INFERENCE FROM STOPPED BERNOULLI SEQUENCES

BY

BETH GLADEN

TECHNICAL REPORT NO. 27
NOVEMBER 22, 1976

PREPARED UNDER THE AUSPICES
OF
PUBLIC HEALTH SERVICE GRANT 5 R01 GM21215-02

DIVISION OF BIOSTATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
INFERENCE FROM STOPPED BERNOLLI SEQUENCES

by

BETH GLADEN

TECHNICAL REPORT NO. 27  
November 22, 1976

PREPARED UNDER THE AUSPICES  
OF  
PUBLIC HEALTH SERVICE GRANT 5 RO1 GM21215-02

DIVISION OF BIOSTATISTICS  
STANFORD UNIVERSITY  
STANFORD, CALIFORNIA

Also issued as Technical Report No. 93 under the Auspices of National Science Foundation Grant MPS74-21416 - Department of Statistics Stanford University.
ACKNOWLEDGMENTS

I would like to thank Paul Switzer for many useful suggestions during the course of this research. I would like to thank Byron Brown for suggesting the spontaneous abortion problem to me. I would like to thank Thomas Sager for his comments on the manuscript. Finally, I would like to thank the National Science Foundation and the National Institutes of Health for financial support throughout the period of this research.
<table>
<thead>
<tr>
<th>CONTENTS</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I INTRODUCTION ................................................. 1</td>
<td></td>
</tr>
<tr>
<td>A. Example ..................................................... 1</td>
<td></td>
</tr>
<tr>
<td>B. Possible difficulties ....................................... 3</td>
<td></td>
</tr>
<tr>
<td>C. Outline ...................................................... 4</td>
<td></td>
</tr>
<tr>
<td>II MODELS ........................................................... 6</td>
<td></td>
</tr>
<tr>
<td>A. Comparison of models ........................................ 10</td>
<td></td>
</tr>
<tr>
<td>B. Calculations for stopping rule A .......................... 12</td>
<td></td>
</tr>
<tr>
<td>1. Model a: ( P\text{(success)} = p ) .................. 12</td>
<td></td>
</tr>
<tr>
<td>2. Model c: ( P\text{(success)} = p \sim F ) .............. 14</td>
<td></td>
</tr>
<tr>
<td>3. Model d: ( P\text{(success on ith trial)} = p_i ) .... 14</td>
<td></td>
</tr>
<tr>
<td>C. Calculations for stopping rule B .......................... 18</td>
<td></td>
</tr>
<tr>
<td>1. Model a: ( P\text{(success)} = p ) .................. 18</td>
<td></td>
</tr>
<tr>
<td>2. Model c: ( P\text{(success)} = p \sim F ) .............. 20</td>
<td></td>
</tr>
<tr>
<td>3. Model d: ( P\text{(success on ith trial)} = p_i ) .... 21</td>
<td></td>
</tr>
<tr>
<td>4. Model e: Markov case ......................................... 26</td>
<td></td>
</tr>
<tr>
<td>D. Calculations for stopping rule C .......................... 28</td>
<td></td>
</tr>
<tr>
<td>E. Calculations for stopping rule D .......................... 36</td>
<td></td>
</tr>
<tr>
<td>III CONDITIONAL TESTS .......................................... 38</td>
<td></td>
</tr>
<tr>
<td>IV USE OF NATURAL STATISTICS .................................. 41</td>
<td></td>
</tr>
<tr>
<td>V CHI-SQUARED TESTS .............................................. 44</td>
<td></td>
</tr>
<tr>
<td>A. General chi-squared theory ................................ 45</td>
<td></td>
</tr>
<tr>
<td>B. Application to zero-one sequences ........................ 47</td>
<td></td>
</tr>
<tr>
<td>1. General considerations .................................... 47</td>
<td></td>
</tr>
<tr>
<td>2. Grouping of sequences according to the total number of successes ................................ 49</td>
<td></td>
</tr>
<tr>
<td>3. Other groupings .............................................. 50</td>
<td></td>
</tr>
<tr>
<td>C. Numerical results ............................................ 53</td>
<td></td>
</tr>
<tr>
<td>D. Simulations .................................................. 58</td>
<td></td>
</tr>
<tr>
<td>E. Other considerations ......................................... 66</td>
<td></td>
</tr>
<tr>
<td>APPENDIX: TABLES ................................................ 68</td>
<td></td>
</tr>
<tr>
<td>REFERENCES ....................................................... 84</td>
<td></td>
</tr>
</tbody>
</table>
I. INTRODUCTION

Consider data sets which consist of many short sequences of zeroes and ones. The probability structures of such sequences may be regarded as having two components: one, the generating mechanism, determines whether each element is a zero or a one; the other, the stopping mechanism, determines the length of the sequence. Most of the work that has been done on sequences of zeroes and ones makes the assumption that the length of the sequence goes to infinity, and very little attention seems to have been paid to inference from short sequences. In this paper, when asymptotic theory is required, we will make the assumption that the number of sequences, rather than the length of each, goes to infinity.

The situation we would like to consider is one where the interest is in the generating model, and the stopping rule is a nuisance mechanism. We have a certain model in mind for the generating mechanism, and we would like to test the fit of this model. We have at best a vague idea of the stopping rule, and would like to develop procedures which are fairly insensitive to which stopping rule actually obtains.

A. Example

As motivation for the problem, and to provide a frame of reference, let us consider the question of spontaneous abortion rates. (See [3] for a fuller discussion of this subject.) The available data, for each woman,
consists of a string of zeroes and ones representing her pregnancy
history; "one" represents success, or live birth, and "zero" represents
failure, or spontaneous abortion (miscarriage). Such sequences will
necessarily be short. The goals of the investigation may be, for example,
to construct an estimator of the probability that the next pregnancy will
abort, based on the past history of the woman, or to answer questions
such as: "Are some women more prone than others to abort?" or "Does the
probability of abortion increase with age (or with its surrogate,
pregnancy number)?" A natural way to proceed is to build a model for
pregnancy outcome. An example of a model which has appeared in the
medical literature is the following: each woman has a constant abortion
probability, her pregnancies are a sequence of independent Bernoulli
trials, and abortion probabilities have a 2-point distribution over all
women [7]. Estimation of parameters in such a model, theoretically if
not computationally, is trivial, but testing whether the model fits
is not. It is clear that the decision of a woman to have another
pregnancy might be strongly influenced by the number of living children;
i.e., lengths of sequences are not independent of outcomes. Some account
of this must be taken when testing the fit of the model, but it is a
nuisance mechanism, difficult to model, and we would prefer tests of
fit which are independent of the stopping rule. Failing this, we would
like a procedure which is reasonably insensitive to it.
B. Possible difficulties

Suppose that we have a data set of the structure that has been described. If we make the assumption that the stopping rule is simply time censoring independent of the elements of the sequence, and this assumption is not true, the distortions introduced into many standard procedures can be quite dramatic.

As an illustration, consider the spontaneous abortion problem, and consider testing the hypothesis that abortion probabilities are constant across women and across trials. Under this hypothesis, conditional on the number of trials, the number of successes has a binomial distribution, and testing for goodness of fit becomes relatively simple.

However, a stopping rule which might be operating is the following: a woman continues to have pregnancies until she has had one live-born child, and then stops. Since, at the time of data collection, some women would not yet have been successful, the effect of this would be that a woman with \( n \) pregnancies must have either \( n \) or \( n-1 \) failures, even though the probability of success were quite high. Clearly the conditional distributions will be highly non-binomial, even though the hypothesis is true. Of course, this stopping rule is unrealistic, and sufficiently dramatic that it would not go undetected. But a generalization of the above, where women desire not one child, but some random number, is certainly plausible as a reasonable approximation to reality, and the conditional distributions that would be produced would still be distorted.
As a second example, suppose that the observed proportion of successes on the \( i \)th trial decreases with \( i \). We might be tempted to conclude that the probability of success for any given woman decreased with trial number. However, this same pattern of decreasing success proportions will be seen if the stopping rule described above holds, and the probability of success varies from woman to woman but is constant from trial to trial for a given woman. For, as time goes on, the low-risk women will have all the children they desire and will stop having pregnancies. Thus the women still having pregnancies will include a greater and greater proportion of high-risk women. While this second model does produce decreasing marginal probabilities of success, the implications are very different, and we would like to be able to distinguish differences within women from differences between women.

C. Outline

The remainder of this paper is concerned with investigating several approaches to testing goodness of fit for such zero-one sequences, with an eye to both the power of the test when the hypothesized model is not true, and to the insensitivity of the test to changes in the stopping rule.

Section II describes the various models and stopping rules which will play a role in the sequel. Probabilities of the possible sequences under the various combinations of models are derived, and the differences between the models are illustrated.
Section III describes the use of conditional tests of goodness of fit. These are particularly good as descriptive measures, but not as successful for formal testing, as they result in a collection of dependent test statistics.

Section IV describes the use of the distribution of natural descriptive measures such as number of successes. There seems to be no particular aspects of the distribution which is sensitive only to departures from the model and not to departures from the stopping rule.

Section V examines the use of standard chi-squared goodness of fit statistics. In particular, the effect of various schemes for classifying the possible sequences of zeroes and ones into a small number of cells is examined. Various schemes which seem natural are shown to have poor properties, and others are proposed which do well against several classes of alternatives.
II. MODELS

Let us suppose that we observe the random variables

\[ X_i = (X_{1i}, \ldots, X_{Ni}) \text{ for } i = 1, 2, \ldots, M. \]

Here \( X_{ij} \) is the result of the \( j^{th} \) trial of the \( i^{th} \) individual, and takes values in \{0, 1\}; \( N_i \) is the length of the \( i^{th} \) sequence and takes values in \{1, 2, \ldots\}.

Define \( K_i = \sum_{j=1}^{N_i} X_{ij} \), the number of successes for the \( i^{th} \) individual.

In the remainder of this section, we will consider various models for the distribution of the \( X_i \).

The stopping rules that will be considered can be summarized as follows:

A) \( P(\text{continue}) = q \)

Rule A specifies that the probability of continuation, for any individual at any point, is some constant, \( q \). This may well be appropriate for some situations, but is not appropriate for data sets like the spontaneous abortion data, where the probability of continuation can be clearly seen to depend on the results of previous trials.

B) \( P(\text{continue/preceding success}) = q_1 \)

\[ P(\text{continue/preceding failure}) = q_0 \]

Rule B allows the determination of whether or not to continue to depend on the result of the immediately preceding trial.
C) \( P(\text{continue/K preceding successes}) = q_k = \alpha q^k \)

Rule C allows the continuation probability to depend, not on the immediately preceding trial, but on the total number of preceding successes. This rule has, of course, an infinite number of parameters, so the further restriction that \( q_k = \alpha q^k \) will be used for simplicity. Any such decreasing sequence would correspond, in the spontaneous abortion case, to the desires of women to limit family size; the more successes a woman has, the less likely she is to become pregnant again.

D) \( P(\text{continue/no preceding failures}) = \pi_0 \)

\( P(\text{continue/any preceding failures}) = \pi_1 \)

Rule D allows the continuation probabilities to depend on whether or not there have been any failures so far. In our abortion example, if \( \pi_0 \) is greater than \( \pi_1 \), this might correspond to "once bitten, twice shy;" a woman who has been through an abortion may be more reluctant or less able to become pregnant again.

In any realistic situation, the stopping rule observed by the statistician will be some combination of the characteristics of human behavior and of time censoring induced by the sampling scheme. The stopping rules above make no attempt to distinguish between these, but rather are intended to subsume both of them.

The models for success and failure that will be considered are the following:

a) Model a assumes that all trials from all individuals are independent with a constant success probability \( p \).
b) Model b is a generalization of model a. Within individuals, all trials are independent with a constant success probability, but individuals are drawn from a two-point distribution: their associated success probabilities are either $p_1$ or $p_2$, with probabilities $\gamma$ and $1-\gamma$, respectively.

c) Model c is a generalization of model b; instead of individuals being drawn from a two-point distribution, they are drawn from an arbitrary distribution $F$.

d) Model d assumes that individuals are alike, but trials are different; the probability of success on the $i^{th}$ trial in a sequence is $p_i$. In particular, we will focus on the cases where the $p_i$'s form an increasing or decreasing sequence.

e) Model e is a one-step stationary Markov model; the probability of success on any trial depends on the result of the immediately preceding trial. All individuals are assumed to be identical and independent. The transition probabilities are:

\[
P(\text{success/preceding success}) = a \\
P(\text{success/preceding failure}) = b \\
P(\text{success on first trial}) = p = b/(1-a+b)
\]

The stationary assumption $p = ap + b(1-p)$ is made, yielding the expression for probability of success on the first trial given above.

The entire probability structure of the sequences $X_i$ is described when a stopping rule and a model have been specified. Throughout this paper, probability structures will be denoted by a
capital letter and a small letter, signifying stopping rule and model respectively. For example, B-a denotes the model where

\[
P(\text{success}) = p
\]
\[
P(\text{continue/success}) = q_1
\]
\[
P(\text{continue/failure}) = q_0
\].

The notation for these models is summarized below for reference.

**Stopping rules**

A) \( P(\text{continue}) = q \)

B) \( P(\text{continue/preceding success}) = q_1 \)
\[
P(\text{continue/preceding failure}) = q_0
\]

C) \( P(\text{continue/K preceding successes}) = q^K \) \( (= \alpha q^k) \)

D) \( P(\text{continue/no preceding failures}) = \tau_0 \)
\[
P(\text{continue/any preceding failures}) = \tau_1
\]

**Success-failure models**

a) \( P(\text{success}) = p \)

b) \( P(\text{success}) = \begin{cases} p_1 & \text{with probability } \gamma \\ p_2 & \text{with probability } 1-\gamma \end{cases} \)

c) \( P(\text{success}) = p \sim F \)

d) \( P(\text{success on trial i}) = p_i \)

e) \( P(\text{success/preceding success}) = a \)
\[
P(\text{success/preceding failure}) = b
\]
\[
P(\text{success on first trial}) = p = b/(1-a+b) \)
A. **Comparison of models**

The table at the end of this section gives an example comparing some of these models. The probabilities of various sequences were computed under model B-b with \((p_1, p_2, q_0, q_1) = (.6, .4, .2, .7, .5)\). Then the probabilities were recomputed under various other models, with parameters chosen so that some of the probabilities matched.

Model B-a has three parameters, so the probabilities for the sequences 0, 1 and 00 were matched, giving \( (p, q_0, q_1) = (.432, .704, .491) \). For model C-a the parameters were \( (p, q) = (.432, .704, .697) \) and for model D-a they were \( (p, q_0, q_1) = (.432, .491, .704) \). Model B-d has five parameters \( (p_1, p_2, p_3, q_0, q_1) \). After matching the same three probabilities, it was impossible to match simultaneously the fourth, fifth, or sixth, so the arbitrary choice \( p_1 = .5 \) was made. Matching the seventh sequence to get a value for \( p_3 \) gave \( (p_1, p_2, p_3, q_0, q_1) = (.5, .398, .389, .664, .560) \). For model B-e, matching the probabilities of the sequences 0, 1, 00 and 01 gave \( (p, q_0, q_1) = (.627, .390, .656, .570) \).

