NASA CR-
140385

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF HOUSTON HOUSTON, TEXAS

(NASA-CR-140385) SOME QUALITATIVE REMARKS ON THE VARIATION OF THE PROBABILITY OF ERROR (Houston Univ.) 18 p HC $3.25 CSCL 12A N75-13576 Unclas G3/65 04901

SOME QUALITATIVE REMARKS ON THE VARIATION OF THE PROBABILITY OF ERROR.
BY HOMER F. WALKER.
AUGUST 1974. REPORT #34

PREPARED FOR
EARTH OBSERVATION DIVISION, JSC
UNDER
CONTRACT NAS-9-12777

3801 CULLEN BLVD.
HOUSTON, TEXAS 77004

Some Qualitative Remarks on the Variation of the Probability of Error

August 1974

Homer F. Walker
Department of Mathematics
University of Houston
Houston, Texas

Report # 34

Some Qualitative Remarks on the Variation of the Probability of Error

by

Homer F. Walker
Department of Mathematics
University of Houston

1. Introduction.

In this note, we discuss the qualitative behavior of the probability of misclassifying observations in $\mathbb{R}^n$ from two normally distributed populations as the classification regions are varied in a prescribed way. This discussion is intended to provide a preliminary generalization of the results obtained by Walton [4] for the case of normally distributed observations in $\mathbb{R}^1$ with varying *a priori* probabilities. We hope to provide quantitative analogues of these results in subsequent reports.

We assume that observations from two populations π_1 and π_2 are known to have *a priori* probabilities α_{01} and α_{02} and normal density functions

$$p_{0i}(x) = \frac{1}{(2\pi)^{n/2}\,|\Sigma_{0i}|^{1/2}}\; e^{-1/2(x-\mu_{0i})^T\Sigma_{0i}^{-1}(x-\mu_{0i})},\quad i = 1,2,$$

for $x = (x_1,\dots,x_n)^T \in \mathbb{R}^n$. We further assume that these observations are classified, not by using the true Bayes optimal (maximum likelihood) classification scheme for π_1 and π_2, but by using the Bayes optimal classification scheme defined by a priori probabilities $\alpha_1(t)$ and $\alpha_2(t)$ and density functions

$$p_i(x,t) = \frac{1}{(2\pi)^{n/2}|\Sigma_i(t)|^{1/2}} \; e^{-1/2(x-\mu_i(t))^T \Sigma_i(t)^{-1}(x-\mu_i(t))}, \quad i = 1,2,$$

where the functions $\alpha_i(t)$, $\mu_i(t)$, and $\Sigma_i(t)$ are continuously differentiable functions of the parameter t in a neighborhood of $t = 0$. This is to say that and observation $x \in \mathbb{R}^n$ is classified as coming from π_i if and only if $\alpha_i(t)\ p_i(x,t) = \max_{j=1,2} \alpha_j(t)\ p_j(x,t)$. (We assume that $p_1(x,t) \not\equiv p_2(x,t)$ as a function of x in a neighborhood of $t = 0$).

Under these assumptions, the probability of error in classifying an observation is a function of t in a neighborhood of $t = 0$, given by

$$P_e\ (t) = \int_{R_{1(t)}} \alpha_{02} p_{02}(x)dx + \int_{R_{2(t)}} \alpha_{01} p_{01}(x)dx,$$

where the regions $R_1(t)$ and $R_2(t)$ are defined as follows. Let

$$F(x,t) = \log \frac{\alpha_1(t)p_1(x,t)}{\alpha_2(t)p_2(x,t)}$$

$$= \log \frac{\alpha_1(t)|\Sigma_2(t)|^{1/2}}{\alpha_2(t)|\Sigma_2(t)|^{1/2}} - \frac{1}{2}(x-\mu_1(t))^T\Sigma_1(t)^{-1}(x-\mu_1(t)) + \frac{1}{2}(x-\mu_2(t))^T\Sigma_2(t)^{-1}(x-\mu_2(t)).$$

Then $R_1(t) = \{x \in \mathbb{R}^n : F(x,t) \geq 0\}$ and $R_2(t) = \{x \in \mathbb{R}^n : F(x,t) < 0\}$. (For a more thorough discussion of the probability of error, see Anderson [1].)

