The Power of Conservative Tests of Significance in the Analysis of Variance

Joseph L. Fleiss

Division of Biostatistics,
Columbia University, U.S.A.

SUMMARY

In some experimental designs the only reasonable procedure for testing a hypothesis using the mean squares of the analysis of variance table is to calculate the ratio of the mean square for the hypothesis to a linear combination of independent mean squares. The distribution under the hypothesis of the ratio is not the standard F distribution, however. A conservative test is based on assigning to the linear combination the minimum of the degrees of freedom associated with its components. Often the design is such that an exact F ratio may be calculated by computing an error term from the raw data; the degrees of freedom for the denominator of the exact ratio are usually equal to those in the conservative test. It is shown that, no matter what the design, there are always values for the parameters such that the conservative test is more powerful than the exact test.

INTRODUCTION

A number of experimental designs are such that no exact tests of some hypotheses can be based on the sufficient statistics. In the analysis of a split-plot design, for example, it may be of interest to compare the mean responses to one of the levels of the split-plot factor across the levels of the whole-plot factor. No exact test based on the mean squares for blocks and error appears to exist (Kemphorne [1952, pp. 377–378]), but an exact test results from ignoring all data not on the split-plot treatment of interest and viewing the remaining data as forming a one-way layout.

In the Behrens-Fisher problem of comparing the means of two normal popu-

1. This work was supported in part by grant DE R01 00793 from the National Institute of Dental Research, U. S. Public Health Service, and forms part of the author's Ph. D. thesis at Columbia University.
lations, the ratio of whose variances is unknown, a rejection region of size exactly \( \alpha \) in the space of the four sufficient statistics has been constructed (Palamodov [1966]), but it is so complicated as to be virtually unusable. A simple exact test results from randomly pairing the observations in the two samples, as in the Scheffe solution [1943].

Assume more generally that the mean squares \( MS_1 \) and \( MS_2 \) are independently distributed, \( MS_i \) as \( \sigma_i^2 \chi^2_{n_i} / n_i \), \( i = 1, 2 \), and that the mean square \( MS_1 \) is distributed as \( (c_1 \sigma_1^2 + c_2 \sigma_2^2) \chi^2_{n, \lambda} / n \), independently of \( MS_1 \) and \( MS_2 \), where \( c_1 \) and \( c_2 \) are known constants. \( \chi^2_v \) denotes a central chi square variate with \( v \) degrees of freedom and \( \chi^2_{v, \phi} \) a noncentral chi square variate with \( v \) degrees of freedom and noncentrality parameter \( \phi \). The definition of the noncentrality parameter is that of Graybill [1961, section 4.2]. An intuitively appealing statistic for testing the hypothesis \( H \) that \( \lambda = 0 \) is

\[
F^* = \frac{MS_1}{c_1 MS_1 + c_2 MS_2}
\]

(1)

The distribution of \( F^* \) is not that of a standard \( F \) variate, however.

Calculations by Satterthwaite [1946], Box [1954] and Welch [1956] indicate that the distribution of \( F^* \) is well approximated by that of \( F_{n, n^*, \lambda} \), i.e., by that of a noncentral \( F \) variate with degrees of freedom \( n \) and \( n^* \) and noncentrality parameter \( \lambda \), where

\[
n^* = \frac{(1+\theta)^2}{\frac{1}{n_1} + \frac{\theta^2}{n_2}}
\]

(2)

with \( \theta = c_2 \sigma_2^2 / c_1 \sigma_1^2 \). For the usual significance levels the critical \( F \) value \( F_{n, m}(\alpha) \) is a decreasing function of \( m \). Since \( n^* \geq n_0 = \min(n_1, n_2) \), it appears that the test

\[
\text{reject } H \text{ if and only if } F^* > F_{n, n_0}(\alpha)
\]

(3)
is conservative, i.e., has a probability of rejecting $H$ when true less than or equal to the nominal $\alpha$ for all values of $\theta$.

The experimental design is often such that a quantity $MS_H$ may be so calculated from the raw data that the ratio

$$F = \frac{MS_T}{MS_H}$$  \hspace{1cm} (4)

is distributed exactly as $F_{n, n_0, \lambda}$. A test with significance level exactly $\alpha$ is then

$$\text{reject } H \text{ if and only if } F > F_{n, n_0} (\alpha).$$  \hspace{1cm} (5)

Kempthorne, in discussing the problem associated with the split-plot design cited above, mentions the possibility of performing a conservative test of the type given in (3) [1952, pp. 378 and 382], but recommends instead an exact test of the type given in (5) when feasible [loc. cit., p. 378]. The idea behind this recommendation is perhaps that, since the conservative test is less likely to reject $H$ when true, it is expected to be less likely to reject $H$ when false, i.e., it is expected to be uniformly less powerful than the exact test. It is shown below, however, that this expectation is unwarranted.

