LIKELIHOOD RATIO COMPUTATIONS OF OPERATING CHARACTERISTICS[1]

By William D. Lawing and H. T. David

Research Triangle Institute and Iowa State University

1. Summary and introduction. The point of this paper is that likelihood ratios, especially in conjunction with suitable symmetry, yield at least partial operating characteristic (OC) information for procedures other than the sequential probability ratio test (SPRT). Unpublished work of W. J. Hall ([12], [13]) apparently contains related observations to this effect. We find it useful to discuss OC computations in terms of mutually "conjugate" parametric values, which are considered at some length in Section 2; this term appears first in [17][2], and the concept also elsewhere ([11], [4], [6]), all in the context of the SPRT. In Section 3 the ideas are applied to the Wiener process, for the sorts of two-decision "wedge" procedures discussed in [15], [8][3], [1][3], and [12], and, in Section 4, to modifications of the sorts of three-decision procedures discussed in [2], [20], and [3]. Section 5 contains certain special absorption probabilities for Brownian motion; Section 6 is devoted to the binomial case.

2. Straight-line boundary segments and mutually conjugate parametric values. Consider a one-parameter family of densities $f(x; \theta) = \exp[V(x) + a(\theta)T(x) + b(\theta)]$, and define $T_m \equiv \sum_{j=1}^{m} T(x_j)$, $\lambda_m = \prod_{j=1}^{m} f(x_j; \theta_1)/f(x_j; \theta_0)$, and $c_r(\theta) = ra(\theta) + b(\theta)$; also call *mutually conjugate with respect to* r any θ_0 and θ_1 such that $c_r(\theta_0) = c_r(\theta_1)$. Then, for any sample $(x_1, \cdots, x_m)$ with

$$T_m \doteq h + rm, \tag{1}$$

$\lambda_m \doteq \exp[h(a(\theta_1) - a(\theta_0))]$. Hence, if E is any event, or any sum of events, specifying any sorts of conditions on $(x_1, \cdots, x_m)$ that include the condition $T_m \doteq h + rm$, then

$$\Pr\{E \mid \theta_1\} \doteq \exp[h(a(\theta_1) - a(\theta_0))] \Pr\{E \mid \theta_0\}. \tag{2}$$

The question of the existence of mutually conjugate parametric values has been considered in [5] and [19]; one may note in this connection that if, in an obvious notation, some open θ-interval I is such that (i) $0 < V_\theta(T) < +\infty$ and $E_\theta[T]$ is differentiable and monotone increasing on I, and (ii) $E_{\hat\theta}[T] = r$ for some $\hat\theta \varepsilon I$, then, at least locally (about $\hat\theta$), θ-values will be paired uniquely into mutually conjugate pairs. This because, borrowing from the theory of the Cramér-Rao inequality, $c_r'(\theta)$ can then be written in the form $E_\theta'(T)V_\theta^{-1}(T)[r - E_\theta(T)]$, showing a local maximum for $c_r(\theta)$ at $\theta = \hat\theta$. At

Received 14 January 1966; revised 28 June 1966.

[1] Research supported in part by NSF Research Grant GP-1801.

[2] We owe this reference to J. A. Lechner.

[3] We owe these references to T. W. Anderson.

any rate, in the binomial and normal (as well as other) cases, this pairing is not merely local: *all* p-values in $(0, 1)$ are paired into pairs (p_0, p_1) in accordance with

$$p_0^{r}(1 - p_0)^{1-r} = p_1^{r}(1 - p_1)^{1-r}, \tag{3}$$

and all μ-values in $(-\infty, +\infty)$ into pairs (μ_0, μ_1) in accordance with

$$\mu_0 + \mu_1 = 2r. \tag{4}$$

Relation (2) reduces, in these two cases, respectively to

$$\Pr\{E \mid p_1\} \doteq [p_1(1 - p_0)/p_0(1 - p_1)]^h \Pr\{E \mid p_0\} \tag{5}$$

and

$$\Pr\{E \mid \mu_1\} \doteq \exp[h(\mu_1 - \mu_0)] \Pr\{E \mid \mu_0\}, \tag{6}$$

relation (6) written for $\sigma = 1$. T_m is $\sum_{i=1}^{m} x_i$ in both cases, and is usually denoted by d_m in the first.

As is indicated in [10], and also made use of in [17] in the context of the SPRT, the above remarks for the normal case apply as well to the Wiener process, and (6) in fact then holds with strict equality; in other words, consider a line $b(t) = h + rt$, for example with $h < 0$; also let $X(\cdot\,; \mu)$ be a separable Wiener process with unit dispersion parameter and drift parameter μ; finally let $K(s)$ be a condition on the behavior of $X(\cdot\,; \mu)$ on the interval $(0, s)$. Then, for sufficiently regular $K(s)$, if one defines

(7) $E(t)$: for some s in $(0, t)$, $K(s)$ obtains and $X(\cdot\,; \mu)$ crosses $b(\cdot)$ at s,

then, for μ_0 and μ_1 satisfying (4), one has

$$\Pr\{E(t) \mid \mu_1\} = \exp[h(\mu_1 - \mu_0)] \Pr\{E(t) \mid \mu_0\}. \tag{8}$$

Note that the argument $t = +\infty$ is admitted in (8), since $\Pr\{E(t) \mid \mu\}$ is non-decreasing and hence converges.

