BOUNDS ON THE DISTRIBUTION FUNCTIONS OF THE BEHRENS-FISHER STATISTIC[1]

By M. Ray Mickey and Morton B. Brown

University of California, Los Angeles

1. Introduction. It is commonly accepted in the case of two independently distributed normal variables that the distribution function of the Behrens-Fisher statistic is bounded, for all values of the variance ratio σ_1^2/σ_2^2, by the distribution functions of the Student-t variates with $(n_1 + n_2 - 2)$ and min $(n_1 - 1, n_2 - 1)$ degrees of freedom (df). By the Behrens-Fisher statistic we mean

$$V = [\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)]/(s_1^2/n_1 + s_2^2/n_2)^{\frac{1}{2}}$$

where $x_{11}, \cdots, x_{1i}, \cdots, x_{1n_1}$ and $x_{21}, \cdots, x_{2j}, \cdots, x_{2n_2}$ are samples from the two independent Gaussian distributions $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ respectively and

$$\bar{x}_i = \textstyle\sum_j x_{ij}/n_i;$$

$$s_i^2 = \textstyle\sum_j (x_{ij} - \bar{x}_i)^2/(n_i - 1).$$

The purpose of this note is to supply an analytical proof of the above proposition.

This result has certain implications. If a critical value for the Behrens-Fisher statistic is specified that is constant for all ratios of the observed sample variances, then it should lie between those of the Student-t variates with $(n_1 + n_2 - 2)$ df and with min $(n_1 - 1, n_2 - 1)$ df at the desired level of significance. In the "equivalent degrees of freedom" approaches, it is reasonable that $(n_1 + n_2 - 2)$ and min $(n_1 - 1, n_2 - 1)$ be bounds on the number of degrees of freedom with which to enter the Student-t table; also we may then put limits on the tail probability. However a constant critical value is not desirable in this problem and effective use of prior knowledge may yield critical values which are not bounded by those of the Student-t variate with $(n_1 + n_2 - 2)$ and min $(n_1 - 1, n_2 - 1)$ df [1].

2. Development. A formal statement of the basic proposition is as follows:

Theorem. *Let X be normally distributed with zero mean and unit variance, and let f_1W_1 and f_2W_2 be distributed as chi-square variates with f_1 and f_2 degrees of freedom* (df) *respectively, such that X, W_1, and W_2 are mutually independently distributed. Then for all γ in the interval* $0 \leqq \gamma \leqq 1$,

$$P\{|T_1| < v\} \leqq P\{|V_\gamma| < v\} \leqq P\{|T_2| < v\} \tag{1}$$

where

$$V_\gamma = X/(\gamma W_1 + (1 - \gamma)W_2)^{\frac{1}{2}}$$

Received 21 October 1965; revised 7 January 1966.

[1] Research supported in part by National Institutes of Health Grant No. FR-3.

and T_1 and T_2 are Student-t variables with min (f_1, f_2) *and* $(f_1 + f_2)$ *degrees of freedom respectively.*

PROOF. It is more convenient to work in terms of T_1^2, V^2, and T_2^2. The variable V_γ^2 can be expressed as:

$$
\begin{aligned}
V_\gamma^2 &= X^2/[\gamma W_1 + (1-\gamma)W_2] \\
&= [X^2(f_1+f_2)/(f_1W_1+f_2W_2)]/ \\
&\quad [(f_1+f_2)(\gamma W_1 + (1-\gamma)W_2)/(f_1W_1+f_2W_2)] \\
&= T^2/Z_\gamma ,
\end{aligned} \tag{2}
$$

where

$$T = X/[(f_1W_1+f_2W_2)/(f_1+f_2)]^{\frac{1}{2}} \tag{3}$$

and

$$
\begin{aligned}
Z_\gamma &= \gamma[(f_1+f_2)/f_1]f_1W_1/(f_1W_1+f_2W_2) \\
&\quad + (1-\gamma)[(f_1+f_2)/f_2]f_2W_2/(f_1W_1+f_2W_2) \\
&= \gamma Y/g + (1-\gamma)(1-Y)/(1-g)
\end{aligned} \tag{4}
$$