**REGION OF SUPERIORITY OF THE CONSERVATIVE TEST**

The following theorem is proved in the appendix.

**Theorem.** Let $X$ be distributed as $F_{c, a, \varphi}$ and $Y$ as $F_{c, b, \varphi}$, with $a < b$. If $f$ is such that $\Pr \{ Y > f | \varphi = 0 \} < \Pr \{ X > f | \varphi = 0 \}$, there exists a finite value $\varphi^*$ for the noncentrality parameter such that $\Pr \{ Y > f | \varphi \} > \Pr \{ X > f | \varphi \}$ if and only if $\varphi > \varphi^*$.

The statistic $F^*$ (1) is distributed as

$$\frac{(1 + \theta) X^2}{\chi^2_{n_1, \lambda} / n_1 + \theta \chi^2_{n_2, n_2}}.$$

For $\theta = 0$, $n_2 / n_1$ and $\omega$, $F^*$ is distributed exactly as a noncentral $F$ variate, with
denominator degrees of freedom $n_1$, $n_1 + n_2$ and $n_2$ respectively. Assume that $n_0 = \min(n_1, n_2) = n_2$ and let $P(\theta, \lambda)$ denote the power of the conservative test (3). Assuming for the moment that $\lambda$ varies independently of $\theta$, the power of the exact test (5) is $P(\infty, \lambda)$. It follows from the theorem, with $c = n_1$, $a = n_2$, $f = F_{n_1, n_2}^\infty(\alpha)$, $b = n_1$ for $\theta = 0$ and $b = n_1 + n_2$ for $\theta = n_2/n_1$ that, for these two values of $\theta$,

there exists a $\lambda^*(\theta)$ such that $P(\theta, \lambda) > P(\infty, \lambda)$ if and only if $\lambda > \lambda^*(\theta)$.

(6)

If $n_1 = n_2$, then $P(0, \lambda) = P(\infty, \lambda)$ for all values of $\lambda$. The validity of (6) has not been proved for other values of $\theta$, but that it is expected to hold follows from the closeness of the approximation to the distribution of $F^*$ by that of $F_{n_1, n_2}^\infty$, where $n^*$ is given by (2).

Table 1 gives values $\lambda^*$ which are conjectured, on the basis of a number of numerical checks, to be the maxima over $\theta$ of $\lambda^*(\theta)$. The probabilities $P(\theta, \lambda)$ for values of $\theta$ other than 0, $n_2/n_1$ and $\infty$ were computed by numerical integration using the result that the conditional distribution of $F^*$, given that $MS_1/MS_2 = r$, is that of a constant depending on $r$ and $\theta$ times a random variable distributed as $F_{n_1, n_1 + n_2}^\infty$ (Cochran [1951]). The degrees of freedom $n_1$ and $n_2$ were so chosen that Tang’s finite series for the noncentral $F$ distribution [1938, section V] could be used. The power of the exact test at $\lambda^*$ appears to be in the interval (.60, .75) for all degrees of freedom and $\alpha$ either .01 or .05.

--- Table 1 about here ---

It has thus far been assumed that the power of the exact test is independent of $\theta$. This would not be the case if the appropriate measure of departure from the null hypothesis were a quantity other than $\lambda$. The noncentrality parameter may be written
\[ \lambda = \frac{1}{2} \frac{\tau^2}{c_1 \sigma_1^2 + c_2 \sigma_2^2} \]

where \( \tau^2 \) is obtained by replacing each observation in \( SS_I = nMS_I \) by its expected value (see, e.g., Scheffé [1959, p. 39]). A more appropriate measure of departure from the null hypothesis might be

\[ \gamma = \frac{1}{2} \frac{\tau^2}{c_1' \sigma_1^2 + c_2' \sigma_2^2} \]  

with \( c_1' \) and \( c_2' \) possibly different from \( c_1 \) and \( c_2 \).