A typical condition $K(s)$ required in our applications is as follows: Consider three continuous functions $b_1(\cdot)$, $b_2(\cdot)$ and $b_3(\cdot)$ satisfying $b_1(0) \geqq b_3(0) > 0 > b_2(0) \geqq b(0)$, $b_1(\cdot) \geqq b_3(\cdot) > b_2(\cdot) \geqq b(\cdot)$ on $(0, \tau)$, $\tau > 0$, and $b_1(\cdot) > b(\cdot)$ on $(0, t)$. These serve to define

$K(s)$: for some u in $(0, \min(\tau, s))$, (i) $X(\cdot\,; \mu)$ crosses $b_3(\cdot)$ at u, (ii) for all u' in $(0, u)$, neither $X(u'; \mu) > b_3(u')$ nor $X(u'; \mu) < b_2(u')$, (iii) for no s' in (u, s), $X(s'; \mu) > b_1(s')$,

so that, in accordance with its earlier definition (7), $E(t)$ is the crossing of b_3 before b_2 somewhere on $(0, \tau)$, followed by the crossing of b before b_1 somewhere on $(0, t)$. For this $E(t)$, relation (8) may be verified by passing to the limit of discrete-time approximations. The above-mentioned regularity of the general $K(s)$ is intended to insure the validity of this operation.

The argument extends without trouble to the k-dimensional case. Thus, let $X(\cdot\,; \mu) \equiv (X_1(\cdot\,; \nu), \cdots, X_k(\cdot\,; \theta))$ be a vector process whose components

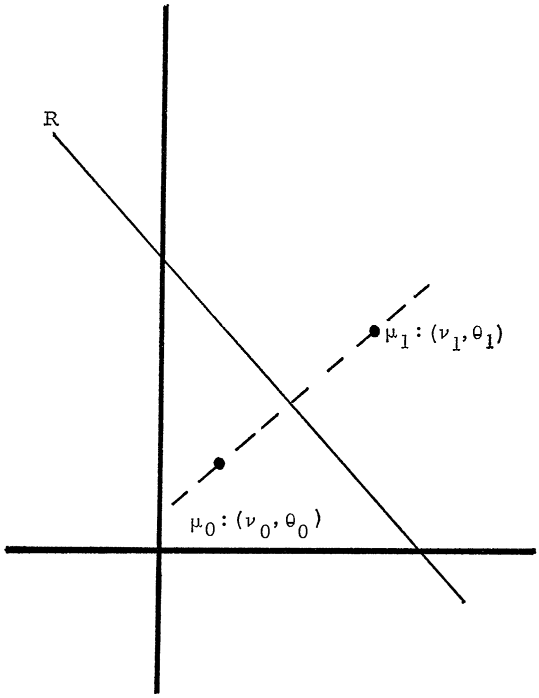

FIG. 1. Bivariate normal conjugacy.

are mutually independent separable Wiener processes with unit dispersion parameters and drift parameters respectively $\nu, \cdots, \theta$; X is continuous with probability 1 by virtue of the continuity of its components. Suppose that the vectors μ_0 and μ_1 are "conjugate" with respect to the plane

$$R = d_1\nu + \cdots + d_k\theta + d_0 = 0 \tag{9}$$

in μ-space, in the sense that μ_0 and μ_1 are at the same distance from, and on the same perpendicular to, R(as illustrated for $k = 2$ in Figure 1). Consider as well the plane

$$R^*: d_1x_1 + \cdots + d_kx_k + d_0t = 1 \tag{10}$$

in (t, x)-space, and a "sufficiently regular" condition $K(s)$ on the behavior of X on the interval $(0, s)$. Then, if one defines

(11) $E(t)$: for some s in $(0, t), K(s)$ obtains, and $X(\cdot\,; \mu)$ crosses R^* at s,

one has

$$\Pr\{E(t) \mid \mu_1\} = \lambda(\mu_0, \mu_1) \Pr\{E(t) \mid \mu_0)\}, \tag{12}$$

where $\lambda(\mu_0, \mu_1) = \exp[(\nu_1 - \nu_0)/d_1] = \cdots = \exp[(\theta_1 - \theta_0)/d_k]$.

3. Wedge boundaries for the Wiener process. Consider first a wedge-shaped region bounded by the line $b(t) = h_0 + r_1 t$ (the "Accept" line) and the line $b_1(t) = h_1 + r_0 t$ (the "Reject" line), with $h_0 < 0 < h_1$ and

$$r_1 \geqq r_0 . \tag{13}$$

Letting $K(s)$ be the condition that $X(s'; \mu) > b_1(s')$ for no s' in $(0, s)$, the event $E((h_1 - h_0)/(r_1 - r_0))$, as defined by (7), specifies absorption somewhere on $b(\cdot)$, i.e. "Accept". Hence relation (5), in view of the usual definition of the OC, yields

$$\text{OC }(\mu) = \exp\,[2h_0(\mu - r_1)]\,\text{OC }(2r_1 - \mu) \tag{14}$$

and similarly, since absorption on one or the other boundary is assured,

$$1 - \text{OC }(\mu) = \exp\,[2h_1(\mu - r_0)]\,(1 - \text{OC }(2r_0 - \mu)). \tag{15}$$

If the OC is known at some point $\tilde{\mu}$, say computed in accordance with [1] or [8], relations (14) and (15) determine the OC on the two uniformly spaced grids $\tilde{\mu} + 2k(r_1 - r_0)$ and $-\tilde{\mu} + (2k + 1)(r_1 - r_0) + (r_1 + r_0)$, $k = 0, \pm 1, \pm 2, \cdots$. When $h_0 = -h_1 \equiv -h$, symmetry provides such a point directly, namely $\tilde{\mu} = (r_0 + r_1)/2 \equiv \bar{r}$, with $\text{OC}(\bar{r}) = \frac{1}{2}$, in which case the OC is determined on the single uniformly spaced grid $\bar{r} + k(r_1 - r_0)$, $k = 0, \pm 1, \pm 2, \cdots$. Defining $\delta = r_1 - r_0$, alternating applications of (14) and (15) then yield

$$\begin{aligned}
&\text{OC }(\bar{r} + \delta) = e^{-\delta h}\,\text{OC }(\bar{r}) = e^{-\delta h}/2, \qquad 1 - \text{OC }(\bar{r} - \delta) \\
&\quad = e^{-\delta h}(1 - \text{OC }(\bar{r})) \text{ or OC }(\bar{r} - \delta) = 1 - e^{-\delta h}/2, \\
&\text{OC }(\bar{r} + 2\delta) = e^{-3\delta h}\,\text{OC }(\bar{r} - \delta) = e^{-3\delta h}(1 - e^{-\delta h}/2), \\
&1 - \text{OC }(\bar{r} - 2\delta) = e^{-3\delta h}(1 - \text{OC }(\bar{r} + \delta)) \text{ or OC }(\bar{r} - 2\delta) \\
&\quad = 1 - e^{-3\delta h}(1 - e^{-\delta h}/2) \cdots
\end{aligned} \tag{16}$$

