DOCUMENT RESUME

ED 175 903 TM 009 248

AUTHOR Myers, Barbara E.; Pohlmann, John T.
TITLE The Null Hypothesis as the Research Hypothesis.
PUB DATE [Apr 79]
NOTE 26p.; Paper presented at the Annual Meeting of the American Educational Research Association (63rd, San Francisco, California, April 8-12, 1979)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Educational Research; *Hypothesis Testing; Mathematical Models; *Research Design; *Statistical Analysis; *Statistical Studies; Statistics; *Tests of Significance

ABSTRACT

A procedure was developed within hypothesis-testing logic that allows researchers to support a hypothesis that has traditionally been the statistical or null hypothesis. Four activities involved in attainment of this goal were discussed: (1) development of statistical logic needed to define the sampling distribution associated with the hypothesis to be supported; (2) examination of the role of Type I and II errors and power; (3) definition of effect size; and (4) generation of critical values on noncentral sampling distributions of a test statistic. Differences between the traditional procedure of rejecting the null hypothesis and the new procedure allowing support of hypotheses were discussed. Specifically, the new procedure specifies that the researcher must choose a trivial effect size; that the level of significance the experimenter controls for is the Type II (not Type I) error rate; and that a noncentral sampling distribution is defined by the index of effect size and the degrees of freedom upon which the critical region is located. (Author/RD)

THE NULL HYPOTHESIS AS THE RESEARCH HYPOTHESIS

Barbara E. Myers John T. Pohlmann

Introduction

Hypothesis-testing is often used to make statistical inferences from sample data to population characteristics. The logic of hypothesis tests is that of negative inference or indirect proof. In statistical hypothesis-testing a statement is made about a hypothetical population with a known parameter from which a specific sample is said to have been drawn. This statement was called the "null hypothesis" by Fisher (1949) and must be stated in terms of a parameter being set equal to some value. The purpose of this null or statistical* hypothesis is to define the midpoint of the hypothetical sampling distribution against which a test statistic's value will be compared to determine the probability of that value occurring by chance alone within that sampling distribution.

In the classical Neyman-Pearson (1933) hypothesis-testing procedures the researcher initially states two hypotheses about some parameter that are mutually exclusive: the statistical hypothesis, H_o, and the alternative hypothesis, H_1. A probability value is chosen a priori to serve as a criterion against which the probability of the test statistic's value will be compared. If the probability of the sample value is smaller than or equal to the level of significance the researcher concludes that the sample value did not occurr by chance alone but, rather, falls within a different sampling distribution so the H_o is rejected as

*The terms null and statistical hypothesis will be used interchangeably in this paper.

false. Two types of errors may occur as a result of the statistical decision of a hypothesis test: a Type I error occurs when a true H_o is rejected, a Type II error occurs when a false H_o is retained.

Two approaches to the interpretation of the statistical decision have been taken by theorists who follow the Fisherian and the Neyman-Pearson logic of hypothesis-testing. Fisher stated that the statistical decision could have been only to reject H_o when the test statistic falls in the critical region or reserve judgement until further evidence has been gathered (1949). Neyman-Pearson use the phrase "reject or accept" the statistical hypothesis as the two decision choices (1933). Both viewpoints have been discussed at length by such authors as Bakkan (1966), Binder (1963), Grant (1962), and Rozeboom (1960) with no universally accepted conclusion. If one were to read the original articles of Neyman-Pearson and others such as Lindgren (1968) and Winer (1971) as to the nature of the statistical decision it would become clear that one cannot prove or disprove a hypothesis but can only support or fail to support it.

Statement of the Problem

Traditionally, the goal of a statistical hypothesis test has been to reject the statistical hypothesis and to support the alternative hypothesis. This goal is reflected in the setting of a low alpha level to protect against Type I errors and by maintaining a high probability of correctly rejecting a false H_o which is called power. A problem arises when the researcher wants to support a statistical hypothesis. That is, there are several instances when a researcher would want to support a hypothesis that the difference between the real and hypothesized parameter values is zero (the traditional H_o). There currently exist no techniques within statistical hypothesis-testing that allow such a hypothesis to be supported with the researcher being able to control the probability of a Type II error or the probability of a correct decision. It was the purpose of this investigation to develop a procedure within hypothesis-testing logic that allowed the traditional H_o to be supported. The attainment of this goal involved (1) developing the statistical logic needed to define the sampling distribution associated with the hypothesis to be supported, (2) examining the role of Type I and II errors and power, (3) defining effect size and demonstrating its key role in this procedure, and (4) generating critical values on noncentral sampling distributions of a test statistic. Each of these activities will be described in the remainder of this paper.

Circumstances When the Null Hypothesis is to be Supported

Error terms for *t* and *F* tests are often pooled without a statistical test of the underlying assumptions. Even when a statistical test is performed to test the assumption of homogeneity of variance, for example, confusion arises as to how to conclude that the assumption has not been violated. The statistical hypothesis in this particular test would be:

$$H_o: \sigma^2_1 = \sigma^2_2 \ \sigma^2_3 = \ldots\ldots \sigma^2_k$$

that samples come from populations with equal variances which the experimenter would want to support in order to conclude that the homogeneity assumption had been satisfied. Authors such as Kirk (1968) and Winer (1971) suggest that a level of significance, $\alpha = .25$ be used so that the probability of making a Type II error is lowered. Computer software packages such as SPSS (Nie, et al, 1975) and SAS (Barr, et al, 1976) automatically include a test of homogeneity of variance in t-test as suggested by Hays (1963) but no guidelines are included for the appropriate level of significance to be used.