Notice that there is still a great deal of difference among the models. Many of the sequences show a two to three-fold difference in the possible probabilities associated with them.
<table>
<thead>
<tr>
<th>sequence</th>
<th>B-b</th>
<th>B-a</th>
<th>C-a</th>
<th>D-a</th>
<th>B-d</th>
<th>B-e</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.1680</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>1</td>
<td>.2200</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>00</td>
<td>.0672</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>01</td>
<td>.0840</td>
<td>.0880</td>
<td>.0880</td>
<td>.0511</td>
<td>.0581</td>
<td>*</td>
</tr>
<tr>
<td>10</td>
<td>.0360</td>
<td>.0356</td>
<td>.0613</td>
<td>.0356</td>
<td>.0567</td>
<td>.0362</td>
</tr>
<tr>
<td>11</td>
<td>.0500</td>
<td>.0466</td>
<td>.0602</td>
<td>.0466</td>
<td>.0245</td>
<td>.0787</td>
</tr>
<tr>
<td>000</td>
<td>.0273</td>
<td>.0269</td>
<td>.0269</td>
<td>.0269</td>
<td>*</td>
<td>.0269</td>
</tr>
<tr>
<td>001</td>
<td>.0329</td>
<td>.0352</td>
<td>.0352</td>
<td>.0204</td>
<td>.0227</td>
<td>.0172</td>
</tr>
<tr>
<td>010</td>
<td>.0141</td>
<td>.0142</td>
<td>.0245</td>
<td>.0204</td>
<td>.0152</td>
<td>.0091</td>
</tr>
<tr>
<td>100</td>
<td>.0141</td>
<td>.0142</td>
<td>.0171</td>
<td>.0142</td>
<td>.0230</td>
<td>.0145</td>
</tr>
<tr>
<td>011</td>
<td>.0185</td>
<td>.0187</td>
<td>.0241</td>
<td>.0155</td>
<td>.0126</td>
<td>.0300</td>
</tr>
<tr>
<td>101</td>
<td>.0185</td>
<td>.0187</td>
<td>.0168</td>
<td>.0108</td>
<td>.0191</td>
<td>.0120</td>
</tr>
<tr>
<td>110</td>
<td>.0079</td>
<td>.0076</td>
<td>.0117</td>
<td>.0076</td>
<td>.0128</td>
<td>.0129</td>
</tr>
<tr>
<td>111</td>
<td>.0118</td>
<td>.0099</td>
<td>.0103</td>
<td>.0099</td>
<td>.0107</td>
<td>.0281</td>
</tr>
<tr>
<td>all others</td>
<td>.2297</td>
<td>.2292</td>
<td>.1686</td>
<td>.2856</td>
<td>.2621</td>
<td>.1952</td>
</tr>
</tbody>
</table>

* indicates that the probabilities of these sequences were forced to match model B-b.
B. Calculations for stopping rule A

What follows are details of the calculations of sequence probabilities under different models. Distributions and moments of various summary statistics are also derived. For each stopping rule, various models will be considered. We will begin with the simplest.

1. Model a: \( P(\text{success}) = p \)

Consider the model A-a. Under this model, the distribution of the sequence from any individual is:

\[
P(X_i = (x_1, \ldots, x_n)) = \prod_{j=1}^{n} \frac{x_j}{p^j(1-p)} \left(\frac{1-x_j}{q}\right)^{n-1}(1-q) \quad \text{for} \quad x_i \in \{0, 1\}, \quad n = 1, 2, \ldots
\]

\[
= p^k(1-p)^{n-k} q^{n-1}(1-q)
\]

where \( k = \sum_{j=1}^{n} x_j \). Thus sufficient statistics for an individual are the pairs \((K_i, N_i)\), and their distribution is:

\[
P((K_i, N_i) = (k, n)) = \binom{n}{k} p^k(1-p)^{n-k} q^{n-1}(1-q) \quad \text{for} \quad k = 0, 1, \ldots, n
\]

\[
n = 1, 2, \ldots
\]

\( N_i \) has a geometric distribution, and \( K_i \), conditional on \( N_i \), has a binomial distribution.

If we now consider all of the individuals simultaneously, we obtain as the joint distribution:
\[
P((K_1, N_1) = (x_i, n_i) \text{ for } i = 1, 2, \ldots, M) \\
= \left[ \prod_{i=1}^{M} \binom{n_i}{k_i} \right] p^{\Sigma_{k_i}} (1-p)^{\Sigma_{n_i} - \Sigma_{k_i}} q^{\Sigma_{n_i} - M} \\
\quad \text{for } k_i = 0, 1, \ldots, n_i \quad n_i = 1, 2, \ldots \\
= \left[ \prod_{i=1}^{M} \binom{n_i}{k_i} \right] p^{M(t-s)} q^{M(t-l)} (1-q)^M
\]

where \( t = \sum_{i=1}^{M} n_i / M \) and \( s = \sum_{i=1}^{M} x_i / M \). \( S = \sum_{i=1}^{M} K_i / M \) and \( T = \sum_{i=1}^{M} N_i / M \) are sufficient for the entire problem, and their distribution is:

\[
P((S, T) = (s, t)) = \binom{M}{s} p^{Mt} q^{M(t-s)} (1-p)^{M(t-l)} (1-q)^M
\]

for \( s = 0, 1/M, \ldots, t \)
\[
t = 1, (M+1)/M, (M+2)/M, \ldots
\]

That is, MT has a negative binomial distribution, and MS, conditional on T, has a binomial distribution. To estimate the parameters of the model, we may use the maximum likelihood estimates \( \hat{p} = S/T \) and \( \hat{q} = (T-1)/T \).

Under this model, we have the following moments:

\[
E(K_i) = p/(1-q) \\
E(N_i) = 1/(1-q) \\
\text{Var } K_i = p(1-q) + p^2(2q-1)/(1-q)^2 \\
\text{Var } N_i = q/(1-q)^2 \\
\text{Cov}(K_i, N_i) = pq/(1-q)^2
\]
2. **Model C:** \( P(\text{success}) = p \sim F \)

Now consider the more general model A-c. The results given in the preceding section can be viewed as the conditional distributions given \( p \). Hence we simply need to integrate to get the unconditional probability of a sequence:

\[
P(X_i = (x_1, \ldots, x_n)) = \int p^k(1-p)^{n-k} \, dp \, q^{n-1}(1-q)
\]

\[
P((K_i, N_i) = (k,n)) = \binom{n}{k} \int p^k(1-p)^{n-k} \, dp \, q^{n-1}(1-q)
\]

Here the pairs \((K_i, N_i), i = 1,2,\ldots,M\), are sufficient. \( N_i \) still has a geometric distribution, but the distribution of \( K_i \) is more complex. Letting \( A_0 = \int p \, dF(p) \) and \( B_0 = \int p^2 \, dF(p) \), we have

\[
E K_i = A_0/(1-q)
\]

\[
E N_i = 1/(1-q)
\]

\[
\text{Var} K_i = A_0/(1-q) + 2B_0q/(1-q)^2 - A_0/(1-q)^2
\]

\[
\text{Var} N_i = q/(1-q)^2
\]

\[
\text{Cov}(K_i, N_i) = A_0q/(1-q)^2
\]

3. **Model d:** \( P(\text{success on } i^{\text{th}} \text{ trial}) = p_i \)

Thirdly, let us consider the model A-d. In this case, we have

\[
P(X_i = (x_1, \ldots, x_n)) = \left[ \prod_{i=1}^{n} p_i^{x_i}(1-p_i)^{1-x_i} \right] q^{n-1}(1-q)
\]
\[ P((k_i, N_i) = (k, n)) = \sum_{x_1, \ldots, x_n} \left[ \prod_{i=1}^{n} p_i^{x_i} (1-p_i)^{1-x_i} \right] q^{n-1}(1-q) \]

\[ = A(k, n) \cdot q^{n-1}(1-q) \]

where

\[ A(k, n) = \begin{cases} 
\sum_{x_1, \ldots, x_n} \left[ \prod_{i=1}^{n} p_i^{x_i} (1-p_i)^{1-x_i} \right] & \text{if } k = 0, 1, \ldots, n \\
1 & \text{if } k = n = 0 \\
0 & \text{otherwise} 
\end{cases} \]

We also have the recursion relations:

\[ A(k, n) = p_n A(k-1, n-1) + (1-p_n) A(k, n-1) \quad \text{for } n \geq 1 \]

\[ A(0, 0) = 1 \]

Notice also that

\[ \sum_{k=0}^{n} A(k, n) = \sum_{k=0}^{n} \left[ p_n A(k-1, n-1) + (1-p_n) A(k, n-1) \right] \]

\[ = p_n \sum_{k=0}^{n-1} A(k, n-1) + (1-p_n) \sum_{k=0}^{n-1} A(k, n-1) \]

\[ = \sum_{k=0}^{n-1} A(k, n-1) \]

By induction, we have

\[ \sum_{k=0}^{n} A(k, n) = \sum_{k=0}^{1} A(k, 1) = p_1 + (1-p_1) = 1 \]

Similarly,
\[
\sum_{k=0}^{n} kA(k,n) = \sum_{k=0}^{n} k(p_n A(k-1,n-1) + (1-p_n) A(k,n-1)) \\
= p_n \sum_{k=0}^{n} (k-1) A(k-1,n-1) + p_n \sum_{k=0}^{n} A(k-1,n-1) \\
+ (1-p_n) \sum_{k=0}^{n} kA(k,n-1) \\
= \sum_{k=0}^{n-1} kA(k,n-1) + p_n \\

\text{By induction,} \\
\sum_{k=0}^{n} kA(k,n) = p_1 + \cdots + p_n.
\]

This model is a special case of model B-d, and it follows from results given in Section II.C.3 that

\[
P(N_1 = n) = q^{n-1}(1-q)
\]

\[
P(K_1 = k) = \begin{cases} 
T(k-1) - T(k) & \text{if } k \geq 1 \\
1 - T(0) & \text{if } k = 0
\end{cases}
\]

where \( T(i) = \sum_{m=1}^{i} A(i,m) q^m p_{m+1} \). Thus

\[
E(K_1) = \sum_{k=0}^{\infty} T(k) \\
= \sum_{m=0}^{\infty} \sum_{k=0}^{m} A(k,m) q^m p_{m+1} \\
= \sum_{m=0}^{\infty} q^m p_{m+1}
\]
\[ E_{K_1}^2 = \sum_{k=1}^{\infty} k^2 [T(k-1) - T(k)] \]

\[ = T(0) + \sum_{k=1}^{\infty} (k+1)^2 T(k) - \sum_{k=1}^{\infty} k^2 T(k) \]

\[ = T(0) + \sum_{k=1}^{\infty} (2k+1) T(k) \]

\[ = \text{EK}_1 + 2 \sum_{k=1}^{\infty} k T(k) \]

\[ = \text{EK}_1 + 2 \sum_{m=1}^{\infty} q^m p_{m+1} \sum_{i=1}^{m} p_i \]

\[ \text{EK}_1 N_1 = \sum_{n=1}^{\infty} \sum_{k=0}^{n} k n A(k,n) q^{n-1} (1-q) \]

\[ = \sum_{n=1}^{\infty} n q^{n-1} (1-q) \sum_{k=1}^{n} p_k \]

\[ = \sum_{k=1}^{\infty} p_k \sum_{n=k}^{\infty} n q^{n-1} (1-q) \]

\[ = \sum_{k=1}^{\infty} p_k (1-q) \left[ \frac{q^k}{(1-q)^2} + k q^{k-1}/(1-q) \right] \]

\[ = q E_{K_1}/(1-q) + \sum_{k=1}^{\infty} k p_k q^{k-1} \]

\[ \text{Cov}(K_1, N_1) = \sum_{k=1}^{\infty} k p_k q^{k-1} - \text{EK}_1 \]

\[ = \sum_{k=1}^{\infty} (k-1) p_k q^{k-1} + \sum_{k=1}^{\infty} p_k q^{k-1} - \sum_{k=1}^{\infty} p_k q^{k-1} \]

\[ = \sum_{k=1}^{\infty} k p_{k+1} q^k \]
C. Calculations for stopping rule B

1. Model a: \( P(\text{success}) = p \)

Now consider the more complicated stopping rule, where the probability of continuation depends on the result of the immediately preceding trial. Let us begin with model B-a. Here we have:

\[
P(X_i = (x_1, \ldots, x_n)) = \prod_{j=1}^{n} \frac{x_j}{p_j} \left(1-p_j\right)^{1-x_j} \prod_{j=1}^{n-1} \left[q_j q_0^j \left(1-q_j\right)^{n-1} \left(1-q_0\right)^x\right]^{1-x_n}
\]

for \( x_j \in \{0,1\}, n = 1, 2, \ldots \)

\[
= p^n \left(1-p\right)^{n-k} q_1^{k-x} q_0^{n-1-k+x} \left(1-q_1\right)^x \left(1-q_0\right)^{1-x_n}
\]

Thus the sufficient statistic for an individual is \((K_i, N_i, X_i N_i)\), and its distribution is:

\[
P((K_i, N_i, X_i N_i) = (k, n, x)) = \binom{n-1}{k-x} p^n \left(1-p\right)^{n-k} q_1^{k-x} q_0^{n-1-k+x} \left(1-q_1\right)^x \left(1-q_0\right)^{1-x}
\]

for \( x \in \{0,1\}; k = x, x+1, \ldots, n; n = 1, 2, \ldots \),

where we use the convention \( \binom{0}{0} = 1 \). If we observe the sequences for \( M \) individuals, the sufficient statistic is

\[
(M \sum_{i=1}^{M} K_i, M \sum_{i=1}^{M} N_i, M \sum_{i=1}^{M} X_i N_i) = (MS, MT, ML),
\]

where

\[
L = \sum_{i=1}^{M} X_i N_i / M.
\]
The maximum likelihood estimates of the parameters are:

\[ \hat{p} = \frac{S}{T} \]
\[ \hat{q}_1 = \frac{(S-L)}{S} \]
\[ \hat{q}_0 = \frac{(T-L-S+L)}{(T-S)} \]

The marginal distributions of the sufficient statistics are:

\[ P((K_i, N_i) = (k,n)) = p^{(1-p)^{n-k}} q_1^{k-1} q_0^{n-k-1} \left( \binom{n-1}{k} q_1(1-q_0) + \binom{n-1}{k-1} q_0(1-q_1) \right) \]

for \( k = 0,1,\ldots,n; \ n = 1,2,\ldots \).