where Y and g are defined as

$$Y = f_1W_1/(f_1W_1+f_2W_2), \qquad g = f_1/(f_1+f_2). \tag{5}$$

Equation (2) is a rearrangement of an equivalent expression given by Fisher [3]. Since $(f_1W_1 + f_2W_2)$ is distributed as chi-square with $f_1 + f_2$ degrees of freedom, independently of X, T is distributed as Student-t with $(f_1 + f_2)$ degrees of freedom. Also Y is distributed as a $\beta(f_1/2, f_2/2)$ variate independently of $(f_1W_1 + f_2W_2)$ and hence independently of T. Since $E\{Y\} = f_1/(f_1 + f_2) = g$, $E\{Z_\gamma\} = 1$. As a consequence of the independence of T and Z_γ, $P\{|V_\gamma| \leqq v\}$ can be expressed as

$$
\begin{aligned}
P\{|V_\gamma| \leqq v\} &= P\{V_\gamma^2 \leqq v^2\} \\
&= P\{T^2 \leqq v^2 Z_\gamma\} \\
&= E\{G(v^2Z_\gamma\,; 1, f_1+f_2)\}
\end{aligned} \tag{6}
$$

in which $G(F; 1, m)$ denotes the cumulative distribution function of the Snedecor-Fisher F with 1 and m degrees of freedom. From the concavity of $G(F)$, i.e., $G''(F) < 0$, and the linearity of Z_γ with respect to γ, it follows that for $0 < \gamma < 1$, $v > 0$,

$$(\partial^2/\partial\gamma^2)P\{|V_\gamma| < v\} = E\{G''(v^2Z_\gamma\,; 1, f_1+f_2)(\partial v^2 Z_\gamma/\partial\gamma)^2\} < 0. \tag{7}$$

(The propriety of differentiation under the expectation sign follows from the dominated convergence theorem, [5], p. 126, since for $0 < \gamma_0 \leqq \gamma \leqq \gamma_1 < 1$, $v > 0$, $\partial^2 G(v^2Z_\gamma)/\partial\gamma^2$ exists, is integrable in Y for each γ and is bounded over the domain $0 \leqq Y \leqq 1$ uniformly in γ.) Equation (7) implies that the minimum of $P\{|V_\gamma| \leqq v\}$ over the interval $0 \leqq \gamma \leqq 1$ is at either $\gamma = 0$ or $\gamma = 1$. At either end point V_γ is distributed as Student-t so that the minimum is the smaller of $G(v^2; 1, f_1)$ and $G(v^2; 1, f_2)$. Since $G(F; 1, m)$ is an increasing function of m [2], the lower bound

$$P\{|T_1| < v\} = G(v^2; 1, \min(f_1, f_2)) \leqq P\{|V_\gamma| \leqq v\}$$

is established for all γ in the interval $0 \leqq \gamma \leqq 1$.

The upper bound is established by applying Jensen's inequality [4] to (6), from which the concavity of G implies the inequality

$$E\{G(v^2 Z_\gamma)\} \leqq G(v^2 E(Z_\gamma)) \tag{8}$$

and since $E\{Z_\gamma\} = 1$ for all γ, (8) becomes the upper bound

$$P\{|V_\gamma| \leqq v\} \leqq G(v^2; 1, f_1 + f_2) = P\{|T_2| \leqq v\}, \qquad v \geqq 0, \tag{9}$$

of the inequality (1).

The limits cannot be improved because the upper bound is attained at $\gamma = g = f_1/(f_1 + f_2)$, as is readily seen from Equations (4) and (6), and the lower bound is attained at either $\gamma = 0$ or $\gamma = 1$. Q.E.D.

A referee has noted that the theorem could be restated as applying to the distribution of convex combinations of F ratios that are based on the same denominator and have independent numerator mean squares. The result is stated here as a corollary to the theorem.

