EXACT UNCONDITIONAL TESTS FOR 2×2 CONTINGENCY TABLES

BY

SAMY SALOMON SUISSA

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL
OF THE UNIVERSITY OF FLORIDA IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1982

To Nicole, Daniel

and

my parents

ACKNOWLEDGEMENTS

I am deeply grateful to Professor Jonathan Shuster for being an instrumental part of my education. His understanding and confidence were important in my entry to this graduate program. Throughout my studies, his patience, help and guidance were invaluable to my work. His contribution to this dissertation was crucial.

I am thankful to my family for their constant support: My wife Nicole, for her patience, love and understanding, my son Daniel, for the joys and the change of pace, and my parents, brother and sisters who were always there with love and encouragement.

My appreciation extends to Professors Ronald Randles, Mark Hale, Ramon Littell and Ken Portier for their full cooperation on such short notice.

TABLE OF CONTENTS

APPENDICES

Abstract of Dissertation Presented to the Graduate Council
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

EXACT UNCONDITIONAL TESTS FOR
2×2 CONTINGENCY TABLES

by

Samy Salomon Suissa

August 1982

Chairman: Jonathan J. Shuster
Major Department: Statistics

The so-called "exact" conditional tests are very popular for testing hypotheses in the presence of nuisance parameters. However, in the context of discrete distributions, they must be supplemented with randomization to become exactly of size α, the nominal significance level. This practice is undesirable since irrelevant events should not affect one's decision. Consequently, the conditional test without randomization, while still called "exact," becomes conservative.

As an unconditional alternative, a methodology is developed to compute the exact size of any test when the null power function is of a given form. This approach is a way of catering to the worst possible configuration of the nuisance parameter by maximizing the null power function over the domain of the nuisance parameter. As special cases, the

2×2 contingency table to compare two independent proportions and the 2×2 contingency table to compare two correlated proportions are considered.

For the equal sample size case, exact critical values of the Z-test for comparing two independent proportions are computed and tabulated for $n=10(1)150$, $\alpha=.025$ and $\alpha=.05$. Sample size requirements based on the exact unconditional one-sided Z-test, with $\alpha=.05$ and 80% power, were never larger than the corresponding sample size requirements based on the "exact" conditional test, namely Fisher's exact test. In fact, the proposed Z-test is uniformly more powerful than Fisher's exact test for $n=10(1)150$, $\alpha=.025$ and $\alpha=.05$.

For comparing two correlated proportions, exact critical values of the Z-test, which is the appropriate square root of McNemar's chi-square test, are computed and tabulated for $N=10(1)200$, $\alpha=.025$ and $\alpha=.05$, where N is the total number of matched pairs. Here again, sample size determinations based on the exact unconditional one-sided Z-test, with $\alpha=.05$ and 80% power, were never larger than the corresponding sample size requirements based on the "exact" conditional test. In fact, the proposed Z-test is uniformly more powerful than the "exact" conditional test for the cases considered, namely $N=10(1)200$, $\alpha=.025$ and $\alpha=.05$.

CHAPTER 1

INTRODUCTION

1.1 The problem

The problem of testing a simple hypothesis about a parameter θ, the index of a known probability distribution P_θ, against a simple alternative hypothesis, for a specified significance level α, is straightforward. Its optimal solution is given by the fundamental lemma of Neyman and Pearson (Lehmann, 1959:65). However, many statistical models involve more than the parameter being tested. The distribution could be indexed by the parameters (θ,ν), where ν is an additional unknown parameter. If inference is required only about θ, the parameter ν is called a nuisance parameter. The problem then becomes one of eliminating this nuisance parameter from the model.

Over the years, many solutions to this problem have been put forward. Basu (1977), in reviewing two such elimination methods, classified all solutions that have been proposed into ten categories. In this dissertation, three of the most common of these methods will be discussed with applications to 2×2 contingency tables. The first method, often used in asymptotic theory, is that of replacing the nuisance parameter by an estimate. The second is the conditional method, which can evolve from two different approaches. The third method considered is the maximization method which essentially caters to the worst possible configuration of the null power function of a test by maximizing this function over the domain of the nuisance parameter. This is the definition of the size of a test as given in Lehmann (1959) or Ferguson (1967). These three methods of eliminating nuisance parameters will be presented in section 1.2, with an emphasis on their advantages and drawbacks. Greater emphasis will be put on the third method, namely the maximization method, since it is the proposed solution to the problems considered in this dissertation. In section 1.3, some intuitive reasons, accentuated by an example from a 2×2 contingency table, will be given as the motivating force for considering the maximization method as a potential competitor to the other two. In section 1.4, the proposed approach will be discussed in the context of the two problems to which it is applied, namely comparing two independent proportions and comparing two correlated proportions. Furthermore, a preview of the results will be given.

1.2 Some Methods of Eliminating Nuisance Parameters

The estimation method. A classical method of eliminating the nuisance parameter ν is to replace it by an estimate $\hat{\nu}$. This practice is popular in asymptotic theory where all the parameters, except the parameter being tested, are often replaced by consistent estimators. This method usually leads to simple test statistics that are referred to well-tabulated distributions. For example, the chi-square test for comparing two independent proportions and McNemar's test for comparing two correlated proportions, both have an asymptotic chi-square null distribution with one degree of freedom. However, because all results are asymptotic or approximate, these methods should not be used for small samples. Further, it is usually unclear what size constitutes a "large enough" sample. This method should be avoided in any study where asymptotic theory is questionable.

The conditional method. The conditional method of eliminating nuisance parameters may evolve from two different approaches. The first of these approaches, due to R.A. Fisher, is described in Kendall and Stuart (1967, vol. 2) as follows. Let (T,U) be sufficient statistics for (θ,ν). If the marginal distribution of U is independent of θ, then U is called an ancillary statistic. The conditionality principle of Fisher is stated as follows: If U is an ancillary statistic (distributed free of θ), then the conditional distribution of $T|U$ is all that is needed to test a hypothesis about θ. This principle is clearly acceptable when U is a sufficient statistic for ν (when θ is known) because the likelihood will then factor into two distributions, namely that of $T|U$ with parameter θ and that of U with parameter ν. The problem arises, however, when U is not sufficient for ν. It is then doubtful whether this principle leads to "optimal" tests. In fact, Welch (1939) gave an example in which the conditional test based on $T|U$ may be uniformly less powerful than an alternative unconditional test.

The conditional method can be obtained through another approach. By restricting the test procedure to a smaller class of tests, namely α-similar or unbiased tests, the nuisance parameter may be eliminated. In particular, if the distribution of (T,U), the sufficient statistics for (θ,ν), is in the exponential family, the distribution of $T|U$ will give UMPU (uniformly most powerful unbiased) tests for hypotheses about θ (Lehmann, 1959:134). Because of the richness of the exponential family, UMPU tests can be readily obtained for a large number of problems through a conditional distribution which is free of the nuisance parameter. However, if $T|U$ is discrete, a randomization supplementing the experiment will be required to make the test exactly of level α. This randomization is undesirable since irrelevant events should not affect one's decision. Hence, the test without randomization becomes conservative and consequently loses power in the process.

The maximization method. The third method of eliminating nuisance parameters is the maximization method. It is based on the direct utilization of the definition of the size of a test, as defined in Lehmann (1959:61), which requires that the null power function of any test to be maximized over the domain of the nuisance parameter. In other words, the maximization method is based on the premise that the worst possible configuration of the nuisance parameter could take place and that the testing procedure should always be protected against this eventuality. This method is easily explained to laymen because they are already familiar with the test of hypothesis for a given value of the nuisance parameter ν. Therefore, by letting ν vary, they will have an intuitive feeling for the maximization method. The major drawback to this method is the complexity of maximizing the null power function. With the advances in computer technology, this method should be investigated more fully.

1.3 Numerical Example

The problem of comparing two independent proportions, called the 2×2 comparative trial, is an example of a test of hypothesis in the presence of a nuisance parameter. If θ_1 and θ_2 are the success rates of the first and second populations respectively, then the equality of these success rates

$\theta_1=\theta_2$ (=θ, unspecified) involves a nuisance parameter, namely θ, the unknown common proportion of success in the two populations. A numerical example is given to present the discrepancies that occur between the first two methods of eliminating that nuisance parameter. Let the outcome of independent random samples of size n=10 each from two binary populations Π_1 and Π_2 be

Table 1.1

	S	F	totals
Π_1	4	6	10
Π_2	8	2	10
totals	12	8	20

where S=success and F=failure. The question is: Is the success rate of Π_1 less than the success rate of Π_2 at significance level α=.05 ?

In the spirit of the estimation method, θ may be replaced by $\hat{\theta}$ = 12/20 in the Z-test statistic with pooled variance estimator (section 3.2) to give

$$Z_p = \frac{\sqrt{20}\ (8-4)}{(12\times 8)^{\frac{1}{2}}} = 1.83$$

which, upon referring to the standard normal distribution, gives an attained significance level (p-value) of .0336, which is less than $\alpha=.05$.

If the conditional argument is used, Fisher's exact test (section 3.3) is the solution to both approaches. By restricting the sample space to only those outcomes with the same marginal totals as Table 1.1, namely

Table 1.2

	S	F	totals
Π_1			10
Π_2			10
totals	12	8	20

the p-value is given by the sum of the hypergeometric probabilities of all tables with the above marginal totals which are more extreme than the observed one (i.e. ≤ 4 successes in Π_1) in the direction of the alternative hypothesis. Accordingly, Fisher's exact test for Table 1.1 has p-value = .0849 > .05.

The statistician is now faced with a dilemma. He would conclude that $\theta_1<\theta_2$ by the first method, but could not by the second method at $\alpha=.05$. This discrepancy raises several questions: How accurate is the normal approximation to the Z-test ? How conservative is Fisher's exact test without randomization ? If the conditional method is preferred, how does one explain to the layman that the sample space is restricted only to the tables with 12 total successes ?

The answers to these questions lie in the null power function, a function of the nuisance parameter θ. For each of the two tests, the exact null power function is plotted on the basis of the attained significance levels as the new nominal significance levels. For the Z-test, the plot of the exact null power function, based on the Z-value of 1.83 (nominal significance .0336), is given in Figure 1.1. It is seen that the exact significance would be greater than the nominal significance level of .0336 when $.13<\theta<.87$. In that range, the Z-test is liberal in the sense that it would reject the null hypothesis when it should not at level .0336. However, notice that the exact null power function never exceeds the original significance level of $\alpha=.05$, the maximum being .047. For the conditional approach, the plot of the exact null power function based on the attained significance level of Fisher's exact test (.0849) is given in Figure 1.2. The conservativeness of this conditional test is obvious. In fact, its null power function is never larger than the original nominal significance level of $\alpha=.05$, the maximum being .045.

Further insight into the reasons for such largely different results can be obtained by inspecting the critical regions of each test. Once more, the Z-test will be based on the nominal significance level .0336 and Fisher's exact

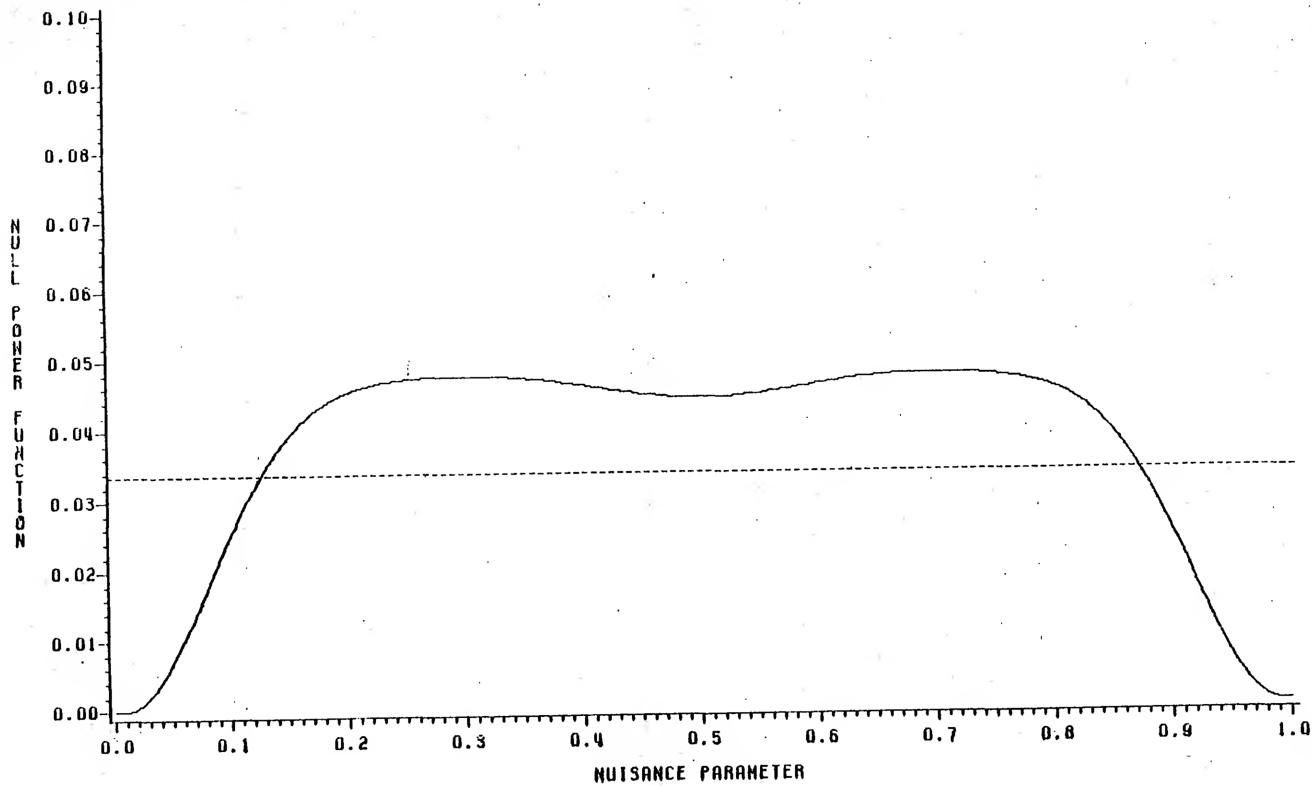

Figure 1.1 Exact null power function of Z-test ($n=10$, $z=1.83$, $\alpha=.0336$) for comparing two independent proportions.

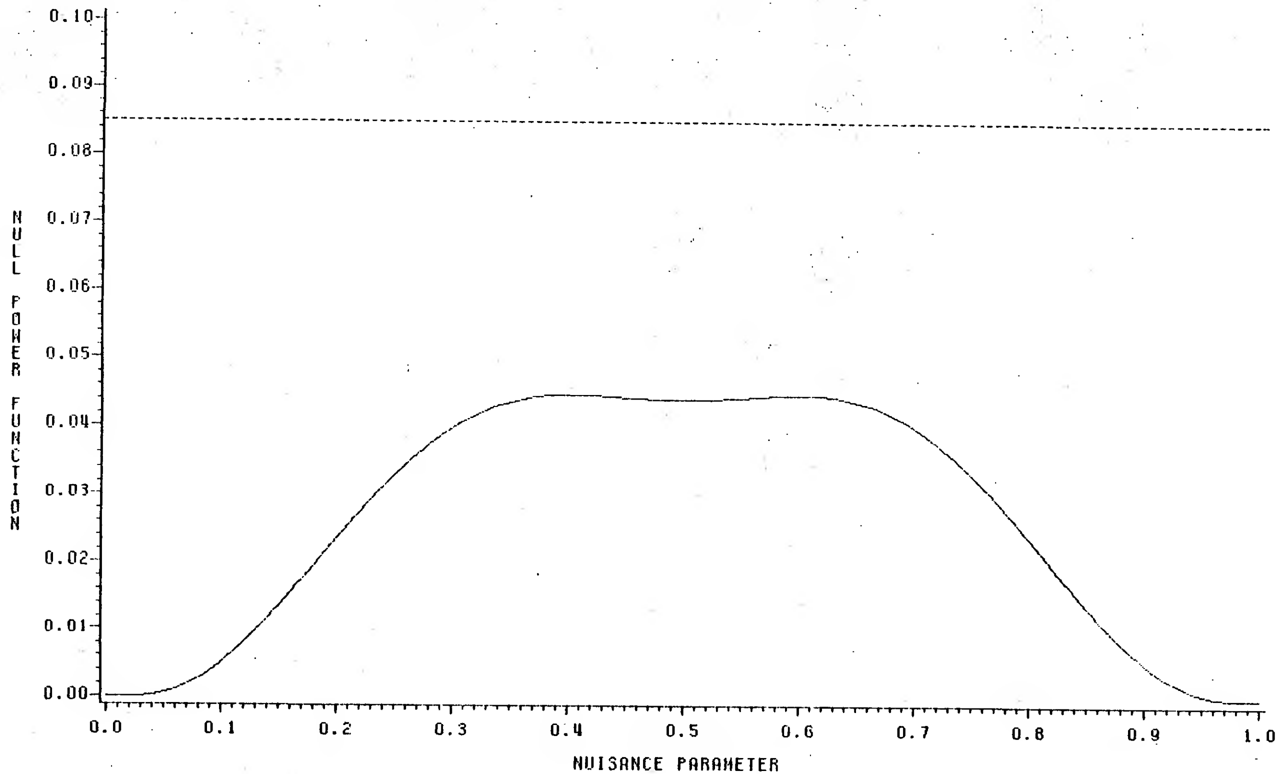

Figure 1.2 Exact null power function of Fisher's exact test with n=10 and nominal significance level .0849.

test on the level .0849. By representing all the possible results of the experiment in the form of points of a lattice diagram, given in Figure 1.3, the critical regions are simply given by marked subsets of these points.

S_2 =10	⊗	⊗	⊗	⊗	⊗	⊗	⊗	O	.	.	.
9	⊗	⊗	⊗	⊗	⊗	⊗	.	.	.	.	.
8	⊗	⊗	⊗	⊗	⊗	.	.	.	.	.	.
7	⊗	⊗	⊗	.	.	.	.	.	.	.	.
6	⊗	⊗	⊗	.	.	.	.	.	.	.	.
5	⊗	⊗	.	.	.	.	.	.	.	.	.
4	⊗	.	.	.	.	.	.	.	.	.	.
3	O	.	.	.	.	.	.	.	.	.	.
2	.	.	.	.	.	.	.	.	.	.	.
1	.	.	.	.	.	.	.	.	.	.	.
0	.	.	.	.	.	.	.	.	.	.	.
	0	1	2	3	4	5	6	7	8	9	10 = S_1

Figure 1.3 Critical regions of Z-test and Fisher's exact test.

In Figure 1.3, S_1= number of successes from Π_1 and S_2 = number of successes from Π_2. The sample points in the lattice diagram marked by an "O" belong to the critical region defined by the Z-test (nominal significance level .0336) and those marked by an "×" belong to the critical region defined by Fisher's exact test (nominal significance level .0849). Although the nominal significance level of the Z-test (.0336) is much smaller than that of Fisher's exact test (.0849), notice that the critical region of the Z-test contains the critical region of Fisher's exact test. This is a flagrant example of the conservativeness of Fisher's exact test and of the liberalness of the Z-test.

These discrepancies suggest that the maximization method might be more appropriate, if not more exact, than the first two methods of eliminating the nuisance parameter. Therefore, by either adjusting the Z-test or the unconditional Fisher's exact test, the maximization method could lead to significance levels which are closer to the nominal levels.

1.4 Proposed Approach and Preview

In this dissertation, the maximization method will be used in conjunction with tests derived from the estimation method to develop an exact unconditional testing procedure as an alternative to the popular, but conservative, conditional method. This procedure will be applied to two problems of the 2×2 contingency table, namely that of comparing two independent proportions and of comparing two correlated proportions. For this purpose, a methodology for computing the supremum of a null power function of a certain general form is developed in Chapter 2. The form of this null power function includes, as particular cases, the null power functions of the two problems considered in this dissertation. In Chapter 3, the supremum of the null power function of the Z-test (the size of the Z-test) for comparing two independent proportions, will be computed for the equal sample size case. Consequently, exact critical values and minimum required sample sizes will be obtained and tabulated for the design and analysis of such a trial. Comparisons with the conditional method (Fisher's exact test) will show that the exact unconditional method leads to smaller (or equal) sample sizes for α=.05, 80% power and common sample size n=10(1)150.

In Chapter 4, the size of the Z-test for comparing two correlated proportions will be computed via the methodology developed in Chapter 2. As a result, exact critical values and sample size determinations will be obtained and tabulated. This exact unconditional test will produce smaller (or equal) sample sizes than the conditional test (the sign test) for α=.05, 80% power and the number of paired observations N=10(1)200.

Furthermore, a comparision of the critical regions will show that the exact unconditional Z-tests are uniformly more powerful than their conditional counterparts in the range considered, namely α=.025, α=.05 and n=10(1)150 for the independent case, and N=10(1)200 for the correlated case.

CHAPTER 2

METHODOLOGY FOR COMPUTING THE SIZE OF A TEST

2.1 Introduction

In this chapter, a methodology is developed to compute the unconditional size of any test of hypothesis T when:

a. There exists a nuisance parameter p such that $0<p<1$

and

b. The null power function of T is a linear combination of a finite number of binomial terms in p.

Thus, we assume that the null power function of T can be written as

$$\pi(p) = \sum_{i \varepsilon C} a_i \, p^{b_i} (1-p)^{c_i} , \tag{2.1.1}$$

where a_i, b_i and c_i are ≥ 0, i is an indexing subscript over the whole sample space S defined by the sampling scheme, C is the set of subscripts for which the related sample points belong to the critical region defined by the testing procedure T and p is a nuisance parameter on the unit interval.

The size of the test T is given by $\sup_p \pi(p)$ and, because only discrete null distributions will be considered, is restricted to a finite collection of possible values, namely, the natural levels of test T as referred to as in Randles and Wolfe (1979). The remainder of this chapter deals with the technique of computing $\sup_p \pi(p)$. The method used is based on the mean value theorem of differential calculus applied to successive subintervals of the unit interval.

The use of this methodology will be illustrated in two important cases. The case of comparing two independent proportions is given in Chapter 3 and that of comparing two correlated proportions in Chapter 4.

2.2 Local Bound for $\pi'(p)$

The applicability of the mean value theorem of differential calculus, as stated in Courant and John (1965), requires primarily a bound on the derivative of $\pi(p)$ for each subinterval. The task of finding such a bound is facilitated by the form of $\pi(p)$, namely that it is a linear combination of binomial terms. In this section, a method of computing this bound for any subinterval is given. First, note that the derivative of $\pi(p)$,

$$\pi'(p) = \sum_{i \varepsilon C} a_i \{ b_i \, p^{b_i - 1} (1-p)^{c_i} - c_i \, p^{b_i} (1-p)^{c_i - 1} \} , \tag{2.2.1}$$

is also a linear combination of binomial terms of the form

$$h(p) = p^r \, (1-p)^{s-r} ,$$

so that for any given subinterval $I=(a,b)$ with $0<a<b<1$,

$$\begin{aligned} \sup_{p\varepsilon I} h(p) &= h(b) && \text{if } \hat{p} > b \\ &= h(a) && \text{if } \hat{p} < a \\ &= h(\hat{p}) && \text{if } \hat{p} \in I \end{aligned} \tag{2.2.2}$$

and

$$\inf_{p\varepsilon I} h(p) = \min\ (\ h(a),\ h(b)\ ) \tag{2.2.3}$$

where $\hat{p} = r/s$.

