UNIVERSITY OF NORTH CAROLINA
Department of Statistics
Chapel Hill, N. C.

SEVERAL TEST PROCEDURES (SEQUENTIAL AND
NON-SEQUENTIAL) FOR NORMAL MEANS

by
Michael D. Hogan

December 1963

This research was supported by Public Health Service
Research Grant GM-10397-01, from the Division of
General Medical Sciences, National Institutes of
Health.

Institute of Statistics
Mimeo Series No. 379
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>CONTENTS</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1. Summary</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2. Statement of Problem and Background</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Discussion</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3. General Intersection Test Procedure</td>
<td>10</td>
</tr>
<tr>
<td>II</td>
<td>DEVELOPMENT OF THE THREE-DECISION PROCEDURES</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>1. The Sobel-Wald Procedure</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>2. The Fixed Sample Size Test</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>3. The Stein Two-Stage Procedure</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>4. The Baker-Hall Sequential Analog of Stein's Two-Stage Procedure</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>5. The Minimum Probability Ratio Test Procedure</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>6. Hall's Modification of the M.P.R.T. for the Case of Unknown Variance</td>
<td>44</td>
</tr>
<tr>
<td>III</td>
<td>EMPIRICAL EVALUATIONS</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>BIBLIOGRAPHY</td>
<td>63</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

I am deeply indebted to Professor William Jackson Hall, not only for proposing my thesis topic and conscientiously directing my research, but also for guiding me through the more trying moments of my graduate studies. He has been a great source of personal warmth and inspiration for me.

I would also like to thank Professor Normal L. Johnson for his many valuable comments and suggestions. I am very fortunate to have had the opportunity to avail myself of his abundant talent.

Finally, I would like to thank my father, for without his continued support and encouragement, I would never have reached this phase of my academic career.
1. Summary

In this paper the problem of choosing one of three hypotheses about the unknown mean of a normal distribution when there are specified bounds on the error probabilities is considered. Three-decision analogs of the fixed sample size test, Wald sequential probability ratio test, Hall minimum probability ratio test, Stein two-stage test, Baker-Hall sequential analog of Stein's two-stage test, and the modified minimum probability ratio test are developed by use of an intersection technique suggested by Sobel and Wald (1949). The first three procedures are applicable when the variance is known; the latter three are of use in the case of unknown variance. In addition either exact or approximate formulas for the operating characteristic functions are presented for several of the procedures. Finally, the results of an empirical sampling investigation of the terminal sample sizes required in three different test cases by each of the six procedures mentioned above are presented. The results indicate that, as in the case of the two-decision procedures, (i) the three-decision sequential procedures lead to a substantial savings in the average amount of required sampling,
(ii) the truncated sequential procedures (i.e., the M.P.R.T. and its modification for the case of unknown variance) are much less variable than the non-truncated or two-stage procedures, and

(iii) all of the procedures tend to be somewhat conservative w.r.t. their error bounds.

2. **Statement of Problem and Background Discussion.**

Suppose that one is sampling from a normal distribution with unknown mean, $\theta$, and that one wishes to test the composite hypotheses

$$H^-: \theta \leq \theta_1 \text{ vs. } H_0: \theta_2 \leq \theta \leq \theta_3 \text{ vs. } H^+: \theta \geq \theta_4,$$

where

(i) $\theta_1 < \theta_2 < \theta_3 < \theta_4$, and

(ii) $\theta_2 - \theta_1 = \theta_4 - \theta_3 = \Delta$, say .

Thus, one is dealing with a three-decision problem, i.e., one of the three hypotheses is to be "accepted" and the remaining two "rejected". Alternatively, this problem is frequently treated as a two-sided, two-decision problem by merging $H^-$ and $H^+$ into a single hypothesis $H_1$ to be tested against $H_0$.

Suppose, further, that it is desired that the test be constructed so as to ensure that the following bounds on the error probabilities will be satisfied:
Pr \{\text{rejecting } H^-|\theta \} \leq \gamma_1, \text{ for } \theta \leq \theta_1;

(1) \quad \text{Pr} \{\text{rejecting } H_0|\theta \} \leq \gamma_2, \text{ for } \theta_2 \leq \theta \leq \theta_3; \text{ and}

\quad \text{Pr} \{\text{rejecting } H^+|\theta \} \leq \gamma_3, \text{ for } \theta \geq \theta_4,

for specified \( \gamma_1, \gamma_2, \gamma_3 \). No requirements are placed on the decision probabilities for \( \theta \)-values between \( \theta_1 \) and \( \theta_2 \) and between \( \theta_3 \) and \( \theta_4 \); these intervals might thus be termed "indifference regions".

To illustrate how such a problem could arise in a practical situation, it might be illuminating to consider an example from the field of quality control. Suppose that a contractor who is engaged in the production of concrete for highway construction wants to determine the mean breaking point, \( \theta \), of his product. Suppose, also, that past experience indicates that (i) if the mean breaking point falls below 1800 lbs./in\(^2\) the composition of the concrete will be materially altered so as to render it unsuitable for normal highway traffic, (ii) if the mean breaking point lies between 1900 lbs./in\(^2\) and 2100 lbs./in\(^2\), the concrete will be able to withstand the stresses of a normal load of traffic, and (iii) if the mean breaking point exceeds 2200 lbs./in\(^2\), the concrete will contain an unnecessarily large proportion of cement, thereby raising production costs. Under such conditions it seems likely that the contractor's quality control department might want to test the hypotheses

\[ H^-: \theta \leq 1800 \text{ vs. } H_0: 1900 \leq \theta \leq 2100 \text{ vs. } H^+: \theta \geq 2200, \]
with the bounds $\gamma_1, \gamma_2, \gamma_3$, on the error probabilities being determined by various economic considerations.

In this paper we shall consider six different procedures for testing $H^- \text{ vs. } H^+_0 \text{ vs. } H^+$ so that the guarantees in (1) are satisfied. Three of these procedures are suitable for the case of known variance $\sigma^2$; the remaining three being applicable when $\sigma^2$ is unknown. Of the first three, one is non-sequential and two are strictly sequential (one of which is truncated). The latter three are analogs of the first three procedures, except that an initial sample is taken in each largely for the purpose of estimating the unknown variance. Each of these procedures is developed from a uniform approach (see Chapter I, Section 3), which may be described simply as the intersection of two one-sided test procedures. This approach is essentially due to Sobel and Wald [8].

The Sobel-Wald procedure (see Chapter II, Section 1) is a sequential test procedure that is applicable to the preceding general problem when $\sigma^2$ is known. Essentially this technique is based on the simultaneous performance of two Wald S.P.R.T.'s, one of $\theta_1 \text{ vs. } \theta_2$ and the other of $\theta_3 \text{ vs. } \theta_4$. The final decision with respect to $H^-, H^+_0$, and $H^+$ depends on the outcomes of the respective sub-tests, which are designed so as to ensure the error bounds in (1).

*Numbers in square brackets refer to the bibliography.
It should be noted that the proposed procedures all tend to give a weak inequality for the bound on Pr \{rejecting $H_0$\} over the interval $[\theta_2, \theta_3]$ unless the difference, $\theta_3 - \theta_2$, is quite small (see Chapter I, Section 3). Thus, in this respect, they are rather conservative. A less conservative, three-decision fixed-sample size test can be constructed for the case where $(\theta_3 - \theta_2)$ is significantly greater than zero (see Chapter II, Section 2) by departing from the general method of development employed throughout this paper.

In section 1 of Chapter II the Sobel-Wald technique, denoted by I.S.P.R.T., is reviewed in some detail; and in the remaining sections of Chapter II three-decision analogs are successively developed for the fixed-sample size test, denoted by I.F.S.S.T., Stein's two-stage test [9], designated by I.S.T.S.T.\textsuperscript{1}, the Baker-Hall sequential analog of Stein's two-stage test [2, 5], designated by I.S.P.R.T.\textsubscript{(m)}, the minimum probability ratio test [4], denoted by I.M.P.R.T., and Hall's modification of the minimum probability ratio test for the case of unknown variance [5], designated by I.M.P.R.T.\textsuperscript{(m)}. Thus, three-decision analogs of the following one-sided test procedures are treated:

<table>
<thead>
<tr>
<th>Sampling procedure</th>
<th>$\sigma^2$ known</th>
<th>$\sigma^2$ unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>one-or two-stage</td>
<td>F.S.S.T.</td>
<td>S.T.S.T.\textsubscript{(m)}</td>
</tr>
<tr>
<td>sequential</td>
<td>S.P.R.T.</td>
<td>S.P.R.T.\textsubscript{(m)}</td>
</tr>
<tr>
<td></td>
<td>M.P.R.T.</td>
<td>M.P.R.T.\textsubscript{(m)}</td>
</tr>
</tbody>
</table>

\textsuperscript{1}The three-decision analog of Stein's two-stage procedure is denoted by I.S.T.S.T.\textsubscript{(m)}, since it is formed by intersecting two one-sided S.T.S.T.\textsubscript{(m)} procedures and since the unknown variance is estimated from an initial sample of size $m$. 
Finally, Chapter III presents the results of an empirical sampling investigation of the terminal sample sizes required in three different test cases by each of the six procedures under consideration.

To see when the above test procedures might be applied to a practical problem, let's reconsider the contractor engaged in the mixing of concrete for highway construction. If the variance of the concrete's breaking point is known (perhaps the quality control department has at its disposal a substantial quantity of pervious inspection data), then an I.F.S.S.T., I.S.P.R.T., or I.M.P.R.T. \(^2\) could be utilized to test \(H^- vs H_0 vs. H^+\). On the other hand if the variance is not known, then an I.S.T.S.T.(m), I.S.P.R.T.(m), or I.M.P.R.T.(m) could be employed. The relative merits of these test procedures are not known explicitly. Since the error probabilities have been controlled in the same way for each of the procedures, they do not provide much basis for choosing among the various tests. Furthermore, although the minimum sample size needed to ensure that the guarantees stated in (1) will be satisfied can readily be determined for the one- and two-stage procedures, little is known about the average sample number functions (A.S.N.) for the sequential and modified sequential techniques. [Sobel and Wald were able to obtain a lower bound on the A.S.N. for the three-decision S.P.R.T. which seems to be a fairly close approximation to the A.S.N. except over the range \((c_2, c_2)\). By considering a slightly

\(^2\)It should be noted that the I.M.P.R.T. and I.M.P.R.T.(m) are applicable only when \(\gamma_1 = \gamma_2 = \gamma_2/2\).
less general problem they were also able to derive a rather complicated expression for an upper bound on the A.S.N. that appears to provide a reasonably close approximation when certain restrictions are satisfied.] Numerical and theoretical studies that have been made of the one-sided test procedures listed in the table on page 5 seem to indicate that if one of the (two) hypotheses being tested is true, then one can expect substantial savings on the average by using a suitable analog of a S.P.R.T. instead of the comparable one- or two-stage procedures. However, even though the average sample numbers (sizes) will not be too large with such a procedure, actual sample numbers will in fact be random and also unbounded. On the other hand the truncated sequential procedures, i.e., the M.P.R.T. and the M.P.R.T.(m), tend to provide more control over the variability and the maximum amount of sampling than do the non-truncated procedures. Furthermore, their A.S.N.'s tend to be a little flatter than those of the S.P.R.T.'s -- slightly lower between the hypotheses being tested and slightly higher when one of the hypotheses is true. Hence, the truncated sequential tests may be preferable to the non-truncated techniques, especially if neither of the hypotheses under consideration is true.

Now, the empirical investigations summarized in Chapter III indicate that these same general conclusions can be drawn for the various three-decision procedures being discussed in this paper. A portion of the data contained in Chapter III is presented below.
Here, "s.d." refers to the standard deviations of the required sample sizes. Each of these numbers is based on 165 repetitions of the tests, so all of the figures are reasonably precise. (The standard deviations of the averages range from about 0.5 to about 2.5.)