Our goal is to examine qualitatively the rate at which $P_e(t)$ varies as t varies in a neighborhood of zero. In our main result, the exact rate of variance of $P_e(t)$ is seen to depend on a number of factors. However, an inequality of the form

$$|P_e(t) - P_e(0)| \leq K|t|^\alpha$$

is obtained in every case. In other words, $P_e(t)$ is always Hölder continuous at $t = 0$. In the following, the exponent α is determined precisely in each case. The constant K is merely asserted to exist; no estimate of its size is given. Unfortunately, to implement such an inequality in practice, one must know both the size of K and the range of t for which the inequality holds.

In the sequel, large constants are denoted generically by K, K', etc. Distinguished constants are subscripted. The common boundary of $R_1(t)$ and $R_2(t)$ is denoted by $S(t)$.

2. The variation of $P_e(t)$.

Our objective is to prove the following theorem.

Theorem: If $\nabla F(x,0) \neq 0$ on $S(0)$, then there exists a constant K such that

$|P_e(t) - P_e(0)| \le K|t|$ for small t. If $\nabla F(x,0)$ vanishes somewhere on $S(0)$, then there exists a constant K such that $|P_e(t) - P_e(0)| \le K|t|^{\frac{m}{m+1}}$ for small t, where m is the number of non-zero eigenvalues (counting multiplicity) of $\Sigma_2(0)^{-1} - \Sigma_1(0)^{-1}$.

Remarks: If $\nabla F(x,0)$ vanishes anywhere, then the assumption $p_1(x,0) \not\equiv p_2(x,0)$ implies that $m > 0$. Thus $P_e(t)$ is Hölder continuous at $t = 0$ with exponent at least $\frac{1}{2}$. In the special case in which $\alpha_i(0) = \alpha_{0i}$, $\mu_i(0) = \mu_{0i}$, $\Sigma_i(0) = \Sigma_{0i}$, $i = 1,2$, exponents of Hölder continuity larger than those specified above can be obtained. The determination of these exponents is not carried out here.

Before beginning the proof of the theorem, we establish several lemmas. For a subset $X \subseteq \mathbb{R}^n$, define

$$d(x,X) = \begin{cases} \inf\limits_{y \in X} |x-y| & \text{if } X \neq \emptyset \\ \infty & \text{if } X = \emptyset. \end{cases}$$

Let $T = \{x \in \mathbb{R}^n : \nabla F(x,0) = 0\}$.

For non-negative p and q and positive r and s, define

$$L_{p,r}(t) = \left\{ x \in \mathbb{R}^n : |x| \ge r|t|^{-p} \right\},$$

$$M_{p,q,r,s}(t) = \left\{ x \in \mathbb{R}^n : |x| \le r|t|^{-p} \text{ and } d(x,T) \ge s|t|^q \right\},$$

$$N_{p,q,r,s}(t) = \left\{ x \in \mathbb{R}^n : |x| \le r|t|^{-p} \text{ and } d(x,T) \le s|t|^q \right\}.$$

When there is no danger of ambiguity, we will omit the subcripts $p,q,r,$ and s, i.e., $L_{p,r}(t) = L(t)$, etc.

Lemma 1: Suppose that $0 \le q < 1$ and $0 \le p < 1-q$.
Then there exists a constant K, independent of p,q,r, and s, such that, if t is sufficiently small, then $|\nabla F(x,t)| \ge Ks|t|^q$ for all $x \in M(t)$.

Proof: Writing $F(x,t) = x^T A(t)x + B(t)x + C(t)$, one obtains $\nabla F(x,t) = 2A(t)x + B(t)^T$ and $\frac{\partial}{\partial t} \nabla F(x,t) = 2 \frac{d}{dt}A(t)x + \frac{d}{dt} B(t)^T$. From this, it is seen that there exist constants K' and K'', independent of p,q,r, and s, such that $|\nabla F(x,0)| \ge K's|t|^q$ and $|\frac{\partial}{\partial t} \nabla F(x,0)| \le K''(1 + r)|t|^{-p}$ for $x \in M(t)$ and t small. It follows that there exists a constant K, independent of p,q,r, and s, such that, for $x \in M(t)$,

$$|\nabla F(x,t)| \ge K's|t|^q - K''(1 + r)|t|^{1-p} \ge Ks|t|^q$$

whenever t is small.

