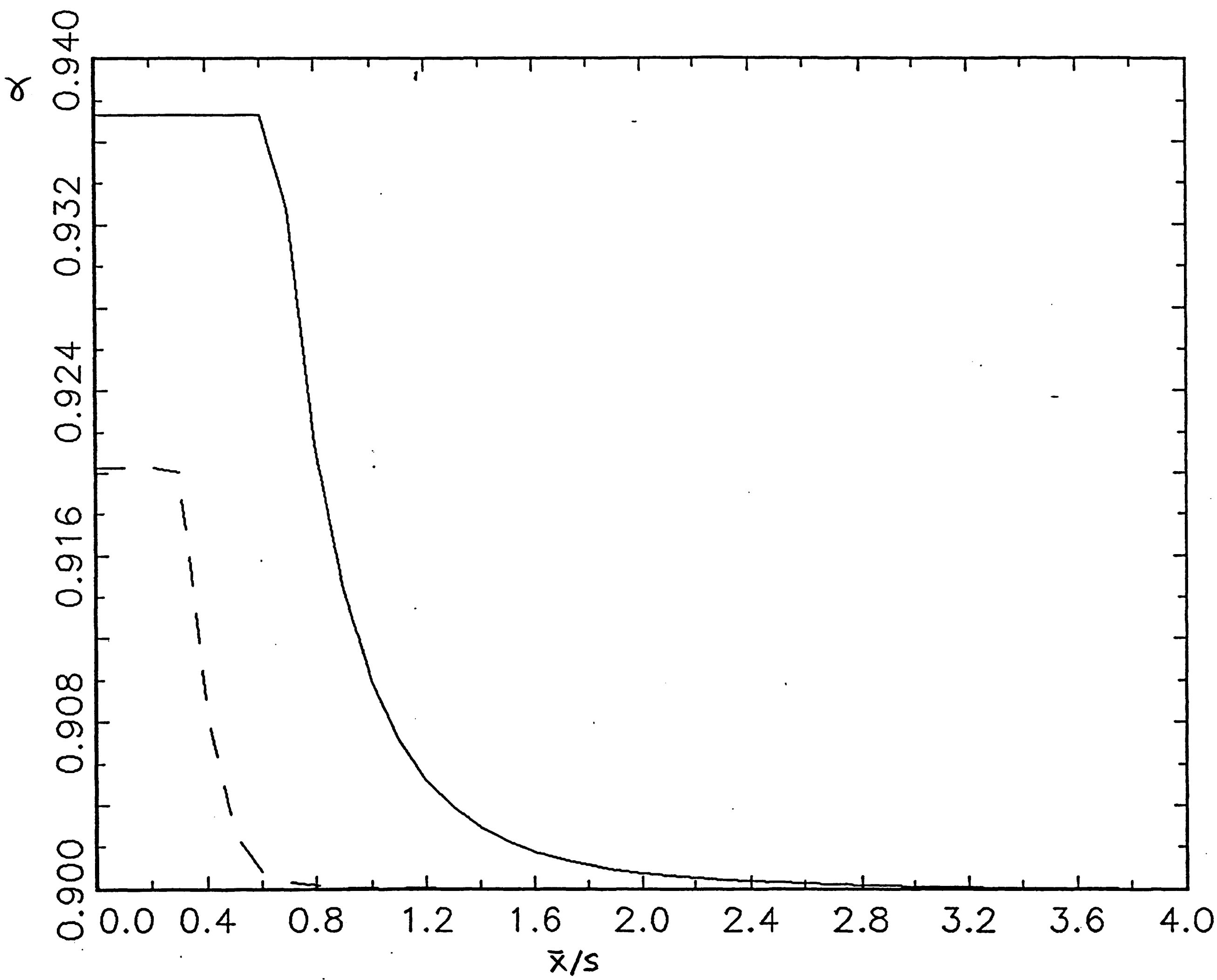

Figure 3. Difference in risks $E_{\mu,\sigma^2}\left[\left\{\gamma_c(\bar{X}/S) - I_{C(\bar{X},S)}(\mu)\right\}^2\right] - E_{\mu,\sigma^2}\left[\left\{1-\alpha - I_{C(\bar{X},S)}(\mu)\right\}^2\right]$ for n=5 (solid line), n=9 (dotted line) and 1-α=.90.