MULTIPLE DECISION SEQUENTIAL PROCEDURES

by

William Dennis Lawing, Jr.

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of

The Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State University
Of Science and Technology
Ames, Iowa

1965

# TABLE OF CONTENTS

# I. INTRODUCTION

## A. The Sequential Probability Ratio Test

The sequential probability ratio test (abbreviated SPRT in the sequel) was developed by Wald (1947). The theory for the SPRT is especially tractable when testing a simple null hypothesis $H_o$ against a simple alternative hypothesis $H_1$. As indicated below, further simplifications occur when these two hypotheses concern the unknown parameter $\theta$ of a distribution of Koopman-Pitman exponential type.

The test is conducted by successively comparing with two constants A and B ($B < 1 < A$) the ratio $\lambda_m$ of the likelihoods of the first m observations under $H_1$ and under $H_o$. So long as $B < \lambda_m < A$ the sampling continues, with termination and acceptance of $H_1(H_o)$ as soon as $\lambda_m \geq A$ ($\lambda_m \leq B$). The constants A and B are chosen to insure that the two error probabilities do not exceed specified levels. Wald (1947) shows that $A \leq (1-\beta)/\alpha$ and $B \geq \beta/(1-\alpha)$, where $\alpha$ = Prob. [accept $H_1|H_o$] and $\beta$ = Prob. [accept $H_o|H_1$.

Consider a one parameter ($\theta$) family of distributions admitting a derivative with respect to some measure on the real line; suppose that this derivative has the exponential form (Koopman, 1936, Pitman, 1936)

$$f(x;\theta) = \exp[U(x) + T(x)a(\theta) + b(\theta)].$$

Then, as noted by Bartlett (1946) and Tweedie (1946) in the discussion of a paper by Barnard (1946), the SPRT may be described in terms of a pair of parallel lines in the plane. Sampling is continued as long as the sample path resulting from plotting (m, $T_m$) remains between these lines, with termination and acceptance of $H_o(H_1)$ as soon as the lower (upper) line is crossed. Here m is the sample size and $T_m = \sum_{i=1}^{m} T(x_i)$.

The characteristic feature of sequential testing procedures in general and the SPRT in particular is that the sample size n required to reach a decision is not fixed in advance, but is itself a random variable. Among all test procedures (sequential or otherwise) with $\alpha \leq \alpha'$ and $\beta \leq \beta'$, and SPRT with error probabilities $\alpha'$ and $\beta'$ minimizes the expected sample size E[n] both when $H_1$ is true and when $H_o$ is true. For a proof and further discussion see Wald and Wolfowitz (1948, 1950), Lehmann (1959), and Blackwell and Girshick (1954).

## B. Multiple Decision Procedures

The problem of designing a sequential procedure for choosing one of k (k > 2) competing simple hypotheses has been studied from several points of view. Wald (1947) outlines the general nature of such a problem. Let $H_1, H_2, \ldots, H_k$ be the hypotheses under consideration; also let $R_m$ be m-dimensional sample space. Then a multiple decision sequential procedure corresponds to successively specified regions $R_{m,1}, \ldots R_{m,k+1}$, $m=1,2,\ldots$, with the property that $R_{m+1,1} + \ldots + R_{m+1,k+1} = R_{m,k+1} \times R_1$. As sampling proceeds, $(x_1, \ldots, x_m)$ is observed for successive m; sampling stops with acceptance of $H_i$, $i=1, \ldots k$, as soon as $(x_1, \ldots, x_m)$ falls into $R_{m,i}$. The problem then is to select the successive regions $R_{m,i}$ such that the performance of the procedure can be evaluated, and, if possible, such that some optimality criterion is satisfied.

In analogy to the two-decision case, Wald considers k performance characteristics; these are k-1 operating characteristic (OC) functions $L_i(\theta) = \text{Prob.}[\text{accept } H_i|\theta]$, and the ASN function, $E[n|\theta]$.

Wald also defines weight functions $W_i(\theta)$ $(i=1,2, \ldots k)$ which measure the loss resulting from deciding for $H_i$ when $\theta$ is the value of the parameter. Also defined is the expected loss or risk $r(\theta) = \sum_{i=1}^{k} L_i(\theta)W_i(\theta)$ which, in a sense, is the summarization into a single function of the k-1 OC functions $L_i(\theta)$. One question then is to find a large family of procedures for which $r(\theta)$ and $E[n|\theta]$ can be computed; another problem is to find a large family of procedures, and an optimality criterion involving $r(\theta)$ and $E[n|\theta]$, such that the optimum member of this family can be found.

Wald carries this problem to a partial solution, using weight functions which are essentially characteristic functions for certain zones of preference for the several hypothesized values of $\theta$. Given these weight functions it is possible to construct a family of procedures with bounded risk; ASN optimization within this class is not carried through. Specifically, the procedures of the family in question involve construction of successive confidence sets for $\theta$ of fixed conservative level $1-r_o$. Sampling stops as soon as a confidence set is entirely within a zone of preference, with acceptance of the corresponding hypothesis. Assuming sure termination, it is clear that $r(\theta) \leq r_o$. For example, if $\theta$ is in the zone of preference for $H_1{:}\theta=\theta_1$, $L_1(\theta)\geq 1-r_o$, $1-L_1(\theta) \leq r_o$, $W_1(\theta) = 0$, and $W_i(\theta)\leq 1$, $i\neq 1$, from which it follows that $r(\theta)\leq r_o$. A similar procedure is given by Paulson (1963) for testing the mean of a normal distribution both with known and unknown variance.

Illustrating prior work with emphasis on computing performance characteristics rather than optimization, Sobel and Wald (1949) give a sequential procedure for choosing one of three hypotheses concerning the mean of a normal distribution with known variance. This procedure consists of two SPRT's conducted simultaneously and may be described geometrically in terms of two pairs of parallel lines. If the sample path crosses the upper line before the lower in both tests, accept $H_3:\theta$ is "large". If the path crosses the lower lines of both tests before the upper, accept $H_1:\theta$ is "small". Otherwise, accept $H_2:\theta$ is "medium". The OC functions $L_i(\theta)$ $(i=1,2,3,)$, are derived by exploiting the fact that $L_1(\theta)$ and $L_3(\theta)$ are OC functions of ordinary two-decision SPRT's; $L_2(\theta)$ is then found by subtraction, since the procedure terminates with probability one. These OC calculations are, of course, exact only in the case of no excess over the stop boundaries.

This procedure is used by several authors for distributions other than the normal. DeBoer (1953) applies it to the testing of a binomial parameter. An interesting feature of DeBoer's paper is an exact representation of the ASN function $E[n|p]$ which can be generalized as follows: Let $n$ = the sample size at termination of the process, $N^*$ = the sample size on first crossing the initial wedge, $N_{12}$ = the sample size at termination of the lower SPRT, $N_{32}$ = the sample size at termination of the upper SPRT, then $E[n|p] = E[N_{12}|p] + E[N_{32}|p] - E[N^*|p]$, and $E[N_{12}|p]$ and $E[N_{32}|p]$ are simply the ASN functions for the two SPRT's involved and may be easily found. The function $E[N^*|p]$ is the ASN for the initial wedge which is not known in general, but which may be calculated by summing probabilities for the finite number of paths involved in the binomial case treated by DeBoer.

A similar scheme is given by Armitage (1947), who constructs a two-sided sequential test of Student's hypotheses, based on the binomial distribution of the number of exceedences of $\mu_o$, the hypothesized value of the normal mean. Armitage (1957, 1960) also constructs three-decision sequential procedures for testing two-sided hypotheses concerning the mean of a normal distribution with known variance and the proportion of a binomial distribution; these plans, called restricted procedures, have the outer boundaries of the procedures above, but replace the two inner boundaries by a single truncating line. Schneiderman and Armitage (1962) introduce yet another class of procedures, called closed sequential procedures, for two-sided testing of a normal mean with known variance. These tests of two-sided hypotheses again are three decision procedures. Again the usual outer boundaries are retained, while the inner boundaries are curved lines intersecting the outer ones, forming two wedge-shaped regions. The closed procedures are constructed with the same outer boundaries as the equivalent restricted procedures, but the truncation point is moved out to compensate for the decreased probabilities of crossing the outer boundaries caused by moving the center of the middle boundary towards the origin. The OC functions for both the restricted and closed procedures are approximated for the Wiener process. The approximation involves considering the two outer boundaries independently, ignoring the effect of one upon the other.

Lechner and Ginsburg[1] (1963) also combine two SPRT's to form a three-decision procedure; the application is to the exponential distribution. They further suggest combining k SPRT's into a (k+1)-decision procedure, an idea discussed in Chapter III.

Blackwell and Girshick (1954) characterize Bayes sequential procedures for choosing one of k hypotheses both for the truncated and untruncated case. Such procedures depend on a partitioning of the k-simplex $\Sigma$, the space of probability distributions over the parameter space; sampling is terminated with acceptance of $H_i$ as soon as the posterior (or prior) distribution enters a region of $\Sigma$ corresponding to $H_i$. These regions vary with sample size in the truncated case, but are fixed in the untruncated case. Explicit construction of these regions for $k > 2$, especially in the untruncated case, seems difficult; however, a particular example for k=3, reduceable to the case k=2, is solved.

C. Sequential Procedures with Wedge Shaped Continuation Regions

Although eventual termination of the SPRT is assured, there is no guaranteed upper limit for the number of samples required to make a decision; in addition, the sample size may be quite large for certain intermediate values of $\theta$. Thus, the usefulness of the SPRT is limited in situations where samples are expensive or difficult to procure. One remedy is truncation; however, work on optimization of truncated procedures is not complete (Mengido, 1963).

---

[1]Lechner, J. A., Amherst, Mass. Amherst discussion. Private communication. 1964.

Another remedy is to use members of a large class of procedures introduced by Weiss (1953). These procedures, the generalized sequential probability ratio tests (GSPRT), whose completeness is discussed by Sobel (1953), Kiefer and Weiss (1957), and DeGroot (1961) replace the constants B and A of the SPRT with two constants $B_m$ and $A_m$ depending on m. When the $A_m$ and $B_m$ lie on two intersecting lines, the GSPRT has the following characteristics: 1. The sample size required for decision is bounded. 2. The procedure satisfies a necessary condition for admissability due to Kiefer and Weiss (1957, page 60). 3. In the case of an exponential family, the procedure is amenable to OC computations which exploit terminal likelihood ratios. Such procedures are called "wedge procedures" in this thesis.

Anderson (1960) considers the Wiener process for a wedge procedure with possible truncation before the intersection point.[1] Anderson derives the OC and ASN functions for these wedges as infinite series involving Mill's ratio. The OC is given as a definite integral for symmetric untruncated wedges. Anderson indicates that, in those cases where equal error rates for the two hypotheses are desired, it may be sufficient to restrict attention to wedges which are symmetric in a sense to be discussed in Chapter IV. As indicated in that chapter such symmetric wedges also offer additional opportunities for OC computations.

Donnelly (1957) also treats the wedge for the Wiener process.[1] His approach is to solve the diffusion equation with boundary conditions

---

[1]Anderson, T. W., Amherst, Mass. Amherst discussion. Private communication. 1964.

corresponding to the wedge; he obtains a series for the probability of accepting $H_o$ on a differential portion of the boundary.

Hall (1961) has developed the minimum probability ratio test (MPRT), which specializes to a wedge procedure, to construct a test of $\theta_1$ vs. $\theta_2$, Hall introduces a third parameter value $\theta_o$ $(\theta_1 < \theta_o < \theta_2)$; wedge boundaries arise from the requirement that the terminal ratio $f_{2m}/f_{om}$ or $f_{1m}/f_{om}$ have specified values. The MPRT bounds a given weighted average of two risks. Fukushima (1961) applies the MPRT to the Poisson distribution.

## D. Summary

Chapter II of this thesis reviews the idea of conjugacy for exponential families; the discussion includes consideration of conjugacy for multidimensional Wiener processes.

Multidecision procedures based on wedges are discussed in Chapter III. Hypotheses are successively eliminated as the random walk based on a sufficient statistic traverses a system of wedges. The next two chapters discuss OC computations for such multidecision procedures, all of these computations being based on terminal likelihood ratios. Some simplified wedge systems are given, for which special OC computations of this type are available.

Chapter IV treats the Wiener process. For two-decision procedures, bounds for the OC function are given and its asymptotic behavior described; this asymptotic behavior, for the symmetric case, is derivable from Anderson (1960), relation 4.63.

Asymptotic behavior of the OC functions for multidecision procedures is considered next; this behavior is then exploited in Chapter VI in connection with a characterization theorem. For certain symmetric wedge systems exact OC's are derived for an infinite grid of parameter points. Chapter IV also discusses generalizations to the multiparameter case; here we obtain OC expressions for certain special symmetric tubes analagous to the SPRT bands of the one-parameter case; these multiparameter OC computations furnish, in fact, certain absorption probabilities for Brownian motion in many dimensions.

The binomial and Poisson cases are discussed in Chapter V. Approximate OC's, ignoring excess, are given. In the case of symmetric wedges, a special method yields OC values on an infinite grid analagous to that arising in the normal case; such symmetric procedures are especially relevant for comparisons of two binomial populations when the hypotheses are symmetric about 1/2. Also treated is the three-decision binomial case, with wedges specialized into SPRT bands. In both cases the problem of excess is treated in detail. Risk bounds are computed which are analagous to those obtained by Wald for the SPRT.

The last chapter considers some questions of the performance of such procedures. It is shown, in a certain asymptotic sense, that, in the case of the Wiener process, the best multidecision procedure is the SPRT when performance at only the two extreme parameter points is considered.

# II. PARAMETRIC CONJUGACY FOR THE SPRT

## A. Conjugacy for Exponential Families

It was noted in Chapter I that, for a family of the Koopman-Pitman exponential type, the SPRT may be characterized by a pair of parallel lines. The SPRT for $H_o : \theta = \theta_o$ vs. $H_1 : \theta = \theta_1$, with error probabilities not exceeding given levels $\alpha$ and $\beta$, depends on two constants $A = (1-\beta)/\alpha$ and $B = \beta/(1-\alpha)$. For successive values of m the likelihood ratio $\lambda_m$ is compared to the constants A and B. If

$$\lambda_m \geq A, \tag{2.1}$$

stop and accept $H_1$; if

$$\lambda_m \leq B, \tag{2.2}$$

stop and accept $H_o$; otherwise continue sampling.

Now, if the family of probability distributions admits a derivative with respect to some measure on the line, we have

$$\lambda_m = \frac{\prod_{j=1}^{m} f(x_j ; \theta_1)}{\prod_{j=1}^{m} f(x_j ; \theta_o)} .$$

Further, if this derivative is of the form

$$f(x ; \theta) = \exp[U(x) + T(x)\, a(\theta) + b(\theta)], \tag{2.3}$$

it is seen that

$$\log \lambda_m = [a(\theta_1) - a(\theta_o)] \sum_{j=1}^{m} T(x_j) + m[b(\theta_1)-b(\theta_o)] = aT_m + bm,$$

where

$$T_m = \sum_{j=1}^{m} T(x_j),$$

$$a = a(\theta_1) - a(\theta_o), \tag{2.4}$$

and

$$b = b(\theta_1) - b(\theta_o). \tag{2.5}$$

The test may then be conducted by comparing $\log \lambda_m$ with log A and log B. The relations 2.1 and 2.2 are respectively equivalent to

$$T_m \geq h_1 + sm \tag{2.6}$$

and

$$T_m \leq h_o + sm, \tag{2.7}$$

where $h_1 = 1/a \log A$, $h_o = 1/a \log B$ and $s = -b/a$, and a and b are defined in 2.4 and 2.5. The SPRT is now conducted by plotting $T_m$ vs. m for successive m until one of conditions 2.6 or 2.7 is met, and then making the proper decision.

The fact that this geometric characterization depends on only three parameters ($h_o$, $h_1$, and s), while the likelihood ratio characterization depends on four parameters ($\theta_o$, $\theta_1$, $\alpha$ and $\beta$), has been exploited by Girshick (1946), Baker (1950), David (1952), Davies (1954), Blasbalg (1957), and Lechner (1964) to yield the OC function.

Let $(\theta_o, \theta_1)$ be any pair of parameter points satisfying

$$c(\theta_1) \equiv s\ a(\theta_1) + b(\theta_1)$$

$$= s\ a(\theta_o) + b(\theta_o) \equiv c(\theta_o). \qquad (2.8)$$

Such a pair will be called conjugate[1] with respect to s. Then for any sample sequence $(x_1 \ldots x_m)$ terminating near the line of 2.6, we have, in view of the equivalance of 2.6 and 2.1, that

$$\lambda_{1m} \doteq e^{h_1 a} \qquad (2.9)$$

Similar reasoning shows that, for $(\theta_1, \theta_o)$ satisfying 2.8 and a sample sequence $(x_1 \ldots x_m)$ terminating near the line of 2.7,

$$\lambda_{om} \doteq e^{h_o a} \qquad (2.10)$$

From 2.9 and 2.10 and the fact that the procedure terminates with probability one, it follows, by summing over all sample sequences terminating near these respective lines, that

$$L(\theta_1) \doteq e^{h_o a}\ L(\theta_o) \qquad (2.11)$$

and

$$1-L(\theta_1) \doteq e^{h_1 a}\ [1-L(\theta_o)] \qquad (2.12)$$

Equality holds in 2.11 and 2.12 only if there is no excess over the boundaries at termination, i.e. only if every sample sequence terminates

---

[1]This usage is due to Lechner (1964).

exactly on one of the lines; this zero excess is realized, for example, if instead of a discrete sampling scheme, we consider a stochastic process with continuous parameter such as the Wiener process (Dvoretzky, Kiefer, and Wolfowitz, 1953).