A similar situation occurs when the assumption of homogeneity of regression is tested when the analysis of covariance is used. Kerlinger and Pedhazur (1973) and Kirk (1968) caution against the use of ANCOVA if this assumption is not met. In this test the hypothesis to be supported would be:

$$H_0: \beta_1 = \beta_2 = \beta_3 = \ldots\ldots \beta_k$$

Again, no guidelines exist for the researcher to follow to confidently support this statistical hypothesis.

Chi-square goodness-of-fit tests present the same type of dilemma to the experimenter if he wants to support the hypothesis that the distribution of sample values does not differ significantly from the hypothesized distribution.

There are other instances when the researcher may want to support a hypothesis that a parameter (μ, σ^2, ρ, ρ^2, β, σ, P) is equal to a specific value. These instances occur in path analyses, significance tests of correlation coefficients used to indicate test-retest or intra-, inter-judge reliability, replication and validation studies, and in particular experimental designs such as the ABA design often used in behavior modification studies. In the ABA design, for example, it must be shown that the response rate in the A_1 and A_2 periods are the same in order to conclude that a change in response rate during the B period is due only to the independent variable manipulated by the experimenter. This stipulation implies the support of either or both of the following hypotheses:

$$H_0: \mu_{A_1} - \mu_{A_2} = 0$$

$$H_0: \beta_{A_1} - \beta_{A_2} = 0$$

McNeil, Kelly, and NcNeil (1975) suggest testing these hypotheses at a level of significance as high as $\alpha = .60$ "because the acceptance of the no difference hypothesis is here desired" (p. 434).

These examples suggest that there are several circumstances when one might want to support a hypothesis that has traditionally been the null hypothesis. The classical hypothesis-testing logic does not provide the methodology for such a conclusion that allows the researcher to control the probability of a correct decision.

Traditional Location of the Critical Region

In the traditional hypothesis test the goal is to reject H_o and the following rational. is used:

1. H_o and H_1 are stated in terms of a parameter, P.

   For example, H_o: $P = 0$

   H_1: $P > 0$

2. A level of significance is chose a priori such that

   $$p(s \geq Sc \mid H_o) = \alpha$$

   where s = a test statistic's value

   Sc = a critical value of s that defines a critical region for H_o.

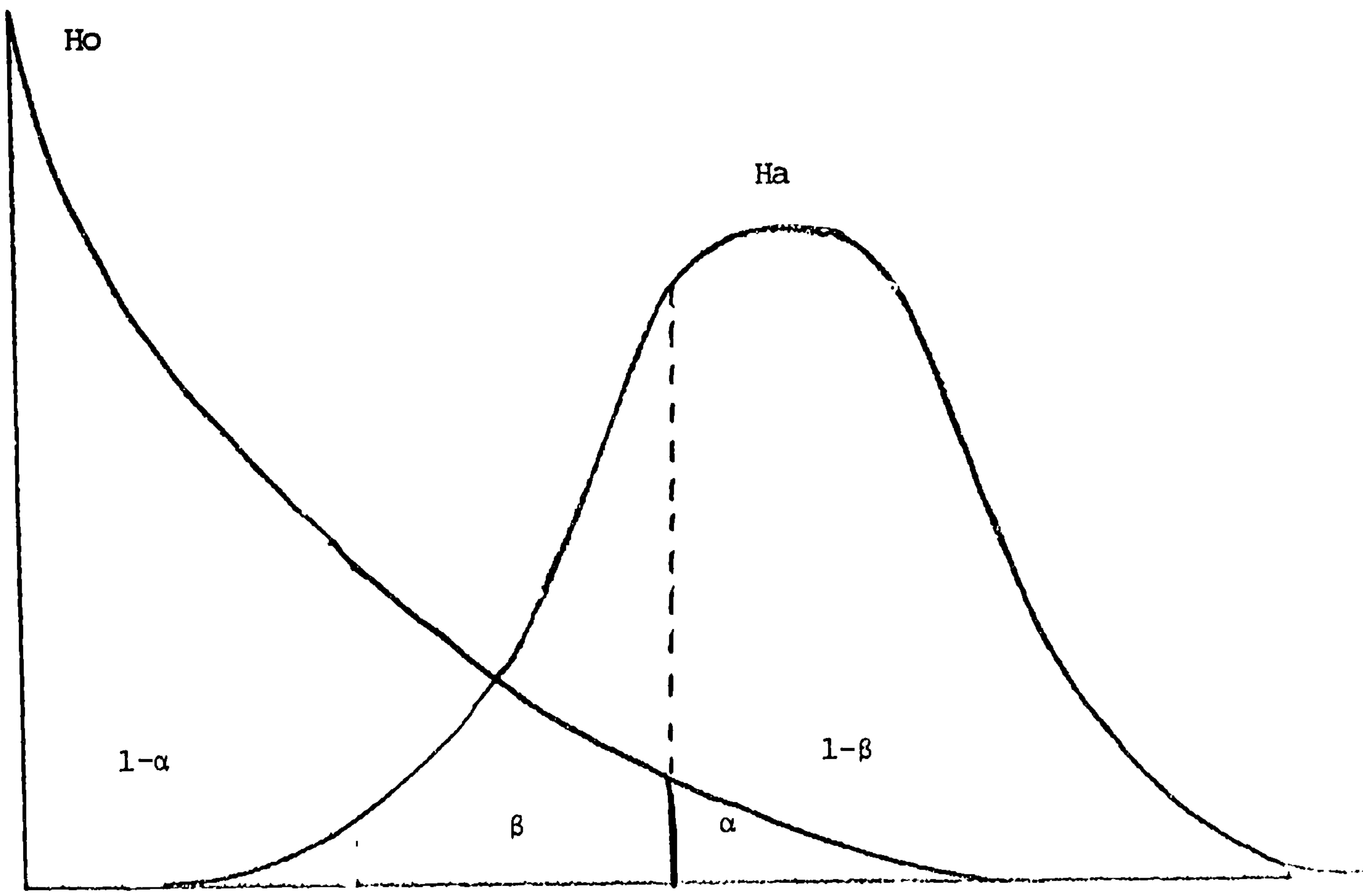
   α = the probability of erroneously rejecting a true H_o.