Letting \( Q_C \) and \( Q_E \) denote the powers of the conservative and exact tests when \( \gamma \) is the appropriate measure, it is seen that

\[ Q_C(\theta, \gamma) = P \left( \theta, \frac{c_1' c_2 + c_1 c_2' \theta}{c_1 c_2 (1+\theta)} \gamma \right) \]

and that

\[ Q_E(\theta, \gamma) = P \left( \infty, \frac{c_1' c_2 + c_1 c_2' \theta}{c_1 c_2 (1+\theta)} \gamma \right). \]

Thus, for a fixed value of \( \gamma > \lambda^* c_2 / c_2' \), the conservative test (test C, say) is more powerful than the exact test (test E, say) provided \( \theta > \theta^* \), where

\[ \theta^* = \frac{c_2}{c_1} \frac{\lambda^*}{c_1' \gamma} - \frac{c_1'}{c_2' \lambda^*} \gamma \]

and, for a fixed value of \( \gamma < \lambda^* c_2 / c_2' \), test C is more powerful than test E provided \( \theta < \theta^* \).

**EXAMPLE**

Consider a two period change-over design with unequal residual treatment.
effects, and suppose that the direct treatment effects are to be tested for equality. Grizzle [1965, Table 4] gives the complete analysis of variance table, from which it is seen that no exact F test based on the sufficient statistics exists. An approximate criterion is 
\[ F^* = 2 \cdot \frac{\text{SSS}}{\text{MSS} + \text{MSE}}, \]
MSS and MSE denoting the mean square for subjects and the mean square for error.

If \( N_1 \) subjects are given the two treatments in one order and \( N_2 \) in the second order, then the two mean squares in the denominator of \( F^* \) both have \( N_1 + N_2 - 2 \) degrees of freedom, and a conservative test rejects the hypothesis if and only if 
\[ F^* > F_{1, N_1 + N_2 - 2}(\alpha). \]
Grizzle recommends instead an exact test using only the data from the first period; the exact criterion has \( N_1 + N_2 - 2 \) degrees of freedom in the denominator.

The noncentrality parameter is
\[ \lambda = \frac{N_1 N_2}{2(N_1 + N_2)} \frac{(\varphi_1 - \varphi_2)^2}{\sigma_e^2 + \sigma_s^2}, \]

where \( \sigma_e^2 \) is the component of variance due to error, \( \sigma_s^2 \) the component of variance due to subject effects, and \( \varphi_1 - \varphi_2 \) the difference between the population effects. Letting 
\[ \sigma_1^2 = \sigma_e^2 = E(\text{MSE}) \] and 
\[ \sigma_2^2 = \sigma_e^2 + 2\sigma_s^2 = E(\text{MSS}), \]
it is seen that \( c_1 = c_2 = 1/2 \). There are three reasonable criteria for measuring departures from the null hypothesis. One is 
\[ \Delta_1 = (\varphi_1 - \varphi_2)^2/(\sigma_e^2 + \sigma_s^2), \]
so that test C is more powerful than test E for all values of \( \theta \) provided \( \Delta_1 > \Delta^* \), where
\[ \Delta^* = \frac{2\lambda^*(N_1 + N_2)}{N_1 N_2} \]
and \( \lambda^* \) is obtained from Table 1 with \( n = 1 \) and \( n_1 = n_2 = N_1 + N_2 - 2 \).
A second possible criterion is \( \Delta_2 = (\varphi_1 - \varphi_2)^2 / \sigma_s^2 \), so that, in (7), \( c_1^* = -1/2 \) and \( c_2^* = 1/2 \). Thus \( \theta^* = (\Delta_2 + \Delta^*)/(\Delta_2 - \Delta^*) \) and test C is more powerful than test E provided \( \rho > \Delta^*/\Delta_2 \), where \( \rho = (\theta - 1)/(\theta + 1) = \sigma_s^2/(\sigma_e^2 + \sigma_s^2) \), the intraclass correlation coefficient. A final possible criterion is \( \Delta_3 = (\varphi_1 - \varphi_2)^2 / \sigma_e^2 \), so that \( c_1^* = 1, c_2^* = 0 \) and \( \theta^* = 2\Delta_3 / \Delta^* - 1 \). Thus test C is more powerful than test E provided \( \rho < 1 - \Delta^*/\Delta_3 \). Table 2 summarizes these results.