## B. The Binomial, Normal and Poisson Cases

We give in this section the explicit formulation of the conjugacy relation 2.8 for three cases.

### 1. The binomial case

The useful expressions of Section A in the binomial case are as below.

$$T_m = \sum_{j=1}^{m} x_j \ ,$$

$$a(p) = \log \frac{p}{1-p} \ ,$$

$$b(p) = \log (1-p) \ ,$$

$$c(p) = s \log \frac{p}{1-p} + \log (1-p) \ , \tag{2.13}$$

and

$$\lambda_{im} = \left[\frac{p_1(1-p_o)}{p_o(1-p_1)}\right]^{h_i} , \ i = 1, 2. \tag{2.14}$$

### 2. The Poisson case

The expressions of Section A for the Poisson case are,

$$T_m = \sum_{j=1}^{m} x_j \ ,$$

$$a(\theta) = \log \theta \ ,$$

$$b(\theta) = - \theta \ ,$$

$$c(\theta) = s \log \theta - \theta \ , \qquad (2.15)$$

and

$$\lambda_{im} = \left[\frac{\theta_1}{\theta_o}\right]^{h_i} . \qquad (2.16)$$

It can be shown that relations 2.15 and 2.16 apply as well to the Poisson process.

3. The normal case

In the normal case we find,

$$T_m = \sum_{j=1}^{m} x_j \ ,$$

$$a(\mu) = \mu,$$

$$b(\mu) = - 1/2 \ (\mu^2 + \log 2\pi),$$

$$c(\mu) = s \ \mu - 1/2 \ (\mu^2 + \log 2\pi), \qquad (2.17)$$

and

$$\lambda_{im} = e^{h_i(\mu_1 - \mu_o)} . \qquad (2.18)$$

## C. The Wiener Process

Consider the one-dimensional Wiener process X(t) discussed in Section A of Chapter IV. Consider as well two drift parameters $\mu_o$ and $\mu_1$ which are conjugate with respect to s in the sense that $c(\mu_1) = c(\mu_o)$, where $c(\cdot)$ is given by 2.17; this implies $s=(1/2)(\mu_o+\mu_1)$. Also consider any one-dimensional boundary that includes a linear portion 1: $X = h+st$, and let $g_\mu(\cdot)$ be the density postulated for that linear portion in Section A of Chapter IV. Then if $p = (t, x)$ is any point on 1,

$$\frac{g_{\mu_1}(p)}{g_{\mu_o}(p)} = \frac{\phi\left(\frac{x-\mu_1 t}{\sqrt{t}}\right)}{\phi\left(\frac{x-\mu_o t}{\sqrt{t}}\right)} = e^{h(\mu_1-\mu_o)} , \qquad (2.19)$$

which is 2.18.

Extending to the two-dimensional case, consider the Wiener process Z(t) discussed in Section A of Chapter IV. Consider as well two drift parameters $(\theta_1, \theta_2)$ and $(\nu_1, \nu_2)$ which are conjugate with respect to the line

$$a_1\mu_1 + a_2\mu_2 + b = 0 \qquad (2.20)$$

in the sense that $(\theta_1, \theta_2)$ and $(\nu_1, \nu_2)$ are at the same distance d from and on the same perpendicular to the line 2.20 (see Figure 1). Also consider any two-dimensional stop boundary which includes a planar portion

$$a_1X_1 + a_2 X_2 + bt = 1, \qquad (2.21)$$

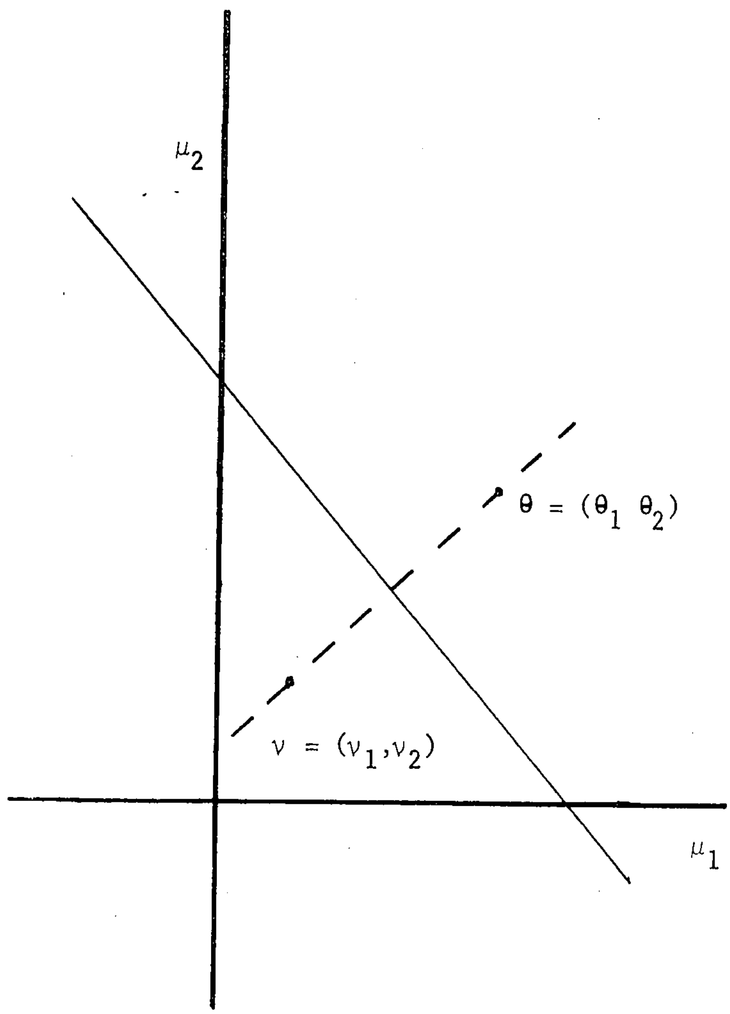


Figure 1. Bivariate normal conjugacy

and let $g_{\mu_1,\mu_2}(\cdot)$ be the density postulated for that planar portion in Section A of Chapter IV; then, if $p = (t, x_1, x_2)$ is any point on 2.21,

$$\frac{g_{\theta_1,\theta_2}(p)}{g_{\nu_1,\nu_2}(p)} = \frac{\phi\left(\frac{x_1-\theta_1 t}{\sqrt{t}}\right)\ \phi\left(\frac{x_2-\theta_2 t}{\sqrt{t}}\right)}{\phi\left(\frac{x_1-\nu_1 t}{\sqrt{t}}\right)\ \phi\left(\frac{x_2-\nu_2 t}{\sqrt{t}}\right)}$$

$$= e^{\frac{1}{a_i}(\theta_i-\nu_i)}, \quad i = 1 \text{ or } 2. \tag{2.22}$$

Analogues of 2.20, 2.21 and 2.22 for higher dimensions are as follows:

$$a_1\mu_1 + a_2\mu_2 + \ldots + a_r\mu_r + b = 0. \tag{2.23}$$

$$a_1X_1 + a_2X_2 + \ldots + a_rX_r + bt = 1. \tag{2.24}$$

$$\frac{g_{\theta_1,\theta_2,\ldots\theta_r}(p)}{g_{\nu_1,\nu_2,\ldots\nu_r}(p)} = e^{\frac{1}{a_i}(\theta_i-\nu_i)}, \quad i = 1 \text{ or } 2 \text{ or } \ldots \text{ or } r. \tag{2.25}$$

D. Conjugacy Apart from the SPRT

As implied in the previous section parametric conjugacy, though apparently only considered heretofore within the framework of the SPRT, should be useful whenever paths terminate on suitable boundaries, e.g., straight lines in the case of exponential families. This is the case, for example, when paths must terminate on one of two intersecting straight line boundaries. The resulting continuation regions are the wedges discussed in Section C of Chapter I.

Hall's MPRT (1961) apparently also utilizes terminal likelihood ratios in the analysis of wedges; however, exploiting parametric conjugacy in the manner of this thesis seems not to have been done previously.

# III. MULTIPLE DECISION SEQUENTIAL PROCEDURES WHOSE BOUNDARIES ARE WEDGES

## A. Introductory Remarks

In this chapter we exhibit some sequential procedures for deciding among k alternative hypotheses concerning a single unknown parameter of a distribution of Koopman-Pitman form. Such a procedure is best envisioned as a sequence of k-1 stages, each stage being terminated by the crossing of one of a pair of intersecting lines (we allow that any sequence of terminal stages may feature parallel lines). We give two general OC relations which, under certain circumstances or assumptions, provide explicit OC functions; these are considered in Chapters IV and V. Some simplified versions of the k-decision procedure are presented as well.

## B. Two- and Three-Decision Procedures

Two-decision procedures were discussed in Section C of Chapter I. A typical three-decision wedge system is pictured in Figure 2. To explain the decision procedure we introduce the following notation. Let the symbol $[WX|YZ]$ denote the event: the sample path crosses line WX before crossing line YZ. Then define the following events:

$$U \equiv [h_1A|h_oA],$$

$$L \equiv [h_oA|h_1A],$$

$$UU \equiv U \text{ and } [h_{U1}A_U|h_{Uo}A_U],$$

$$UL \equiv U \text{ and } [h_{Uo}A_U|h_{U1}A_U],$$

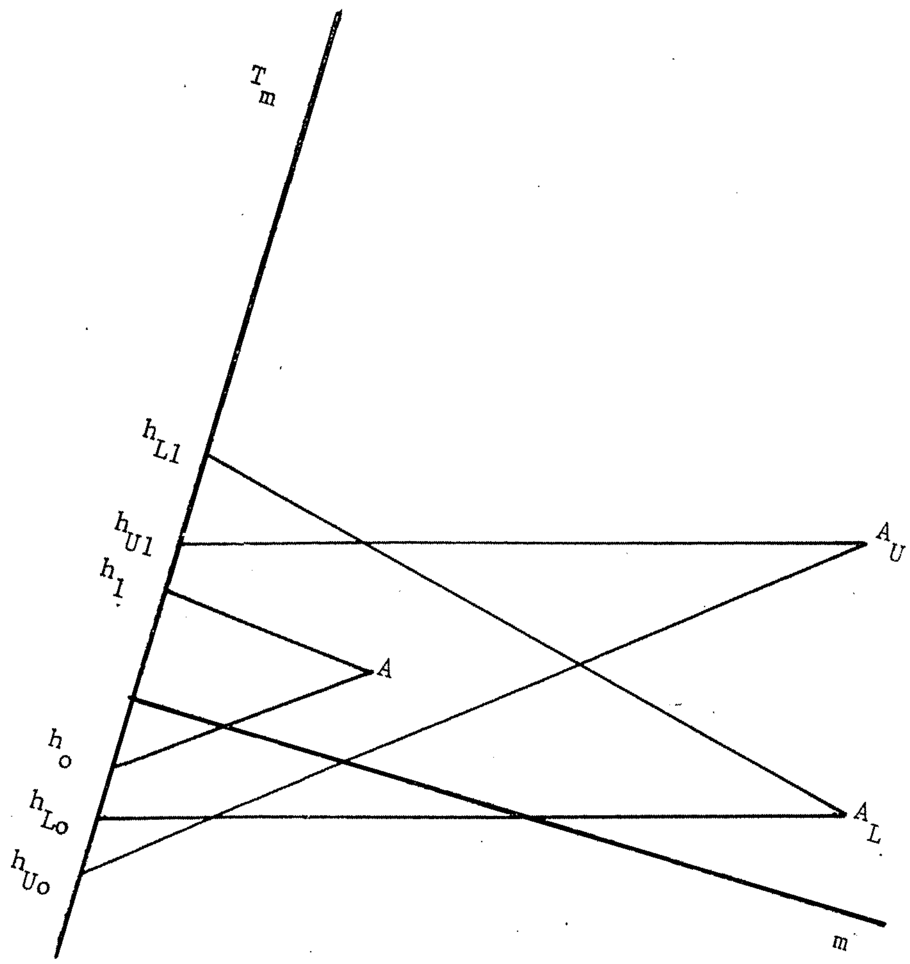


Figure 2. Three-decision wedge system

$$LU \equiv L \text{ and } [h_{L1}A_L | h_{Lo}A_L],$$

$$LL \equiv L \text{ and } [h_{Lo}A_L | h_{L1}A_L].$$

The event UU results in acceptance of $H_3 : \theta = \theta_3$ ; either UL or LU results in acceptance of $H_2 : \theta = \theta_2$, and LL leads to acceptance of $H_1 : \theta = \theta_1$, where $\theta_3 > \theta_2 > \theta_1$.

Let W denote the wedge $h_1Ah_o$, $W_U$ the wedge $h_{U1}A_Uh_{Uo}$, and $W_L$ the wedge $h_{L1}A_Lh_{LO}$. The wedge system is so arranged that occurrence of event U places the sample path within the wedge $W_U$, while occurrence of event L places the sample path within $W_L$. Finally, let $P(L|\theta)$, $P(LL|\theta)$, $P(LU|\theta)$ etc. be the probabilities of the events L, LL, LU etc. when the value of the parameter is $\theta$. We will now derive some useful likelihood ratio relations using the parametric conjugacy developed in Chapter II.

If $L_i(\theta)$ is the probability of accepting $H_i$ given $\theta$, we see that $L_3(\theta) = P(UU|\theta)$, $L_2(\theta) = P(UL|\theta) + P(LU|\theta)$, and $L_1(\theta) = P(LL|\theta)$. The OC problem then is reduced to computing the functions $P(XX|\theta)$.

Let $s_{xo}$ and $s_{x1}$ be the slopes of the upper and lower boundaries respectively in the wedge $W_x$, then for example $s_o$ is slope of $h_1A$, $s_1$ the slope of $h_oA$, $s_{Uo}$ the slope of $h_{U1}A_U$, etc. For a given parameter value $\theta$, let $\theta'(s)$ be its conjugate relative to the slope s and let $a(s) = a(\theta) - a(\theta')$ (see Equation 2.4). Then, by 2.11 or 2.12 and Section D of Chapter II, we have

$$P(L|\theta) \doteq e^{h_o a(s_1)} P(L|\theta'(s_1)) \tag{3.1}$$

$$1-P(L|\theta) \doteq e^{h_1 a(s_o)} [1-P(L|\theta'(s_o))]. \tag{3.2}$$

Relations 3.1 and 3.2 can be made in certain cases to yield solutions (or approximations) for $P(L|\theta)$. Assuming $P(L|\theta)$, and employing 2.11 2.12 once more, this time for wedge $W_L$, we have

$$P(LL|\theta) = e^{h_{Lo}a(s_{L1})} P(LL|\theta'(s_{L1})), \tag{3.3}$$

$$P(L|\theta) - P(LL|\theta) = e^{h_{L1}a(s_{Lo})}[P(L|\theta'(s_{Lo})) - P(LL|\theta'(s_{Lo}))]. \tag{3.4}$$

In addition, using $P(U|\theta) = 1-P(L|\theta)$ and 2.11 or 2.12 for $W_U$,

$$P(UL|\theta) = e^{h_{Uo}a(s_{U1})} P(UL|\theta'(s_{U1})), \tag{3.5}$$

$$P(U|\theta) - P(UL|\theta) = e^{h_{U1}a(s_{Uo})}[P(U|\theta(s_{Uo})) - P(UL|\theta'(s_{Uo}))]. \tag{3.6}$$

Again, in some cases we will be able to obtain $P(LL|\theta)$ and $P(LU|\theta)$ from 3.3 and 3.4, and in a similar manner $P(UL|\theta)$, and $P(UU|\theta)$ from 3.5 and 3.6.

A simplification results if we let $h_o = h_{Uo} = h_{Lo}$, $h_1 = h_{U1} = h_{L1}$, $s_o = s_{Lo}$, and $s_1 = s_{U1}$; then the wedge system consists of two primary wedges $W_U : h_1A_Uh_o$, and $W_L : h_1A_Lh_o$, with the first stage wedge being formed by the intersection at the lower boundary of $W_U$ and the upper boundary of $W_L$. Then, since $UU = [h_1A_U|h_oA_U]$ and $LL = [h_oA_L|h_1A_L]$,

$$L_3(\theta) = e^{h_1a(s_{Uo})} L_3(\theta'(s_{Uo})), \tag{3.7}$$

$$L_1(\theta) = e^{h_oa(s_{L1})} L_1(\theta'(s_{L1})), \tag{3.8}$$

$$1-L_3(\theta) = e^{h_oa(s_1)}[1-L_3(\theta'(s_1))], \tag{3.9}$$

$$1-L_1(\theta) = e^{h_1 a(s_o)}[1-L_1(\theta'(s_o))]. \qquad (3.10)$$

If in addition $s_1 = s_{U1} = s_{Uo}$ and $s_o = s_{Lo} = s_{L1}$, the wedge system becomes two pairs of parallel lines as in the three-decision systems of Armitage (1947) and Sobel and Wald (1949). These simplifications are amenable to OC computations discussed below.

Still another simplification results if all the slopes are the same. The wedges $W_1$, $W_U$, and $W_L$ then become SPRT bands and OC's are obtainable as the probabilities of compound events, e.g. $L_1(\theta) = P(LL|\theta) = P(L|\theta) \times P(LL|L;\theta)$; $P(L|\theta)$ is the OC for the SPRT W and $P(LL|L;\theta)$ is the OC for the SPRT $W_L$ conditional upon the path starting on the lower boundary of W.