3. The sample data are analyzed and Ss is calculated.

4. The statistical decision to reject H_o if $Ss \geq Sc$ or not to reject H_o if $Ss < Sc$, where Ss is the sample value of s.

If the H_o is rejected either a correct decision was made or a Type I error occurred. Since α was probably chosen to be small the researcher concludes that the decision was correct and he therefore, finds support for H_1. The probability of making a correct decision is equal to $1 - \beta$ and is found on the noncentral sampling distribution of the test statistic defined by effect size and the degrees of freedom.

Figure I. Location of the Critical Region in a Traditional H_o Test



In the traditional hypothesis test, if the researcher did not want to reject H_0 and the statistical decision was not to reject H_0 two possibilities again exist: a Type II error or a correct decision. This time the experimenter could not conclude that a correct decision had been made because he was not able to control for the probability of a Type II error. Thus, the traditional hypothesis-testing logic does not allow one to support the statistical hypothesis with the same confidence provided by control of statistical errors. A different approach to hypothesis-testing must be taken if the goal of the experiment is to support a hypothesis that has been the null hypothesis traditionally. One approach will be described in this paper, though alternative approaches have been offered by Myers (1978) and in the logic of Bayesian statistics.

Proposed Procedure to Support H_0

In this procedure the same hypotheses could be tested:

1. H_0: $P = 0$

   H_1: $P > 0$

2. The probability of an incorrect decision is chosen a priori such that:

   $p(s \leq Sc \mid P = X) = \beta$

Since the desired outcome is to reject H_1 a noncentral sampling distribution must be defined so that the critical region may be located on it rather than on the usual central distribution. This noncentral sampling distribution of the test statistic is defined by the parameters of effect size and the appropriate degrees of freedom.

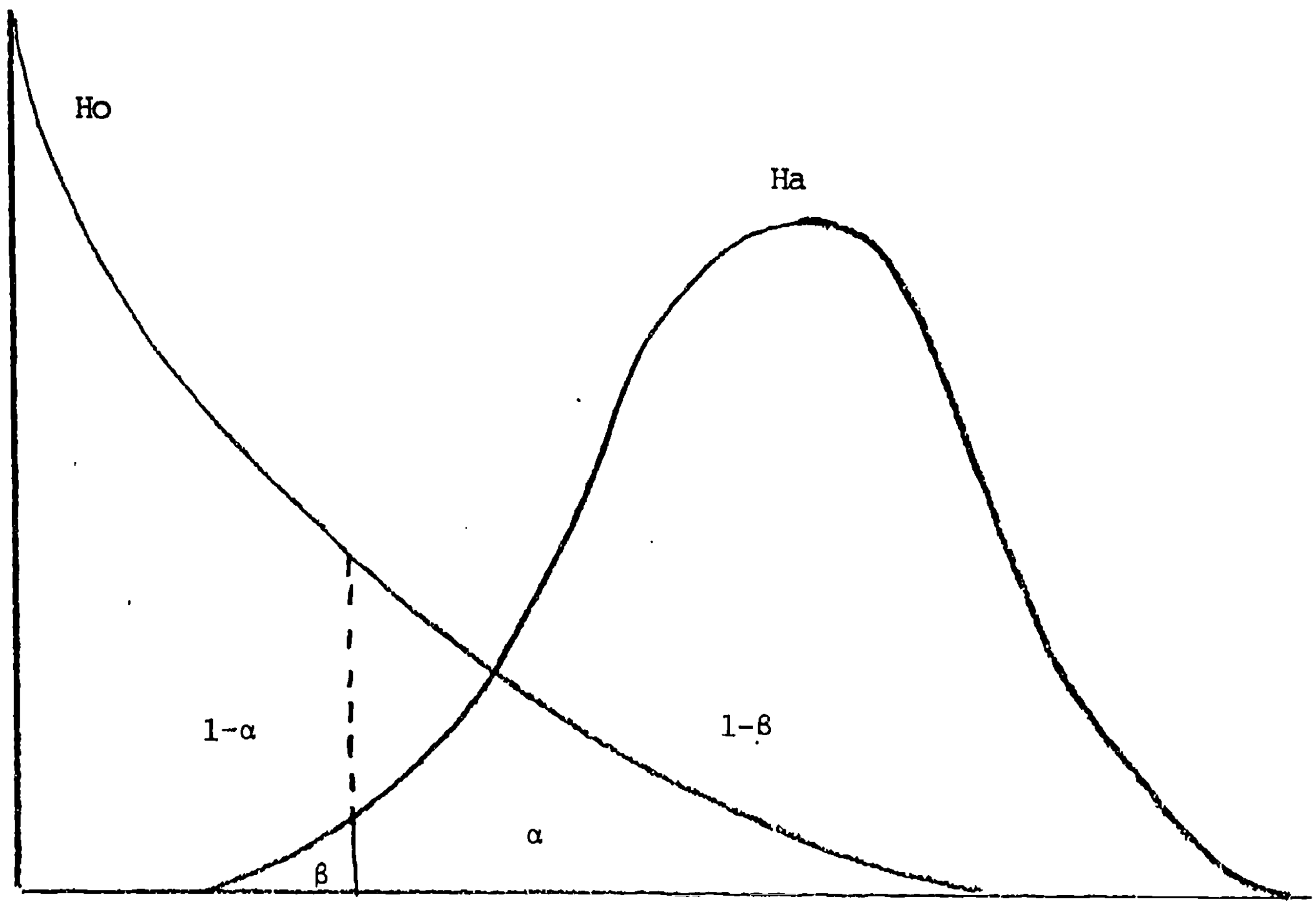
3. A "trivial" effect size, X, is chosen that is the smallest difference between the true and hypothesized parameter values stated in H_0 deemed to have no practical significance.

