Asymptotic properties of sequences of positive kernels

Patricia Sullivan Conn
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Statistics and Probability Commons

Recommended Citation
https://lib.dr.iastate.edu/rtd/3564

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
CONN, Patricia Sullivan, 1935-
ASYMPTOTIC PROPERTIES OF SEQUENCES
OF POSITIVE KERNELS.

Iowa State University, Ph.D., 1969
Statistics

University Microfilms, Inc., Ann Arbor, Michigan
ASYMPTOTIC PROPERTIES OF SEQUENCES OF POSITIVE KERNELS

by

Patricia Sullivan Conn

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
Ames, Iowa

1969
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>I. INTRODUCTION</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>II. SEQUENCES OF POSITIVE KERNELS</strong></td>
<td>10</td>
</tr>
<tr>
<td>A. Convergence Theorems</td>
<td>10</td>
</tr>
<tr>
<td>B. Rates of Convergence and Error Bounds</td>
<td>36</td>
</tr>
<tr>
<td>C. Asymptotic Tendency to Kernels of Rank One</td>
<td>48</td>
</tr>
<tr>
<td>D. Convergence Theorems for Characteristic Roots and Functions</td>
<td>55</td>
</tr>
<tr>
<td><strong>III. APPLICATIONS TO STATISTICS</strong></td>
<td>63</td>
</tr>
<tr>
<td>A. Sequential Probability Ratio Tests</td>
<td>63</td>
</tr>
<tr>
<td>B. Generalized Sequential Probability Ratio Tests</td>
<td>75</td>
</tr>
<tr>
<td><strong>IV. LITERATURE CITED</strong></td>
<td>87</td>
</tr>
<tr>
<td><strong>V. ACKNOWLEDGEMENTS</strong></td>
<td>90</td>
</tr>
</tbody>
</table>
I. INTRODUCTION

The major motivation for this thesis is the study of the asymptotic properties of non-stationary random walks between two absorbing barriers. Random walks are special cases of discrete time Markov processes. These in turn may be studied in terms of the iteration of non-negative matrices and kernels. It is within this larger framework that asymptotic behavior is investigated.

A stochastic process \( \{X_n: n = 0, 1, 2, \ldots\} \) is called a Markov process if

\[
\Pr[\{X_n(w) \leq \lambda/X_0, X_1, \ldots, X_{n-1}\}] = \Pr[\{X_n(w) \leq \lambda/X_{n-1}\}].
\]

When the state space, the set \( X \) of values that the random variables \( X_n \) can assume, is finite or denumerably infinite, the process is called a Markov chain. This work considers only finite chains and those discrete time Markov processes having some finite interval \([a, b]\) of real numbers as state space.

An \( N \) state Markov chain is defined by specifying an initial probability vector \( p_0 = (p_{01}, p_{02}, \ldots, p_{0N}) \) and a sequence \( \{P_n\} = \{p_{ij}^n\} \) of non-negative transition matrices satisfying the condition

\[
\sum_{i=1}^{N} p_{ij}^n = 1 \quad \text{(1.1)}
\]
The probability distribution for $X_n$ is given by

$$p_n = p_np_{n-1} \quad n = 1, 2, \ldots \quad (1.2)$$

Similarly, a large class of discrete time Markov processes with an interval $[a, b]$ as state space can be defined by specifying an initial density function $p_0(x)$ and a sequence $\{P_n(x,y)\}$ of non-negative, integrable kernels on $[a, b]^2$ with the property

$$\int_a^b P_n(x,y)dx = 1 \quad (1.3)$$

These considerations of ergodic behavior are extended to the case of "dishonest" discrete time Markov processes characterized by the Generalizations 1.5 and 1.7 below of Equations 1.1 and 1.3. In particular, random walk between two absorbing barriers is studied from the point of view of the asymptotic behavior of the sequence of distributions over the non-absorbing states, conditional on no prior absorption. These probability distributions are given for discrete walks with $N$ non-absorbing states by

$$p'_n = [p_{n1}/(p_n, 1), p_{n2}/(p_n, 1), \ldots, p_{nN}/(p_n, 1)] \quad (1.4)$$

where $p_n = p_np_{n-1}$ and $1 = (1, 1, \ldots, 1)$; the matrices $P_n$ satisfy
\[ \sum_{i=1}^{\infty} p_{ij}^{n} \leq 1 \quad n = 1, 2, \ldots \quad (1.5) \]