Lemma 2: Suppose that $0 \le q \le \frac{1}{2}$ and $0 \le p \le \frac{1}{2} - q$.
If t and $\frac{(1 + r)}{s}$ are sufficiently small, or if t is sufficiently small and $0 \le p < \frac{1}{2} - q$, then, for $|t_0| \le |t|$ and $x \in M(t) \cap S(t_0)$, the solution $y(x,\tau)$ of the initial-value problem

$$\frac{d}{d\tau} y(x,\tau) = - \frac{\frac{\partial}{\partial t} F(x,\tau)}{|\nabla F(y,\tau)|^2} \nabla F(y,\tau)$$

$$y(x,t_0) = x$$

exists and is continuously differentiable in x and τ for $|\tau| \le |t|$.

Remarks: Note that, wherever $y(x,\tau)$ exists, $y(x,\tau) \in S(\tau)$ and $\frac{d}{d\tau}y(x,\tau)$ is normal to $S(\tau)$. It is seen in the proof that $|y(x,\tau) - x| \le K \frac{(1+r)^2}{s} |t|^{1-2p-q}$, where the constant K is independent of $p,q,r,s,$ and t for small t.

Proof: From Lemma 1 and the fact that $\frac{\partial}{\partial t}F(x,t)$ is quadratic in x, one sees that, if t is sufficiently small, then there exists a constant K, independent of $p,q,r,s,$ and t, such that

$$\frac{\left|\frac{\partial}{\partial t} F(x,\tau)\right|}{|\nabla F(x,\tau)|} \le \frac{K(1+r)^2}{s} |t|^{-2p-q}$$

for $|\tau| \le |t|$ and $x \in M_{p,q,2r,s/2}(t)$. Consequently, $|y(x,\tau) - x| \le 2K \frac{(1+r)^2}{s} |t|^{1-2p-q}$ whenever τ lies in the domain of existence of $y(x,\tau)$. If $\frac{1+r}{s}$ is so small that $2K \frac{(1+r)}{s} < r$ and $2K \frac{(1+r)^2}{s} < \frac{s}{2}$, then, since $1-2p-q \ge q$ and $1-2p - q \ge -p$, we have

(1) $$|y(x,\tau) - x| < r\,|t|^{-p} \quad \text{and} \quad |y(x,\tau) - x| < \frac{s}{2}|t|^{q}$$

whenever $x \in M_{p,q,2r,s/2}(t)$ and τ lies in the domain of existence of $y(x,\tau)$. If $0 \le p < \frac{1}{2} - q$, then $1-2p-q > q$ and $1-2p-q > -p$, and one easily verifies that (1) again holds for small t. Thus (1) holds under the hypotheses of the lemma.

Suppose that the hypotheses of the lemma are satisfied (so that (1) holds) and that there exists a t_0, $|t_0| \le |t|$, and $x \in M_{p,q,r,s}(t) \cap S(t_0)$ such that $y(x,\tau)$ does not exist for all τ, $|\tau| \le |t|$. Then one can find a t_1, $|t_1| \le |t|$, for which $y(x,t_1) \in \partial M_{p,q,2r,s/2}(t)$ [2]. But this contradicts

(1), and the lemma is proved.

Lemma 3: Suppose that $0 \le q \le \frac{1}{2}$ and $0 \le p \le \frac{1}{2} - q$.
If t and $\frac{1+r}{s}$ are sufficiently small, or if t is sufficiently small and $0 \le p < \frac{1}{2} - q$, then

(i) $R_1(t)\Delta R_1(0) \subseteq \bigcup_{|\tau|\le|t|} S(\tau)$,

(ii) $[R_1(t)\Delta R_1(0)]\cap M_{p,q,r,s}(t) \subseteq \{y(x,\tau)\in\mathbb{R}^n : |\tau| \le |t|$ and $x \in S(0)\cap M_{p,q,2r,s/2}(t)\}$,

(iii) $\{y(x,\tau) \in\mathbb{R}^n : |\tau| \le |t|$ and $x \in S(0)\cap M_{p,q,2r,\frac{s}{2}}(t)\} \subseteq \bigcup_{|\tau|\le|t|} S(\tau)\cap M_{p,q,3r,\frac{s}{3}}(t)$.