An upper bound for $\pi'(p)$ can be obtained on (a,b) by substituting the right hand side of (2.2.2) for each positive term of (2.2.1) and the right hand side of (2.2.3) for each negative term of (2.2.1). Similarly, a lower bound for $\pi'(p)$ can be obtained on (a,b) by reversing these substitutions. Finally, a bound M for $|\pi'(p)|$ on (a,b) is taken as the larger of the two bounds, in absolute value.

2.3 A Least Upper Bound for $\pi(p)$

Since local bounds for $|\pi'(p)|$ can now be computed for any subinterval (a,b), the mean value theorem can be applied to the successive subintervals $I_1=(0,.01)$, $I_2=(.01,.02),\ldots,$ $I_{100}=(.99,1)$ of the unit interval to obtain an upper bound for $\pi(p)$ in each I_j. An upper bound for $\pi(p)$, $0<p<1$, is simply the maximum of these local upper bounds.

First, for each I_j, $j=1(1)100$, it is now possible to find a M_j such that

$$|\pi'(\theta_j)| < M_j \,, \qquad \text{for all } \theta_j \varepsilon I_j \,,$$

where the M_j's are obtained by the method of section 2.2. By the mean value theorem of differential calculus, it can be concluded that

$$\pi(\theta_j) \; \varepsilon \; \left(\pi(p_j)-.005M_j \,,\; \pi(p_j)+.005M_j \right), \qquad (2.3.1)$$

for all $\theta_j \varepsilon I_j$, where $p_j=(j-.5)/100$, the midpoint of I_j. For each I_j, $j=1(1)100$, the local maximum is given by

$$\pi^+(p_j) = \pi(p_j) + .005M_j .$$

It can then be deduced that, for p on the unit interval (0,1),

$$\pi(p) < \max_j \{ \pi^+(p_j);\; j=1(1)100 \}. \qquad (2.3.2)$$

Therefore, the right hand side of (2.3.2) is greater than $\sup_p \pi(p)$, the size of T.

Because M_j can be quite large for some values of j, it is possible that the bound on the right hand side of (2.3.2) will be a conservative upper bound for the function $\pi(p)$. This bound can be improved upon to produce a least upper bound of precision δ in the following manner. Any interval I_j for which

$$\pi^+(p_j) > \max_i \{ \pi(p_i);\ i=1(1)100 \} + \delta$$

must be iteratively subdivided into 2^{m_j} subintervals $\{ I_{jk_j};\ k_j=1,2,\ldots,\ 2^{m_j},\ m_j\geq 1 \}$, where m_j is the smallest integer such that

$$\pi(p_{jk_j})+.005\ M_{jk_j}/2^{m_j} < \max\{\max_i\{\pi(p_i)\},\ \max_{u_j}\{\pi(p_{ju_j})\}\}+\delta \tag{2.3.3}$$

for all k_j, and where

$$p_{jk_j} = (j-1)/100 + (2k_j-1)/(100\times 2^{m_j-1})$$

is the midpoint of I_{jk_j} and M_{jk_j} is the local bound for $|\pi'(p)|$ in I_{jk_j}. The inequality (2.3.3) is clearly attainable since, from (2.2.1),

$$M_{jk_j} \leq \sum_{i\varepsilon C} a_i \max(b_i,c_i). \tag{2.3.4}$$

Moreover, using the right hand side of (2.3.4) rather than M_{jk_j} is inefficient and would make computing costs prohibitively large.

By the methodology developed in this section, it is now possible to compute the size (with precision δ) of any test for which the null power function is of the form (2.1.1).

These computations will then shed a light on the extent of conservativeness or liberalness of the tests that are used in the elimination of a nuisance parameter in the present context.

2.4 Stability of the Null Power Function

In the ideal case of the absence of a nuisance parameter, the null power function is constant over the null parameter space. However, when a nuisance parameter is present, the magnitude of its effect on the null power function $\pi(p)$ can be of interest. A relevant feature is the stability or flatness of $\pi(p)$. An indication of this stability can be created using

$$\inf_p \{ \pi(p);\ p \epsilon (p_a, p_b) \} \tag{2.4.1}$$

where p_a and p_b are chosen appropriately for each problem. The difference between $\sup_p \pi(p)$ and (2.4.1) is then an indicator of this stability. The whole unit interval is not used in (2.4.1) because $\pi(0)=\pi(1)=0$. A lower bound for the set in (2.4.1) can be computed from (2.3.1) by

$$\min_j \{ \pi(p_j) - .005M_j;\ j=j_a(1)j_b \}$$

where $j_a = \text{int}(100p_a) + 1$,

$j_b = \text{int}(100p_b) + 1$

and int(.) is the integer function, p_j being as in (2.3.1).

2.5 Choice of the Test Statistic

The choice of the testing procedure is quite arbitrary since the goal here is to compare unconditional to conditional tests. The test statistic that is derived from a testing procedure will simply be a means of dividing the unconditional sample space S into a critical region C and an acceptance region S-C. The choice should then be based on optimal procedures that produce test statistics that are powerful, simple to compute, and intuitively appealing. Three such procedures are the likelihood ratio test criterion, the chi-square goodness-of-fit test and a Z-test based on the asymptotically standardized maximum likelihood estimator, often a function of the sufficient statistic, of the parameter being tested.

The likelihood ratio criterion is given by

$$R = \frac{\sup_{H_0} L(\theta)}{\sup L(\theta)}$$

where $L(\theta)$ is the likelihood function of the sample. If large sample theory applies, the more convenient equivalent statistic

$$X^2 = -2 \log(R) \qquad (2.5.1)$$

can be used because of its limit chi-squared null distribution. The chi-square goodness-of-fit test, based on the statistic

$$X^2 = \sum_i \frac{(O_i - E_i)^2}{E_i} \tag{2.5.2}$$

is especially apropos because of its applicability to multinomial data which lead to null power functions of the form (2.1.1). The asymptotic null distribution of (2.5.2) is also chi-square. The third test is based on the asymptotic normality of $\hat{\theta}_n$, the maximum likelihood estimator, if one exists, of the parameter, call it θ, being tested. The statistic

$$Z = \frac{\hat{\theta}_n - \theta}{s(\hat{\theta}_n)}\ , \tag{2.5.3}$$

where $s^2(\hat{\theta}_n)$ is the asymptotic variance of $\hat{\theta}_n$, or a consistent estimator thereof, has a standard normal asymptotic null distribution under the regularity conditions of likelihood theory. For the two problems considered in Chapters 3 and 4, the maximum likelihood estimator is a function of the sufficient statistic. This statistic and the chi-square goodness-of-fit statistic are the most appealing since (2.5.1) could be computationally laborious.

A further advantage of using the test statistics (2.5.1), (2.5.2) and (2.5.3) is their well-known and well-tabulated asymptotic distributions. These tables can be used to find reasonable starting points for the critical values. Moreover, these values can be compared to the percentage points of the asymptotic distributions and thus provide a study of the accuracy of the large sample approximations.

CHAPTER 3

THE 2×2 TABLE FOR INDEPENDENT PROPORTIONS

3.1 Introduction

A classical problem is the one of comparing two proportions from independent samples. This seemingly simple problem that involves only four numbers, has generated a large amount of literature and has been the subject of much controversy about the use of conditional tests. Since Fisher (1935) proposed the "exact" test, Barnard (1947) and Pearson (1947) started a conflict that has not yet been resolved, as can be seen in the recent articles by Berkson (1978), Barnard (1979), Basu (1979), Corsten and de Kroon (1979) and Kempthorne (1979). Because of the complexity of the power function, only partial attempts have been made in order to resolve the argument. The statement of the problem follows.

Let X and Y be independent binomial random variables with parameters (n,p_1) and (n,p_2) respectively. An experiment that compares p_1 and p_2 is called a 2×2 comparative trial by Barnard (1947), the outcome of which is represented in the form of Table 3.1,

Table 3.1

	S	F	totals
T_1	x	n-x	n
T_2	y	n-y	n

where $p_i = P(S|T_i) = 1-P(F|T_i)$, i=1,2. The labels S and F represent the binary outcomes (S=success, F=failure) and T_1 and T_2 represent the two populations being compared.

The problem is to test, at level α, the null hypothesis $H_0:p_1=p_2$ against the alternative hypothesis $H_a:p_1<p_2$. Because the other one-sided alternative $H_a:p_1>p_2$ and the two-sided alternative $H_a:p_1\neq p_2$ are treated in a similar manner, only this one-sided case will be considered here. Furthermore, only the case of equal sample sizes will be considered because of its optimality under equal sampling costs (Lehmann, 1959:146). The probability of observing the outcome in Table 3.1 is

$$P(X=x,Y=y) = \binom{n}{x} p_1^x (1-p_1)^{n-x} \binom{n}{y} p_2^y (1-p_2)^{n-y}$$

and is, under the null hypothesis $H_0:p_1=p_2$ (=p say),

$$P(X=x,Y=y) = \binom{n}{x}\binom{n}{y} p^{x+y} (1-p)^{2n-x-y} ,$$

a function of the nuisance parameter p, the unspecified common value of p_1 and p_2 under H_0.

Because of this dependence on a nuisance parameter, either approximate tests based on asymptotic results or exact conditional tests are used. Few attempts, however, have been made to compute the exact unconditional size of any of these tests. Barnard (1947) proposed an unconditional test based on $\sup_p \pi(p)$, the size. The criterion that he suggested was intricate and no methodology for computing the size was given. McDonald, Davis and Milliken (1977) tabulated critical regions based on the unconditional size of Fisher's exact test for $n \leq 15$ and $\alpha=.01$ and $.05$. Again, no formal methodology for computing the size was given. Furthermore, no sample size tables or power calculations based on an exact unconditional test exist.

In this chapter, the most common tests are presented. For the asymptotic case, two normal tests and some sample size formulae are given. For the general case, and in particular for small sample sizes, two derivations that both lead to Fisher's exact test are presented. As an alternative to these tests, the results of chapter 2 are used to compute and tabulate the exact unconditional size of two simple statistics as well as the required sample sizes for a significance level of $\alpha=.05$ and a power of $1-\beta=.80$. It is also shown that these tests are uniformly more powerful than Fisher's exact test in the range considered, namely $\alpha=.025$, $\alpha=.05$ and $n=10(1)150$.

3.2 Asymptotic Tests and Sample Size Formulae

A way of circumventing the effect of the nuisance parameter is through the use of asymptotic tests. These approximate tests are appealing because they are usually based on simple test statistics for which the limiting distributions are well tabulated. They are, however, approximations and should not be used when the sample sizes are small.

When n is relatively large, the most widely used tests for the hypothesis of interest are the normal tests. The first one, based on the inversion of the asymptotic confidence interval for p_2-p_1, is the Z-test with an unpooled estimator of the variance and is given by

$$Z_u = \frac{\sqrt{n}\,(\hat{p}_2 - \hat{p}_1)}{(\hat{p}_2\hat{q}_2 + \hat{p}_1\hat{q}_1)^{\frac{1}{2}}} \qquad (3.2.2)$$

where $\hat{p}_1 = x/n = 1-\hat{q}_1$, $\hat{p}_2 = y/n = 1-\hat{q}_2$, with x, y and n as in Table 3.1. The second one, based on the asymptotic null distribution of $\hat{p}_2-\hat{p}_1$, is the Z-test with a pooled variance estimator and is given by

$$Z_p = \frac{\sqrt{n}\,(\hat{p}_2 - \hat{p}_1)}{(\ 2\ \hat{p}\ \hat{q}\ )^{\frac{1}{2}}} \qquad (3.2.2)$$

where $\hat{p}_1$ and $\hat{p}_2$ are as in (3.1.1) and $\hat{p} = (x+y)/2n = 1-\hat{q}$. The limiting distribution of both Z_u and Z_p is the standard normal distribution and an approximate test of size α is based on the percentage points of Φ, the standard normal distribution function. The test statistic Z_p is most frequently used through an equivalent test statistic, the chi-square goodness of fit statistic given by

$$X_p^2 = Z_p^2 \tag{3.2.3}$$

which has a χ_1^2 limiting distribution. Because this chi-square test deals with two-sided alternatives, the statistic Z_p is preferred in the present context of a one-sided test.

The accuracy of the approximation was studied by various authors. The nominal significance level α was compared to the actual significance level by computing (3.4.3) for some values of p. Between the papers by Pearson(1947) and Berkson (1978), numerous studies have shown that for Z_p, the actual level could be larger than the nominal level for some values of p, making this test a liberal one.

To determine the sample size required in each group, two formulae are often used. The first one, based on Z_p and derived in Fleiss (1980), determines the sample size in each group by

$$n_p = \frac{[\ z_\alpha\ (2\bar{p}\bar{q})^{\frac{1}{2}} - z_{1-\beta}\ (p_1q_1+p_2q_2)^{\frac{1}{2}}\ ]^2}{(p_2-p_1)^2} \tag{3.2.4}$$

where z_γ is the upper 100γ percentile of the standard normal distribution, α and β are the type I and type II error probabilities, p_1 and p_2 are the desired alternatives, and $\bar{p} = (p_1+p_2)/2 = 1-\bar{p}$. The second formula, based on the variance stabilizing property of the arcsine transformation on proportions, is given in Cochran and Cox (1957) by

$$n_{as} = \frac{(z_\alpha + z_\beta)^2}{2(\sin^{-1}\sqrt{p_1} - \sin^{-1}\sqrt{p_2})^2} \tag{3.2.5}$$

Other formulae have been derived and are mostly corrected versions of (3.2.4). Kramer and Greenhouse (1959), arguing that the test based on Z_p was too liberal, adjusted (3.2.4) and found

$$n_c = n_p \frac{\{ 1+[1+8(p_2-p_1)/n_p]^{\frac{1}{2}} \}^2}{4(p_2 - p_1)^2} , \tag{3.2.6}$$

where n_p is the sample size found from (3.2.4). More recently, namely since sample size tables based on the exact conditional test were computed, a further adjustment to (3.2.4) was suggested by Casagrande, Pike and Smith (1978a) in order to arrive closer to results based on the exact conditional test. They proposed the formula

$$n_r = n_p \frac{\{ 1+[1+4(p_2p_1)/n_p]^{\frac{1}{2}} \}^2}{4(p_2 - p_1)^2} , \tag{3.2.7}$$

the derivation of which was based on a slight deviation from the derivation of (3.2.6).

3.3 Fisher's Exact Test

The "exact" method of eliminating the nuisance parameter is based on a conditional argument and can be obtained via two different approaches. The first approach, put forward by Fisher (1935), is that of a permutation test. The permutation test argument is a conditional one in that the critical region is constructed on a space conditional on some information from the data. Fisher argues that, because the marginal totals of Table 3.1 alone do not supply any information about the equality of p_1 and p_2, it is reasonable to test conditionally. Thus, given x+y and under $H_0:p_1=p_2$, the probability of Table 3.1 is given by the hypergeometric distribution, namely

$$P(x,y|x+y) = \frac{\binom{n}{x}\binom{n}{y}}{\binom{2n}{x+y}} . \tag{3.3.1}$$

This is Fisher's exact test, the size of which is based on the tail areas of (3.3.1).

The second approach is based on the Neyman-Pearson lemma for testing hypotheses, a thorough treatment of which is given in Lehmann (1959:134) in the case for which a nuisance parameter is present. In the current case, the probability of Table 3.1 is given by

$$P(x,y) = \binom{n}{x} p_1^x (1-p_1)^{n-x} \binom{n}{y} p_2^y (1-p_2)^{n-y}$$

$$= \binom{n}{x}\binom{n}{y} (1-p_1)^n (1-p_2)^n$$

$$\times \exp\{x \log[p_1/(1-p_1)]+y \log[p_2/(1-p_2)]\}$$

$$= \binom{n}{x}\binom{n}{y} (1-p_1)^n (1-p_2)^n$$

$$\times \exp\{x \log\left(\frac{p_1/(1-p_1)}{p_2/(1-p_2)}\right) +(x+y)\log[p_2/(1-p_2)]\}.$$

By Lemma 2 of Lehmann (1959:139), the uniformly most powerful unbiased (UMPU) level α test for comparing p_1 and p_2 is based on the conditional distribution of $X(=x)$ given $T=X+Y(=t)$ and has the form

$$\begin{aligned} \phi(x,t) &= 1 && \text{when } x < C(t) \\ &= \gamma(t) && \text{when } x = C(t) \\ &= 0 && \text{when } x > C(t) \end{aligned}$$

where C and γ are determined by

$$E_{H_0} [\phi(X,T) | T=t] = \alpha$$

for all t, that is

$$\alpha = P_{H_0}[X<C(t)|T=t] + \gamma(t)\ P_{H_0}[X=C(t)|T=t] \ .$$

The conditional distribution is

$$P(X=x|T=t) = \frac{\binom{n}{x}\binom{n}{t-x}\rho^x}{\sum_{u=0}^{t}\binom{n}{u}\binom{n}{t-u}\rho^u}$$

where $\rho = \dfrac{p_1/(1-p_1)}{p_2/(1-p_2)}$ is the odds ratio,

and under H_0, the distribution is given by

$$P_{H_0}(X=x|T=t) = \frac{\binom{n}{x}\binom{n}{t-x}}{\binom{2n}{t}}, \quad x=0,1,\ldots,t,$$

the same hypergeometric distribution found by Fisher's permutation method. Here, $C(t)$ is taken to be the largest value such that

$$\sum_{x=0}^{C(t)-1} \frac{\binom{n}{x}\binom{n}{t-x}}{\binom{2n}{t}} \leq \alpha.$$

In practice, the nonrandomized version of ϕ is used, that is ϕ without the random element $\gamma(t)$. Therefore, the conditional test always has size $\leq \alpha$ and, unlike ϕ, is not UMPU of level α.

3.4 An Exact Unconditional Test

In this section, the methodology of Chapter 2 is used to compute the size of Z_u, the normal test statistic with unpooled variance estimator. This statistic was chosen on the basis of its computational simplicity and its intuitively appealing form. It is given by

$$Z_u = \frac{\sqrt{n}\ (\hat{p}_2 - \hat{p}_1)}{(\hat{p}_2\hat{q}_2 + \hat{p}_1\hat{q}_1)^{\frac{1}{2}}} \tag{3.4.1}$$

where $\hat{p}_1 = x/n = 1-\hat{q}_1$, $\hat{p}_2 = y/n = 1-\hat{q}_2$, with x, y and n as in Table 3.1. The asymptotic null distribution (the standard normal) of Z_u is frequently used in this problem to approximate its actual size. The results of this chapter can thus be used to verify the accuracy of this approximation.

Since x and y are outcomes of independent binomial random variables with parameters (n,p_1) and (n,p_2) respectively , the power function of any test is given by

$$\Pi(p_1,p_2) = \sum_{(x,y)\varepsilon C}\sum \binom{n}{x} p_1^x (1-p_1)^{n-x} \binom{n}{y} p_2^y (1-p_2)^{n-y} \tag{3.4.2}$$

and under $H_0{:}p_1=p_2$ (=p say), the null power function, also denoted by π, is given by

$$\pi(p) = \sum_{(x,y)\varepsilon C}\sum \binom{n}{x}\binom{n}{y} p^{x+y} (1-p)^{2n-x-y} \tag{3.4.3}$$

where p is the nuisance parameter and C is the critical region defined by the test statistic. For the one-sided test of interest, the critical region defined by Z_u is given by

$$C = \{(x,y): Z_u > z_u;\ x,y=0(1)n,\ z_u \geq 0\}. \tag{3.4.4}$$

For an α level test, the critical value of Z_u, namely z_u^*, satisfies the equation

$$z_u^* = \inf\{z_u: \sup_p \pi(p) \leq \alpha\}. \tag{3.4.5}$$

Since (3.4.3) has the form of (2.1.1), the methodology of Chapter 2 can be utilized to find α^*, a value at most δ above $\sup_p \pi(p)$ in (3.4.5).

First, to simplify the computations, (3.4.3) can be reduced to a single summation by solving the inequality $Z_u > z_u$, namely

$$y - x > z_u \left(\frac{y(n-y) + x(n-x)}{n}\right)^{\frac{1}{2}}. \tag{3.4.6}$$

After squaring both sides of (3.4.6), the larger root for y in

$$n\ (y-x)^2 = z_u^2\ [y(n-y) + x(n-x)]$$

is found to be

$$y = h(x) = \frac{b + (b^2-4ac)^{\frac{1}{2}}}{2a}$$

where $a = 1 + z_u^2/n$,

$b = 2x + z_u^2$,

and $c = ax^2 - xz_u^2$.

Hence (3.4.4) reduces to

$$C = \{ (x,y):\ y > h(x)\ ;\ x,y=0(1)n \}. \qquad (3.4.7)$$

Next, (3.4.3) can be written as

$$\pi(p) = \sum_{x=0}^{v} \sum_{y>h(x)} \binom{n}{x}\binom{n}{y} p^{x+y} (1-p)^{2n-x-y}$$

$$= \sum_{x=0}^{v} \binom{n}{x} p^x (1-p)^{n-x} \sum_{y>h(x)} \binom{n}{y} p^y (1-p)^{n-y}$$

$$= \sum_{x=0}^{v} f(x)\ [\ 1 - F(h(x))\ ] \qquad (3.4.8)$$

where $v = \mathrm{int}[\ n^2/(n+z_u^2)\ ]$, int[.] is the integer function, f(.) is the binomial probability mass function with parameters (n,p) and F(.) is its cummulative distribution function.

The derivative of $\pi(p)$, which can also be reduced significantly to simplify the computations, is given by

$$\pi'(p) = \sum_C \binom{n}{x}\binom{n}{y} (x+y)\, p^{x+y-1} (1-p)^{2n-x-y}$$

$$- \sum_C \binom{n}{x}\binom{n}{y} (n-x+n-y)\, p^{x+y} (1-p)^{2n-x-y-1}$$

where $\sum_C$ denotes the double summation $\sum\sum_{(x,y)\in C}$.