Insert Table 2 about here

DISCUSSION

It is seen that the choice of one of the two competing test procedures requires, at least, the definition of a measure of departure from the null hypothesis which is appropriate to the aims of the study and the specification of the minimum value for this measure which is considered to be of practical or theoretical importance. For some measures knowledge is also required of the ratio of the two population variances. The value of the ratio of variances is rarely known exactly (if it were, then the statistic

\[
F^* = \frac{MS_1}{\frac{1+\theta}{n_1+n_2} \left( c_1 n_1 MS_1 + \frac{1}{\theta} c_2 n_2 MS_2 \right)}
\]

which is distributed exactly as \( F_{n_1, n_1+n_2, \lambda} \) would be used). A rough idea of its magnitude is often available from previous experience, however.

If the experiment is such that these requirements cannot all be satisfied, then it is recommended as a general rule that the exact test be employed if even slight departures from the null hypothesis, no matter how measured, are considered important, and that the conservative test be employed if only large departures are considered important.
REFERENCES


APPENDIX

The following lemmas are required.

Lemma 1. Define 
\[ I_x(p, q) = \frac{1}{b(p, q)} \int_0^x w^{p-1}(1-w)^{q-1} \, dw, \]
the incomplete beta function. If \( x < 1 \), \( \lim_{p \to \infty} I_x(p, q) = 0 \).

Proof. If \( X \) is a nonnegative random variable with expectation \( \mu \), Chebycheff's inequality states that \( \Pr\{X > c\} \leq \mu/c \). Since the expectation of a random variable with the distribution function \( I_x(a, b) \) is \( a/(a+b) \), therefore 
\[ I_x(p, q) = 1 - I_{1-x}(q, p) \leq \frac{q}{p+q} \frac{1}{1-x} \rightarrow 0 \]
as \( p \to \infty \), provided \( x < 1 \).

Lemma 2. Let \( a, b, c \) and \( f \) be any positive numbers, with \( a < b \). Define 
\[ s_k = \left( \frac{a+cf}{b+cf} \right)^k \quad \text{if } b+c+2(i-1) \]
and 
\[ C = \frac{c(f-1)}{2}. \]
The sequence \( \{s_k\} \) is such that 
\[ s_k < s_{k+1} \quad \text{if } k < C, \]
\[ = s_{k+1} \quad \text{if } k = C, \]
\[ > s_{k+1} \quad \text{if } k > C. \]
Further, \( \lim_{k \to \infty} s_k = 0 \).

Proof. Since \( a, b, c \) and \( f \) are positive, \( s_k > 0 \) for all \( k \). Now,
\[ \frac{s_{k+1}}{s_k} = \frac{a+cf}{b+cf} \cdot \frac{b+c+2k}{a+c+2k} \]

and (10) is easily verified. Since \( \{s_k\} \) is monotone decreasing for \( k > C \), and is bounded below, it has a limit (Courant [1937, p. 61]). Rewrite \( s_k \) as

\[ s_k = \prod_{i=1}^{k} \left( 1 - \frac{b-a}{b+cf} t_i \right), \]

where

\[ t_i = \frac{2(i-1) - c(f-1)}{a + c + 2(i-1)}. \]

Since \( \lim_{i \to \infty} t_i = 1 \), therefore, given any \( \epsilon > 0 \), there exists an integer \( l_0 \) such that \( i > l_0 \) implies \( t_i > 1 - \epsilon \). Thus,

\[ \lim_{k \to \infty} \prod_{i=1}^{l_0} \left( 1 - \frac{b-a}{b+cf} t_i \right) \cdot \lim_{k \to \infty} \left( 1 - (1-\epsilon) \frac{b-a}{b+cf} \right)^{k-l_0} = 0. \]

Proof of theorem. Define

\[ \Delta(\phi) = \Pr\{X > f(\phi)\} - \Pr\{Y > f(\phi)\}, \]

so that

\[ \Delta(\phi) = \sum_{k=0}^{\infty} \frac{\phi^k}{k!} D_k, \]

where

\[ D_k = \frac{1}{a+cf} \left( \frac{1}{2} c + k, \frac{1}{2} b \right) - \frac{1}{b+cf} \left( \frac{1}{2} c + k, \frac{1}{2} a \right) \]

(see, e.g., Graybill [1961, p. 79]). It follows from Lemma 1 that \( \lim_{k \to \infty} D_k = 0 \). Since

\[ l_1(p, q) = l_x(p+1, q) + \frac{1}{p \tilde{D}(p, q)} x^p (1-x)^q \]
(Jordan [1950, p. 84]), therefore

\[
D_k = \frac{1}{2} \cdot \frac{1}{b+c+k} \cdot \frac{1}{b+cf} \cdot \left(\frac{1}{b+cf}\right)^{\frac{1}{2}c+k} \cdot \left(\frac{1}{b+cf}\right)^{\frac{1}{2}b} 
\]

\[
-\frac{1}{2} \cdot \frac{1}{b+c+k} \cdot \frac{1}{a+cf} \cdot \left(\frac{1}{a+cf}\right)^{\frac{1}{2}c+k} \cdot \left(\frac{1}{a+cf}\right)^{\frac{1}{2}a} + D_{k+1}.
\]