COROLLARY 1. *Let $f_0 W_0$, $f_1 W_1$, $f_2 W_2$ be independently distributed as chi-square with degrees of freedom f_0, f_1, and f_2. Let $F_i = W_i/W_0$, $i = 1, 2$. Then if $f_0 = 1$, for all γ in the interval $0 \leqq \gamma \leqq 1$,*

$$G(F; f_1 + f_2, 1) \leqq P\{\gamma F_1 + (1 - \gamma)F_2 < F\} \leqq G(F; \min(f_1, f_2), 1), \tag{10}$$

where $G(F; a, b)$ denotes the cumulative distribution function of the F distribution with a (b) degrees of freedom for the numerator (denominator).

PROOF. The result follows directly from (1) by defining W_0 as $W_0 = X^2$ and forming the reciprocal

$$1/V_\gamma^2 = [\gamma W_1 + (1 - \gamma)W_2]/W_0 = \gamma F_1 + (1 - \gamma)F_2 .$$

Our proof of the theorem does not generalize the corollary to more than two degrees of freedom for the denominator; the extension to convex combinations of more than two F ratios presents no difficulties.

The original problem concerning the distribution of the Behrens-Fisher statistics may be stated as:

COROLLARY 2. *Let $x_{11}, \cdots, x_{1i}, \cdots, x_{1n_1}$ and $x_{21}, \cdots, x_{2i}, \cdots, x_{2n_2}$ be samples from two independent Gaussian distributions $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ respectively and $\bar{x}_i$ and s_i^2 ($i = 1, 2$) be the sample means and variances. Then the distribution function of the Behrens-Fisher statistic V is bounded by those of the Student-t variates* *with* $(n_1 + n_2 - 2)$ *and* min $(n_1 - 1, n_2 - 1)$ *degrees of freedom where*

$$V = [\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)]/(s_1^2/n_1 + s_2^2/n_2)^{\frac{1}{2}}. \tag{11}$$

Proof. The result is established by the correspondence:

$$\begin{aligned} X &= [(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)]/[\sigma_1^2/n_1 + \sigma_2^2/n_2]^{\frac{1}{2}}; \\ W_i &= s_i^2/\sigma_i^2, \qquad i = 1, 2; \\ \gamma &= (\sigma_1^2/n_1)/[\sigma_1^2/n_1 + \sigma_2^2/n_2]; \\ f_i &= n_i - 1, \qquad i = 1, 2, \end{aligned} \tag{12}$$

where X, W_i, γ, and f_i are the notation of the previous theorem.

Then V_γ becomes

$$V_\gamma = X/[\gamma W_1 + (1 - \gamma)W_2]^{\frac{1}{2}} = [(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)]/[s_1^2/n_1 + s_2^2/n_2]^{\frac{1}{2}}, \tag{13}$$

which is the Behrens-Fisher statistic.

Corollary 3. *The Behrens-Fisher statistic is asymptotically normal as* min $(n_1, n_2) \to \infty$, *and the approximation error is bounded as*

$$|P\{|V| < v\} - \Phi(v)| \leqq |P\{|T_1| < v\} - \Phi(v)| \tag{14}$$

uniformly in $\theta = \sigma_1^2/\sigma_2^2$, *where* $v \geqq 0$,

$$\Phi(v) = \int_{-v}^{v} 1/(2\pi)^{\frac{1}{2}} e^{-x^2/2}\, dx \tag{15}$$

and where T_1 *is the Student-t variate with* min $(n_1 - 1, n_2 - 1)$ df.

Proof. Since $P\{|T| < v\}$ is an increasing function of the degrees of freedom of T [2], and approaches $\Phi(v)$ as $m \to \infty$, the inequality (14) is implied by (1). Since V is distributed symmetrically about zero, the asymptotic normality follows from the limiting value zero of the right hand side of (14) as min $(n_1, n_2) \to \infty$.

REFERENCES

[1] Brown, Morton B. (1965). A secondarily Bayes approach to the two-means problem. Ph.D. Thesis, Princeton University.

[2] Dunn, Olive Jean (1965). A property of the multivariate t distribution. *Ann. Math. Statist.* **36** 712–714.

[3] Fisher, R. A. (1939). The comparison of samples with possibly unequal variances. *Ann. Eugenics* **9** 174–180.

[4] Jensen, J. L. W. V. (1906). Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Math.* **30** 175–193.

[5] Loève, M. (1963). *Probability Theory* (3rd ed.). Van Nostrand, Princeton.