It can be rewritten as

$$\pi'(p) = \sum_C n \binom{n-1}{x-1}\binom{n}{y} p^{x+y-1} (1-p)^{2n-x-y}$$

$$+ \sum_C n \binom{n}{x}\binom{n-1}{y-1} p^{x+y-1} (1-p)^{2n-x-y}$$

$$- \sum_C n \binom{n-1}{x}\binom{n}{y} p^{x+y} (1-p)^{2n-x-y-1}$$

$$- \sum_C n \binom{n}{x}\binom{n-1}{y} p^{x+y} (1-p)^{2n-x-y-1}, \qquad (3.4.9)$$

where $\binom{n-1}{-1} = \binom{n-1}{n} = 0$.

Consider the boundary of the critical region C defined by

$$W = \{(x,y) : (x,y)\in C \text{ and } (x+1,y)\notin C\}.$$

The sum of the first and third terms of (3.4.9) becomes, after cancellation of opposing signed identical contributions,

$$\pi_1'(p) = -\sum_W n \binom{n-1}{x}\binom{n}{y} p^{x+y} (1-p)^{2n-x-y-1} . \qquad (3.4.10)$$

The other boundary of C is defined by

$$V = \{(x,y) : (x,y)\varepsilon C \text{ and } (x,y-1)\not\varepsilon C\}.$$

Then the sum of the second and fourth terms of (3.4.9) becomes

$$\pi_2'(p) = \sum_V n \binom{n}{x}\binom{n-1}{y-1} p^{x+y-1} (1-p)^{2n-x-y} . \qquad (3.4.11)$$

Upon combining (3.4.10) and (3.4.11), the derivative of $\pi(p)$ is given by

$$\pi'(p) = \pi_1'(p) + \pi_2'(p) ,$$

and can be further reduced by noticing, from (3.4.4) and (3.4.6), that

$$(x_0,y_0)\varepsilon C \quad \text{iff} \quad (n-y_0,n-x_0)\varepsilon C$$

so that

$$(x_1,y_1)\varepsilon V \quad \text{iff} \quad (n-y_1,n-x_1)\varepsilon C \text{ and } (n-y_1+1,n-x_1)\not\varepsilon C$$

$$\text{iff} \quad (n-y_1,n-x_1)\varepsilon W.$$

The derivative of $\pi(p)$ can finally be written as

$$\pi'(p) = \sum_V n \left[\binom{n}{x}\binom{n-1}{y-1} p^{x+y-1} (1-p)^{2n-x-y} - \binom{n-1}{n-y}\binom{n}{n-x} p^{2n-x-y} (1-p)^{x+y-1} \right]$$

$$= \sum_V n \binom{n}{x}\binom{n-1}{y-1} \left[p^{x+y-1}(1-p)^{2n-x-y} - p^{2n-x-y} (1-p)^{x+y-1} \right] ,$$

a summation over the set of sample points that form a boundary of the critical region C and which are directly obtained from the reduction (3.4.7).

The methodology developed in Chapter 2 can now be utilized to find α^*, the size (of precision δ=.001) of Z_u for any value z_u. For a test of significance of level α, the critical value z_u^* can then be obtained by equation (3.4.5). This is done by using the 100α percentile point of the standard normal distribution as a starting of z_u. This value is then incremented or decremented until $\alpha^* \leq \alpha$ and z_u^* is taken as the smallest value which satisfies this inequality. This procedure was implemented in a FORTRAN computer program, listed in Appendix C.1.

For n=10(1)150 and α=.05 and .025, z_u^* , the exact critical values and α^*, the size (of precision δ=.001) of Z_u, were computed and are given in Table A.1. Furthermore, Table A.1 also contains α_1, a lower bound for $\{\pi(p);\ .05<p<.95\}$

and α_2, a lower bound for $\{\pi(p);\ .10<p<.90\}$, as indicators of the stability of the null power function. The critical values of Z_p, the Z statistic with pooled variance estimator discussed in the next section, are also given in Table A.1 and are denoted by z_p^* .

The null power function, $\pi(p)$, of each test procedure was plotted for some values of n and a nominal significance level of $\alpha=.05$. For n=10, Figure B.1 contains the plot of $\pi(p)$ based on the normal approximation of $Z_u (z_u=1.645)$ and Figure B.2 is based on Z_p $(z_p=1.645)$. Although both graphs exceed the ideal value of $\alpha=.05$ for some values of p, showing that the normal approximation produces a liberal test in this case, it is clear that Z_p gives a better approximation than Z_u when referred to a standard normal distribution. This point is discussed in length in the next section. Figure B.3 is based on the unconditional critical region defined by Fisher's exact test, the criterion used by McDonald, Davis and Milliken (1977) in their unconditional approach. The conservativeness of Fisher's exact test is evident; its actual size in this case being approximately .02. In Figure B.4, the value $z_u^*=1.96$ (or equivalently $z_p^*=1.80$) of Table A.1 is used to plot $\pi(p)$. This plot is seen to perform the best at approaching the nominal level $\alpha=.05$ without exceeding it.

For larger values of n, the null power function of the exact unconditional test Z_u (or equivalently Z_p) is seen to behave better. For n=20, Figure B.5 is the plot of $\pi(p)$

based on $z_u^* = 1.85$ (or $z_p^* = 1.78$) and Figure B.6 is based on $n=30$ and $z_u^* = 1.77$ (or $z_p^* = 1.73$).

3.5 Relation to the Chi-square Goodness-of-fit Test

The other test statistic given in (3.2.2), namely Z_p, the square of which is the chi-square goodness-of-fit test statistic, has a functional relationship with Z_u in the present equal sample size case. This relationship can be derived by first noticing that

$$Z_u = \frac{\sqrt{n}\ (y - x)}{[y(n-y) + x(n-x)]^{\frac{1}{2}}} \tag{3.5.1}$$

and

$$Z_p = \frac{\sqrt{n}\ (y - x)}{[\frac{1}{2}(x+y)(2n-x-y)]^{\frac{1}{2}}} \ . \tag{3.5.2}$$

The square of the denominator of Z_u

$$yn - y^2 + xn - x^2 \tag{3.5.3}$$

can be compared to the square of the denominator of Z_p,

$$yn + xn - \tfrac{1}{2}(x+y)^2 \ . \tag{3.5.4}$$

By rewriting (3.5.3) as

$$yn + xn - \tfrac{1}{2}(x+y)^2 - \tfrac{1}{2}(y-x)^2 ,$$

it is clear that Z_p and Z_u satisfy the relation

$$\frac{1}{Z_u^2} = \frac{1}{Z_p^2} - \frac{1}{2n} \tag{3.5.5}$$

which can be rewritten as

$$Z_p^2 = \frac{2n\ Z_u^2}{2n + Z_u^2}$$

so that the critical value z_p^* of Z_p given in Table A.1 are obtained by

$$Z_p = \mathrm{sgn}(y-x) \left(\frac{2n\ Z_u^2}{2n + Z_u^2}\right)^{\frac{1}{2}}, \tag{3.5.6}$$

where

$$\begin{aligned} \mathrm{sgn}(u) &= 1 \qquad \text{if } u > 0 \\ &= 0 \qquad \text{if } u = 0 \\ &= -1 \qquad \text{if } u < 0 \ . \end{aligned}$$

Robbins (1977) has noted that $|Z_u| \geq |Z_p|$ for the equal sample size case and has posed the question as to which of Z_u or Z_p is more powerful. This question was investigated by Eberhardt and Fligner (1977). They noticed, via a computational argument, that the increase in the significance level for Z_u is compensated fairly well by an increase in power. Moreover, they suggested that Z_u should not be used for small samples because Z_p is closer to a standard normal random variable. In view of the relation (3.5.5), Z_u and Z_p are monotonic increasing functions of each other and are therefore equivalent in the sense that Z_u, with some nominal significance level α, is equivalent to Z_p with some lower level α. Thus, for the same nominal level α, Z_u will reject H_0 more often than Z_p will.

3.6 Power and Sample Sizes.

Given the critical values of Table A.1, it is now possible to compute the exact power by (3.4.2) for $\alpha=.025$ and $\alpha=.05$ and various values of p_1 and p_2. The minimum sample size required per group to attain a power of $1-\beta$ and significance level of α can thus be computed by solving the equation

$$n^* = \min \{n: \Pi(p_1,p_2) \geq 1-\beta\}$$

where the critical region that defines $\Pi(p_1,p_2)$ is based on z_u^*, a function of n.

This equation was solved for $\alpha=.05$ and $1-\beta=.80$ and the results are given in Table A.2 for various combinations of p_1 and p_2. Table A.2 also contains the critical values z_u^*, the size α^* (of precision $\delta=.001$) and the attained power $1-\beta^*$. This table is thus sufficient for both the design and analysis of the 2×2 comparative trial.

Table A.3 compares the results of Table A.2 to the exact conditional test sample sizes [n_e] found in Gail and Gart (1973), Haseman (1978) and Casagrande, Pike, and Smith (1978b). Furthermore, the approximate formulae given in section 3.2 are also computed and compared to n^* and n_e in Table A.3. For the configurations considered, it is seen that n^* tend to be smaller than n_e, the sample sizes determined by Fisher's exact test. Furthermore, the sample sizes based on the arcsine formula [n_{as}] and those based on Z_p, the pooled Z-test [n_p], tend to co-agree quite well and to be, in general, slightly smaller than n^*. The other formulae discussed in section 3.2, namely n_c and n_r are seen to exceed n^* and n_e.

A direct comparision of the critical regions defined by Fisher's exact test and the exact Z-tests was performed numerically. It showed that, for all the cases considered, the critical region defined by Fisher's exact test is contained in the critical region defined by the exact Z-tests. Therefore, the exact Z-tests are uniformly more powerful than Fisher's exact test for the cases n=10(1)150 and $\alpha=.05$ and .025.

CHAPTER 4

THE 2×2 TABLE FOR CORRELATED PROPORTIONS

4.1 Introduction

When the dichotomous responses for each of two regimens are sampled in pairs, either by measuring the same experimental unit under each regimen or by pairing experimental units with respect to some common characteristic, the problem of comparing the success rate of these two regimens involves two correlated proportions. Prior to 1947, this type of data was incorrectly analyzed as if they were independent binomial samples. McNemar (1947) derived the variance of the difference between two correlated binary random variables under the null hypothesis of equal success rate and consequently, using an asymptotic approach, derived the well-known "McNemar's test". This problem, like the independent binomial case, falls into the realm of testing a hypothesis in the presence of a nuisance parameter. Analogous to the independent binomial case (Chapter 3), the most common methods of tackling this problem are based on asymptotic approximations or on the conditional approach. The problem is formulated as follows.

Let (R,S) represent a pair of binary random variables with joint distribution

$$P(R=i,S=j) = p_{ij}\ , \quad i,j=0,1, \quad \sum_{ij}\sum p_{ij} = 1.$$

The outcome of a random sample of N such matched pairs is usually displayed in the form of a 2×2 contingency table such as Table 4.1,

Table 4.1

		S 0	1	totals
R	0	u	x	u+x
	1	y	v	y+v
totals		u+y	x+v	N

where {u,x,y,v} are the frequencies.

The problem is to test, at level α, the null hypothesis H_0:P(R=1) = P(S=1) against one of the alternative hypotheses H_a:P(R=1) < P(S=1), H_a:P(R=1) > P(S=1) or H_a:P(R=1) $\neq$ P(S=1). For the sake of illustration, only the alternative hypothesis H_a:P(R=1) < P(S=1) will be considered. Note that

$$P(R=1) = P(R=1,S=0) + P(R=1,S=1)$$

and $\quad P(S=1) = P(R=0,S=1) + P(R=1,S=1)$

so that the problem becomes that of testing $H_0: p_{01} = p_{10}$ against $H_a: p_{01} > p_{10}$. The likelihood of the sample is given by

$$P(u,x,y,v) = \binom{N}{u\ x\ y\ v} p_{00}^{u}\, p_{01}^{x}\, p_{10}^{y}\, p_{11}^{v}$$

the quadrinomial distribution with probabilities $\{ p_{ij} ; i,j=0,1\}$. Under the null hypothesis $H_0: p_{01} = p_{10}$ (=p say), the likelihood of the sample becomes

$$P_{H_0}(u,x,y,v) = \binom{N}{u\ x\ y\ v} p_{00}^{u}\, p^{x+y}\, (1-p_{00}-2p)^{v},$$

a function of the unspecified common proportion p and an unknown probability p_{00}.

The problem was first tackled by McNemar (1947) who used the asymptotic approach of the standardized sufficient statistic. Cochran (1950), by an intuitive argument, reduced the problem to a sign test, which is the exact conditional test obtained by the Neyman-Pearson approach to the elimination of nuisance parameters. Bennett (1967) has computed the chi-square goodness-of-fit test statistic and observed that it coincides with McNemar's test. A point to note about these asymptotic tests is that they are also conditional in the sense that they only involve x and y, and not N. It turns out that they simply evolve from the asymptotic null distribution of the exact conditional test. No attempts have been made to compute the size of any of these tests, although Bennett and Underwood (1970), in assessing the adequacy of

McNemar's test against its continuity-corrected form, have computed their null power functions for three values of the nuisance parameter p, namely p=.10, .50 and .90. Beyond this investigation, researchers have completely relied upon these asymptotic approximations and the conditional test. It is surprising that conditional tests were not contested in this problem, in light of the fact that they are solely based on the number of discordant pairs x and y, and not at all on the number of concordant pairs u and v. That these tests do not involve N could be disturbing. Lehmann (1959:147), discussing in the context of the sign test with ties, has hinted that N enters the picture through the parameter p_{00} when the unconditional power is computed.

Approximate power calculations and derivations of sample size formulae were made by Miettinen (1968). Bennett and Underwood (1970) compared the exact and approximate powers of McNemar's test and its continuity-corrected form for alternatives close to the null state. Schork and Williams (1980) tabulated the required sample sizes based on the exact power function of the conditional test.

In this chapter, McNemar's test and other asymptotic-type tests will be presented. The approximate sample size formulae will also be given. The exact conditional test will be derived via the Neyman-Pearson approach. The results of Chapter 2 will then be used in section 4.4 to compute and tabulate the size of McNemar's test for the one-sided case.

In section 4.5, the exact unconditional critical values obtained in 4.4 will be used to tabulate the required sample sizes for a significance level of α=.05 and a power of $1-\beta$=.80. It is also shown that this exact unconditional test is uniformly more powerful than the exact conditional sign test for the cases considered, namely α=.05, α=.025 and N=10(1)200.

4.2 McNemar's Test, other Asymptotic Tests and Sample Size Formulae

McNemar (1947) derived the mean and variance of S-R (as defined in Table 4.1) under the null hypothesis and thus proposed the asymptotic test statistic

$$X^2 = \frac{(x - y)^2}{x + y} \tag{4.2.1}$$

for the two-sided alternative. This statistic has an asymptotic χ_1^2 null distribution. Cochran (1950) reduced the problem to a sign test, using the statistic

$$X^2 = \frac{(x - \frac{1}{2}n)^2}{\frac{1}{2}n} + \frac{(y - \frac{1}{2}n)^2}{\frac{1}{2}n}$$

$$= \frac{(x - y)^2}{x + y}$$

where n=x+y, the total number of discordant pairs. Bennett (1967) used the chi-square goodness-of-fit test, applied to the quadrinomial frequencies of Table 4.1, to find

$$X^2 = \frac{(u - N\hat{p}_{00})^2}{N\hat{p}_{00}} + \frac{(x - N\hat{p}_{01})^2}{N\hat{p}_{01}} + \frac{(y - N\hat{p}_{10})^2}{N\hat{p}_{10}} + \frac{(v - N\hat{p}_{11})^2}{N\hat{p}_{11}}$$

$$= \frac{(x - y)^2}{x + y}$$

where the $\hat{p}_{ij}$'s are the maximum likelihood estimators of the p_{ij}'s under H_0. The three methods lead to the same test statistic, namely McNemar's, and therefore have the same asymptotic null distribution.

To determine the required sample size, Miettinen (1968) derived two formulas. The first one, based on an approximation to the asymptotic unconditional power function of X^2 (McNemar's test statistic) gives, for a one-sided test of significance level α and power $1-\beta$, the required sample size as

$$N_{a_1} = \frac{\{ z_\alpha \psi + z_\beta (\psi^2 - \Delta^2)^{\frac{1}{2}} \}^2}{\psi \Delta^2} \tag{4.2.2}$$

where $\psi=p_{01}+p_{10}$, $\Delta=p_{10}-p_{01}$ and z_γ is the upper 100γ percentile of the standard normal distribution. The second formula is based on a more precise approximation to the asymptotic unconditional power function of X^2 and is given by

$$N_{a_2} = \frac{\{ z_\alpha\psi + z_\beta \, [\psi^2 - \tfrac{1}{4}\Delta^2(3+\psi)]^{\frac{1}{2}} \}^2}{\psi \, \Delta^2} . \qquad (4.2.3)$$

For the purpose of comparision with exact conditional and exact unconditional results, these formulas were computed and are given in Table A.6. These comparisions are discussed in section 4.5.

4.3 The Exact Conditional Test

The exact conditional test is obtained by the Neyman-Pearson approach described in Lehmann (1959). The probability of the sample is given by

$$P(u,x,y,v) = \binom{N}{u\ x\ y\ v} p_{00}^{u}\, p_{01}^{x}\, p_{10}^{y}\, p_{11}^{v}$$

and can be written in the exponential family form as

$$P(u,x,y,v) = \binom{N}{u\ x\ y\ v} \exp\{ u \log(p_{00}) + x \log(p_{01})$$

$$+ (N-u-x-v) \log(p_{10}) + v \log(p_{11})\}.$$

It can be reparametrized as

$$P(u,x,y,v) = \binom{N}{u\ x\ y\ v} \exp\{\, u \log(p_{00}/p_{10})$$

$$+ x \log(p_{01}/p_{10}) + v \log(p_{11}/p_{10})$$

$$+ N \log(p_{10}) \,\}.$$

The new parameters are, in the notation of Lehmann (1959),

$$\theta = \log(p_{01}/p_{10})$$

$$\underline{\nu} = (\ \log(p_{00}/p_{10})\ ,\ \log(p_{11}/p_{10})\ ),$$

and the hypothesis to be tested becomes $H_0{:}\theta=0$ against $H_a{:}\theta>0$. The sufficient statistics are

$$X = \sum_{i=1}^{N} (1-R_i)S_i$$

$$\underline{T} = (U,V) = \left(\ \sum_{i=1}^{N} (1-R_i)(1-S_i)\ ,\ \sum_{i=1}^{N} R_i\ S_i\ \right).$$

Therefore the UMPU test is given by

$$\begin{aligned} \phi(x,\underline{t}) &= 1 && \text{when } x > C(\underline{t}) \\ &= \gamma(\underline{t}) && \text{when } x = C(\underline{t}) \\ &= 0 && \text{when } x < C(\underline{t}) \end{aligned}$$

where C and γ are such that

$$E_{H_0}\{\ \phi(X,U,V)\ |\ U=u,\ V=v\ \} = \alpha\ ,\ \text{all } u,v.$$

To find this conditional expectation, first notice that the distribution of (U,V) is

$$P(u,v) = \binom{N}{u\ v\ N-u-v} p_{00}^{u}\ p_{11}^{v}\ (1-p_{00}-p_{11})^{N-u-v}$$

so that the distribution of X given U=u and V=v is

$$P(x|u,v) = \binom{N-u-v}{x\ y} p_{01}^{x}\ p_{10}^{y} \ /\ (p_{01}+p_{10})^{N-u-v}$$

$$= \binom{n}{x} (p_{01}/(p_{01}+p_{10}))^{x}\ \ (p_{10}/(p_{01}+p_{10}))^{y}$$

$$= \binom{n}{x} p^{x}\ (1-p)^{n-x}$$

where $n=x+y$ and $p = p_{01}/(p_{01}+p_{10})$. Therefore, the null hypothesis $H_0{:}p_{01}=p_{10}$ reduces to $H_0{:}p=\frac{1}{2}$, the usual sign test problem based only on n, the total number of discordant pairs. Because this conditional distribution of X is discrete, the test ϕ needs the randomization element γ to become UMPU of level α. However, since the practice of using γ is rare, the test without randomization will be a conservative one and not UMPU of level α.

4.4 An Exact Unconditional Test

As in the case of two independent proportions, the choice of the test statistic is based on the standardization of the sufficient statistic for the parameter being tested, namely $p_{01}-p_{10}$. This statistic is the square root of McNemar's test statistic and is given by

$$Z_c = \frac{x - y}{(x+y)^{\frac{1}{2}}} \tag{4.4.1}$$

where x and y are as in Table 4.1. This statistic is often written in terms of n (=x+y), the total number of discordant pairs, as

$$Z_c = \frac{x - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}} , \tag{4.4.2}$$

the approximation to the sign test, referred to the standard normal distribution. In this section, the methodology of Chapter 2 is used to compute the size of Z_c and the exact critical values based on Z_c. These values will then provide a means of assessing the accuracy of the normal approximation.

The power function of Z_c is given by

$$\Pi(p_{01},p_{10}) = \sum_{(x,y)\varepsilon C}\sum \sum_{u} \binom{N}{u\ x\ y\ v} p_{00}^{u}\, p_{01}^{x}\, p_{10}^{y}\, p_{11}^{v} ,$$

a function only of p_{01} and p_{10} since it is based on the marginal distribution of (X,Y) which is obtained as

$$P(x,y) = \sum_{u=0}^{N-x-y} \binom{N}{u\ x\ y\ v} p_{00}^{u}\, p_{01}^{x}\, p_{10}^{y}\, p_{11}^{N-x-y-u}$$

$$= \binom{N}{x\ y\ N-x-y} p_{01}^{x}\, p_{10}^{y}\, (1-p_{01}-p_{10})^{N-x-y} \, .$$

The power function of Z_c can then be written in terms of n $(=x+y)$ as

$$\Pi(p_{01},p_{10}) = \sum_{(x,y)\varepsilon C}\sum \binom{N}{x\ n-x\ N-n} p_{01}^{x}\, p_{10}^{n-x}$$

$$\times\ (1-p_{01}-p_{10})^{N-n} \, , \qquad (4.4.3)$$

where C is the critical region defined by Z_c. Under the null hypothesis $H_0:p_{01}=p_{10}$ (=θ say), the null power function of Z_c is given by

$$\Lambda(\theta) = \sum_{(x,n)\varepsilon C}\sum \binom{N}{x\ n-x\ N-n} \theta^{n}\, (1-2\theta)^{N-n}$$

or equivalently, if $p = 2\theta$, by

$$\pi(p) = \sum_{(x,n)\varepsilon C}\sum \binom{N}{x\ n-x\ N-n} \tfrac{1}{2}^{n}\, p^{n}\, (1-p)^{N-n} \, , \qquad (4.4.4)$$

a function of the nuisance parameter p, the probability of a discordant pair. For the one-sided test of interest, namely for the alternative hypothesis $H_a:p_{01}>p_{10}$, the critical region C defined by Z_c is given by

$$C = \{(x,n): Z_c > z_c;\ x=0(1)n,\ n=0(1)N,\ z_c>0\}. \qquad (4.4.5)$$

For an α level test, the critical value of Z_c, namely z_c^*, satisfies the equation

$$z_c^* = \inf\{z_c: \sup_p \pi(p) \le \alpha\}. \qquad (4.4.6)$$

Note that $\pi(p)$, as defined in (4.4.4), is a function of p as well as of z_c through (4.4.5). Since (4.4.4) has the form of (2.1.1), the methodology developed in Chapter 2 can be utilized to solve (4.4.6) and thus find α^*, a value at most δ above $\sup_p \pi(p)$.