Test Case 1

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Correct Hypothesis</th>
<th>Median Sample Size</th>
<th>s.d.</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.F.S.S.T.</td>
<td>( H^+ )</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>I.S.P.R.T.</td>
<td>( H^- )</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>I.M.P.R.T.</td>
<td>( H^- )</td>
<td>23</td>
<td>11</td>
</tr>
<tr>
<td>I.S.T.S.T.(_m)</td>
<td>( H^+ )</td>
<td>43</td>
<td>19</td>
</tr>
<tr>
<td>I.S.P.R.T.(_m)</td>
<td>( H^- )</td>
<td>24</td>
<td>21</td>
</tr>
<tr>
<td>I.M.P.R.T.(_m)</td>
<td>( H^- )</td>
<td>25</td>
<td>15</td>
</tr>
</tbody>
</table>

Test Case 2

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Correct Hypothesis</th>
<th>Median Sample Size</th>
<th>s.d.</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.F.S.S.T.</td>
<td>( H_0 )</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>I.S.P.R.T.</td>
<td>( H_0 )</td>
<td>23</td>
<td>12</td>
</tr>
<tr>
<td>I.M.P.R.T.</td>
<td>( H_0 )</td>
<td>24</td>
<td>7</td>
</tr>
<tr>
<td>I.S.T.S.T.(_m)</td>
<td>( H_0 )</td>
<td>43</td>
<td>19</td>
</tr>
<tr>
<td>I.S.P.R.T.(_m)</td>
<td>( H_0 )</td>
<td>26</td>
<td>17</td>
</tr>
<tr>
<td>I.M.P.R.T.(_m)</td>
<td>( H_0 )</td>
<td>27</td>
<td>13</td>
</tr>
</tbody>
</table>

\( m = 16 \)
### Test Case 3

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Correct Hypothesis</th>
<th>Median Sample Size</th>
<th>s.d.</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.F.S.S.T.</td>
<td>none</td>
<td>44</td>
<td>---</td>
</tr>
<tr>
<td>I.S.P.R.T.</td>
<td>none</td>
<td>26</td>
<td>22</td>
</tr>
<tr>
<td>I.M.P.R.T.</td>
<td>none</td>
<td>29</td>
<td>12</td>
</tr>
<tr>
<td>I.S.T.S.T.(m)</td>
<td>none</td>
<td>43</td>
<td>19</td>
</tr>
<tr>
<td>I.S.P.R.T.(m)</td>
<td>none</td>
<td>34</td>
<td>31</td>
</tr>
<tr>
<td>I.M.P.R.T.(m)</td>
<td>none</td>
<td>31</td>
<td>19</td>
</tr>
</tbody>
</table>

It should be noted that the I.S.T.S.T.(m), I.S.P.R.T.(m), and the I.M.P.R.T.(m) all require the choice of an initial sample size of $m$. This choice is arbitrary, but in every case the estimation of $\sigma^2$ is based solely on this sample. If a small value is chosen for $m$, then $\sigma^2$ may be poorly estimated and the rules for continuing sampling badly chosen as a consequence; on the other hand, if a large value is chosen for $m$, one may be doing more sampling than is necessary. Some discussion of this choice of the initial sample size for the S.T.S.T.(m) procedure appears in [7]. As a rule of thumb, one might guess at the true value of $\sigma^2$ and then choose $m$ to be roughly $1/4$ to $1/2$ the required sample size for the corresponding I.F.S.S.T.

Finally, an alternative and commonly used approach to testing $H^- \text{ vs. } H_0 \text{ vs. } H^+$ which may be employed in place of the
I.S.T.S. T.\(t(m)\), the I.S.P.R.T.\(t(m)\), or I.M.P.R.T.\(t(m)\) is to "re-formulate" the hypotheses in terms of (unknown) standard deviation units, i.e.,

\[ H^- : \theta/\sigma \leq \theta_1^t \text{ vs. } H_0 : \theta_2^t \leq \theta/\sigma \leq \theta_3^t \text{ vs. } H^+ : \theta/\sigma \geq \theta_4^t, \]

for specified \(\theta_1^t, \theta_2^t, \theta_3^t\), and \(\theta_4^t\), or else to permit the error probabilities to vary with the unknown \(\sigma^2\). Then a three-decision sequential (or non-sequential) \(t\)-test could be used. Neither of these approaches may be completely satisfactory, and they will not be considered further here. It should be noted, however, that non-sequential tests meeting the guarantees in (1) are not possible when \(\sigma^2\) is unknown unless one makes such a reformulation of the problem.


In the preceding section it was noted that each of the procedures considered in this paper was to be developed from a uniform approach -- namely, the intersecting of two two-decision tests. More explicitly, this intersection technique may be thought of as a four-step process.

First, the original variables, \(X_1, X_2, \ldots\), are transformed by setting

\[ X_1^- = (X_1 - \theta_2)/(\theta_1 - \theta_2), \text{ and } X_1^+ = (X_1 - \theta_3)/(\theta_4 - \theta_3). \]

Thus, the p.d.f. of \(X_1^- = \mathcal{N}((\theta - \theta_2)/(\theta_1 - \theta_2), \sigma^2/(\theta_1 - \theta_2)^2) = \mathcal{N}(\theta^-, (\sigma^2)^2), \text{ say,}\]
and

\[ p.d.f. \ of \ x_1^+ = N(\theta - \theta_3)/(\theta_4 - \theta_3),\ \sigma^2/(\theta_4 - \theta_3)^2) = N(\theta^+, (\sigma^+)^2), \]  

Next, two sub-tests \( T^- \) and \( T^+ \) are constructed so that \( T^- \) is a test, whose specific form depends on the particular three-decision procedure under consideration, of \( \theta^- = 0 \) vs. \( \theta^- = 1 \), with pre-assigned strength \((\alpha^-, \beta^-)\), and \( T^+ \) is a test of \( \theta^+ = 0 \) vs. \( \theta^+ = 1 \), with pre-assigned strength \((\alpha^+, \beta^+)\). [Because of the monotonicity of the power function for a normal mean test, it follows that an \( \alpha^- \)-level test of \( \theta^- = 0 \) vs. \( \theta^- = 1 \) is also an \( \alpha^- \)-level test of \( \theta^- \leq 0 \) vs. \( \theta^- \geq 1 \). Similarly, an \( \alpha^+ \)-level test of \( \theta^+ = 0 \) vs. \( \theta^+ = 1 \) is also an \( \alpha^+ \)-level test of \( \theta^+ \leq 0 \) vs. \( \theta^+ \geq 1 \).]

Then sub-tests \( T^- \) and \( T^+ \) are performed simultaneously, and the testing procedure is continued until decision \( d^-_0 (\theta^- = 0) \) or decision \( d^-_1 (\theta^- = 1) \) has been made under \( T^- \) and decision \( d^+_0 (\theta^+ = 0) \) or decision \( d^+_1 (\theta^+ = 1) \) has been made under \( T^+ \).

Finally, an overall test \( T \) is defined as the intersection of sub-tests \( T^- \) and \( T^+ \), with sampling rule as described in the preceding paragraph and with terminal decisions \( D^- (\theta \leq \theta_1), D_0 (\theta_2 \leq \theta \leq \theta_3), \) and \( D^+ (\theta \geq \theta_4) \). Since \( T = T^- \cap T^+ \), the following combinations of sub-decisions could arise with test \( T \):

\[(d^-_1 \cap d^+_0), \quad (d^-_0 \cap d^+_1), \quad (d^-_0 \cap d^+_0), \quad (d^-_1 \cap d^+_1).\]

Expressing the various sub-decisions in their corresponding equivalent forms of inequalities involving \( \theta \), one can see that making
decision

\[ \begin{align*}
 d^-_1 (\theta \leq \theta_1) & \quad \left\{ \begin{array}{l}
 d^+_o (\theta \leq \theta_3) \\
 d^-_o (\theta \geq \theta_2) \\
 d^+_o (\theta \geq \theta_2)
\end{array} \right\} \\
 d^-_o (\theta \geq \theta_2)
\end{align*} \]

with test \( T^- \), and making decision \( d^-_1 (\theta \leq \theta_1) \)

\[ \begin{align*}
 d^+_o (\theta \leq \theta_3) & \quad \left\{ \begin{array}{l}
 d^+_o (\theta \leq \theta_3) \\
 d^+_1 (\theta \geq \theta_4)
\end{array} \right\}
\end{align*} \]

with test \( T^+ \), would logically lead to a terminal decision of

\[ \begin{align*}
 d^-_1 (\theta \leq \theta_1) & \quad \left\{ \begin{array}{l}
 d^+_o (\theta \leq \theta_3) \\
 d^+_0 (\theta \leq \theta_3) \\
 d^+_1 (\theta \geq \theta_4)
\end{array} \right\} \\
 d^+_1 (\theta \geq \theta_4)
\end{align*} \]

for test \( T \). Furthermore, it is apparent that the sub-tests \( T^- \) and \( T^+ \) should be designed so as to render the making of both decision \( d^-_1 \) and decision \( d^+_1 \) under test \( T \) impossible, since no reasonable conclusion can be drawn from such a combination. With this restriction on the sub-tests, the terminal decisions for \( T \) can be expressed as

\[ D^- = d^-_1, \quad D_o = d^-_o \cap d^+_o, \quad \text{and} \quad D^+ = d^+_1. \]

Now,

\[ \Pr[D^- \text{ using } T | \theta] = \Pr[d^-_1 \text{ using } T^- | \theta] \geq 1 - \beta^-, \text{ for } \theta \leq \theta_1; \]

\[ \Pr[D^+ \text{ using } T | \theta] = \Pr[d^+_1 \text{ using } T^+ | \theta] \geq 1 - \beta^+, \text{ for } \theta \geq \theta_4; \] and

\[ \Pr[D_o \text{ using } T | \theta] = \Pr[(d^-_o \cap d^+_o) | \theta] = 1 - \Pr[d^-_1 \text{ using } T^- | \theta] \]

\[ - \Pr[d^+_1 \text{ using } T^+ | \theta] \geq 1 - \alpha^- - \alpha^+, \text{ for } \theta_2 \leq \theta \leq \theta_3. \]
Thus it is obvious that the bounds on the pre-assigned error probabilities given in (1) will be obtained using $T$ if $\beta^-, \beta^+, \alpha^-, \alpha^+$ are chosen so that

$$\beta^- \leq \gamma_1, \beta^+ \leq \gamma_3, \text{ and } \alpha^- + \alpha^+ \leq \gamma_2, \text{ i.e.,}$$

$$\alpha^+ \leq \gamma_2^+, \text{ where } \gamma_2^- + \gamma_2^+ = \gamma_2, \text{ and } \gamma_2^+ \geq 0.$$

Now, even though there is some loss of generality, we shall for convenience take

$$\gamma_2^- = \gamma_2^+ = \gamma_2/2$$

throughout this paper. It should be noted that the inequality on

$$\Pr[D_0 \text{ using } T\theta]$$

is weak if $\theta_2 < \theta_3$, since

$$\max_{\theta_2 \leq \theta \leq \theta_3} \Pr[d^-_1 \text{ using } T^-|\theta] = \Pr[d^-_1 \text{ using } T^-|\theta = \theta_2] = \alpha^-,$$

and

$$\max_{\theta_2 \leq \theta \leq \theta_3} \Pr[d^+_1 \text{ using } T^+|\theta] = \Pr[d^+_1 \text{ using } T^+|\theta = \theta_3] = \alpha^+.$$

Finally, analogous to the power function or operation characteristic function of a two-decision test procedure, one can obtain exact or approximate expressions for the respective decision probabilities, $\Pr[D^- \text{ using } T\theta], \Pr[D^+ \text{ using } T\theta], \text{ and } \Pr[D_0 \text{ using } T\theta]$, if the sub-test power functions for the three-decision procedures under consideration are available, i.e.,
\[
\begin{align*}
\Pr(D^- \text{ using } T|\theta) &= \Pr(d^-_1 \text{ using } T^-|\theta), \\
\Pr(D^+ \text{ using } T|\theta) &= \Pr(d^+_1 \text{ using } T^+|\theta), \text{ and} \\
\Pr(D_0 \text{ using } T|\theta) &= 1 - \Pr(d^-_1 \text{ using } T^-|\theta) - \Pr(d^+_1 \text{ using } T^+|\theta).
\end{align*}
\]
CHAPTER II
DEVELOPMENT OF THE THREE-DECISION PROCEDURES

1. The Sobel-Wald Procedure

The I.S.P.R.T. [8] is a three-decision sequential procedure which is applicable to the general problem described in Section 2 of Chapter I when the variance, \( \sigma^2 \), is known. In the Sobel-Wald procedure we let \( T^- \) denote the S.P.R.T. of \( \theta^- = 0 \) vs. \( \theta^- = 1 \), based on \( X_1, X_2, \ldots \), with pre-assigned strength \((\alpha^-, \beta^-)\) and termination boundaries \( B^- \) and \( A^- \); and we let \( T^+ \) denote the S.P.R.T. of \( \theta^+ = 0 \) vs. \( \theta^+ = 1 \), based on \( X_1^+, X_2^+, \ldots \), with pre-assigned strength \((\alpha^+, \beta^+)\) and termination boundaries \( B^+ \) and \( A^+ \), where

\[
\begin{align*}
    b^- &= \log B^-, \quad a^- = \log A^-, \quad b^+ = \log B^+, \quad a^+ = \log A^+; \\
    (2) \quad (A^-)(B^+) \geq 1, \quad \text{and} \quad (A^+)(B^-) \geq 1; \quad \text{and} \\
    B^- &\leq \gamma_1, \quad B^+ \leq \gamma_3, \quad A^- \geq 2/\gamma_2, \quad \text{and} \quad A^+ \geq 2/\gamma_2.
\end{align*}
\]

Then, computing the relevant probability ratios, one finds that the terminal decision rule for test \( T^- \) is given by:

stop sampling and make decision \( d^- (\theta^- = 0) \) if

\[
\sum_{i=1}^{n} x_i \geq (b^-) \sigma^2/\theta_1 + \sigma^2/\theta_2 + n(\theta_1 + \theta_2)/2 ,
\]
stop sampling and make decision \(d^- (\theta^- = 1)\) if
\[
\sum_{i=1}^{n} x_i \leq (a^-) \sigma^2 / (\theta_1 - \theta_2) + n(\theta_1 + \theta_2)/2 ,
\]
and continue sampling if
\[
(a^-) \sigma^2 / (\theta_1 - \theta_2) + n(\theta_1 + \theta_2)/2 < \sum_{i=1}^{n} x_i < (b^-) \sigma^2 / (\theta_1 - \theta_2) + n(\theta_1 + \theta_2)/2.
\]