## C. Multiple Decision Procedures

The procedures discussed in the preceding section for three decisions are generalized for k decisions; the notation and terminology will be the same. A typical procedure for k = 4 is illustrated in Figure 3. A procedure for choosing one of k hypothesized values, $\theta_1\ \theta_2\ .\ .\ .\ \theta_k$, is a sequence of k - 1 stages; the first stage is a wedge $W : h_1 A h_o$, the second stage is one of the wedges $W_U : h_{U1} A_U h_{Uo}$ or $W_L : h_{L1} A_L h_{Lo}$, depending on the outcome of the first stage, the jth stage is one of $2^{j-1}$ wedges, etc. We will denote the events relating to successive boundaries crossed by a set of symbols (X . . . X) where the X in the jth position is either a U or an L depending on which boundary was crossed at the jth stage. At the termination of the (j-1)th stage the path

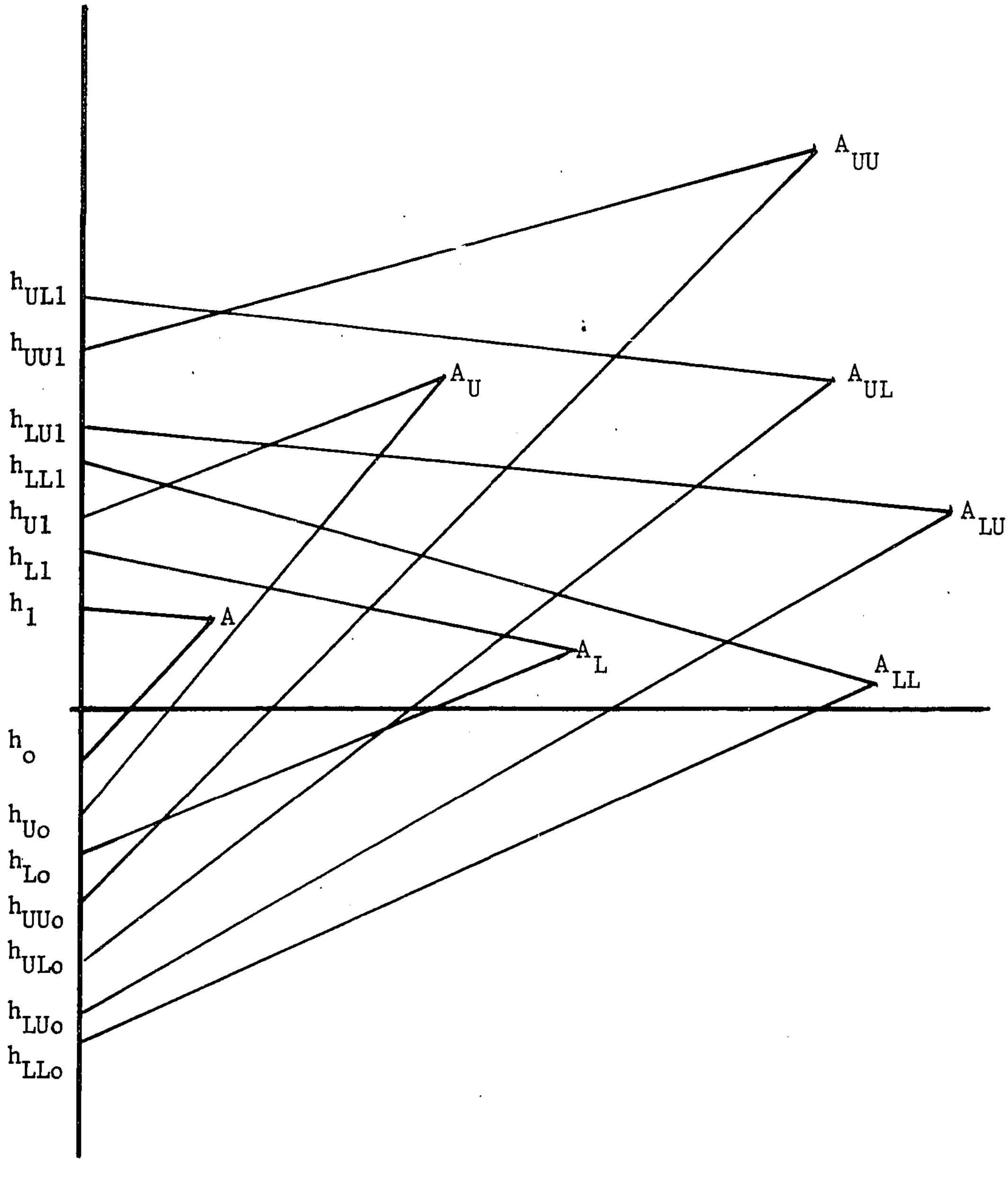


Figure 3. Four-decision wedge system

history is a set of (j-1) symbols, U or L; hence there are $2^{j-1}$ possible path types. The $2^{j-1}$ wedges of the $j^{th}$ stage are subscripted as follows: $W_{X\ldots X} : h_{X\ldots X1}\, A_{X\ldots X}\, h_{X\ldots Xo}$ where (X . . . X) is the path history leading to this wedge. The decision process may now be expressed as follows: accept $H_i : \theta = \theta_i$ if an event (X . . . X) occurs featuring (i-1) U's and (k-i) L's among the (k-1) X's. Thus a path leads to acceptance of $H_i : \theta = \theta_i$ if and only if it terminates on the upper boundary $h_{X\ldots X1}\, A_{X\ldots X}$ of a wedge $W_{X\ldots X}$ featuring (k-i) L's and (i-2) U's among the (k-2) X's, or on the lower boundary $h_{X\ldots Xo}\, A_{X\ldots X}$ of a wedge $W_{X\ldots X}$ featuring (k-i-1) L's and (i-1) U's among the (k-2) X's. For example, if k = 4 we would accept $H_3 : \theta = \theta_3$ if the path history (XXX) had 2 U's and 1 L among the 3 symbols; there are 3 such events: UUL, ULU, and LUU. These are paths terminating on the upper boundaries of wedges $W_{UL}$ and $W_{LU}$ and on the lower boundary of $W_{UU}$. Define $P_j$ $(X\ldots X|\theta)$ to be the probability of the event (X . . . X) consisting of (j-1) symbols when $\theta$ is the value of the parameter. Then $L_i(\theta) = \text{Prob.}[\text{accept } H_i|\theta] = \Sigma\, P_k(X\ldots X|\theta)$, where the summation is over all (X . . . X) having (i-1) U's and (k-i) L's among the (k-1) X's. Again we see that the OC problem is reduced to that of computing the functions $P_k(X\ldots X|\theta)$. We give the following likelihood ratios

$$P_{j+1}(X\ldots XL|\theta) = e^{h_{X\ldots Xo}\, a(s_{X\ldots X1})}\, P_{j+1}(X\ldots XL|\theta'(s_{X\ldots X1})) \tag{3.11}$$

$$P_j(X \ldots X|\theta) - P_{j+1}(X \ldots XL|\theta) = \qquad (3.12)$$

$$e^{h_{X \ldots X1}\, a(s_{X \ldots Xo})}[P\,(X \ldots X|\theta'(s_{X \ldots Xo})$$

$$- P_{j+1}(X \ldots XL|\theta'(s_{X \ldots Xo}))]$$

which in certain cases allow computation of the $P_k(X \ldots X|\theta)$ by induction on j.

## D. A Simplification

Underlying relations 3.11 and 3.12 is the fact that imbedded in a k-decision procedure based on $(2^{k-1} - 1)$ wedges $W$, $W_U$, $W_L$, . . . $W_{U \ldots U}$, . . . $W_{L \ldots L}$, are a succession of 2,3, . . . j, . . ., and (k-1) decision procedures based respectively on $W$; $W$, $W_U$, $W_L$ ; . . . and $W$, $W_U$, $W_L$,. . .,$W_{U \ldots U}$, . . . $W_{L \ldots L}$. If any of these imbedded decision procedures is degenerate in a special way, the OC computations become simplified; if the k-decision procedure itself is so degenerate the computations become especially simple. The type of degeneracy to which we refer involves the coincidence of all the intercepts of the upper lines for each wedge and of all the lower lines, as illustrated for j = 4 in Figure 4. The degeneracy also manifests itself in part by the coincidence of some wedge boundaries involved in the k-decision procedure; but the main point is that all the wedges are formed by intersections of (j - 1) primary wedges $W_{j-1}$,$W_{j-2}$, . . . $W_1$, all based on the same two intercepts $h_1$ and $h_o$.

First we consider the substantial simplification which ensues if

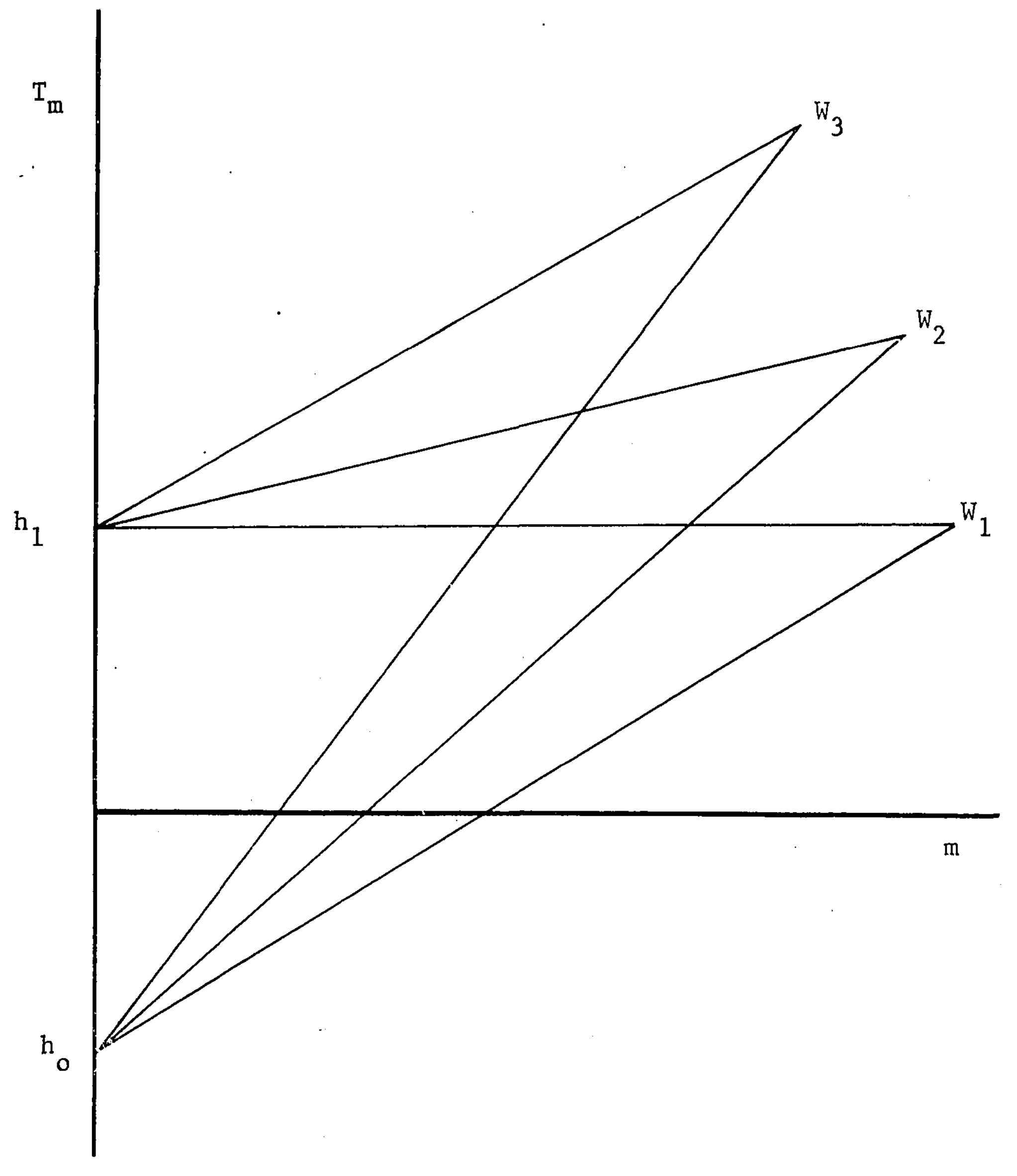


Figure 4. Simplified four-decision wedge system

$j = k$, i.e., if the k-decision procedure itself is so degenerate. Let $g_i(\theta)$ be the probability that the sample path crosses the upper boundary of principle wedge $W_i$ before the lower boundary when $\theta$ is the value of the parameter. Then we have

$$L_i(\theta) = g_{i-1}(\theta) - g_i(\theta), \tag{3.13}$$

where $g_k(\theta) \equiv 0$, $g_o(\theta) \equiv 1$ and the remaining $g_i(\theta)$ are computable as for a simple two-decision wedge. As an illustration consider the four-decision procedure of Figure 4. Then $L_4(\theta) = g_3(\theta) - 0$, or the probability of accepting $H_4 : \theta = \theta_4$ is the probability of crossing the upper boundary of $W_3$ before the lower boundary. A check reveals that precisely, the paths leading to acceptance of $H_4$, (UUU), do indeed cross $h_1 A_{UU}$ before crossing $h_o A_{UU}$. Also, $L_3(\theta) = g_2(\theta) - g_3(\theta)$, or the paths leading to acceptance of $H_3 : \theta = \theta_3$ are those which cross the upper boundary of $W_2$ before the lower except the ones crossing the upper boundary of $W_3$ before the lower. These paths are (UUL), (ULU), and (LUU) as they should be. It may be noted that the special case wherein the principle wedges degenerate into pairs of parallel lines is the case mentioned by Lechner and Ginsburg (1963), and for $k = 3$, is just the Sobel-Wald (1949) procedure.

We next consider a lesser simplification, pertaining to the case where $2 < j < k$, i.e. where only an embedded j-decision procedure is degenerate.

We first note that if the k-decision procedure is degenerate the rule of Section C for assigning decisions to events states that $H_i$ is accepted when the k-decision process terminates on the upper boundary of $W_{i-1}$ or on the lower boundary of $W_i$. Hence, walks terminating a degenerate j-decision procedure (imbedded in a k-decision procedure) on the same boundary of the same wedge have histories (X . . . X) with the same number of U's and L's. It is thus reasonable to assume of the k-decision procedure ($R_k$) in which our degenerate j-decision procedure ($R_j$) is imbedded that paths terminating $R_j$ on the same boundary of the same wedge should meet the same additional wedge system in $R_k$ beyond $R_j$. Under this assumption it is possible to begin the iteration of 3.11 and 3.12 with functions $u_i(\theta)$ and $l_i(\theta)$ which are defined as follows: $u_i(\theta) =$ Prob.[$R_j$ terminates on the upper boundary of $W_i|\theta$], $l_i(\theta)$ = Prob[$R_j$ terminates on the lower boundary of $W_i|\theta$]. The functions $u_i(\theta)$ and $l_i(\theta)$ are computed in a manner analagous to 3.13:

Consider the (j-1)-decision procedure $R_{j-1}$ imbedded in $R_j$; $R_{j-1}$ is degenerate and its hypotheses and principle wedges will be denoted respectively by $G_i$ and $V_i$. Recalling that $g_i(\theta)$ is the probability that a path crosses the upper boundary of $W_i$ before the lower, and defining $h_i(\theta)$ as the probability that a path crosses the upper boundary of $V_i$ before the lower for parameter value $\theta$, we have from 3.13 and the decision rule of Section C applied to $R_{j-1}$

$$
\begin{aligned}
h_{i-1}(\theta) - h_i(\theta) &= \text{Prob.}[R_{j-1} \text{ terminates with acceptance of } G_i|\theta] \\
&= \text{Prob. } [R_{j-1} \text{ terminates on upper boundary of } V_{i-1} \\
&\qquad \text{or on lower boundary of } V_i|\theta] \\
&= \text{Prob. } [R_j \text{ terminates on a boundary of } W_i|\theta] \\
&= u_i(\theta) + l_i(\theta) \qquad (3.14)
\end{aligned}
$$

Now using 3.13 and the decision rule of Section C on $R_j$

$$g_{i-1}(\theta) - g_i(\theta) = \text{Prob.}[R_j \text{ terminates with acceptance of } H_i|\theta]$$

$$= l_i(\theta) + u_{i-1}(\theta) \tag{3.15}$$

Relations 3.14 and 3.15 now yield $u_i(\theta)$ and $l_i(\theta)$.

# IV. COMPUTATIONS FOR THE WIENER PROCESS

## A. Introductory Remarks

We shall consider a Wiener process $[X(t); t \geq 0]$ with drift $\mu$; its properties to be used first are as follows: (1) $X(t)$ is normal with mean $\mu t$ and variance $t$ for all $t \geq 0$; (2) $[X(t) ; t \geq 0]$ is continuous with probability one, which allows us to assume that the process terminates on any reasonably regular one-dimensional stop boundary without excess; (3) given that the process terminates on such a boundary, we postulate the existence of a density $g_\mu(\cdot)$ defined on such a boundary with the property that: (a) the probability that the process terminates on any Borel set of the boundary is obtainable by integrating $g_\mu(\cdot)$, and (b) $g_\mu(p)/g_{\mu'}(p) = \phi(\frac{x-\mu t}{\sqrt{t}})/\phi(\frac{x-\mu' t}{\sqrt{t}})$, where $p = (t,x)$ and $\phi(.)$ is the standard normal density function.