Similarly, the conditional distribution at time \( n \) for a continuous walk with an interval \([a, b]\) as its set of non absorbing states is given by

\[ p_{n}^{\prime}(x) = \frac{p_{n}(x)}{(p_{n}(x), l(x))} \quad (1.6) \]

where \( p_{n}(x) = \int_{a}^{b} P_{n}(x, y) p_{n-1}(y)dy \) and \( l(x) = 1 \) for all \( x \) in \([a, b]\).

The kernels \( P_{n}(x, y) \) satisfy

\[ \int_{a}^{b} P_{n}(x, y)dx \leq 1 \quad n = 1, 2, \ldots \quad (1.7) \]

Only those processes for which density kernels exist will be considered. The density function for \( X_{n} \) is, then, given by

\[ p_{n}(x) = \int_{a}^{b} P_{n}(x, y) p_{n-1}(y)dy \quad n = 1, 2, \ldots \quad (1.8) \]

A process which is such that \( P_{n} = P \) for all \( n \) is called homogeneous or stationary. The asymptotic behavior of \( p_{n} \) as \( n \) tends to infinity has been studied extensively for stationary processes. In this case Equations 1.2 and 1.8 have, respectively, the forms
\[ r_n = p^n r_0 \]

and

\[ p_n(x) = \int_a^b p^{(n)}(x,y) \ p_0(y) \, dy \]

where \( p^n \) represents the \( n \)th power of the transition matrix \( P \)
and \( p^{(n)}(x,y) \) the \( n \)th iterate of the kernel \( P(x,y) \). Conditions
on the matrix \( P \) which insure that the sequence \( \{p_n\} \) converges
to a limit independent of \( p_0 \) are given by many authors, e.g.
Doob (1953) and Gantmacher (1959b); Doob's conditions apply as
well to the case where the state space is a finite interval.

One problem considered here is that of generalizing these
ergodic results to processes which are not stationary. It is
simplified if consideration is limited to those whose transition
matrices or kernels are positive. Positivity, while
overly restrictive for this purpose, does eliminate the
necessity of coping with decomposability and periodicity. Two
types of limiting behavior will be distinguished for non-
stationary processes in accordance with definitions stated by
Hajnal (1956). A process will be called "ergodic in the
strong sense" if the sequence \( p_n \) tends to a limit independent
of \( p_0 \). The sequence \( \{p_n\} \) may fail to converge but still lose
all trace of the initial distribution; such a process will be
called "ergodic in the weak sense."
Much of the work in the body of this thesis is based on a special property of the spectra of positive valued matrices and kernels. The following theorem, stated by Gantmacher (1959b), was proved by Perron in 1907: A positive matrix A always has a real and positive characteristic value \( r \) which is a simple root of the characteristic equation and exceeds the moduli of all the other characteristic values. To this maximal characteristic value \( r \) there corresponds a characteristic vector \( z \) with positive coordinates.

Jentzsch, in 1912, obtained the same result for positive square summable kernels which fulfill certain continuity conditions. A statement of Jentzsch's theorem appears in Schmeidler (1950). A similar theorem is given by Sarymsakov (1949), who requires that the kernels be continuous and non-negative and that for any \((x,t)\) in \([a,b]^2\) there exist iterates \( K^{(m)} \) and \( K^{(n)} \) such that \( K^{(m)}(x,t) > 0 \) and \( K^{(n)}(t,x) > 0 \). Sarymsakov applies his result to show convergence of the sequence of iterates of a continuous kernel which is the transition probability density kernel for a Markov process.