Similarly, the terminal decision rule for test \(T^+\) is given by:

stop sampling and make decision \(d^+ (\theta^+ = 0)\) if
\[
\sum_{i=1}^{n} x_i \leq (b^+) \sigma^2 / (\theta_4 - \theta_3) + n(\theta_3 + \theta_4)/2 ,
\]
stop sampling and make decision \(d^+ (\theta^+ = 1)\) if
\[
\sum_{i=1}^{n} x_i \geq (a^+) \sigma^2 / (\theta_4 - \theta_3) + n(\theta_3 + \theta_4)/2 ,
\]
and continue sampling if
\[
(b^+) \sigma^2 / (\theta_4 - \theta_3) + n(\theta_3 + \theta_4)/2 < \sum_{i=1}^{n} x_i < (a^+) \sigma^2 / (\theta_4 - \theta_3) + n(\theta_3 + \theta_4)/2
\]

Now, a necessary and sufficient condition for the impossibility of making both decision \(d^-\) and decision \(d^+\) with test \(T\) is that at \(n=1\), the following inequalities should hold (see Figure 1):

\[
\text{Acceptance number (of } \theta^- = 1\text{) for } T^- \leq \text{Acceptance number (of } \theta^+ = 0\text{) for } T^+ , \text{ and}
\]

\[
(3) \quad \text{Rejection number (of } \theta^- = 1\text{) for } T^- \leq \text{Rejection number (of } \theta^+ = 0\text{) for } T^+ , \text{ i.e.,}
\]
These inequalities can be written as
\[ (a^{-}) \sigma^2/(\theta_1 - \theta_2) + 1(\theta_1 + \theta_2)/2 \leq (a^{+}) \sigma^2/(\theta_4 - \theta_3) + 1(\theta_3 + \theta_4)/2, \]
and
\[ (b^{-}) \sigma^2/(\theta_1 - \theta_2) + 1(\theta_1 + \theta_2)/2 \leq (a^{+}) \sigma^2/(\theta_4 - \theta_3) + 1(\theta_3 + \theta_4)/2. \]

These inequalities can be written as
\[ (1/A^{-} \cdot B^{+}) \frac{\sigma^2}{\Delta} \leq e^{1/2[(\theta_4 + \theta_3) - (\theta_1 + \theta_2)]}, \]
and
\[ (1/B^{-} \cdot A^{+}) \frac{\sigma^2}{\Delta} \leq e^{1/2[(\theta_4 + \theta_3) - (\theta_1 + \theta_2)]}. \]

Now, since
\[ \Delta/2 \sigma^2 [(\theta_4 + \theta_3) - (\theta_1 + \theta_2)] > 0, \]
the inequalities in (3) will certainly be satisfied if
\[ (A^{-} \cdot B^{+}) \geq 1, \text{ and } (B^{-} \cdot A^{+}) \geq 1. \]
\( \sum_{i=1}^{n} x_i \)

- **Accept \( H^+ \)**
- **Accept \( H_0 \)**
- **Accept \( H^- \)**

\( (d^+): \quad \Sigma x_1 = (a^+) \sigma^2/\Delta + n(\theta_3 + \theta_4)/2 \),

\( (d^+_0): \quad \Sigma x_1 = (b^+) \sigma^2/\Delta + n(\theta_3 + \theta_4)/2 \),

\( (d^-): \quad \Sigma x_1 = -(b^-)\sigma^2/\Delta + n(\theta_1 + \theta_2)/2 \), and

\( (d^-_1): \quad \Sigma x_1 = -(a^-)\sigma^2/\Delta + n(\theta_1 + \theta_2)/2 \).

**FIGURE 1**.
[Obviously, for any given problem the actual shape of the graphic representation of the Sobel-Wald test procedure will depend on the values of \( b^+, a^+, \sigma^2 \), and \( \Delta \), which jointly determine the distance between the parallel decision lines, and on the quantities \((\theta_1 + \theta_2)\) and \((\theta_3 + \theta_4)\), which determine the slopes of the (pairs of) decision lines.]

Now, recalling Wald's conservative bounds on the error probabilities of a S.P.R.T. [10] and referring to (2), one can readily see that

\[
\begin{align*}
  (i) \quad \beta^- & \leq B^- \leq \gamma_1, \\
  (ii) \quad \beta^+ & \leq B^+ \leq \gamma_3, \text{ and} \\
  (iii) \quad \alpha^+ & \leq \gamma_2/2.
\end{align*}
\]

Thus, the bounds on the overall error probabilities for test \( T \) are satisfied, as was shown in section 3 of Chapter I.

Finally, recalling Wald's approximation formula for the operating characteristic function of a S.P.R.T. (see p. 50 [10]), namely

\[
(4) \quad L(\theta) = \Pr[\text{accepting the null hypothesis }|\theta] = \frac{1 - e^{\alpha h(\theta)}}{e^{\beta h(\theta)} - e^{\alpha h(\theta)}},
\]

where

\[
\text{h}(\theta) \text{ is a non-zero solution of }
\int_{-\infty}^{\infty} \left[ \frac{f_{\theta_1}(x)}{f_{\theta_0}(x)} \right]^{\alpha}(x) e^{-\beta h(\theta)} (x) dF_\theta(x) = 1,
\]

and "\( e \) implies that "excess" over the boundaries at the termination of the procedure is neglected,
one can obtain the following approximations to the various decision probabilities for test T:

\[
\Pr[D^- \text{ using } T|\theta] \approx \frac{1 - \exp(-h^-(-\epsilon))}{\exp(b^-h^-(\theta)) - \exp(a^-h^-(\theta))},
\]

\[
\Pr[D_0 \text{ using } T|\theta] \approx \frac{1 - \exp(-h^-(\theta))}{\exp(b^-h^-(\theta)) - \exp(a^-h^-(\theta))} - \frac{\exp(b^+h^+(\theta)) - 1}{\exp(b^+h^+(\theta)) - \exp(a^+h^+(\theta))},
\]

and

\[
\Pr[D^+ \text{ using } T|\theta] \approx \frac{\exp(b^+h^+(\theta)) - 1}{\exp(b^+h^+(\theta)) - \exp(a^+h^+(\theta))},
\]

where

\[
h^-(\theta) = \frac{2\theta - (\theta_1 + \theta_2)}{\Delta}, \text{ and } h^+(\theta) = \frac{(\theta_3 + \theta_4) - 2\theta}{\Delta}.
\]

Now, at \(\epsilon = (\theta_1 + \theta_2)/2\) and at \(\epsilon = (\theta_3 + \theta_4)/2\) [i.e., \(h^-(\theta) = 0\), or \(h^+(\theta) = 0\)], (4) can be replaced by

\[
L(\theta) \approx \frac{a}{(a-b)},
\]

where (5) is obtained by taking the limits (as \(h(\theta) \to 0\)) in (4) and using l'Hopital's rule.

2. The Fixed Sample Size Test

For the I.F.S.S.T. procedure, which can be used to solve the general problem under consideration in this paper when \(\sigma^2\) is known, \(T^\sim\) denotes a F.S.S.T. of \(\theta^\sim = 0\) vs. \(\theta^\sim = 1\), based on a sample of \((x_1^-, x_2^-, \ldots, x_N^-)\) with pre-assigned strength \((\alpha^-, \beta^-)\). Similarly,
$T^+$ denotes a F.S.S.T. of $\theta^+ = 0$ vs. $\theta^+ = 1$, based on a sample of $(X_1^+, X_2^+, \ldots, X_N^+)$, with pre-assigned strength $(\alpha^+, \beta^+)$, where

\[
\beta^- \leq \gamma_1, \beta^+ \leq \gamma_3, \text{ and } \alpha^+ \leq \gamma_2/2
\]

\[
N = \max, \left\{ \left[ \frac{\sigma^2(z^- + z^-)^2}{\alpha^-} + 1 \right], \left[ \frac{\sigma^2(z^+ + z^+)^2}{\beta^+} + 1 \right] \right\}
\]

(6) $[\ ]$ is used symbolically to denote "the largest integer in";

and $z_\mu$ is that number which is exceeded by a standard normal random variable with probability $\mu$.

[Before we consider the respective terminal decision rules for sub-tests $T^-$ and $T^+$, it should be noted that by defining the overall sample size, $N$, as in (6), we are actually ensuring that

\[
\Pr[\text{rejecting } H^-|\theta] \leq \min. (\gamma_1, \gamma_3), \text{ for } \theta \leq \theta_1
\]

and

\[
\Pr[\text{rejecting } H^+|\theta] \leq \min. (\gamma_1, \gamma_3), \text{ for } \theta \geq \theta_4
\]

Hence if $\gamma_1 \neq \gamma_3$, test $T$ will be more conservative than is necessary.]

Now, the terminal decision rules for sub-tests $T^-$ and $T^+$ can be stated as

make decision $d^-_0(\theta^- = 0)$ if $\overline{x}_N \geq -\Delta c^- + \theta_2$,

make decision $d^-_1(\theta^- = 1)$ if $\overline{x}_N < -\Delta c^- + \theta_2$,

and

make decision $d^+_0(\theta^+ = 0)$ if $\overline{x}_N \leq \Delta c^+ + \theta_3$,

make decision $d^+_1(\theta^+ = 1)$ if $\overline{x}_N > \Delta c^+ + \theta_3$,
respectively, where

\[ c^- = (z_{\alpha^-})\sigma/\sqrt{N} \text{, and } c^+ = (z_{\alpha^+})\sigma/\sqrt{N} \].

Since \( \theta_2 \leq \theta_3 \), and since \( c^- \) and \( c^+ \) are always non-negative (i.e., \( \gamma_2/2 \) is always \( \leq 1/2 \)), it is obvious that

Acceptance no. (of \( e^- = 1 \)) for \( T^- \leq \) Acceptance no. (of \( e^+ = 0 \)) for \( T^+ \), and

Rejection no. (of \( e^- = 1 \)) for \( T^- \leq \) Rejection no. (of \( e^+ = 0 \)) for \( T^+ \), i.e.,

\[-\Delta c^- + \theta_2 \leq \Delta c^+ + \theta_3.\]

Hence, it is impossible to make both decisions \( d^- \) and \( d^+ \) using test \( T \).

Furthermore, since

(i) \( \beta^- \leq \gamma_1 \),
(ii) \( \beta^+ \leq \gamma_3 \), and
(iii) \( \alpha^+ \leq \gamma_2/2 \).

as was stipulated in (6), the bounds on the error probabilities [given in (1)] for test \( T \) are satisfied.

A less conservative, three-decision F.S.S.T. procedure can be obtained by choosing \( c^- \), \( c^+ \), and \( N \) so that

(i) \( \inf_{\theta \leq \theta_1} \text{Pr}\{\text{making decision } D^-|\theta\} \geq 1 - \gamma_1 \),
(ii) \( \inf_{\theta \geq \theta_4} \text{Pr}\{\text{making decision } D^+|\theta\} \geq 1 - \gamma_3 \), and
(iii) \( \inf_{\theta_2 \leq \theta \leq \theta_3} \text{Pr}\{\text{making decision } D_0|\theta\} \geq 1 - \gamma_2 \).
where the terminal decision rule is of the form

\[
\begin{align*}
D^-; & \text{ if } x_N \leq -\Delta c^- + \theta_2, \\
D^+; & \text{ if } x_N \geq \Delta c^+ + \theta_3, \\
D_0; & \text{ otherwise.}
\end{align*}
\]

Thus, for \( N \) fixed and equality in (i) and (ii), we find, using the monotonicity in \( \theta \) of the probabilities, that

\[
c^- = 1 - z_{\gamma_1} \cdot \frac{\sigma}{\Delta} / \sqrt{N},
\]

and

\[
c^+ = 1 - z_{\gamma_3} \cdot \frac{\sigma}{\Delta} / \sqrt{N}.
\]

Observing that

\[
\Pr[\text{making decision } D_o | \theta] = \Phi\left(\frac{\overline{M}}{\sigma} \left(\Delta c^+ + \theta_3 - \theta\right)\right)
\]

\[- - \Phi\left(\frac{\overline{M}}{\sigma} \left(-\Delta c^- + \theta_2 - \theta\right)\right),
\]

we find, upon taking first and second derivatives w.r.t. \( \theta \), that \( \Pr[\text{making decision } D_o | \theta] \) attains its maximum over the interval \([\theta_2, \theta_3]\) at

\[
\theta_o = \frac{\Delta(c^+ - c^-) + (\theta_2 + \theta_3)}{2},
\]

and its minimum at either \( \theta_2 \) or \( \theta_3 \). If the minimum occurs at \( \theta = \theta_2 \), then equality in (iii) implies that

\[
\gamma_2 = \Phi\left(z_{\gamma_1} - \frac{\sqrt{N}(\theta_2 - \theta_1)}{\sigma}\right) + \Phi\left(z_{\gamma_3} - \frac{\sqrt{N}(\theta_3 - \theta_1)}{\sigma}\right).
\]
A trial-and-error solution for N in the above expression, say N', can be obtained by referring to a table of the standard normal distribution. Similarly, if Pr[making decision D_0 | θ] attains its minimum over [θ_2, θ_3] at θ = θ_3, then equality in (iii) implies that

\[ γ_2 = \Phi \left[ z_3 - \frac{\sqrt{N} (θ_2 - θ)}{σ} \right] + \Phi \left[ z_1 - \frac{\sqrt{N} (θ_3 - θ)}{σ} \right] \]

Denoting the trial-and-error solution for N in the above by N^+, we can define the overall sample size as the first integer that exceeds max(N', N^+). This N is certainly less than the conservative one obtained for the I.F.S.S.T. which was, in effect, chosen by replacing the minimization in (iii), where the probability to be minimized may be written as the sum of two different terms, with the sum of the minimizations of these separate terms. (A more detailed discussion of the above test procedure appears in the Appendix of [3].)