Proof:

(i) Suppose that $t > 0$ and $x \in R_1(t) - R_1(0)$. Set $t_0 = \inf\{\tau : x\in R_1(\tau) - R_1(0)\}$. Clearly, $x \in S(t_0)$. The other cases follow similarly.

(ii) If $w \in [R_1(t)\Delta R_1(0)]\cap M_{p,q,r,s}(t)$, then $w \in S(t_0)$ for some t_0, $|t_0| \le |t|$. If t and $\frac{1+r}{s}$ are small, or if t is small and $0 \le p < \frac{1}{2} - q$, then, by Lemma 2, $y(w,\tau)$ exists for $|\tau| \le |t|$. In particular, $x = y(w,0)$ satisfies $y(x,t_0) = w$. Now, by the remarks after Lemma 2, $|x-w| \le K\frac{(1+r)^2}{s}|t|^{1-2p-q}$ for a constant K independent of p,q,r,s, and t. If $\frac{1+r}{s}$ and t are sufficiently small, or if t is small and $0 \le p < \frac{1}{2} - q$, then one sees that $|x-w| < r\,|t|^{-p}$ and $|x-w| < \frac{s}{2}|t|^{q}$. Consequently, $x \in S(0)\cap M_{p,q,2r,\frac{s}{2}}(t)$.

(iii) If $x \in S(0)\cap M_{p,q,2r,\frac{s}{2}}(t)$, then, as in the proof of (ii), one uses an inequality $|y(x,\tau) - x| \le K\frac{(1+r)^2}{s}|t|^{1-2p-q}$ to obtain $y(x,\tau) \in M_{p,q,3r,\frac{s}{3}}(t)$ for $|\tau| \le |t|$, if t and $\frac{(1+r)}{s}$ are sufficiently small, or if t is small and $0 \le p < \frac{1}{2} - q$.

Proof of the theorem:

One has

$$P_e(t) - P_e(0) = \int_{R_1(t)} \alpha_{02}p_2(x)dx + \int_{R_2(t)} \alpha_{01}p_1(x)dx - \int_{R_1(0)} \alpha_{02}p_2(x)dx - \int_{R_2(0)} \alpha_{01}p_1(x)dx$$

$$= \int_{R_1(t)-R_1(0)} G(x)dx - \int_{R_1(0)-R_1(t)} G(x)dx,$$

where $G(x) = [\alpha_{01}P_1(x) - \alpha_{02}p_2(x)]$. Thus

$$|P_e(t) - P_e(0)| \le \int_{R_1(t)\Delta R_1(0)} |G(x)|dx,$$

and, for given $p,q,r,$ and s, we obtain

$$(2) \quad |P_e(t) - P_e(0)| \le \int_{[R_1(t)\Delta R_1(0)]\cap L(t)} |G| + \int_{[R_1(t)\Delta R_1(0)]\cap M(t)} |G| + \int_{[R_1(t)\Delta R_1(0)]\cap N(t)} |G| .$$

We consider the following cases:

(1) $\nabla F(x,0)$ never vanishes on $S(0)$,

(2) $\nabla F(x,0)$ vanishes somewhere on $S(0)$ and $m > 1$,

(3) $\nabla F(x,0)$ vanishes somewhere on $S(0)$ and $m = 1$.

Case 1: First, the following lemma is needed.

Lemma 4: Suppose $\nabla F(x,0) \neq 0$ on $S(0)$ and $0 \le p < \frac{1}{2}$.
If t and s are sufficiently small, then $S(\tau) \cap N(t) = \emptyset$ for $|\tau| \le |t|$

Proof: Suppose that the lemma is false. Then there exist sequences $\{s_j\}$, $\{t_j\}$, $\{\tau_j\}$, and $\{x_j\}$ with $s_j \to 0$, $t_j \to 0$, $|\tau_j| \le |t_j|$, and $x_j \in S(\tau_j) \cap N(t_j)$. Note that $\nabla F(x_j,\tau_j) \to 0$, since $|\nabla F(x_j,\tau_j)| \le Ks_j|t_j|^q$.