Hence \( N_i \) has a geometric distribution with parameter \( v = pq_1 + (1-p)q_0 \):

\[ P(N_i = n) = v^{n-1}(1-v) \]

\[ E[N_i] = \frac{1}{(1-v)} \]

\[ Var[N_i] = \frac{1}{(1-v)^2} \]

Let \( w = \frac{pq_1}{(1-q_0(1-p))} \). Then we can easily derive:

\[ P(K_i = k) = \begin{cases} 
  w^{k-1}/q_1 & \text{if } k \neq 0 \\
  1 - w/q_1 & \text{if } k = 0 
\end{cases} \]

Hence we have
\[
E_{K_i} = \frac{w}{q_1(1-w)} = p/(1-v)
\]

\[
\text{Var } K_i = E_{K_i} + (2q_1 - 1)E_{K_i}^2
\]

\[
\text{Cov}(K_i, N_i) = pq_1/(1-v)^2
\]

\(X_{iN_1}\) is a Bernoulli variate with

\[
P(X_{iN_1} = 1) = \frac{p(1-q_1)}{(1-v)}
\]

\[
= 1 - P(X_{iN_1} = 0)
\]

2. Model c: \(P(\text{success}) = p \sim F\)

Consider the more general model B-c, where the fixed success probability \(p\) for a sequence has the distribution \(F\) over sequences. Again, integration of the preceding results gives the probabilities and expectations for this case. Thus, for example:

\[
E_T = E_{N_1} = \int (1-v)^{-1} dF(p)
\]

\[
= \int \sum_{\alpha=0}^{\infty} v^\alpha dF(p)
\]

\[
= \sum_{\alpha=0}^{\infty} \int v^\alpha dF(p)
\]
\[ \text{ES} = \text{EK}_i = \int p(1-v)^{-1} \, d\mathcal{F}(p) \]
\[ = \int \frac{v - q_0}{(q_1 - q_0)(1-v)} \, d\mathcal{F}(p) \]
\[ = \frac{1}{q_1 - q_0} \left[ \int \frac{1}{1-v} \, d\mathcal{F}(p) - q_0 \int \frac{1}{1-v} \, d\mathcal{F}(p) \right] \]
\[ = \frac{1}{q_1 - q_0} [\text{ET}(1-q_0) - 1] \]

\[ \text{EL} = \text{E}_{iN_1} = (1-q_1) \, \text{ES}. \]

Model B-b is the special case where \( \mathcal{F} \) is a two-point distribution, i.e.,
\[ p = \begin{cases} 
  p_1 & \text{with probability } \gamma \\
  p_2 & \text{with probability } 1-\gamma
\end{cases} \]

Let \( v_i = p_1 q_1 + (1-p_1) q_0 \) for \( i = 1, 2 \). Then, for example,
\[ \int v^i \, d\mathcal{F}(p) = \gamma v_1^i + (1-\gamma) v_2^i \]
\[ \text{ET} = \gamma/(1-v_1) + (1-\gamma)/(1-v_2). \]

3. Model d: \( P(\text{success on } i^{\text{th}} \text{ trial}) = p_i \)

Next, let us consider the model B-d, where the probability of success varies from trial to trial. Here
\[ P(X_i = (x_1, \ldots, x_n)) = \prod_{j=1}^{n} \left[ p_j^j (1-p_j)^{1-x_j} \right] \prod_{j=1}^{n-l} \left[ q_0^j q_0^{n-l-x_j} \right] (1-q_1)^x (1-q_0)^{1-x_n} \]

Again let

\[ A(k, n) = \begin{cases} 
\sum_{x=1}^{n} \prod_{i=1}^{x} p_i (1-p_i) x_i^{1-x_i} & \text{if } k = 0, 1, \ldots, n \\
0 & \text{if } n = 1, 2, \ldots,
\end{cases} \]

Recall that we have the recursion relations:

\[ A(k, n) = p_n A(k-l, n-l) + (1-p_n) A(k, n-l) \]
\[ A(0, 0) = 1 \]

Then

\[ P((K, N_i, X = x) = (k, n, x)) = A(k-x, n-l) p_n^{x} (1-p_n)^{1-x} q_0^{k-x} q_0^{n-l-k+x} (1-q_1)^x (1-q_0)^{1-x} \]

\[ P((K, N_i) = (k, n)) = A(k, n-l) (1-p_n) q_0^{n-l-k} (1-q_0) + A(k-l, n-l) p_n q_1^{k-1} q_0^{n-k} (1-q_1) \]

\[ P(N_i = n) = \sum_{k=0}^{n-l} A(k, n-l) (1-p_n) q_0^{n-l-k} (1-q_0) + \sum_{k=1}^{n} A(k-l, n-l) p_n q_1^{k-1} q_0^{n-k} (1-q_1) \]
\[ \sum_{k=0}^{n-1} A(k, n-1) q_1^k q_0^{n-1-k} \left( (1-q_1) p_n + (1-q_0) p_n \right) \]

\[ = S(q_1/q_0, n-1) q_0^{n-1} (1-v_n) \]

where \( v_i = p_i q_1 + (1-p_i) q_0 \) and \( S(b,n) = \sum_{i=0}^{n} A(i,n)b^i \). Note that

\[ S(b,n) = \sum_{i=0}^{n} A(i,n)b^i \]

\[ = \sum_{i=0}^{n} \left[ p_n A(i-1, n-1) + (1-p_n) A(i, n-1) \right] b^i \]

\[ = (1-p_n + p_n b) S(b, n-1) \]

\[ = \Pi_{i=1}^{n} \left( 1-p_i + p_i b \right) \quad \text{for } n \geq 1. \]

Furthermore, \( S(b,0) = 1 \). Thus,

\[ P(N_i = n) = q_0^{n-1} (1-v_n) \prod_{i=1}^{n-1} \left( 1-p_i + p_i q_1/q_0 \right) \]

\[ = (1-v_n) \prod_{i=1}^{n-1} v_i \]

and thus

\[ \mathbb{E}N_i = \sum_{n=1}^{\infty} \prod_{i=1}^{n-1} v_i. \]

\( X_{iN_i} \) is, as before, a Bernoulli variate, and its distribution is:
\[ P(X_{iN_i} = 0) = \sum_{n=1}^{\infty} \sum_{k=0}^{n-1} A(k, n-1)(1-p_n) q_0^{n-k} (1-q_0) \]

\[ = \sum_{n=1}^{\infty} (1-p_1) q_0^{n-1} (1-q_0) S(q_{1}/q_0, n-1) \]

\[ = \sum_{n=1}^{\infty} (1-q_0)(1-p_1) \prod_{i=1}^{n-1} v_i \]

\[ = \sum_{n=1}^{\infty} (1-q_0)(q_1 - v_n)^{-1} \prod_{i=1}^{n-1} v_i \]

\[ = (1-q_0)(q_1 - q_0)^{-1} \left[ 1 - EN_i(1-q_0) \right] \]

\[ P(X_{iN_i} = 1) = (1-q_1)(q_1 - q_0)^{-1} \left[ EN_i(1-q_0) - 1 \right]. \]

To obtain the distribution of \( K_1 \), define

\[ R(i) = \sum_{m=1}^{\infty} A(i, m) q_0^m \]

\[ T(i) = \sum_{m=1}^{\infty} A(i, m) q_0^m p_{m+1} \]

Note that

\[ R(i) = \sum_{m=1}^{\infty} [p_m A(i-1, m-1) + (1-p_m) A(i, m-1)] q_0^m \]

\[ = q_0 T(i-1) + q_0 R(i) - q_0 T(i) \]

Thus

\[ R(i) = q_0 (1-q_0)^{-1} [T(i-1) - T(i)] \]

Now
\[ P(K_i = k) = \sum_{n=k+1}^{\infty} A(k, n-1) (1-p_n) q_1^{n-1-k} q_0^{k-n+1-k} (1-q_0) \\
+ \sum_{n=k}^{\infty} A(k-1, n-1) p_n q_1^{n-k} q_0^{k-n} (1-q_1) \\
= R(k)(q_1/q_0)^k (1-q_0) - (q_1/q_0)^k (1-q_0) T(k) \\
+ (q_1/q_0)^{k-1} (1-q_1) T(k-1) \\
= (q_1/q_0)^{k-1} T(k-1) - (q_1/q_0)^k T(k) \quad \text{for } k \geq 1 \\
\]

Also
\[ P(K_i = 0) = \sum_{n=1}^{\infty} \left[ \prod_{i=1}^{n-1} (1-p_i) \right] q_0^{n-1}(1-q_0) \\
= \sum_{n=1}^{\infty} A(0, n) q_0^{n-1}(1-q_0) \\
= (1-p_1)(1-q_0) + \sum_{n=2}^{\infty} A(0, n-1) q_0^{n-1}(1-q_0)(1-p_n) \\
= (1-p_1)(1-q_0) + q_0 P(K_i = 0) - (1-q_0)[T(0) - p_1]. \]

Thus \( P(K_i = 0) = 1 - T(0) \). Finally,
\[ EK_i = \sum_{k=1}^{\infty} P(K_i \geq k) = \sum_{k=1}^{\infty} (q_1/q_0)^{k-1} T(k-1) \\
= \sum_{k=0}^{\infty} (q_1/q_0)^k \sum_{m=k}^{\infty} A(k, m) q_0^m p_{n+1} = \sum_{m=0}^{m} q_0^m p_{n+1} S(q_1/q_0, m) \\
= \sum_{m=0}^{\infty} p_{m+1} \prod_{i=1}^{m} v_i \\
= \sum_{m=0}^{\infty} (v_{m+1} - q_0)(q_1 - q_0)^{-1} \prod_{i=1}^{m} v_i \\
= \frac{E_{N_1}(1-q_0) - 1}{q_1 - q_0} \]

\[ 25 \]
Notice that we have the same relationship among the three expectations here as in model B-c.

4. Model e: Markov case

Finally, let us consider the stationary Markov case B-e:

\[ P(\text{continue/success}) = q_1 \]
\[ P(\text{continue/failure}) = q_0 \]
\[ P(\text{success/success}) = a \]
\[ P(\text{success/failure}) = b \]

\[ P(\text{success on first trial}) = p = ap + b(1-p) . \]

Thus we have:

\[
P(X_1 = (x_1, \ldots, x_n))
\]
\[ = p^{1-p} \prod_{i=2}^{n-1} q_{i-1}^{x_{i-1}} (1-a)^{1-x_i} a^{1-x_i} (1-a)^{1-x_i} (1-b)^{1-x_i} (1-b)^{1-x_i} \]
\[ \times \prod_{i=1}^{n-1} q_{i}^{q_{i}} (1-q_{i})^{1-x_{i}} (1-q_{i})^{1-x_{i}} \]
\[ = p^{1-p} \prod_{i=1}^{n-1} a^{1-x_i} n_{11}(1-a)^{n_{01}} n_{00}^{n_{10}} n_{10}^{n_{01}} n_{01}^{n_{00}} (1-q_{i})^{1-x_{i}} (1-q_{i})^{1-x_{i}} \]

where \( n_{i,j} \) is the number of transitions from \( i \) to \( j \) in the sequence.

Note that \( x_1 + n_{11} + n_{01} = n_{10} + n_{11} + x_n = k \). Also
\[ n_{00} + n_{01} + n_{10} + n_{11} = n-1. \]
In deriving the distribution of $K_1$, general formulae are difficult to obtain. Let us proceed by cases.

\[
P(K_1 = 0) = \sum_{n=1}^{\infty} (1-p)(1-b)^{n-1} q_0^{n-1} (1-q_0)
\]

\[
= \frac{(1-p)(1-q_0)}{1-q_0(1-b)}
\]

\[
= 1 - H
\]

where we make the definitions:

\[
H = p(1 + R q_0)
\]

\[
R = (1-a)/(1-\lambda)
\]

\[
\lambda = q_0(1-b)
\]

\[
F = q_1 - R q_1 (1-q_0) = q_1 (a + b q_0 R)
\]

Now

\[
P(K_1 = 1) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} P(0 \cdots 0 \ 1 \ 0 \cdots 0)
\]

\[
= P(1) + \sum_{m=1}^{\infty} P(1 \ 0 \cdots 0) + \sum_{n=1}^{\infty} P(0 \cdots 0 \ 1)
\]

\[
+ \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} P(0 \cdots 0 \ 1 \ 0 \cdots 0)
\]

\[
= p(1-q_1) + \sum_{m=1}^{\infty} pq_1 (1-a)[q_0(1-b)]^{m-1} (1-q_0)
\]

\[
+ \sum_{n=1}^{\infty} (1-p) q_0 [(1-b)q_0]^{n-1} b(1-q_1)
\]
\[
\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} (1-p)q_0 [(1-b)q_0]^{n-1} b q_1 (1-a) [q_0 (1-b)]^{m-1} (1-q_0)
\]

\[
P(1-q_1) + pq_1 (1-a)(1-q_0)/(1-\lambda) + (1-p)q_0 b (1-q_1)/(1-\lambda)
\]

\[
+ (1-p)q_0^2 b q_1 (1-a)(1-q_0)/(1-\lambda)^2
\]

\[
= p(1-q_1) + pR[q_1 (1-q_0) + q_0 (1-q_1)] + pR^2 q_0 q_1 (1-q_0)
\]

\[
= p(1 + R q_0)(1-q_1 + R q_1 (1-q_0))
\]

\[
= H(1-F).
\]

Similar laborious calculations give:

\[P(K_i=2) = H(1-F)F\]

\[P(K_i=3) = H(1-F)F^2\]

\[P(K_i=4) = H(1-F)F^3\]

General expressions are elusive, but it seems reasonable to conclude that the distribution of \(K_i\) in this case is the same as in model B-a, with \(w\) corresponding to \(F\), and \(q_1\) corresponding to \(F/H\).

If so, then we have:

\[P(K_i=k) = H(1-F)F^{k-1}\]

\[E K_i = H/(1-F)\]

D. Calculations for stopping rule C.

Now we turn to the third stopping rule, where the probability of continuation depends on the number of previous successes. Consider the simplest case (model C-a):
\[ P(\text{continue/k successes}) = q_k \quad k = 0, 1, 2, \ldots \]

\[ P(\text{success}) = p \]

Here

\[ P(X_i = (x_1, \ldots, x_n)) = \prod_{i=1}^{n-1} \left[ p^{i-1} (1-p) \right] q^{x_1+\ldots+x_i}_i \]

\[ = p^{k-1} (1-p)^{n-k} \prod_{i=1}^{n-1} q_{T_i} \]

where \( T_i = \sum_{j=1}^{i} x_i \). Let

\[ Q(k, n) = \begin{cases} \sum_{x_1=1}^{n-1} \prod_{i=1}^{n} q_{T_i} & \text{if } k = 0, \ldots, n \\ \sum_{x=k}^{n} q_{x} & n = 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases} \]

Then \( P(K_1, N_1 = (k, n)) = p^{k-1} (1-p)^{n-k} q_{k} Q(k, n) \). Note we have the recursion relations:

\[ Q(k, n) = q_{k-1} Q(k-1, n-1) + q_{k} Q(k, n-1) \]

\[ Q(0, 1) = Q(1, 1) = 1. \]

Let us obtain the marginal distribution of \( K_1 \):

\[ P(K_1 = 0) = \sum_{n=1}^{\infty} (1-p)^n q_{0}^{n-1} (1-q_{0}) \]

\[ = (1-p)(1-q_{0})/(1 - q_{0}(1-p)) \]

\[ = 1 - p/R_0 \]
where we define  \( R_i = 1 - q_i(1-p) \). For  \( k \neq 0 \),

\[
P(K_i=k) = \sum_{n=k}^{\infty} p^{k}(1-p)^{n-k}(1-q_k)Q(k,n)
\]

\[
= p^{k}(1-q_k)Q(k,k) + p^{k}(1-q_k)\sum_{n=k+1}^{\infty}(1-p)^{n-k}
\]

\[
\times [q_{k-1}Q(k-1,n-1) + q_kQ(k,n-1)]
\]

\[
= p^{k}(1-q_k)Q(k,k) + p^{k}(1-q_k)\sum_{m=k}^{\infty}(1-p)^{m-k+1}q_{k-1}Q(k-1,m)
\]

\[
+ p^{k}(1-q_k)\sum_{m=k}^{\infty}(1-p)^{m-k+1}q_kQ(k,m)
\]

Note that

\[
P(K_i=k-1) = \sum_{n=k-1}^{\infty} p^{k-1}(1-p)^{n-k+1}(1-q_{k-1})Q(k-1,n)
\]

\[
= p^{k-1}(1-q_{k-1})Q(k-1,k-1)
\]

\[
+ \sum_{n=k}^{\infty} p^{k-1}(1-p)^{n-k+1}(1-q_{k-1})Q(k-1,n)
\]

Solving this relation and substituting into the previous equation gives:

\[
P(K_i=k)
\]

\[
= p^{k}(1-q_k)Q(k,k) + p^{k}(1-q_{k-1})q_{k-1}\left[ \frac{P(K_i=k-1)}{p^{k-1}(1-q_{k-1})} - Q(k-1,k-1) \right]
\]

\[
+ (1-p)q_k P(K_i=k).
\]

Noting that  \( Q(k,k) = q_{k-1}Q(k-1,k-1) \) gives:
\[ P(K_i=k) = \frac{p(1-q_i)q_{k-1}P(K_i=k-1)}{R_k(1-q_{k-1})} \]

\[ = \frac{pq_{k-1}(1-q_i)}{R_k(1-q_{k-1})} \cdot \frac{pq_{k-2}(1-q_{k-1})}{R_{k-1}(1-q_{k-2})} \cdot P(K_i=k-2) \]

\[ = \frac{p^{k-1} \prod_{i=1}^{k-1} q_i (1-q_i) P(K_i=1)}{\prod_{i=2}^{k} \prod_{i=1}^{k} R_i (1-q_i)} \]

But similar calculations give:

\[ P(K_i=1) = \frac{p(1-q_i)}{R_0 R_1} = \frac{p(1-c_i)}{R_0} \]

where \( c_k = \frac{p q_k}{R_k} \). Hence

\[ P(K_i=k) = p^{k-1} \prod_{i=1}^{k-1} q_i (1-q_k) / \prod_{i=1}^{k-1} R_i \]

for \( k \geq 1 \)

\[ = p \prod_{i=1}^{k-1} c_i (1-c_k) / R_0 \]

\[ P(K_i \geq k) = p \prod_{i=1}^{k-1} c_i / R_0 \]