Note that Anderson's integral OC expression (equation 4.63 of [1]) for the symmetric case is thereby evaluated on a grid. Specifically, for $h_0 = -h_1 = -h$ and $r_0 = -r_1 \equiv -r$, Anderson gives

$$\begin{aligned}
1 - \text{OC }(\mu) = {}& (2\pi)^{-\frac{1}{2}} \textstyle\int_{-\infty}^{\infty} \exp\,[-(z - \mu(h/r)^{\frac{1}{2}})^2/2] \\
& \exp\,[2(rh)^{-\frac{1}{2}}z][1 + \exp\,(2(rh)^{\frac{1}{2}}z)]^{-1}\,dz,
\end{aligned} \tag{17}$$

which, on the grid $\bar{r} + k\delta$, i.e. the grid $2kr$, reduces, with $a = 2(rh)^{\frac{1}{2}}$, to

$$\begin{aligned}
1 - \text{OC }(2kr) = (2\pi)^{-\frac{1}{2}} \textstyle\int_{-\infty}^{\infty} \{&\exp\,[-(z - ka)^2/2] \exp\,(az)/ \\
&(1 + \exp\,(az))\}\,dz,
\end{aligned}$$

of which the left hand side is given by (16), with $\bar{r} = 0$ and $\delta = 2r$.

For increasing $|\mu|$, the marginal utility of knowing one OC $(\tilde{\mu})$ in addition to (14) and (15) declines, and these relations, by themselves, provide increasingly sharper bounds for the OC. For example, substituting $2r_1 - \mu$ for μ in (15), followed by substituting the resulting expression for OC $(2r_1 - \mu)$ in (14), yields

$$\text{OC}(\mu) = \exp[2h_0(\mu - r_1)]\{1 - \exp[2h_1(r_1 + \delta - \mu)][1 - \text{OC}(\mu - 2\delta)]\}, \tag{18}$$

so that

$$\exp[2h_0(\mu - r_1)][1 - \exp[2h_1(r_1 + \delta - \mu)]] \leqq \text{OC}(\mu) \leqq \exp[2h_0(\mu - r_1)], \tag{19}$$

which is informative on the right for $\mu > r_1$ and on the left for $\mu > r_1 + \delta$, and bounds OC (μ) progressively more sharply as $\mu \to +\infty$, to the extent that one may conclude, for μ large, that

$$\text{OC}(\mu) \cong \exp[2h_0(\mu - r_1)]. \tag{20}$$

Similar relations hold for μ small, in which case

$$1 - \text{OC}(\mu) \cong \exp[2h_1(\mu - r_0)]. \tag{21}$$

Iterating (18) provides additional information, since substituting $\mu - 2\delta$ for μ in the left hand side of (18), followed by substituting the resulting expression for OC $(\mu - 2\delta)$ back into (18), yields

$$\begin{aligned}\text{OC}(\mu) = \exp[2h_0(\mu - r_1)]\{1 &- \exp[2h_1(r_1 + \delta - \mu)] \\ &+ \exp[2h_1(r_1 + \delta - \mu)]\exp[2h_0(\mu - r_1 - 2\delta)] \\ &\cdot[1 - \exp[2h_1(r_1 + 3\delta - \mu)] + \exp[2h_1(r_1 + 3\delta - \mu)]\,\text{OC}(\mu - 4\delta)]\},\end{aligned}$$

so that

$$\begin{aligned}&\exp[2h_0(\mu - r_1)]\{1 - \exp[2h_1(r_1 + \delta - \mu)] + \exp[2h_1(r_1 \\ &\quad + \delta - \mu)]\exp[2h_0(\mu - r_1 - 2\delta)][1 - \exp[2h_1(r_1 + 3\delta - \mu)]]\} \\ &\quad \leqq \text{OC}(\mu) \leqq \exp[2h_0(\mu - r_1)]\{1 - \exp[2h_1(r_1 + \delta - \mu)] \\ &\quad + \exp[2h_1(r_1 + \delta - \mu)]\exp[2h_0(\mu - r_1 - 2\delta)]\},\end{aligned} \tag{22}$$

where the upper and lower bounds bound OC (μ) more sharply than does (19), respectively for $\mu > r_1 + 2\delta$ and $\mu > r_1 + 3\delta$. A further iteration yields a still sharper upper bound for $\mu > r_1 + 4\delta$ and sharper lower bound for $\mu > r_1 + 5\delta$, and so on. An analogous argument for the left tail thus leads to an upper bound function $U(\mu)$, valid for all μ, of changing analytic form over successive intervals $(r_0 + (2k - 1)\delta, r_1 + 2k\delta)$, and a lower bound function $L(\mu)$ of changing analytic form over successive intervals $(r_0 + 2k\delta, r_1 + (2k + 1)\delta)$.