First to simplify the computation of (4.4.4), notice that the inequality $Z_c > z_c$ is in fact

$$x > \tfrac{1}{2}z_c\sqrt{n} + \tfrac{1}{2}n$$

so that the critical region C reduces to

$$C = \{(x,n): x > h(n);\ x=0(1)n,\ n=0(1)N\},$$

where $h(n) = \tfrac{1}{2}\{z_c\sqrt{n} + n\}$. The null power function (4.4.4) becomes

$$\pi(p) = \sum_{n=0}^{N} \sum_{x>h(n)} \binom{N}{n}\binom{n}{x} \tfrac{1}{2}^n p^n (1-p)^{N-n}$$

$$= \sum_{n=0}^{N} \binom{N}{n} p^n (1-p)^{N-n} \sum_{x>h(n)} \binom{n}{x} \tfrac{1}{2}^n$$

$$= \sum_{n=k}^{N} \binom{N}{n} p^n (1-p)^{N-n} [1 - F_n(i_n)]$$

where $k = \text{int}[z_c^2 + 1]$, $i_n = \text{int}[h(n)]$, int[.] is the integer function and $F_n(.)$ is the binomial cumulative distribution with parameters $(n,\tfrac{1}{2})$. Notice that, since $i_n \geq \tfrac{1}{2}n$, it is more efficient to compute

$$\sum_{x=i_n+1}^{n} \binom{n}{x} \tfrac{1}{2}^n$$

instead of $1-F_n(i_n)$. Then, by the symmetry of the binomial distribution with $(n,\tfrac{1}{2})$, the null power function of Z_c can be rewritten as

$$\pi(p) = \sum_{n=k}^{N} \binom{N}{n} p^n (1-p)^{N-n} F_n(n-i_n-1), \qquad (4.4.7)$$

in the form of (2.1.1). The derivative of the null power function is

$$\pi'(p) = \sum_{n=k}^{N} F_n(n-i_n-1) \binom{N}{n} [\, n\, p^{n-1} (1-p)^{N-n}$$

$$- (N-n)\, p^n (1-p)^{N-n-1} \,]$$

so that the methodology developed in Chapter 2 can now be utilized to find α^*, the size (of precision $\delta=.001$) of Z_c for any value z_c.

In this problem, the size was taken as $\sup\{\pi(p):0<p<.99\}$ because of the behaviour of $\pi(p)$ when p approaches 1. The null power function is dominated by the last term of the summation, namely $p^N F_N(N-i_N-1)$, when p tends to 1. From the practical point of view, the fact that p tends to 1 implies that almost no concordant pairs will be observed, so that the problem virtually reduces to a problem with no nuisance parameter, namely the sign test with no ties. It then seems reasonable to compute the size on the interval $p \in (0,.99)$.

For a test of significance of level α, the critical value z_c^* can then be obtained by equation (4.4.6). Because Z_c is asymptotically normal, the 100α upper percentile point of the standard normal distribution can be used as a starting value of z_c. This value is then incremented until $\alpha^* \le \alpha$ and z_c^* is taken as the smallest value of z_c which satisfies $\alpha^* \le \alpha$. The FORTRAN computer program used to implement this procedure is given in Appendix C.2.

For N=10(1)200 and α=.05 and .025, z_c^* , the exact critical values and α^*, the size (of precision δ=.001) of Z_c, were computed and are given in Table A.4. Furthermore, Table A.4 also contains α_1, a lower bound for $\{\pi(p): p > 5/N\}$ and α_2, a lower bound for $\{\pi(p): p > 10/N\}$ as indicators of the stability of the null power function. These lower bounds on p are obtained for expected number of discordant pairs of at least 5 and 10 respectively.

The null power function, $\pi(p)$, of the exact conditional test, as well as of the exact and approximate unconditional tests, was plotted for some values of N and a nominal level of significance of α=.05. For N=10, Figure B.7 contains the plot of $\pi(p)$ based on the normal approximation of Z_c (critical value z_c=1.645). It is apparent that using the normal approximation in this case induces a liberal test for that range of the nuisance parameter p where the null power function exceeds the nominal level α=.05, namely $.30<p<.47$ and $p>.74$. Figure B.8 is the plot based on the unconditional critical region defined by the exact conditional test, namely the sign test. Here, the test is very conservative, its actual size being approximately .013. In Figure B.9, the exact unconditional critical value z_c^*=1.90 of Table A.4 is used to plot $\pi(p)$. From these plots, the exact Z-test (Figure B.9) is seen to perform best at approaching the nominal significance level without exceeding it, although, because of the sparsity of its natural levels its size is only .0265.

The null power function of the exact unconditional test Z_c is seen to behave better for larger values of N. For N=30, Figure B.10 is the plot of $\pi(p)$ based on $z_c^*=1.74$ and Figure B.11 is the plot of $\pi(p)$ based on $z_c^*=1.68$ and N=40.

4.5 Power and Sample Sizes

Now that the exact critical values of Z_c have been computed (Table A.4), the exact power in (4.4.3) can be readily obtained for $\alpha=.025$, $\alpha=.05$, N=10(1)200 and various values of p_{01} and p_{10}. Consequently, the minimum sample size required to achieve a power of $1-\beta$ and a significance level of α for a combination of (p_{01},p_{10}) can be computed by solving the equation

$$N^* = \min \{ N: \Pi(p_{01},p_{10}) > 1-\beta \} , \qquad (4.5.1)$$

where the critical region that defines $\Pi(p_{01},p_{10})$ is based on z_c^* , a function of N.

Because all other sample size results are given in terms of the parameters $\psi=p_{01}+p_{10}$ and $\Delta=p_{10}-p_{01}$, equation (4.5.1) was solved in terms of these parameters for the purpose of comparability. For $\alpha=.05$ and $1-\beta=.80$, and various combinations of ψ and Δ, the minimum sample sizes from (4.5.1) are given in Table A.5. This table also contains the critical values z_c^* , the size (of precision $\delta=.001$) of Z_c and the attained power $1-\beta^*$. Therefore, Table A.5 is sufficient for both the design and the analysis of the 2×2 table for comparing two correlated proportions.

In Table A.6, the exact unconditional sample sizes [N^*] of Table A.5 are compared to the exact conditional sample sizes [N_e] found in Schork and Williams (1980). Furthermore, the approximate formulae derived by Miettinen (1968), namely N_{a_1} and N_{a_2} of section 4.2, are computed and also compared to N^* and N_e in Table A.6. The exact unconditional sample sizes N^* are seen to be smaller than N_e, the sample sizes based on the exact conditional test, for all except some combinations of ψ and Δ. This seems to happen for larger values of ψ and Δ. The approximate sample sizes N_{a_1} and N_{a_2} are almost equal to each other, much smaller than N_e and slightly smaller than N^*.

Because these results suggest that the exact unconditional test might be more powerful than the exact conditional test, the critical regions of each test were compared numerically. This comparision showed that, for all the cases considered, the critical region defined by the exact conditional test (sign test) is contained in the critical region defined by the exact Z-test. Therefore, the exact Z-test is uniformely more powerful than the exact conditional test for the cases considered, namely N=10(1)200, α=.025 and α=.05.

APPENDIX A

TABLES

These tables contain critical values and sample size determinations for the problems of comparing two independent proportions and of comparing two correlated proportions. For one-sided tests, the tables of critical values are produced for significance levels α=.05 and .025, and the sample size tables for a level of α=.05 and 80% power. The legend for these tables is given below.

<u>Legend for Tables A.1, A.2 and A.3: two independent proportions</u>

- n = sample size in each group
- α = nominal significance level
- α_1 = lower bound for $\{\pi(p): .05<p<.95\}$
- α_2 = lower bound for $\{\pi(p): .10<p<.90\}$
- $\pi(p)$ = null power function
- z_u^* = exact one-sided critical values of Z_u, the Z-test with unpooled variance estimator
- z_p^* = exact one-sided critical values of Z_p, the Z-test with pooled variance estimator
- p_i = probability of success in group i, i=1,2.

n^* = sample size determined by exact Z-tests

α^* = size of Z-tests for n and z_u^* or z_p^*

$1-\beta^*$ = attained power for n^*, z_u^* or z_p^*, p_1 and p_2

The following are the sample sizes of Table A.2 as determined by:

n_e = Fisher's exact test, Casagrande et al. (1978b)

n_c = the corrected χ^2 approximation (3.2.6)

n_r = the recorrected χ^2 approximation (3.2.7)

n_p = the uncorrected χ^2 approximation (3.2.4)

n_{as} = the arcsine formula (3.2.5)

n^* = the exact Z-tests.

Legend for Tables A.4, A.5 and A.6: two correlated proportions

N = sample size (number of matched pairs)

α = nominal significance level

α_1 = lower bound for $\{\pi(p): p > 5/N\}$

α_2 = lower bound for $\{\pi(p): p > 10/N\}$

$\pi(p)$ = null power function

z_c^* = exact one-sided critical values of Z_c, the Z-test

Δ = $p_{10}-p_{01}$

ψ = $p_{10}+p_{01}$

p_{01} = P(R=0,S=1) (see section 4.1)

p_{10} = P(R=1,S=0) (see section 4.1)

N^* = sample size determined by the exact Z-test

α^* = size of Z-test for N and z_c^*

$1-\beta^*$ = attained power for N^*, z_c^*, Δ and ψ

The following are the sample sizes of Table A.5 as determined by:

N_{a_1} = McNemar's test, first approximation (4.2.2)

N_{a_2} = McNemar's test, second approximation (4.2.3)

N_e = the exact conditional test (sign test), Schork and Williams (1980)

N^* = the exact Z-test.

Table A.1 Critical Values and Sizes of Z-tests for Comparing Two Independent Proportions.

	$\alpha = .05$					$\alpha = .025$				
n	α_1	α_2	α^*	z_u^*	z_p^*	α_1	α_2	α^*	z_u^*	z_p^*
10	.0068	.0251	.0476	1.96	1.80	.0005	.0044	.0212	2.17	1.96
11	.0086	.0292	.0504	1.84	1.72	.0007	.0058	.0208	2.40	2.14
12	.0105	.0329	.0471	1.86	1.74	.0010	.0073	.0225	2.26	2.06
13	.0125	.0363	.0484	1.81	1.71	.0014	.0089	.0200	2.26	2.07
14	.0146	.0394	.0495	1.77	1.68	.0018	.0106	.0209	2.19	2.03
15	.0168	.0422	.0505	1.74	1.66	.0024	.0122	.0218	2.14	2.00
16	.0188	.0289	.0421	1.92	1.82	.0030	.0137	.0252	2.10	1.97
17	.0209	.0308	.0426	1.90	1.81	.0036	.0152	.0233	2.21	2.07
18	.0230	.0335	.0430	1.88	1.80	.0042	.0166	.0241	2.14	2.02
19	.0251	.0368	.0435	1.86	1.78	.0049	.0174	.0246	2.14	2.03
20	.0271	.0339	.0438	1.85	1.78	.0057	.0192	.0252	2.10	2.00
21	.0290	.0351	.0445	1.84	1.77	.0064	.0183	.0251	2.17	2.06
22	.0308	.0360	.0484	1.83	1.77	.0071	.0175	.0245	2.17	2.07
23	.0325	.0367	.0447	1.84	1.78	.0078	.0179	.0239	2.17	2.07
24	.0341	.0373	.0458	1.81	1.76	.0086	.0178	.0222	2.15	2.06
25	.0342	.0342	.0448	1.80	1.75	.0094	.0184	.0233	2.13	2.04
26	.0361	.0361	.0449	1.79	1.74	.0102	.0183	.0217	2.12	2.04
27	.0369	.0369	.0450	1.79	1.74	.0109	.0191	.0225	2.11	2.03
28	.0386	.0386	.0458	1.78	1.74	.0117	.0194	.0233	2.10	2.03
29	.0393	.0393	.0458	1.78	1.74	.0124	.0195	.0242	2.09	2.02
30	.0395	.0395	.0467	1.77	1.73	.0034	.0128	.0219	2.15	2.08
31	.0398	.0398	.0494	1.77	1.73	.0138	.0198	.0243	2.11	2.04
32	.0387	.0387	.0462	1.80	1.76	.0145	.0187	.0236	2.09	2.03
33	.0395	.0395	.0476	1.77	1.73	.0151	.0193	.0245	2.07	2.01
34	.0404	.0405	.0493	1.75	1.72	.0157	.0193	.0238	2.06	2.00
35	.0376	.0376	.0476	1.75	1.72	.0163	.0204	.0243	2.06	2.00
36	.0380	.0380	.0467	1.75	1.72	.0168	.0206	.0245	2.06	2.00
37	.0398	.0398	.0467	1.74	1.71	.0173	.0208	.0249	2.05	2.00
38	.0408	.0414	.0476	1.74	1.71	.0063	.0141	.0225	2.13	2.07
39	.0409	.0410	.0469	1.74	1.71	.0169	.0169	.0233	2.10	2.05
40	.0410	.0419	.0470	1.73	1.70	.0186	.0203	.0242	2.08	2.03
41	.0411	.0422	.0491	1.73	1.70	.0186	.0186	.0237	2.07	2.02
42	.0406	.0406	.0481	1.78	1.75	.0193	.0193	.0244	2.05	2.00
43	.0411	.0413	.0487	1.76	1.73	.0196	.0197	.0241	2.05	2.00
44	.0412	.0420	.0498	1.74	1.71	.0199	.0208	.0243	2.04	2.00
45	.0396	.0396	.0492	1.73	1.71	.0202	.0217	.0249	2.03	1.99
46	.0383	.0383	.0482	1.73	1.71	.0204	.0215	.0247	2.03	1.99
47	.0395	.0395	.0492	1.72	1.70	.0176	.0177	.0251	2.06	2.02
48	.0404	.0404	.0478	1.72	1.70	.0097	.0174	.0233	2.09	2.05
49	.0417	.0417	.0486	1.72	1.70	.0176	.0180	.0241	2.07	2.03

Table A.1 -- continued

n	α = .05 α_1	α_2	α^*	z_u^*	z_p^*	α = .025 α_1	α_2	α^*	z_u^*	z_p^*
50	.0419	.0429	.0494	1.71	1.69	.0177	.0181	.0248	2.05	2.01
51	.0420	.0426	.0487	1.72	1.70	.0177	.0182	.0245	2.04	2.00
52	.0435	.0435	.0489	1.71	1.69	.0178	.0184	.0242	2.04	2.00
53	.0440	.0444	.0499	1.71	1.69	.0178	.0185	.0247	2.04	2.01
54	.0394	.0394	.0481	1.76	1.74	.0178	.0187	.0236	2.04	2.01
55	.0396	.0396	.0482	1.75	1.73	.0178	.0189	.0249	2.03	2.00
56	.0413	.0413	.0498	1.73	1.71	.0178	.0191	.0246	2.03	2.00
57	.0398	.0398	.0495	1.72	1.70	.0178	.0192	.0248	2.03	2.00
58	.0389	.0389	.0494	1.72	1.70	.0126	.0180	.0222	2.09	2.06
59	.0398	.0398	.0499	1.71	1.69	.0178	.0196	.0236	2.07	2.04
60	.0405	.0405	.0500	1.71	1.69	.0178	.0207	.0250	2.05	2.02
61	.0416	.0416	.0501	1.70	1.68	.0178	.0198	.0249	2.04	2.01
62	.0409	.0409	.0502	1.72	1.70	.0178	.0194	.0247	2.03	2.00
63	.0428	.0428	.0503	1.70	1.68	.0178	.0197	.0246	2.03	2.00
64	.0406	.0406	.0485	1.72	1.70	.0178	.0204	.0249	2.02	1.99
65	.0406	.0406	.0486	1.72	1.70	.0178	.0208	.0249	2.03	2.00
66	.0407	.0407	.0491	1.72	1.70	.0178	.0207	.0239	2.04	2.01
67	.0327	.0378	.0473	1.75	1.73	.0179	.0211	.0240	2.03	2.00
68	.0406	.0406	.0483	1.74	1.72	.0179	.0212	.0247	2.03	2.00
69	.0383	.0383	.0483	1.73	1.72	.0147	.0175	.0231	2.08	2.05
70	.0387	.0387	.0487	1.72	1.71	.0152	.0195	.0243	2.06	2.03
71	.0392	.0392	.0489	1.71	1.70	.0179	.0196	.0242	2.05	2.02
72	.0393	.0393	.0486	1.71	1.70	.0197	.0199	.0247	2.04	2.01
73	.0400	.0400	.0485	1.71	1.70	.0193	.0193	.0248	2.03	2.01
74	.0409	.0409	.0486	1.71	1.70	.0194	.0194	.0249	2.03	2.01
75	.0412	.0412	.0486	1.71	1.70	.0198	.0200	.0251	2.02	2.00
76	.0412	.0417	.0487	1.70	1.69	.0198	.0204	.0250	2.02	2.00
77	.0413	.0418	.0488	1.70	1.69	.0198	.0210	.0251	2.02	2.00
78	.0414	.0419	.0489	1.70	1.69	.0198	.0211	.0252	2.02	2.00
79	.0415	.0420	.0489	1.70	1.69	.0198	.0210	.0238	2.03	2.01
80	.0415	.0437	.0490	1.70	1.69	.0198	.0213	.0246	2.03	2.01
81	.0430	.0430	.0496	1.70	1.69	.0161	.0186	.0243	2.07	2.05
82	.0327	.0378	.0480	1.74	1.73	.0162	.0187	.0240	2.06	2.04
83	.0327	.0378	.0480	1.73	1.72	.0160	.0176	.0240	2.05	2.03
84	.0396	.0396	.0491	1.72	1.71	.0194	.0194	.0244	2.04	2.02
85	.0398	.0398	.0491	1.71	1.70	.0196	.0196	.0249	2.03	2.01
86	.0393	.0393	.0491	1.71	1.70	.0191	.0191	.0242	2.03	2.01
87	.0399	.0399	.0496	1.70	1.69	.0195	.0195	.0246	2.03	2.01
88	.0403	.0403	.0492	1.70	1.69	.0200	.0200	.0243	2.03	2.01
89	.0410	.0410	.0494	1.69	1.68	.0200	.0205	.0248	2.02	2.00

Table A.1 -- continued

n	$\alpha = .05$					$\alpha = .025$				
	α_1	α_2	α^*	z_u^*	z_p^*	α_1	α_2	α^*	z_u^*	z_p^*
90	.0418	.0418	.0504	1.69	1.68	.0200	.0207	.0249	2.02	2.00
91	.0423	.0423	.0496	1.69	1.68	.0201	.0211	.0250	2.01	1.99
92	.0425	.0425	.0494	1.69	1.68	.0201	.0208	.0242	2.02	2.00
93	.0431	.0431	.0495	1.69	1.68	.0202	.0209	.0245	2.02	2.00
94	.0436	.0436	.0500	1.68	1.67	.0170	.0184	.0228	2.07	2.05
95	.0442	.0442	.0503	1.68	1.67	.0171	.0185	.0232	2.06	2.04
96	.0440	.0440	.0499	1.68	1.67	.0171	.0186	.0236	2.05	2.03
97	.0445	.0445	.0501	1.68	1.67	.0153	.0185	.0244	2.04	2.02
98	.0450	.0450	.0504	1.68	1.67	.0191	.0191	.0249	2.03	2.01
99	.0329	.0389	.0486	1.72	1.71	.0153	.0186	.0244	2.03	2.01
100	.0382	.0391	.0504	1.71	1.70	.0191	.0191	.0247	2.02	2.00
101	.0382	.0392	.0491	1.71	1.70	.0192	.0192	.0246	2.02	2.00
102	.0382	.0393	.0503	1.70	1.69	.0192	.0192	.0244	2.02	2.00
103	.0382	.0393	.0496	1.70	1.69	.0192	.0192	.0247	2.02	2.00
104	.0382	.0394	.0493	1.70	1.69	.0193	.0193	.0250	2.01	1.99
105	.0407	.0407	.0495	1.69	1.68	.0193	.0193	.0249	2.01	1.99
106	.0413	.0413	.0495	1.69	1.68	.0211	.0211	.0250	2.01	2.00
107	.0419	.0419	.0495	1.69	1.68	.0212	.0212	.0245	2.01	2.00
108	.0421	.0421	.0495	1.69	1.68	.0176	.0176	.0225	2.07	2.05
109	.0422	.0422	.0495	1.69	1.68	.0183	.0183	.0232	2.06	2.04
110	.0427	.0427	.0495	1.69	1.68	.0184	.0184	.0236	2.05	2.03
111	.0426	.0426	.0496	1.69	1.68	.0185	.0192	.0243	2.04	2.03
112	.0437	.0437	.0497	1.68	1.67	.0182	.0193	.0247	2.03	2.02
113	.0435	.0435	.0495	1.68	1.67	.0182	.0189	.0243	2.03	2.02
114	.0439	.0439	.0496	1.68	1.67	.0182	.0194	.0249	2.02	2.01
115	.0443	.0443	.0499	1.68	1.67	.0182	.0196	.0248	2.02	2.01
116	.0391	.0400	.0484	1.72	1.71	.0199	.0199	.0251	2.01	2.00
117	.0380	.0380	.0484	1.72	1.71	.0203	.0203	.0250	2.01	2.00
118	.0392	.0398	.0486	1.71	1.70	.0205	.0205	.0248	2.01	2.00
119	.0402	.0402	.0500	1.70	1.69	.0206	.0206	.0245	2.01	2.00
120	.0396	.0396	.0494	1.70	1.69	.0209	.0209	.0247	2.01	2.00
121	.0405	.0405	.0520	1.69	1.68	.0209	.0209	.0247	2.01	2.00
122	.0405	.0405	.0502	1.69	1.68	.0212	.0212	.0247	2.01	2.00
123	.0407	.0407	.0497	1.69	1.68	.0215	.0215	.0250	2.00	1.99
124	.0411	.0411	.0497	1.69	1.68	.0213	.0213	.0248	2.01	2.00
125	.0418	.0418	.0502	1.68	1.67	.0181	.0181	.0232	2.05	2.04
126	.0419	.0419	.0496	1.69	1.68	.0187	.0187	.0239	2.04	2.03
127	.0426	.0426	.0500	1.68	1.67	.0185	.0193	.0243	2.03	2.02
128	.0427	.0428	.0497	1.68	1.67	.0186	.0200	.0250	2.02	2.01
129	.0424	.0424	.0497	1.69	1.68	.0186	.0193	.0249	2.02	2.01