Thus

\[ E K_i = p \sum_{k=1}^{\infty} \prod_{i=1}^{k-1} c_i / R_0 \]

\[ = p / R_0 \left( 1 + c_1 + c_1 c_2 + c_1 c_2 c_3 + \cdots \right) \]

In the special case where \( q_k = c q^k \), we may obtain the distribution of \( N_1 \) as follows:
Let

\[ B(n, a) = \sum_{k=0}^{n} a^k q(k, n) \]

\[ = \sum_{k=0}^{n} a^k [q_{k-1} Q(k-1, n-1) + q_k Q(k, n-1)] \]

\[ = \alpha a B(n-1, a) + \alpha B(n-1, aq) \]

\[ = \alpha (1 + a) B(n-1, aq) \]

\[ = \alpha(1 + a) \alpha(1 + qa) B(n-2, aq^2) \]

\[ = \alpha^n \prod_{i=1}^{n-1} (1 + aq^{i-1}) B(1, aq^{n-1}) \]

\[ = \alpha^n \prod_{i=0}^{n-1} (1 + aq^i) \]

\[ P(N_1 = n) = \sum_{k=0}^{n} p^k (1-p)^{n-k} (1-q_k) Q(k, n) \]

\[ = (1-p)^n \left[ B(n, \frac{p}{1-p}) - \alpha B(n, \frac{pq}{1-p}) \right] \]

\[ = (1-p)^n \alpha^{n-1} \left[ \prod_{i=0}^{n-1} \left( 1 + \frac{pq}{1-p} \right) \right] - \alpha \prod_{i=0}^{n-1} \left( 1 + \frac{pq}{1-p} \right) \]

\[ = \alpha^{n-1} \prod_{i=1}^{n-1} (1-p+pq^i) - \alpha^n \prod_{i=1}^{n} (1-p+pq^i) \]

\[ P(N_1 \geq n) = \alpha^{n-1} \prod_{i=1}^{n-1} (1-p+pq^i) \]

\[ E N_1 = \sum_{n=1}^{\infty} \alpha^{n-1} \prod_{i=1}^{n-1} (1-p+pq^i) \]

\[ = 1 + \sum_{n=1}^{\infty} \prod_{i=1}^{n} v_i \]

\[ = 1 + v_1 + v_1 v_2 + v_1 v_2 v_3 + \cdots \]
where \( v_1 = \alpha(1-p) + \alpha q^i \).

We would like to show that \( E_k = p^{\infty} \). Let

\[
x = R_0(1 + v_1 + v_1 v_2 + \ldots).
\]

It suffices to show that \( x = 1 + c_1 + c_1 c_2 + \ldots \).

\[
x = R_0(1 + v_1 + \sum_{i=2}^{\infty} v_1 \ldots v_i)
\]

\[
= (1 - \alpha(1-p)) (1 + v_1 + \sum_{i=2}^{\infty} v_1 \ldots v_i)
\]

\[
= 1 + v_1 + \sum_{i=2}^{\infty} v_1 \ldots v_i - \alpha(1-p) - \alpha(1-p) \sum_{i=1}^{\infty} v_1 \ldots v_i
\]

\[
= 1 + v_1 + \sum_{i=2}^{\infty} v_1 \ldots v_{i-1}(\alpha(1-p) + \alpha q^i)
\]

\[
- \alpha(1-p) - \alpha(1-p) \sum_{i=1}^{\infty} v_1 \ldots v_i
\]

\[
= 1 + v_1 + \alpha p \sum_{i=2}^{\infty} v_1 \ldots v_{i-1} q^i - \alpha(1-p)
\]

\[
= 1 + \alpha p q + \alpha p \sum_{i=1}^{\infty} v_1 \ldots v_i q^{i+1}
\]

\[
= 1 + \alpha p q(1 + \sum_{i=1}^{\infty} v_1 \ldots v_i q^i)
\]

\[
= 1 + c_1(1 - \alpha q(1-p))(1 + \sum_{i=1}^{\infty} v_1 \ldots v_i q^i).
\]

Define \( f(n) = (1 - \alpha q^{n-1}(1-p))(1 + \sum_{i=1}^{\infty} v_1 \ldots v_i q^{(n-1)i}) \). Then we have shown \( x = f(1) = 1 + c_1 f(2) \). The same argument shows, in general, that \( f(n) = 1 + c_n f(n+1) \). Thus
\[ x = f(1) \]
\[ = 1 + c_1 f(2) \]
\[ = 1 + c_1 + c_1 c_2 f(3) \]
\[ = 1 + \sum_{i=1}^{n} c_1 \ldots c_i + c_1 c_2 \ldots c_{n+1} f(n+2) \]

\[ |x - (1 + \sum_{i=1}^{n} c_1 \ldots c_i)| = |c_1 \ldots c_{n+1} f(n+2) - \sum_{i=n+1}^{\infty} c_1 \ldots c_i| \]
\[ \leq |c_1 \ldots c_{n+1}| |f(n+2)| + |\sum_{i=n+1}^{\infty} c_1 \ldots c_i| \]
\[ \leq c_1^n |f(n+2)| + c_1^{n+1} / (1-c_1) \]

since the c's form a decreasing sequence. Since the v's decrease also,

\[ \sum_{i=1}^{\infty} v_1 \ldots v_i q^{(n-1)i} \leq \sum_{i=1}^{\infty} v_1 q^{(n-1)i} \]
\[ = v_1 q^{n-1} / (1-v_1 q^{n-1}) \to 0 \]

Thus \( f(n) \to 1 \) as \( n \to \infty \). Therefore the difference between \( x \) and \( 1 + \sum_{i=1}^{\infty} c_1 \ldots c_i \) can be shown to be arbitrarily small, and the claim has been established.

Finally, to get the distribution of \( X_{iN_1} \), we observe, in the special case \( q_k = aq^k \):
\[ \text{EX}_{1N_1} = P(X_{1N_1} = 1) \]

\[ = P(1) + \sum_{n=2}^{\infty} \sum_{x_1 \ldots x_{n-1}} P(x_1 \ldots x_{n-1}) \]

\[ = P(1) + \sum_{n=2}^{\infty} \sum_{k=1}^{n} \sum_{x_1 \ldots x_{n-1}} P(k = k) Q(k-1, n-1) \]

\[ = p(1-q_1) + \sum_{n=2}^{\infty} \sum_{k=1}^{n} p^k (1-p)^{n-k} q_{k-1} (1-q_k) Q(k-1, n-1) \]

\[ = p(1-\alpha q) + \sum_{n=2}^{\infty} \alpha p (1-p)^{n-1} \left\{ B(n-1, pq/(1-p)) - \alpha q B(n-1, pq^2/(1-p)) \right\} \]

\[ = p(1-\alpha q) + \sum_{n=2}^{\infty} \alpha p (1-p)^{n-1} \left\{ \prod_{i=0}^{n-2} (1 + \frac{pq}{1-p})^{i+1} \right\} \]

\[ = p(l-\alpha q) + \sum_{n=2}^{\infty} \frac{\prod_{i=1}^{n-1} (1 + \alpha q) \prod_{i=2}^{n} v_i}{\prod_{i=2}^{n} v_i} \]

\[ = p(l-\alpha q) + p(EN_1 - l) - \frac{p\alpha q(EN_1 - l - v_1)}{v_1} \]

\[ = \frac{pEN_1(1-p)(1-q) + pq}{1 - p + pq} \]
E. Calculations for stopping rule D

Finally, consider the stopping rule where the probability of continuation depends on whether or not there have been any failures to that point. We will consider only the simplest case, D-a:

\[ P(\text{continue/no failures}) = \pi_0 \]
\[ P(\text{continue/any failures}) = \pi_1 \]
\[ P(\text{success}) = p \]

Then

\[ P(X_1=(x_1,\ldots,x_n)) = \begin{cases} 
  p^{k(1-p)} \pi_0^{-k} \pi_1^{m-l} \pi_0^{-m-l} (1-\pi_1) & \text{if } m \leq n \\
  p^{n} \pi_0^{-n} (1-\pi_0) & \text{if } m = n+1
\end{cases} \]

where

\[ m = \begin{cases} 
  \inf\{ i \in \{1,\ldots,n\} : \prod_{j=1}^{i} x_j = 0 \} & \text{if } \prod_{j=1}^{n} x_j = 0 \\
  n+1 & \text{if } \prod_{j=1}^{n} x_j = 1
\end{cases} \]

i.e., \( m \) is the time of the first failure.

Simple algebra gives the following:

\[ P(N_1=n) = \frac{(1-p)(1-\pi_1)\pi_1^n}{\pi_1 - p\pi_0} + \frac{(\pi_1 - \pi_0)(1-p\pi_0)(p\pi_0)^n}{\pi_0(\pi_1 - p\pi_0)} \]

\[ EN_1 = \frac{1 - p\pi_1}{(1-\pi_1)(1-p\pi_0)} \]
\[ P(K_1 = k) = \begin{cases} 
\frac{(1-p)(1-p_{11})p_{11}}{D} \left( \frac{p_{11}}{D} \right)^k + \frac{(p_{11}-p_{10})(1-p_{10})}{p_{10}(p_{11}-p_{10}D)} \left( p_{10} \right)^k & \text{for } k \neq 0 \\
(1-p_{11})(1-p)/D & \text{for } k = 0 
\end{cases} \]

\[ EK_1 = pE_{N_1} \]

\[ P(X_{i_{N_1}} = 1) = 1 - (1-p)(1-p_{11})E_{N_1} = 1 - P(X_{i_{N_1}} = 0) \]

where \[ D = 1 - p_{11}(1-p) \].
III. CONDITIONAL TESTS

Suppose that we are faced with a data set of the structure that has been described above, and we wish to test the fit of some model. Two basic ways of dealing with the stopping rule are possible. One is to try to take it into account by modeling. Another is to devise procedures which are independent of the stopping rule. The latter possibility is more attractive if the interest is focused on the values of the parameters measuring risk of abortion and not on demographic properties such as expected family size. However, it is difficult to achieve tests of fit that have properties that do not depend on the stopping rule.

One scheme which achieves the independence is to look at each trial in the sequence conditioned on the results of the previous trials and on the decision to continue. Thus, under any of the models, all first trials have a constant success probability, independent of the stopping rule. All second trials, conditional on the first trial being a success, have another constant success probability, while second trials conditional on the first being a failure have a possibly different success probability. Both, however, are independent of the stopping rule. Thus, for example, under model b (i.e., a two-point distribution of success probabilities), we have:
\[ P(1) = \gamma p_1 + (1-\gamma)p_2 \]
\[ P(1/0) = \frac{\gamma p_1 (1-p_1) + (1-\gamma) p_2 (1-p_2)}{\gamma (1-p_1) + (1-\gamma) (1-p_2)} \]
\[ P(1/1) = \frac{\gamma p_1^2 + (1-\gamma)p_2^2}{\gamma p_1 + (1-\gamma)p_2} \]

regardless of the stopping rule, where \( P(1/0) \) is an abbreviation for the probability that the second trial is a success, given that the first trial was a failure and that a second trial was made, and the other terms are defined similarly. Define the statistics \( n_0, n_1, n_{0/0}', n_{1/0}', n_{0/1}', n_{1/1}', \ldots \), as the number of first trials that failed, the number of first trials that succeeded, the number of second trials that failed among all those whose first trial was a failure, etc.

Then, given the total \( n_0 + n_1 \), \( n_1 \) has a binomial distribution \( B(n_0 + n_1, P(1)) \). Given \( n_0 \) and \( n_1 \), \( n_{0/0}', n_{1/0}', n_{0/1}', n_{1/1}' \) are independent with respective binomial distributions \( B(n_0, P(1/0)) \) and \( B(n_1, P(1/1)) \).

Comparing any of these statistics to their conditional expectations certainly provides a test of the fit of the model. Unfortunately, these statistics are dependent, so the individual tests cannot be combined in any simple way into an overall test of the fit of the model.

The dependence arises because the same individuals may enter into more than one of the statistics; i.e., each individual contributes multiple trials. Thus, independence of the statistics can be achieved by using only one trial from each individual. However, besides the problem of loss of information, there is considerable difficulty in
finding a scheme for picking trials which will simultaneously allow non-trivial power and be independent of the stopping rule. For example, picking the first trial of each sequence will not work; these trials contribute to the estimation of the parameters of the various models, but cannot be used to discriminate between them. Picking the second trial of each involves the difficulty that not all sequences have a second trial, and any scheme which chooses the first or second randomly has the same defect. Picking the last trial of each will give discriminatory power between various models, but the distribution of successes on last trials is not independent of the stopping rule. For example, under model B-a, the probability that the last trial is a success is not p, but \[ \frac{p(1-q)}{p(1-q) + (1-p)(1-q)} \]. (In this particular model, the probability is independent of the results of the previous trials.) Thus any scheme which picks trials by choosing the last trial or otherwise counting back from the end of the sequence will involve the stopping rule. It is difficult if not impossible to find a scheme for picking one trial from each individual which avoids all these pitfalls. Thus the statistics \( n_0, n_1 \), etc., while useful descriptive statistics, seem to have limited utility for any test of fit of a model.
IV. USE OF NATURAL STATISTICS

What are some of the possible statistics we might use to test the fit of a model? The natural descriptive statistics for each sequence are the number of trials, \(N_i\), and the number of successes, \(K_i\). Can these be used to distinguish models? Suppose that we wish to test the hypothesis that the probability of success is constant across individuals and across trials (i.e., model a). Let us first consider the case where we know the form of the stopping rule, and in fact it is the simplest, rule A, where the probability of stopping is independent of the results of previous trials. For simplicity, let us confine our attention to the moments of \(K_i\) and \(N_i\). For each individual \(i\), denote the mean of the vector \((K_i, N_i)\) by \(\mu\) and the covariance by \(\Sigma\). Consider first the null hypothesis, model A-a.

Then we have already shown:

\[
\mu_1 = \frac{p}{(1-q)}
\]

\[
\mu_2 = \frac{1}{(1-q)}
\]

\[
\sigma_{11} = \frac{p}{(1-q)} + \frac{p^2(2q-1)}{(1-q)^2}
\]

\[
\sigma_{12} = \frac{pq}{(1-q)^2}
\]

\[
\sigma_{22} = \frac{q}{(1-q)^2}
\]

Note that the values of \(\mu_1\) and \(\mu_2\) are not fixed by the hypotheses.
but that, given \( \mu \), the covariance matrix is determined, and is given by the functions \( \sigma_{ij}^* \) defined as follows:

\[
\begin{align*}
\sigma_{11}^*(\mu) &= \mu_1 - \mu_1^2 + 2\mu_1(\mu_2 - 1)/\mu_2 \\
\sigma_{12}^*(\mu) &= \mu_1(\mu_2 - 1) \\
\sigma_{22}^*(\mu) &= \mu_2(\mu_2 - 1).
\end{align*}
\]

Now consider the model A-c; here we have

\[
\begin{align*}
\mu_1 &= A_0/(1-q) \\
\mu_2 &= 1/(1-q) \\
\sigma_{11} &= A_0/(1-q) - A_0^2/(1-q)^2 + 2A_0 q/(1-q)^2 \\
&= \sigma_{11}^*(\mu) + 2q[\text{Var}_F]/(1-q)^2 \\
\sigma_{12} &= A_0 q/(1-q)^2 = \sigma_{12}^*(\mu) \\
\sigma_{22} &= q/(1-q)^2 = \sigma_{22}^*(\mu)
\end{align*}
\]

Thus the effect of this type of departure from the null hypothesis, model a, is seen only in the variance of \( K_1 \), and the effect is to increase the variance from our expectations under the null hypothesis by an amount depending on both the value of \( q \) and the variance of the distribution \( F \).