In the symmetric case ($h_0 = -h_1$), Anderson (Corollary 4.5 of [1]) gives an approximation, call it $A(\mu)$, intended for moderate values of μ, to the OC function (17). Since the tails of this approximation are of order $\mu^{-1}\exp[-c\mu^2]$, it will be possible to divide the approximation task between $A(\mu)$ and $(L(\mu), U(\mu))$ in accordance with the intersection of $A(\mu)$ with $U(\mu)$ on the left, and with $L(\mu)$ on the right. Another aspect of the symmetric case is that, within the family of symmetric wedge procedures, the SPRT's (i.e., wedge procedures with $r_0 = r_1$) minimize risk for fixed AST, in a certain asymptotic sense. Specifically, suppose that we consider, for some large fixed μ, all symmetric wedge procedures with

$$\text{AST}(\mu) \doteq \text{AST}(-\mu) \doteq \tau, \qquad \tau \text{ small.} \tag{23}$$

Now, if T denotes the random time to decision, $E[T \mid \mu](\mu - s_0) = E[X(\cdot; \mu) - s_0 T \mid \mu] = E[X(\cdot; \mu) - s_0 T \mid \text{Accept}; \mu]\cdot \text{OC}(\mu) + h_1(1 - \text{OC}(\mu))$, where the first equality follows from the fundamental identity for the Wiener process [10]. Hence $\lim_{\mu\to\infty} \mu\, \text{AST}(\mu) = h_1$, and, similarly, $\lim_{\mu\to\infty} \mu\, \text{AST}(-\mu) = -h_0$, so that (23) amounts, approximately, to restricting consideration to symmetric wedge procedures with $-h_0 = h_1 = \mu\tau \equiv h$, and, in view of (20) and (21), one has, for such wedges,

$$\lim_{\mu\to\infty} \exp[4\mu h]\, \text{OC}(\mu)(1 - \text{OC}(-\mu)) = \exp(2h\delta),$$

or

$$\text{OC}(\mu)(1 - \text{OC}(-\mu)) \doteq \exp[2\mu\tau(r_1 - r_0) - 4\mu^2\tau].$$

But, as specified by (13), wedges restrict us to $r_1 - r_0 \geqq 0$; hence, within the class of symmetric wedges satisfying (23), the risk product on the left is approximately minimum when $r_0 = r_1$.

4. K-decision procedures. A certain 3-decision procedure based on two SPRT's is suggested in [2], [20] and [3]; as pointed out in [18],[4] this is easily extended to a K-decision procedure based on $K - 1$ component SPRT's, whose K OC functions are differences between the OC functions of successive component SPRT's. The $K - 1$ component SPRT's may be replaced by $K - 1$ wedges with common intercepts (h_0, h_1), as illustrated by Figure 2 for the case $K = 4$, and the K OC functions of the resulting K-decision procedure can be computed in special cases, again by subtraction. For example, when all $K - 1$ wedges are symmetric, i.e., when $h_0 = -h_1 \equiv -h$, and the apexes of the $K - 1$ wedges all are of the form $(T, 2lh)$, then the $K - 1$ grids coincide, all being of form $2lh$, $l{:}0, \pm1, \pm2, \cdots$, and all K OC functions are therefore computable on this grid.

Another sort of extension of the original proposals in [2], [20] and [3] is to complicate the geometry in a manner indicated by Figure 3 for the case $K = 3$: There are now three (rather than two—in general $(2^{K-1} - 1)$ rather than $(K - 1)$) wedges, call them W, W_U and W_L, with W nested within both W_U and W_L. Decision d_3 corresponds to the event UU: absorption on the upper boundary of W, followed by absorption on the upper boundary of W_U, decision

[4] We owe this reference to J. A. Lechner.

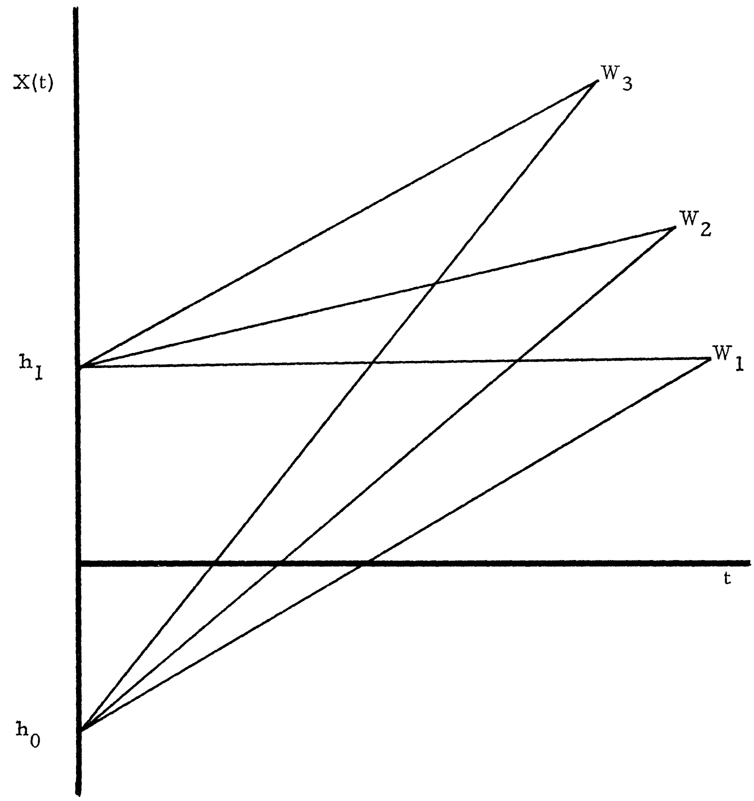

FIG. 2. Four-decision wedge system.

d_1 to LL, and d_2 to UL and LU. Hence, in an obvious notation, the three OC functions are given by $g_{UU}(\mu)$, $g_{UL}(\mu) + g_{LU}(\mu)$, and $g_{LL}(\mu)$.