Birkhoff (1957) utilizes projective metrics to obtain various extensions of Jentzsch's theorem among which are results on the limiting behavior of the iterates of a uniformly positive, bounded linear transformation \( P \) of a vector lattice into itself. Integral operators with positive valued kernels are included as a special case.
The properties of integral operators with primitive kernels are discussed by Harris (1964), who proves the following useful theorem: Let \( M(x,A) = \int_A m(x,y) dV(y) \). If there exists an integer \( n_o \) such that \( 0 < a \leq m_{n_0}(x,y) \leq b < \infty \), then \( m(x,y) \) has a simple positive characteristic value \( \rho \) larger in magnitude than any other characteristic value and corresponding to right and left characteristic functions \( u \) and \( v \) which are bounded and uniformly positive. Furthermore, if \( u \) and \( v \) are normalized so that \( (u,v) = 1 \), then
\[
m^n(x,y) = \rho^n u(x)v(y)[1 + o(\Delta^n)], \quad 0 < \Delta < 1, \quad n \to \infty
\]
where the bound \( \Delta \) can be taken independently of \( x \) and \( y \).

The unique, simple characteristic value which exceeds all others in modulus, the existence of which is assured for positive matrices and kernels by the theorems quoted above, will henceforth be referred to as the "dominant" characteristic value.

Ergodic results for stationary Markov chains are, as previously indicated, widely available in the literature. Research on the ergodic properties of non-stationary processes has been largely directed at establishing conditions which will insure asymptotic independence of the initial distribution and/or asymptotic normality of the standardized sums.

Sarymsakov (1953) gave the following sufficient condition for weak ergodicity in a non-stationary Markov chain. Let
Let \( p^k = \left( \frac{c^k}{\sum_{i,j=1}^{N} a_{ij}^k} \right) \) be the \( k \)th transition matrix and let \( C \) be the class of \( N \) dimensional primitive stochastic matrices \( A \) with the property that \( AB \) is primitive whenever \( B \) is. Then if it is possible to find a sequence of matrices in \( C \), \( \{ g^k_{ij} \} \), \( k \geq 1 \), such that for some \( c > 0 \) and all \( i, j, k \) with \( g^k_{ij} > 0 \) the inequality \( p^k_{ij} > c \) is satisfied, the chain is weakly ergodic. Further results on sufficient conditions for weak ergodicity and asymptotic normality of standardized sums have been obtained by Sarymsakov (1961).

The central limit theorem and conditions insuring weak ergodicity for non-stationary Markov chains have also been studied by Dobrushin (1956). It should be noted that Dobrushin and several others use the term "ergodic" to describe a chain which is here called "ergodic in the weak sense" and the term "strongly ergodic" to describe a sequence \( \{ P_i \} \) of transition functions with the property that any sequence \( \{ P_i \} \) containing \( \{ P_i \} \) as a subsequence is, according to the definitions used here, "ergodic in the weak sense." Dobrushin does not consider the property which is described in this work as "ergodic in the strong sense."

Hajnal (1956) gives sufficient conditions for both strong and weak ergodicity for finite, non-stationary Markov chains which are neither periodic nor decomposable. Hajnal does not require positivity, but the sufficient conditions for strong
ergodicity which he obtains are more restrictive than those given here.

Section A of Chapter II provides necessary and sufficient conditions for strong ergodicity in the case of non-stationary Markov processes with finite or finite interval state spaces and uniformly positive transition matrices or kernels. Applied to a random walk between absorbing barriers, the results of this section imply that the sequence of conditional densities converges if and only if the sequence of right characteristic functions associated with the dominant characteristic values of the transition kernels converges.

Section B examines convergence rates for the sequence of conditional densities to which Theorem 2.1.1 applies and relates the rate of convergence of this sequence to that of the sequence of right characteristic functions associated with the dominant roots of the transition kernels. Section C applied to Markov processes gives the known result that positivity implies weak ergodicity. Applied to random walks between absorbing barriers, it states that initial information is asymptotically lost, although the sequence of conditional densities may fail to converge, if the sequence of left characteristic functions associated with the dominant roots of the transition kernels converges in an appropriate manner. It is established in Section D that the convergence of a sequence of
positive kernels implies the convergence of the sequences of their dominant roots and associated characteristic functions. Some applications to sequential analysis are discussed in Chapter III.
II. SEQUENCES OF POSITIVE KERNELS