Is this effect maintained by other sorts of alternatives? Consider model A-d, where the probability of success varies from trial to trial. Then
\[ \mu_1 = \sum_{m=0}^{\infty} q^m p_{m+1} \]
\[ \mu_2 = \frac{1}{1-q} \]
\[ \sigma_{11} = \mu_1 - \frac{2}{1-q} + 2 \sum_{m=1}^{\infty} \sum_{i=1}^{m} p_i p_{m+1} q^m \]
\[ = \sigma_{11}^*(\mu) + 2 \sum_{m=1}^{\infty} \sum_{i=1}^{m} q^m p_i [p_{m+1} - p_{m+1-i}] \]
\[ \sigma_{12} = \sum_{k=1}^{\infty} k q^k p_{k+1} \]
\[ = \sigma_{12}^*(\mu) + \sum_{m=1}^{\infty} \sum_{k=m}^{\infty} q^k (p_{k+1} - p_m) \]
\[ \sigma_{22} = \frac{q}{(1-q)^2} = \sigma_{22}^*(\mu). \]

If the \( p_i \)'s form an increasing sequence, then the two additional terms in \( \sigma_{11} \) and \( \sigma_{12} \) are positive, and so the variance of \( K \) and the covariance are increased, as above. However, if the \( p_i \)'s form a decreasing sequence, the opposite occurs, and these quantities are decreased.

Thus, even when we restrict ourselves to a given stopping rule, the effects of departures from the hypothesis on the moments of the simple summary statistics \( (K, N) \) are inconsistent.
V. CHI-SQUARED TESTS

Another possibility for testing models is to use the standard chi-squared test by proceeding in the following manner. We classify all possible sequences of zeroes and ones into a small number of cells. We have, a priori, some model we wish to test; we choose, arbitrarily, some stopping rule. Using the specified model and stopping rule, the total probability of each cell can be ascertained as a function of the various parameters. We now use the data to estimate the parameters and thus obtain estimated cell probabilities. Finally, we compare observed to estimated expected occurrences of sequences from the various cells in the standard way, using the chi-squared goodness-of-fit statistic.

What properties will such a test have? If the hypothesized model and the assumed stopping rule both obtain, then standard chi-squared theory tells us that the power will be equal to the level of significance. If an alternate model holds, but the assumed stopping rule still obtains, then, as we would desire, the power will approach one as the sample size goes to infinity. Unfortunately, if the hypothesized model holds but a different stopping rule obtains, the power will still approach one, and that is undesirable. However, by designing the test around a judicious choice of a stopping rule
and a grouping of the possible data sequences, it may be feasible to construct a test for which the power will increase rapidly in the former situation and slowly in the latter. Such a test would have reasonable power against disturbances of the model, but would be robust against disturbances of the stopping rule. The remainder of this section is an exploration of such possibilities.

A. General chi-squared theory

Let us review the general theory of chi-squared tests. Suppose that we have $k$ cells, and that we observe $x_1, \ldots, x_k$ in cells $1, \ldots, k$ respectively, where $\sum_1^k x_i = m$. Let $x$ represent the vector $(x_1, \ldots, x_k)$. Let $y_i = x_i/m$ and let $y$ represent the vector $(y_1, \ldots, y_k)$.

Suppose first that we wish to test the simple hypothesis

$$H_0 : P(\text{cell } i) = f_i, \quad i = 1, \ldots, k.$$ 

The standard chi-squared statistic is $T = \text{mD}(y, f)$ where

$$D(y, f) = \sum_{i=1}^k (y_i - f_i)^2/f_i.$$ 

It is well known (see e.g. [1]) that, under $H_0$, $T$ converges in distribution to a chi-squared distribution with $k-1$ degrees of freedom. Furthermore, for any fixed alternative

$$H_1 : P(\text{cell } i) = g_i, \quad i = 1, \ldots, k$$

where $f_i \neq g_i$ for some $i$, and for any fixed non-trivial level of significance, the test is consistent. Now consider the sequence of local alternatives
\[ H_m : P(cell\ i) = g_{mi} = f_i + \frac{c_i}{\sqrt{m}}, \quad i = 1, \ldots, k \]

where \( \sum_1^k c_i = 0 \). Under this sequence of distributions, \( T \) converges to a non-central chi-squared distribution with \( k-1 \) degrees of freedom and non-centrality parameter \( \sum_1^k c_i^2 / f_i = mD(g_m, f) \). Thus under an alternative \( g \), the distribution of \( T \) may be approximated by a chi-squared distribution with \( k-1 \) degrees of freedom and non-centrality parameter \( mD(g, f) \).

Suppose the hypothesis to be tested is a composite one:

\[ H : P(cell\ i) = f_i(\Theta) \text{ for some } \Theta \in \Theta \]

where \( \Theta \) is an \( r \)-dimensional space. Suppose we use the statistic \( T = mD(y, f(\hat{\Theta})) \), where \( \hat{\Theta} \) is an asymptotically efficient estimate of \( \Theta \), such as the maximum likelihood estimate or the minimum chi-squared estimate. Then results similar to those above can be obtained [6]. Under a local alternative \( g \), the distribution of \( T \) may be approximated by a non-central chi-squared distribution with \( k-r-1 \) degrees of freedom and non-centrality parameter \( \inf_\Theta mD(g, f(\Theta)) \).

What is the effect of changing the grouping of possible outcomes from one set of cells to another? Suppose we got from one grouping to a coarser one by a process of combining several cells into one. Such a coarsening will result in a reduction in both the degrees of freedom and the non-centrality parameter. (The former is due to the reduction in the number of cells, while the latter is due to the algebraic identity...
\[
\frac{(a-b)^2}{a} + \frac{(c-d)^2}{c} \geq \frac{(a+c-b-d)^2}{a+c}
\]

It is well-known (see e.g. [2]) that, for fixed non-centrality parameter, reducing the degrees of freedom will increase the power, while for fixed degrees of freedom, reducing the non-centrality parameter will decrease the power. Thus these two changes will tend to have opposite effects, and the final effect on asymptotic power can be either an increase or a decrease.

B. Application to zero-one sequences

1. General considerations. Returning to our particular situation, let us consider testing model a, that all trials from all individuals are independent with a common probability \( p \) of success. Choose, arbitrarily, the stopping rule \( B \), that the probability of continuation is either \( q_0 \) or \( q_1 \) depending on the result of the immediately preceding trial. We will investigate whether judicious choice of grouping will enable a chi-squared test to be sensitive to departures from the model while being robust under departures from the stopping rule.

In order to judge a grouping's performance against a given alternative \( g \), the following criterion will be used. Suppose that the grouping has \( k \) cells, and that the probabilities of the \( i \)th cell under the null hypothesis and under the particular alternative hypothesis are, respectively, \( f_i(p, q_0, q_1) \) and \( g_i \). Let

\[
\lambda = \inf_{p, q_0, q_1} D(g, f(p, q_0, q_1)).
\]

Then a chi-squared statistic based on \( m \) observations will be distributed, under \( g \), approximately as
\[ \chi^2_{k-4}(m\lambda) \]. Rather than considering arbitrary sample sizes, we will characterize the power against each alternative by the sample size required to obtain power .50 when the level of the test is .05. Based on the asymptotic approximation above, this sample size will be inversely proportional to \( \lambda \), with the constant of proportionality depending on the degrees of freedom. Tables of the non-central chi-squared distribution \([4]\) give the following:

<table>
<thead>
<tr>
<th>d.f.</th>
<th>( m\lambda )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.84</td>
</tr>
<tr>
<td>2</td>
<td>4.96</td>
</tr>
<tr>
<td>3</td>
<td>5.76</td>
</tr>
<tr>
<td>4</td>
<td>6.42</td>
</tr>
<tr>
<td>5</td>
<td>7.00</td>
</tr>
<tr>
<td>6</td>
<td>7.50</td>
</tr>
<tr>
<td>7</td>
<td>7.97</td>
</tr>
<tr>
<td>8</td>
<td>8.40</td>
</tr>
<tr>
<td>9</td>
<td>8.81</td>
</tr>
<tr>
<td>10</td>
<td>9.19</td>
</tr>
<tr>
<td>11</td>
<td>9.56</td>
</tr>
</tbody>
</table>

For an alternative which represents a disturbance of the hypothesized model, but not the stopping rule, we would like these sample sizes \( m \) to be small (i.e. \( \lambda \) large), whereas if the stopping rule but not the model has been disturbed, we would like these sample sizes \( m \) to be large (\( \lambda \) small).
2. Grouping of sequences according to the total number of successes.

One grouping which seems natural is based on the number of successes; i.e., the first cell consists of all sequences with no successes, the second of all with one success, and so on, with the last consisting of all with some specified number \( J+1 \) or more. Under model B-a, we have:

\[
P(0 \text{ successes}) = 1 - R
\]

\[
P(k \text{ successes}) = R^k w^{k-1} (1-w), \quad k = 1, \ldots, J
\]

\[
P(J+1 \text{ or more}) = R^J w^J
\]

where

\[
w = \frac{p q_1}{(1-q_0)(1-p)}
\]

\[
R = \frac{w}{q_1}
\]

Consider an alternative with probabilities \( A_0, A_1, \ldots, A_J, A_{J+1} \).

Then the non-centrality parameter \( \lambda \) can be obtained as follows:

\[
\lambda = \min_{w,R} \left[ \frac{A_0^2}{1-R} + \sum_{k=1}^{J} \frac{A_k^2}{R^k w^{k-1} (1-w)} + \frac{A_{J+1}^2}{R^J w^J} - 1 \right]
\]

\[
= \min_{w,R} \left[ \frac{A_0^2}{1-R} + \frac{Q(w)}{R} - 1 \right]
\]

where

\[
Q(w) = \sum_{k=1}^{J} \frac{A_k^2}{R^k w^{k-1} (1-w)} + \frac{A_{J+1}^2}{R^J w^J}
\]

The minimizing value of \( R \) can easily be seen to be
\[ R = \frac{\sqrt{Q(w)}}{\sqrt{Q(w)} + A_0} \]

and thus

\[ \lambda = \min \left( (A_0 + \sqrt{Q(w)})^2 - 1 \right) \]

\[ = (A_0 + \min w \{Q(w)\}^{1/2})^2 - 1 \]

Now, on the range of \( w \), which is the unit interval, \( Q \) is a sum of
convex functions and therefore convex, so a unique minimum exists and
can easily be found numerically. The results given below, for example,
were found by a binary search of the unit interval for the point at
which the derivative crosses zero. Notice that the number of degrees
of freedom for the chi-squared distribution is \( J-1 \).

This grouping is discussed further below, where it can be
seen to be fairly unsuccessful. One particular case which should be
noted here is the Markov case: the distribution of the number of
successes is the same in the Markov case, model B-e, as in the hypothesized
model B-a. (See Section II.C.4.) Thus the test based on this grouping
has no power against the Markov family of alternatives.

3. Other groupings. All other groupings discussed in this paper
are obtained in the following manner: a base grouping of 15 cells is
obtained by putting the 14 sequences of length three or less into
individual cells and grouping all sequences of length four or more into
one cell; other groupings are coarsening of this partition (i.e. they
are obtained from the base by combining several cells into one). Since there are three parameters in the hypothesized model B-a, the minimum number of cells is five.

Hopefully, by looking at groupings of the above form, the importance of and relationships between individual sequences can be discovered. By combining attention to the first \( l^4 \) sequences, this task is cut down to manageable proportions. Furthermore, in the original application to spontaneous abortions, most of the observed sequences had lengths less than four, so that these \( l^4 \) were the ones of primary interest.

The non-centrality parameters of the chi-squared approximations to the distributions of these test statistics under alternative distributions can only be found numerically; for a particular alternative \( g \) and a particular grouping (reflected in \( f \)) it is necessary, as in Section V.B.1, to find the values of \( p, q_0 \) and \( q_1 \) which minimize \( D(g,f(p,q_0,q_1)) \). For the base grouping, the following procedure was used.

As a first approximation, the probability limits of the maximum likelihood estimates of \( p, q_0 \) and \( q_1 \) were obtained. Letting \( y_1, \ldots, y_{15} \) represent the observed proportions in the 15 cells of the base grouping (ordered as in the appendix), the maximum likelihood estimates of \( p, q_0 \) and \( q_1 \) are calculated as follows:
\[ x_1 = y_1 + y_3 + y_5 + y_7 + y_9 + y_{10} + y_{13} \]
\[ x_2 = y_5 + y_6 + 2y_7 + 2y_8 + y_9 + y_{10} + y_{11} + y_{12} \]
\[ x_3 = y_2 + y_4 + y_6 + y_8 + y_{11} + y_{12} + y_{14} \]
\[ = 1 - x_1 - y_{15} \]
\[ x_4 = y_5 + y_6 + y_9 + y_{10} + y_{11} + y_{12} + 2y_{13} + 2y_{14} \]
\[ x_5 = 3y_{15} \]
\[ A = x_2(x_2 + x_4 + x_5) \]
\[ B = x_1(x_2 + x_4) \]
\[ C = x_4(x_2 + x_4 + x_5) \]
\[ D = x_5(x_2 + x_4) \]
\[ \hat{p} = (C + D)/(A + B + C + D) \]
\[ \hat{q}_0 = A/(A + B) \]
\[ \hat{q}_1 = C/(C + D) \]

Replacing \( y \) by \( g \) throughout gives the probability limits of these estimates. Using these values as a starting point, an approximate minimum to the function \( D(g, f(p, q_0, q_1)) \) was found by two iterations of a Newton-Rephson-like technique. The value of the minimum gives the non-centrality parameter for the base grouping. The point \((p^*, q_0^*, q_1^*)\) at which it was obtained was used in turn as a starting point for finding non-centrality parameters corresponding to other groupings.

These other groupings were handled by the following search procedure. Suppose we start at some point \((p^*, q_0^*, q_1^*)\); we examine
it and the 26 points on the cube around it of the form
\[(p^* + \delta_1 d, q_0^* + \delta_2 d, q_1^* + \delta_3 d)\] where \(\delta_i \in \{-1,0,1\}\) for \(i = 1, 2, 3\).
The function \(D(g, f(p, q_0, q_1))\) is evaluated at all 27 points, and
we move to the point giving the minimum value. This is repeated until
the minimum is obtained at the center point. We then shrink the value
of \(d\) and repeat. This was done for \(d = .02, .005\) and \(.001\). If
the search ever led out of the unit cube, it was terminated and
the value on the boundary was accepted as the minimum. (This only
occurred for groupings which performed badly.)

C. Numerical results

The numerical results for various selected groupings are given
in the tables in the appendix. All results are given relative to the
base grouping of 15 cells discussed in Section V.B.3. Thus the column
labeled "base sample size" gives, for the base grouping and the particular
alternative hypothesis, the asymptotically approximated sample sizes
required to achieve power .50. For the other groupings, the entry is
not the sample size itself, but the ratio of the sample size required
by that particular grouping to the size required by the base grouping.
As will be seen, the base grouping itself is not particularly good,
so these relative numbers must be interpreted with some caution; a
grouping that is better than the base grouping is not necessarily
good.

The base grouping is most sensitive to the increasing/decreasing
probabilities cases (model B-d), where the required sample sizes
(i.e. number of sequences) are generally 300 or less, and to the Markov case (model B-e), where they range up to 500. However, it is highly insensitive to the two-point model (B-b), where 20,000 sequences are sometimes needed. The two cases which represent disturbances of the stopping rule (models C-a and D-a) give intermediate results. Thus the base grouping fails to separate departures from the model from departures from the stopping rule, and also does very poorly in terms of power against the two-point model.

Grouping 1 is a particular case of grouping by the number of successes, as discussed in Section V.B.2. There are seven cells, consisting of all sequences with 0, 1, 2, 3, 4, 5, and more than 5 successes, respectively. This is somewhat worse than the base grouping. Sensitivity is decreased everywhere; sample sizes are often increased by an order of magnitude, and sometimes by several. Thus power is definitely decreased. The alternatives representing disturbances of the stopping rule roughly maintain their position vis-a-vis the other alternatives, so no clearer separation is seen. The one case where a drastic effect is seen is the Markov case: as shown above (see Section V.B.2) this grouping has no power at all against Markov alternatives.