Finally, by performing some simple algebraic manipulations one sees that the decision probabilities for test T can be written as

\[
\begin{align*}
\Pr[D^- \text{ using } T | θ] &= \Pr[Z < -z_{α^-} + \frac{\sqrt{N} (θ_2 - θ)}{σ}] \ , \\
\Pr[D_0 \text{ using } T | θ] &= \Pr[Z \geq -z_{α^-} + \frac{\sqrt{N} (θ_2 - θ)}{σ}] \ , \\
&- \Pr[Z > z_{α^+} + \frac{\sqrt{N} (θ_3 - θ)}{σ}] ,
\end{align*}
\]
and

\[ \Pr[D^+ \text{ using } T_0] = \Pr[Z > \frac{z_{\alpha^+} + \sqrt{N}(\alpha_3 - \alpha)}{\sigma}] , \]

where

\( N \) is defined as in (6),

and

\( Z \) is a standard normal random variable.

3. The Stein Two-Stage Procedure

Stein's test [9] is a two-stage procedure that can be applied to the general problem introduced in Section 2 of Chapter I when the variance, \( \sigma^2 \), is unknown. Under this technique the sub-tests \( T^- \) and \( T^+ \) are constructed as follows: Let \( (T^*)^T \) represent the Stein two-stage test, here denoted by \( S.T.S.T.(m) \), of \( \theta^+ = 0 \) vs. \( \theta^+ = 1 \), based on samples \( \{X_1, X_2, \ldots, X_m\} \) and \( \{X_{m+1}, X_{m+2}, \ldots, X_{N^-}\} \), with initial sample size of \( m \) (a fixed integer \( \geq 2 \)) and pre-assigned strength \( (\alpha^+, \beta^+) \), where

\[ N^- = \max\{m, [s^2_m(t^- + t_-)^2/\Delta^2] + 1\} ; \]

\[ \beta^- \leq \min (1/2, \gamma_1), \beta^+ \leq \min (1/2, \gamma_3), \text{ and } \alpha^+ \leq \min (1/2, \gamma_2/2) ; \]

\[ (\bar{x}_m = \frac{\Sigma_{i=1}^m x_i}{m}, \text{ and } s^2_m = \frac{\Sigma_{i=1}^m (x_i - \bar{x}_m)^2}{(m-1)} ; \text{ and} \]

\[ t_\mu \text{ is that number which is exceeded by a Student's } t \text{ random variable (with } m-1 \text{ d.f.) with probability } \mu. \]
Then sub-tests $T^-$ and $T^+$ are equivalent to tests $(T^*)^-$ and $(T^*)^+$, respectively, except that the sample sizes $N^-$ and $N^+$ are replaced by $N = \max(N^-, N^+)$. 

[As was noted in the discussion of the I.F.S.S.T. in the preceding section, defining $N$ as above leads to a more conservative test $T$ than is actually required if $\gamma_1 \neq \gamma_3$.] 

Now, the terminal decision rules for sub-tests $T^-$ and $T^+$ are given, respectively, by

- **Make decision $d_0^-$** ($\theta^- = 0$) if $x_N > -\Delta k^- + \theta_2$,
- **Make decision $d_1^-$** ($\theta^- = 1$) if $x_N < -\Delta k^- + \theta_2$,

and

- **Make decision $d_0^+$** ($\theta^+ = 0$) if $x_N \leq \Delta k^+ + \theta_3$,
- **Make decision $d_1^+$** ($\theta^+ = 1$) if $x_N > \Delta k^+ + \theta_3$,

where

- $k^- = \frac{t^-}{\alpha^-/(\alpha^- + \beta^-)}$, and $k^+ = \frac{t^+}{\alpha^+/(\alpha^+ + \beta^+)}$.

Since it is given that $\theta_2 \leq \theta_3$ and since (7) implies that $c^-$ and $c^+$ are always non-negative, it obviously follows that

- Acceptance no. (of $\theta^- = 1$) for $T^- \leq$ Acceptance no. (of $\theta^+ = 0$) for $T^+$, and
- Rejection no. (of $\theta^- = 1$) for $T^- \leq$ Rejection no. (of $\theta^+ = 0$) for $T^+$, i.e.,

$$-\Delta k^- + \theta_2 \leq \Delta k^+ + \theta_3.$$
Therefore, a situation in which both decision $d^-_1$ and decision $d^+_1$ are made will never arise when test $T$ is employed.

Finally, since

(i) $\beta^- \leq \gamma_1$ ,

(ii) $\beta^+ \leq \gamma_3$, and

(iii) $\alpha^+ \leq \gamma_2/2$ ,

as stated in (7), it follows from Section 3 of Chapter I that the bounds on the pre-assigned error probabilities for test $T$ are satisfied.

Now, in order to derive expressions for the various decision probabilities under test $T$ we shall make use of the fact that

$$\frac{\sqrt{n}}{s_m} (\bar{X}_N - \theta)$$

has a Student's$\cdot$-$t$ distribution with $(m-1)$ d.f. (see [9]). With this result in mind one can readily see that

$$\text{Pr}[D^- \text{ using } T|\theta] = \text{Pr}[T(m-1) < (\bar{X}_N/s_m) [ - \Delta k^- + (\theta_2 - \theta)]] ,$$

$$\text{Pr}[D_0 \text{ using } T|\theta] = \text{Pr}[T(m-1) \geq (\bar{X}_N/s_m) [- \Delta k^- + (\theta_2 - \theta)]]$$

$$= \text{Pr}[T(m-1) > (\bar{X}_N/s_m) [ \Delta k^+ + (\theta_3 - \theta)]] ,$$

and

$$\text{Pr}[D^+ \text{ using } T|\theta] = \text{Pr}[T(m-1) > (\bar{X}_N/s_m) [ \Delta k^+ + (\theta_3 - \theta)]] ,$$

where $T(m-1)$ is, marginally, a Student's$\cdot$-$t$ random variable with $(m-1)$ degrees of freedom.
Now, replacing the random variable \( \bar{N}/m \) by its lower bound, 
\[
\max_{\alpha^- \beta^-} \{(t_+ + t_-)/\Delta, (t_+ + t_-)/\Delta\},
\]
we can obtain either an upper bound or a lower bound on each of the various decision probabilities, depending on the value of \( \theta \). For example, if
\[
[-\Delta k^- + (\theta_2 - \theta^*)] < 0,
\]
then
\[
\Pr[D^- \text{ using } T|\theta^*] \leq \Pr[T(m-1) < -t_+ + (\theta_2 - \theta^*)(t_+ + t_-)/\Delta];
\]
and if
\[
[-\Delta k^- + (\theta_2 - \theta^*)] > 0,
\]
then
\[
\Pr[D^- \text{ using } T|\theta^*] \geq \Pr[T(m-1) < -t_+ + (\theta_2 - \theta^*)(t_+ + t_-)/\Delta].
\]
These bounds are approximate equalities if the probability of continuing sampling beyond the initial sample is close to unity.

4. The Baker-Hall Sequential Analog of Stein's Two-Stage Procedure.

The Baker-Hall procedure \([2, 5]\) can be thought of as a sequential analog of Stein's two-stage test, because it is, essentially, a S.P.R.T. with the unknown variance replaced by an estimate based on an initial sample. It tends to be somewhat less restrictive than the Stein procedure since it can be applied directly to the general problem set forth in the Introduction rather than to the slightly diluted version that was, in effect, considered in the preceding section, i.e., where \( T \) was designed so that
\[
\Pr[\text{rejecting } H^-|\theta] \leq \min_\gamma (\gamma_1, \gamma_2), \text{ for } \theta \leq \theta_1,
\]
and

$$\Pr\{\text{rejecting } H^+ | \theta \} \leq \min \{ \gamma_1, \gamma_2 \}, \text{ for } \theta \geq \theta_{h+}.$$ 

Furthermore, the Baker-Hall technique may result in a substantial savings of the total number of observations required to reach a final decision, just as, in the case of known $$\sigma^2$$, the I.S.P.R.T. may result in savings over the I.F.S.S.T.

Now, in order to apply the Baker-Hall procedure the sub-tests $$T^-$$ and $$T^+$$ should be constructed as follows: Let $$T^-$$ denote a S.P.R.T. of $$\theta^*=0$$ vs. $$\theta^*=1$$, based on $$\bar{X}_m, X_{m+1}^-, X_{m+2}^-, \ldots$$, with pre-assigned strength $$(\alpha^-,\beta^-)$$, termination boundaries $$B^{-}_m$$ and $$A^{-}_m$$, and with $$(\sigma^-)^2$$ [known] replaced by $$(s^-_m)^2$$.

Similarly, let $$T^+$$ denote a S.P.R.T. of $$\theta^+=0$$ vs. $$\theta^+=1$$, based on $$\bar{X}_m^+, X_{m+1}^+, X_{m+2}^+, \ldots$$, with pre-assigned strength $$(\alpha^+,\beta^+)$$, termination boundaries $$B^+_m$$ and $$A^+_m$$, and with $$(\sigma^+)^2$$ [known] replaced by $$(s^+_m)^2$$, where

$$(A^-_m)(B^+_m) \geq 1, \text{ and } (B^-_m)(A^+_m) \geq 1;$$

$$b^-_m = \log B^-_m \leq -\nu (\gamma_{1}^{2/\nu} - 1)/2, \quad \nu = m - 1,$$

(8)

$$b^+_m = \log B^+_m \leq -\nu (\gamma_{3}^{2/\nu} - 1)/2,$$

$$a^+_m = \log A^+_m \geq \nu [(\gamma_2/2)^{2/\nu} - 1]/2; \text{ and}$$

$$(s^-_m)^2 = s^2_m/(\theta_1 - \theta_2)^2 = s^2_m/A^2 = s^2_m/(\theta_{4} - \theta_{3})^2 = (s^+_m)^2.$$

Then, computing the relevant probability ratios, one finds that the terminal decision rules for sub-tests $$T^-$$ and $$T^+$$ are given by
stop sampling and make decision \( d_{o}^{-}(\theta^{-}=0) \) if \( r_{n}(s_{m}^{-}) \leq b_{m}^{-} \),
stop sampling and make decision \( d_{1}^{-}(\theta^{-}=1) \) if \( r_{n}(s_{m}^{-}) \geq a_{m}^{-} \),
continue sampling if \( b_{m}^{-} < r_{n}(s_{m}^{-}) < a_{m}^{-} \),

and

stop sampling and make decision \( d_{o}^{+}(\theta^{+}=0) \) if \( r_{n}(s_{m}^{+}) \leq b_{m}^{+} \),
stop sampling and make decision \( d_{1}^{+}(\theta^{+}=1) \) if \( r_{n}(s_{m}^{+}) > a_{m}^{+} \),
continue sampling if \( b_{m}^{+} < r_{n}(s_{m}^{+}) < a_{m}^{+} \),

respectively, where

\[
    r_{n}(s_{m}^{\pm}) = \sum_{i=1}^{n} \frac{(x_{i}^{\pm} - \bar{x})^{2}}{s_{m}^{\pm}} , \quad n \geq m .
\]

Equivalent forms of the above terminal decision rules, which greatly facilitate numerical computations, can be stated as follows:

stop sampling and make decision \( d_{o}^{-} \) if
\[
    \sum_{i=1}^{n} x_{i}^{-} \geq \frac{(b_{m}^{-}) s_{m}^{-}/(-\Delta) + n(\theta_{1} + \theta_{2})/2}{1} ,
\]
stop sampling and make decision \( d_{1}^{-} \) if
\[
    \sum_{i=1}^{n} x_{i}^{-} \leq \frac{(a_{m}^{-}) s_{m}^{-}/(-\Delta) + n(\theta_{1} + \theta_{2})/2}{1} ,
\]
continue sampling if
\[
    \frac{(a_{m}^{-}) s_{m}^{-}/(-\Delta) + n(\theta_{1} + \theta_{2})/2}{1} < \sum_{i=1}^{n} x_{i}^{-} < \frac{(b_{m}^{-}) s_{m}^{-}/(-\Delta) + n(\theta_{1} + \theta_{2})/2}{1} ,
\]

and

stop sampling and make decision \( d_{o}^{+} \) if
\[
    \sum_{i=1}^{n} x_{i}^{+} \leq \frac{(b_{m}^{+}) s_{m}^{+}/\Delta + n(\theta_{3} + \theta_{4})/2}{1} ,
\]
stop sampling and make decision \( d_{1}^{+} \) if
\[
    \sum_{i=1}^{n} x_{i}^{+} \geq \frac{(a_{m}^{+}) s_{m}^{+}/\Delta + n(\theta_{3} + \theta_{4})/2}{1} ,
\]
continue sampling if

\[(b_m^+)s_m^2/\Delta + n(\theta_3 + \theta_4)/2 < \sum_{i=1}^{n} x_i < (a_m^+)s_m^2/\Delta + n(\theta_3 + \theta_4)/2 \cdot \]