Computing the probabilities g by the methods of this paper again is possible only in special cases; for example, as illustrated by Figure 3, when W is symmetric and W_U and W_L are degenerate (i.e., SPRT's), with slopes restricted to a certain grid depending on the slopes of the boundaries of W. Specifically, let the SPRT's W_U and W_L have slopes respectively $R = (r_0 + r_1)/2 + m(r_1 - r_0)/2 \equiv \bar{r} + m\delta/2$ and $S = \bar{r} + n\delta/2$. Then, denoting the OC of W by OC (μ), the iterative construction (16) yields OC (μ), hence $g_{LL}(\mu) + g_{LU}(\mu)$ and $g_{UL}(\mu) + g_{UU}(\mu)$, on the grid $\bar{r} + k\delta$, and in particular at the points $\mu_1 \equiv \bar{r} + k\delta$ and $\mu_0 \equiv \bar{r} + (m - k)\delta$ mutually conjugate with respect to $R = \bar{r} + m\delta/2$.

Moreover, recall the particular $K(s)$ defined following (8), and identify symbols as indicated in Figure 3. If this $K(s)$ is substituted in (7), the event $E(t)$ of (7), with $t = +\infty$, is seen in fact to be the event UL, and (8) yields

$$g_{UL}(\mu_1) = \exp [H_0\delta(2k - m)]g_{UL}(\mu_0)$$

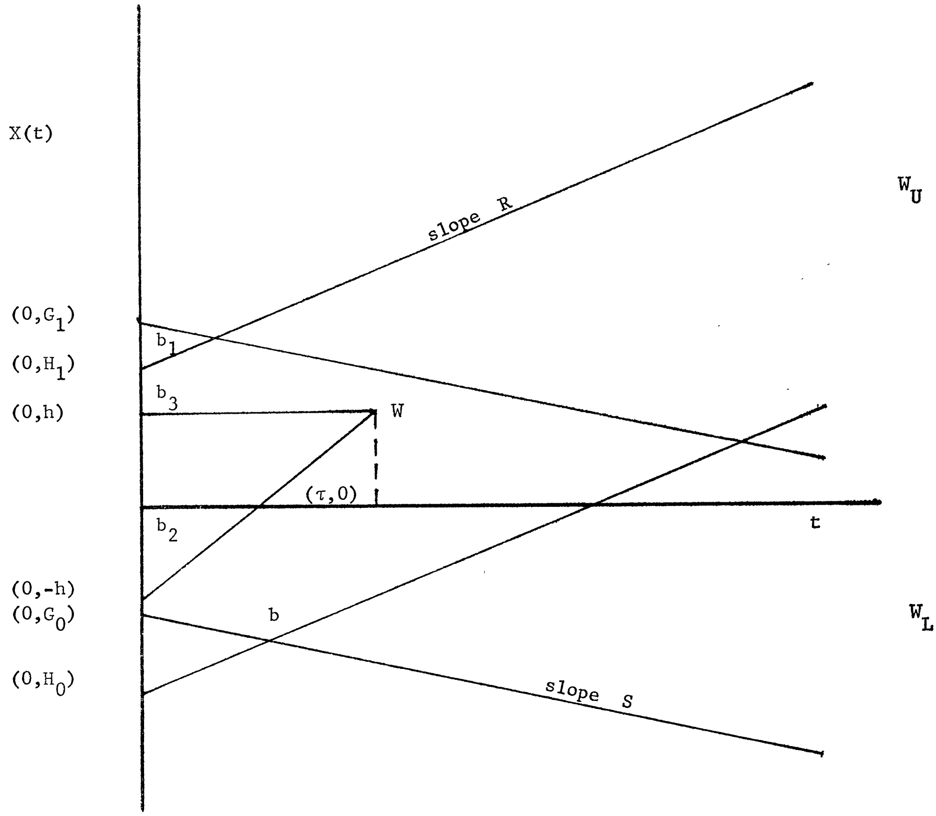

FIG. 3. Special three-decision procedure.

and, similarly,

$$g_{UU}(\mu_1) = \exp [H_1\delta(2k - m)]g_{UU}(\mu_0),$$

these relations being informative, except when $k = m/2$, i.e. when $\mu_0 = \mu_1 = R$. But, according to the previous paragraph, $g_{UL}(\mu_0) + g_{UU}(\mu_0)$ and $g_{UL}(\mu_1) + g_{UU}(\mu_1)$ are known; hence we have in fact four equations in the four unknowns $g_{UL}(\mu_0)$, $g_{UL}(\mu_1)$, $g_{UU}(\mu_0)$ and $g_{UU}(\mu_1)$, and therefore know $g_{UL}(\mu)$ and $g_{UU}(\mu)$ at every point of the grid $\bar{r} + k\delta$, excepting $\mu = \bar{r} + m\delta/2$ when m is even. A similar argument applies for g_{LL} and g_{LU}, and the three OC functions are thus determined for $\mu = \bar{r} + k\delta$, except possibly for $k = m/2$, $n/2$.

5. Higher-dimensional Wiener processes. Consider a cylinder in (t, x_1, x_2)-space with equilateral triangular base and planar boundary portions A, B and C as indicated in Figure 4. Let $X(t) = (X_1(t), X_2(t))$ be a two-dimensional Wiener process, with drift parameters (ν, θ), starting at the origin. Of interest are the probabilities $A(\nu, \theta)$, $B(\nu, \theta)$ and $C(\nu, \theta)$ that $X(t)$ first traverses the