4. The sample statistic's value Ss is found.

5. The H_1 is rejected if Ss $\leq$ Sc and the H_o is supported.

When the H_o is supported using this method two possibilities exist: a Type II error has occurred or a correct decision $(1 - \alpha)$* has been made.

Figure 2. Location of Critical Region in Proposed Method



* $1 - \beta$ = p(rejecting a false H_o) is found on the noncentral sampling distribution defined by effect size and df in a traditional hypothesis test and is called power. $1 - \alpha$ = p(retaining a true H_o) found on the central distribution in this method has never been named. Myers (1978) has proposed the use of the terms of power I and power II for $1 - \beta$ and $1 - \alpha$ respectively.

To summarize this new approach to hypothesis-testing, when the desired outcome is to support the statistical hypothesis: the researcher chooses a priori a small level of significance to protect against Type II error and chooses a "trivial" effect size which is the smallest difference between the true and hypothesized parameter values judged to have to practical meaning. This effect size and the appropriate degrees of freedom define a noncentral sampling distribution of the test statistic upon which is located the critical region. The test statistic's sample value Ss is calculated and is compared to the Sc value. H_1 is rejected if Ss $\leq$ Sc and the H_o is said to be supported.

Choice of a Test Statistic

The authors decided to choose one test statistic for which a noncentral sampling distribution could be defined and which is used to test several different statistical hypotheses across research applications.

The F statistic possesses these characteristics and was chosen as the test statistic in this investigation. The central F distribution is used commonly to test the following hypotheses:

1. $H_o: \mu_1 = \mu_2 = \mu_3 = \ldots\ldots\mu_k$ — as in ANOVA
2. $H_o: \mu_1 - \mu_2 = 0$ — since t^2 with df = n equals F with $df_1 = 1$, $df_2 = n$
3. $H_o: \sigma^2_1 = \sigma^2_2 = \ldots\ \sigma^2_k$ — to test homogeneity of variance
4. $H_o: \beta_1 = \beta_2 = \ldots\ldots\beta_k$ — to test homogeneity of regression
5. $H_o: \beta = 0 \quad H_o: \beta =$ value — to test regression weights
6. $H_o: \rho^2 = 0 \quad H_o: \rho^2 =$ value — to test population squared multiple correlation

   $H_o: \rho^2_1 = \rho^2_2 = \ldots\rho^2_k$

Indices of Effect Size

Effect size is the degree to which a phenomenon exists in a population that is judged to have some practical importance, or the difference between the true and hypothesized parameter values that has some meaning. The noncentral sampling distribution of a test statistic is defined by effect size and degrees of freedom.

The population squared multiple correlation, ρ^2, was chosen as the index of effect size (ES) for this investigation. The ρ^2 parameter may be interpreted as the proportion of variance accounted for in a dependent variable by a linear model. Several other indices of ES have been developed for power analyses of the F test and are shown to be related to ρ^2 below:

Cohen (1969) $f = \sqrt{\frac{\rho^2}{1-\rho^2}}$

Winer (1971) $\lambda = \frac{n\rho^2}{1-\rho^2}$

Tiku (1967) $\phi = \sqrt{\frac{n\rho^2}{(k+1)(1-\rho^2)}}$ where $k=df_1$ in F ratio

Hays (1963) $\eta^2 = \rho^2$

Definition of a Trivial Effect Size

Cohen (1969) has suggested a convention to define small, medium, and large effect sizes in terms of his f index. The three f values he suggests are f = .10, .25, and .40 to represent small, medium, and large effect sizes. When these f values are converted into ρ^2 values they become ρ^2 = .0099, .0588, and .1379, respectively. These values chosen by Cohen were chosen on the basis of the standardized difference between group means in the population not on their relative frequency of occurrence in the research literature (Cohen, 1969, pp. 277 - 281).

