The Annals of Statistics
1987, Vol. 15, No. 1, 462–468

A CHARACTERIZATION OF THE FIELLER SOLUTION

BY MARTIN A. KOSCHAT

Bell Communications Research

We consider the problem of finding α-level confidence intervals for the ratio of two normally estimated means. We show that there is no procedure that with probability 1 gives bounded α-level confidence intervals for the ratio, and we show that within a large class of sensible procedures the Fieller solution is the only one with exact coverage probability.

1. Introduction. Setting limits on the ratio of two normally estimated means is a frequent problem of statistical practice. In its abstract form the problem presents itself as follows. X_1 and X_2 are observed, normally distributed random variables with distribution $N((\mu_1, \mu_2)^t, \sigma^2 I)$. Here I is a known matrix, which, without loss of generality, may be chosen to be the 2×2 identity matrix; μ_1, μ_2 and σ are unknown parameters. Also given is $\hat{\sigma}$, an estimator of σ, which is distributed independently of X_1, X_2. The variable $\nu\hat{\sigma}^2/\sigma^2$ has a χ^2-distribution with ν degrees of freedom. The problem of interest is to bracket the ratio μ_2/μ_1. The frequentist solution most commonly found in textbooks is referred to as the Fieller solution. It was given by Fieller (1940) in a paper on the standardization of insulin, and is derived as a fiducial solution. For a historical account and a discussion of alternative solutions of this problem see Wallace (1980). The Fieller solution has some well-documented features. The confidence region for the ratio may consist of a finite interval, two disjoint semi-infinite intervals or the whole real line. The aim of this note is twofold. First, we will show that for any positive α there is no procedure that gives bounded α-level confidence intervals with probability 1. Second, we will show that within a large class of solutions the Fieller solution is the only one that gives exact coverage probability for all parameters.

2. Bounded confidence intervals. It was pointed out by James, Wilkinson and Venables (1974) that the problem can be conveniently addressed in polar coordinates. Let the random variables R and Θ and the parameters r_0 and θ_0 be defined by the relations

$$X_1 = R\cos\Theta, \qquad X_2 = R\sin\Theta,$$

$$\mu_1 = r_0\cos\theta_0, \qquad \mu_2 = r_0\sin\theta_0.$$

R and r_0 are allowed to vary over the whole real line, while Θ and θ_0 are unoriented angles. Since $\tan\theta_0 = \mu_2/\mu_1$, questions about θ_0 correspond to questions about the ratio μ_2/μ_1 and vice versa.

Received October 1985; revised June 1986.
AMS 1980 *subject classifications*. Primary 62F25; secondary 62J05.
Key words and phrases. Fieller solution, uniqueness of the Fieller solution, inverse regression, regression.

Whenever confidence intervals for θ_0 contain the angle $\pi/2$, the corresponding confidence intervals for μ_2/μ_1 are unbounded. In this section let an unoriented angle be represented by the angle in $(-\pi/2, \pi/2]$. A confidence interval for the ratio is bounded if and only if the corresponding confidence region for θ_0, based on $R, \Theta, \hat{\sigma}$, is an interval (C_L, C_U) of the form $-\pi/2 < C_L \le C_U < \pi/2$. We now show that bounds with this property for all R, Θ, $\hat{\sigma}$ and some given positive coverage probability do not exist.

PROPOSITION 1. *For the Fieller problem there are no bounds $C_L(r, \theta, s)$ and $C_U(r, \theta, s)$ such that*

(1) $$-\pi/2 < C_L(r, \theta, s) \le C_U(r, \theta, s) < \pi/2, \quad \textit{for all } r, \theta, s$$

and

(2) $$\operatorname{pr}\big(C_L(R, \Theta, \hat{\sigma}^2) \le \theta_0 \le C_U(R, \Theta, \hat{\sigma}^2)\big) \ge \alpha,$$
for all r_0, θ_0, σ and some positive α.

PROOF. If the unoriented angle is represented by the corresponding angle in $(-\pi/2, \pi/2]$, the joint density for R and Θ may be written as

$$f_{(R,\Theta)}(r, \theta; r_0, \theta_0) = k\sigma^{-2}|r|\exp\big(-\big(r^2 + r_0^2 - 2rr_0\cos(\theta - \theta_0)\big)/2\sigma^2\big),$$

where k is a normalizing constant. Note that

$$f_{(R,\Theta)}(r, \theta; r_0, \theta_0) \le k\sigma^{-2}|r|\exp\big(-(|r| - |r_0|)^2/2\sigma^2\big).$$