We will also utilize independent multivariate Wiener processes; in particular, in the bivariate case, this means a process $[Z(t) = X_1(t), X_2(t); t \geq 0]$ with the following properties: (1) $Z(t)$ is an independent bivariate normal with mean $(\mu_1 t, \mu_2 t)$ and covariance matrix $\Sigma = It$, for any $t \geq 0$; (2) as implied by the continuity of its two components, $[Z(t); t \geq 0]$ is continuous with probability one, which allows us to assume that the process terminates on any reasonably regular two-dimensional boundary without excess; (3) given that the process terminates on such a boundary, we postulate the existence of a density $g_{\mu_1,\mu_2}(\cdot)$ defined on such a boundary with the property that: (a) the probability that the process terminates on any Borel set of the boundary is obtainable by integrating $g_{\mu_1,\mu_2}(\cdot)$, and (b) $g_{\mu_1,\mu_2}(p)/g_{\mu_1',\mu_2'}(p) =$

$$\phi(\frac{x_1-\mu_1 t}{\sqrt{t}})\phi(\frac{x_2-\mu_2 t}{\sqrt{t}})/\phi(\frac{x_1-\mu_1' t}{\sqrt{t}})\phi(\frac{x_2-\mu_2' t}{\sqrt{t}}), \text{ where } p = (t_1\ x_1,\ x_2).$$

In Section B of this chapter we give a method for obtaining the OC functions for certain symmetric wedge systems. This method, based on conjugacy, yields the OC for parameter points on a certain "grid"; the OC thus found is exact for the Wiener process.

For the case of a single symmetric wedge, Anderson (1960) has given the OC in the form of an integral. Equating this integral to the OC obtained on the grid allows evaluation of the integral for the set of agruments corresponding to the grid.

In Section C, multivariate conjugacy yields OC computations for certain symmetric tube systems which are analogues of systems of SPRT's in the one-dimensional case; the results are, in fact, absorption probabilities for multidimensional Brownian motion.

Section D is devoted to the asymptotic behavior of the OC of an abritrary two-decision wedge. It is shown first that Anderson's integral form may be used for this purpose in the symmetric case; lacking symmetry, the procedure is to derive the asymptotic behavior from likelihood ratios. Bounds for the OC are also derived.

Asymptotic OC's for multiple wedge systems are treated in Section E; these results will be of use in Chapter VI.

The final section acknowledges the inherent limitations in any argument based exclusively on terminal likelihood ratios.

## B. Symmetric Wedge Systems

For a symmetric wedge, $h_1 = -h_o = h$, where $h_o$ and $h_1$ are the respective intercepts of the lower and upper lines. If $s_1$ and $s_o$ are the respective slopes of the lower and upper lines, define $\bar{s} = (1/2)(s_0+s_1)$, $\delta = s_1-s_o$.

Symmetry requires that, when $\mu = \bar{s}$, the probability that a sample path crosses the lower line before the upper line is 1/2, or $P(L|\bar{s}) = 1/2$. This can be seen by noting that Prob.[X(t) touches $-h + s_1t$ before $h + s_ot$] = Prob. [$Y(t) \equiv X(t) - \mu t$ touches $-h + (s_1-\mu)t$ before $h + (s_o-\mu)t$]. In other words, for $\mu = \bar{s}$ , we have a process Y(t) without drift with a boundary symmetric in the ordinary sense, and under Y(t) all paths touching the lower line have mirror images touching the upper line.

Starting with $P(L|\bar{s}) = 1/2$, we are now able to compute the OC on the grid $\mu = \bar{s} + \delta r$, where r ranges over the integers; the conjugates of $\bar{s} + \delta r$ relative to $s_1$ and $s_o$ respectively are $\bar{s} - \delta(r-1)$ and $\bar{s} - \delta(r+1)$ Now consider the density $g_\mu(\cdot)$ assumed in Section A to exist on the lower boundary of the wedge. Then using 2.19 and integrating over the lower boundary, we find:

$$P(L|\bar{s}+\delta r) = e^{-h\delta(2r-1)}P(L|\bar{s}-\delta(r-1)). \qquad (4.1a)$$

Similarly, integrating over the upper boundary

$$1-P(L|\bar{s}+\delta r) = e^{h\delta(2r+1)}[1-P(L|\bar{s}-\delta(r+1))]. \qquad (4.1b)$$

Successive application of 4.1a and 4.1b for $r = 0, \pm 1, \pm 2, \ldots$ yields the OC for the grid. As an example consider the wedge with $h = 1$,

$s_1 = -s_o = 1$; then $\bar{s} = 0$ and $\delta = 2$. We have by 4.1 and 4.2,

$$\begin{aligned} P(L|0) &= 1/2 \text{ by symmetry,} \\ P(L|2) &= e^{-2}P(L|0) = (1/2)e^{-2}, \\ 1-P(L|-2) &= e^{-2}[1-P(L|0)] = (1/2)e^{-2}, \\ P(L|4) &= e^{-6}\ P(L|-2) = e^{-6}[1-(1/2)e^{-2}], \\ 1-P(L|-4) &= e^{-6}[1-P(L|2)] = e^{-6}[1-(1/2)e^{-2}], \text{ etc.} \end{aligned} \tag{4.2}$$

We now give two examples illustrating calculation of OC functions on a grid for multiple wedge systems.

If a k-decision procedure is degenerate in the sense described in Section C of Chapter III, and if the (k-1) principle wedges of this procedure are symmetric, and further, if their grids have common points, we are then able to compute the functions $g_i(\mu)$ of relation 3.13 for each wedge $W_i$ at each value of $\mu$ on the common grid. Such procedures are similar to those proposed by Sobel and Wald (1949) and Lechner and Ginsburg (1963). One way to insure the existence of a common grid for all the wedges in such a system is to construct all the wedges with apex points whose abcissas are the same and whose ordinates differ by some even integral multiples of h. Such a system is given in Figure 5 for k = 5. Here h = 1/2 and the r principle wedges $W_4$, $W_3$, $W_2$, and $W_1$ respectively have apexes (1, 4), (1, 3), (1, 1), and (1, -1). The grids for these wedges are respectively $4 \pm r$, $3 \pm r$, $1 \pm r$, and $-1 \pm r$ (r= 0, 1, . . .). Hence we can compute $g_i(\mu)$ for each wedge for $\mu$ an

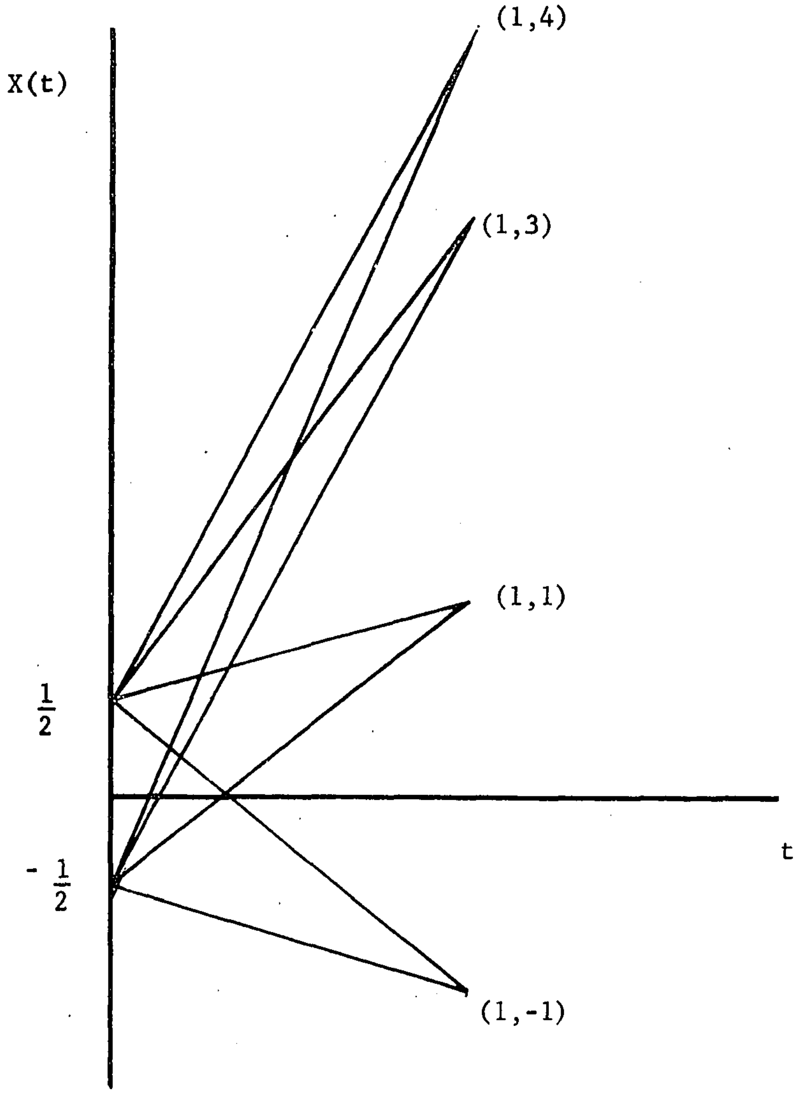


Figure 5. Symmetric five-decision procedure

integer, thereby getting $L_i(\mu)$ ($i= 1,2,3,4$) from 3.13 for integer $\mu$

Another example is the system of Figure 6. For wedge W, $h = 1$, $s_o = -1/2$, $s_1 = 1/2$, $\delta= 1$, $\bar{s} = 0$ and the grid is $\mu = \pm r$ ($r= 1,2, \ldots$). The wedges $W_U$ and $W_L$ are simple SPRT bands as in the Sobel-Wald procedure; for $W_U$, $h_{Uo} = -2$, $h_{U1} = 2$ $s_U = 1$, and for $W_L$, $h_{Lo} = -1$, $h_{L1} = 2$, $s_L = -1/2$. We are able to compute $P(L|\mu)$ for $\mu$ = an integer by symmetry for $\mu = 0$ and application of 4.1 and 4.2.

$$P(L|0) = 1/2,$$

$$P(L|1) = e^{-1}P(L|0) = (1/2)e^{-1} = 1-P(L|-1),$$

$$P(L|2) = e^{-3}P(L|-1) = e^{-3}(1-(1/2)e^{-2})=1-P(L|-2),$$

$$P(L|3) = e^{-5}P(L|-2) = e^{-5}[1-e^{-3}(1-(1/2)e^{-2})]= 1-P(L|-3), \text{ etc.}$$

Now using relations 3.3 and 3.4 and the fact that $W_L$ is an SPRT band we have

$$P(LL|0) = e^{-2}P(LL|-1)$$

$$P(L|0) - P(LL|0) = e^{2}[P(L|-1) - P(LL|-1)],$$

which, in view of the fact that $P(L|0)$ and $P(L|-1)$ are known, may be solved to yield

$$P(LL|-1) = \frac{e^{2}-1}{e^{2}-e^{-2}}$$

$$P(LL|0) = \frac{1-e^{-2}}{e^{2}-e^{-2}} \quad .$$

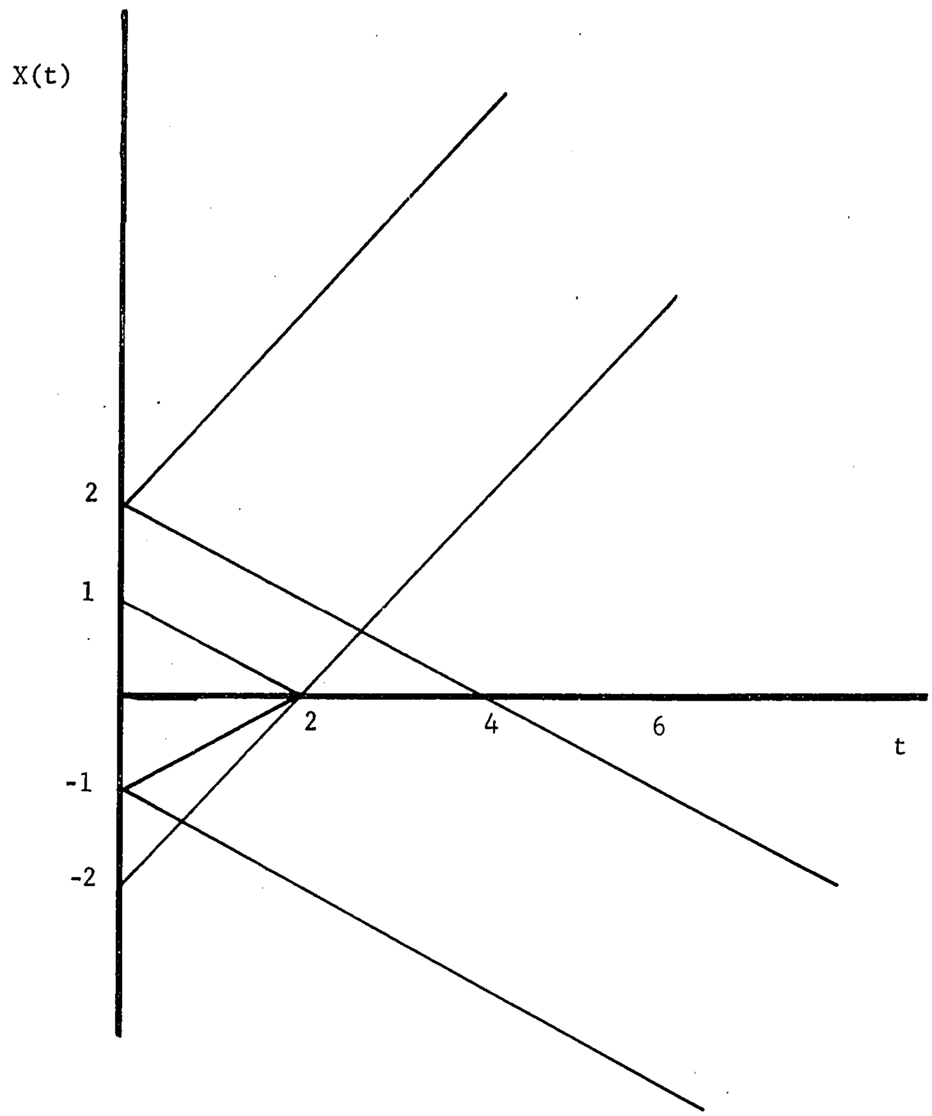


Figure 6. Special three-decision procedure

Using relations 3.5 and 3.6 and the fact that $W_U$ is an SPRT band we have

$$P(UL|0) = e^{2}(P(UL|2)$$

$$P(U|0) - P(UL|0) = e^{-2}[P(U|2) - P(UL|2)],$$

which, in view of the fact that $P(U|0)$ and $P(U|2)$ are known, yield

$$P(UL|2) = \frac{e^{-2}-e^{-7}+(1/2)(e^{-9}-1)}{e^{-2}-e^{2}}$$

$$P(UL|0) = \frac{1-e^{-5}+(1/2)(e^{-7}-e^{2})}{e^{-2}-e^{2}}.$$

We then have

$$L_1(0) = P(LL|0) = \frac{1-e^{-2}}{e^{2}-e^{-2}} = 0.1192,$$

$$L_3(0) = P(UU|0) = P(U|0) - P(UL|0)$$

$$= \frac{1+(1/2)(e^{2}+e^{-7})-e^{-5}-e^{-2}}{2(e^{2}-e^{-2})} = .6277,$$

and

$$L_2(0) = 1-L_1(0) - L_3(0) = 0.2531.$$

We can in this manner find $L_i(\mu)$ for any integer $\mu$.