A special note on the increasing/decreasing probabilities model is needed. The parameters given in the tables specify only the first three elements of the sequence of probabilities. For all other groupings considered, these suffice, but for this one, it is necessary to know the entire sequence to determine the distributions.
Two possible completions are used in the tables: A) all unspecified elements equal to $p_3$ and B) all unspecified elements equal to 1 in the increasing case and to 0 in the decreasing case. It is interesting to notice the lack of symmetry. The increasing case, where ratios of several hundred are seen, seems to behave much more badly for this grouping than the decreasing one, where the ratios are on the order of 50 to 100.

Grouping 2 is another natural one to consider. Here sequences with both the same number of successes and the same number of failures are combined. Thus the following 10 cells are created: [0], [1], [00], [01, 10], [11], [000], [001, 010, 100], [011, 101, 110], [111], and (all other sequences). This performs slightly better than the base groupings; the enormous sample sizes of the two point model are brought down by factors of roughly .8, which is an improvement, though still not good. Those for the increasing/decreasing probabilities cases roughly triple, but they were low to begin with, so this is not such an exorbitant price to pay. The Markov case figures are essentially unchanged. The figures for the final two cases, where only the stopping rule has been disturbed, improve somewhat (the sample sizes generally double) but this improvement is not nearly sufficient to give this grouping any usefulness.

Grouping 3 represents a guess based on preliminary exploration of individual sequences. For a few representative alternatives, the 15 cells of the base grouping were reduced to 14 cells (by combining two into one) in all possible ways. On the basis of the patterns that emerged, the sequences were identified as sensitive or non-sensitive
to the different types of alternatives. Thus, for example, the sequences 000 and 111 are very sensitive to the two-point alternative, while 001 and 011 seem to be sensitive to the increasing probabilities alternative. The suggested grouping was \{0, 001, 011\}, \{00, 11, 000, 111\}, 
\{01, 10\}, \{1\}, and \{all other sequences\}. The results, while incomplete, show clearly that this grouping is a failure. The patterns are somewhat similar to grouping 1. This one is included only to show the possible dangers of devising groupings from such principles.

Grouping 4 is reasonably successful. It is again a 5-cell grouping, as follows: \{1\}, \{000\}, \{001, 011\}, \{111\}, and \{all other sequences\}. The sample sizes for the two-point model are again brought down by roughly .8. Those for the increasing/decreasing probabilities cases, expect for a few wild cases, go up by factors of 5 to 10, while the Markov case ones go up by 1 to 5. So from the point of view of power, the grouping is not outstanding. But sample sizes for the alternatives where the stopping rule has been disturbed are increased by factors like 50; thus, in contrast to the other groupings we have seen, the sample sizes for these two cases tend to be larger than for the two-point model. Sensitivity has finally been ordered correctly; the test is least sensitive to alternatives where only the stopping rule has been changed.

Grouping 5, which consists of these 5 cells: \{000,111\}, 
\{001, 011\}, \{1, 10\}, \{0, 01\} and \{all other sequences\}, behaves rather like grouping 4. It does better against the increasing probabilities alternatives, but less well against the two-point model.
However, the performance against stopping rule D is not good, as some of the sample sizes are back down in the 200 to 300 range.

Groupings 6 through 10 are related. They consist of the following cells:

6) \{000\}, \{111\}, \{001\}, \{100\} and \{all other sequences\}
7) \{000,00\}, \{111,11\}, \{001\}, \{100\}, and \{all other sequences\}
8) \{000,00\}, \{111,11\}, \{001,011\}, \{100,110\} and \{all other sequences\}
9) \{000,00\}, \{111,11\}, \{001,011,01\}, \{100,110,10\} and \{all other sequences\}
10) \{000,00,0\}, \{111,11,1\}, \{001,011,01\}, \{100,110,10\} and \{all other sequences\}

Generally, as more sequences are shifted to the first four cells, there is an increase in sensitivity. For the two-point and Markov cases, the trend is striking until grouping 10 is reached; at this point, when the sequences 0 and 1 are shifted, a decrease in sensitivity occurs. Against the alternatives with stopping rules C and D, the trend is not as striking, but discernible. Typically one of the five groupings is "out of order," but which one it is varies. The fact that sensitivity decreases for both kinds of alternatives presents a problem when choosing among these five; for power against disturbances of the model, we would like to use grouping 10, while for robustness against disturbances of the stopping rule, we would like to use grouping 6, where sensitivity is low. In fact, examination of the results in the appendix shows that grouping 6 is the best choice overall;
the gains in power which are possible by using others are small compared to the decrease in robustness. Notice that grouping 6 distinguishes only \( \frac{1}{4} \) sequences, and groups all others together; as a general rule, the more successful groupings of those that were tried similarly tended to distinguish only a small number of sequences.

D. Simulations

Throughout this paper, the non-central chi-squared distribution has been used to approximate the power of the chi-squared tests. This approximation is an asymptotic local one, so one needs to address the question of whether or not the approximation is generally useful. To check this, some simulations were run. The results are displayed in tables at the end of this section. Using various representative alternative hypotheses, 300 or 500 sequences were generated; and the chi-squared statistic was computed, both for the base grouping and for grouping \( \frac{1}{4} \). This process was replicated, usually 50 times. For each grouping, the proportion of these statistics above various cutoff points (corresponding to various sizes of tests) are tabulated and compared to the expected proportions. The results are generally quite good; the expected and observed distributions are often strikingly close. In particular, this is true when the power is near \( .5 \), which is the place where efficiency comparisons were made. There are a few cases where the discrepancies are large (see especially case \( 4 b \)), but the proportion of such cases is so low that it seems safe to generally rely on these approximations as good guides to reality.
Since the sample sizes involved are quite large by ordinary standards, the usefulness of asymptotic approximations is perhaps not surprising.

SIMULATION RESULTS

1) two-point model (B-b)

a) \((p_1, p_2, \gamma, q_0, q_1) = (.8, .4, .5, .7, .5)\)

sample size = 500, replications = 50

<table>
<thead>
<tr>
<th>size</th>
<th>power: base grouping</th>
<th>power: No. 4 grouping</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>expected</td>
<td>observed</td>
</tr>
<tr>
<td>.10</td>
<td>.58</td>
<td>.50</td>
</tr>
<tr>
<td>.05</td>
<td>.45</td>
<td>.40</td>
</tr>
<tr>
<td>.025</td>
<td>.34</td>
<td>.36</td>
</tr>
<tr>
<td>.01</td>
<td>.22</td>
<td>.26</td>
</tr>
<tr>
<td>.005</td>
<td>.16</td>
<td>.24</td>
</tr>
<tr>
<td>.001</td>
<td>.07</td>
<td>.14</td>
</tr>
</tbody>
</table>

b) \((p_1, p_2, \gamma, q_0, q_1) = (.8, .2, .2, .5, .3)\)

sample size = 300, replications = 50

<table>
<thead>
<tr>
<th>size</th>
<th>power: base grouping</th>
<th>power: No. 4 grouping</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>expected</td>
<td>observed</td>
</tr>
<tr>
<td>.10</td>
<td>.65</td>
<td>.56</td>
</tr>
<tr>
<td>.05</td>
<td>.52</td>
<td>.46</td>
</tr>
<tr>
<td>.025</td>
<td>.41</td>
<td>.36</td>
</tr>
<tr>
<td>.01</td>
<td>.29</td>
<td>.22</td>
</tr>
<tr>
<td>.005</td>
<td>.22</td>
<td>.16</td>
</tr>
<tr>
<td>.001</td>
<td>.10</td>
<td>.10</td>
</tr>
</tbody>
</table>
c) \((p_1, p_2, r, q_0, q_1) = (.6, .2, .5, .7, .5)\)

sample size = 300, replications = 50

<table>
<thead>
<tr>
<th>size</th>
<th>expected</th>
<th>observed</th>
<th>expected</th>
<th>observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>.10</td>
<td>.38</td>
<td>.40</td>
<td>.46</td>
<td>.44</td>
</tr>
<tr>
<td>.05</td>
<td>.24</td>
<td>.36</td>
<td>.34</td>
<td>.34</td>
</tr>
<tr>
<td>.025</td>
<td>.18</td>
<td>.30</td>
<td>.25</td>
<td>.22</td>
</tr>
<tr>
<td>.01</td>
<td>.11</td>
<td>.22</td>
<td>.15</td>
<td>.20</td>
</tr>
<tr>
<td>.005</td>
<td>.07</td>
<td>.14</td>
<td>.10</td>
<td>.16</td>
</tr>
<tr>
<td>.001</td>
<td>.02</td>
<td>.02</td>
<td>.04</td>
<td>.08</td>
</tr>
</tbody>
</table>

2) increasing/decreasing probabilities case (B-d)

a) \((p_1, p_2, p_3, q_0, q_1) = (.2, .5, .6, .7, .3)\)

sample size = 500, replications = 50

<table>
<thead>
<tr>
<th>size</th>
<th>expected</th>
<th>observed</th>
<th>expected</th>
<th>observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>.10</td>
<td>1</td>
<td>1</td>
<td>.86</td>
<td>.80</td>
</tr>
<tr>
<td>.05</td>
<td>1</td>
<td>1</td>
<td>.77</td>
<td>.74</td>
</tr>
<tr>
<td>.025</td>
<td>1</td>
<td>1</td>
<td>.69</td>
<td>.66</td>
</tr>
<tr>
<td>.01</td>
<td>1</td>
<td>1</td>
<td>.57</td>
<td>.54</td>
</tr>
<tr>
<td>.005</td>
<td>1</td>
<td>1</td>
<td>.47</td>
<td>.44</td>
</tr>
<tr>
<td>.001</td>
<td>1</td>
<td>1</td>
<td>.29</td>
<td>.22</td>
</tr>
</tbody>
</table>
b) $$(p_1, p_2, p_3, q_0, q_1) = (.2, .3, .4, .5, .3)$$

sample size = 300, replications = 50

<table>
<thead>
<tr>
<th>size</th>
<th>power: base grouping</th>
<th>power: No. 4 grouping</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>expected</td>
<td>observed</td>
</tr>
<tr>
<td>.10</td>
<td>.65</td>
<td>.64</td>
</tr>
<tr>
<td>.05</td>
<td>.52</td>
<td>.48</td>
</tr>
<tr>
<td>.025</td>
<td>.41</td>
<td>.44</td>
</tr>
<tr>
<td>.01</td>
<td>.29</td>
<td>.38</td>
</tr>
<tr>
<td>.005</td>
<td>.22</td>
<td>.30</td>
</tr>
<tr>
<td>.001</td>
<td>.10</td>
<td>.10</td>
</tr>
</tbody>
</table>

c) $$(p_1, p_2, p_3, q_0, q_1) = (.6, .4, .2, .7, .5)$$

sample size = 300, replications = 50

<table>
<thead>
<tr>
<th>size</th>
<th>power: base grouping</th>
<th>power: No. 4 grouping</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>expected</td>
<td>observed</td>
</tr>
<tr>
<td>.10</td>
<td>.99</td>
<td>1</td>
</tr>
<tr>
<td>.05</td>
<td>.97</td>
<td>1</td>
</tr>
<tr>
<td>.025</td>
<td>.95</td>
<td>1</td>
</tr>
<tr>
<td>.01</td>
<td>.89</td>
<td>.92</td>
</tr>
<tr>
<td>.005</td>
<td>.85</td>
<td>.88</td>
</tr>
<tr>
<td>.001</td>
<td>.72</td>
<td>.64</td>
</tr>
</tbody>
</table>
3) Markov case (model B-e)
   a) \((a, b, q_0, q_1) = (.4, .6, .5, .3)\)

   \[
   \text{sample size} = 500, \text{replications} = 50
   \]

   \[
   \begin{array}{ccc|ccc}
   \text{size} & \text{expected} & \text{observed} & \text{power: base grouping} & \text{expected} & \text{observed} \\
   .10 & .57 & .74 & .30 & .34 \\
   .05 & .44 & .52 & .20 & .16 \\
   .025 & .33 & .38 & .13 & .16 \\
   .01 & .22 & .26 & .07 & .10 \\
   .005 & .16 & .18 & .04 & .02 \\
   .001 & .07 & .14 & .01 & 0 \\
   \end{array}
   \]

   b) \((a, b, q_0, q_1) = (.6, .2, .7, .5)\)

   \[
   \text{sample size} = 300, \text{replications} = 50
   \]

   \[
   \begin{array}{ccc|ccc}
   \text{size} & \text{expected} & \text{observed} & \text{power: base grouping} & \text{expected} & \text{observed} \\
   .10 & .96 & .98 & .89 & .88 \\
   .05 & .93 & .92 & .83 & .78 \\
   .025 & .89 & .84 & .75 & .74 \\
   .01 & .81 & .84 & .65 & .70 \\
   .005 & .75 & .80 & .54 & .68 \\
   .001 & .58 & .70 & .35 & .54 \\
   \end{array}
   \]

   c) \((a, b, q_0, q_1) = (.2, .6, .7, .3)\)

   \[
   \text{sample size} = 300, \text{replications} = 50
   \]

   \[
   \begin{array}{ccc|ccc}
   \text{size} & \text{expected} & \text{observed} & \text{power: No. 4 grouping} & \text{expected} & \text{observed} \\
   .10 & .95 & .94 & .54 & .60 \\
   .05 & .90 & .90 & .41 & .34 \\
   .025 & .85 & .84 & .31 & .22 \\
   .01 & .76 & .80 & .20 & .04 \\
   .005 & .69 & .70 & .14 & 0 \\
   .001 & .51 & .48 & .06 & 0 \\
   \end{array}
   \]
4) stopping rule C (model C-a)

   a) \((p, \alpha, q) = (0.2, 0.8, 0.4)\)

   sample size = 500, replications = 98

<table>
<thead>
<tr>
<th>size</th>
<th>expected</th>
<th>observed</th>
<th>expected</th>
<th>observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.89</td>
<td>0.73</td>
<td>0.10</td>
<td>0.05</td>
</tr>
<tr>
<td>0.05</td>
<td>0.81</td>
<td>0.69</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>0.025</td>
<td>0.73</td>
<td>0.60</td>
<td>0.025</td>
<td>0.02</td>
</tr>
<tr>
<td>0.01</td>
<td>0.62</td>
<td>0.53</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>0.005</td>
<td>0.53</td>
<td>0.50</td>
<td>0.005</td>
<td>0</td>
</tr>
<tr>
<td>0.001</td>
<td>0.35</td>
<td>0.37</td>
<td>0.001</td>
<td>0</td>
</tr>
</tbody>
</table>

b) \((p, \alpha, q) = (0.6, 0.8, 0.6)\)

   sample size = 300, replications = 50

<table>
<thead>
<tr>
<th>size</th>
<th>expected</th>
<th>observed</th>
<th>expected</th>
<th>observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.96</td>
<td>0.40</td>
<td>0.10</td>
<td>0.02</td>
</tr>
<tr>
<td>0.05</td>
<td>0.93</td>
<td>0.34</td>
<td>0.05</td>
<td>0</td>
</tr>
<tr>
<td>0.025</td>
<td>0.89</td>
<td>0.32</td>
<td>0.03</td>
<td>0</td>
</tr>
<tr>
<td>0.01</td>
<td>0.81</td>
<td>0.22</td>
<td>0.01</td>
<td>0</td>
</tr>
<tr>
<td>0.005</td>
<td>0.75</td>
<td>0.22</td>
<td>0.01</td>
<td>0</td>
</tr>
<tr>
<td>0.001</td>
<td>0.58</td>
<td>0.16</td>
<td>0.00</td>
<td>0</td>
</tr>
</tbody>
</table>
c) \((p, \alpha, q) = (.6, .6, .2)\)

sample size = 300, replications = 95

<table>
<thead>
<tr>
<th>size</th>
<th>power: base grouping</th>
<th>power: No. 4 grouping</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>expected</td>
<td>observed</td>
</tr>
<tr>
<td>.10</td>
<td>.14</td>
<td>.26</td>
</tr>
<tr>
<td>.05</td>
<td>.07</td>
<td>.19</td>
</tr>
<tr>
<td>.025</td>
<td>.04</td>
<td>.15</td>
</tr>
<tr>
<td>.01</td>
<td>.02</td>
<td>.08</td>
</tr>
<tr>
<td>.005</td>
<td>.01</td>
<td>.06</td>
</tr>
<tr>
<td>.001</td>
<td>.00</td>
<td>.04</td>
</tr>
</tbody>
</table>