Table A.1 -- continued

n	$\alpha = .05$					$\alpha = .025$				
	α_1	α_2	α^*	z_u^*	z_p^*	α_1	α_2	α^*	z_u^*	z_p^*
130	.0427	.0432	.0497	1.68	1.68	.0187	.0193	.0243	2.02	2.01
131	.0427	.0431	.0497	1.68	1.68	.0187	.0197	.0251	2.01	2.00
132	.0427	.0435	.0498	1.68	1.68	.0187	.0198	.0242	2.02	2.01
133	.0427	.0438	.0499	1.68	1.68	.0188	.0199	.0241	2.02	2.01
134	.0428	.0436	.0500	1.68	1.68	.0188	.0204	.0246	2.01	2.00
135	.0383	.0383	.0486	1.72	1.71	.0189	.0205	.0244	2.01	2.00
136	.0402	.0402	.0486	1.71	1.70	.0189	.0208	.0246	2.01	2.00
137	.0388	.0388	.0485	1.71	1.70	.0189	.0208	.0244	2.01	2.00
138	.0403	.0403	.0499	1.70	1.70	.0190	.0207	.0245	2.01	2.00
139	.0394	.0394	.0500	1.70	1.70	.0190	.0210	.0244	2.01	2.00
140	.0395	.0395	.0499	1.70	1.70	.0191	.0208	.0245	2.01	2.00
141	.0402	.0402	.0500	1.69	1.69	.0191	.0210	.0249	2.01	2.00
142	.0406	.0406	.0500	1.69	1.69	.0185	.0185	.0236	2.04	2.03
143	.0409	.0409	.0501	1.69	1.69	.0191	.0191	.0243	2.03	2.02
144	.0413	.0413	.0500	1.69	1.69	.0192	.0203	.0251	2.02	2.01
145	.0415	.0415	.0500	1.69	1.69	.0191	.0191	.0246	2.02	2.01
146	.0422	.0422	.0501	1.68	1.68	.0192	.0192	.0247	2.02	2.01
147	.0423	.0423	.0498	1.68	1.68	.0194	.0194	.0242	2.02	2.01
148	.0427	.0427	.0500	1.68	1.68	.0194	.0197	.0246	2.01	2.00
149	.0427	.0427	.0502	1.68	1.68	.0195	.0199	.0246	2.01	2.00
150	.0427	.0427	.0501	1.68	1.68	.0195	.0203	.0252	2.00	1.99

Table A.2 Minimum Sample Sizes to Achieve 80% Power and $\alpha \leq .05$ for One-sided Z-tests for Comparing Two Independent Proportions.

p_1	p_2	n^*	z_u^*	z_p^*	α^*	$1-\beta^*$
.05	.15	107	1.69	1.68	.0495	.8009
	.20	56	1.73	1.71	.0498	.8016
	.25	38	1.74	1.71	.0476	.8098
	.30	28	1.78	1.74	.0458	.8095
	.35	22	1.83	1.77	.0484	.8095
	.40	18	1.88	1.80	.0430	.8190
	.45	13	1.81	1.71	.0484	.8142
.10	.25	79	1.70	1.69	.0489	.8026
	.30	49	1.72	1.70	.0486	.8071
	.35	35	1.75	1.72	.0476	.8063
	.40	26	1.79	1.74	.0449	.8088
	.45	21	1.84	1.77	.0445	.8057
	.50	17	1.90	1.81	.0426	.8213
	.55	13	1.81	1.71	.0484	.8016
	.60	10	1.96	1.80	.0476	.8016
.15	.30	95	1.68	1.67	.0503	.8023
	.35	59	1.71	1.69	.0499	.8115
	.40	40	1.73	1.70	.0470	.8077
	.45	29	1.78	1.74	.0458	.8001
	.50	23	1.84	1.78	.0447	.8134
	.55	18	1.88	1.80	.0430	.8033
	.60	14	1.77	1.68	.0495	.8016
	.65	13	1.81	1.71	.0484	.8366
.20	.35	111	1.69	1.68	.0496	.8005
	.40	68	1.74	1.72	.0483	.8038
	.45	44	1.74	1.71	.0498	.8017
	.50	32	1.80	1.76	.0462	.8060
	.55	26	1.79	1.74	.0449	.8298
	.60	20	1.85	1.78	.0438	.8153
	.65	15	1.74	1.66	.0505	.8273
	.70	13	1.81	1.71	.0484	.8239
.25	.40	123	1.69	1.68	.0497	.8017
	.45	71	1.71	1.70	.0489	.8017
	.50	48	1.72	1.70	.0478	.8061
	.55	33	1.77	1.73	.0476	.8051
	.60	26	1.79	1.74	.0449	.8006
	.65	20	1.85	1.78	.0438	.8091
	.70	15	1.74	1.66	.0505	.8232
	.75	13	1.81	1.71	.0484	.8205

Table A.2 -- continued

p_1	p_2	n^*	z_u^*	z_p^*	α^*	$1-\beta^*$
.30	.45	132	1.68	1.68	.0498	.8015
	.50	77	1.70	1.69	.0488	.8021
	.55	50	1.71	1.69	.0494	.8029
	.60	37	1.74	1.71	.0467	.8145
	.65	27	1.79	1.74	.0450	.8064
	.70	20	1.85	1.78	.0438	.8075
.35	.50	136	1.71	1.70	.0486	.8016
	.55	79	1.70	1.69	.0489	.8070
	.60	51	1.72	1.70	.0487	.8085
	.65	37	1.74	1.71	.0467	.8111
.40	.55	144	1.69	1.69	.0500	.8021
	.60	79	1.70	1.69	.0489	.8057

Table A.3 Comparision of Sample Sizes to Achieve 80% Power and $\alpha \leq .05$ for One-sided Tests for Comparing Two Independent Proportions.

p_1	p_2	n_e	n_c	n_r	n_p	n_{as}	n^*
.05	.15	126	148	130	111	105	107
	.20	67	84	72	59	55	56
	.25	45	57	48	39	35	38
	.30	34	42	36	28	25	28
	.35	25	33	28	21	19	22
	.40	20	27	22	17	15	18
	.45	17	23	19	14	12	13
.10	.25	89	104	92	79	76	79
	.30	56	67	58	49	47	49
	.35	39	49	42	34	32	35
	.40	30	37	31	25	24	26
	.45	24	30	25	20	19	21
	.50	19	25	20	16	15	17
	.55	16	21	17	13	12	13
	.60	13	18	14	11	10	10
.15	.30	106	120	108	95	94	95
	.35	65	76	67	57	56	59
	.40	46	54	46	39	38	40
	.45	34	40	35	28	28	29
	.50	26	32	27	22	21	23
	.55	22	26	22	17	17	18
	.60	17	22	18	14	13	14
	.65	15	18	15	11	11	13
.20	.35	121	134	122	109	108	111
	.40	73	83	74	64	64	68
	.45	49	58	50	43	42	44
	.50	36	43	37	31	30	32
	.55	27	34	28	23	23	26
	.60	23	27	23	18	18	20
	.65	17	22	18	14	14	15
	.70	15	19	15	12	12	13
.25	.40	132	145	133	120	119	123
	.45	78	89	79	70	69	71
	.50	54	61	53	46	46	48
	.55	37	45	39	32	32	33
	.60	30	35	30	24	24	26
	.65	23	28	23	19	19	20
	.70	18	23	19	15	15	15
	.75	15	19	15	12	12	13

Table A.3 -- continued

p_1	p_2	n_e	n_c	n_r	n_p	n_{as}	n^*
.30	.45	142	154	141	128	128	132
	.50	84	92	83	74	73	77
	.55	55	63	55	48	48	50
	.60	41	46	40	33	33	37
	.65	31	35	30	25	25	27
	.70	23	28	23	19	19	20
.35	.50	143	159	147	134	134	136
	.55	85	95	86	76	76	79
	.60	56	64	56	49	49	51
	.65	41	46	40	34	34	37
.40	.55	144	162	149	136	136	144
	.60	85	96	86	77	77	79

Table A.4 Critical Values and Sizes of Z-test for Comparing Two Correlated Proportions.

	$\alpha = .05$				$\alpha = .025$			
N	α_1	α_2	α^*	z_c^*	α_1	α_2	α^*	z_c^*
10	.0114		.0265	1.90	.0114		.0208	2.01
11	.0322	.0345	.0395	1.74	.0062	.0062	.0197	2.12
12	.0205	.0205	.0373	1.74	.0133	.0204	.0233	2.01
13	.0316	.0328	.0430	1.74	.0120	.0120	.0213	2.01
14	.0303	.0303	.0371	1.74	.0069	.0069	.0126	2.14
15	.0187	.0187	.0370	1.81	.0127	.0186	.0211	2.01
16	.0312	.0349	.0454	1.74	.0114	.0114	.0209	2.01
17	.0261	.0261	.0413	1.74	.0125	.0175	.0225	2.01
18	.0313	.0353	.0446	1.74	.0128	.0163	.0208	2.01
19	.0307	.0333	.0399	1.74	.0103	.0103	.0208	2.07
20	.0220	.0220	.0395	1.79	.0126	.0187	.0246	2.01
21	.0310	.0356	.0459	1.74	.0142	.0142	.0251	1.97
22	.0278	.0278	.0425	1.74	.0200	.0200	.0249	1.97
23	.0309	.0357	.0427	1.74	.0182	.0182	.0247	1.97
24	.0307	.0332	.0412	1.74	.0121	.0121	.0213	2.05
25	.0305	.0357	.0494	1.74	.0124	.0166	.0211	2.05
26	.0301	.0357	.0434	1.74	.0154	.0154	.0246	1.97
27	.0275	.0275	.0413	1.74	.0204	.0204	.0246	1.97
28	.0307	.0349	.0407	1.74	.0185	.0185	.0246	1.97
29	.0302	.0315	.0407	1.74	.0120	.0128	.0207	2.05
30	.0303	.0357	.0450	1.74	.0126	.0166	.0207	2.05
31	.0303	.0357	.0407	1.74	.0155	.0155	.0246	1.98
32	.0262	.0262	.0407	1.77	.0196	.0196	.0246	1.98
33	.0301	.0357	.0458	1.74	.0179	.0179	.0246	1.98
34	.0299	.0299	.0435	1.74	.0125	.0128	.0207	2.06
35	.0298	.0356	.0429	1.74	.0208	.0208	.0249	1.98
36	.0304	.0327	.0425	1.74	.0123	.0151	.0226	2.01
37	.0302	.0356	.0449	1.74	.0127	.0182	.0226	2.01
38	.0299	.0355	.0419	1.74	.0119	.0168	.0226	2.01
39	.0400	.0411	.0501	1.68	.0123	.0192	.0242	2.01
40	.0399	.0400	.0500	1.68	.0192	.0192	.0250	1.98
41	.0305	.0305	.0500	1.72	.0120	.0143	.0226	2.04
42	.0303	.0354	.0433	1.74	.0208	.0208	.0250	1.98
43	.0298	.0323	.0428	1.74	.0163	.0163	.0247	1.99
44	.0375	.0375	.0500	1.68	.0186	.0186	.0247	1.99
45	.0352	.0352	.0501	1.68	.0175	.0175	.0248	1.99
46	.0395	.0395	.0501	1.68	.0198	.0198	.0248	1.99
47	.0382	.0382	.0501	1.68	.0202	.0202	.0248	1.99
48	.0290	.0300	.0427	1.74	.0115	.0150	.0229	2.03
49	.0396	.0408	.0501	1.70	.0118	.0167	.0228	2.03

Table A.4 -- continued

	$\alpha = .05$				$\alpha = .025$			
N	α_1	α_2	α^*	z_c^*	α_1	α_2	α^*	z_c^*
50	.0325	.0325	.0500	1.70	.0167	.0167	.0248	1.99
51	.0352	.0352	.0500	1.70	.0183	.0183	.0247	1.99
52	.0333	.0333	.0500	1.70	.0174	.0174	.0247	1.99
53	.0365	.0365	.0500	1.70	.0194	.0194	.0246	1.99
54	.0359	.0359	.0501	1.70	.0200	.0200	.0246	1.99
55	.0387	.0387	.0501	1.70	.0120	.0151	.0230	2.03
56	.0299	.0360	.0453	1.74	.0123	.0163	.0227	2.03
57	.0309	.0309	.0501	1.73	.0187	.0187	.0248	1.99
58	.0390	.0407	.0501	1.68	.0177	.0177	.0248	1.99
59	.0311	.0311	.0501	1.73	.0168	.0168	.0247	1.99
60	.0347	.0347	.0501	1.70	.0186	.0186	.0247	1.99
61	.0329	.0329	.0501	1.70	.0190	.0190	.0246	1.99
62	.0294	.0359	.0445	1.74	.0197	.0197	.0250	1.99
63	.0298	.0359	.0427	1.74	.0200	.0200	.0246	1.99
64	.0298	.0358	.0470	1.74	.0125	.0185	.0234	2.01
65	.0294	.0358	.0453	1.74	.0116	.0189	.0231	2.01
66	.0297	.0356	.0442	1.74	.0119	.0186	.0230	2.01
67	.0382	.0395	.0498	1.70	.0201	.0201	.0248	1.99
68	.0378	.0378	.0501	1.70	.0200	.0200	.0247	1.99
69	.0334	.0334	.0499	1.70	.0200	.0200	.0247	1.99
70	.0381	.0398	.0499	1.68	.0199	.0200	.0247	1.99
71	.0346	.0346	.0495	1.70	.0119	.0180	.0234	2.02
72	.0378	.0393	.0501	1.68	.0199	.0199	.0250	1.99
73	.0377	.0401	.0500	1.68	.0194	.0194	.0247	1.99
74	.0374	.0374	.0501	1.68	.0198	.0199	.0247	1.99
75	.0376	.0384	.0500	1.68	.0191	.0191	.0247	1.99
76	.0375	.0391	.0501	1.68	.0197	.0198	.0247	1.99
77	.0374	.0395	.0501	1.67	.0181	.0181	.0247	1.99
78	.0373	.0396	.0501	1.67	.0186	.0186	.0247	1.99
79	.0371	.0393	.0500	1.69	.0196	.0197	.0251	1.98
80	.0370	.0397	.0499	1.69	.0112	.0185	.0234	2.02
81	.0369	.0388	.0501	1.67	.0195	.0195	.0250	1.99
82	.0369	.0388	.0500	1.67	.0194	.0195	.0247	1.99
83	.0368	.0389	.0501	1.67	.0194	.0195	.0247	1.99
84	.0367	.0390	.0500	1.67	.0192	.0192	.0247	1.99
85	.0365	.0390	.0501	1.67	.0193	.0194	.0247	1.99
86	.0364	.0391	.0500	1.67	.0192	.0193	.0247	1.99
87	.0362	.0380	.0500	1.67	.0192	.0192	.0247	1.99
88	.0361	.0381	.0500	1.67	.0191	.0192	.0251	1.98
89	.0360	.0381	.0500	1.67	.0190	.0192	.0251	1.98

Table A.4 -- continued

N	α = .05 α_1	α_2	α^*	z_c^*	α = .025 α_1	α_2	α^*	z_c^*
90	.0359	.0382	.0500	1.69	.0190	.0191	.0248	1.99
91	.0358	.0382	.0499	1.69	.0189	.0191	.0247	2.00
92	.0357	.0383	.0500	1.67	.0189	.0190	.0247	2.00
93	.0356	.0383	.0500	1.67	.0189	.0189	.0247	2.00
94	.0354	.0384	.0499	1.67	.0188	.0190	.0246	2.00
95	.0353	.0385	.0499	1.67	.0188	.0189	.0247	2.00
96	.0351	.0371	.0499	1.67	.0188	.0188	.0251	1.99
97	.0350	.0372	.0500	1.67	.0187	.0187	.0250	1.98
98	.0348	.0372	.0499	1.67	.0187	.0187	.0249	1.98
99	.0347	.0372	.0500	1.67	.0187	.0187	.0248	1.98
100	.0346	.0373	.0500	1.67	.0117	.0194	.0236	2.01
101	.0344	.0371	.0500	1.70	.0118	.0190	.0236	2.01
102	.0343	.0371	.0500	1.70	.0185	.0185	.0248	1.99
103	.0342	.0410	.0499	1.68	.0199	.0199	.0248	1.99
104	.0341	.0396	.0500	1.68	.0192	.0192	.0247	1.99
105	.0341	.0383	.0500	1.68	.0198	.0198	.0247	1.99
106	.0340	.0394	.0500	1.68	.0199	.0199	.0247	1.99
107	.0339	.0396	.0500	1.67	.0199	.0199	.0247	1.99
108	.0338	.0409	.0499	1.67	.0198	.0198	.0246	1.99
109	.0338	.0410	.0499	1.67	.0198	.0198	.0250	1.99
110	.0337	.0411	.0499	1.67	.0198	.0198	.0249	1.99
111	.0336	.0412	.0499	1.67	.0203	.0203	.0246	2.00
112	.0395	.0413	.0499	1.67	.0190	.0190	.0246	2.00
113	.0395	.0413	.0499	1.67	.0203	.0203	.0251	1.99
114	.0395	.0399	.0499	1.69	.0202	.0202	.0248	1.99
115	.0382	.0382	.0499	1.69	.0192	.0192	.0248	1.99
116	.0394	.0404	.0499	1.68	.0197	.0197	.0251	1.99
117	.0394	.0409	.0499	1.68	.0203	.0203	.0247	1.99
118	.0380	.0380	.0499	1.68	.0205	.0205	.0247	1.99
119	.0388	.0388	.0499	1.68	.0205	.0205	.0251	1.98
120	.0390	.0390	.0499	1.68	.0205	.0205	.0251	1.98
121	.0392	.0403	.0499	1.68	.0205	.0205	.0251	1.99
122	.0391	.0406	.0499	1.68	.0204	.0205	.0245	2.00
123	.0391	.0406	.0500	1.67	.0204	.0204	.0245	2.00
124	.0391	.0407	.0499	1.67	.0182	.0182	.0245	2.00
125	.0390	.0407	.0499	1.67	.0204	.0204	.0250	1.98
126	.0390	.0408	.0499	1.67	.0204	.0204	.0250	1.98
127	.0390	.0409	.0498	1.69	.0198	.0198	.0251	1.98
128	.0389	.0409	.0499	1.69	.0199	.0199	.0251	1.98
129	.0389	.0410	.0498	1.68	.0203	.0203	.0251	1.98

Table A.4 -- continued

	$\alpha = .05$				$\alpha = .025$			
N	α_1	α_2	α^*	z_c^*	α_1	α_2	α^*	z_c^*
130	.0388	.0403	.0499	1.68	.0203	.0203	.0251	1.98
131	.0388	.0403	.0498	1.68	.0203	.0203	.0251	1.98
132	.0387	.0404	.0498	1.68	.0202	.0203	.0251	1.98
133	.0387	.0389	.0499	1.68	.0202	.0203	.0245	2.00
134	.0386	.0397	.0499	1.68	.0202	.0202	.0245	2.00
135	.0386	.0397	.0500	1.67	.0202	.0202	.0245	2.00
136	.0386	.0398	.0499	1.67	.0202	.0202	.0250	1.98
137	.0385	.0398	.0499	1.67	.0202	.0202	.0250	1.98
138	.0385	.0399	.0499	1.67	.0201	.0202	.0250	1.98
139	.0385	.0399	.0499	1.67	.0201	.0201	.0250	1.98
140	.0384	.0400	.0499	1.67	.0199	.0199	.0250	1.98
141	.0383	.0400	.0498	1.67	.0201	.0201	.0250	1.98
142	.0383	.0401	.0497	1.68	.0200	.0201	.0251	1.98
143	.0382	.0401	.0497	1.68	.0200	.0201	.0251	1.98
144	.0382	.0402	.0499	1.67	.0094	.0197	.0241	2.01
145	.0381	.0402	.0499	1.67	.0095	.0197	.0237	2.01
146	.0381	.0403	.0499	1.67	.0200	.0200	.0249	1.99
147	.0380	.0403	.0498	1.67	.0199	.0200	.0246	1.99
148	.0380	.0404	.0498	1.67	.0199	.0200	.0246	1.99
149	.0379	.0395	.0498	1.67	.0199	.0199	.0246	1.99
150	.0379	.0395	.0498	1.67	.0199	.0199	.0246	1.99
151	.0379	.0395	.0498	1.67	.0198	.0199	.0250	1.98
152	.0379	.0396	.0499	1.67	.0198	.0199	.0250	1.98
153	.0378	.0396	.0499	1.67	.0198	.0199	.0250	1.98
154	.0377	.0387	.0498	1.67	.0198	.0198	.0250	1.98
155	.0376	.0387	.0498	1.67	.0198	.0198	.0250	1.98
156	.0376	.0388	.0499	1.67	.0197	.0197	.0250	1.98
157	.0375	.0388	.0497	1.68	.0197	.0197	.0246	2.00
158	.0374	.0388	.0496	1.68	.0197	.0197	.0246	2.00
159	.0374	.0389	.0500	1.67	.0197	.0197	.0246	2.00
160	.0373	.0389	.0498	1.67	.0196	.0197	.0249	1.99
161	.0373	.0389	.0498	1.67	.0196	.0197	.0247	1.99
162	.0372	.0390	.0498	1.67	.0196	.0197	.0246	1.99
163	.0372	.0390	.0498	1.67	.0196	.0196	.0246	1.99
164	.0372	.0390	.0498	1.67	.0195	.0196	.0246	1.99
165	.0371	.0390	.0498	1.67	.0195	.0196	.0246	1.99
166	.0371	.0391	.0498	1.67	.0195	.0196	.0246	1.99
167	.0370	.0391	.0498	1.67	.0195	.0196	.0246	1.99
168	.0370	.0391	.0498	1.67	.0194	.0195	.0246	1.99
169	.0369	.0392	.0498	1.67	.0194	.0195	.0247	1.99

Table A.4 -- continued

N	$\alpha = .05$				$\alpha = .025$			
	α_1	α_2	α^*	z_c^*	α_1	α_2	α^*	z_c^*
170	.0368	.0392	.0498	1.67	.0194	.0195	.0251	1.98
171	.0367	.0393	.0498	1.67	.0194	.0195	.0245	2.00
172	.0367	.0393	.0498	1.68	.0193	.0195	.0245	2.00
173	.0366	.0393	.0498	1.68	.0193	.0195	.0251	1.99
174	.0365	.0382	.0498	1.67	.0193	.0194	.0249	1.99
175	.0365	.0382	.0498	1.67	.0193	.0193	.0247	1.99
176	.0364	.0383	.0498	1.67	.0192	.0193	.0246	1.99
177	.0363	.0383	.0498	1.67	.0192	.0193	.0246	1.99
178	.0363	.0383	.0498	1.67	.0192	.0193	.0246	1.99
179	.0362	.0383	.0498	1.67	.0191	.0193	.0246	1.99
180	.0362	.0383	.0498	1.67	.0191	.0193	.0251	1.98
181	.0361	.0384	.0498	1.67	.0191	.0192	.0251	1.98
182	.0361	.0372	.0498	1.67	.0191	.0191	.0247	1.99
183	.0360	.0372	.0498	1.67	.0191	.0191	.0250	1.99
184	.0360	.0372	.0498	1.67	.0190	.0190	.0245	2.00
185	.0360	.0372	.0498	1.67	.0190	.0190	.0250	1.99
186	.0359	.0373	.0498	1.67	.0190	.0190	.0251	1.99
187	.0359	.0373	.0499	1.67	.0190	.0190	.0249	1.99
188	.0359	.0373	.0501	1.67	.0190	.0190	.0247	1.99
189	.0358	.0373	.0497	1.68	.0189	.0189	.0246	1.99
190	.0358	.0373	.0498	1.68	.0189	.0189	.0246	1.99
191	.0358	.0373	.0498	1.68	.0189	.0189	.0246	1.99
192	.0357	.0374	.0500	1.67	.0189	.0189	.0251	1.98
193	.0357	.0374	.0498	1.67	.0189	.0189	.0250	1.98
194	.0357	.0374	.0498	1.67	.0189	.0189	.0250	1.98
195	.0356	.0374	.0498	1.67	.0188	.0188	.0250	1.98
196	.0356	.0374	.0498	1.67	.0188	.0188	.0251	1.98
197	.0355	.0374	.0498	1.67	.0188	.0188	.0250	1.98
198	.0355	.0374	.0498	1.67	.0188	.0188	.0251	1.98
199	.0354	.0374	.0498	1.67	.0188	.0188	.0247	1.99
200	.0354	.0375	.0498	1.67	.0187	.0187	.0247	1.99

Table A.5 Minimum Sample Sizes to Achieve 80% Power and $\alpha \leq .05$ for One-sided Z-test for Comparing Two Correlated Proportions.