It remains, as an ancillary exercise, to evaluate, for special arguments, the integral

$$P_1(T) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(z-\mu\sqrt{T})^2/2} \cdot \frac{e^{(2c/\sqrt{T})z}}{1+e^{(2c/\sqrt{T})z}}\, dz \qquad (4.3)$$

appearing in Equation 4.63 of Anderson (1960). Let W be a wedge symmetric about the t-axis, whose slopes have absolute value $s = c/T$ and whose intercepts have absolute value c. Then $P_1(T)$ is the probability that the process touches the upper boundary of W before the lower. In other words, in the notation of this thesis, 4.3 becomes

$$P(U|\mu) = 1-P(L|\mu)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-(z-\mu\sqrt{h/s})^2}{2}} \cdot \frac{e^{2\sqrt{sh}\,z}}{1+e^{2\sqrt{sh}\,z}}\, dz. \tag{4.4}$$

In the case of W, $\bar{s} = 0$ and $\delta = 2s$, so that the grid has the form 2sr ($r=0, \pm 1, \pm 2, \ldots$), and 4.4 becomes

$$1-P(L|2sr) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(z-2r\sqrt{hs})^2/2} \cdot \frac{e^{2\sqrt{sh}\,z}}{1+e^{2\sqrt{sh}\,z}}\, dz,$$

or, letting $a = 2\sqrt{sh}$,

$$1-P(L|2sr) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(z-ra)^2/2} \cdot \frac{e^{az}}{1+e^{az}}\, dz. \tag{4.5}$$

In view of our solutions of 4.2, the integral on the right of 4.5 can be solved for r an integer. For example, for $r = 0, -1, -2, -3,$

$$1-P(L|0) = 1/2,$$

$$1-P(L|-2s) = e^{2hs}[1-P(L|0)] = (1/2)e^{2hs}$$

$$= (1/2)e^{a^2/2} = P(L|2s),$$

$$1-P(L|-4s) = e^{6hs}\ [1-P(L|2s)]$$

$$= e^{3a^2/2}\ [1-(1/2)e^{a^2/2}]$$

$$= P(L|4s),$$

and

$$1-P(L|-6s) = e^{10hs}[1-P(L|4s)]$$

$$= e^{5a^2/2}(1-e^{3a^2/2}+(1/2)^{2a^2}).$$

C. Multidimensional Brownian Motion

Consider a cylinder set in $(X_1, X_2, t)$-space with triangular base and planar boundary portions A, B, and C as indicated in Figure 7. We shall compute the probability that a two-dimensional Wiener process $[Z(t) = (X_1(t), X_2(t); t \geq 0]$ starting at the origin first touches the cylindrical boundary on a particular one of these three planar boundary portions comprising the cylinder. Define $A(\mu_1,\mu_2)$ to be the probability that the process first touches boundary A when $(\mu_1,\mu_2)$ is the drift parameter, and similarly for $B(\mu_1,\mu_2)$ and $C(\mu_1,\mu_2)$. We now observe that the parameter points $(\nu_1,\nu_2)$ and $(-\nu_1,-\nu_2)$ of Figure 8 are conjugate with respect to the line $\beta$. Consider the densities $g_{\nu_1,\nu_2}(\cdot)$ and $g_{-\nu_1,-\nu_2}(\cdot)$

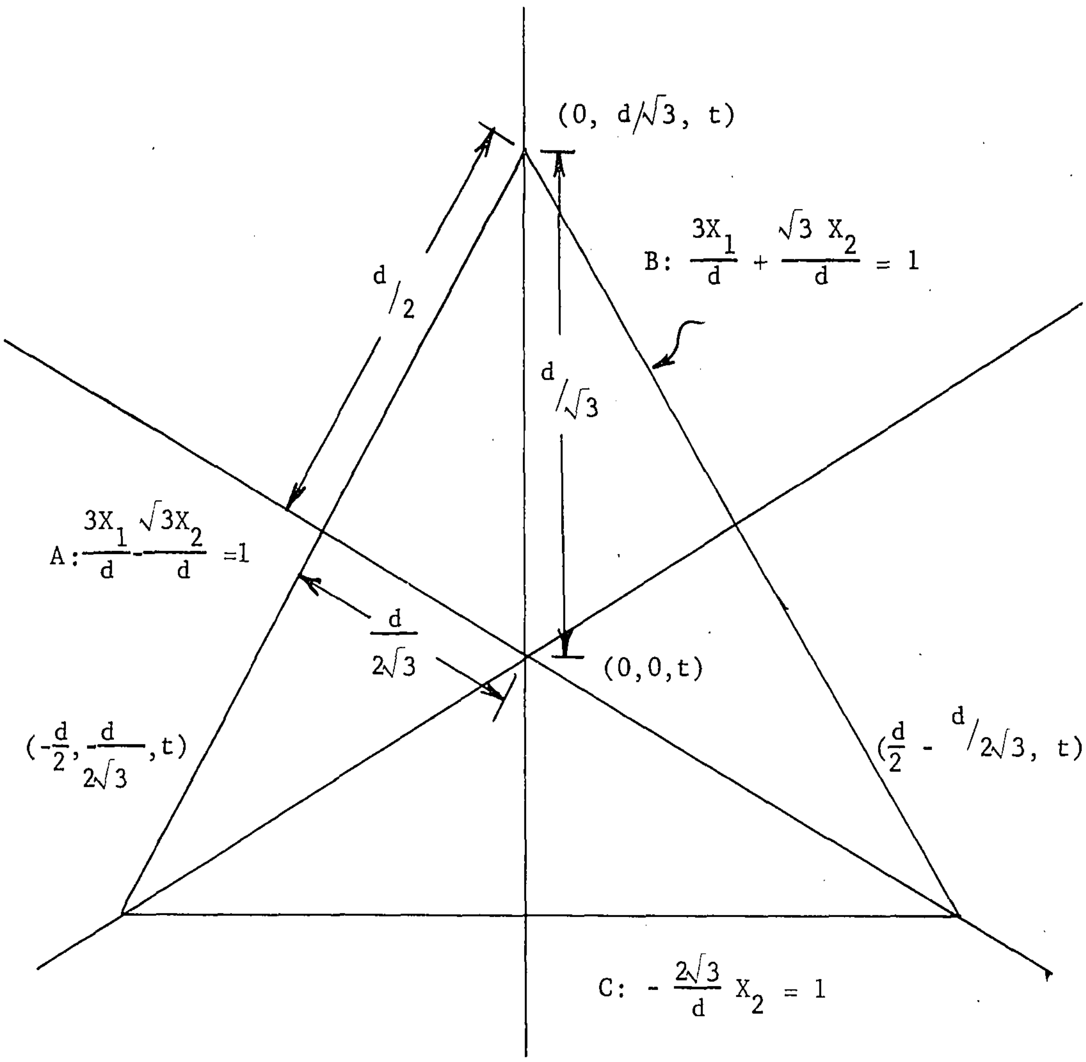


Figure 7. Isosceles triangular boundary in $(X_1(t), X_2(t))$ - plane

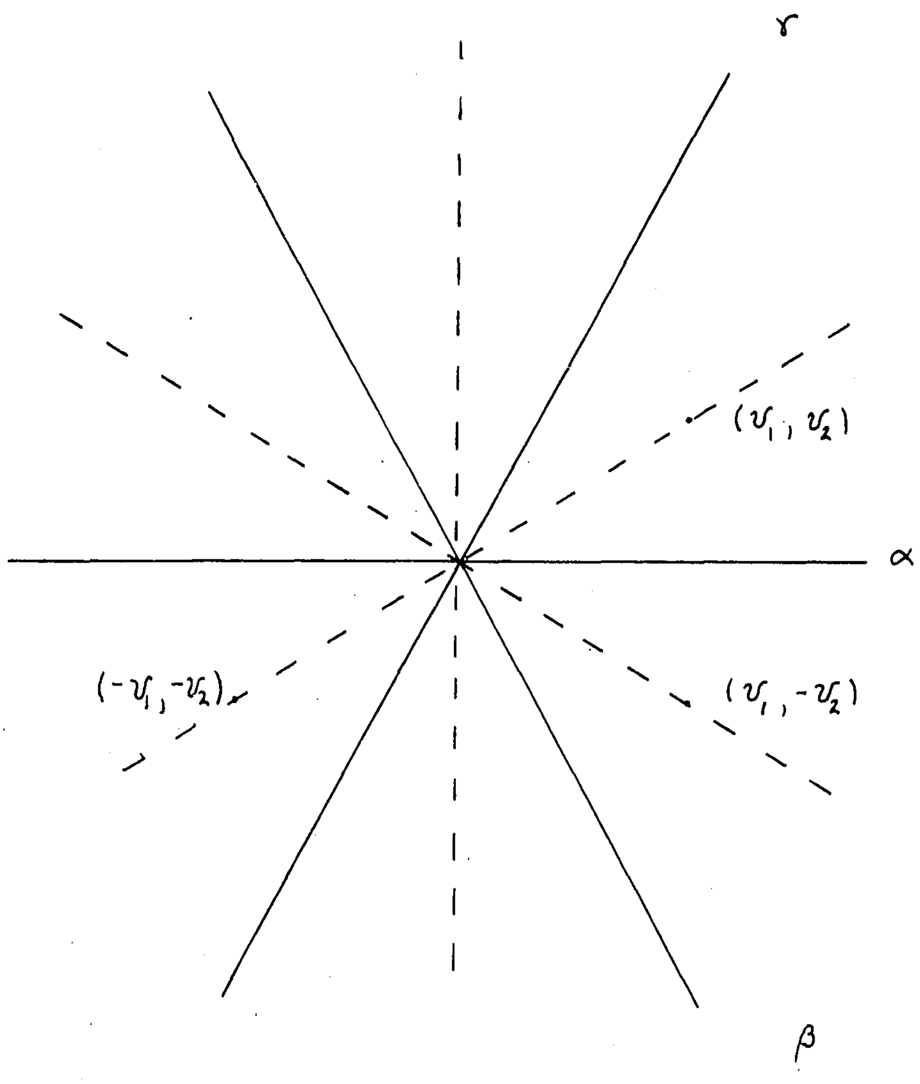


Figure 8. Conjugacy and symmetry lines in the $(\mu_1, \mu_2)$-plane

postulated in Section A for the planar boundary portion B. If p is any point on B, it is clear, referring to 2.20, 2.21 and 2.22, that the ratio $g_{\nu_1,\nu_2}(p)/g_{-\nu_1,-\nu_2}(p)$ is $e^{2d\nu_2/\sqrt{3}}$. Hence, integrating over B, we obtain

$$\frac{B(\nu_1,\nu_2)}{B(-\nu_1-\nu_2)} = e^{2d\nu_2/\sqrt{3}} . \tag{4.6}$$

A similar argument for the points $(\nu_1,\nu_2)$ and $(\nu_1,-\nu_2)$ and planar boundary portion A yields

$$\frac{A(\nu_1,\nu_2)}{A(\nu_1,-\nu_2)} = e^{-d\nu_2/\sqrt{3}} . \tag{4.7}$$

In addition, symmetry yields

$$2A(\nu_1,\nu_2) + B(\nu_1,\nu_2) = 1, \tag{4.8}$$

$$2A(\nu_1,-\nu_2) + B(-\nu_1,-\nu_2) = 1. \tag{4.9}$$

Equations 4.6, 4.7, 4.8 and 4.9 now may be solved to get

$$A(\nu_1,-\nu_2) = \frac{e^{d\nu_2/\sqrt{3}} - e^{\sqrt{3}\,d\nu_2}}{2[1-e^{\sqrt{3}\,d\nu_2}]} , \tag{4.10}$$

and

$$A(\nu_1,\nu_2) = \frac{1-e^{d\nu_2/\sqrt{3}}}{2[1-e^{\sqrt{3}d\nu_2}]} . \tag{4.11}$$

Relations 4.10 and 4.11 bear resemblence to Wald's OC function for the normal SPRT.

Consider next a cylinder set in $(X_1,X_2,t)$-space with square base and planar boundary portions A, B, C and D as in Figure 9. We shall compute the probability that a two-dimensional Wiener process $Z(t) = [X_1(t), X_2(t)];\ t \geq 0$ starting of the origin first touches the cylindrical boundary on a particular one of these four planar boundary portions comprising the cylinder. Define $A(\mu_1,\mu_2)$, $B(\mu_1,\mu_2)$, $C(\mu_1,\mu_2)$ and $D(\mu_1,\mu_2)$ as before. We now observe that the parameter points $(\nu,\nu)$ and $(\nu, -\nu)$ of Figure 10 are conjugate with respect to the line $\gamma$. Consider the densities $g_{\nu,\nu}(\cdot)$ and $g_{\nu,-\nu}(\cdot)$ postulated in Section A for the planar boundary portion C. Then in analogy to the previous example, if p is any point on C, $g_{\nu,\nu}(p)/g_{\nu,-\nu}(p) = e^{d\nu}$. Hence, integrating over C we have

$$\frac{C(\nu,\nu)}{C(\nu,-\nu)} = e^{d\nu}. \tag{4.12}$$

Symmetry also yields

$$2C(\nu\ \nu) + 2C(\nu,-\nu) = 1. \tag{4.13}$$

Relations 4.12 and 4.13 now may be solved for

$$C(\nu,-\nu) = \frac{1}{2(1+e^{d\nu})} \tag{4.14}$$

and

$$C(\nu,\nu) = \frac{e^{d\nu}}{2(1+e^{d\nu})} . \tag{4.15}$$

Relations 4.14 and 4.15 extend to $^1/k\ (1+e^{d\nu})$ and $e^{d\nu}/k\ (1+e^{d\nu})$ in k-dimensional space.

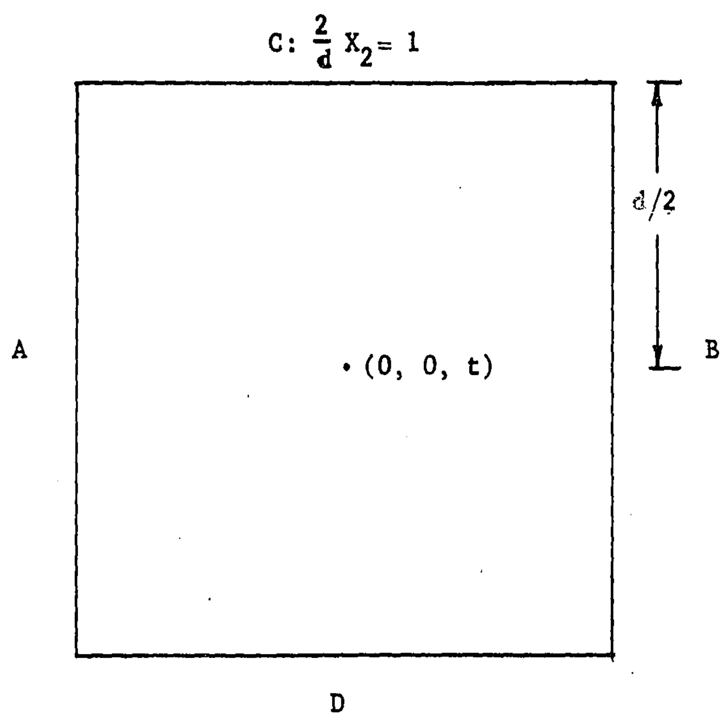


Figure 9. Square boundaries in $(X_1(t), X_2(t))$-plane

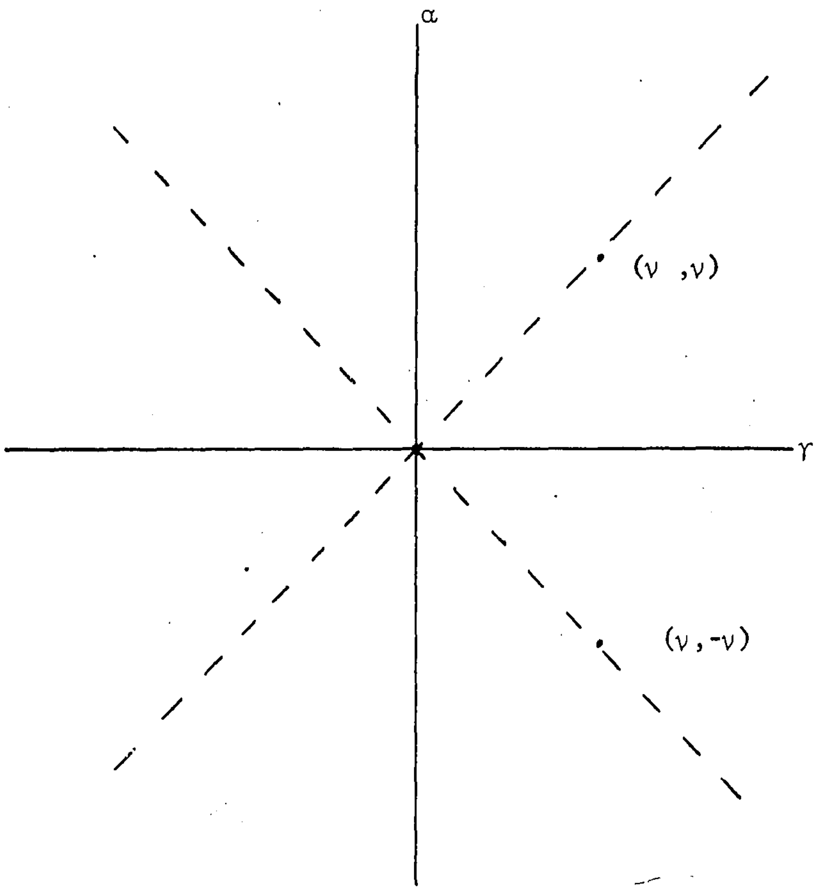


Figure 10. Conjugacy and symmetry lines in the $(\mu_1,\mu_2)$-plane

D. Asymptotic OC for Single Wedges

As pointed out in Section B, Anderson's integral form (Anderson, 1960) for the OC function of a symmetric wedge is given by

$$1-P(L|\mu) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(z-\mu\sqrt{h/s})^2/2} \cdot \frac{e^{2\sqrt{sh}\, z}}{1+e^{2\sqrt{sh}\, z}}\, dz .$$

Utilizing a familiar saddle point argument discussed for example in Abramowitz (1954) and David and Kruskal (1956), we note that the integrand is maximized at $\tilde{z} = \mu\sqrt{h/s}$, which suggests the change of variable $y = z - \tilde{z}$, transforming the integral into

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} \cdot \frac{e^{2\sqrt{sh}\,(y+\tilde{z})}}{1+e^{2\sqrt{sh}\,(y+\tilde{z})}}\, dy$$

$$= \frac{e^{2\sqrt{sh}\,\tilde{z}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{-(1/2)(y-2\sqrt{sh})^2 + 2sh}}{1+e^{2\sqrt{sh}\,(y+\tilde{z})}}\, dy .$$

But the integrand is bounded by zero and its integrable numerator, to which it tends from below as $\tilde{z}$ becomes small. Hence, by Lebesgue's Theorem (Cramér, 1946, page 66), integration with respect to y and taking the limit with respect to $\tilde{z}$ may be interchanged, yielding:

$$\lim_{\tilde{z} \to -\infty} \frac{1-P(L|\mu)}{e^{2\sqrt{sh}\,\tilde{z}}}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(1/2)(y-2\sqrt{sh})^2+2sh} \, dy,$$

or

$$\lim_{\mu \to -\infty} \frac{1-P(L|\mu)}{e^{2h\mu}} = e^{2sh} ,$$

or, for small $\mu$,

$$1-P(L|\mu) \simeq e^{2h(s+\mu)}. \tag{4.16}$$

Relation 4.16 and its analogues for the non-symmetric case follow at once from consideration of terminal likelihood ratios for conjugate pairs. Consider a wedge with intercepts $h_1$ and $h_o$ and slopes $s_1$ and $s_o$; consider also the $g_\mu(\cdot)$ assumed in Section A to exist on the lower boundary. Then using 2.19 and integrating over the lower boundary we find:

$$P(L|\mu) = e^{2h_o(\mu-s_1)} P(L|2s_1-\mu). \tag{4.17}$$

Similarly integration over the upper boundary gives:

$$1-P(L|\mu) = e^{2h_1(\mu-s_o)} [1-P(L|2s_o-\mu)] . \tag{4.18}$$

We also have:

$$\lim_{\mu \to \infty} P(L|2s_1-\mu) = \lim_{\mu \to -\infty} [1-P(L|2s_o-\mu)] = 1 \tag{4.19}$$

Now combining 4.17 and 4.19 we see that

$$P(L|\mu) \simeq e^{2h_o(\mu-s_1)} \tag{4.20}$$

for $\mu$ large, and combining 4.18 and 4.19,

$$P(U|\mu) = 1-P(L|\mu) \simeq e^{2h_1(\mu-s_o)} \tag{4.21}$$

for small $\mu$, which specializes to 4.16.