If we denote by c_ν the density of a χ^2-distribution with ν degrees of freedom, then the joint density for $R, \Theta, \hat{\sigma}^2$ is found as

$$f_{(R,\Theta,\hat{\sigma}^2)}(r, \theta, s; r_0, \theta_0, \sigma) = \nu\sigma^{-2}c_\nu(\nu s/\sigma^2) f_{(R,\Theta)}(r, \theta; r_0, \theta_0).$$

For a given pair of bounds C_L, C_U satisfying (1) define the family of sets

$$M_{\theta_0} = \{(r, \theta, s): C_L(r, \theta, s) \le \theta_0\}.$$

It is easily verified that

(3)
(a) $M_{\theta_0} \supseteq \{(r, \theta, s): C_L(r, \theta, s) \le \theta_0 \le C_U(r, \theta, s)\}$;
(b) if $\theta_0' \le \theta_0''$, then $M_{\theta_0'} \subseteq M_{\theta_0''}$;
(c) $\lim_{\theta_0 \to -\pi/2} M_{\theta_0} = \bigcap_{\theta_0} M_{\theta_0} = \varnothing$.

Consequently,

$$\begin{aligned} &\operatorname{pr}\big(C_L(R, \Theta, \hat{\sigma}^2) \le \theta_0 \le C_U(R, \Theta, \hat{\sigma}^2)\big) \\ &\quad \le \iiint_{M_{\theta_0}} f_{(R,\Theta,\hat{\sigma}^2)}(r, \theta, s; r_0, \theta_0, \sigma)\, dr\, d\theta\, ds \\ &\quad \le k' \iiint_{M_{\theta_0}} c_\nu(\nu s/\sigma^2)|r|\exp\big(-(|r| - |r_0|)^2/2\sigma^2\big)\, dr\, d\theta\, ds. \end{aligned}$$

In view of (3) the last integral converges to 0 for any fixed r_0 and σ as θ_0 converges to $-\pi/2$. Therefore

$$\lim_{\theta_0 \to -\pi/2} \operatorname{pr}\big(C_L(R,\Theta,\hat{\sigma}^2) \le \theta_0 \le C_U(R,\Theta,\hat{\sigma}^2)\big) = 0,$$

and the bounds C_L, C_U cannot satisfy (2). □

3. A uniqueness property of the Fieller solution. To actually find bounds for θ_0 which have a certain prescribed coverage probability, consider the coordinate transformation

$$\begin{pmatrix} X_1' \\ X_2' \end{pmatrix} = \begin{pmatrix} \cos\theta_0 & \sin\theta_0 \\ -\sin\theta_0 & \cos\theta_0 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}.$$

The variables X_1', X_2' are normally distributed with distribution $N((r_0, 0)^t, \sigma^2 I)$, and the corresponding polar representation Θ' and R' is related to Θ and R by $R' = R$ and $\Theta' = \Theta - \theta_0$. Denote by $\tilde{R} = R/\hat{\sigma}$ and $\tilde{R}' = R'/\hat{\sigma}$. Then

$$\tilde{R}^2 \sin^2(\Theta - \theta_0) = (\tilde{R}' \sin\Theta')^2 = (X_2'/\hat{\sigma})^2.$$

Therefore $\tilde{R}^2 \sin^2(\Theta - \theta_0)$ has an F distribution with 1 and ν degrees of freedom. To find an α percent confidence interval, choose for given $\tilde{R}$ and Θ all the θ satisfying $\tilde{R}^2 \sin^2(\Theta - \theta) \le F^{\alpha}_{1,\nu}$, where $F^{\alpha}_{1,\nu}$ is the upper percentage point of an F distribution with 1 and ν degrees of freedom. The resulting bounds comprise the *Fieller solution*. If $\tilde{R}^2 \le F^{\alpha}_{1,\nu}$, no angle is ruled out; this is known as the *Fieller paradox*. If σ is known, $\tilde{R}$ is replaced by R and the F distribution is replaced by a χ^2-distribution with 1 degree of freedom.

The distribution of Θ' is symmetric around 0 and depends only on $\tilde{r}$, where $\tilde{r} = r_0/\sigma$. The variance of Θ' decreases as $|\tilde{r}|$ increases. $(\Theta, |\tilde{R}|)$ is the maximum likelihood estimator of $(\theta_0, |\tilde{r}|)$. It is therefore intuitively appealing to look for bounds on θ_0 that are centered at Θ and whose width is a decreasing function of $|\tilde{R}|$. In other words, it is reasonable to look for bounds for the unoriented angle that can be written as $\Theta \pm w(|\tilde{R}|)$, where w is a decreasing function. The Fieller bounds have precisely this form. It turns out that the Fieller bounds are the only ones of this form that have exact coverage probability for all parameters. This fact is the content of the next proposition and its corollary.