Several unbiased estimates of ρ^2 were calculated from statistical tests reported in educational research articles in order to select a "trivial" effect size in this investigation. A formula suggested by Cohen and Cohen (1975, p. 106) was used to obtain unbiased estimates of ρ^2. This formula is the same as the shrinkage formula for R^2 presented by Kerlinger and Pedhazur (1973, p. 283).

$$\hat{\rho}^2 = \frac{df_1(F-1)}{df_2 + (df_1 F)}$$

where $\hat{\rho}^2$ =the unbiased estimate of ρ^2

df_1=the degrees of freedom numerator of the *F* ratio
df_2=the degrees of freedom denominator of the *F* ratio
F=the sample *F* value

A total of 58 unbiased estimates of $\hat{\rho}^2$ were computed from articles in eight issues of the American Educational Research Journal, and randomly selected issues of the Journal of Educational Measurement, and Experimental Child Psychology. No more than three *F* and/or *t* values were evaluated from any one study so that any massive study would not overwhelm the results. Table I presents the centile values for the frequency distribution of the 58 estimates including the f values for comparison.

Table I. Centile Values for a Distribution of 58 Unbiased Estimates of $\hat{\rho}^2$ and f Taken from Educational Research Journals

Percentile rank	$\hat{\rho}^2$	f
10	.06	.25
25	.12	.37
50	.17	.45
75	.3[illegible]	.69
90	.59	1.20

The median value was .17, reflecting that 17% of the variance was the typical amount accounted for by the linear model used in the statistical test. This median value is larger than Cohen's "large" effect size of $\hat{\rho}^2$ =.1379 and his "medium" effect size of $\hat{\rho}^2$ =.0588 fell at the tenth percentile in this distribution. On the basis of the literature review and Cohen's suggestions two "trivial" effect sizes were chosen to define the noncentral F distributions in this study: $\hat{\rho}^2$ -.0099 and .06.

Generation of Critical F Values on the Noncentral Distribution

The noncentral beta form of the noncentral F distribution was used in this process (Winer, 1971, p. 832). When this noncentral beta density function was integrated the cummulative distribution function for fixed factors was:

$$p(R^2 < R_o^2 | \rho^2, a, b) = e^{-k} \sum \frac{k^i}{i!} I_x R_o^2 (\tfrac{1}{2}a+i, \tfrac{1}{2}b)$$

where R_o^2 = an observed sample coefficient of determination

ρ^2 = the population coefficient of determination
a = the degrees of freedom numerator
b = the degrees of freedom denominator
$k = \tfrac{1}{2}\lambda$, with being the noncentral parameter
I_x = the incomplete beta function

This equation was used to solve for $F(a,b,\lambda)$ iterating on the sample r^2 value until the error was less than 10^{-4}*. The computer software requires the user to set the "trivial" effect size value in the program and then to supply the desired level of significance and df to get the appropriate critical F value on the noncentral F distribution. The obtained values of R^2 were expressed as F values using the following formula:

$$F(a,b) = \frac{R_o^2/a}{(1-R_o^2)/b}$$

Four tables of critical F values are presented at the end of this paper for the noncentral F distributions defined by two effect sizes, $\hat{\rho}^2$ =.0099 and .06, and by various df at two levels of significance (β =.05 and .01). In order to support H_o using this method one would want to find a sample F value equal to or smaller than the critical F value.

Examples of Ho Support Procedure

Namboodiri, Carter, & Blalock (1975,p.99) present an example of a one-way ANOVA in which three groups each of fifteen subjects are compared to determine the effectiveness of three teaching techniques. If the experimenter had wanted to support the hypothesis that all three groups, on the average, did equally well on the achievement test he would need to support the traditional Ho. For the purposes of this paper, assume that the investigator had determined a priori that the smallest amount of variance accounted for by knowledge of group membership in predicting achievement scores was 6% or r^2=.06 and that a significance level of .05 was desired. Thus, the effect size would be ρ^2=.06 and β=.05. Using the tables included with this paper, the critical F value using the conservative degrees of freedom of 2,30 would be .143 (the exact F value at df=2,42 is .203). The observed F value in this example was .10 which is smaller than the critical value needed to retain Ho at the .05 level of significance. The experimenter would have ' en able to conclude that the three teaching techniques yielded equal means on the achievement test with Type II error controlled at .05.

A second example could be taken from Kohout's 1974 study in which an ABA design was used. Kohout investigated the effects of positive verbal stimuli on the response rate of part-word repetitions in stutterers. No stimulation was presented by the experimenter during the A_1 or A_2 baserate and extinction periods and the verbal stimuli were presented contingent upon the occurrence of a part-word repetition during the experimental (B) period. As described previously, the A_1 and A_2 response rates must be shown to be equivalent in order to conclude that a change during the experimental period was due only to the experimental variable.

If Kohout were to re-do her statistical analysis using the Ho support procedure instead of concluding no difference on the basis of a non-significant traditional Ho test she would be able to retain Ho:

$$Ho: \mu_{A_1} - \mu_{A_2} = 0$$
$$Ha: \mu_{A_1} - \mu_{A_2} \neq 0$$

Let us assume that Kohout decided that a difference of one-half of one standard deviation was the smallest difference of consequence. Cohen (1969) presents an effect size index of d for the difference between means where:

$$d = \frac{/m_1 - m_2/}{s}$$

This d index can be converted into the amount of variance accounted for by a linear model using the equation:

$$\rho^2 = \frac{d^2}{(d^2+4)}$$

If Kohout chose one-half of one standard deviation as her effect size the $\hat{\rho}^2$ value would equal .0583 which is close to the tabled effect size of .06. Using the .06 value and a level of significance of .01 the tabled critical F value is .001 at df=1,16 (the actual degrees of freedom should be 1,17 since there were 18 measures per period). The observed F value taken from the Kohout data is .00039 for subject 1 which is smaller than the critical value needed to retain Ho so Kohout would have been able to conclude that the average response rate of part-word repetitions during the baserate and extinction periods was equal.