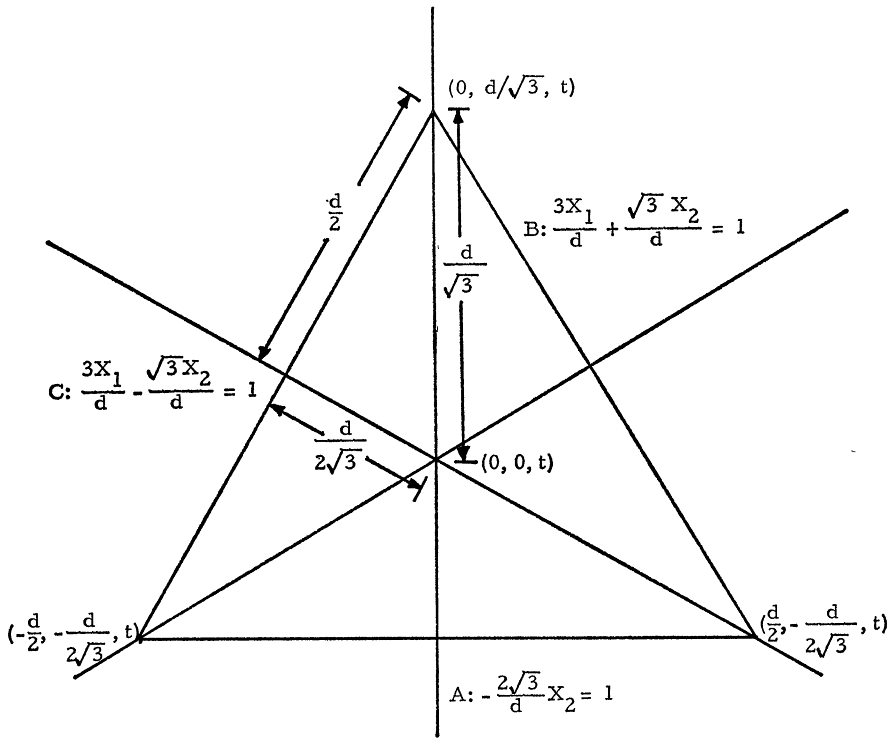

FIG. 4. Equilateral triangular boundary in $(X_1(t), X_2(t))$-plane.

cylindrical boundary on a particular one of the three planar boundary portions A, B, and C. Note to this end that the parameter points (ν, θ) and $(-\nu, -\theta)$ of Figure 5 are conjugate with respect to the line β in the sense of (9), with

$$d_1 = 3/d, \qquad d_2 = 3^{\frac{1}{2}}/d, \qquad d_0 = 0, \tag{24}$$

and that, with (24), R^* of (10) is in fact the plane containing B. Note also, if $K(s)$ of (11) is taken to be the obvious condition precluding the crossing of either plane A or plane C in $(0, s)$, that $\Pr\{E(+\infty) \mid \mu)\}$ is then $B(\nu, \theta)$, so that (12) yields

$$B(\nu, \theta) = \exp(2\,d\theta/3^{\frac{1}{2}})B(-\nu, -\theta). \tag{25}$$

An analogous argument yields

$$A(\nu, \theta) = \exp(-d\theta/3^{\frac{1}{2}})A(\nu, -\theta). \tag{26}$$

Symmetry also requires that

$$2A(\nu, \theta) + B(\nu, \theta) = 2A(\nu, -\theta) + B(-\nu, -\theta) = 1, \tag{27}$$

and solving (25), (26) and (27) yields

$$A(\nu, \theta) = (1 - \exp(d\theta/3^{\frac{1}{2}}))/2(1 - \exp(3^{\frac{1}{2}}\,d\theta))$$

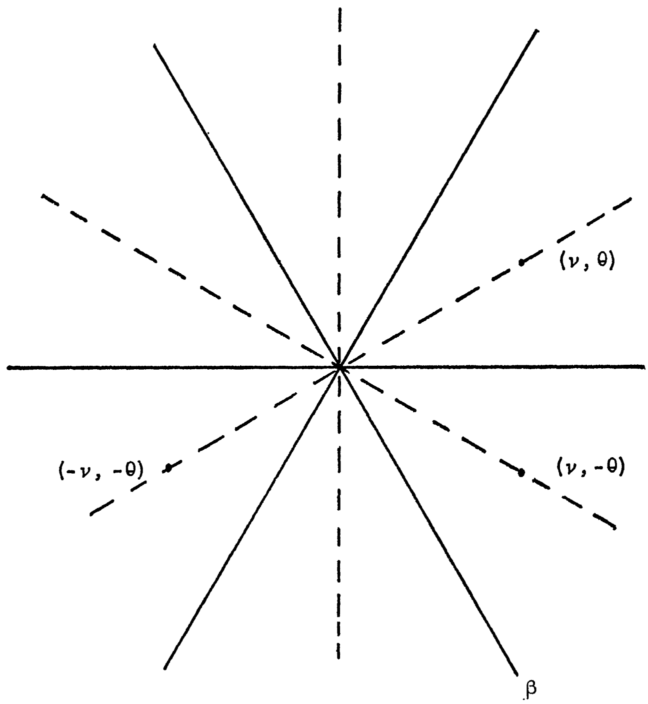

FIG. 5. Conjugacy and symmetry lines in the (ν, θ)-plane.

and

$$A(\nu, -\theta) = (\exp(d\theta/3^{\frac{1}{2}}) - \exp(3^{\frac{1}{2}} d\theta))/2(1 - \exp(3^{\frac{1}{2}} d\theta)),$$

so that, by symmetry, $A(\nu, \theta)$, $B(\nu, \theta)$ and $C(\nu, \theta)$ are evaluated on the three dotted lines of Figure 5.

Similar arguments yield $A(\nu, \theta)$, $B(\nu, \theta)$, $C(\nu, \theta)$ and $D(\nu, \theta)$ on the two dotted lines of Figure 7 when the base of the cylindrical region is the square indicated in Figure 6. In particular, $C(\nu, \nu)$ and $C(\nu, -\nu)$ are then $e^{d\nu}/2(1 + e^{d\nu})$ and $1/2(1 + e^{d\nu})$ respectively; these expressions extending to $e^{d\nu}/k(1 + e^{d\nu})$ and $1/k(1 + e^{d\nu})$ for k dimensions.