PROPOSITION 2. *If for the Fieller problem w and v are two decreasing left continuous functions with $w(0) \le \pi/2$, $v(0) \le \pi/2$, and such that*

$$\begin{aligned} &\operatorname{pr}\big(\Theta - w(|\tilde{R}|) \le \theta_0 \le \Theta + w(|\tilde{R}|)\big) \\ &\quad = \operatorname{pr}\big(\Theta - v(|\tilde{R}|) \le \theta_0 \le \Theta + v(|\tilde{R}|)\big), \quad \textit{for all } \theta_0, r_0, \sigma, \end{aligned} \tag{4}$$

then $w = v$.

PROOF. We will use the notation already introduced. Again, let the unoriented angle be represented by the corresponding angle in $(-\pi/2, \pi/2]$. Our interest is in the joint distribution of Θ' and $\tilde{R}$. For our purpose it is sufficient to note that $(\tilde{R}, \Theta')$ is the polar representation of the variable $(\tilde{X}_1', \tilde{X}_2')$, where $\tilde{X}_1' = X_1'/\hat{\sigma}$, $\tilde{X}_2' = X_2'/\hat{\sigma}$ and X_1' and X_2' are defined as before. The joint distribution of $\tilde{X}_1'$ and $\tilde{X}_2'$ is a noncentral t. Denoting r_0/σ by $\tilde{r}$, its density may be written as

$$f_{(\tilde{X}_1', \tilde{X}_2')}(x_1, x_2) = k\int_{R_+} s^{\nu/2}\exp\Big(-\big((x_1\sqrt{s} - \tilde{r})^2 + x_2^2 s + \nu s\big)/2\Big)\, ds.$$

If v and w are two decreasing functions satisfying (4), then

$$\mathrm{pr}\big(\theta - w(|\tilde{R}|) \le \theta_0 \le \theta + w(|\tilde{R}|)\big) - \mathrm{pr}\big(\theta - v(|\tilde{R}|) \le \theta_0 \le \theta + v(|\tilde{R}|)\big)$$

$$= \int_{-\pi/2}^{\pi/2}\int_{R_+} f_{(\tilde{R}, \Theta')}(r, \theta) D(r, \theta)\, dr\, d\theta,$$

where

$$\begin{aligned} D(r,\theta) &= -1, \quad \text{for } w(|r|) \le |\theta| < v(|r|),\\ &= 1, \quad \text{for } v(|r|) \le |\theta| < w(|r|),\\ &= 0, \quad \text{else.}\end{aligned} \tag{5}$$

Therefore (4) can be rewritten in Cartesian form as

$$\int_R\int_R f_{(\tilde{X}_1', \tilde{X}_2')}(x_1, x_2) D\big(r(x_1, x_2), \theta(x_1, x_2)\big)\, dx_1\, dx_2 = 0, \quad \text{for all } \tilde{r},$$

or after substituting for $f_{(\tilde{X}_1', \tilde{X}_2')}$,

$$\int_R\int_R\int_{R_+} \exp\Big(-\big((x_1\sqrt{s} - \tilde{r}) + x_2^2 s + \nu s\big)/2\Big) s^{\nu/2}$$
$$\times D\big(r(x_1, x_2), \theta(x_1, x_2)\big)\, ds\, dx_1\, dx_2 = 0, \qquad \text{for all } \tilde{r}.$$

This becomes after the transformation

$$x_1\sqrt{s} \to x_1,\ x_2 \to x_2,\ s \to s,\ \sqrt{s}\, ds\, dx_1\, dx_2 \to ds\, dx_1\, dx_2$$

and some rearrangement,

$$\int_R\int_R\int_{R_+} D\big(r(x_1/\sqrt{s}, x_2), \theta(x_1/\sqrt{s}, x_2)\big) s^{(\nu-1)/2}\exp\big(-(x_2^2 s + \nu s)/2\big)\, ds\, dx_2$$
$$\times \exp(-x_1^2/2)\exp(\tilde{r}x_1)\, dx_1 = 0, \qquad \text{for all } \tilde{r}.$$

The integral on the left-hand side exists for every $\tilde{r}$. This allows an application of the uniqueness theorem of the bilateral Laplace transform [see, e.g., Widder (1941), Chapter 6, Section 6, Theorem 6h]. Consequently,

$$\begin{aligned}&\int_R\int_R D\big(r(x_1/\sqrt{s}, x_2), \theta(x_1/\sqrt{s}, x_2)\big) s^{(\nu-1)/2}\\ &\quad\times \exp\big(-(x_2^2 s + \nu s)/2\big)\, ds\, dx_2 = 0, \qquad \text{for almost all } x_1.\end{aligned} \tag{6}$$