Δ	ψ	N^*	z_c^*	α^*	$1-\beta^*$
.10	.30	185	1.67	.0498	.8006
	.22	134	1.68	.0499	.8014
	.14	80	1.69	.0499	.8038
.20	.98	153	1.67	.0499	.8003
	.94	146	1.67	.0499	.8014
	.90	139	1.67	.0499	.8014
	.86	135	1.67	.0500	.8066
	.82	129	1.68	.0498	.8018
	.78	122	1.68	.0499	.8018
	.74	116	1.68	.0499	.8042
	.70	108	1.67	.0499	.8016
	.66	103	1.68	.0499	.8026
	.62	96	1.67	.0499	.8027
	.58	89	1.67	.0500	.8013
	.54	82	1.67	.0500	.8025
	.50	77	1.67	.0501	.8070
	.46	70	1.68	.0499	.8005
	.42	67	1.70	.0498	.8219
	.38	58	1.68	.0501	.8193
	.34	51	1.70	.0500	.8058
	.30	44	1.68	.0500	.8186
	.26	38	1.74	.0419	.8081
	.22	30	1.74	.0450	.8124
.30	.88	63	1.74	.0427	.8029
	.84	58	1.68	.0501	.8081
	.80	57	1.73	.0501	.8088
	.76	52	1.70	.0500	.8033
	.72	49	1.70	.0501	.8035
	.68	46	1.68	.0501	.8029
	.64	44	1.68	.0500	.8117
	.60	39	1.68	.0501	.8020
	.56	39	1.68	.0501	.8324
	.52	35	1.74	.0429	.8003
	.48	33	1.74	.0458	.8120
	.44	30	1.74	.0450	.8123
	.40	26	1.74	.0434	.8075
	.36	22	1.74	.0425	.8076
	.32	19	1.74	.0399	.8195
.40	.98	39	1.68	.0501	.8209
	.94	38	1.74	.0419	.8089
	.90	35	1.74	.0429	.8037
	.86	34	1.74	.0435	.8046

Table A.5 -- continued

Δ	ψ	N^*	z_c^*	α^*	$1-\beta^*$
.40	.82	33	1.74	.0458	.8090
	.78	31	1.74	.0407	.8114
	.74	29	1.74	.0407	.8117
	.70	27	1.74	.0413	.8108
	.66	25	1.74	.0494	.8067
	.62	23	1.74	.0427	.8003
	.58	22	1.74	.0425	.8120
	.54	20	1.79	.0395	.8134
	.50	17	1.74	.0413	.8002
	.46	15	1.81	.0370	.8030
	.42	14	1.74	.0371	.8365
.50	.88	21	1.74	.0459	.8029
	.84	21	1.74	.0459	.8201
	.80	19	1.74	.0399	.8062
	.76	18	1.74	.0446	.8049
	.72	17	1.74	.0413	.8031
	.68	16	1.74	.0454	.8119
	.64	14	1.74	.0371	.8020
	.60	13	1.74	.0430	.8166
	.56	12	1.74	.0373	.8325
	.52	11	1.74	.0395	.8517
.60	.98	18	1.74	.0446	.8522
	.94	16	1.74	.0454	.8414
	.90	16	1.74	.0454	.8525
	.86	15	1.81	.0370	.8196
	.82	13	1.74	.0430	.8023
	.78	12	1.74	.0373	.8084
	.74	11	1.74	.0395	.8107
	.70	11	1.74	.0395	.8504
	.66	11	1.74	.0395	.8935

Table A.6 Comparision of Sample Sizes to Achieve 80% Power and $\alpha \leq .05$ for One-sided Tests for Comparing Two Correlated Proportions.

Δ	ψ	N_{a_1}	N_{a_2}	N_e	N^*
.10	.30	179	180	199	185
	.22	127	128	146	134
	.14	70	74	94	80
.20	.98	150	150	159	153
	.94	143	143	152	146
	.90	137	137	147	139
	.86	131	131	141	135
	.82	125	125	135	129
	.78	118	118	129	122
	.74	112	112	122	116
	.70	106	106	116	108
	.66	99	99	110	103
	.62	93	93	103	96
	.58	86	87	96	89
	.54	80	80	90	82
	.50	73	74	83	77
	.46	67	67	76	70
	.42	60	61	70	67
	.38	53	54	63	58
	.34	46	48	56	51
	.30	39	41	50	44
	.26	31	33	44	38
	.22	22	25	36	30
.30	.88	58	59	65	63
	.84	56	56	61	58
	.80	53	53	59	57
	.76	50	50	56	52
	.72	47	47	53	49
	.68	44	44	50	46
	.64	41	41	47	44
	.60	38	38	44	39
	.56	35	35	41	39
	.52	32	32	38	35
	.48	29	29	35	33
	.44	25	26	32	30
	.40	22	23	30	26
	.36	18	20	27	22
	.32	14	16	23	19
.40	.98	36	36	42	39
	.94	34	35	38	38
	.90	33	33	37	35
	.86	31	31	35	34

Table A.6 -- continued

Δ	ψ	N_{a_1}	N_{a_2}	N_e	N^*
.40	.82	29	30	34	33
	.78	28	28	33	31
	.74	26	26	31	29
	.70	24	25	29	27
	.66	23	23	27	25
	.62	21	21	25	23
	.58	19	19	24	22
	.54	17	18	22	20
	.50	15	16	21	17
	.46	13	14	19	15
	.42	10	11	17	14
.50	.88	20	20	23	21
	.84	19	19	22	21
	.80	17	18	21	19
	.76	16	16	20	18
	.72	15	15	19	17
	.68	14	14	18	16
	.64	13	13	17	14
	.60	11	12	16	13
	.56	10	10	14	12
	.52	8	9	13	11
.60	.98	15	15	18	18
	.94	14	14	17	16
	.90	13	13	16	16
	.86	13	13	15	15
	.82	12	12	15	13
	.78	11	11	14	12
	.74	10	10	13	11
	.70	9	9	12	11
	.66	8	8	11	11

APPENDIX B

PLOTS OF THE NULL POWER FUNCTION

In this appendix, plots of $\pi(p)$, the null power function, are given for the two problems considered here. The dotted line represents the nominal significance level on which $\pi(p)$ is based. These plots are referred to in section 3.4 for the two independent proportions case, and in section 4.4 for the two correlated proportions case.

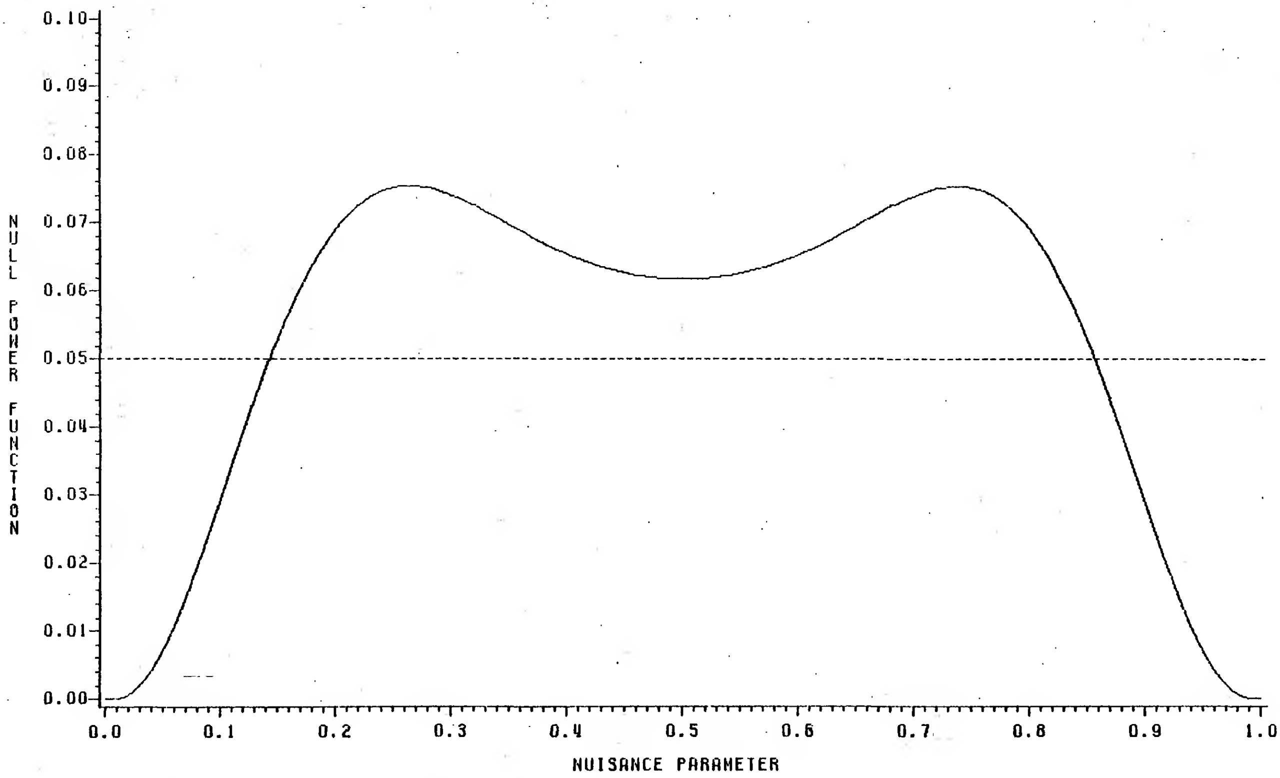

Figure B.1 Exact null power function of Z-test with unpooled variance ($n=10$, $z=1.645$, $\alpha=.05$) for comparing two independent proportions.

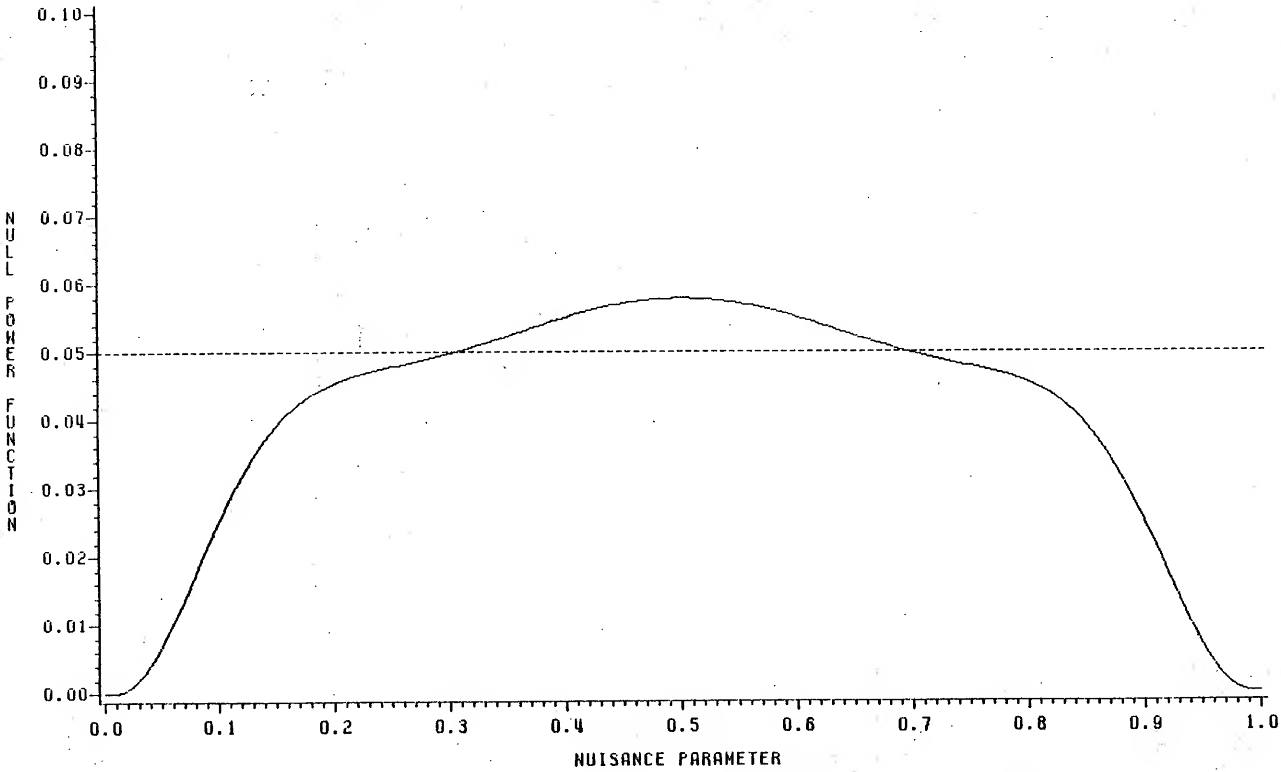

Figure B.2 Exact null power function of Z-test with pooled variance ($n=10$, $z=1.645$, $\alpha=.05$) for comparing two independent proportions.

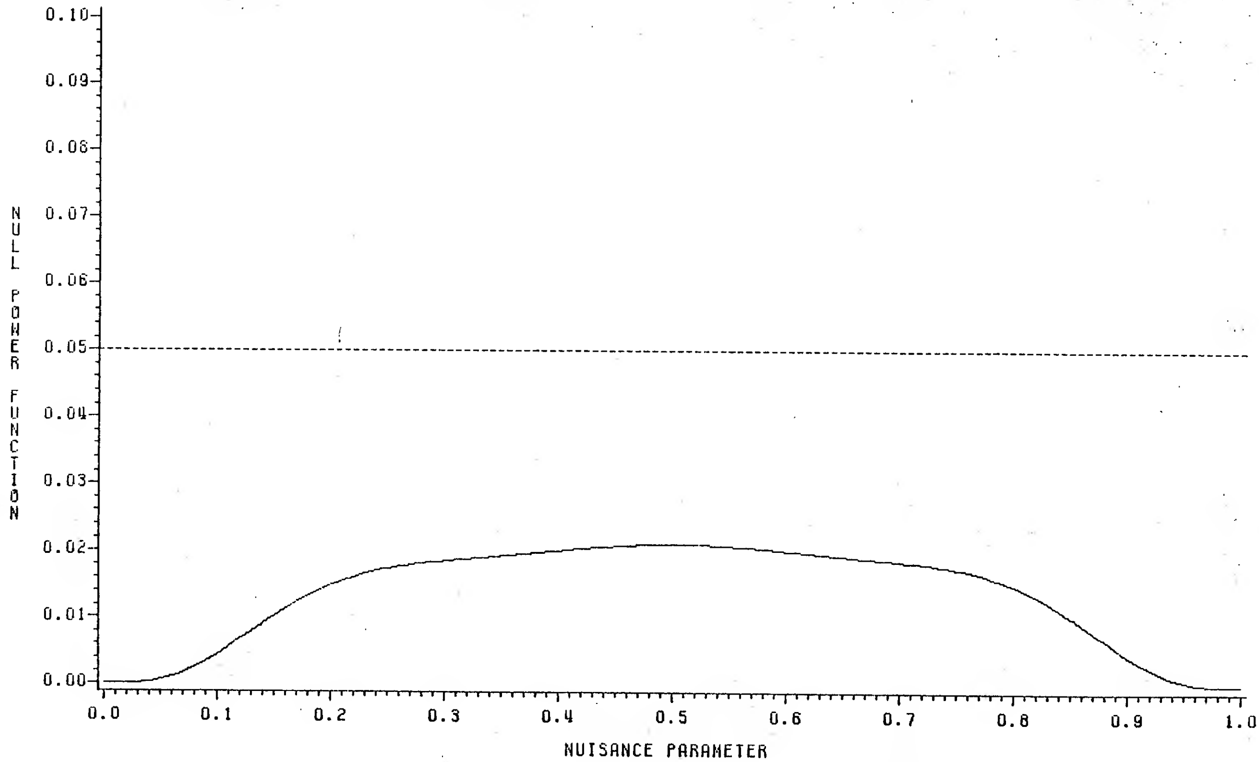

Figure B.3 Exact null power function of Fisher's exact test with n=10 and nominal significance level .05.

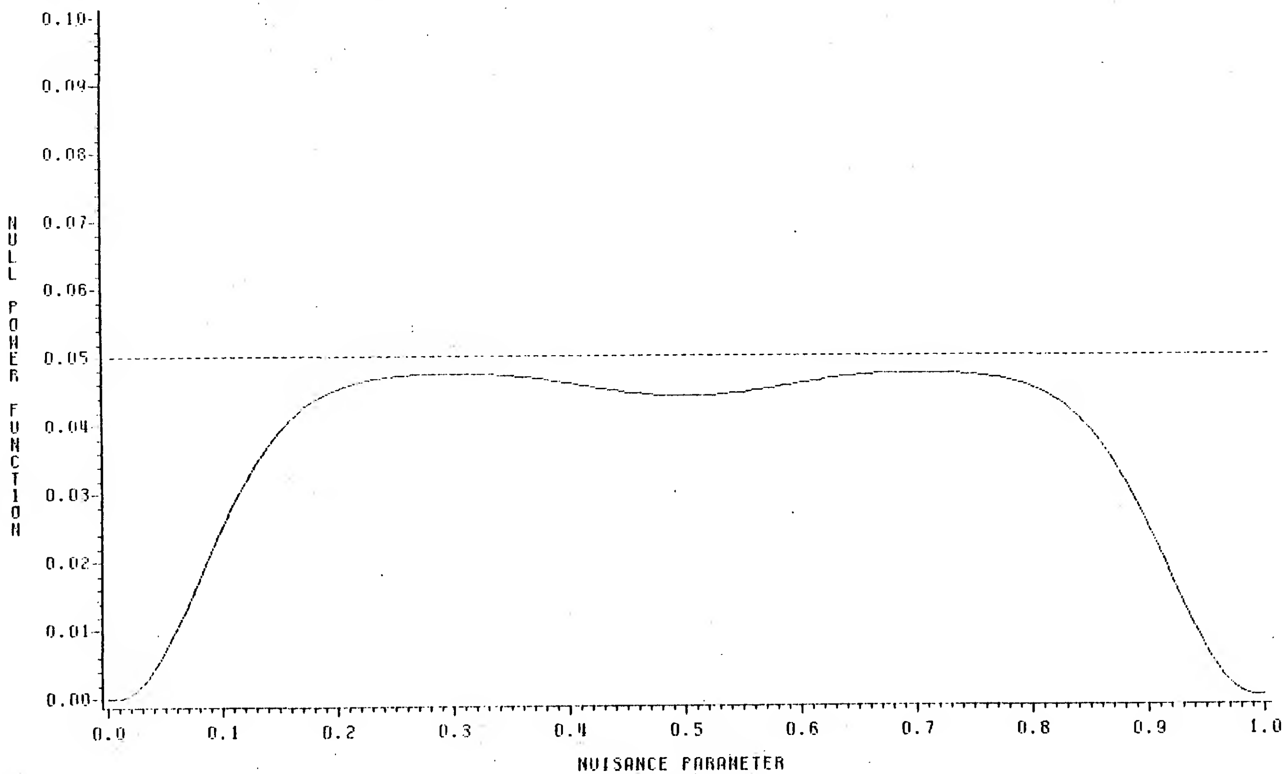

Figure B.4 Exact null power function of the exact Z-test ($n=10$, $\alpha=.05$, $z_u^*=1.96$ or $z_p^*=1.80$) for comparing two independent proportions.