Conclusion

A procedure was developed within hypothesis-testing logic that allows the researcher to support a hypothesis that traditionally has been the statistical or null hypothesis. The major differences between this procedure and the traditional test when the goal is to reject H_o is that (1) the researcher must choose a "trivial" effect size, (2) the level of significance controlled for by the experimenter is the Type II error rate not a Type I error, and (3) a non-central sampling distribution is defined by the index of effect size and the df upon which the critical region is located.

Readers are encouraged to review Cohen's work on power analysis (1969) in which he provides several examples of expressing effect size in terms of f or for various research designs. The authors of this paper are currently conducting a massive review of research literature in various fields for the compilation of tabled frequency distributions of effect sizes cataloged by area of study, measurement level of the variables, and the research design. Hopefully, tables of this type would serve as useful guidelines to the choise of a trivial effect size for researchers.

The procedure presented in this paper could be applied in any research setting when hypothesis tests are conducted and the desired outcome is to support a hypothesis that states that the difference between the true and hypothesized parameter values is zero.

Table 2

Critical Non Central F Values

Effect Size: $\rho^2 = .06$

5th Percentile

df_1

df_2	1	2	3	4	5	6	7	8	9	10	12	16	20	30	60
1	0.008	0.061	0.110	0.143	0.165	0.182	0.194	0.203	0.211	0.217	0.226	0.239	0.246	0.256	0.266
2	0.006	0.062	0.119	0.161	0.191	0.214	0.231	0.244	0.255	0.264	0.278	0.296	0.308	0.323	0.339
3	0.006	0.063	0.125	0.172	0.207	0.233	0.254	0.270	0.283	0.294	0.311	0.334	0.348	0.367	0.387
4	0.006	0.065	0.130	0.180	0.218	0.247	0.270	0.288	0.303	0.316	0.335	0.361	0.377	0.400	0.424
5	0.007	0.067	0.134	0.187	0.227	0.258	0.283	0.303	0.319	0.332	0.353	0.382	0.400	0.425	0.452
6	0.007	0.069	0.138	0.193	0.235	0.267	0.293	0.314	0.331	0.345	0.368	0.399	0.418	0.446	0.475
7	0.008	0.071	0.142	0.198	0.241	0.275	0.302	0.324	0.342	0.357	0.381	0.413	0.434	0.463	0.495
8	0.008	0.073	0.146	0.203	0.247	0.282	0.309	0.332	0.351	0.366	0.391	0.425	0.447	0.478	0.512
9	0.008	0.076	0.149	0.207	0.253	0.288	0.316	0.339	0.359	0.375	0.401	0.436	0.459	0.491	0.527
10	0.009	0.078	0.153	0.212	0.258	0.294	0.323	0.346	0.366	0.382	0.409	0.445	0.469	0.503	0.540
12	0.010	0.083	0.160	0.221	0.267	0.304	0.334	0.358	0.379	0.396	0.423	0.461	0.486	0.522	0.562
16	0.013	0.094	0.174	0.237	0.285	0.323	0.354	0.379	0.400	0.418	0.447	0.487	0.513	0.552	0.596
20	0.016	0.106	0.190	0.253	0.302	0.340	0.371	0.396	0.418	0.436	0.465	0.507	0.534	0.575	0.622
30	0.030	0.143	0.233	0.296	0.344	0.382	0.412	0.436	0.457	0.475	0.504	0.545	0.573	0.615	0.666
60	0.174	0.325	0.400	0.448	0.483	0.510	0.531	0.549	0.565	0.578	0.600	0.631	0.654	0.690	0.737
100	0.828	0.735	0.712	0.704	0.703	0.703	0.706	0.709	0.712	0.716	0.721	0.733	0.743	0.762	0.793
200	3.762	2.269	1.773	1.527	1.378	1.281	1.214	1.162	1.122	1.091	1.046	0.990	0.960	0.922	0.895
500	16.048	8.464	5.936	4.674	3.917	3.412	3.048	2.779	2.570	2.404	2.152	1.837	1.650	1.402	1.160
1000	40.155	20.540	14.024	10.750	8.786	7.477	6.543	5.842	5.297	4.861	4.208	3.392	2.901	2.249	1.599