6. The binomial case. This section treats binomial sampling for the symmetric wedge procedures of Section 3: Stop as soon as $d_m < -h + r_1 m$ or $d_m > h + r_0 m$, "accepting" in the first case and "rejecting" in the second. The event "Accept" thus specifies sample sequences $(x_1, \cdots, x_m)$ all of which satisfy (1), so that, in view of (5), defining $\lambda(p_0, p_1, g) \equiv [p_1(1 - p_0)/p_0(1 - p_1)]^g$,

$$\text{OC}(p_1) \doteq \lambda(p_0, p_1, -h)\ \text{OC}(p_0) \tag{28}$$

for p_0 and p_1 conjugate with respect to r_1 in the sense of (3), and a similar argument yields

$$1 - \text{OC}(p_1) \doteq \lambda(p_0, p_1, h)(1 - \text{OC}(p_0)) \tag{29}$$

for p_0 and p_1 conjugate with respect to r_0.

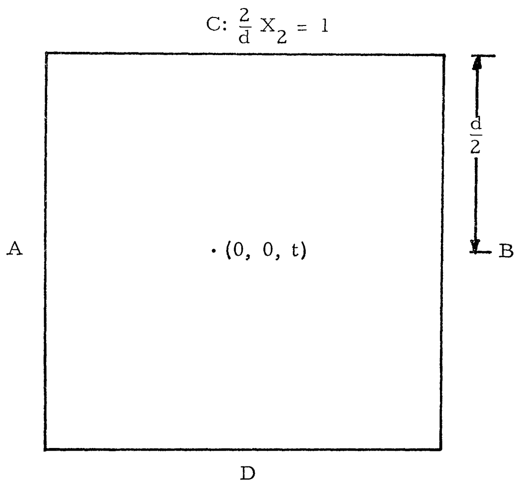

FIG. 6. Square boundary in $(X_1(t), X_2(t))$-plane.

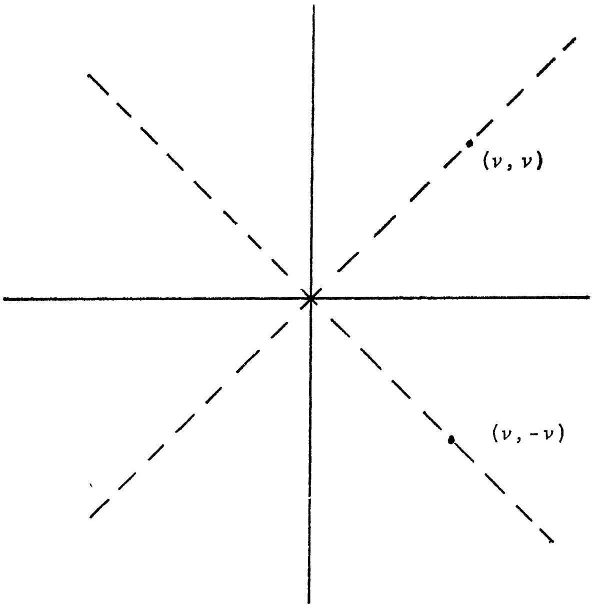

FIG. 7. Conjugacy and symmetry lines in the (ν, θ)-plane.

If symmetry is to be counted on again for evaluating the OC, one needs the further condition $r_0 + r_1 = 1$, which insures OC $(\frac{1}{2}) = \frac{1}{2}$. This restriction still accommodates procedures suitable for testing $H_0{:}p < \frac{1}{2}$ vs. $H_1{:}p > \frac{1}{2}$ with symmetric losses due to wrong decision; hence, by Wald's ([21], p. 107) treatment of the double dichotomy, procedures suitable for testing $H_0: p_0 < p_1$ vs. $H_1: p_0 > p_1$.

The grid on which the OC is computed now consists of the point $\frac{1}{2}$, a sequence of points $\frac{1}{2} > \pi_1 > \pi_2 > \cdots$ converging to zero, and the complementary sequence $\frac{1}{2} < \rho_1 < \rho_2 < \cdots$,

$$\pi_i + \rho_i = 1. \tag{30}$$

Specifically, π_1 is the conjugate of $\frac{1}{2}$ with respect to r_0, ρ_2 the conjugate of π_1 with respect to r_1, π_3 the conjugate of ρ_2 with respect to r_0, and so on. A similar construction applies to $\rho_1, \pi_2, \rho_3, \cdots$. The OC is then computed by repeated application of (28) and (29); thus

$$\begin{aligned} \text{OC}(\rho_1) &\doteq \lambda(\tfrac{1}{2}, \rho_1, -h)\, \text{OC}(\tfrac{1}{2}) = \tfrac{1}{2}((1-\rho_1)/\rho_1)^h, \\ 1 - \text{OC}(\pi_2) &\doteq \lambda(\rho_1, \pi_2, h)\cdot(1 - \text{OC}(\rho_1)), \\ \text{OC}(\rho_3) &\doteq \lambda(\pi_2, \rho_3, -h)\, \text{OC}(\pi_2), \end{aligned} \tag{31}$$

and the OC on $\pi_1, \rho_2, \pi_3, \cdots$ is gotten in similar fashion, or simply from the fact that symmetry and (30) imply that OC (π_i) + OC $(\rho_i) = 1$.