The left-hand side of (6) is a continuous function in x_1. The function $D(r(x_1/\sqrt{s}, x_2), \theta(x_1/\sqrt{s}, x_2))$ is for fixed x_1 symmetric in x_2. Therefore (6) is equivalent to

$$\int_{R_+}\int_{R_+} D\left(r\left(x_1/\sqrt{s}, x_2\right), \theta\left(x_1/\sqrt{s}, x_2\right)\right) s^{(\nu-1)/2} \times \exp\left(-\left(x_2^2 s + \nu s\right)/2\right) ds\, dx_2 = 0, \qquad \text{for all } x_1. \tag{7}$$

Next we make the transformation $z_1 = x_1/\sqrt{s}$, $z_2 = s(x_2^2 + \nu)/(2x_1^2)$, $ds\,dx_2 = 2x_1^2 z_1^{-1}(2z_1^2 z_2 - \nu)^{-1/2} dz_1\, dz_2$. To simplify the notation let us define the functions K, r^*, θ^* and D^* as

$$K(z_1, z_2) = 1, \quad \text{if } z_2 > \nu/\left(2z_1^2\right), \text{ and 0 otherwise},$$

$$r^*(z_1, z_2) = r\left(z_1, \left(2z_1^2 z_2 - \nu\right)^{1/2}\right),$$

$$\theta^*(z_1, z_2) = \theta\left(z_1, \left(2z_1^2 z_2 - \nu\right)^{1/2}\right),$$

$$D^*(z_1, z_2) = D\left(r^*(z_1, z_2), \theta^*(z_1, z_2)\right).$$

Then (7) becomes

$$\int_{R_+}\int_{R_+} K(z_1, z_2) D^*(z_1, z_2)\left(2z_1^2 z_2 - \nu\right)^{-1/2} z_1^{-\nu}\, dz_1 \times \exp\left(-x_1^2 z_2\right) dz_2 = 0, \qquad \text{for all } x_1.$$

Applying the uniqueness theorem for the unilateral Laplace transform [see, e.g., Widder (1941), Chapter 2, Section 9, Corollary 9.3.b], we may then conclude that

$$\int_{R_+} K(z_1, z_2) D^*(z_1, z_2)\left(2z_1^2 z_2 - \nu\right)^{-1/2} z_1^{-\nu}\, dz_1 = 0, \quad \text{for almost all } z_2. \tag{8}$$

For a given z_2, D^* as a function of z_1 is either nonnegative or nonpositive. To prove this it is sufficient to show that the function values for any pair of numbers z_1', z_1'', such that $z_1' < z_1''$, are either both nonnegative or either both nonpositive. If $D^*(z_1', z_2) = 0$, there is nothing else to show. If $D^*(z_1', z_2) = 1$, then by definition of D

$$w\left(r^*(z_1', z_2)\right) \le \theta^*(z_1', z_2) < v\left(r^*(z_1', z_2)\right).$$

Note that for a given z_2, r^* and θ^* are increasing functions in $|z_1|$. Hence

$$w\left(r^*(z_1'', z_2)\right) \le w\left(r^*(z_1', z_2)\right) \le \theta^*(z_1', z_2) \le \theta^*(z_1'', z_2),$$

which implies that $D^*(z_1'', z_2) \ge 0$. Similarly one can show that if $D^*(z_1', z_2) = -1$, then $D^*(z_1'', z_2) \le 0$. Therefore for a given z_2 the integrand in (8) is either nonnegative or nonpositive, and (8) can be satisfied if and only if $D^*(z_1, z_2) = 0$ almost everywhere. This implies that $D(r, \theta) = 0$ almost everywhere. Consequently, $w(r) = v(r)$ almost everywhere, and because of the continuity constraint imposed on w and v, $w = v$. □

The last proposition immediately yields the following uniqueness property of the Fieller bounds.