Somewhat less obvious but also using the terminal likelihood ratios for conjugate pairs are bounds for the OC from which 4.20 and 4.21 could be derived. Letting the parameter of 4.18 be $2s_1-\mu$

$$1-P(L|2s_1-\mu) = e^{2h_1(s_1+\delta-\mu)}[1-P(L|\mu-2\delta)], \tag{4.22}$$

where $\delta = s_1-s_o$. Then using 4.22 in 4.17

$$P(L|\mu) = e^{2h_o(\mu-s_1)}\left\{1-e^{2h_1(s_1+\delta-\mu)}[1-P(L|\mu-2\delta)]\right\}, \tag{4.23}$$

so that,

$$e^{2h_o(\mu-s_1)}[1-e^{2h_1(s_1+\delta-\mu)}] \leq P(L|\mu) \leq e^{2h_o(\mu-s_1)}, \tag{4.24}$$

which bounds $P(L|\mu)$ progressively more sharply as $\mu \to +\infty$. Then the asymptotic form of 4.20 follows for large $\mu$. Iteration of 4.23 provides still sharper bounds; for example letting the parameter in the left hand side of 4.23 be $\mu-2\delta$

$$P(L|\mu-2\delta) = e^{2h_o(\mu-s_1-2\delta)}\left\{1-e^{2h_1(s_1+3\delta-\mu)}[1-P(L|\mu-4\delta)]\right\}. \qquad (4.25)$$

Now using 4.25 in 4.23

$$P(L|\mu) = e^{2h_o(\mu-s_1)}\left\{1-e^{2h_1(s_1+\delta-\mu)} + e^{2h_1(s_1+\delta-\mu)}e^{2h_o(\mu-s_1-2\delta)}\right.$$

$$\left.[1-e^{2h_1(s_1+3\delta-\mu)} + e^{2h_1(s_1+3\delta-\mu)}P(L|\mu-4\delta)]\right\},$$

so that

$$e^{2h_o(\mu-s_1)}\left\{1-e^{2h_1(s_1+\delta-\mu)} + e^{2h_1(s_1+\delta-\mu)}e^{2h_o(\mu-s_1-2\delta)}[1-e^{2h_1(s_1+3\delta-\mu)}]\right\}$$

$$\leq P(L|\mu) \leq e^{2h_o(\mu-s_1)}\left\{1-e^{2h_1(s_1+\delta-\mu)} + e^{2h_1(s_1+\delta-\mu)}e^{2h_o(\mu-s_1-2\delta)}\right\},$$

which bounds $P(L|\mu)$ more sharply than does 4.24 for $\mu > 3\delta + s_1$.

## E. Asymptotic OC for Multiple Wedge Systems

In this section we prove a theorem which gives the asymptotic forms for the OC's in a general multiple wedge system. We begin with the proof of a lemma.

Lemma 4.1: Let the events X . . . X, X . . . XU, and X . . . XL be defined as in Chapter III. Then

$$\lim_{\mu \to \infty} P(X\ .\ .\ .\ XU|X\ .\ .\ .\ X;\ \mu) = 1, \qquad (4.26)$$

$$\lim_{\mu \to -\infty} P(X\ .\ .\ .XL|X\ .\ .\ .\ X;\ \mu) = 1\ . \qquad (4.27)$$

Proof: Let $P(h_o, h_1, \delta, \mu - s_o)$ be the probability that a Wiener process starting from the origin first reaches the upper boundary of a wedge parameterized by $h_o$, $h_1$, $s_o$, and $s_1 = s_o + \delta$ when the drift parameter is $\mu$. Let $g_\mu(\sigma) = P(h_o(\sigma), h_1(\sigma), s_1 - s_o, \mu - s_o)$. We postulate the existence of a conditional density $f_\mu(\sigma)$ on the lower boundary AB of wedge $W_{X \,.\,.\,.\, X}$ in Figure 11 with the property that

$$P(X \,.\,.\,.\, XLU | X \,.\,.\,.\, XL;\mu) = \int_B^A g_\mu(\sigma) f_\mu(\sigma) d\sigma. \tag{4.28}$$

Now by 4.24 we have (except possibly at A)

$$1-e^{2(h_o' - h_1')(\mu - s_1)} \leq 1-e^{2h_o(\sigma)[\mu - s_1]} \leq g_\mu(\sigma), \tag{4.29}$$

so that the integral on the right hand side of 4.28 is bounded below by $1-e^{2(h_o'-h_1')(\mu-s_1)} \int_B^A f_\mu(\sigma)d\sigma = 1-e^{2(h_o'-h_1')(\mu-s_1)}$ which tends to one as $\mu$ becomes large. This completes the proof of the lemma for the case $X \,.\,.\,.\, XXU = X \,.\,.\,.\, XLU$; the argument for the other three cases is similar.

We next apply the lemma to a two-stage system of the general type discussed in Section B of Chapter III. Let $P(XX;\mu \ll 0)$ and $P(XX;\mu \gg 0)$ be the asymptotic forms of the probabilities of event XX for $\mu$ small and $\mu$ large respectively. From lemma 4.1

$$P(LL;\mu \ll 0) = 1, \tag{4.30}$$

and from 2.19 and 3.3

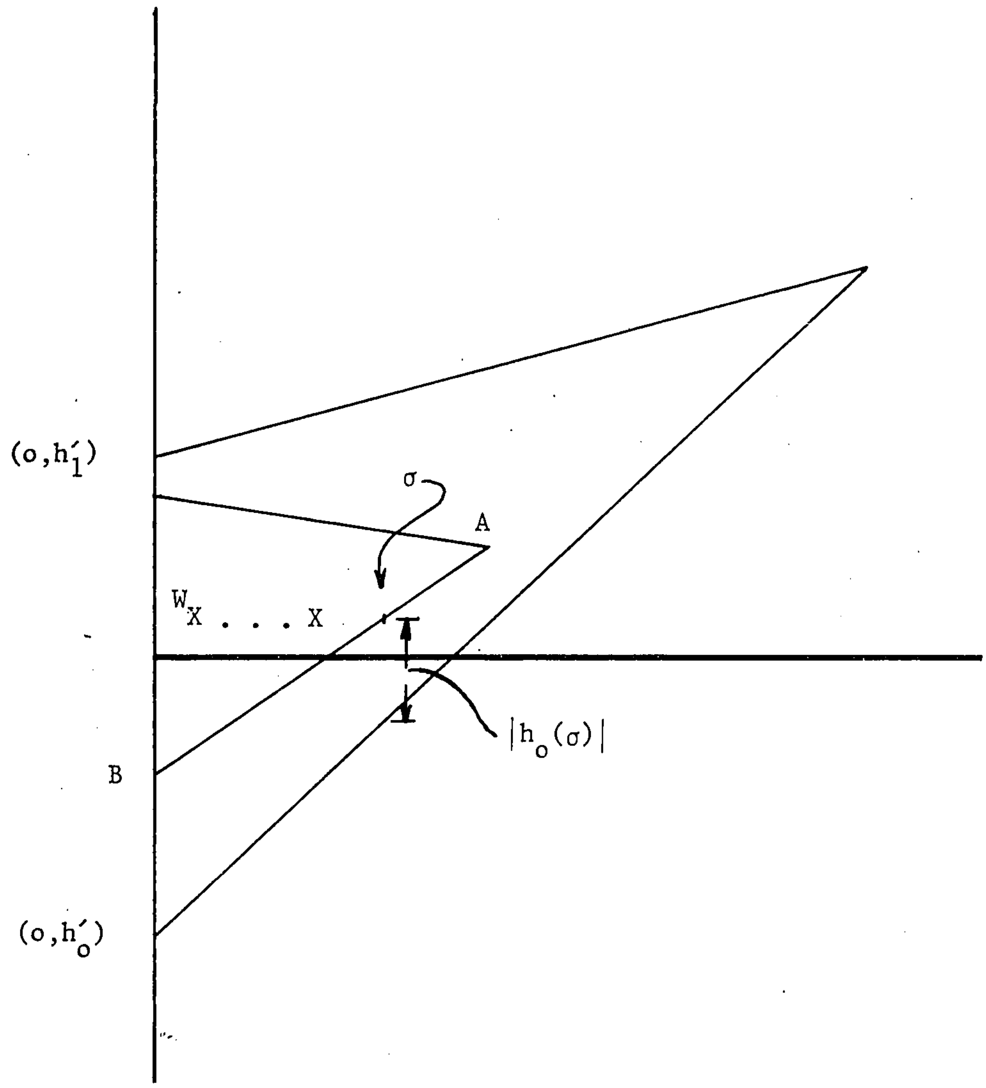


Figure 11. Figure accompanying proof of lemma 4.1

$$P(LL;\mu) = e^{2h_{Lo}(\mu - s_{L1})} P(LL; 2s_{L1} - \mu),$$

so that

$$P(LL;\mu \gg 0) = e^{2h_{Lo}(\mu - s_{L1})}. \tag{4.31}$$

Also,

$$P(UU;\mu \gg 0) = 1, \tag{4.32}$$

and

$$P(UU;\mu) = e^{2h_{U1}(\mu - s_{Uo})} P(UU; 2s_{Uo} - \mu),$$

so that

$$P(UU;\mu \ll 0) = e^{2h_{U1}(\mu - s_{Uo})}. \tag{4.33}$$

Now

$$P(LU;\mu) = P(LU|L;\mu)P(L;\mu),$$

and so by lemma 4.1

$$\begin{aligned} P(LU;\mu \gg 0) &= (1)\ P(L;\mu \gg 0) \\ &= e^{2h_o(\mu - s_1)}, \end{aligned} \tag{4.34}$$

also

$$P(LU;\mu) = e^{2h_{L1}(\mu - s_{Lo})} P(LU; 2s_{Lo} - \mu),$$

so by 4.34

$$\begin{aligned} P(LU;\mu \ll 0) &= e^{2h_{L1}(\mu - s_{Lo})} e^{-2h_o(\mu - 2s_{Lo} - s_1)} \\ &= e^{2\mu(h_{L1} - h_o) - 2(h_{L1}s_{Lo} + h_o s_o - 2h_o s_{Lo})} \end{aligned} \tag{4.35}$$

Similar reasoning gives

$$P(UL;\mu \gg 0) = e^{-2\mu(h_1 - h_{Uo}) - 2(h_{Uo}s_{U1} + h_1 s_o - 2h_1 s_{U1})} \tag{4.36}$$

$$P(UL;\mu \ll 0) = e^{2h_1(\mu - s_o)} \tag{4.37}$$

It is clear on examining these relations that the only steps in a sequence X . . . X which affect the limiting forms of the probabilities are changes from U to L (L to U) for $\mu$ becoming large (small). As will be shown in the theorem the effect of such a step on the asymptotic form is measured by an exponential function of the difference in the height at which such a step begins and the height at which it ends. We next introduce some terminology and notation.

We will use the term "L-inversion" to indicate either of the following:

(a) An L or a sequence of L's beginning a sequence X . . . X,

or

(b) An L or a sequence of L's immediately following a U in a sequence X . . . X.

We will use the term "U-inversion" to indicate either of the following:

(a) A U or a sequence of U's beginning a sequence X . . . X,

or

(b) A U or a sequence of U's immediately following an L in a sequence X . . . X.

We introduce for each sequence X . . . X two index sets of integers. I(X . . . X) = [1,2, . . . r] is the set whose elements count the L-inversions of X . . . X; J(X . . . X) = [1,2, . . . v] is the set whose elements count the U-inversions of X . . . X.

For the $i^{th}$ L-inversion in a sequence, let $h_{i1}$ and $s_{io}$ be the slope and intercept of the last upper boundary crossed before the inversion and let $h_{io}$ and $s_{i1}$ the corresponding parameters for the last lower boundary crossed in the L-inversion. If the first L-inversion occurs at the origin $h_{11} = s_{1o} = 0$. Let $h'_{jo}$ and $s'_{j1}$ be the intercept and slope for the last lower boundary crossed before the $j^{th}$ U-inversion and $h'_{j1}$ and $s'_{jo}$ be the parameters for the last upper boundary crossed in the $j^{th}$ U-inversion. If the first U-inversion occurs at the origin $h'_{1o} = s'_{11} = 0$.

Then define for any sequence X . . . X,

$$\Delta_i = (h_{i1}-h_{io}) \quad \text{for } i \in I(X \ldots X), \tag{4.38}$$

$$\gamma_{i1} = h_{i1}[s_{io} + 2\sum_{k>i}^{r} s_{ko} - 2\sum_{d=i}^{r} s_{d1}],\ i \in I(X \ldots X), \tag{4.39}$$

$$\gamma_{io} = h_{io}[s_{i1} + 2\sum_{k>i}^{r} s_{k1} - 2\sum_{d>i}^{r} s_{do}],\ i \in I(X \ldots X), \tag{4.40}$$

$$\Delta'_j = (h'_{jo}-h'_{j1}),\ j \in J(X \ldots X), \tag{4.41}$$

$$\gamma'_{jo} = h'_{jo}[s'_{j1} + 2\sum_{k>j}^{v} s'_{k1} - 2\sum_{d=j}^{v} s'_{do}],\ j \in J(X \ldots X), \tag{4.42}$$

$$\gamma'_{j1} = h'_{j1}[s'_{jo} + 2\sum_{k>j}^{v} s'_{ko} - 2\sum_{d>j}^{v} s'_{d1}],\ j \in J(X \ldots X). \tag{4.43}$$

Some useful relations involving these parameters are now given.

If the sequence X . . . X is of the form L . . . X then the first L-inversion occurs at the origin and the end of the $k^{th}$ L-inversion is the beginning of the $k^{th}$ U-inversion, so that $h_{ko} = h'_{ko}$, $h_{k1} = h'_{k-1,1}$ $s_{k1} = s'_{k1}$ and $s_{ko} = s'_{k-1,o}$. Similarly if X . . . X is of the form UX . . . S, $h'_{k1} = h_{k1}$ $h'_{ko} = h_{k-1,o}$, $s'_{ko} = s_{ko}$, and $s'_{k1} = s_{k-1,1}$.

If X . . . X is of the form X . . . XU, then the last inversion is a U-inversion and

$$\sum_{i=1}^{r} \Delta_i = - \sum_{j=1}^{v} \Delta_j' - h_{v1}'. \tag{4.44}$$

If the last inversion is an L-inversion, X . . . X is of the form X . . . XL, and

$$\sum_{i=1}^{r} \Delta_i = - \sum_{j=1}^{v} \Delta_j' + h_{ro}. \tag{4.45}$$

If X . . . X is of the form X . . XU,

$$\sum_{i=1}^{r} \gamma_{i1} = \sum_{j=1}^{v} \gamma_{j1}' - s_{vo}'(2 \sum_{j=1}^{v} h_{j1}' - h_{v1}'), \tag{4.46}$$

$$\sum_{i=1}^{r} \gamma_{io} = \sum_{j=1}^{v} \gamma_{jo}' + 2s_{vo}' \sum_{j=1}^{v} h_{jo}'; \tag{4.47}$$

while if X . . . X is of the form X . . . XL,

$$\sum_{i=1}^{r} \gamma_{i1} = \sum_{j=1}^{v} \gamma_{j1}' - 2s_{r1} \sum_{i=1}^{r} h_{i1}, \tag{4.48}$$

$$\sum_{i=1}^{r} \gamma_{io} = \sum_{j=1}^{v} \gamma_{jo}' + s_{r1}(2 \sum_{i=1}^{r} h_{io} - h_{ro}) \tag{4.49}$$

<u>Theorem</u> 4.1: For an m-stage system,

$$P(X \ldots X;\mu \gg 0) = \exp\left[-2\mu \sum_{i=1}^{r} \Delta_i - 2 \sum_{i=1}^{r} \gamma_{i1} - 2 \sum_{i=1}^{r} \gamma_{io}\right], r \neq 0 \tag{4.50}$$

$$= 1, \quad r = 0,$$

$$P(X\ .\ .\ .\ X;\mu \ll 0) = \exp\left[-2\mu \sum_{j=1}^{v} \Delta_j' - 2 \sum_{j=1}^{v} \gamma_{j1}' - 2 \sum_{j=1}^{v} \gamma_{jo}'\right],\ v \neq 0 \quad (4.51)$$

$$= 1,\ v = 0.$$

Proof: Relations 4.30 through 4.37 demonstrate the theorem for $m = 2$. For $m > 2$ the proof is by induction:

$$P(X\ .\ .\ .\ XU;\mu) = P(X\ .\ .\ .\ XU|X\ .\ .\ .\ X;\mu)P(X\ .\ .\ .\ X;\mu),$$

or

$$P(X\ .\ .\ .\ XU;\mu \gg 0) = (1)\ P(X\ .\ .\ .\ X;\mu \gg 0)$$

by lemma 4.1, and since $I(X\ .\ .\ .\ XU) \equiv I(X\ .\ .\ .\ X)$ the validity of 4.50 for $X\ .\ .\ .\ XU$ follows.