Now, a necessary and sufficient condition to ensure that both decisions \(d_1^-\) and \(d_1^+\) will not be made under test \(T\) is that at \(n = m\),

Acceptance no. (of \(\theta^- = 1\)) for \(T^- \leq \) Acceptance no. (of \(\theta^+=0\)) for \(T^+\), and

Rejection no. (of \(\theta^- = 1\)) for \(T^- \leq \) Rejection no. (of \(\theta^+=0\)) for \(T^+\), i.e.,

\[(a_m^-)s_m^2/(-\Delta) + m(\theta_1 + \theta_2)/2 \leq (b_m^+)s_m^2/\Delta + m(\theta_3 + \theta_4)/2 , \]

and

\[(b_m^-)s_m^2/(-\Delta) + m(\theta_1 + \theta_2)/2 \leq (a_m^+)s_m^2/\Delta + m(\theta_3 + \theta_4)/2 . \]

Now, these inequalities can be expressed in the equivalent form

\[\left(1/A_m^- \cdot B_m^+\right)s_m^2/\Delta \leq e^{-m/2[(\theta_3 + \theta_4) - (\theta_1 + \theta_2)]} , \]

and

\[\left(1/A_m^- \cdot B_m^+\right)s_m^2/\Delta \leq e^{-m/2[(\theta_3 + \theta_4) - (\theta_1 + \theta_2)]} . \]

Thus, since

\[m \Delta/2 \cdot s_m^2 [(\theta_3 + \theta_4) - (\theta_1 + \theta_2)] > 0 , \]

they will certainly be satisfied if

\[(A_m^-) (B_m^+) \geq 1, \text{ and } (B_m^-) (A_m^+) \geq 1 . \]
In order to determine whether or not test $T$ meets the bounds on the pre-assigned error probabilities given in (1), it is necessary to consider in some detail a conditional S.P.R.T., $T(s_m, \sigma)$, applied to the general problem introduced in Section 2 of Chapter I. So, for given $(s_m, \sigma)$ let $T^-(s_m, \sigma)$ denote the conditional S.P.R.T. of $\theta^- = 0$ vs. $\theta^- = 1$, based on $X_m^- , X_{m+1}^- , X_{m+2}^- , \ldots$, with termination boundaries $\overline{b}_m^-$, $\overline{a}_m^-$. Similarly, let $T^+(s_m, \sigma)$ denote the conditional S.P.R.T. of $\theta^+ = 0$ vs. $\theta^+ = 1$, based on $X_m^+ , X_{m+1}^+ , X_{m+2}^+ , \ldots$, with termination boundaries $\overline{b}_m^+$ and $\overline{a}_m^+$, where

$$\overline{b}_m^- = \log \overline{b}_m^- = \log \left( \frac{b_m^-}{(s_m^-)^2 / \sigma^-^2} \right) = \log \left( \frac{b_m^-}{s_m^- / \sigma^-} \right) = \log \frac{b_m^-}{s_m^- / \sigma^-} ,$$

$$\overline{b}_m^+ = \log \overline{b}_m^+ = \log \left( \frac{b_m^+}{s_m^+ / \sigma^+} \right) ,$$

$$\overline{a}_m^- = \log \overline{a}_m^- = \log \left( \frac{a_m^-}{s_m^- / \sigma^-} \right) ,$$

$$\overline{a}_m^+ = \log \overline{a}_m^+ = \log \left( \frac{a_m^+}{s_m^+ / \sigma^+} \right) .$$

Then, the terminal decision rules for sub-tests $T^-$ and $T^+$ can be expressed as

stop sampling and make decision $d^-(\theta^- = 0)$ if $r_n(\sigma^-) \leq \overline{b}_m^-$,

stop sampling and make decision $d^-(\theta^- = 1)$ if $r_n(\sigma^-) \geq \overline{a}_m^-$, continue sampling if $\overline{b}_m^- < r_n(\sigma^-) < \overline{a}_m^-$,

and

stop sampling and make decision $d^+(\theta^+ = 0)$ if $r_n(\sigma^+) \leq \overline{b}_m^+$,

stop sampling and make decision $d^+(\theta^+ = 1)$ if $r_n(\sigma^+) \geq \overline{a}_m^+$, continue sampling if $\overline{b}_m^+ < r_n(\sigma^+) < \overline{a}_m^+$.
It can be shown by straightforward algebraic manipulation that the necessary and sufficient condition for the impossibility of making both decision \( d^- \) and decision \( d^+ \) using test \( T(s_m, \sigma) \) is the same as that for test \( T \).

Furthermore, since
\[
\frac{r_n(\sigma^+)}{r_n(s_m^+)} = \frac{s_m^2}{r_n(s_m^+)/\sigma^2},
\]
\[
\frac{b^+}{s_m^2} = \frac{s_m^2(b^+)+/s^2}, \quad \text{and} \quad \frac{a^+}{s_m^2} = \frac{s_m^2(a^+)+/s^2},
\]
it is obvious that \( T(s_m, \sigma) \) has precisely the same decision at each stage as does test \( T \). Since the terminal decision rules for \( T(s_m, \sigma) \) and \( T \) are equivalent, it follows that

\[
E_{S_m} \{ \Pr(D* \text{ using } T(s_m, \sigma) | S_m, \sigma, \theta) \} = E_{S_m} \{ \Pr(D* \text{ using } T | S_m, \sigma, \theta) \},
\]

where \( D* \) represents some specified decision.

Now, consider the random variable, \( \omega \), where
\[
\omega = \begin{cases} 
1; & \text{if decision } D* \text{ is made using test } T, \\
0; & \text{otherwise}
\end{cases}
\]

Clearly,
\[
E[\omega | S_m, \sigma, \theta] = \Pr(D* \text{ using } T | S_m, \sigma, \theta),
\]
and
\[
E_{S_m} \{ E[\omega | S_m, \sigma, \theta] \} = E[\omega | \sigma, \theta] = \Pr(D* \text{ using } T | \sigma, \theta).
\]
Thus,

\[(10) \quad E_S \{ \Pr(D^* \text{ using } T(s_m, \sigma) | S_m, \sigma, \theta) \} = \Pr(D^* \text{ using } T|\sigma, \theta) . \]

Now, making use of Wald's conservative bounds on the error probabilities of a S.P.R.T., we have

\[\Pr(d_o^+ \text{ using } T^+(s_m, \sigma) | s_m, \sigma, \theta^+ = 1) \leq \frac{F_m^+}{A_m^+} = \exp \left( b_m^+ s_m^2/\sigma^2 \right),\]

and

\[\Pr(d_o^- \text{ using } T^-(s_m, \sigma) | s_m, \sigma, \theta^- = 0) \leq \frac{1}{A_m^+} = \exp \left( -s_m^+ s_m^2/\sigma^2 \right).\]

Hence, it follows from (10) that

\[(11) \quad \Pr(d_o^- \text{ using } T^- | \sigma, \theta^- = 1) \leq E\{ \exp \left( b_m^- s_m^2/\sigma^2 \right) \} = (1 - 2b_m^-/v)^{-v/2}, \text{ for all } \sigma^2 , \]

since

\[v s_m^2/\sigma^2 = \chi_v^2, \text{ and } E\{ \exp \left( t \chi_v^2 \right) \} = (1 - 2t)^{-v/2} . \]

Similarly,

\[(12) \quad \Pr(d_o^+ \text{ using } T^+ | \sigma, \theta^+ = 1) \leq E\{ \exp \left( b_m^+ s_m^2/\sigma^2 \right) \} = (1 - 2b_m^+/v)^{-v/2} , \]
\[(13) \quad \Pr\{d_{m}^{-} \text{ using } T^{-}[\sigma, \theta^{-} = 0] \leq E \{\exp \{-a_{m}^{-} s_{m}^{2}/\sigma^{2}\}\} \]
\[= (1 + 2a_{m}^{-}[\nu]^{-\nu/2}) ,\]

and

\[(14) \quad \Pr\{d_{m}^{+} \text{ using } T^{+}[\sigma, \theta^{+} = 0] \leq E \{\exp \{-a_{m}^{+} s_{m}^{2}/\sigma^{2}\}\} \]
\[= (1 + 2a_{m}^{+}[\nu]^{-\nu/2}) ,\]

for all \(\sigma^{2}\). Thus, by substituting for \(\nu_{m}^{+}\) and \(a_{m}^{+}\) from (8) in (11) - (14), one finds that

\[(i) \quad \beta^{-} \leq \gamma_{1} ,\]
\[(ii) \quad \beta^{+} \leq \gamma_{3} , \text{ and}\]
\[(iii) \quad \alpha^{+} \leq \gamma_{2}/2 ,\]

which guarantees that the bounds on the error probabilities given in (1) are satisfied by test \(T\).

Now, in order to develop approximation formulae \(h^{x}\) for the various decision probabilities for the I.S.P.R.T. \(T_{(m)}\) it is necessary, once again, to refer to Wald's approximation formula for the operating characteristic function of a S.P.R.T. Thus, for the conditional S.P.R.T., \(T(s_{m}, \sigma)\), one has

\[(15) \quad \Pr[D^{-} \text{ using } T(s_{m}, \sigma)|s_{m}, \sigma, \theta] \]
\[= \Pr[d_{m}^{-} \text{ using } T^{-}[s_{m}, \sigma)|s_{m}, \sigma, \theta] \leq 1 - \frac{(\bar{A}_{m}^{-})h^{-}(\theta) - 1}{(\bar{A}_{m}^{-})h^{-}(\theta) - (\bar{B}_{m}^{-})h^{-}(\theta)} ,\]

The approximation formulae for the various decision probabilities which are derived above are taken from \([5]\), and the reader who is interested in a more complete discussion of the mathematics involved in their development should consult this paper.
where \( h^-(\varnothing) = \frac{[2\varnothing - (\varnothing_1 + \varnothing_2)]}{\Delta} \). Taking expectations with respect to \( S_m^2 \) in (15), substituting from (9), and ignoring the subscripts on \( a_m, b_m \) and \( S_m^2 \), we obtain

\[
(16) \quad \Pr[D^- \text{ using } T|\sigma, \varnothing] = \mathbb{E} \left\{ 1 - \frac{\exp\left[ -a^- h^- S^2 / \sigma^2 \right]}{1 - \exp\left[ -(a^- - b^-) h^- S^2 / \sigma^2 \right]} \right\} \\
= 1 - \mathbb{E}\left[ 1 - \exp\left( -a^- h^- S^2 / \sigma^2 \right) \right] \sum_{i=1}^{\infty} \exp\left( -(\varnothing - a^-) h^- S^2 / \sigma^2 \right)
\]

for \( h^-(\varnothing) > 0 \), i.e., \( \varnothing > (\varnothing_1 + \varnothing_2)/2 \), which does not depend on \( \sigma \).

Expanding the RHS of (16) and then taking expectations term-by-term, one has

\[
(17) \quad \Pr[D^- \text{ using } T|\sigma, \varnothing] \leq (1 + 2a^- h^- / \nu)^{-\nu/2} \\
-\left[ 1 + 2(a^- - b^-) h^- / \nu \right]^{-\nu/2} + [1 + 2(2a^- - b^-) h^- / \nu]^{-\nu/2} \\
-\left[ 1 + 4(a^- - b^-) h^- / \nu \right]^{-\nu/2} + \ldots
\]

Similarly, for \( h^-(\varnothing) < 0 \), i.e., \( \varnothing < (\varnothing_1 + \varnothing_2)/2 \), we have

\[
(18) \quad \Pr[D^- \text{ using } T|\sigma, \varnothing] \leq 1 - (1 + 2b^- h^- / \nu)^{-\nu/2} \\
+ [1 + 2(b^- - a^-) h^- / \nu]^{-\nu/2} - [1 + 2(2b^- - a^-) h^- / \nu]^{-\nu/2} \\
+ [1 + 4(b^- - a^-) h^- / \nu]^{-\nu/2} - \ldots
\]

Analogously, for \( h^+(\varnothing) = [(\varnothing_3 + \varnothing_4) - 2\varnothing] / \Delta > 0 \), we obtain

\[
(19) \quad \Pr[D^+ \text{ using } T|\sigma, \varnothing] \leq (1 + 2a^+ h^+ / \nu)^{-\nu/2} \\
-\left[ 1 + 2(a^+ - b^+) h^+ / \nu \right]^{-\nu/2} + [1 + 2(2a^+ - b^+) h^+ / \nu]^{-\nu/2} \\
-\left[ 1 + 4(a^+ - b^+) h^+ / \nu \right]^{-\nu/2} + \ldots
\]
and for \( h^+(\theta) < 0 \), i.e., \( \theta > (\theta_3 + \theta_4)/2 \),

\[
(20) \quad \Pr[D^+ \text{ using } T|\sigma, \theta] \leq 1 - (1 + 2b^+ h^+/\nu)^{-\nu/2} \\
+ [1 + 2(b^+ - a^+) h^+/\nu]^{-\nu/2} - \ldots
\]

Now, if \( h^+(\theta) = 0 \), then, as was noted in Section 1 of Chapter II, 
\[
\Pr[D^+ \text{ using } T|\sigma, \theta] \leq b^+/(b^+ - a^+)
\]
by l'Hopital's rule.