It is also possible to bound the OC in essentially the manner suggested by Wald in the case of the SPRT ([21], p. 164): All sample sequences leading to "Accept" in fact satisfy $T_m < -h + r_1 m$, so that, for

$$p_1 > p_0, \tag{32}$$

relation (28) can be sharpened to

$$\text{OC}(p_1) < \lambda(p_0, p_1, -h)\, \text{OC}(p_0), \tag{33}$$

and, similarly, relation (29) to

$$1 - \text{OC}(p_1) > \lambda(p_0, p_1, h)(1 - \text{OC}(p_0)). \tag{34}$$

In addition, every sample sequence $(x_1, \cdots, x_m)$ leading to "Accept" satisfies (a) $T_{m-1} \geqq -h + r_1(m-1)$, which, under (32), implies

$$\Pr\{x_1, \cdots, x_{m-1} \mid p_1\} \geqq \lambda(p_0, p_1, -h) \Pr\{x_1, \cdots, x_{m-1} \mid p_0\} \tag{35}$$

and (b) $x_m = 0$, which, together with (35), implies that

$$\Pr\{x_1, \cdots, x_m \mid p_1\} \geqq (1-p_1)/(1-p_0)\lambda(p_0, p_1, -h) \Pr\{x_1, \cdots, x_m \mid p_0\}.$$

Hence, under (32), (33) can be complemented by

$$\text{OC}(p_1) \geqq (1-p_1)/(1-p_0)\lambda(p_0, p_1, -h)\text{OC}(p_0) \tag{36}$$

and, similarly, (34) by

$$1 - \text{OC}(p_1) \leqq (p_1/p_0)\lambda(p_0, p_1, h)(1 - \text{OC}(p_0)), \tag{37}$$

so that, in view of (33), (34), (36) and (37), the computations (31) can be expanded on as follows:

(a) By (33) and (36), $(1 - \rho_1)^{h+1}/\rho_1^h = (1 - \rho_1)\lambda(\frac{1}{2}, \rho_1, -h) \leqq \mathrm{OC}(\rho_1) < \frac{1}{2}\cdot\lambda(\frac{1}{2}, \rho_1, -h) = \frac{1}{2}\cdot((1 - \rho_1)/\rho_1)^h$.

(b1) By (34) and (a), $1 - \mathrm{OC}(\pi_2) < \lambda(\rho_1, \pi_2, h)(1 - \mathrm{OC}(\rho_1)) \leqq \lambda(\rho_1, \pi_2, h)\cdot(1 - (1 - \rho_1)^{h+1}/\rho_1^h)$

(b2) By (37) and (a), $1 - \mathrm{OC}(\pi_2) \geqq (\pi_2/\rho_1)\lambda(\rho_1, \pi_2, h)(1 - \mathrm{OC}(\rho_1)) > (\pi_2/\rho_1)\lambda(\rho_1, \pi_2, h)(1 - \frac{1}{2}((1 - \rho_1)/\rho_1)^h)$,

and so on, successively, to bounds of decreasing relative sharpness for the OC at all other points π_i and ρ_i.

REFERENCES

[1] ANDERSON, T. W. (1960) A modification of the sequential probability ratio test to reduce sample size *Ann. Math. Statist.* **31** 165–197.

[2] ARMITAGE, P. (1947). Some sequential tests of Student's hypothesis. *Suppl. J. Roy. Statist. Soc.* **9** 250–263.

[3] ARMITAGE, P. (1960). *Sequential medical trials.* Blackwell. Oxford, England.

[4] BAKER, A. G. (1950) Properties of some tests in sequential analysis. *Biometrika* **37** 334–346.

[5] BLASBALG, H. (1957). Transformations of the fundamental relationships in sequential analysis. *Ann. Math. Statist.* **28** 1024–1028.

[6] DAVID, H. T. (1952). Approximate OC functions for the first Rushton approximations to the WAGR test. Unpublished memorandum SRC-20205D12. Statistics Research Center, Univ. of Chicago.

[7] DAVID, H. T. and LAWING, WILLIAM D. (1964). Multi-decision sequential procedures (preliminary report). Abstract. *Ann. Math. Statist.* **35** 1835.

[8] DONNELLY, T. G. (1957). A family of truncated sequential tests. Unpublished Ph.D. thesis. Library, Univ. of North Carolina.

[9] DOOB, J. L. (1953). *Stochastic process.* Wiley, New York.

[10] DVORETZKY, J., KIEFER, J., and WOLFOWITZ, J. (1953). Sequential decision problems for processes with continuous time parameter [I] Testing hypotheses. *Ann. Math. Statist.* **24** 254–264.

[11] GIRSHICK, M. A. (1946). Contributions to the theory of sequential analysis II, III. *Ann. Math. Statist.* **17** 282–298.

[12] HALL, W. J. (1961). Sequential analysis. Unpublished lecture notes, Univ. of North Carolina. Original not available for examination; cited in Fukushima, Kozo. 1961.

[13] HALL, W. J. (1961). Some sequential tests for symmetric problems. Invited address at Eastern Regional Meeting of Institute of Mathematical Statistics, Ithaca, N. Y.

[14] HALMOS, PAUL R. (1950). *Measure theory.* van Nostrand, New York.

[15] KIEFER, J. and WEISS, LIONEL. (1957). Some properties of generalized sequential probability ratio tests. *Ann. Math. Statist.* **28** 57–75.

[16] LAWING, WILLIAM DENNIS, JR. (1965). Multiple decision sequential procedures. Unpublished Ph.D. thesis. Library, Iowa State Univ.

[17] LECHNER, J. A. (1964). Optimality and the OC curve for the Wald SPRT. *Amer. Statist. Assoc. J.* **59** 464–468.

[18] 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.

[19] SAVAGE, LEONARD J. (1957). When different pairs of hypotheses have the same family of likelihood ratio test regions. *Ann. Math. Statist.* **28** 1028–1032.

[20] SOBEL, MILTON and WALD, A. (1949). A sequential decision procedure for choosing one of three hypotheses concerning the unknown mean of a normal distribution. *Ann. Math. Statist.* **20** 502–522.

[21] WALD, ABRAHAM. (1947). *Sequential analysis.* Wiley, New York.