Table 3

Critical Non Central F Values

Effect Size: ρ^2 = .06

1st Percentile

df_2 \ df_1	1	2	3	4	5	6	7	8	9	10	12	16	20	30	60
1	0.000	0.012	0.033	0.052	0.067	0.079	0.089	0.096	0.102	0.107	0.115	0.126	0.132	0.141	0.151
2	0.000	0.012	0.037	0.062	0.083	0.101	0.115	0.126	0.136	0.143	0.156	0.173	0.184	0.199	0.215
3	0.000	0.012	0.039	0.068	0.093	0.114	0.131	0.145	0.157	0.167	0.183	0.204	0.218	0.238	0.259
4	0.000	0.013	0.041	0.072	0.100	0.123	0.142	0.158	0.172	0.183	0.202	0.228	0.244	0.268	0.293
5	0.000	0.013	0.043	0.076	0.105	0.130	0.151	0.168	0.183	0.196	0.217	0.246	0.264	0.291	0.321
6	0.000	0.013	0.044	0.078	0.109	0.136	0.158	0.177	0.193	0.206	0.229	0.260	0.281	0.311	0.344
7	0.000	0.014	0.046	0.081	0.113	0.140	0.164	0.184	0.201	0.215	0.239	0.273	0.295	0.327	0.363
8	0.000	0.014	0.047	0.083	0.116	0.145	0.169	0.189	0.207	0.223	0.248	0.283	0.307	0.342	0.381
9	0.000	0.015	0.048	0.085	0.119	0.148	0.173	0.195	0.213	0.229	0.256	0.293	0.318	0.354	0.396
10	0.000	0.015	0.049	0.087	0.122	0.152	0.178	0.200	0.219	0.235	0.262	0.301	0.327	0.365	0.409
12	0.000	0.016	0.052	0.091	0.127	0.158	0.185	0.208	0.228	0.245	0.274	0.315	0.343	0.385	0.432
16	0.001	0.018	0.057	0.098	0.136	0.169	0.198	0.222	0.243	0.262	0.293	0.337	0.368	0.414	0.468
20	0.001	0.021	0.062	0.106	0.145	0.179	0.209	0.234	0.256	0.275	0.307	0.354	0.386	0.436	0.495
30	0.001	0.029	0.077	0.125	0.167	0.203	0.234	0.260	0.283	0.303	0.337	0.386	0.421	0.475	0.541
60	0.009	0.072	0.141	0.197	0.241	0.278	0.308	0.334	0.356	0.375	0.408	0.456	0.490	0.544	0.616
100	0.089	0.213	0.286	0.333	0.370	0.398	0.421	0.441	0.458	0.473	0.498	0.535	0.563	0.609	0.672
200	1.593	1.101	0.947	0.871	0.826	0.799	0.782	0.770	0.761	0.754	0.747	0.741	0.740	0.746	0.769
500	11.058	5.928	4.219	3.365	2.853	2.512	2.265	2.083	1.943	1.830	1.662	1.450	1.328	1.164	1.010
1000	31.965	16.438	11.263	8.659	7.111	6.067	5.323	4.772	4.337	3.996	3.473	2.829	2.441	1.925	1.413

Table 4

Critical Non Central F Values

Effect Size: $\rho^2 = .0099$

5th Percentile

df_2 \ df_1	1	2	3	4	5	6	7	8	9	10	12	16	20	30	60
1	0.006	0.055	0.100	0.132	0.154	0.169	0.181	0.190	0.198	0.204	0.213	0.225	0.232	0.242	0.252
2	0.005	0.054	0.107	0.146	0.176	0.197	0.214	0.227	0.238	0.247	0.261	0.279	0.290	0.305	0.321
3	0.005	0.054	0.110	0.155	0.188	0.214	0.234	0.250	0.263	0.273	0.290	0.313	0.327	0.346	0.366
4	0.005	0.054	0.113	0.160	0.197	0.225	0.247	0.265	0.280	0.292	0.311	0.337	0.353	0.376	0.400
5	0.005	0.054	0.114	0.164	0.202	0.233	0.257	0.276	0.292	0.305	0.327	0.355	0.374	0.400	0.427
6	0.005	0.054	0.116	0.167	0.207	0.238	0.264	0.285	0.302	0.316	0.339	0.370	0.390	0.418	0.449
7	0.005	0.054	0.117	0.169	0.211	0.243	0.270	0.291	0.309	0.325	0.349	0.382	0.403	0.434	0.467
8	0.005	0.055	0.118	0.171	0.214	0.247	0.275	0.297	0.316	0.332	0.357	0.392	0.415	0.447	0.482
9	0.005	0.055	0.118	0.173	0.216	0.251	0.279	0.302	0.321	0.338	0.364	0.400	0.424	0.458	0.496
10	0.005	0.055	0.119	0.174	0.218	0.253	0.282	0.306	0.326	0.343	0.370	0.408	0.433	0.468	0.508
12	0.005	0.056	0.121	0.177	0.221	0.258	0.288	0.313	0.333	0.351	0.380	0.420	0.446	0.485	0.528
16	0.004	0.057	0.123	0.180	0.227	0.265	0.296	0.322	0.344	0.363	0.394	0.437	0.466	0.509	0.558
20	0.005	0.057	0.125	0.183	0.231	0.270	0.302	0.329	0.352	0.372	0.404	0.450	0.480	0.526	0.580
30	0.006	0.060	0.130	0.190	0.239	0.279	0.313	0.341	0.365	0.386	0.420	0.469	0.503	0.554	0.615
60	0.009	0.071	0.145	0.206	0.257	0.298	0.333	0.363	0.387	0.409	0.446	0.498	0.535	0.592	0.665
100	0.009	0.084	0.165	0.229	0.279	0.321	0.355	0.384	0.409	0.431	0.467	0.520	0.557	0.616	0.693
200	0.030	0.137	0.227	0.290	0.339	0.378	0.410	0.436	0.459	0.478	0.511	0.559	0.594	0.650	0.726
500	0.382	0.466	0.505	0.540	0.562	0.576	0.591	0.603	0.615	0.623	0.642	0.667	0.688	0.722	0.776
1000	2.294	1.515	1.256	1.111	1.037	0.998	0.961	0.934	0.913	0.902	0.882	0.857	0.846	0.837	0.844