Finally, recalling that
\[
\Pr[D^+ \text{ using } T|\sigma, \theta] = 1 - \Pr[d^- \text{ using } T^-|\sigma, \theta] \\
- \Pr[d^+ \text{ using } T^+|\sigma, \theta]
\]
we have

\[
(21a) \quad \Pr[D^+ \text{ using } T|\sigma, \theta] \leq (1 + 2b^- h^-/\nu)^{-\nu/2} \\
- (1 + 2a^+ h^+/\nu)^{-\nu/2} - [1 + 2(b^- - a^-) h^-/\nu]^{-\nu/2} \\
+ [1 + 2(a^+ - b^+) h^+/\nu]^{-\nu/2} - \ldots, \text{ for } \theta < (\theta_1 + \theta_2)/2 ;
\]

\[
(21b) \quad \Pr[D^+ \text{ using } T|\sigma, \theta] \leq 1 - (1 + 2a^- h^-/\nu)^{-\nu/2} \\
- (1 + 2a^+ h^+/\nu)^{-\nu/2} + [1 + 2(a^- - b^-) h^-/\nu]^{-\nu/2} \\
+ [1 + 2(a^- - b^+) h^+/\nu]^{-\nu/2} - \ldots, \text{ for } (\theta_1 + \theta_2)/2 < \theta < (\theta_3 + \theta_4)/2 ;
\]

\[
(21c) \quad \Pr[D^+ \text{ using } T|\sigma, \theta] \leq (1 + 2b^+ h^+/\nu)^{-\nu/2} \\
- (1 + 2a^- h^-/\nu)^{-\nu/2} - [1 + 2(b^+ - a^+) h^+/\nu]^{-\nu/2} \\
+ [1 + 2(a^- - b^-) h^-/\nu]^{-\nu/2} - \ldots, \text{ for } \theta > (\theta_3 + \theta_4)/2 .
\]
The Minimum Probability Ratio Test Procedure

The minimum probability ratio test, denoted by M.P.R.T., is a modified version of Wald's S.P.R.T. with converging straight-line boundaries. The three-decision analog of the M.P.R.T., designated by I.M.P.R.T., is only applicable to the symmetric version (i.e., \( \gamma_1 = \gamma_3 = \frac{\gamma_2}{2} = \gamma \), say) of the general problem described in Section 2 of Chapter I when the variance is known. For this special case it is, presumably, intermediate between the I.F.S.S.T. and the I.S.P.R.T. with respect to the area of best performance.

In order to develop the I.M.P.R.T., the sub-tests \( T^- \) and \( T^+ \) should be constructed as follows: Let \( T^- \) denote a M.P.R.T. of \( \theta = 0 \) vs. \( \theta = 1 \), based on \( X_1^-, X_2^-, \ldots \), with pre-assigned strength \((\alpha, \alpha)\); and let \( T^+ \) denote a M.P.R.T. of \( \theta = 0 \) vs. \( \theta = 1 \), based on \( X_1^+, X_2^+, \ldots \), with pre-assigned \((\alpha, \alpha)\), where \( \alpha \leq \min(\gamma, 1/2) \). Then, if

\[
y_1^+ = x_1^+ - 1/2
\]

so that

\[
\text{p.d.f. of } y_1^+ = N[(\theta^+ - 1/2), (\sigma^+)^2] \\
= N[\mu^+, (\sigma^+)^2], \text{ say}
\]

the decision rules for sub-tests \( T^- \) and \( T^+ \) can be stated as

---

Anderson [1] proposed a number of test procedures for the case of known \( \sigma^2 \) as modifications of Wald's S.P.R.T., the purpose being to effect a reduction in the required sample size at values of the unknown mean intermediate between those being tested. His methods were approximate, being based on a Weiner process approximation to the sequence of random variables under study. Hall [4], utilizing some methods due to Hoeffding [6], showed that one of Anderson's procedures could be treated exactly. He called this procedure the sequential minimum probability ratio test.
stop sampling if \[ \left| \sum_{i=1}^{n} y_i^- \right| \geq 2 c(\sigma^-)^2 - n/4 \]
and
make decision \( d^- (\theta^- = 0) \) if \( \sum_{i=1}^{n} y_i^- < 0 \),

make decision \( d^- (\theta^- = 1) \) if \( \sum_{i=1}^{n} y_i^- > 0 \),

(ARBITRARY if \( \sum_{i=1}^{n} y_i^- = 0 \), or

continue sampling if \( \left| \sum_{i=1}^{n} y_i^- \right| < 2c (\sigma^-)^2 - n/4 \),

and

stop sampling if \[ \left| \sum_{i=1}^{n} y_i^+ \right| \geq 2 c(\sigma^+)^2 - n/4 \],
and

make decision \( d^+ (\theta^+ = 0) \) if \( \sum_{i=1}^{n} y_i^+ < 0 \),

make decision \( d^+ (\theta^+ = 1) \) if \( \sum_{i=1}^{n} y_i^+ > 0 \),

(ARBITRARY if \( \sum_{i=1}^{n} y_i^+ = 0 \), or

continue sampling if \( \left| \sum_{i=1}^{n} y_i^+ \right| < 2c (\sigma^+)^2 - n/4 \),

respectively, where \( c = \log \left( \frac{1}{2 \alpha} \right) \). Equivalent forms for the above decision rules are given by

stop sampling if \[ \left| \sum_{i=1}^{n} x_i + n(\theta_1 + \theta_2)/2 \right| \geq 2 \sigma^2 c/\Delta - n\Delta/4 \],
and

make decision \( d^- \) if \( \sum_{i=1}^{n} x_i > n(\theta_1 + \theta_2)/2 \),

make decision \( d^- \) if \( \sum_{i=1}^{n} x_i < n(\theta_1 + \theta_2)/2 \),

(ARBITRARY if \( \sum_{i=1}^{n} x_i = n(\theta_1 + \theta_2)/2 \),
continue sampling if \[ \left| \sum_{i=1}^{n} x_i + n(\theta_1 + \theta_2)/2 \right| < 2\sigma^2 c/\Delta - n\Delta/4 \], and

stop sampling if \[ \left| \sum_{i=1}^{n} x_i - n(\theta_3 + \theta_4)/2 \right| > 2\sigma^2 c/\Delta - n\Delta/4 \], and

make decision \( d_0 \) if \( \sum_{i=1}^{n} x_i < n(\theta_3 + \theta_4)/2 \),

make decision \( d_1 \) if \( \sum_{i=1}^{n} x_i > n(\theta_3 + \theta_4)/2 \),

(arbitrary if \( \sum_{i=1}^{n} x_i = n(\theta_3 + \theta_4)/2 \)),

continue sampling if \[ \left| \sum_{i=1}^{n} x_i - n(\theta_3 + \theta_4)/2 \right| < 2\sigma^2 c/\Delta - n\Delta/4 \],

respectively.

From the preceding decision rules for sub-tests \( T^- \) and \( T^+ \) it can readily be seen that \( Y_{n+1}^- \) (or \( Y_{n+1}^+ \)) is observed if and only if the inequality

\[ \left| \sum_{i=1}^{n} y_i^- \right| < 2c (\sigma^2)/\Delta - n/4 \]

holds. Furthermore, it is obvious that this inequality will be violated if

\[ 2c (\sigma^2)/\Delta^2 - n/4 < 0 \],

i.e., if

\[ n > 8c (\sigma^2)/\Delta^2 \].

Thus, an upper bound on the total sample size for test \( T \) is given by

\[ N = \lceil 8c (\sigma^2)/\Delta^2 \rceil + 1 \].
Now, a necessary and sufficient condition to ensure that both decision $d^-_1$ and decision $d^+_1$ will not be made with test $T$ is that the following inequalities always hold:

Acceptance no. (of $\theta^- = 1$) for $T^- \leq$ Acceptance no. (of $\theta^+ = 0$) for $T^+$, and

Rejection no. (of $\theta^- = 1$) for $T^- \leq$ Rejection no. (of $\theta^+ = 0$) for $T^+$.

Since the sub-tests have converging straight-line boundaries, the above inequalities will certainly hold for all values of $n > 1$ if they hold at $n = 1$ (see Figure 2), i.e., if

$$-2\sigma^2(c)/\Delta + l(3\theta_2 + \theta_1)/4 \leq -2\sigma^2(c)/\Delta + l(3\theta_4 + \theta_3)/4,$$

and

$$2\sigma^2(c)/\Delta + l(3\theta_1 + \theta_2)/4 \leq 2\sigma^2(c)/\Delta + l(3\theta_3 + \theta_4)/4,$$

or equivalently, if

$$3\theta_2 + \theta_1 \leq 3\theta_4 + \theta_3,$$

and

$$3\theta_1 + \theta_2 \leq 3\theta_3 + \theta_4,$$

which follows immediately from the initial restrictions imposed on the $\theta$-values in the Introduction.
(d^+): \( \Sigma x_i = 2\sigma^2(c)/\Delta + n(3\theta_3 + \theta_4)/4 \)

(d^0): \( \Sigma x_i = -2\sigma^2(c)/\Delta + n(3\theta_4 + \theta_3)/4 \)

(d^-): \( \Sigma x_i = 2\sigma^2(c)/\Delta + n(3\theta_1 + \theta_2)/4 \)

(d^-): \( \Sigma x_i = -2\sigma^2(c)/\Delta + n(3\theta_2 + \theta_1)/4 \)

and

FIGURE 2.
Now, let $R_n^-$ denote the acceptance region for sub-test $T^-$, i.e., let
\[
\{y(n) \in R_n^- \} \sim \{N = n, \text{ and decision } d_o^- \text{ is made} \},
\]
where
\[
y(n) = (y_1^-, y_2^-, \ldots, y_n^-),
\]
and
\[
N = \text{ the terminal sample size for sub-test } T^-.
\]
Next, let
\[
f_n^+ = [(2\pi)^{-n/2}(\sigma^-)^{-n}] e^{-1/2} \sum_{i=1}^{n} (y_i^- - 1/2)^2/(\sigma^-)^2
data_n^+ = \int_{R_n^-} f_n^+ dy(n)
\]
and
\[
f_n^0 = [(2\pi)^{-n/2}(\sigma^-)^{-n}] e^{-1/2} \sum_{i=1}^{n} (y_i^-)^2/(\sigma^-)^2
\]
Then,
\[
Pr(\text{rejecting } H^- | H^- \text{ is true}) = Pr(d_o^- \text{ using } T^- | \mu^- = 1/2) = \sum_n \int_{R_n^-} f_n^+ dy(n) = \sum_n \int_{R_n^-} \left[ \left\{ e^{-y_i^-/2(\sigma^-)^2} - n/2(\sigma^-)^2 \right\} f_n^0 dy(n) \right]
\]
\[
= \sum_n \int_{R_n^-} \left[ (-2c(\sigma^-)^2 + n/4)/2(\sigma^-)^2 - n/2(\sigma^-)^2 \right] f_n^0 dy(n)
\]
\[
= \sum_n \int_{R_n^-} \left[ e^{-c} \right] f_n^0 dy(n) = \sum_n \int_{R_n^-} (2\alpha) f_n^0 dy(n)
\]
\[
= (2\alpha)(1/2) = \alpha \leq \gamma.
\]
since
\[ \sum_{n} \int_{\hat{d}_n^{+}}^{\hat{d}_n^{-}} f_n \, dy(n) = \Pr\{d_0^{-} \text{ using } T^{-} | \mu = 0\} = 1/2, \]
because of the symmetry of the test procedure. Similarly, it can be shown that

\[ \Pr\{\text{rejecting } H^+ | H^+ \text{ is true}\} \leq \gamma, \]
and
\[ \Pr\{\text{rejecting } H_0 | H_0 \text{ is true}\} \leq 2\gamma. \]

Little can be said about the various decision probabilities under test T other than the basic guarantees stated above, since little is known about the operating characteristic functions for the sub-tests T^- and T^+ other than that they are monotone. However, this monotonicity does imply that \( \Pr\{D^+ \text{ using } T|\theta\} \) and \( \Pr\{D^- \text{ using } T|\theta\} \) are monotone. Also, by the symmetry of the formulation of the procedure T, it is readily seen that \( \Pr\{D_0 \text{ using } T|\theta\} \) is symmetric about the point \( \theta^* = (\theta_2 + \theta_3)/2 \) and at least approximately monotone (for \( \gamma \) small) on either side of \( \theta^* \).