We also have

$$P(X\ .\ .\ .\ XU;\mu) = e^{2h_{X\ .\ .\ .\ X1}(\mu - s_{X\ .\ .\ .\ Xo})} P(X\ .\ .\ .\ XU; 2s_{X\ .\ .\ .\ Xo} - \mu). \quad (4.52)$$

Let $\theta = 2s_{X\ .\ .\ .\ Xo} - \mu$ in 4.52, then by 4.50,

$$P(X\ .\ .\ .\ XU;\theta \gg 0) = \exp\left[-2\theta \sum_{i=1}^{r} \Delta_i - 2 \sum_{i=1}^{r} \gamma_{i1} - 2 \sum_{i=1}^{r} \gamma_{io}\right], \quad (4.53)$$

where $i$ ranges over $I(X\ .\ .\ .\ XU) = [1,2,\ .\ .\ .\ r]$.

Now, from 4.52 and 4.53,

$$P(X\ .\ .\ .\ XU;\mu \ll 0) = \exp\Big[-2(2s_{x\ .\ .\ .\ xo} - \mu) \sum_{i=1}^{r} \Delta_i$$

$$-2 \sum_{i=1}^{r} \gamma_{i1} - 2 \sum_{i=1}^{r} \gamma_{io} + 2h_{X\ .\ .\ .\ X1}(\mu - s_{X\ .\ .\ .\ Xo})\Big]. \quad (4.54)$$

Now using 4.44, 4.46, and 4.47

$$\sum_{i=1}^{r} \Delta_i = - \sum_{j=1}^{v} \Delta'_j - h_{X \ldots X1}, \tag{4.55}$$

$$\sum_{i=1}^{r} \gamma_{i1} = \sum_{j=1}^{v} \gamma'_{j1} - s_{X \ldots Xo}\left(2 \sum_{j=1}^{v} h'_{j1} - h_{X \ldots X1}\right), \tag{4.56}$$

$$\sum_{i=1}^{r} \gamma_{io} = \sum_{j=1}^{v} \gamma'_{jo} + 2s_{X \ldots Xo} \sum_{j=1}^{v} h'_{jo}, \tag{4.57}$$

where $j \in J(X \ldots XU) = [1, 2, \ldots v]$. Now using 4.55, 4.56 and 4.57 in 4.54,

$$P(X \ldots XU);\mu \ll 0) = \exp[-2\mu \sum_{j=1}^{v} \Delta'_j - 2 \sum_{j=1}^{v} \gamma'_{j1} - 2 \sum_{j=1}^{v} \gamma'_{jo}]$$

which is 4.51. This completes the proof of the validity of 4.50 and 4.51 for X . . . XU; the proof for X . . . XL is similar.

## F. Comments on the Limitations of Likelihood Ratios for Computing OC Functions

Although we are able to compute OC functions for single symmetric wedges on a grid, and for certain special multiple wedge systems on a grid, we are not able to find complete OC's for the general wedge systems. This is due to the fact that the likelihood ratio equations for conjugate pairs on the upper and lower boundaries provide two equations in three unknowns. Without some further information (such as knowledge of the OC at one point for a symmetric wedge) not much more can be given about the

OC than the bounds and asymptotic behavior. If there is one non-increasing function $f(\cdot)$ satisfying

$$f(\mu) = e^{2h_o(\mu - s_1)} f(2s_1 - \mu), \tag{4.58}$$

$$1 - f(\mu) = e^{2h_1(\mu - s_o)} [1 - f(2s_o - \mu)], \tag{4.59}$$

$$\lim_{\mu \to \infty} f(\mu) = \lim_{\mu \to -\infty} [1 - f(\mu)] = 0, \tag{4.60}$$

$$0 \leq f(\mu) \leq 1, \text{ for all } \mu, \tag{4.61}$$

then there are infinitely many, even with the additional assumption that $f(\cdot)$ is analytic. This is seen by observing that if $f(\cdot)$ is an analytic function satisfying 4.58 through 4.61, then so is

$$g(\mu) = f(\mu) - ke^{h_o(\mu - s_o)^2/2\delta} \cdot e^{-h_1(\mu - s_1)^2/2\delta}$$

for k sufficiently small, where 4.24 is used in verifying 4.61 for $g(\cdot)$.

# V. THE BINOMIAL AND POISSON CASES

## A. Introductory Remarks

In this chapter we consider the OC functions for wedge procedures in the binomial and Poisson cases.

The OC is found for values of the binomial parameter on a grid for a single wedge symmetric about 1/2. It is pointed out that such a procedure may be used to compare two binomial populations. Bounds for the OC on the grid are also found.

It is shown that OC bounds may be computed for a special 3-decision procedure by using linear programming methods.

Methods are given for obtaining approximate wedge OC's for binomial sampling and the Poisson process.

## B. Symmetric Binomial Wedges--Double Dichotomies

Wald (1947, Chapter 6) has shown that a comparison of $p_1$ and $p_2$, the parameters of two binomial populations, is equivalent to testing a hypothesis about $p = p_1(1-p_2)/ p_1[1-p_2)+p_2(1-p_1)]$. If $p_1 = p_2$, $p = 1/2$, while if $p_1 > p_2$, $p < 1/2$, and if $p_1 < p_2$, $p > 1/2$. In those cases where the two losses are symmetric about $p = 1/2$, a reasonable procedure would be to use a test whose OC function is symmetric about $p = 1/2$. We will evaluate the OC on a certain grid for such a procedure based on a symmetric wedge.

Consider a symmetric wedge in the $(m-V_m)$ -plane, where m = number of observations and $V_m$ = (number of successes) - (number of failures) =

$2\,T_m$-m. Let the upper and lower boundaries respectively be $V_m = 2h\text{-}sm$ and $V_m = -2h + sm$. Then $p(L \mid 1/2) = 1/2$ since each path crossing the lower boundary has a mirror image with equal probability crossing the upper boundary. Transforming this wedge into the $(m\text{-}T_m)$-plane gives $T_m = h_1{+}s_o m$ and $T_m = h_o + s_1 m$ as upper and lower boundaries respectively, where $h_1 = -h_o = h$, $s_1 = (s{+}1)/2$, $s_o = (1\text{-}s)/2$ and, further, $s_1{+}s_o = 1/2$.

Since the wedge is symmetric we have exactly $L(1/2) = P(L \mid 1/2) = 1/2$. Now consider a pair $(p_1,p_o)$ with $p_1$ conjugate to 1/2 relative to $s_1$ and $p_o$ conjugate to 1/2 relative to $s_o$; then by 2.13 $c(p_1) \overset{s_1}{=} c(1/2)$[1] $\overset{s_o}{\doteq} c(p_o)$, or $p_1/p_o = (1\text{-}p_o)/(1\text{-}p_1)$; in fact we see that either $p_1 = 1\text{-}p_o$ or trivially, $p_1 = p_o = 1/2$. Then by symmetry

$$L(p_1) = 1\text{-}L(p_o). \tag{5.1}$$

By 2.14

$$L(p_1) \doteq \left[\frac{p_1}{1\text{-}p_1}\right]^{-h} L(1/2), \tag{5.2}$$

and by 2.14 and 5.1

$$L(p_o) \doteq \left[\frac{p_1}{1\text{-}p_1}\right]^{h} [1\text{-}L(1/2)]. \tag{5.3}$$

Now at $p_2$, the conjugate of $p_1$ with respect to $s_o$, we evaluate $L(p_2)$; then at $p_3$, the conjugate of $p_2$ with respect to $s_1$, we find $L(p_3)$. Proceeding in this manner we may compute the OC function on a grid of points which approach the extremes of the interval [0,1]. As an example consider a wedge with $h = 2$, $s_o = 1/4$, $s_1 = 3/4$. We have $L(1/2) = 1/2$ by symmetry and $p_1 = (1\text{-}p_o) \doteq .92$. Then $L(.92) \doteq [.92/.08]^{-2}(1/2) =$

[1]Meaning $p_i$ is conjugate to 1/2 relative to $s_1$.

$.0378 \doteq 1\text{-}L(.08)$.

It is also possible to get bounds for the OC at these grid points. We know that if a sample sequence first crosses the lower boundary at $m = n$

$$\lambda_{n-1} > \left[\frac{p_1}{1-p_1}\right]^{-h} , \tag{5.4}$$

and

$$\lambda_n = 2(1-p_1)\lambda_{n-1} \leq \left[\frac{p_1}{1-p_1}\right]^{-h} . \tag{5.5}$$

where $\lambda_m$ is the likelihood ratio under $p = p_1$ and $p = 1/2$ for the first m observations. Then

$$2(1-p_1)B < \lambda_n \leq B \tag{5.6}$$

where $B = \left[\frac{p_1}{1-p_1}\right]^{-h}$ . Summing 5.6 over all sequences crossing the lower boundary before the upper gives the bounds on $L(p_1)$

$$(1-p_1)B < L(p_1) \leq B/2. \tag{5.7}$$

Repeated application of this procedure provides bounds for L(p) for all p on the grid. For example at $p_2$, the conjugate of $p_1$ relative to $s_o$, we have

$$\frac{p_1(1-p_1)}{p_2} AB < \frac{p_1}{p_2} AL(p_1) < L(p_2) \leq AL(p_1) \leq AB/2,$$

where $A = \left[\frac{p_1(1-p_2)}{p_2(1-p_1)}\right]^{h}$ .

## C. OC Bounds for a Special 3-Decision Procedure

Consider a 3-decision binomial procedure in which all slopes equal $s$. Let $p_1$ and $p_o$ be a conjugate pair relative to $s$. Define the following OC's: $L_{11}(p) = P(UU;p)$, $L_{1o}(p) = P(UL;p)$, $L_{o1}(p) = P(LU;p)$ and $L_{oo}(p) = P(LL;p)$. Then we have the following relations from the likelihood ratio considerations of Chapter II.

$$L_{11}(p_1) \geq K^{h_{U1}} L_{11}(p_o)$$

$$L_{1o}(p_1) \leq K^{h_{U0}} L_{1o}(p_o)$$

$$L_{o1}(p_1) \geq K^{h_{L1}} L_{o1}(p_o)$$

$$1-L_{11}(p_1)-L_{1o}(p_1)-L_{o1}(p_1) \leq K^{h_{L0}}[1-L_{11}(p_o)-L_{1o}(p_o)-L_{o1}(p_o)]$$

$$L_{11}(p_1)+L_{1o}(p_1) \geq K^{h_1}[L_{11}(p_o)+L_{1o}(p_o)]$$

$$1-L_{11}(p_1)-L_{1o}(p_1) \leq K^{h_o}[1-L_{11}(p_o)-L_{1o}\ p_o)]$$

where the inequalities stem from the fact that there may be excess over the boundary at termination. The factor $K = p_1(1-p_o)/p_o(1-p_1)$ from 2.14. We can get six corresponding relations by considering the maximum excess over a boundary at termination. The boundaries whose intercepts are $h_1$, $h_{U1}$ and $h_{L1}$ are crossed by a sequence whose terminal observation is a success; hence the likelihood ratios for sequences terminating near these boundaries are at most $(p_1/p_o)K^{h_{X1}}$. Similarly sequences terminating near the boundaries with intercepts $h_{Xo}$ have likelihood ratios no greater than $(p_o/p_1)K^{h_{Xo}}$. We then have

$$L_{11}(p_1) \leq (p_1/p_o)K^{h_{U1}} L_{11}(p_o)$$

$$L_{1o}(p_1) \geq (p_o/p_1)K^{h_{Uo}} L_{1o}(p_o)$$

$$L_{o1}(p_1) \leq (p_1/p_o)K^{h_{L1}} L_{o1}(p_o)$$

$$1-L_{11}(p_1)-L_{1o}(p_1)-L_{\ 1}(p_1) \leq (p_o/p_1)K^{h_{Lo}}[1-L_{11}(p_o)-L_{1o}(p_o)-L_{o1}(p_o)]$$

$$L_{11}(p_1)+L_{1o}(p_1) \leq (p_1/p_o)K^{h_1}[L_{11}(p_o)+L_{1o}(p_o)]$$

$$1-L_{11}(p_1)-L_{1o}(p_o) \geq (p_o/p_1)K^{h_o}[1-L_{11}(p_o)-L_{1o}(p_o)]$$

The set of 12 inequalities in 6 unknowns will determine a convex set in 6-dimensional space in which the point $(L_{11}(p_1), L_{1o}(p_1), \ldots, L_{o1}(p_o))$ will lie. It would be possible, though laborious, to determine the co-ordinates in 6-space of this simplex. A more practical method to find bounds for the OC's would be to use the simplex method of linear programming to maximize and minimize the functions

$$f_1(L_{11}(p_1), \ldots L_{o1}(p_o)) = L_{11}(p_1),\ f_2(L_{11}(p_1) \ldots L_{o1}(p_o)) = L_{11}(p_o)\ ,$$

etc. The linear programming method was tried on a 2-decision procedure with single slope (an SPRT) and found to give the correct OC bounds.

The linear programming technique is useful only in obtaining numerical values for the bounds in specific cases. Explicit forms for the bounds can be derived, as will be demonstrated by the following example.

Consider a 3-decision binomial procedure characterized by slope s and intercepts $h_{U1}$, $h_1 = h_{L1}$, $h_o = h_{Uo}$, and $h_{Lo}$. The two sets of inequalities on pages 61 and 62 are of the same basic form, differing only in the senses of the inequalities and the constants involving K. Let $K_4$ represent either $K^{h_{U1}}$ or $(p_1/p_o)K^{h_{U1}}$; $K_3$ either $K^{h_1}$ or $(p_1/p_o)K^{h_1}$; $K_2$ either $K^{h_o}$ or $(p_o/p_1)K^{h_o}$; and $K_1$ either $K^{h_{Lo}}$ or $(p_o/p_1)K^{h_{Lo}}$. Then, replacing the inequalities with equalities, both sets of 6 relations become

$$\begin{bmatrix} -K_4 & 0 & 0 & 1 & 0 & 0 \\ 0 & -K_3 & 0 & 0 & 1 & 0 \\ 0 & 0 & -K_2 & 0 & 0 & 1 \\ K_1 & K_1 & K_1 & -1 & -1 & -1 \\ K_3 & 0 & K_3 & 1 & 0 & 1 \\ K_2 & 0 & K_2 & -1 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} L_{11}(p_o) \\ L_{o1}(p_o) \\ L_{1o}(p_o) \\ L_{11}(p_1) \\ L_{o1}(p_1) \\ L_{1o}(p_1) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ K_1-1 \\ 0 \\ K_2-1 \end{bmatrix} .$$

This system may be solved for the 6 OC's in terms of the K's:

$$\begin{bmatrix} L_{11}(p_1) \\ L_{o1}(p_o) \\ L_{1o}(p_o) \\ L_{11}(p_1) \\ L_{o1}(p_1) \\ L_{1o}(p_1) \end{bmatrix} = \begin{bmatrix} (1-K_2)/(K_4-K_2) \\ (K_2-K_1)(K_3-1)/(K_3-K_1)(K_3-K_2) \\ (1-K_2)(K_4-K_3)/(K_4-K_2)(K_3-K_2) \\ K_4(1-K_2)/(K_4-K_2) \\ K_3(K_2-K_1)(K_3-1)/(K_3-K_1)(K_3-K_2) \\ K_2(1-K_2)(K_4-K_3)/(K_4-K_2)(K_3-K_2) \end{bmatrix} .$$

Imposing the restrictions $K_4 > K_3 > 1 > K_2 > K_1$ then leads to the following bounds for the OC's:

$$\overline{L_{11}(p_o)} = \frac{1-B_2}{A_4-B_2} ,$$

$$\underline{L_{11}(p_o)} = \frac{1-A_2}{B_4-A_2} ,$$

$$\overline{L_{o1}(p_o)} = \frac{(A_2-B_1)(B_3-1)}{(B_3-B_1)(B_3-A_2)} , \text{ if } \frac{(B_3-1)(A_3-1)}{(1-A_2)(1-B_1)} \leq 1,$$

$$= \frac{(A_2-B_1)(A_3-1)}{(A_3-B_1)(A_3-A_2)}, \quad \text{otherwise},$$

$$\underline{L_{o1}(p_o)} = \frac{(B_2-A_1)(A_3-1)}{(A_3-A_1)(A_3-B_2)}, \text{ if } \frac{(B_3-1)(A_3-1)}{(1-A_1)(1-B_2)} \leq 1,$$

$$= \frac{(B_2-A_1)(B_3-1)}{(B_3-A_1)(B_3-B_2)} , \text{ otherwise},$$

$$\overline{L_{1o}(p_o)} = \frac{(1-B_2)(B_4-A_3)}{(B_4-B_2)(A_3-B_2)} , \text{ if } \frac{(B_4-1)(A_3-1)}{(1-A_2)(1-B_2)} \geq 1,$$

$$= \frac{(1-A_2)(B_4-A_3)}{(B_4-A_2)(A_3-A_2)} , \text{ otherwise},$$

$$\underline{L_{1o}(p_o)} = \frac{(1-A_2)(A_4-B_3)}{(A_4-A_2)(B_3-A_2)} , \text{ if } \frac{(A_4-1)(B_3-1)}{(1-A_2)(1-B_2)} \geq 1,$$

$$= \frac{(1-B_2)(A_4-B_3)}{(A_4-B_2)(B_3-B_2)} , \text{ otherwise},$$

$$\overline{L_{11}(p_1)} = \frac{A_4(1-B_2)}{A_4-B_2} ,$$

$$\underline{L_{11}(p_1)} = \frac{B_4(1-A_2)}{B_4-A_2} ,$$

$$\overline{L_{o1}(p_1)} = \frac{B_3(A_2-B_1)(B_3-1)}{(B_3-B_1)(B_3-A_2)} , \text{ if } \frac{A_3B_3}{B_1A_2} < \frac{A_3+B_3-1}{B_1+A_2-1}$$

$$= \frac{A_3(B_2-A_1)(B_3-1)}{(A_3-B_1)(A_3-A_2)} , \text{ otherwise,}$$

$$\underline{L_{o1}(p_1)} = \frac{B_3(B_2-A_1)(B_3-1)}{(B_3-A_1)(B_3-B_2)} , \text{ if } \frac{A_3B_3}{A_1B_2} > \frac{A_3+B_3-1}{A_1+B_2-1} ,$$

$$= \frac{A_3(B_2-A_1)(A_3-1)}{A_3-A_1)(A_3-B_2)} , \text{ otherwise,}$$

$$\overline{L_{1o}(p_1)} = \frac{B_2(1-B_2)(B_4-A_3)}{(B_4-B_2)(A_3-B_2)} , \text{ if } \frac{A_3B_4}{A_2B_2} > \frac{A_3+B_4-1}{A_2+B_2-1}$$

$$= \frac{A_2(1-B_2)(B_4-A_3)}{(B_4-A_2)(A_3-A_2)} , \text{ otherwise,}$$

$$\underline{L_{1o}(p_1)} = \frac{B_2(1-B_2)(A_4-B_3)}{(A_4-B_2)(B_3-B_2} , \text{ if } \frac{A_4B_3}{A_2B_2} < \frac{A_4+B_3-1}{A_2+B_2-1} ,$$

$$= \frac{A_2(1-A_2)(A_4-B_3)}{(A_4-A_2)(B_3-A_2)} , \text{ otherwise,}$$

Where an underline indicates a lower bound and an overline an upper bound for the OC. Also, $A_1 = K^{h_{Lo}}$, $B_1 = (p_o/p_1)A_1$, $A_2 = K^{h_o}$, $B_2 = (p_o/p_1)A_2$, $A_3 = K^{h_1}$, $B_3 = (p_1/p_o)A_3$, $A_4 = K^{h_{U1}}$, and $B_4 = (p_1/p_o)A_4$.