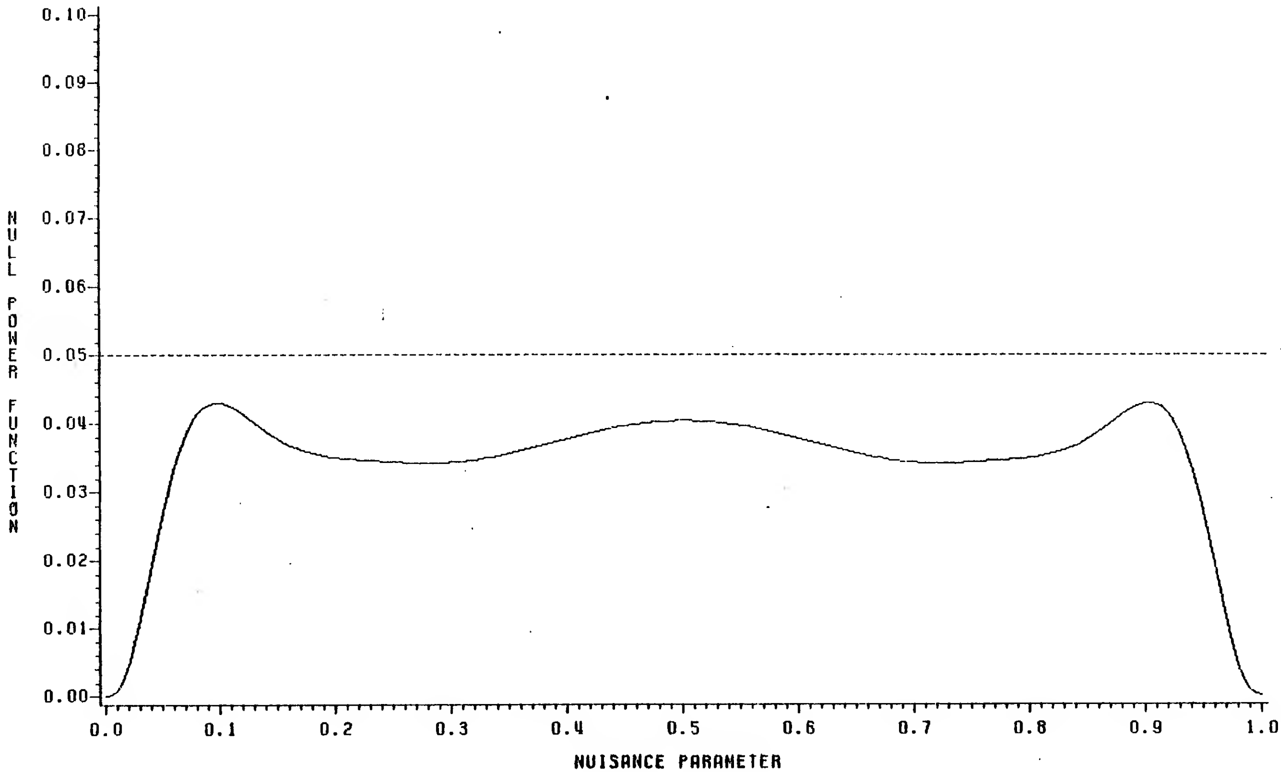

Figure B.5 Exact null power function of the exact Z-test ($n=20$, $\alpha=.05$, $z_u^*=1.85$ or $z_p^*=1.78$) for comparing two independent proportions.

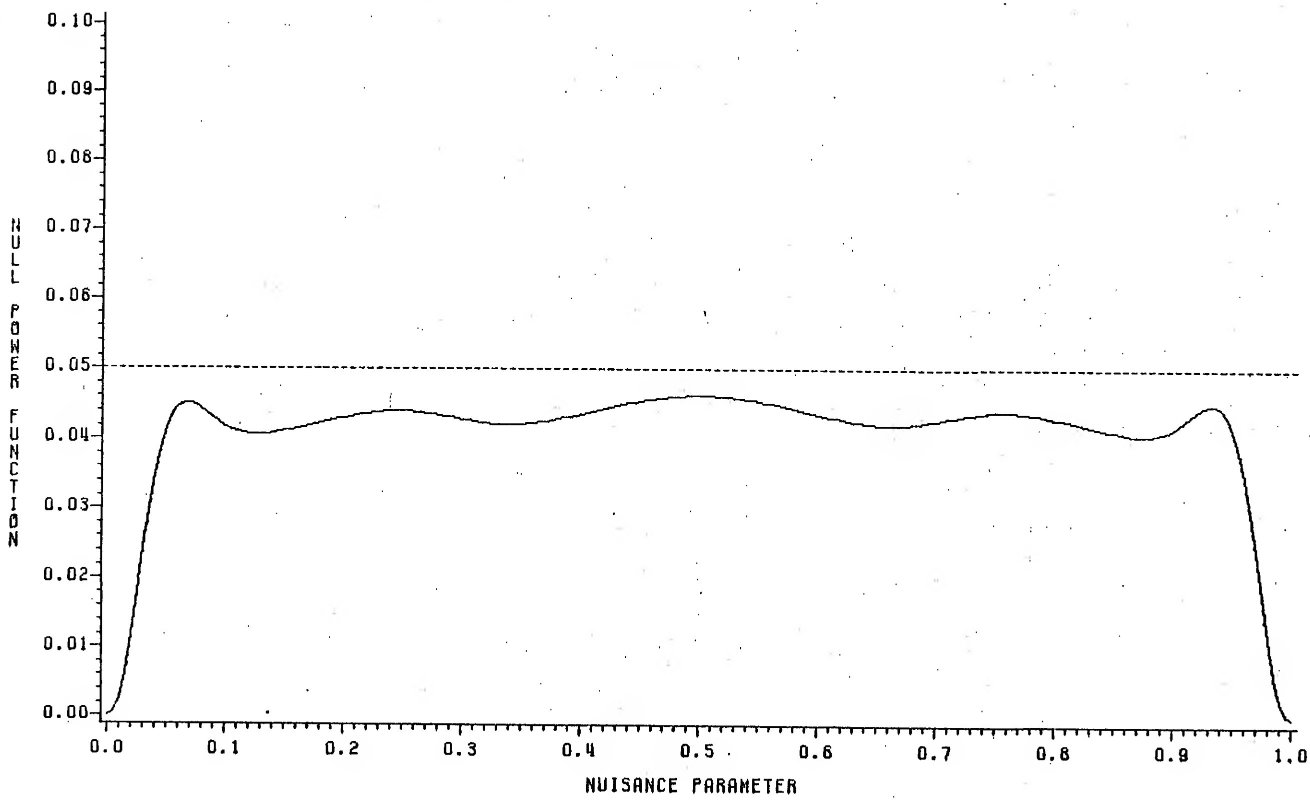

Figure B.6 Exact null power function of the exact Z-test ($n=30$, $\alpha=.05$, $z_u^*=1.77$ or $z_p^*=1.73$) for comparing two independent proportions.

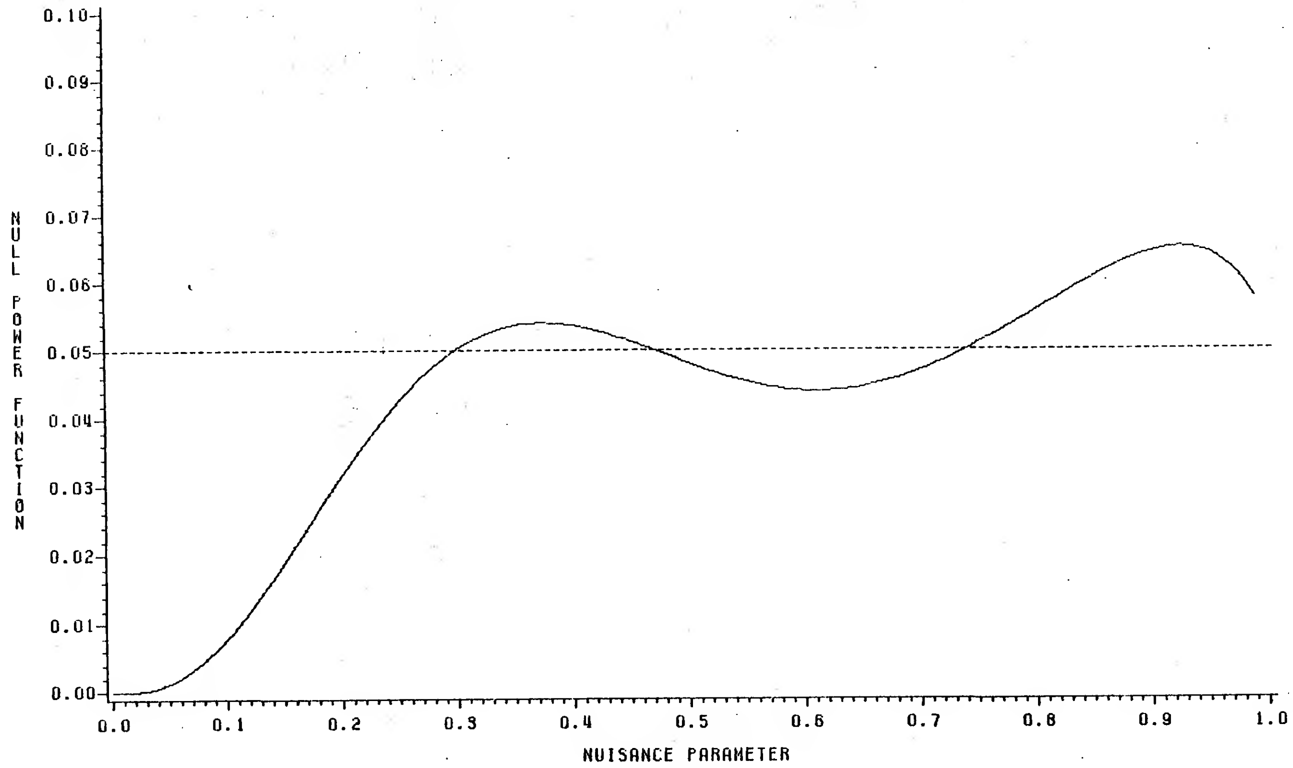

Figure B.7 Exact null power function of the Z-test ($N=10$, $z=1.645$, $\alpha=.05$) for comparing two correlated proportions.

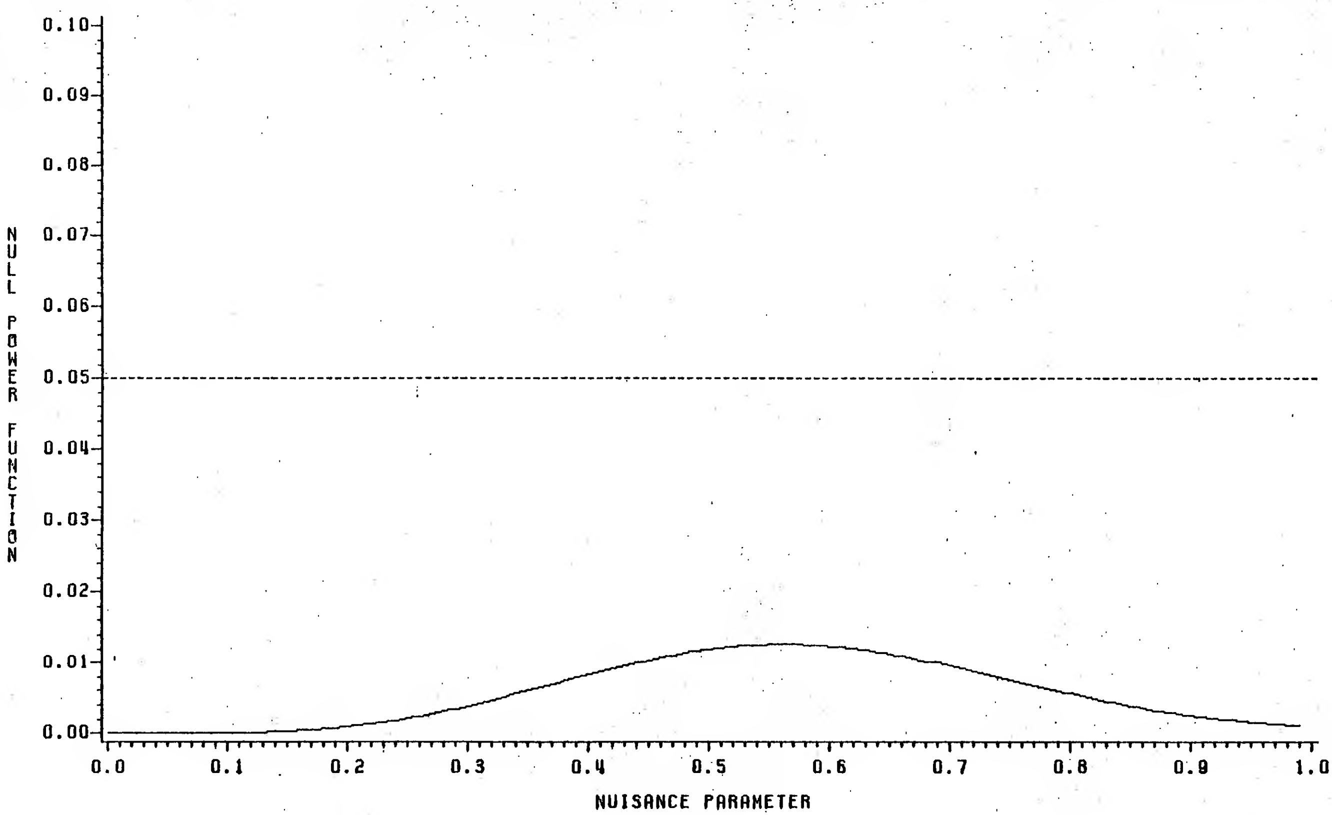

Figure B.8 Exact null power function of the sign test with N=10 and α=.05 for comparing two correlated proportions.

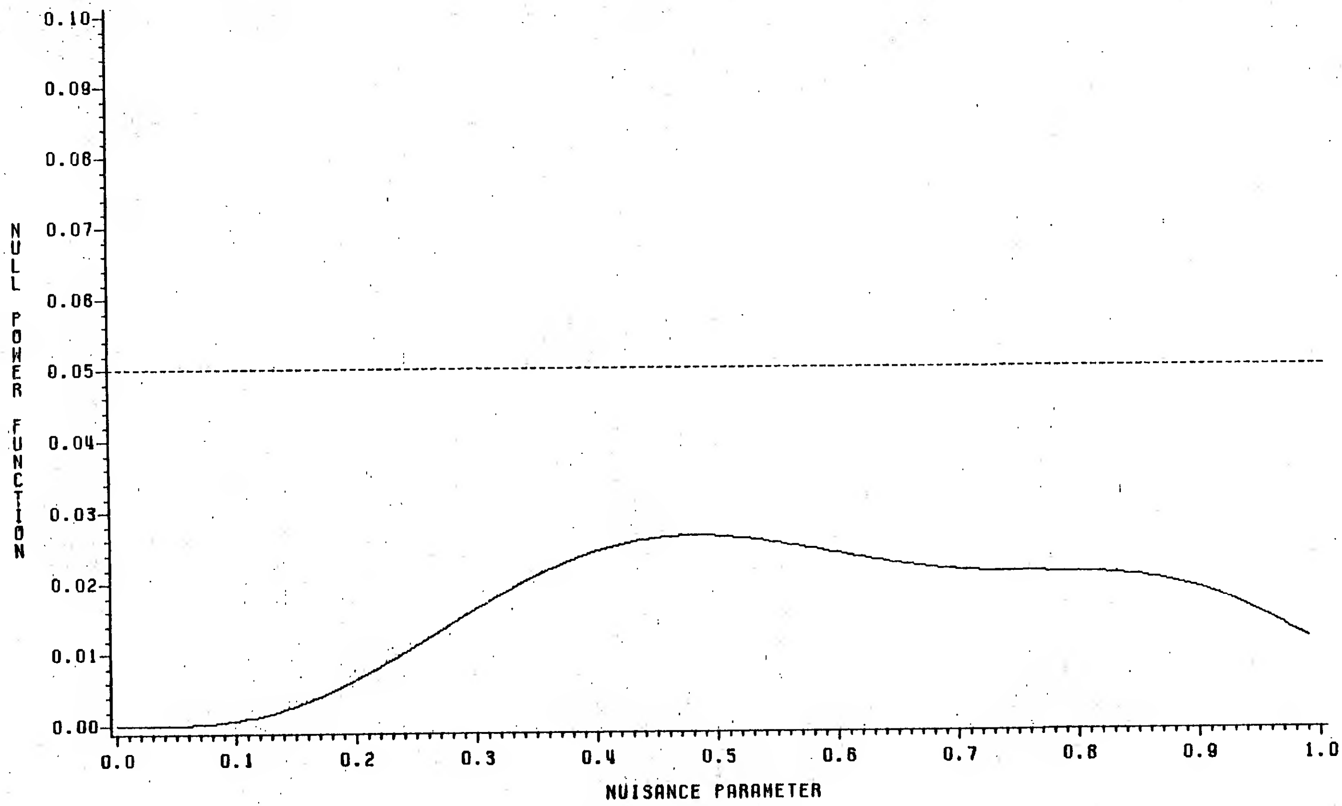

Figure B.9 Exact null power function of the exact Z-test ($N=10$, $\alpha=.05$, $z_c^*=1.90$) for comparing two correlated proportions.

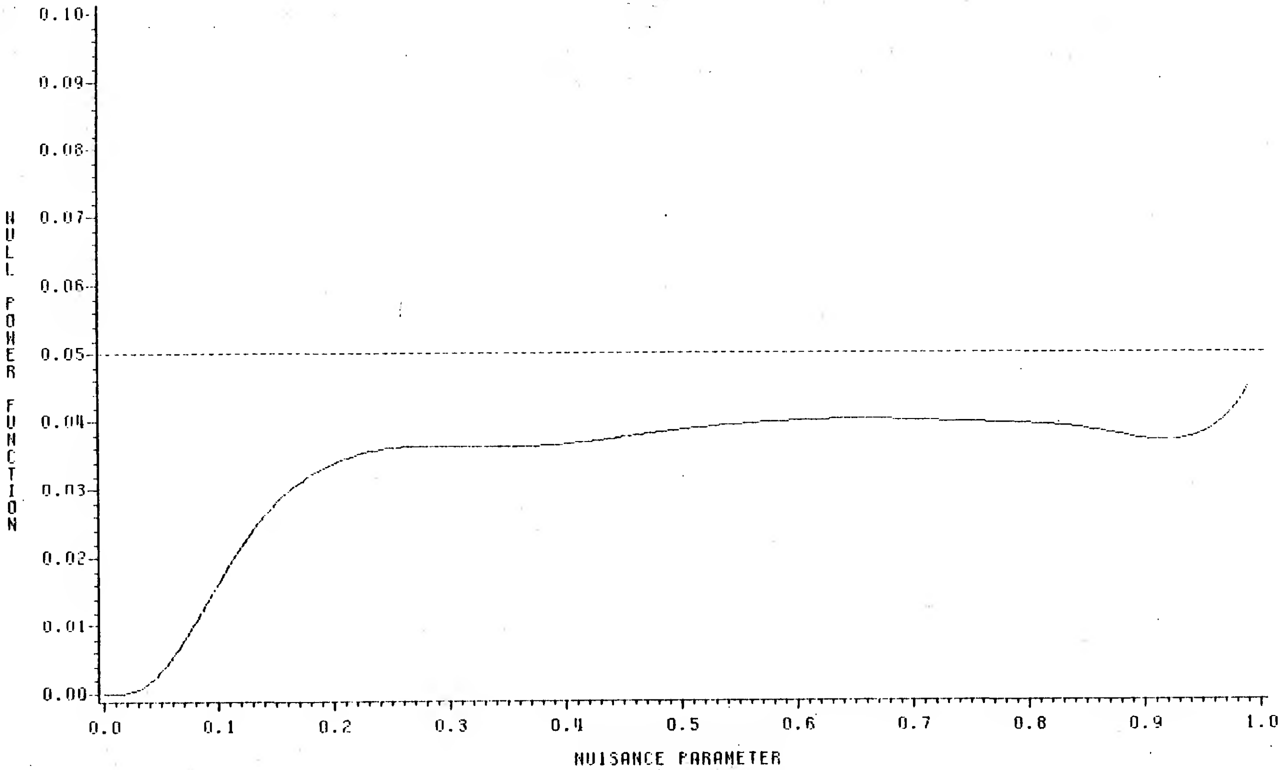

Figure B.10 Exact null power function of the exact Z-test (N=30, α=.05, z_c^*=1.74) for comparing two correlated proportions.

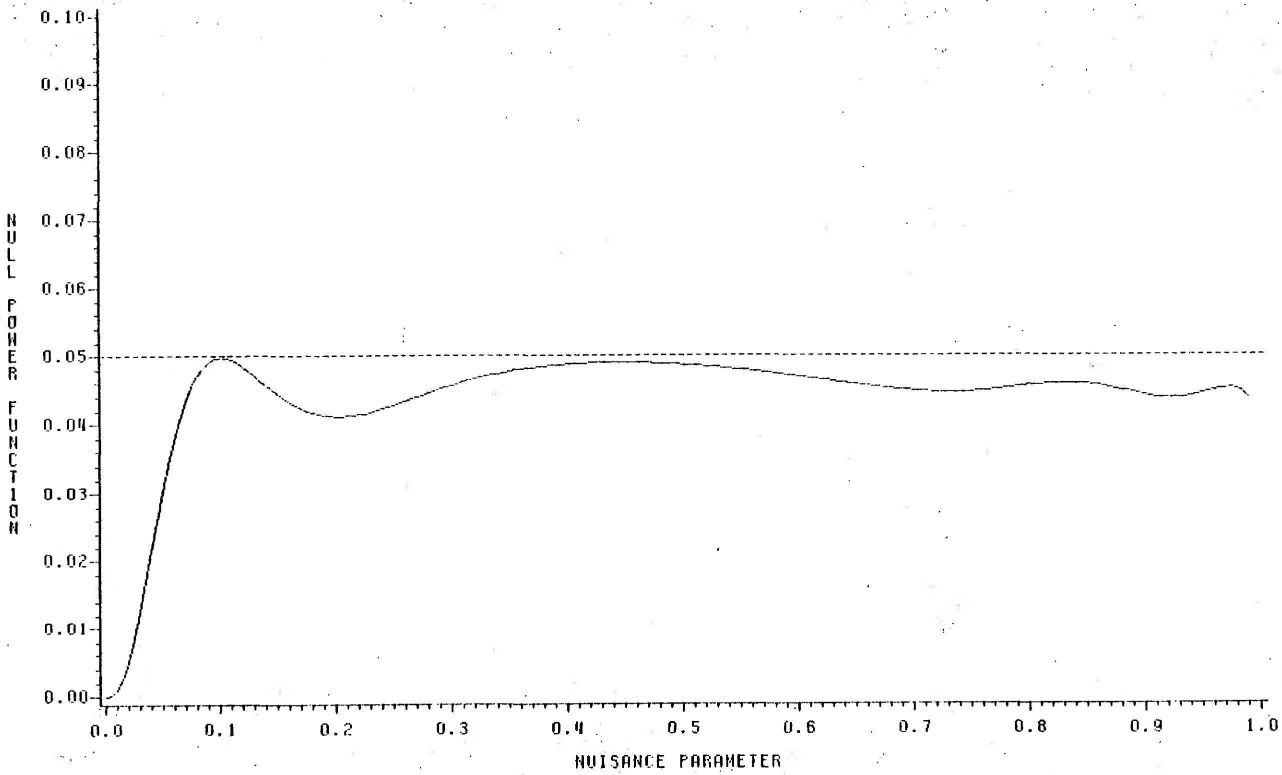

Figure B.11 Exact null power function of the exact Z-test (N=40, α=.05, z_c^*=1.68) for comparing two correlated proportions.

APPENDIX C

COMPUTER PROGRAMS

The listing of the FORTRAN computer programs used to compute the size of the Z-test for each of the two problems consdered is given in this Appendix. For the case of two independent proportions, Appendix C.1 gives the exact p-value for any n=10(1)150 and any value of the Z-test statistic with unpooled variance estimator. In Appendix C.2, the case of two correlated proportions, the exact p-value for N=10(1)200 and any value of the Z-test statistic can be obtained.