Now $F(x,t) = x^T A(t)x + B(t)x + C(t)$, where $A(t) = \frac{1}{2}(\Sigma_2^{-1}(t) - \Sigma_1^{-1}(t))$ is symmetric and $A(t)$ and $B(t)$ are continuously differentiable near $t = 0$. Denote by $\mathcal{N}(A(0))$ the null-space of $A(0)$. Writing $A(t) = A(0) + 0_1(t)$, $B(t) = B(0) + 0_2(t)$, and $x_j = y_j + z_j$, where $y_j \in \mathcal{N}(A(0))^{\perp}$ and $z_j \in \mathcal{N}(A(0))$, one obtains

$$0 = \lim_{j\to 0} \nabla F(x_j,\tau_j) = \lim_{j\to 0}\{2A(0)y_j + 0_1(\tau_j)x_j + B(\tau_j)^T\}.$$

Since $|x_j| \le r\,|t_j|^{-p}$ and $0 \le p < \frac{1}{2}$, $0_1(\tau_j)\,x_j \to 0$. It follows that $y_j \to y^* \in \mathcal{N}(A(0))^{\perp}$, and

$$0 = 2A(0)y^* + B(0)^T = \nabla F(y^*,0).$$

Note that this equation implies that $B(0)^T \in \mathcal{N}(A(0))^{\perp}$. Furthermore, we have

$$0 = F(x_j,\tau_j) = y_j^T A(0)y_j + x_j^T 0_1(\tau_j)x_j + B(0)y_j + 0_2(\tau_j)x_j + C(\tau_j).$$

As before, $x_j^T 0_1(\tau_j)x_j + 0_2(\tau_j)x_j \to 0$ since $|x_j| \le r\,|t_j|^{-p}$ and $0 \le p < \frac{1}{2}$. Consequently,

$$0 = y^{*T} A(0)y^* + B(0)y^* + C(0) = F(y^*,0).$$

This contradicts the assumption that $\nabla F(x,0)$ never vanishes on $S(0)$, and the proof is complete.

Using Lemma 4, we obtain from (2) that

$$(3) \qquad |P_e(t) - P_e(0)| \le \int_{L(t)} |G| + \int_{[R_1(t)\Delta R_1(0)]\cap M(t)} |G|$$

for $0 \le p < \frac{1}{2}$, if t and s are sufficiently small. If $0 < p < \frac{1}{2}$, then the first integral on the right-hand side of (3) approaches zero faster than any power of t as t approaches zero. In addition, we have the following proposition.

<u>Proposition 1:</u> Suppose that $0 \le q \le \frac{1}{2}$ and $0 \le p \le \frac{1}{2} - q$. If t and $\frac{1+r}{s}$ are sufficiently small, or if t is sufficiently small and $0 \le p < \frac{1}{2} - q$, then

$$\int_{[R_1(t)\Delta R_1(0)]\cap M(t)} |G| \le K\,|t|^{1-q},$$

where the constant K is independent of q.

<u>Proof:</u> It follows from Lemma 3 that if t and $\frac{1+r}{s}$ are sufficiently small, or if t is sufficiently small and $0 \le p < \frac{1}{2} - q$, then

$$\int_{[R_1(t)\Delta R_1(0)]\cap M_{p,q,r,s}(t)} |G| \le \int_{|\tau|\le|t|} \int_{S(0)\cap M_{p,q,2r,\frac{s}{2}}(t)} |G(y(x,\tau))| \frac{|\frac{\partial}{\partial t}F(y(x,\tau),\tau)|}{|\nabla F(y(x,\tau),\tau)|}\, dS(\tau)\,d\tau$$

$$\leq \int_{|\tau|\leq|t|} \left\{ \int_{S(\tau)\cap M_{p,q,3r,\frac{s}{3}(t)}} |G(x)| \frac{\left|\frac{\partial}{\partial t} F(x,\tau)\right|}{|\nabla F(x,\tau)|} dS(\tau) \right\} d\tau,$$

where $dS(\tau)$ is the element of surface area on $S(\tau)$. (See Spivak [3] for a discussion of integration on manifolds.) It is easily seen that