D. Polynomial OC Forms--Binomial Case

The OC for a binomial wedge procedure can be found exactly as the sum of probabilities of sequences leading to acceptance of $H_1$. Thus the OC will take the form of a polynomial in p,

$$f(p) = \sum_{n=0}^{N} a_n p^n \tag{5.8}$$

where N is the maximum sample size required for the procedure. This enumeration of sequences is tedious, however. Two methods will be given for approximating the $a_n$ in 5.8.

We have

$$f(p_1) \doteq K_1^{h_o+k_o} f(p) \tag{5.9}$$

$$1-f(p_o) \doteq K_o^{h_1+k_1}[1-f(p)] \tag{5.10}$$

where $p_1$ and $p_o$ are the conjugates of p relative to $s_1$ and $s_o$ respectively, $K_1 = [\frac{p_1(1-p)}{p(1-p_1)}]$ , $K_o = [\frac{p_o(1-p)}{p(1-p_o)}]$ and $k_o$ and $k_1$ are representative excess over the boundaries. Also since f(0) = 1 and f(1) = 0 we have that

$$a_o = 1, \sum_{n=0}^{N} a_n = 0. \tag{5.11}$$

We can get in addition to 5.11, N-1 independent relations involving the unknown coefficients $a_n$;[N/2] from one of 5.9 or 5.10 and the remaining N-[N/2] from the other. These relations can be gotten by successively substituting pairs of conjugate points into 5.9 and 5.10. An alternate method is to differentiate 5.9 and 5.10 with respect to p and substitute

$p = s_o$ or $p = s_1$. Once these N-1 additional relations have been gotten, approximations $\tilde{a}_n$ are found; these may then be rounded to the nearest integer for final approximations $\hat{a}_n$.

Both the method of substitution and differentiation have been tried on sample cases with good results; in both cases $\hat{a}_n = a_n$ for all n. The method of substitution is simpler than successive differentiation; indeed, it is programmable for electronic calculation and seems more practical than path counting.

The polynomial approximation methods can be used with relations 3.11 and 3.12 for multistage procedures; however, application to sample cases have not yielded results as good as in the 2-decision case.

E. Approximate OC Calculations for the Poisson Process

Consider a Poisson process $[X(t);t \geq 0]$ with parameter $\theta$. For any $t_1$ and $t_2$ ($t_1 < t_2$) the probability of exactly n occurrences in the interval from $t_1$ to $t_2$ is

$$P[X(t_2)-X(t_1)=n] = e^{-\theta(t_2-t_1)}[\theta(t_2-t_1)]^n / n! \qquad (5.12)$$

Now consider a wedge with lower and upper boundaries respectively $X(t) = h_o+s_1t$, and $X(t) = h_1+s_ot$. Let N be the maximum number of occurrences observable at termination on the lower boundary. Since X(t) is restricted to increase in unit steps, there are a finite number of points where the process can terminate on the lower boundary; let these points be $(t_o, 0)$, $(t_1, 1)$, . . . . $(t_N, N)$. Then the OC function $f(\theta) = P(L;\theta)$ is of the form

$$f(\theta) = \sum_{n=0}^{N} a_n e^{-\theta t_n} (\theta t_n)^n . \tag{5.13}$$

we also have, from 2.16, for $(\theta_o, \theta_1)$ conjugate with respect to $s_1$

$$f(\theta_1) = (\theta_1/\theta_o)^{h_o} f(\theta_o). \tag{5.14}$$

Relation 2.16 applies as well, approximately, for the upper boundary; hence, for $(\theta_o, \theta_1)$ conjugate with respect to $s_o$,

$$1-f(\theta_1) \doteq (\theta_1/\theta_o)^{h_1}[1-f(\theta_o)].$$

Also, $s_o = f(0) = 1$.

$$a_o = f(0) = 1, \tag{5.15}$$

Substitution of pairs of conjugate points in 5.14 and 5.15 will now yield N relations which provide approximations to the $a_n$.

# VI. A CHARACTERIZATION THEOREM

## A. General Remarks

In this chapter we present a theorem which shows that among a certain class of multiple decision wedge procedures for the Wiener process the SPRT has minimum asymptotic risk in a certain sense.

## B. A Theorem

Consider the class $\Omega$ of all k-decision wedge procedures for choosing one of the hypotheses $H_1, H_2, \ldots H_k$ concerning the drift $\mu$ of a Wiener process with the property that

$$\lim_{\mu \to \infty} \mu E[T|\mu] \leq A \tag{6.1}$$

$$\lim_{\mu \to -\infty} |\mu| E(T|\mu] \leq A, \tag{6.2}$$

where $A > 0$ and $E[T|\mu]$ is the expected time to termination of the procedure.

<u>Lemma</u> 6.1: Conditions 6.1 and 6.2 respectively imply that

$$h_{U \ldots U1} \leq A \tag{6.3}$$

and

$$-h_{L \ldots Lo} \leq A \quad , \tag{6.4}$$

where $h_{U \ldots U1}$ is the intercept of the upper boundary of the terminal wedge reached by a path all of whose steps are U's; crossing of this boundary leads to acceptance of $H_k : \theta = \theta_k$ $(\theta_k > \theta_{k-1} > \ldots > \theta_1)$.

Similarly $h_{L\ldots Lo}$ is the intercept of the terminal boundary leading to acceptance of $H_1:\theta = \theta_1$.

Proof: Let $Z = X(T)$ be the ordinate of the path at termination. Then

$$E[Z-s_{U\ldots Uo}T|\mu] = h_{U\ldots U}P(U\ldots U;\mu)$$

$$+ E[Z-s_{U\ldots Uo}T|\overline{U\ldots U};\mu]\ P(\overline{U\ldots U};\mu), \tag{6.5}$$

where $\overline{X\ldots X}$ is the complimentary event to $X\ldots X$. Dvoretzky, Kiefer, and Wolfowitz (1953) show that Wald's fundamental identity holds for the Wiener process so that

$$E[Z-s_{U\ldots Uo}T|\mu] = E[T|\mu](\mu-s_{U\ldots Uo}),$$

so that, from 6.5,

$$E[T|\mu] = \frac{h_{U\ldots U}P(U\ldots U;\mu) + E[Z-s_{U\ldots Uo}T|\overline{U\ldots U};\mu]\ P(\overline{U\ldots U};\mu)}{\mu-s_{U\ldots Uo}},$$

and

$$\lim_{\mu\to\infty} \mu E[T|\mu] = h_{U\ldots U1}, \tag{6.6}$$

since $\lim_{\mu\to\infty} P(U\ldots U;\mu) = 1$. Now 6.6 and 6.1 together imply 6.3.

Since the wedges are successively nested

$$h_1 \le h_{U1} \le h_{UU1} \le \cdots \le h_{U\ldots U1} \le A \tag{6.7}$$

and

$$h_o \geq h_{Lo} \geq h_{LLo} \geq \cdots \geq h_{L \cdots Lo} \geq -A \tag{6.8}$$

<u>Theorem</u> 6.1: Let $P(\bar{H}_i;\mu)$ be the probability of rejecting $H_i$ given $\mu$. Then, among the class $\Omega$ of k-decision procedures described above. The Wald SPRT has the property that the product of extreme asymptotic risks $P(\bar{H}_k;\mu \gg 0) \cdot P(\bar{H}_1;\mu \ll 0)$ is minimum.

Proof: Let C be any member of the class $\Omega$. Then $P_C(\bar{H}_k;\mu) \geq P_C(LU \ldots U;\mu)$, and $P_C(\bar{H}_1;\mu) \geq P_C(UL \ldots L;\mu)$. But by Theorem 4.1 $P_C(LU \ldots U;\mu \gg 0) = e^{2\mu h_o - 2h_o s_1}$, and $P_C(UL \ldots L;\mu \ll 0) = e^{2\mu h_1 - 2h_1 s_o}$, so that

$$P_C(H_k;\mu \gg 0)P_C(H_1;-\mu \gg 0) \geq e^{-2\mu(h_1-h_o)-2(h_o s_1+h_1 s_o)}. \tag{6.9}$$

To minimize the asymptotic risk then we need to have $h_1$-$h_o$ as large as possible, but from lemma 6.1 we see that $h_1 \leq A$ and $h_o \geq -A$, hence the risk is minimized by $h_1 = A = -h_o$ and relations 6.7 and 6.8 then imply that all h's featured in 6.7 (6.8) are A(-A).

Now given that $h_1 = A = -h_o$ in 6.9, we find that the asymptotic risk is minimized when $s_1$-$s_o = 0$, i.e. when both slopes for the initial wedge are the same; but, since the wedges must be successively nested, all slopes are the same and we have the SPRT with intercepts A and -A and slope s. There may be some lower boundaries with intercepts $h_{X \ldots Xo} < -A$ and some upper boundaries with intercepts $h_{X \ldots X1} >$ A, but these boundaries are superfluous since the procedure terminates upon crossing one of the boundaries with intercepts A or -A.

# VII. LITERATURE CITED

Abramowitz, Milton. 1954. On the practical evaluation of integrals. Society of Industrial and Applied Mathematics Journal 2: 20-35.

Anderson, T. W. 1960. A modification of the sequential probability ratio test to reduce sample size. Annals of Mathematical Statistics 31: 165-197.

Armitage, P. 1947. Some sequential tests of Student's hypothesis. Royal Statistical Society Journal Supplement 9: 250-263.

Armitage, P. 1957. Restricted sequential procedures. Biometrika 44: 9-26.

Armitage, P. 1960. Sequential medical trials. Oxford, England, Blackwell Scientific Publications, Ltd.

Baker, A. G. 1950. Properties of some tests in sequential analysis. Biometrika 37: 334-346.

Barnard, G. A. 1946. Sequential tests in industrial statistics. Royal Statistical Society Journal Supplement 8: 1-21.

Bartlett, M. S. 1946. Discussion on Mr. Barnard's paper. Royal Statistical Society Journal Supplement 8: 22-23.

Blackwell, David and Girshick, M. A. 1954. Theory of games and statistical decisions. New York, N. Y., John Wiley and Sons, Inc.

Blasbalg, H. 1957. Transformations of the fundamental relationships in sequential analysis. Annals of Mathematical Statistics 28: 1024-1028.

Cramér, Harald. 1946. Mathematical methods of statistics. Princeton, New Jersey, Princeton University Press.

David, H. T. 1952. Approximate OC functions for the first Rushton approximations to the WAGR. Unpublished memorandum SRC-20205D12. Chicago, Illinois, Statistics Research Center, University of Chicago.

David, H. T. 1964. Industrial statistics: sampling inspection. Mimeographed lecture notes. Ames, Iowa, Statistical Laboratory, Iowa State University of Science and Technology.

David, H. T. and Kruskal, W. H. 1956. The WAGR sequential t-test reaches a decision with probability one. Annals of Mathemttical Statistics 27: 797-805.

Davies, Owen L. 1954. The design and analysis of industrial experiments. New York, New York, Hafner Publishing Co.

DeBoer, J. 1953. Sequential tests with three possible decisions for testing an unknown probability. Applied Scientific Research Series B, 3: 249-259.

DeGroot, M. H. 1961. The essential completeness of generalized sequential probability ratio tests. Annals of Mathematical Statistics 32: 602-605.

Donnelly, T. G. 1957. A family of truncated sequential tests. Unpublished Ph.D. thesis. Chapel Hill, N. C., Library, University of North Carolina.

Dvoretzky, J., Kiefer, J., and Wolfowitz, J. 1953. Sequential decision problems for processes with continuous time parameter [I.] Testing hypotheses. Annals of Mathematical Statistics 24: 254-264.

Fukushima, Kozo. 1961. A comparison of sequential tests for the Poisson parameter. University of North Carolina Institute of Statistics Mimeographed Series No. 307.

Girshick, M. A. 1946. Contributions to the theory of sequential analysis II, III. Annals of Mathematical Statistics 17: 282-298.

Hall, W. J. 1961. Sequential analysis. Unpublished lecture notes. Chapel Hill, North Carolina, University of North Carolina. Original not available for examination; cited in Fukushima, Kozo. 1961.

Johnson, N. L. 1959. A proof of Wald's theorem on cumulative sums. Annals of Mathematical Statistics 30: 1245-1247.

Kiefer, J. and Weiss, Lionel. 1957. Some properties of generalized sequential probability ratio tests. Annals of Mathematical Statistics 28: 57-75.

Koopman, B. O. 1936. On distributions admitting a sufficient statistic. American Mathematical Society Transactions 39: 399-409.

Lechner, J. A. 1964. Optimality and the OC curve for the Wald SPRT. American Statistical Association Journal 59: 464-468.

Lechner, J. A. and Ginsburg, H. 1963. Sequential 3-way classification of lot quality. National Symposium on Reliability and Quality Control Proceedings 9: 257-269.

Lehmann, E. L. 1959. Testing statistical hypotheses. New York, New York, John Wiley and Sons, Inc.

Mengido, Rudolfo Miguel. 1963. Random walk methods in sequential analysis. Unpublished M.S. thesis. Ames, Iowa, Library, Iowa State University of Science and Technology.

Paulson, Edward. 1963. A sequential decision procedure for choosing one of k hypotheses concerning the unknown mean of a normal distribution. Annals of Mathematical Statistics 34: 549-554.

Pitman, E. J. G. 1936. Sufficient statistics and intrinsic accuracy. Cambridge Philosophical Society Proceedings 32: 567-579.

Savage, Leonard J. 1957. When different pairs of hypotheses have the same family of likelihood ratio test regions. Annals of Mathematical Statistics 28: 1028-1032.

Schneiderman, M. A. and Armitage, P. 1962. A family of closed sequential procedures. Biometrika 49: 41-56.

Sobel, Milton. 1953. An essentially complete class of decision functions for certain standard sequential problems. Annals of Mathematical Statistics 24: 319-337.

Sobel, Milton and Wald, A. 1949. A sequential decision procedure for choosing one of three hypotheses concerning the unknown mean of a normal distribution. Annals of Mathematical Statistics 20: 502-522.

Tweedie, M. C. K. 1946. Discussion on Mr. Barnard's paper. Royal Statistical Society Journal Supplement 8: 23-24.

Wald, Abraham. 1947. Sequential analysis. New York, New York, John Wiley and Sons, Inc.

Wald, Abraham and Wolfowitz, J. 1948. Optimum character of the sequential probability ratio test. Annals of Mathematical Statistics 19: 326-339.

Wald, Abraham and Wolfowitz, J. 1950. Bayes solutions of sequential decision problems. Annals of Mathematical Statistics 21: 82-99.

Weiss, Lionel. 1953. Testing one simple hypothesis against another. Annals of Mathematical Statistics 24: 273-281.

Wolfowitz, J. 1947. The efficiency of sequential estimates and Wald's equation for sequential processes. Annals of Mathematical Statistics 18: 228-229.

# VIII. ACKNOWLEDGMENTS

The author wishes to express deepest gratitude and sincere appreciation for the guidance, excouragement, and inspiration of Professor H. T. David, without whom this thesis would not likely have been written.

Thanks are also due Dr. T. A. Bancroft and Professor J. K. Walkup for their kindness and interest during the author's stay at Iowa State University.

Acknowledgement is made for partial support of this research by the National Science Foundation on NSF Grant GP-1801.

Finally, the author acknowledges the invaluable moral support lent by his wife, Nancy, who bore up so nobly under the strain.