APPENDIX C.1

```
C
C
C
C
C
C
      REAL A,B,C,TBY(151)
      INTEGER X,U,Y,UU,BOUNDY(151),PHI(151)
      DOUBLE PRECISION LFAC(151),P,Q,LP,LQ,PX(151),FX(151),
     *PVALUE,DERBND,DERU,DERL,LCOM,PHAT1,PHAT2,
     *PMAX1,PMAX2,PMIN1,PMIN2,PL,PU,PL1,PL2,PU1,PU2,SUP,INF,
     *C1,C2,T,T3,T4,T5,T6,D1,D2,D3,D4,D5,D6,LPX,LFX,
     *MAXP,MAXIM,MINIM5,MINI10
      DOUBLE PRECISION DLOG,DLOG10,DEXP,DMAX1,DMIN1,DABS
      WRITE(6,40)
40    FORMAT(' ',///,
C
C
     * 'THIS PROGRAM COMPUTES THE P-VALUE OF THE Z-TEST'/
     * 'WITH UNPOOLED VARIANCE ESTIMATOR, TO COMPARE'/
     * 'TWO INDEPENDENT PROPORTIONS FOR'/
     * ' N = SAMPLE SIZE IN EACH GROUP,'/
     * ' Z = NORMAL STATISTIC WITH UNPOOLED VARIANCE.'/
     * 'ENTER N, Z, IN FREE FORMAT'/)
C
C
99    READ(9,*)N,Z
      MAXP=0.0
      N1=N+1
      LFAC(1)=0.D0
      DO 1 J=2,N1
1     LFAC(J)=LFAC(J-1)+DLOG(DBLE(FLOAT(J-1)))
C
C
C
C      SIMPLIFY THE COMPUTATIONS AND FORM THE
C      BOUNDARY OF THE CRITICAL REGION
C
C
      X=0
      A=1+(Z**2)/N
2     B=2*X+Z**2
      C=(X**2)*A-(Z**2)*X
      TBY(X+1)=(B+SQRT((B**2)-(4*A*C)))/(2*A)
      IF (TBY(X+1)-N) 3,3,4
3     X=X+1
      GO TO 2
4     U=X-1
      UU=U+1
      DO 5 X=1,UU
      BOUNDY(X)=INT(TBY(X)+1)
      PHI(X)=INT(TBY(X))
      L=X-1
5     CONTINUE
      MAXIM=0.D0
      MINIM5=1.D0
      MINI10=1.D0
      I9=0
      I5=0
      P=-.005D0
6     IF(P.GT.0.49D0) GO TO 76
      L=1
72    IF(L.EQ.1) P=P+.01D0
      IF(L.EQ.1) PP=P+.005D0
      Q=1-P
      LP=DLOG(P)
      LQ=DLOG(Q)
      DO 7 J=1,N1
      X=J-1
      T=LFAC(N+1)-LFAC(J)-LFAC(N+1-J+1)+X*LP+(N-X)*LQ
      IF (T.LT.-180.0) PX(J)=0.D0
      IF (T.GE.-180.0) PX(J)=DEXP(T)
```

```
      IF (X) 8,9,8
9     FX(J)=PX(J)
      GO TO 10
8     FX(J)=FX(J-1)+PX(J)
10    CONTINUE
7     CONTINUE
      PVALUE=0
      DERU=0
      DERL=0
      DO 30 K=1,UU
      X=K-1
      IF (PX(K).LE.0.D0) GO TO 11
      LPX=DLOG10(PX(K))
      IF ((1-FX(PHI(K)+1)).LE.0.D0) GO TO 11
      LFX=DLOG10(1-FX(PHI(K)+1))
      IF ((LPX+LFX).LT.-76.0) GO TO 11
      PVALUE=PVALUE+PX(K)*(1-FX(PHI(K)+1))
11    Y=BOUNDY(K)
      LCOM=LFAC(N+1)-LFAC(K)-LFAC(N+1-K+1)
     *          +LFAC(N+1)-LFAC(Y)-LFAC(N+1-Y-1+1)
C
C
C      LOCAL BOUND FOR THE DERIVATIVE OF
C      THE NULL POWER FUNCTION
C
C
      C1=FLOAT(X+Y-1)
      C2=FLOAT(2*N-X-Y)
      PHAT1=C1/(C1+C2)
      PHAT2=1-PHAT1
      PL=P-(.005D0/(2**(L-1)))
      PU=P+(.005D0/(2**(L-1)))
      IF (PHAT1.GT.PU) GO TO 12
      IF ((PHAT1.LE.PU).AND.(PHAT1.GE.PL)) GO TO 13
      IF (PHAT1.LT.PL) GO TO 14
12    PMAX1=PU
      PMIN1=PL
      GO TO 15
13    PMAX1=PHAT1
      IF (PL.NE.0.D0) GO TO 21
      PMIN1=PL
      GO TO 15
21    PL1=C1*DLOG(PL)+C2*DLOG(1-PL)
      PU1=C1*DLOG(PU)+C2*DLOG(1-PU)
      IF (PL1.LT.PU1) PMIN1=PL
      IF (PL1.GE.PU1) PMIN1=PU
      GO TO 15
14    PMAX1=PL
      PMIN1=PU
15    CONTINUE
      IF (PHAT2.GT.PU) GO TO 16
      IF ((PHAT2.LE.PU).AND.(PHAT2.GE.PL)) GO TO 17
      IF (PHAT2.LT.PL) GO TO 18
16    PMAX2=PU
      PMIN2=PL
      GO TO 19
17    PMAX2=PHAT2
      IF (PL.NE.0.D0) GO TO 22
      PMIN2=PL
      GO TO 19
22    PL2=C2*DLOG(PL)+C1*DLOG(1-PL)
      PU2=C2*DLOG(PU)+C1*DLOG(1-PU)
      IF (PL2.LT.PU2) PMIN2=PL
      IF (PL2.GT.PU2) PMIN2=PU
      GO TO 19
18    PMAX2=PL
      PMIN2=PU
19    CONTINUE
      T3=LCOM+C1*DLOG(PMAX1)+C2*DLOG(1-PMAX1)
      IF (T3.LT.-180.0) D3=0.D0
      IF (T3.GE.-180.0) D3=DEXP(T3)
      IF (PMIN2.EQ.0.D0) GO TO 23
```

```
      T4=LCOM+C2*DLOG(PMIN2)+C1*DLOG(1-PMIN2)
      IF (T4.LT.-180.0) D4=0.D0
      IF (T4.GE.-180.0) D4=DEXP(T4)
      GO TO 24
23    D4=0.D0
24    IF (PMIN1.EQ.0.D0) GO TO 25
      T5=LCOM+C1*DLOG(PMIN1)+C2*DLOG(1-PMIN1)
      IF (T5.LT.-180.0) D5=0.D0
      IF (T5.GE.-180.0) D5=DEXP(T5)
      GO TO 26
25    D5=0.D0
26    CONTINUE
      T6=LCOM+C2*DLOG(PMAX2)+C1*DLOG(1-PMAX2)
      IF (T6.LT.-180.0) D6=0.D0
      IF (T6.GE.-180.0) D6=DEXP(T6)
      DERU=DERU+D3-D4
      DERL=DERL+D5-D6
30    CONTINUE
      DERBND=DMAX1(DABS(DERU),DABS(DERL))
C
C
C      COMPUTE THE SIZE
C
C
      SUP=PVALUE+(.005D0/(2**(L-1)))*DERBND
      INF=PVALUE-(.005D0/(2**(L-1)))*DERBND
      MAXP=DMAX1(MAXP,PVALUE)
      IF (P.GE.0.1D0) I9=1
      IF (P.GE.0.05D0) I5=1
C
C      CHECK FOR PRECISION OF SIZE, HERE DELTA=.001
C
      IF ((SUP-MAXP).LE.0.001) GO TO 74
      L=L+1
      P=P-(.005D0/(2**(L-1)))
      GO TO 72
74    MAXIM=DMAX1(MAXIM,SUP)
      IF (I5.EQ.1) MINIM5=DMIN1(MINIM5,INF)
      IF (I9.EQ.1) MINI10=DMIN1(MINI10,INF)
      IF (L.EQ.1) GO TO 6
      IF ((PP-PU).LT.0.00001D0) GO TO 75
      P=P+(.005D0/(2**(L-1)))*2
      GO TO 72
75    P=PP-.005D0
      GO TO 6
76    CONTINUE
      WRITE(6,41)
41    FORMAT(' ','PVALUE FOR COMPARING TWO INDEPENDENT',
     * 'PROPORTIONS',/'FOR:N',
     * 4X,'Z',4X,' IS:',2X,'MIN5',6X,'MIN10',5X,'PVALUE'/)
      WRITE(6,33) N,Z,MINIM5,MINI10,MAXIM
33    FORMAT(' ',I4,F8.3,3F11.5)
      GO TO 99
      END
```

APPENDIX C.2

```
C
C
C
C
C
C
      INTEGER T,T1,T2
      DOUBLE PRECISION LFAC(201),P,Q,PP,LP,LQ,PM,AM,C,LP5
     * ,LAM(201),PL,PU,PHAT1,PHAT2,PMAX1,PMAX2,PMIN1,PMIN2,
     * LCOEF,FU1,FU2,FL1,FL2,GU1,GU2,GL1,GL2,DERU,DERL,
     * LPVAL,PVAL,PVALUE,DERBND,SUP,INF,MAXP,MAXIM,MINIM5,
     * MINIM9,PL1,PU1,PL2,PU2,DLOG,DEXP,DABS,DMAX1,DMIN1
      WRITE(6,40)
40    FORMAT(' ',///,
C
C
C
     * 'THIS PROGRAM COMPUTES THE P-VALUE OF THE Z-TEST'/
     * 'FOR COMPARING TWO CORRELATED PROPORTIONS FOR'/
     * ' N = TOTAL NUMBER OF MATCHED PAIRS,'/
     * ' Z = NORMAL STATISTIC.'/
     * 'ENTER N, Z, IN FREE FORMAT'/)
C
C
C
      LP5=DLOG(0.5D0)
99    READ(9,*)N,Z
      N1=N+1
      LFAC(1)=0.D0
      DO 1 J=2,N1
1     LFAC(J)=LFAC(J-1)+DLOG(DBLE(FLOAT(J-1)))
C
C
C      SIMPLIFY THE COMPUTATIONS
C
C
C
      LN=INT((Z-.0000001)**2)+1
      DO 2 M=LN,N
      M1=M+1
      T=INT((Z*SQRT(FLOAT(M))+M)/2)
      T1=T+1
      T2=T+2
      PM=0.D0
      AM=0.D0
      DO 3 J=T1,M
      C=LFAC(M+1)-LFAC(J+1)-LFAC(M-J+1)+M*LP5
      IF (C.LT.-180.) PM=0.D0
      IF (C.GE.-180.) PM=DEXP(C)
3     AM=AM+PM
2     LAM(M)=DLOG(AM)
      MAXP=0.D0
      MAXIM=0.D0
      MINIM5=1.D0
      MINIM9=1.D0
      I9=0
      I5=0
      P=-.005D0
6     IF (P.GT.0.98) GO TO 76
      L=1
72    IF (L.EQ.1) P=P+.01D0
      IF (L.EQ.1) PP=P+.005D0
      Q=1-P
      LP=DLOG(P)
      LQ=DLOG(Q)
      DERU=0.D0
      DERL=0.D0
      PVALUE=0.D0
C
C
C      LOCAL BOUND FOR THE DERIVATIVE
C      OF THE NULL POWER FUNCTION
```

```
C
C
      PL=P-(.005D0/(2**(L-1)))
      PU=P+(.005D0/(2**(L-1)))
      DO 4 M=LN,N
      PHAT1=FLOAT(M-1)/FLOAT(N-1)
      PHAT2=FLOAT(M)/FLOAT(N-1)
      IF (PHAT1.GT.PU) GO TO 12
      IF ((PHAT1.LE.PU).AND.(PHAT1.GE.PL)) GO TO 13
      IF (PHAT1.LT.PL) GO TO 14
12    PMAX1=PU
      PMIN1=PL
      GO TO 15
13    PMAX1=PHAT1
      IF (PL.NE.0.D0) GO TO 21
      PMIN1=PL
      GO TO 15
21    PL1=(M-1)*DLOG(PL)+(N-M)*DLOG(1-PL)
      PU1=(M-1)*DLOG(PU)+(N-M)*DLOG(1-PU)
      IF (PL1.LT.PU1) PMIN1=PL
      IF (PL1.GE.PU1) PMIN1=PU
      GO TO 15
14    PMAX1=PL
      PMIN1=PU
15    CONTINUE
      IF (PHAT2.GT.PU) GO TO 16
      IF ((PHAT2.LE.PU).AND.(PHAT2.GE.PL)) GO TO 17
      IF (PHAT2.LT.PL) GO TO 18
16    PMAX2=PU
      PMIN2=PL
      GO TO 19
17    PMAX2=PHAT2
      IF (PL.NE.0.D0) GO TO 22
      PMIN2=PL
      GO TO 19
22    PL2=M*DLOG(PL)+(N-M-1)*DLOG(1-PL)
      PU2=M*DLOG(PU)+(N-M-1)*DLOG(1-PU)
      IF (PL2.LT.PU2) PMIN2=PL
      IF (PL2.GT.PU2) PMIN2=PU
      GO TO 19
18    PMAX2=PL
      PMIN2=PU
19    CONTINUE
      LCOEF=LAM(M)+LFAC(N+1)-LFAC(M+1)-LFAC(N-M+1)
      FU1=LCOEF+(M-1)*DLOG(PMAX1)+(N-M)*DLOG(1-PMAX1)
      IF (FU1.LT.-180.0) GU1=0.D0
      IF (FU1.GE.-180.0) GU1=M*DEXP(FU1)
      FL2=LCOEF+M*DLOG(PMAX2)+(N-M-1)*DLOG(1-PMAX2)
      IF (FL2.LT.-180.0) GL2=0.D0
      IF (FL2.GE.-180.0) GL2=(N-M)*DEXP(FL2)
      IF (PMIN2.EQ.0.D0) GO TO 23
      FU2=LCOEF+M*DLOG(PMIN2)+(N-M-1)*DLOG(1-PMIN2)
      IF (FU2.LT.-180.0) GU2=0.D0
      IF (FU2.GE.-180.0) GU2=(N-M)*DEXP(FU2)
      GO TO 24
23    GU2=0.D0
24    IF (PMIN1.EQ.0.D0) GO TO 25
      FL1=LCOEF+(M-1)*DLOG(PMIN1)+(N-M)*DLOG(1-PMIN1)
      IF (FL1.LT.-180.0) GL1=0.D0
      IF (FL1.GE.-180.0) GL1=M*DEXP(FL1)
      GO TO 26
25    GL1=0.D0
26    CONTINUE
      DERU=DERU+GU1-GU2
      DERL=DERL+GL1-GL2
      LPVAL=LCOEF+M*LP+(N-M)*LQ
      IF (LPVAL.LT.-180.) PVAL=0.D0
      IF (LPVAL.GE.-180.) PVAL=DEXP(LPVAL)
      PVALUE=PVALUE+PVAL
4     CONTINUE
      DERBND=DMAX1(DABS(DERU),DABS(DERL))
C
```

```
C
C       COMPUTE THE SIZE
C
C
        SUP=PVALUE+(.005D0/(2**(L-1)))*DERBND
        INF=PVALUE-(.005D0/(2**(L-1)))*DERBND
        MAXP=DMAX1(MAXP,PVALUE)
        IF (P.GE.(10./FLOAT(N))) I9=1
        IF (P.GE.(5./FLOAT(N)))  I5=1
C
C       CHECK FOR PRECISION OF SIZE, HERE DELTA=.001
C
        IF ((SUP-MAXP).LE.0.001) GO TO 74
        L=L+1
        P=P-(.005D0/(2**(L-1)))
        GO TO 72
74      MAXIM=DMAX1(MAXIM,SUP)
        IF (I5.EQ.1) MINIM5=DMIN1(MINIM5,INF)
        IF (I9.EQ.1) MINIM9=DMIN1(MINIM9,INF)
        IF (L.EQ.1) GO TO 6
        IF ((PP-PU).LT.0.00001D0) GO TO 75
        P=P+(.005D0/(2**(L-1)))*2
        GO TO 72
75      P=PP-.005D0
        GO TO 6
76      CONTINUE
        WRITE(6,41)
41      FORMAT(' ','PVALUE FOR COMPARING TWO CORRELATED ',
       * 'PROPORTIONS',/'FOR:N',
       * 4X,'Z',4X,'IS:',2X,'MIN5',6X,'MIN10',4X,'PVALUE'/)
        WRITE(6,33)N,Z,MINIM5,MINIM9,MAXIM
33      FORMAT(' ',I4,F8.3,3F11.5)
        GO TO 99
        END
```

REFERENCES

Barnard, G.A. (1947). Significance Tests for 2×2 Tables. Biometrika 34, 123.

Barnard, G.A. (1979). In Contradiction to J. Berkson's Dispraise: Conditional Tests can be more Efficient. Journal of Statistical Planning and Inference 3, 181.

Basu, D. (1977). On the Elimination of Nuisance Parameters. Journal of the American Statistical Association 72, 355.

Basu, D. (1979). Discussion of Joseph Berkson's Paper "in Dispraise of the Exact Test". Journal of Statistical Planning and Inference 3, 189.

Bennett, B.M. (1967). Tests of Hypotheses Concerning Matched Samples. Journal of the Royal Statistical Society B 29, 468.

Bennett, B.M., and Underwood, R.E. (1970). On McNemar's Test for the 2×2 Table and its Power Function. Biometrics 26, 339.

Berkson, J. (1978). In Dispraise of the Exact Test. Journal of Statistical Planning and Inference 2, 27.

Casagrande, J.T., Pike, M.C., and Smith, P.G. (1978a). An Improved Approximate Formula for Calculating Sample Sizes for Comparing Two Binomial Distributions. Biometrics 34, 483.

Casagrande, J.T., Pike, M.C., and Smith, P.G. (1978b). The Power Function of the "Exact" Test for Comparing Two Binomial Distributions. Applied Statistics 27, 176.

Cochran, W.G. (1950). The Comparision of Percentages in Matched Samples. Biometrika 37, 256.

Cochran, W.G., and Cox, G.M. (1957). Experimental Designs, 2nd ed. New York: John Wiley and Sons.

Corsten, L.C.A., and de Kroon, J.P.M. (1979). Comment on J. Berkson's Paper "In Dispraise of the Exact Test". Journal of Statistical Planning and Inference 3, 193.

Courant, R., and John, F. (1965). Introduction to Calculus and Analysis, Volume 1. Interscience, New York: John Wiley and Sons.

Eberhardt, K.R., and Fligner, M.A. (1977). A Comparision of Two Tests for Equality of Two Proportions. The American Statistician 31, 151.

Ferguson, T.S. (1967). Mathematical Statistics: a Decision Theoretic Approach. New York: Academy Press.

Fisher, R.A. (1935). The Logic of Inductive Inference. Journal of the Royal Statistical Society A 98, 39.

Fleiss, J.L. (1980). Statistical Methods for Rates and Proportions, 2nd ed. New York: John Wiley and Sons.

Gail, M. and Gart, J.J. (1973). The Determination of Sample Sizes for Use with the Exact Conditional Test in 2×2 Comparative Trials. Biometrics 29, 441.

Haseman, J.K. (1978). Exact Sample Sizes for Use with the Fisher-Irwin Test for 2×2 Tables. Biometrics 34, 106.

Kempthorne, O. (1979). In Dispraise of the Exact Test: Reactions. Journal of Statistical Planning and Inference 3, 199.

Kendall, M.G., and Stuart, A. (1967). The Advanced Theory of Statistics, Vol. 2, 2nd ed. New York: Hafner Publishing Company.

Kramer, M., and Greenhouse, S.W. (1959). Determination of Sample Size and Selection of Cases, Psychopharmacology: Problems in Evaluation. J.O. Cole and R.W. Gerard(eds), 356. Washington: National Academy of Sciences, National Research Council.

Lehmann, E.L. (1959). Testing Statistical Hypotheses. New York: John Wiley and Sons.

McDonald, L.L., Davis, B.M., and Milliken, G.A. (1977). A Nonrandomized Unconditional Test for Comparing Two Proportions in 2×2 Contigency Tables. Technometrics 19, 145.

McNemar, Q. (1947). Note on the Sampling Error of the Differences Between Correlated Proportions or Percentages. Psychometrika 12, 153.

Miettinen, O.S. (1968). The Matched Pairs Design in the Case of All-or-none Responses, Biometrics 24, 339.

Pearson, E.S. (1947). The Choice of Statistical Tests Illustrated on the Interpretation of Data Classed in a 2×2 Table. Biometrika 34, 139.

Randles, R.H., and Wolfe, D.A. (1979). Introduction to the Theory of Nonparametric Statistics. New York: John Wiley and Sons.

Robbins, H. (1977). A Fundamental Question of Practical Statistics, Letter to the Editor. The American Statistician 31, 97.

Schork, M.A., and Williams, G.W. (1980). Number of Observations Required for the Comparision of Two Correlated Proportions. Communications in Statistics B 9, 349.

Welch, R.H. (1939). On Confidence Limits and Sufficency, with Particular Reference to Parameters of Location. Annals of Mathematical Statistics 10, 58.

BIOGRAPHICAL SKETCH

Samy Suissa was born on July 9, 1954, in Berréchid, a suburb of Casablanca, Morocco. In 1964, Samy's family moved to Montréal, Canada, where they still reside. There, he attended Baron Byng High School, famous for its role in "The Apprenticeship of Duddy Kravitz" by Mordechai Richler, until 1968 and Outremont High School until 1970. He then entered McGill University where he received a Bachelor of Science degree in mathematics in June 1976, and a Master of Science degree in mathematical statistics in November 1977.

Samy came to the University of Florida in the fall of 1977 to pursue a doctoral degree in statistics. While studying at the University of Florida, Samy has worked as a graduate assistant performing statistical consulting duties in the Department of Orthopaedics among other places.

On July 1, 1982, Samy Suissa will join the "Service d'Epidémiologie Clinique" of the Montreal General Hospital's Research Institute in the capacity of biostatistician. He will also have an appointment as assistant professor in the Department of Epidemiology and Health at McGill University. He has recently been awarded a research grant from "Les Fonds de Recherche en Santé du Québec" for 1982-1983.

He is married to the former Nicole Bonenfant of Montréal, Canada, and has a son, Daniel Moshé.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Jon Shuster

Jonathan J. Shuster, chairman
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Ronald H. Randles

Ronald H. Randles
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Mark P. Hale Jr.

Mark P. Hale
Associate Professor of Mathematics

This dissertation was submitted to the Graduate Faculty of the Department of Statistics in the College of Liberal Arts and Sciences and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1982

Dean for Graduate Studies
and Research