Table 5

Critical Non Central F Values

Effect Size: $\rho^2 = .0099$

1st Percentile

df_1

df_2	1	2	3	4	5	6	7	8	9	10	12	16	20	30	60
1	0.000	0.010	0.030	0.048	0.062	0.074	0.083	0.090	0.096	0.101	0.108	0.118	0.125	0.134	0.143
2	0.000	0.010	0.033	0.056	0.076	0.093	0.106	0.117	0.126	0.134	0.146	0.163	0.173	0.188	0.203
3	0.000	0.010	0.035	0.061	0.084	0.104	0.120	0.134	0.145	0.155	0.170	0.191	0.205	0.224	0.245
4	0.000	0.010	0.036	0.064	0.090	0.111	0.130	0.145	0.158	0.169	0.187	0.212	0.229	0.252	0.277
5	0.000	0.010	0.036	0.066	0.093	0.117	0.137	0.153	0.168	0.180	0.200	0.229	0.247	0.274	0.303
6	0.000	0.011	0.037	0.068	0.096	0.121	0.142	0.160	0.175	0.189	0.211	0.242	0.262	0.291	0.324
7	0.000	0.011	0.038	0.069	0.098	0.124	0.146	0.165	0.182	0.196	0.219	0.252	0.274	0.306	0.342
8	0.000	0.011	0.038	0.070	0.100	0.127	0.150	0.169	0.187	0.201	0.226	0.261	0.285	0.319	0.358
9	0.000	0.011	0.038	0.071	0.101	0.129	0.152	0.173	0.191	0.206	0.232	0.269	0.294	0.331	0.372
10	0.000	0.011	0.038	0.071	0.103	0.131	0.155	0.176	0.195	0.211	0.237	0.275	0.302	0.340	0.385
12	0.000	0.011	0.039	0.073	0.105	0.134	0.159	0.181	0.201	0.217	0.246	0.287	0.315	0.357	0.405
16	0.001	0.011	0.040	0.075	0.108	0.138	0.165	0.188	0.209	0.227	0.258	0.303	0.334	0.381	0.438
20	0.001	0.011	0.041	0.076	0.110	0.141	0.169	0.193	0.215	0.234	0.266	0.314	0.347	0.399	0.461
30	0.001	0.012	0.042	0.079	0.114	0.147	0.176	0.202	0.225	0.245	0.280	0.332	0.369	0.428	0.501
60	0.002	0.014	0.047	0.086	0.124	0.158	0.189	0.216	0.241	0.263	0.300	0.358	0.399	0.467	0.556
100	0.003	0.017	0.054	0.096	0.135	0.171	0.202	0.230	0.256	0.278	0.317	0.376	0.419	0.490	0.587
200	0.006	0.027	0.075	0.124	0.165	0.202	0.235	0.263	0.288	0.311	0.348	0.407	0.449	0.521	0.622
500	0.015	0.114	0.188	0.241	0.285	0.319	0.348	0.370	0.391	0.411	0.441	0.489	0.524	0.583	0.671
1000	0.702	0.626	0.601	0.589	0.594	0.597	0.600	0.602	0.603	0.611	0.617	0.638	0.651	0.680	0.731

References

Bakken, D. The test of significance in psychological research. Psychological Bulletin, Vol.66,pp.423-37, 1966.

Barr, A.,Goodnight, J.,Sall, J. and Helwig, J. A User's guide to SAS 76. Raleigh,N.C.:SAS Institute, 1976.

Binder, A. Further consideration on testing the null hypothesis and the strategy and tactics of investigating theoretical models. Psychology Review. Vol. 70(1), pp.107-115, 1963.

Cohen, J. Statistical power analysis for the behavioral sciences. New York: Academic Press, 1969.

Cohen,J. and Cohen, P. Applied multiple regression: correlation analysis for the behavioral sciences. New York: John Wiley & Sons, 1975.

Fisher, R.A. The design of experiments.(5th Ed.) Edinburgh: Oliver &Boyd, 1949.

Grant, D. Testing the null hypothesis and the strategy and tactics of investigating theoretical models . Psychological Review. Vol.69(1),pp.34-61, 1962.

Hays, W. Statistics. New York: Holt, Rhinehart & Winston, Inc., 1963.

Kerlinger, F. and Pedhazur, E. Multiple regression in the behavioral sciences.

Kirk, R. Experimental design: procedures for the behavioral sciences. Belmont,CA: Brooks/Cole Publishing Co., 1968.

Kohout, B.Myers The Contingent and non-contingent effects of positive verbal stimulation on selected aspects of the stuttering moment. Unpublished M.S. Thesis, Southern Illinois University, 1974.

Lindgren, B. Statistical theory. New York: Macmillan Co., 1968.

McNeil, K., Kelly, F., & McNeil, J. Testing research hypotheses using multiple linear regression. Carbondale, Illinois:Southern Illinois University Press. 1975.

Myers, B. Investigating procedures employing effect size and level of significance to test statistical hypotheses. Ph.D. Dissertation, Southern Illinois University, 1978.

Namboodiri,N., Carter, L., and Blalock, H. Applied multivariate analysis and experimental designs. McGraw-Hill Co., 1975.

Neyman, J. and Pearson, K. The testing of statistical hypotheses in relation to probabilities a priori. Proceedings of the Cambridge Philosophical Society. Vol.24,pp.492-510, 1933.

Nie, N.,Hull, G., Jenkins, J. Steinbrenner, K. and Bent, D. SPSS: Statistical Package for the social sciences. (2nd Ed.), New York: McGraw-Hill, 1975.

Rozeboom, W. The fallacy of the null hypothesis test. Psychological Bulletin. Vol. 57(5), pp. 416-428, 1960.

Tiku, M. Tables of power of the *F* test. Journal of the American Statistical Association. Vol. 62, pp. 525-36, 1967.

Winer, B. Statistical principles in experimental design. (2nd ed.), New York: McGraw-Hill, 1971.