Equation (11) yields, after some simplification,

\[
E_k = D_k - D_{k+1} = h_{c, b}^2(f) \cdot \frac{2f}{a+c+2k} \cdot \left(\frac{c_{+cf}}{a+cf}\right)^k \cdot \frac{B(\frac{1}{2}c, k)^{\frac{1}{2}b}}{B(\frac{1}{2}c, k)^{\frac{1}{2}a}} \left(s_k - r(f)\right), \quad (12)
\]

where \(s_k\) is defined by (8), \(r(f) = h_{c, a}^2(f)/h_{c, b}^2(f)\), and \(h_{n, m}(x)\) is the density function of an \(F\) variate with \(n\) and \(m\) degrees of freedom.

Let \(k_0\) denote the smallest integer greater than \(C\) (9). It may be noted that

\(r(f) < s_{k_0}\), for, if not, then \(s_k < s_{k_0} \leq r(f)\) for all \(k\) by (10), so that, by (12), the sequence \(\{E_k\}\) is nonpositive. Thus the sequence \(\{D_k\}\) is nondecreasing. Since \(\lim_{k \to \infty} D_k = 0\), it follows that \(D_k \leq 0\) for all \(k\), which contradicts the assumption that \(D_0 = k \to \infty\).

\(A(0) > 0\).

Thus, either \(1 \leq r(f) < s_{k_0}\) or \(r(f) < 1\). In the first case, there exist integers \(k' < k_0 < k''\) such that \(E_k < 0\) for \(k < k'\), \(E_k > 0\) for \(k' \leq k < k''\), and \(E_k < 0\) for all \(k \geq k''\).

Therefore, \(\{D_k\}\) increases for \(k < k'\), decreases for \(k' \leq k < k''\), and again increases for \(k \geq k''\). There thus exists a finite integer \(k^*\) such that

\[
D_k \geq 0 \text{ if } k < k^*,
\]

\[
< 0 \text{ if } k \geq k^*.
\]

(13)
In case \( r(f) < 1 \), there exists an integer \( k_1 \geq k_0 \) such that \( E_k \geq 0 \) if and only if \( k \leq k_1 \).

Thus \( \{D_k\} \) is nonincreasing for \( k \leq k_1 \) and is increasing for \( k > k_1 \), so that, for this case too, there exists an integer \( k^* \) such that (13) holds.

Rewrite \( \Delta(\varphi) \) as

\[
\Delta(\varphi) = \frac{e^{-\varphi} \varphi^{k^*}}{(k^*)!} G(\varphi),
\]

where

\[
G(\varphi) = (k^*)! \sum_{k=0}^{k^*-1} \frac{\varphi^{-(k^*-k)}}{k!} D_k + (k^*)! \sum_{k=k^*+1}^{\infty} \frac{\varphi^{k-k^*}}{k!} D_k.
\]

It is easily checked that \( \lim_{\varphi \to 0} G(\varphi) = \infty \), that \( \lim_{\varphi \to \infty} G(\varphi) = -\infty \), and that \( G(\varphi) \) is monotone decreasing in \( \varphi \). There thus exists a finite number \( \varphi^* \) such that \( G(\varphi) < 0 \) if and only if \( \varphi > \varphi^* \), and the theorem is proved.
TABLE 1

Conjectured Minimum Values $\lambda^*$ of the Noncentrality Parameter
Such That $\lambda > \lambda^*$ Implies That the Conservative Test is More
Powerful than the Exact Test for all Values of $\theta$.