5) Stopping rule D (model D-a)

a) \((p, \pi_0, \pi_1) = (.6, .8, .4)\)

sample size = 500, replications = 67

<table>
<thead>
<tr>
<th>size</th>
<th>power: base grouping</th>
<th>power: No. 4 grouping</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>expected</td>
<td>observed</td>
</tr>
<tr>
<td>.10</td>
<td>.99</td>
<td>.82</td>
</tr>
<tr>
<td>.05</td>
<td>.99</td>
<td>.76</td>
</tr>
<tr>
<td>.025</td>
<td>.98</td>
<td>.73</td>
</tr>
<tr>
<td>.01</td>
<td>.97</td>
<td>.67</td>
</tr>
<tr>
<td>.005</td>
<td>.95</td>
<td>.64</td>
</tr>
<tr>
<td>.001</td>
<td>.88</td>
<td>.54</td>
</tr>
</tbody>
</table>
b) \((p, \pi_0, \pi_1) = (.6, .6, .2)\)

sample size = 300, replications = 50

<table>
<thead>
<tr>
<th>size</th>
<th>power: base grouping</th>
<th>power: No. 4 grouping</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>expected</td>
<td>observed</td>
</tr>
<tr>
<td>.10</td>
<td>.53</td>
<td>.40</td>
</tr>
<tr>
<td>.05</td>
<td>.39</td>
<td>.30</td>
</tr>
<tr>
<td>.025</td>
<td>.29</td>
<td>.22</td>
</tr>
<tr>
<td>.01</td>
<td>.18</td>
<td>.14</td>
</tr>
<tr>
<td>.005</td>
<td>.13</td>
<td>.10</td>
</tr>
<tr>
<td>.001</td>
<td>.05</td>
<td>.08</td>
</tr>
</tbody>
</table>

c) \((p, \pi_0, \pi_1) = (.2, .2, .6)\)

sample size = 300, replications = 36

<table>
<thead>
<tr>
<th>size</th>
<th>power: base grouping</th>
<th>power: No. 4 grouping</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>expected</td>
<td>observed</td>
</tr>
<tr>
<td>.10</td>
<td>.70</td>
<td>.33</td>
</tr>
<tr>
<td>.05</td>
<td>.59</td>
<td>.14</td>
</tr>
<tr>
<td>.025</td>
<td>.47</td>
<td>.11</td>
</tr>
<tr>
<td>.01</td>
<td>.35</td>
<td>.06</td>
</tr>
<tr>
<td>.005</td>
<td>.27</td>
<td>.03</td>
</tr>
<tr>
<td>.001</td>
<td>.14</td>
<td>0</td>
</tr>
</tbody>
</table>
E. Other considerations

There is a tendency to think that longer sequences provide more information, and thus that alternatives which produce longer sequences can be more easily distinguished from the null hypothesis than others producing shorter sequences. A glance at the tables in the appendix shows that this is not generally so. The expected sequence length for each alternative hypothesis is tabulated in the appendix, immediately after the specification of the parameters. (In the increasing/decreasing probabilities cases, the incomplete specification of the parameters provides bounds instead of an actual value. Expected lengths were not computed for the Markov case.) There is no clear-cut relationship at all between the expected sequence length under a given alternative and the power against that alternative. For example, in Table 1, dealing with the two-point model, B-b, the rank correlation between the expected sequence length and the non-centrality parameter is .098. Most of the other tables show similar patterns.

Attention in this paper has been focused on groupings with five cells. This is primarily because of power considerations. Coarsening a partition produces opposite effects (see Section V.A) which may, in conjunction, increase power or decrease power. However, the most dramatic single effect is achieved when degrees of freedom are reduced to one or two. The non-centrality parameter needs to be reduced very far to overcome this effect. Thus five cells (i.e. one degree of freedom) seems the most promising arrangement to produce good power properties of these tests against disturbances of the model. Clearly, robustness against stopping rule misspecifications will be adversely affected, but we have seen that this problem can be mitigated.
The entries in the tables in the appendix have interpretations as standard efficiencies in the following sense: for two groupings with the same number of cells, the ratio of the tabulated sample sizes is essentially the Hodges-Lehmann efficiency at that alternative [5]. To see this, define $c$ and $w$ by

$$P(\chi^2_k > c(k, \alpha)) = \alpha$$

$$P(\chi^2_k(w(k, \alpha, \gamma)) > c(k, \alpha)) = \gamma$$

i.e., $c$ is the $\alpha$-level rejection point, and $w$ is the non-centrality parameter necessary to achieve power $\gamma$. (Section V.B.1 gives a table of $w$ for $\alpha = .05, \gamma = .50$.) Consider two groupings with $k$ and $k^*$ cells respectively, and fix some alternative. There will be two non-centrality parameters $\lambda$ and $\lambda^*$ corresponding to this alternative for the two groupings. Based on the chi-squared approximation, the numbers of observations needed to achieve power $\gamma$ for level $\alpha$ are $N = w(k, \alpha, \gamma)/\lambda$ and $N^* = w(k^*, \alpha, \gamma)/\lambda^*$, respectively. The ratio

$$\frac{N}{N^*} = \frac{w(k, \alpha, \gamma)}{w(k^*, \alpha, \gamma)} \frac{\lambda^*}{\lambda}$$

is clearly independent of both $\alpha$ and $\gamma$ if $k = k^*$. (Hodges-Lehmann asymptotic efficiency is defined to be $\lim_{\gamma \to 1} N/N^*$. ) Of course, measures such as this are unsatisfactory in that they do depend on the particular alternative.
APPENDIX

The following tables give, for various groupings and various families of alternative hypotheses, the sample sizes required for a chi-squared test of model B-a at significance level .05 to reject with probability .50. The procedures for obtaining these sample sizes are described in Section V-B. The tables refer to the following alternative models:

<table>
<thead>
<tr>
<th>table number</th>
<th>model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>B-b</td>
</tr>
<tr>
<td>2</td>
<td>B-d</td>
</tr>
<tr>
<td>3</td>
<td>B-d</td>
</tr>
<tr>
<td>4</td>
<td>B-e</td>
</tr>
<tr>
<td>5</td>
<td>C-a</td>
</tr>
<tr>
<td>6</td>
<td>D-a</td>
</tr>
</tbody>
</table>

These models are described in Chapter II, and the meanings of the parameters are given there. The entries are first, the parameters of the particular alternative; next, the expected length of an individual sequence under that alternative; and finally, the sample sizes required by tests based on various groupings. The "base" sample size is that required for achieving the power criterion for the grouping of possible zero-one sequences into cells as follows:
0
1
00
01
10
11
000
001
010
100
011
101
110
111
all others

For all other groupings, the entry is not the actual sample size, but the ratio of the actual sample size to the base sample size. The groupings are referred to by number, and the following list gives the full descriptions:

(1) 7 cells consisting of all sequences with 0, 1, 2, 3, 4, 5 and more than 5 successes, respectively.

(2) 10 cells:

0
1
00
01-10
11
000
001-010-100
011-101-110
111
all others

69
(3) 5 cells:
0-001-011
00-11-000-111
01-10
1
all others

(4) 5 cells:
1
000
001-011
111
all others

(5) 5 cells:
000-111
001-011
1-10
0-01
all others

(6) 5 cells:
000
111
001
100
all others

(7) 5 cells:
000-00
111-11
001
100
all others
(8) 5 cells:  
000-00  
111-11  
001-011  
100-110  
all others