Hall's analog of the M.P.R.T. for the case of unknown variance [5] is similar to the Baker-Hall sequential analog of the S.T.S.T.(m) procedure in that an initial sample of size m is drawn in order to estimate \( \sigma^2 \) after which sampling is continued as in the case of the M.P.R.T. procedure. As was true for the I.M.P.R.T., the
three-decision analog of the M.P.R.T. \( (m) \), denoted by I.M.P.R.T. \((m)\), is applicable to the problem stated in the Introduction only if
\[
\gamma_1 = \gamma_3 = \gamma_2/2 = \gamma, \text{ say}.
\]

In order to obtain the I.M.P.R.T. \((m)\), the sub-tests \( T^- \) and \( T^+ \) are constructed as follows: Let \( T^- \) denote a M.P.R.T. of \( \theta^- = 0 \) vs. \( \theta^- = 1 \), based on \( \bar{X}_m, \bar{X}_{m+1}, \bar{X}_{m+2}, \ldots \), with pre-assigned strength \((\alpha_m, \alpha_m)\) and with \((\sigma^-)^2 [\text{known}] \) replaced by \((s^-)^2\). Similarly, let \( T^+ \) denote a M.P.R.T. of \( \theta^+ = 0 \) vs. \( \theta^+ = 1 \), based on \( \bar{X}_m, \bar{X}_{m+1}, \bar{X}_{m+2}, \ldots \), with pre-assigned strength \((\alpha_m, \alpha_m)\), and with \((\sigma^+)^2 [\text{known}] \) replaced by \((s^+)^2\), where

\[
\alpha_m = (1/2) e^{-1/2(c_m)},
\]

(22) and

\[
c_m \geq \sqrt{(2\gamma)^{-2/v} - 1} > 0.
\]

Then, letting \( \bar{Y}_1^+ = \bar{Y}_1^- - 1/2 \) as before and computing the relevant probability ratios, one can derive the following respective decision rules for sub-tests \( T^- \) and \( T^+ \):

- stop sampling if
  \[
  \sum_{i=1}^{n} y_i^- \geq c_m(s^-)^2 - n/4; \quad n \geq m,
  \]
  and

- make decision \( d^- (\theta^- = 0) \) if \( \sum_{i=1}^{n} y_i^- < 0 \),
- make decision \( d^- (\theta^- = 1) \) if \( \sum_{i=1}^{n} y_i^- > 0 \),
- (arbitrary if \( \sum_{i=1}^{n} y_i^- = 0 \), or

- continue sampling if
  \[
  \sum_{i=1}^{n} y_i^- < c_m(s^-)^2 - n/4,
  \]
and

\[ \text{stop sampling if } \left| \sum_{i=1}^{n} y_i^+ \right| \geq c_m(s_m^+)^2 - n/4 , \]

and

\[ \text{make decision } d_0^+ (\theta^+ = 0) \text{ if } \sum_{i=1}^{n} y_i^+ < 0 , \]

\[ \text{make decision } d_1^+ (\theta^+ = 1) \text{ if } \sum_{i=1}^{n} y_i^+ > 0 , \]

(arbitrary if \( \sum_{i=1}^{n} y_i^+ = 0 \)), or

\[ \text{continue sampling if } \left| \sum_{i=1}^{n} y_i^+ \right| < c_m(s_m^+)^2 - n/4 . \]

For computational purposes it is more convenient to express the above decision rules in the following form:

\[ \text{stop sampling if } \left| n x_i + n(\theta_1 + \theta_2)/2 \right| \geq c_m(s_m^2)/\Delta - n\Delta/4 , \]

and

\[ \text{make decision } d_0^- \text{ if } \sum_{i=1}^{n} x_i > n(\theta_1 + \theta_2)/2 , \]

\[ \text{make decision } d_1^- \text{ if } \sum_{i=1}^{n} x_i < n(\theta_1 + \theta_2)/2 , \]

(arbitrary if \( \sum_{i=1}^{n} x_i = n(\theta_1 + \theta_2)/2 \)), or

\[ \text{continue sampling if } \left| -n x_i + n(\theta_1 + \theta_2)/2 \right| < c_m(s_m^2)/\Delta - n\Delta/4 , \]

and

\[ \text{stop sampling if } \left| \sum_{i=1}^{n} x_i - n(\theta_3 + \theta_4)/2 \right| \geq c_m(s_m^2)/\Delta - n\Delta/4 , \]
and

make decision \( d_0^+ \) if \( \sum_{i=1}^{n} x_i < n(\theta_3 + \theta_4)/2 \),

make decision \( d_1^+ \) if \( \sum_{i=1}^{n} x_i > n(\theta_3 + \theta_4)/2 \),

(arbitrary if \( \sum_{i=1}^{n} x_i = n(\theta_3 + \theta_4)/2 \), or

continue sampling if \( \sum_{i=1}^{n} |x_i - n(\theta_3 + \theta_4)/2| < c_m(s_m^2)/\Delta - n\Delta/4 \).

As was observed in the preceding section, simple algebraic manipulations of the above decision rules readily yield an upper bound on the overall sample size for test \( T \), once \( s_m^2 \) is known, namely

\[
N = \max \{ m, [4 c_m(s_m^2)/\Delta^2] + 1 \}.
\]

Since the outcomes of the respective sub-tests depend on the sampling rules as well as the terminal decision rules, a necessary and sufficient condition for the impossibility of making both decisions \( d_0^- \) and \( d_1^- \) using test \( T \) is that of \( n = m \), the following inequalities hold:

Acceptance no. (of \( \theta^- = 1 \)) for \( T^- \leq \) Acceptance no. of \( (\theta^+ = 0) \) for \( T^+ \), and

Rejection no. (of \( \theta^- = 1 \)) for \( T^- \leq \) Rejection no. of \( (\theta^+ = 0) \) for \( T^+ \), i.e.,

\[
-c_m(s_m^2)/\Delta + m(3\theta_2 + \theta_1)/4 \leq -c_m(s_m^2)/\Delta + m(3\theta_4 + \theta_3)/4,
\]

and
which, as before, must necessarily hold because of the initial restrictions placed on the \( \theta \)-values in Section 2 of Chapter I.

Analogous to the derivation of the Baker-Hall analog of Stein's two-stage test, it is necessary, in order to show that \( T \) satisfies the bounds on the pre-assigned error probabilities [see (1)], to consider the properties of a conditional M.P.R.T., \( T(s_m, \sigma) \).

So, for given \((s_m, \sigma)\) let \( T^-(s_m, \sigma) \) denote a conditional M.P.R.T. of \( \theta^- = 0 \) vs. \( \theta^- = 1 \), based on \( \bar{X}_m, X_{m+1}^-, X_{m+2}^-, \ldots \), with pre-assigned strength \((\bar{\alpha}_m, \alpha_m)\). Similarly, let \( T^+(s_m, \sigma) \) denote a conditional M.P.R.T. of \( \theta^+ = 0 \) vs. \( \theta^+ = 1 \), based on \( \bar{X}_m^+, X_{m+1}^+, X_{m+2}^+, \ldots \), with pre-assigned strength \((\bar{\alpha}_m^+, \alpha_m^+)\), where

\[
\bar{\alpha}_m = (1/2) e^{-1/2 \bar{c}_m}
\]

(23) and

\[
\bar{c}_m = c_m(s_m^2)/\sigma^2
\]

Then, computing the relevant probability ratios, one finds that the decision rules for sub-tests \( T^- \) and \( T^+ \) are given by

stop sampling if

\[
\left| \sum_{i=1}^{n} y_i^- \right| \geq \bar{c}_m (\sigma^-)^2 - n/4
\]

\[
= [c_m(s_m^-)^2/(\sigma^-)^2] (\sigma^-)^2 - n/4 = c_m(s_m^-)^2 - n/4; \quad n \geq m
\]
and

make decision $d_0^-$ if $\sum_{i=1}^{n} y_i^- < 0$,

make decision $d_1^-$ if $\sum_{i=1}^{n} y_i^- > 0$,

(artbitrary if $\sum_{i=1}^{n} y_i^- = 0$), or

continue sampling if $|\sum_{i=1}^{n} y_i^-| < C_m(s_m^-)^2 - n/4$,

and

stop sampling if $|\sum_{i=1}^{n} y_i^+| \geq C_m(s_m^+)^2 - n/4; \ n \geq m$,

and

make decision $d_0^+$ if $\sum_{i=1}^{n} y_i^+ < 0$,

make decision $d_1^+$ if $\sum_{i=1}^{n} y_i^+ > 0$,

(artbitrary if $\sum_{i=1}^{n} y_i^+ = 0$), or

continue sampling if $|\sum_{i=1}^{n} y_i^+| < C_m(s_m^+)^2 - n/4$,

respectively.

Now, since $T(s_m, \sigma)$ and $T$ have the same decision rules, it is fairly obvious that the necessary and sufficient condition mentioned earlier to ensure that both decision $d_1^-$ and decision $d_1^+$ will not be made with test $T$ also holds for $T(s_m, \sigma)$.

Furthermore, as was demonstrated in the development of the Baker-Hall procedure,
\[ E_S \{ \Pr(D* \text{ using } T(s_m, \sigma) \mid S_m, \sigma, \theta) \} \]

\[ = E_S \{ \Pr(D* \text{ using } T|S_m, \sigma, \theta) \} = \Pr(D* \text{ using } T|\sigma, \theta), \]

where \( D^* \) denotes some specified decision. Finally, using the same basic line of reasoning that was employed in the preceding section to prove that the M.P.R.T. satisfied the bounds on the error-probabilities given in (1), it can be shown that for \( s_m^2 \) fixed,

\[ \Pr[\bar{d}_0^- \text{ using } T^-(s_m, \sigma)|s_m, \sigma, \mu^- = 1/2] \leq \bar{\alpha}_m, \]

\[ \Pr[\bar{d}_1^- \text{ using } T^-(s_m, \sigma)|s_m, \sigma, \mu^- = -1/2] \leq \bar{\alpha}_m, \]

\[ \Pr[\bar{d}_0^+ \text{ using } T^+(s_m, \sigma)|s_m, \sigma, \mu^+ = 1/2] \leq \bar{\alpha}_m, \text{ and} \]

\[ \Pr[\bar{d}_1^+ \text{ using } T^+(s_m, \sigma)|s_m, \sigma, \mu^+ = -1/2] \leq \bar{\alpha}_m. \]

Thus, taking expectations with respect to \( S_m \) on both sides of the preceding inequalities, one obtains

(24) \[ \Pr[\bar{d}_0^- \text{ using } T^-|\sigma, \mu^- = 1/2] \]

\[ \leq (1/2) E_S \{ e^{-1/2 c_m s_m^2/\sigma^2} \sigma^2 \} = (1 + c_m/\nu)^{-\nu/2}/2, \]

(25) \[ \Pr[\bar{d}_1^- \text{ using } T^-|\sigma, \mu^- = -1/2] \leq (1 + c_m/\nu)^{-\nu/2}/2, \]

(26) \[ \Pr[\bar{d}_0^+ \text{ using } T^+|\sigma, \mu^+ = 1/2] \leq (1 + c_m/\nu)^{-\nu/2}/2, \]

and
(27) \( \Pr\{d^+_1 \text{ using } T^+|\sigma, \mu^+ = -1/2\} \leq (1 + \frac{c_m}{\nu})^{-\nu/2}/2 \),

for all values of \( \sigma^2 \). Now, from (22) we have

\[
(1 + \frac{c_m}{\nu})^{-\nu/2}/2 \leq \left[1 + \left(\frac{2\gamma}{\nu} - 1\right)/\nu\right]^{-\nu/2}/2 = \gamma,
\]

so \( \alpha_m \leq \gamma \) which implies that test \( T \) satisfies the bounds on the error probabilities stated in (1), as was demonstrated in Section 3 of Chapter I.

Once again our knowledge about the decision probability functions is limited to the type of general comments made in the preceding section.
CHAPTER III  
EMPIRICAL EVALUATIONS  

In order to make some empirical evaluation of the relative merits (w.r.t. terminal sample sizes) of the various procedures developed in the preceding sections of this paper, we shall consider the following test situation: Letting  

\[ \gamma_1 = 0.05, \quad \gamma_2 = 0.10, \quad \gamma_3 = 0.05, \]

the test  

\[ H^+: \theta \leq -0.6, \quad \text{vs.} \quad H_0: -0.1 \leq \theta \leq 0.1, \quad \text{vs.} \quad H^+: \theta \geq 0.6, \]

and compare the average sample size of the test procedures at three \( \theta \)-values, namely  

1. \( \theta = -0.6 \) (i.e., \( H^- \) is correct),  
2. \( \theta = 0.0 \) (i.e., \( H_0 \) is correct), and  
3. \( \theta = 0.2 \) (i.e., none of the hypotheses are correct).  

A Univac 1105 was used to generate 165 random samples of 100 observations each from the first 20,000 listings in the Rand table of 100,000 standard normal deviates. Then the terminal boundaries of the six test procedures were transformed so as to make the true mean of the data correspond to the particular \( \theta \)-value.

\[ ^6 \text{Occasionally, an individual trial for a particular test procedure required more than 100 observations to reach a decision. When this situation arose, additional observations were drawn at random.} \]
specified by the example under investigation. Although the variance
was obviously equal to one, it was assumed to be unknown for the
I.S.T.S.T.\(m\), I.S.P.R.T.\(m\), and I.M.P.R.T.\(m\) procedures. Since
all six of the tests were performed on the same set of data for
each example, the results that were obtained are correlated, pre-
sumably positively. Thus differences that occurred (e.g., in
variability of sample size) among the various procedures are proba-
bly more meaningful than they would have been if the trials were
independent; and similarities are less meaningful than they would
have been if the trials were independent.

The results of the empirical investigation are presented in
the tables and figures below. Table 1 (and 2) presents some
characteristics of the terminal sample size distributions for the
six procedures. From the calculated standard deviations it can
be shown that the coefficients of variation of the average sample
size for the various procedures range from about 2 to 5 percent,
which implies that the average (and median) sample sizes given in
the table are quite precise. Figures 3, 4, and 5 present some histo-
grams of the actual sample size distributions. Table 3 contains em-
pirical data on the frequency of the various decisions in each of
the examples. These data serve as estimates of the true decision
probabilities. However, these estimates are not very precise as
can be seen by appealing to the binomial coefficient of variation
formula, \( \frac{q}{np} \). For example, with an \( n \) of 165 and a \( p \) of 0.05,
the true coefficient of variation of a frequency is about 34 per-
cent. Much more empirical sampling would be required to get pre-
cise estimates of the decision probabilities. Some interpretations
of these tables follow.