$$\int_{S(\tau)\cap M_{p,q,3r,\frac{s}{3}(t)}} |G(x)| \left| \frac{\partial}{\partial t} F(x,\tau) \right| dS(\tau)$$

is bounded for $|\tau| \leq |t|$ uniformly for t near zero. Furthermore, for fixed s, Lemma 1 implies that $|\nabla F(x,\tau)| \geq K |t|^q$ for $x \in S(\tau)\cap M_{p,q,3r,\frac{s}{3}(t)}$ and $|\tau| \leq |t|$. Consequently,

$$\int_{[R_1(t)\Delta R_1(0)]\cap M_{p,q,r,s}(t)} |G| \leq K |t|^{1-q}.$$

It is easily verified that K is independent of q, and the proof is complete.

From Proposition 1 and the preceding remarks, one sees that if $0 < p < \frac{1}{2}, q = 0$, and s is sufficiently small, then

$$|P_e(t) - P_e(0)| \leq K |t|$$

for small t, and the theorem is proved in this case.

Case 2: If $0 < p$, then, as before, the first integral on the right-hand side of (2) approaches zero faster than any power of t as t approaches zero. In addition, Proposition 1 remains valid. Thus, if $0 \le q < \frac{1}{2}$, $0 < p \le \frac{1}{2} - q$ and $\frac{1+r}{s}$ is sufficiently small, then the second integral on the right-hand side of (2) is bounded by $K|t|^{1-q}$ as t approaches zero, where the constant K is independent of q. Of course, $S(\tau) \cap N(t) \neq \emptyset$ for all τ, $|\tau| \le |t|$, and we need the following proposition.

Proposition 2: There exists a constant K, independent of q, for which

$$\int_{N(t)} |G| \le K\, |t|^{m\,q} \,,$$

where m is the number of non-zero eigenvalues (counting multiplicity) of $\Sigma_2^{-1}(0) - \Sigma_1^{-1}(0)$.

Proof: We have $F(x,t) = x^T A(t)x + B(t)x + C(t)$, where $A(t) = \frac{1}{2}(\Sigma_2(t)^{-1} - \Sigma_1(t)^{-1})$, and $\nabla F(x,t) = 2A(t)x + B(t)^T$. Denoting the ball of radius ρ about the origin in $\mathbb{R}^n$ by B_ρ, one sees that, if x_0 is any solution of $\nabla F(x_0,0) = 0$, then

$$N(t) \subseteq \{x_0 + y + z : y \in \mathcal{N}(A(0)),\ z \in \mathcal{N}(A(0))^{\perp} \cap B_{s|t|^q}\}.$$

Now the dimension of $\mathcal{N}(A(0))^{\perp}$ is equal to the number of non-zero eigenvalues (counting multiplicity) of $A(0) = \frac{1}{2}(\Sigma_2^{-1}(0) - \Sigma_1^{-1}(0))$. Denoting this dimension by m, we obtain (with a slight abuse of notation)

$$\int_{N(t)} |G| \leq \int_{\mathcal{N}(A(0))^{\perp} \cap B_{s|t|^q}} \left\{ \int_{\mathcal{N}(A(0))} |G(x_0 + y + z)|\,dy \right\} dz$$

$$\leq K\ |t|^{mq}$$

for an appropriate constant K, independent of q, and the proof is complete.

From the above discussion, one sees that the best rate of decay of the right-hand side of (2) is obtained by choosing p and q such that $0 \leq q < \frac{1}{2}$, $0 < p \leq \frac{1}{2} - q$, and $1 - q = mq$. Since $m > 1$, a compatible choice is $q = \frac{1}{m+1}$ and $0 < p \leq \frac{m-1}{2(m+1)}$. This yields the desired inequality