(9) 5 cells:  
000-00  
111-11  
001-011-01  
100-110-10  
all others

(10) 5 cells:  
000-00-0  
111-11-1  
001-011-01  
100-110-10  
all others

Some further notes about Tables 2 and 3 are given in Section V-C.
<table>
<thead>
<tr>
<th>Parameters</th>
<th>EN</th>
<th>base sample size</th>
<th>Groupings</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_1$</td>
<td>$p_2$</td>
<td>$r$</td>
<td>$a_0$</td>
</tr>
<tr>
<td>.8</td>
<td>.6</td>
<td>.8</td>
<td>.5</td>
</tr>
<tr>
<td>.5</td>
<td>1.56</td>
<td>7888</td>
<td>101.66</td>
</tr>
<tr>
<td>.2</td>
<td>1.59</td>
<td>23764</td>
<td>59.36</td>
</tr>
<tr>
<td>.4</td>
<td>1.55</td>
<td>1016</td>
<td>167.40</td>
</tr>
<tr>
<td>.5</td>
<td>1.61</td>
<td>664</td>
<td>55.47</td>
</tr>
<tr>
<td>.2</td>
<td>1.68</td>
<td>1810</td>
<td>21.02</td>
</tr>
<tr>
<td>.2</td>
<td>1.58</td>
<td>228</td>
<td>249.53</td>
</tr>
<tr>
<td>.5</td>
<td>1.68</td>
<td>147</td>
<td>57.23</td>
</tr>
<tr>
<td>.2</td>
<td>1.78</td>
<td>285</td>
<td>13.56</td>
</tr>
<tr>
<td>.6</td>
<td>1.63</td>
<td>26181</td>
<td>48.09</td>
</tr>
<tr>
<td>.5</td>
<td>1.66</td>
<td>11661</td>
<td>26.75</td>
</tr>
<tr>
<td>.2</td>
<td>1.70</td>
<td>28409</td>
<td>15.96</td>
</tr>
<tr>
<td>.2</td>
<td>1.66</td>
<td>1606</td>
<td>73.10</td>
</tr>
<tr>
<td>.5</td>
<td>1.73</td>
<td>687</td>
<td>27.75</td>
</tr>
<tr>
<td>.2</td>
<td>1.80</td>
<td>1239</td>
<td>8.90</td>
</tr>
<tr>
<td>.4</td>
<td>1.74</td>
<td>23754</td>
<td>23.64</td>
</tr>
<tr>
<td>.5</td>
<td>1.78</td>
<td>8527</td>
<td>13.27</td>
</tr>
<tr>
<td>.2</td>
<td>1.82</td>
<td>16196</td>
<td>6.40</td>
</tr>
<tr>
<td>.8</td>
<td>1.68</td>
<td>11794</td>
<td>137.77</td>
</tr>
<tr>
<td>Parameters</td>
<td>base sample size</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>$p_1$ .8</td>
<td>$p_2$ .6</td>
<td>$r$ .5</td>
<td>$q_0$ .7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>2.33</td>
<td>19172</td>
<td></td>
</tr>
<tr>
<td>.4</td>
<td>2.26</td>
<td>867</td>
<td></td>
</tr>
<tr>
<td>.5</td>
<td>2.40</td>
<td>540</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>2.54</td>
<td>1387</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>2.32</td>
<td>191</td>
<td></td>
</tr>
<tr>
<td>.5</td>
<td>2.55</td>
<td>117</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>2.78</td>
<td>211</td>
<td></td>
</tr>
<tr>
<td>.6</td>
<td>2.43</td>
<td>2172</td>
<td></td>
</tr>
<tr>
<td>.5</td>
<td>2.50</td>
<td>9379</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>2.58</td>
<td>22130</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>2.49</td>
<td>1356</td>
<td></td>
</tr>
<tr>
<td>.5</td>
<td>2.66</td>
<td>545</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>2.82</td>
<td>919</td>
<td></td>
</tr>
<tr>
<td>.5</td>
<td>2.78</td>
<td>6584</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>2.87</td>
<td>12017</td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 2: Increasing Probabilities (B-d)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>EN</th>
<th>base sample size</th>
<th>Groupings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>p_1</strong> .4</td>
<td><strong>p_2</strong> .7</td>
<td><strong>p_3</strong> .8</td>
<td><strong>q_0</strong> .5</td>
</tr>
<tr>
<td>.4</td>
<td>.7</td>
<td>.8</td>
<td>.5</td>
</tr>
<tr>
<td>.5</td>
<td>.8</td>
<td>1.64-1.66</td>
<td>147</td>
</tr>
<tr>
<td>.5</td>
<td>.6</td>
<td>1.65-1.67</td>
<td>428</td>
</tr>
<tr>
<td>.2</td>
<td>.5</td>
<td>.6</td>
<td>1.71-1.74</td>
</tr>
<tr>
<td>.4</td>
<td>.8</td>
<td>1.71-1.74</td>
<td>45</td>
</tr>
<tr>
<td>.4</td>
<td>.6</td>
<td>1.72-1.75</td>
<td>79</td>
</tr>
<tr>
<td>.3</td>
<td>.6</td>
<td>1.73-1.77</td>
<td>99</td>
</tr>
<tr>
<td>.3</td>
<td>.4</td>
<td>1.74-1.78</td>
<td>285</td>
</tr>
<tr>
<td><strong>p_1</strong> .4</td>
<td><strong>p_2</strong> .7</td>
<td><strong>p_3</strong> .8</td>
<td><strong>q_0</strong> .7</td>
</tr>
<tr>
<td>.5</td>
<td>.8</td>
<td>1.91-1.95</td>
<td>107</td>
</tr>
<tr>
<td>.5</td>
<td>.6</td>
<td>1.93-1.98</td>
<td>342</td>
</tr>
<tr>
<td>.2</td>
<td>.5</td>
<td>.6</td>
<td>2.07-2.13</td>
</tr>
<tr>
<td>.4</td>
<td>.8</td>
<td>2.08-2.13</td>
<td>32</td>
</tr>
<tr>
<td>.4</td>
<td>.6</td>
<td>2.10-2.17</td>
<td>62</td>
</tr>
<tr>
<td>.3</td>
<td>.6</td>
<td>2.14-2.21</td>
<td>68</td>
</tr>
<tr>
<td>.3</td>
<td>.4</td>
<td>2.17-2.25</td>
<td>220</td>
</tr>
<tr>
<td>Parameters</td>
<td>base sample size</td>
<td>Groupings</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>-----------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>(p_1), (p_2), (p_3), (q_0), (q_1)</td>
<td>EN</td>
<td>1A</td>
<td>1B</td>
</tr>
<tr>
<td>.4, .7, .8, .7, .5</td>
<td>2.15-2.34</td>
<td>61</td>
<td>838.66</td>
</tr>
<tr>
<td>.5, .8</td>
<td>2.19-2.39</td>
<td>117</td>
<td>472.1</td>
</tr>
<tr>
<td>.5, .6</td>
<td>2.20-2.42</td>
<td>379</td>
<td>103.01</td>
</tr>
<tr>
<td>.2, .5, .6</td>
<td>2.28-2.51</td>
<td>54</td>
<td>2760.90</td>
</tr>
<tr>
<td>.4, .8</td>
<td>2.29-2.51</td>
<td>40</td>
<td>272.50</td>
</tr>
<tr>
<td>.4, .6</td>
<td>2.30-2.54</td>
<td>74</td>
<td>438.29</td>
</tr>
<tr>
<td>.3, .6</td>
<td>2.32-2.57</td>
<td>86</td>
<td>157.87</td>
</tr>
<tr>
<td>.3, .4</td>
<td>2.34-2.60</td>
<td>269</td>
<td>2.03</td>
</tr>
<tr>
<td>Parameters</td>
<td>EN</td>
<td>base sample size</td>
<td>1A</td>
</tr>
<tr>
<td>--------------</td>
<td>--------</td>
<td>------------------</td>
<td>-----</td>
</tr>
<tr>
<td>p1 p2 p3 q0 q1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.8 .7 .4 .5 .3</td>
<td>1.51-1.55</td>
<td>164</td>
<td>8.66</td>
</tr>
<tr>
<td>.5 .4</td>
<td>1.53-1.57</td>
<td>72</td>
<td>143.19</td>
</tr>
<tr>
<td>.4 .2</td>
<td>1.54-1.60</td>
<td>37</td>
<td>64.83</td>
</tr>
<tr>
<td>.6 .5 .4</td>
<td>1.59-1.64</td>
<td>481</td>
<td>29.91</td>
</tr>
<tr>
<td>.5 .2</td>
<td>1.60-1.66</td>
<td>169</td>
<td>9.60</td>
</tr>
<tr>
<td>.4 .2</td>
<td>1.61-1.67</td>
<td>117</td>
<td>28.37</td>
</tr>
<tr>
<td>.3 .2</td>
<td>1.62-1.68</td>
<td>76</td>
<td>76.06</td>
</tr>
<tr>
<td>.4 .3 .2</td>
<td>1.68-1.76</td>
<td>386</td>
<td>20.13</td>
</tr>
<tr>
<td>.8 .7 .4 .7 .3</td>
<td>1.62-1.72</td>
<td>164</td>
<td>17.38</td>
</tr>
<tr>
<td>.5 .4</td>
<td>1.67-1.79</td>
<td>75</td>
<td>343.04</td>
</tr>
<tr>
<td>.4 .2</td>
<td>1.71-1.92</td>
<td>39</td>
<td>78.42</td>
</tr>
<tr>
<td>.6 .5 .4</td>
<td>1.81-1.96</td>
<td>421</td>
<td>94.75</td>
</tr>
<tr>
<td>.5 .2</td>
<td>1.83-2.06</td>
<td>138</td>
<td>16.69</td>
</tr>
<tr>
<td>.4 .2</td>
<td>1.86-2.11</td>
<td>101</td>
<td>45.24</td>
</tr>
<tr>
<td>.3 .2</td>
<td>1.89-2.16</td>
<td>70</td>
<td>151.24</td>
</tr>
<tr>
<td>.4 .3 .2</td>
<td>2.04-2.36</td>
<td>300</td>
<td>50.79</td>
</tr>
<tr>
<td>Parameters</td>
<td>EN</td>
<td>base sample size</td>
<td>Groupings</td>
</tr>
<tr>
<td>------------</td>
<td>----</td>
<td>------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>( p_1 ) ( p_2 ) ( p_3 ) ( q_0 ) ( q_1 )</td>
<td>( 2.02-2.33 )</td>
<td>134</td>
<td>( 6.12 ) ( .66 ) ( 3.45 ) ( 10.59 ) ( 16.62 ) ( 1.03 )</td>
</tr>
<tr>
<td>( .8 ) ( .7 ) ( .4 ) ( .7 ) ( .5 )</td>
<td>( 2.06-2.39 )</td>
<td>70</td>
<td>( 105.49 ) ( 2.24 ) ( 5.90 ) ( 9.96 )</td>
</tr>
<tr>
<td>( .4 ) ( .2 )</td>
<td>( 2.09-2.52 )</td>
<td>36</td>
<td>( 20.20 ) ( 6.22 ) ( 5.59 ) ( 11.54 )</td>
</tr>
<tr>
<td>( .6 ) ( .5 ) ( .4 )</td>
<td>( 2.14-2.49 )</td>
<td>425</td>
<td>( 26.25 ) ( .50 ) ( 3.53 ) ( 20.36 )</td>
</tr>
<tr>
<td>( .5 ) ( .2 )</td>
<td>( 2.15-2.60 )</td>
<td>133</td>
<td>( 4.43 ) ( 1.60 ) ( 2.89 ) ( 135.16 )</td>
</tr>
<tr>
<td>( .4 ) ( .2 )</td>
<td>( 2.17-2.63 )</td>
<td>102</td>
<td>( 9.99 ) ( 3.18 ) ( 3.57 ) ( 37.10 )</td>
</tr>
<tr>
<td>( .3 ) ( .2 )</td>
<td>( 2.19-2.67 )</td>
<td>71</td>
<td>( 25.31 ) ( 7.53 ) ( 4.59 ) ( 22.60 )</td>
</tr>
<tr>
<td>( .4 ) ( .3 ) ( .2 )</td>
<td>( 2.27-2.78 )</td>
<td>339</td>
<td>( 7.24 ) ( 2.53 ) ( 3.32 ) ( 9.96 )</td>
</tr>
<tr>
<td>Parameters</td>
<td>base sample size</td>
<td>Groupings</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>a  b q₀ q₁</td>
<td>2 3 4 5 6 7 8 9 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2 0.4 0.7 0.5</td>
<td>468 1.21 10.02 2.32 4.10 2.69 5.32 7.44 0.82 2.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>112 1.37 4.93 2.50 2.75 4.10 12.91 15.06 0.95 3.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>46 1.52 22.25 4.03 3.35 9.98 63.39 35.78 1.16 4.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4 0.2</td>
<td>489 0.97 8.59 1.46 5.76 1.32 1.23 1.20 0.76 1.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>458 1.17 285.11 1.76 1.85 4.40 237.73 145.62 0.90 2.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>108 1.28 121.67 2.91 2.83 12.42 3.77 1.58 1.08 3.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6 0.2</td>
<td>120 0.88 2.01 1.11 2.43 1.11 0.88 0.86 0.63 1.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>468 0.94 6.63 1.19 1.50 1.52 0.87 0.85 0.70 1.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>446 1.11 50.10 2.46 4.99 14.95 4.33 1.25 1.03 3.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.8 0.2</td>
<td>48 0.82 2.90 2.90 2.87 1.04 0.78 0.76 0.58 1.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>112 0.85 15.27 1.01 8.26 1.44 0.83 0.84 0.65 2.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>454 0.90 32.31 1.29 14.77 2.51 0.85 0.95 0.73 3.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2 0.4 0.7 0.3</td>
<td>565 1.29 1.70 2.85 2.91 3.84 12.26 1.23 0.86 1.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>131 1.32 4.60 2.82 2.81 5.39 2.12 1.92 1.02 2.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>52 1.35 21.59 4.37 4.27 11.60 4.63 3.34 1.23 4.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4 0.2</td>
<td>629 1.05 1.97 1.87 3.00 2.06 1.17 1.11 0.72 1.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>523 1.14 217.95 2.02 2.09 5.32 2.87 2.00 1.13 2.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>118 1.17 173.86 3.33 3.68 13.72 7.05 3.17 1.09 4.52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parameters</td>
<td>base sample size</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>------------</td>
<td>------------------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>a b q₀ q₁</td>
<td>152</td>
<td>.91</td>
<td>.91</td>
</tr>
<tr>
<td>.6 .2 .7 .3</td>
<td>551</td>
<td>.96</td>
<td>110.82</td>
</tr>
<tr>
<td>.8</td>
<td>469</td>
<td>1.03</td>
<td>96.45</td>
</tr>
<tr>
<td>.8 .2 .7 .3</td>
<td>59</td>
<td>.83</td>
<td>5.67</td>
</tr>
<tr>
<td>.4</td>
<td>128</td>
<td>.86</td>
<td>20.45</td>
</tr>
<tr>
<td>.6</td>
<td>491</td>
<td>.89</td>
<td>38.06</td>
</tr>
<tr>
<td>.2 .4 .5 .3</td>
<td>597</td>
<td>1.04</td>
<td>6.36</td>
</tr>
<tr>
<td>.6</td>
<td>142</td>
<td>1.14</td>
<td>.86</td>
</tr>
<tr>
<td>.8</td>
<td>59</td>
<td>1.25</td>
<td>6.25</td>
</tr>
<tr>
<td>.4 .2 .5 .3</td>
<td>625</td>
<td>1.94</td>
<td>6.84</td>
</tr>
<tr>
<td>.6</td>
<td>562</td>
<td>1.04</td>
<td>2.71</td>
</tr>
<tr>
<td>.8</td>
<td>133</td>
<td>1.13</td>
<td>4.46</td>
</tr>
<tr>
<td>.6 .2 .5 .3</td>
<td>151</td>
<td>.85</td>
<td>1.61</td>
</tr>
<tr>
<td>.4</td>
<td>569</td>
<td>.89</td>
<td>1.63</td>
</tr>
<tr>
<td>.8</td>
<td>550</td>
<td>1.01</td>
<td>3.51</td>
</tr>
<tr>
<td>.8 .2 .5 .3</td>
<td>61</td>
<td>.81</td>
<td>1.17</td>
</tr>
<tr>
<td>.4</td>
<td>134</td>
<td>.83</td>
<td>1.23</td>
</tr>
<tr>
<td>.6</td>
<td>530</td>
<td>.86</td>
<td>1.56</td>
</tr>
<tr>
<td>Parameters</td>
<td>p</td>
<td>α</td>
<td>q</td>
</tr>
<tr>
<td>------------</td>
<td>----</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>0.6, 0.2, 0.2</td>
<td>1.11</td>
<td>5247</td>
<td>4.27</td>
</tr>
<tr>
<td>0.2</td>
<td>1.20</td>
<td>11052</td>
<td>48.93</td>
</tr>
<tr>
<td>0.6, 0.4</td>
<td>1.14</td>
<td>3783</td>
<td>2.47</td>
</tr>
<tr>
<td>0.2</td>
<td>1.21</td>
<td>9655</td>
<td>22.14</td>
</tr>
<tr>
<td>0.6, 0.6</td>
<td>1.17</td>
<td>4673</td>
<td>1.73</td>
</tr>
<tr>
<td>0.2</td>
<td>1.22</td>
<td>14242</td>
<td>12.83</td>
</tr>
<tr>
<td>0.6, 0.8</td>
<td>1.20</td>
<td>11400</td>
<td>1.29</td>
</tr>
<tr>
<td>0.2</td>
<td>1.23</td>
<td>41828</td>
<td>7.70</td>
</tr>
<tr>
<td>0.6, 0.4, 0.2</td>
<td>1.25</td>
<td>1074</td>
<td>4.46</td>
</tr>
<tr>
<td>0.2</td>
<td>1.49</td>
<td>2555</td>
<td>39.18</td>
</tr>
<tr>
<td>0.6, 0.4</td>
<td>1.31</td>
<td>785</td>
<td>2.28</td>
</tr>
<tr>
<td>0.2</td>
<td>1.52</td>
<td>2292</td>
<td>15.19</td>
</tr>
<tr>
<td>0.6, 0.6</td>
<td>1.39</td>
<td>957</td>
<td>1.44</td>
</tr>
<tr>
<td>0.2</td>
<td>1.56</td>
<td>3465</td>
<td>7.19</td>
</tr>
<tr>
<td>0.6, 0.8</td>
<td>1.50</td>
<td>2217</td>
<td>0.89</td>
</tr>
<tr>
<td>0.2</td>
<td>1.60</td>
<td>10278</td>
<td>3.38</td>
</tr>
<tr>
<td>0.6, 0.2</td>
<td>1.41</td>
<td>349</td>
<td>5.16</td>
</tr>
<tr>
<td>0.2</td>
<td>1.97</td>
<td>914</td>
<td>33.94</td>
</tr>
<tr>
<td>0.6, 0.4</td>
<td>1.53</td>
<td>265</td>
<td>2.30</td>
</tr>
<tr>
<td>0.2</td>
<td>2.03</td>
<td>833</td>
<td>11.27</td>
</tr>
<tr>
<td>Parameters</td>
<td>base sample size</td>
<td>Groupings</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>-----------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>p α q EN</td>
<td>1 2 3 4</td>
<td>5 6 7 8 9 10</td>
<td></td>
</tr>
<tr>
<td>.6 .6 .6 1.69</td>
<td>324 1.25 1.17 55.45 23874.01 829.64 25.15 23.53 3.42 .83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.2 2.12 1285</td>
<td>4.24 1.61 1.31 118.38 16.85 452.52 10.16 10.39 1.64 3.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.6 .8 1.94</td>
<td>720 .62 .92 389.84 11.20 10164.91 3.50 2.69 .85 .89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.6 .8 .2 1.62</td>
<td>119 7.11 4.04 33.10 442.06 88.61 266.39 239.27 2.41 .88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.2 2.88 306</td>
<td>35.88 7.00 8.41 189.46 36.80 249.26 2070.78 1556.68 2.43 4.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.6 .4 1.82</td>
<td>97 2.67 2.14 64.32 1077.49 33.25 93.87 48.74 1.23 .73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.2 3.02 272</td>
<td>10.46 3.17 3.36 427.46 17.01 24.88 61.23 51.33 2.26 2.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.6 .6 2.10</td>
<td>121 1.20 1.35 152.55 48.11 31.03 4.42 2.03 .70 .68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.2 3.24 413</td>
<td>3.12 1.68 1.98 472.74 9.36 7.84 10.38 9.20 2.58 1.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.6 .8 2.63</td>
<td>256 .46 .96 4.67 5.39 4.34 2.60 .90 .85 .76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.2 3.64 1248</td>
<td>.66 1.04 1.43 16.74 5.10 8.35 6.14 5.37 3.96 1.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parameters</td>
<td>base sample size</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>p .6 .4 .2</td>
<td>1.44</td>
<td>1448</td>
<td>133.01</td>
</tr>
<tr>
<td>p .2</td>
<td>1.50</td>
<td>1977</td>
<td>827.67</td>
</tr>
<tr>
<td>p .6 .6</td>
<td>1.71</td>
<td>381</td>
<td>239.80</td>
</tr>
<tr>
<td>p .2</td>
<td>1.36</td>
<td>497</td>
<td>1675.85</td>
</tr>
<tr>
<td>p .6 .8</td>
<td>2.11</td>
<td>126</td>
<td>474.88</td>
</tr>
<tr>
<td>p .2</td>
<td>1.42</td>
<td>182</td>
<td>3356.21</td>
</tr>
<tr>
<td>p .6 .2 .4</td>
<td>1.43</td>
<td>596</td>
<td>4.72</td>
</tr>
<tr>
<td>p .2</td>
<td>1.59</td>
<td>1087</td>
<td>36.01</td>
</tr>
<tr>
<td>p .6 .6</td>
<td>1.97</td>
<td>751</td>
<td>17.88</td>
</tr>
<tr>
<td>p .2</td>
<td>1.74</td>
<td>1133</td>
<td>127.01</td>
</tr>
<tr>
<td>p .6 .8</td>
<td>2.43</td>
<td>136</td>
<td>35.30</td>
</tr>
<tr>
<td>p .6 .2 .6</td>
<td>1.82</td>
<td>230</td>
<td>233.80</td>
</tr>
<tr>
<td>p .2</td>
<td>1.81</td>
<td>129</td>
<td>1.43</td>
</tr>
<tr>
<td>p .2</td>
<td>2.29</td>
<td>252</td>
<td>12.04</td>
</tr>
<tr>
<td>p .6 .4</td>
<td>2.10</td>
<td>557</td>
<td>2.12</td>
</tr>
<tr>
<td>p .2</td>
<td>2.39</td>
<td>969</td>
<td>17.56</td>
</tr>
<tr>
<td>p .6 .8</td>
<td>3.07</td>
<td>337</td>
<td>6.45</td>
</tr>
<tr>
<td>p .2</td>
<td>2.61</td>
<td>621</td>
<td>46.87</td>
</tr>
<tr>
<td>Parameters</td>
<td>base sample size</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>.6 .2 .8</td>
<td>2.95</td>
<td>.76</td>
<td>.89</td>
</tr>
<tr>
<td>.2</td>
<td>4.37</td>
<td>5.39</td>
<td>1.10</td>
</tr>
<tr>
<td>.6 .4</td>
<td>3.42</td>
<td>.87</td>
<td>.90</td>
</tr>
<tr>
<td>.2</td>
<td>4.56</td>
<td>6.91</td>
<td>1.13</td>
</tr>
<tr>
<td>.6 .6</td>
<td>4.06</td>
<td>1.14</td>
<td>.92</td>
</tr>
<tr>
<td>.2</td>
<td>4.77</td>
<td>9.84</td>
<td>1.30</td>
</tr>
</tbody>
</table>
REFERENCES