### TABLE I

**Test Case 1 (\( \theta = -0.6 \))**

<table>
<thead>
<tr>
<th>Procedure</th>
<th>( n(1) )</th>
<th>( n(165) )</th>
<th>Median sample size</th>
<th>A.S.N.</th>
<th>( s_n )</th>
<th>( s_{\bar{n}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.F.S.S.T.</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>I.S.P.R.T.</td>
<td>4</td>
<td>108</td>
<td>19.33</td>
<td>24.02</td>
<td>15.49</td>
<td>1.21</td>
</tr>
<tr>
<td>I.M.P.R.T.(^7)</td>
<td>5</td>
<td>66</td>
<td>22.88</td>
<td>24.42</td>
<td>10.61</td>
<td>0.83</td>
</tr>
<tr>
<td>I.S.T.S.T.(^8)</td>
<td>16</td>
<td>108</td>
<td>42.67</td>
<td>48.46</td>
<td>19.02</td>
<td>1.48</td>
</tr>
<tr>
<td>I.S.P.R.T.(^m)</td>
<td>16</td>
<td>161</td>
<td>24.17</td>
<td>31.35</td>
<td>21.41</td>
<td>1.67</td>
</tr>
<tr>
<td>I.M.P.R.T.(^m)</td>
<td>16</td>
<td>91</td>
<td>24.61</td>
<td>29.92</td>
<td>14.98</td>
<td>1.17</td>
</tr>
</tbody>
</table>

**Test Case 2 (\( \theta = 0.0 \))**

<table>
<thead>
<tr>
<th>Procedure</th>
<th>( n(1) )</th>
<th>( n(165) )</th>
<th>( n(83) )</th>
<th>A.S.N.</th>
<th>( s_n )</th>
<th>( s_{\bar{n}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.F.S.S.T.</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>I.S.P.R.T.</td>
<td>4</td>
<td>96</td>
<td>23.05</td>
<td>26.44</td>
<td>12.03</td>
<td>0.94</td>
</tr>
<tr>
<td>I.M.P.R.T.</td>
<td>11</td>
<td>54</td>
<td>24.29</td>
<td>26.85</td>
<td>7.37</td>
<td>0.57</td>
</tr>
<tr>
<td>I.S.T.S.T.(^m)</td>
<td>16</td>
<td>108</td>
<td>42.67</td>
<td>48.46</td>
<td>19.02</td>
<td>1.48</td>
</tr>
<tr>
<td>I.S.P.R.T.(^m)</td>
<td>16</td>
<td>115</td>
<td>26.36</td>
<td>31.92</td>
<td>16.53</td>
<td>1.29</td>
</tr>
<tr>
<td>I.M.P.R.T.(^m)</td>
<td>16</td>
<td>92</td>
<td>26.75</td>
<td>30.33</td>
<td>13.11</td>
<td>1.02</td>
</tr>
</tbody>
</table>

\(^7\)For the I.M.P.R.T. (\( \sigma^2 \) known), \( N = 74 \) is an upper bound on the
terminal sample size.

\(^8\)\( m = 16. \)

\(^9\)\( s_{\bar{n}} = s_n / \sqrt{165} \), the estimated standard deviation of the A.S.N., which
serves as a measure of the accuracy of the estimated A.S.N.
## Test Case 3 (θ = 0.2)

<table>
<thead>
<tr>
<th>Procedure</th>
<th>( n(1) )</th>
<th>( n(165) )</th>
<th>( n(83) )</th>
<th>A.S.N.</th>
<th>( s_n )</th>
<th>( s_m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.F.S.S.T.</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>I.S.P.R.T.</td>
<td>4</td>
<td>116</td>
<td>26.17</td>
<td>33.69</td>
<td>22.26</td>
<td>1.73</td>
</tr>
<tr>
<td>I.M.P.R.T.</td>
<td>8</td>
<td>64</td>
<td>28.83</td>
<td>31.93</td>
<td>11.67</td>
<td>0.91</td>
</tr>
<tr>
<td>I.S.T.S.T.(m)</td>
<td>16</td>
<td>108</td>
<td>42.67</td>
<td>48.46</td>
<td>19.02</td>
<td>1.48</td>
</tr>
<tr>
<td>I.S.P.R.T.(m)</td>
<td>16</td>
<td>166</td>
<td>33.50</td>
<td>45.16</td>
<td>31.27</td>
<td>2.43</td>
</tr>
<tr>
<td>I.M.P.R.T.(m)</td>
<td>16</td>
<td>109</td>
<td>31.36</td>
<td>36.67</td>
<td>19.27</td>
<td>1.50</td>
</tr>
</tbody>
</table>
FIGURE 3: Distribution of sample sizes in Test Case 1 for the I.S.T.S.T.\(\text{m}\), I.S.P.R.T.\(\text{m}\), and I.M.P.R.T.\(\text{m}\) procedures \(\text{m}=16\).
FIGURE 4: Distribution of sample sizes in Test Case 2 for the I.S.T.S.T.\( (m) \), I.S.P.R.T.\( (m) \), and I.M.P.R.T.\( (m) \) procedures \( (m=16) \).
FIGURE 5: Distribution of sample sizes in Test Case 3 for the I.S.T.S.T.(m), I.S.P.R.T.(m), and I.M.P.R.T.(m) procedures (m=16).
Comments

1.) There is essentially no apparent difference between the I.S.P.R.T. and the I.M.P.R.T. procedures nor between the I.S.P.R.T.(m) and the I.M.P.R.T.(m) procedures in average or median sample size, though they are superior in this respect to the two-stage or non-sequential procedure. However, if we were to employ the less conservative non-sequential procedure proposed in Section 2 of Chapter II, the required sample size would be reduced from 44 to 36 which is more in line with the various sequential procedures, particularly w.r.t. test case 3. (This reduction in average sample size from 44 to 36 serves as an indication of the extent to which the I.F.S.S.T. and, indirectly, the other procedures developed in this paper are conservative when $\theta_2 \neq \theta_3$.)

2.) The I.M.P.R.T. and the I.M.P.R.T.(m) are less variable w.r.t. sample size than are the non-truncated sequential procedures. (See Table 2) Similarly, the I.M.P.R.T.(m) is less variable than the corresponding two-stage procedure except in Case 3 when they are approximately the same.

<table>
<thead>
<tr>
<th>Test Case</th>
<th>I.S.P.R.T. : I.M.P.R.T.</th>
<th>I.S.P.R.T.(m) : I.M.P.R.T.(m)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ratio of sdn.'s: ratios of ranges</td>
<td>ratio of sdn.'s: ratios of ranges</td>
</tr>
<tr>
<td>1</td>
<td>1.5:1.7</td>
<td>1.4:1.9</td>
</tr>
<tr>
<td>2</td>
<td>1.6:2.1</td>
<td>1.5:1.3</td>
</tr>
<tr>
<td>3</td>
<td>1.9:2.0</td>
<td>1.6:1.6</td>
</tr>
</tbody>
</table>

10 Obviously, the greater variability of the I.S.P.R.T. and I.S.P.R.T.(m) procedures may, for an individual trial, be favorable instead of unfavorable, as is indicated by the $n(1)$ values given in Table 1.
3.) The statistics given in Table 1 seem to indicate that the only two situations in which it would be reasonable to make use of a I.F.S.S.T. procedure are (a.) if sequential sampling is not feasible, or (b.) if the major consideration for choosing a test procedure is to obtain the smallest possible upper bound on the sample size. Furthermore, as indicated previously, the three-decision F.S.S.T. developed in Section 2, of Chapter II would be preferable to the I.F.S.S.T. in either of these situations.

4.) It is interesting to note that the mean sample size for the I.S.T.S.T. (m) is only 4.5 observations larger than the required sample size for the I.F.S.S.T. Thus, the cost (in terms of sample units) of not knowing the true variance \( \sigma^2 \) for the I.S.T.S.T. (m) is relatively small. Slightly larger differences were found in the sequential cases.
### TABLE 3

#### Test Case 1 ($\theta = -0.6$)

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Correct hypothesis</th>
<th>Frequency of Decision</th>
<th>Nominal error prob.</th>
<th>Rel. freq. of errors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I.F.S.S.T.</td>
<td>$H^+$</td>
<td>159</td>
<td>6</td>
<td>0.05</td>
</tr>
<tr>
<td>I.S.P.R.T.</td>
<td>$H^+$</td>
<td>159</td>
<td>6</td>
<td>0.05</td>
</tr>
<tr>
<td>I.M.P.R.T.</td>
<td>$H^+$</td>
<td>159</td>
<td>6</td>
<td>0.05</td>
</tr>
<tr>
<td>I.S.T.S.T. $(m)$</td>
<td>$H^+$</td>
<td>158</td>
<td>7</td>
<td>0.05</td>
</tr>
<tr>
<td>I.S.P.R.T. $(m)$</td>
<td>$H^+$</td>
<td>162</td>
<td>3</td>
<td>0.05</td>
</tr>
<tr>
<td>I.M.P.R.T. $(m)$</td>
<td>$H^+$</td>
<td>157</td>
<td>8</td>
<td>0.05</td>
</tr>
</tbody>
</table>

#### Test Case 2 ($\theta = 0.0$)

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Correct hypothesis</th>
<th>Frequency of Decision</th>
<th>Nominal error prob.</th>
<th>Rel. freq. of errors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I.F.S.S.T.</td>
<td>$H_0$</td>
<td>0</td>
<td>162</td>
<td>0.10</td>
</tr>
<tr>
<td>I.S.P.R.T.</td>
<td>$H_0$</td>
<td>4</td>
<td>159</td>
<td>0.10</td>
</tr>
<tr>
<td>I.M.P.R.T.</td>
<td>$H_0$</td>
<td>2</td>
<td>160</td>
<td>0.10</td>
</tr>
<tr>
<td>I.S.T.S.T. $(m)$</td>
<td>$H_0$</td>
<td>2</td>
<td>161</td>
<td>0.10</td>
</tr>
<tr>
<td>I.S.P.R.T. $(m)$</td>
<td>$H_0$</td>
<td>0</td>
<td>164</td>
<td>0.10</td>
</tr>
<tr>
<td>I.M.P.R.T. $(m)$</td>
<td>$H_0$</td>
<td>0</td>
<td>164</td>
<td>0.10</td>
</tr>
</tbody>
</table>

#### Test Case 3 ($\theta = 0.2$)

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Correct hypothesis</th>
<th>Frequency of Decision</th>
<th>Nominal error prob.</th>
<th>Rel. freq. of errors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I.F.S.S.T.</td>
<td>none</td>
<td>0</td>
<td>138</td>
<td>-</td>
</tr>
<tr>
<td>I.S.P.R.T.</td>
<td>none</td>
<td>1</td>
<td>145</td>
<td>-</td>
</tr>
<tr>
<td>I.M.P.R.T.</td>
<td>none</td>
<td>0</td>
<td>142</td>
<td>-</td>
</tr>
<tr>
<td>I.S.T.S.T. $(m)$</td>
<td>none</td>
<td>2</td>
<td>137</td>
<td>-</td>
</tr>
<tr>
<td>I.S.P.R.T. $(m)$</td>
<td>none</td>
<td>0</td>
<td>153</td>
<td>-</td>
</tr>
<tr>
<td>I.M.P.R.T. $(m)$</td>
<td>none</td>
<td>0</td>
<td>149</td>
<td>-</td>
</tr>
</tbody>
</table>

---

For the I.F.S.S.T., we have $Pr$[rejecting $H^+$ | $\theta = -0.6$] $\approx 0.0473$, $Pr$[rejecting $H_0$ | $\theta = 0.0$] $= 0.0210$, $Pr$[accepting $H_0$ | $\theta = 0.2$] $= .8527$, and $Pr$[accepting $H^+$ | $\theta = 0.2$] $= 0.1636$. 

---

\[11\] For the I.F.S.S.T., we have $Pr$[rejecting $H^+$ | $\theta = -0.6$] $\approx 0.0473$, $Pr$[rejecting $H_0$ | $\theta = 0.0$] $= 0.0210$, $Pr$[accepting $H_0$ | $\theta = 0.2$] $= .8527$, and $Pr$[accepting $H^+$ | $\theta = 0.2$] $= 0.1636$. 

---
5.) The observed error probabilities tended, on the whole, to be considerably smaller than the theoretical significance levels. This consideration suggests that slightly larger theoretical error probabilities than are actually desired could be employed in determining the boundary conditions for the various test procedures (particularly in the cases of the Baker-Hall and Sobel-Wald tests, where differences between the observed error probabilities and the significance levels seem to be the greatest). Although some preliminary trial-and-error investigation would be required to determine the maximum allowable "γ" levels, one would expect to obtain smaller average sample sizes and still satisfy the desired significance levels by such a technique.

6.) It should also be noted that truncation did not tend to produce any great change in the decision-making pattern for the sequential procedures. For example, although under the I.M.P.R.T.(m), the truncated counterpart of the I.S.P.R.T.(m), decision $D^+$ was made 16 times in Test Case 3, as opposed to 12 times for the Baker-Hall test, the same 12 trials resulted in decision $D^+$ for both procedures.

---

12 The conservatism observed in Test Case 2 is largely due to the fact that $\theta_2 \neq \theta_3$. 