THEOREM. *For the Fieller problem the Fieller solution is the only procedure that gives confidence intervals for θ_0 that are of the form $\Theta \pm w(|\tilde{R}|)$, where w is a decreasing function, and that have exact coverage probability for all parameters θ_0, r_0, σ.*

Naturally the last proposition and theorem have corresponding versions for the case σ known. It is of some interest to investigate what happens if the assumptions of the last theorem are relaxed. If we no longer require w to be monotonic it is possible to construct exact confidence bounds of the form $\Theta \pm w(|\tilde{R}|)$ other than the Fieller bounds. Since it leaves the essence of the problem intact, we may choose σ to be known. Figure 1 shows the halfwidth of the 95% limits for the unoriented angle for the Fieller solution and an alternative solution. Both solutions have exact coverage probability for all parameters. The second curve differs from the Fieller curve only in the interval $3.3 \leq R \leq 4.32$ and was found by first changing the Fieller solution in the interval $3.3 \leq R \leq 4$ and then correcting for that change in the interval $4 \leq R \leq 4.3$ using an equivalent to (7).

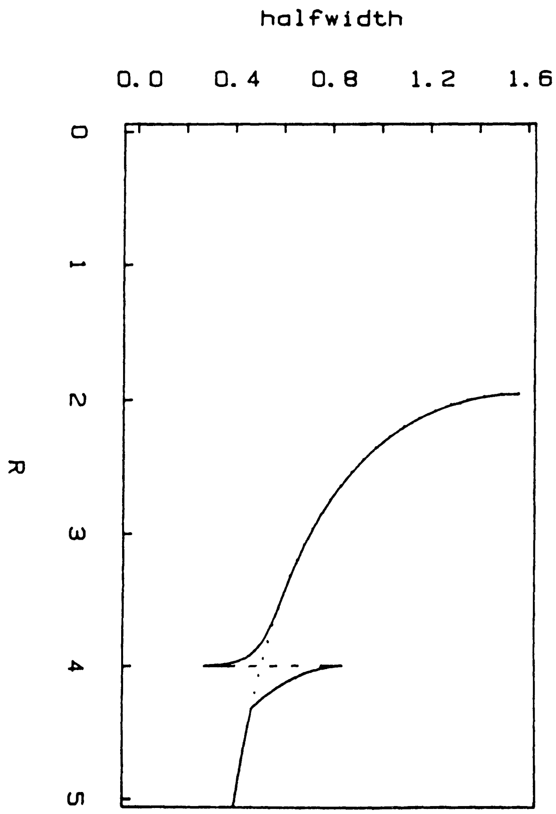

FIG. 1. *Halfwidth profile of the Fieller solution (dotted line) and an alternative (solid line).*

This example demonstrates a characteristic feature of alternative solutions. The negative deviation from the Fieller solution has to be corrected by a brisk increase in bandwidth elsewhere. This property renders them worthless for practical considerations and further enhances the standing of the Fieller solution among the exact bounds.

If we no longer require exact coverage probability for all parameters but are content with conservative bounds for θ_0 we may find a plethora of reasonable alternatives to the Fieller solution. It is not hard to construct intervals of the form $\Theta \pm w(|\tilde{R}|)$, where w is a decreasing function, whose coverage probability exceeds a certain preassigned value for all parameters. In particular, it is possible to find conservative bounds for θ_0 that are nontrivial, i.e., do not enclose all angles with positive probability. We conclude the discussion with the outline of an example. Again assume σ to be known to equal 1, and consider the function v defined on R_+ as

$$\begin{aligned} v(r) &= c_1, && \text{for } 0 \le r \le c_2, \\ &= \arcsin(c_3/r), && \text{for } r > c_2. \end{aligned}$$

The parameters c_1, c_2, c_3 shall be chosen such that v is a continuous function on R_+. It is not hard to verify that the coverage probability of the bounds $\Theta \pm v(|R|)$ increases with r_0. If c_1, c_2, c_3 are chosen such that the coverage probability of these bounds equals α for $r_0 = 0+$, the coverage probability will exceed α for $r_0 > 0$. These bounds are therefore conservative.

Acknowledgment. It is a pleasure to acknowledge the stimulating discussions with Professor C. F. Jeff Wu, my thesis advisor at the University of Wisconsin-Madison, and his support throughout this work.

REFERENCES

FIELLER, E. C. (1940). The biological standardization of insulin. *J. Roy. Statist. Soc.* (*Supp.*) **7** 1–64.

JAMES, A. T., WILKINSON, G. N. and VENABLES, W. N. (1974). Interval estimation for a ratio of means. *Sankhyā Ser. A* **36** 177–183.

WALLACE, D. L. (1980). The Behrens–Fisher and Fieller–Creasy problems. In *R. A. Fisher: An Appreciation* (S. E. Fienberg and D. V. Hinkley, eds.) **1** 119–147. Springer, New York.

WIDDER, D. V. (1941). *The Laplace Transform*. Princeton Univ. Press.

BELL COMMUNICATIONS RESEARCH
290 W. MT. PLEASANT AVENUE
LIVINGSTON, NEW JERSEY 07039-2729