$$|P_e(t) - P_e(0)| \leq K\ |t|^{\frac{m}{m+1}}$$

Case 3: In this case, one sees that $S(0)$ is an $(n-1)$-dimensional hyperplane in $\mathbb{R}^n$ and that $\nabla F(x,0) = 0$ if and only if $x \in S(0)$. By performing a translation of co-ordinates followed by a unitary transformation on $\mathbb{R}^n$ if necessary, we may assume that $S(0) = \{x = (0,x_2,\ldots,x_n)^T \in \mathbb{R}^n\}$. Then $F(x,0) = x^T Ax$, where A has a non-zero entry in the upper left-hand corner and only zero entries elsewhere, i.e., $F(x,0) = \lambda x_1^2$. We will use the sets $L_{p,r}(t)$, $M_{p,q,r,s}(t)$, and $N_{p,q,r,s}(t)$ in the following with $p = 0$ and $q = \frac{1}{2}$, and we set

$$K_c(t) = \{x = (x_1,\ldots,x_n)^T \in \mathbb{R}^n : |x_1| \leq c\sqrt{|t|}\ \sqrt{x_2^2 + \ldots + x_n^2}\}.$$

Proposition 3: If $\frac{1+r}{s}$ is sufficiently small and c is sufficiently large, then $R_1(t)\Delta R_1(0) \subseteq N(t) \cup K_c(t)$ whenever t is small.

This proposition follows from the two lemmas below.

Lemma 5: If $\frac{1+r}{s}$ is sufficiently small, then $M(t)\cap[R_1(t)\Delta R_1(0)] = \emptyset$ whenever t is small.

Proof: In $M(t)$, $F(x,0) \geq \lambda s^2|t|$ and $|F_t(x,\tau)| \leq K(1+r)^2$ for $|\tau| \leq |t|$, where the constant K is independent of $r,s,$ and t for small t. So, for $|\tau| \leq |t|$,

$$|F(x,\tau)| \geq |\lambda s^2 - K(1+r)^2||t|$$

in $M(t)$ whenever t is small. If $\frac{1+r}{s} < \sqrt{\frac{\lambda}{K}}$, one sees that $F(x,\tau) \neq 0$ in $M(t)$ for $|\tau| \leq |t|$. Since $R_1(t)\Delta R_1(0) \subseteq \bigcup_{|\tau|\leq|t|} S(\tau)$, the lemma follows.

Lemma 6: Suppose that r is given. If C is sufficiently large, then $L(t)\cap[R_1(t)\Delta R_1(0)] \subseteq K_c(t)$ whenever t is small.

Proof: In $L(t)$, $F(x,0) = \lambda x_1^2$ and $|F_t(x,\tau)| \leq K|x|^2$ for $|\tau| \leq |t|$, where the constant K depends on r but is independent of t for small t. So, for $|\tau| \leq |t|$, one has

$$|F(x,\tau)| \geq (\lambda - K|t|)x_1^2 - K|t|(\sum_{i=2}^{n} x_i^2)$$

in $L(t)$. If $x \in L(t)-K_c(t)$, then

$$\frac{x_1^2}{c^2|t|} > \sum_{i=2}^{n} x_i^2$$

and

$$|F(x,\tau)| \geq [\lambda - K|t| - \frac{K}{c^2}]x_1^2.$$

If c is sufficiently large, then the right-hand side is positive for small t. Consequently,

$$[L(t)-K_c(t)] \cap [R_1(t) \Delta R_1(0)] \subseteq [L(t)-K_c(t)] \cap [\bigcup_{|\tau| \leq |t|} S(\tau)] = \emptyset,$$

and the proof is complete.

From Proposition 3, one sees that if $\frac{1+r}{S}$ is sufficiently small and c is sufficiently large, then

$$|P_e(t) - P_e(0)| \leq \int_{K_c(t)} |G| + \int_{N(t)} |G|$$

for small t. The two integrals on the right-hand side are easily seen to be bounded by $K\sqrt{|t|}$, and the proof of the theorem is complete.

REFERENCES

1. T. W. Anderson, An Introduction to Multivariate Statistical Analysis, John Wiley and Sons, Inc., New York, 1958.
2. F. John, Ordinary Differential Equations, New York University (Courant Institute of Mathematical Sciences), New York, 1966.
3. M. Spivak, Calculus on Manifolds, W. A. Benjamin, Inc., Menlo Park, California, 1965.
4. J. Walton, A Mathematical Question Incident to the Signature Extension Problem of Bayesian Classification, NASA-ASEE Faculty Fellow Report, Summer, 1974.