<table>
<thead>
<tr>
<th>$n_2$</th>
<th>$n_1$</th>
<th>$n$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>13.4</td>
<td>21.2</td>
<td>28.6</td>
<td>35.9</td>
<td>50.3</td>
<td>64.5</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td></td>
<td>12.8</td>
<td>20.2</td>
<td>27.2</td>
<td>34.1</td>
<td>47.7</td>
<td>61.1</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td></td>
<td>7.2</td>
<td>10.0</td>
<td>12.4</td>
<td>14.7</td>
<td>19.1</td>
<td>23.4</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td></td>
<td>7.1</td>
<td>9.8</td>
<td>12.2</td>
<td>14.4</td>
<td>18.8</td>
<td>23.0</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td></td>
<td>6.1</td>
<td>7.9</td>
<td>9.5</td>
<td>11.0</td>
<td>13.7</td>
<td>16.4</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td></td>
<td>6.0</td>
<td>7.9</td>
<td>9.4</td>
<td>10.9</td>
<td>13.7</td>
<td>16.3</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td></td>
<td>5.6</td>
<td>7.1</td>
<td>8.3</td>
<td>9.5</td>
<td>11.6</td>
<td>13.7</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td></td>
<td>5.6</td>
<td>7.1</td>
<td>8.3</td>
<td>9.5</td>
<td>11.6</td>
<td>13.7</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td></td>
<td>5.3</td>
<td>6.6</td>
<td>7.7</td>
<td>8.7</td>
<td>10.5</td>
<td>12.2</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td></td>
<td>5.3</td>
<td>6.6</td>
<td>7.7</td>
<td>8.7</td>
<td>10.5</td>
<td>12.3</td>
</tr>
</tbody>
</table>

$\alpha = .01$

<table>
<thead>
<tr>
<th>$n_2$</th>
<th>$n_1$</th>
<th>$n$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5.7</td>
<td>8.7</td>
<td>11.5</td>
<td>14.2</td>
<td>19.7</td>
<td>25.0</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td></td>
<td>5.5</td>
<td>8.3</td>
<td>10.9</td>
<td>13.5</td>
<td>18.6</td>
<td>23.7</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td></td>
<td>4.1</td>
<td>5.4</td>
<td>6.7</td>
<td>7.8</td>
<td>10.1</td>
<td>12.3</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td></td>
<td>4.0</td>
<td>5.4</td>
<td>6.6</td>
<td>7.7</td>
<td>9.9</td>
<td>12.1</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td></td>
<td>3.7</td>
<td>4.7</td>
<td>5.6</td>
<td>6.4</td>
<td>8.0</td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td></td>
<td>3.7</td>
<td>4.7</td>
<td>5.6</td>
<td>6.4</td>
<td>7.9</td>
<td>9.4</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td></td>
<td>3.5</td>
<td>4.4</td>
<td>5.1</td>
<td>5.8</td>
<td>7.0</td>
<td>8.2</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td></td>
<td>3.5</td>
<td>4.4</td>
<td>5.1</td>
<td>5.8</td>
<td>7.0</td>
<td>8.2</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td></td>
<td>3.4</td>
<td>4.2</td>
<td>4.8</td>
<td>5.4</td>
<td>6.5</td>
<td>7.5</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td></td>
<td>3.4</td>
<td>4.2</td>
<td>4.9</td>
<td>5.5</td>
<td>6.6</td>
<td>7.6</td>
</tr>
</tbody>
</table>

$\alpha = .05$
### TABLE 2

Regions of Superiority of Exact and Conservative Tests of the Equality of Direct Treatment Effects in a Two-Period Change-Over Design.

<table>
<thead>
<tr>
<th>Measure of Departure from H: ( \phi_1 - \phi_2 = 0 )</th>
<th>Conservative Test</th>
<th>Exact Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta_1 = \frac{(\phi_1 - \phi_2)^2}{\sigma_e^2 + \sigma_s^2} )</td>
<td>( \Delta_1 &gt; \Delta^* )</td>
<td>( \Delta_1 &lt; \Delta^* )</td>
</tr>
<tr>
<td>( \Delta_2 = \frac{(\phi_1 - \phi_2)^2}{\sigma_s^2} )</td>
<td>( \Delta_2 &gt; \Delta^* ) and ( \rho &gt; \frac{\Delta^*}{\Delta_2} )</td>
<td>( \Delta_2 &lt; \Delta^* ) or ( \Delta_2 &gt; \Delta^* ) and ( \rho &lt; \frac{\Delta^*}{\Delta_2} )</td>
</tr>
<tr>
<td>( \Delta_3 = \frac{(\phi_1 - \phi_2)^2}{\sigma_s^2} )</td>
<td>( \Delta_3 &gt; \Delta^* ) and ( \rho &lt; 1 - \frac{\Delta^*}{\Delta_3} )</td>
<td>( \Delta_3 &lt; \Delta^* ) or ( \Delta_3 &gt; \Delta^* ) and ( \rho &gt; 1 - \frac{\Delta^*}{\Delta_3} )</td>
</tr>
</tbody>
</table>