A. Convergence Theorems

Throughout this chapter attention is focused on a sequence 
\[
\{M_n(x,y)\}
\]
of positive, measurable kernels on \([a,b]^2\) and on a 
sequence \(\{f_n(x)\}\) of functions on \([a,b]\), beginning with an 
arbitrary non-negative function \(f_0(x)\), \(0 < \int_a^b f_0(x) < \infty\), and 
generated iteratively according to
\[
f_{n+1}(x) = \int_a^b M_{n+1}(x,y)f_n(y)\,dy \quad n = 0, 1, 2, \ldots \quad (2.1.1)
\]

It will be assumed that there exist numbers \(\overline{m}\) and \(\overline{M}\) such that
\[
0 < \overline{m} \leq M_n(x,y) \leq \overline{M} < \infty \quad (2.1.2)
\]

Let \(\lambda_n\) denote the dominant characteristic root of the kernel 
\(M_n(x,y)\) and let \(\phi_n(x)\) and \(\psi_n(x)\) be the associated right and 
left characteristic functions, respectively, chosen positive 
on \([a,b]\) and normed so that
\[
\int_a^b \phi_n(x)\,dx = \int_a^b \phi_n(x)\psi_n(x)\,dx = 1 \quad (2.1.3)
\]

Bounds 2.1.4, 2.1.5 and 2.1.6 below follow from Equations 
2.1.2 and 2.1.3.

\[
\lambda_n \phi_n(x) = \int_a^b M_n(x,y)\phi_n(y)\,dy \leq \overline{M} \int_a^b \psi_n(x)\,dx = \overline{M}
\]
and

$$\lambda_n \psi_n(x) = \int_a^b M_n(x,y) \psi_n(y) \, dy \geq \frac{\bar{m}}{\bar{M}} \int_a^b \psi_n(x) \, dx = \bar{m}$$

Thus

$$\bar{m} \leq \lambda_n \psi_n(x) \leq \bar{M}$$

Integrating over the interval \([a,b]\),

$$\bar{m}(b-a) \leq \lambda_n \leq \bar{M}(b-a) \quad (2.1.4)$$

Then

$$\frac{\bar{m}}{\bar{M}(b-a)} \leq \frac{\bar{m}}{\lambda_n} \leq \frac{\lambda_n}{\bar{M}} \leq \frac{\bar{M}}{\bar{m}(b-a)} \quad (2.1.5)$$

Also

$$\bar{m} \int_a^b \psi_n(x) \, dx \leq \lambda_n \psi_n(y) \leq \bar{M} \int_a^b \psi_n(x) \, dx$$

and it follows from Equations 2.1.3 and 2.1.5 that

$$\frac{\lambda_n}{\bar{M}} \leq \int_a^b \psi_n(x) \, dx \leq \frac{\lambda_n}{\bar{m}}$$

Thus

$$\frac{\bar{m}}{\bar{M}} \leq \psi_n(y) \leq \frac{\bar{M}}{\bar{m}} \quad (2.1.6)$$
The restrictions on \( f_n \) and the positivity of the kernels insure that \( \int_a^b f_n(x)dx > 0 \); thus it is possible to define

\[
f'_n(x) = \frac{f_n(x)}{\int_a^b f_n(x)dx} \tag{2.1.7}
\]

Theorem 2.1.1 relates the convergence of the sequence \( \{f'_n(x)\} \) to that of the sequence \( \{g_n(x)\} \). Two lemmas precede its statement and proof. Lemma 1 establishes that K fold convolutions of uniformly bounded stochastic kernels tend at a geometric rate toward a kernel which is a function of only one argument. Its proof is a direct generalization of that given by Doob (1953) for powers of stochastic matrices. Lemma 2 relates the K fold convolution of uniformly positive kernels whose right characteristic functions tend pairwise to each other to the K fold convolution of the stochastic kernels obtained by the well known Transformation 2.1.9 below.

**Lemma 2.1.1**: Let \( P = \{P(x,y) : a < x, y < b\}^3 \)

- \( H - 1 \) \( 0 < \delta \leq P(x,y) \leq \Delta \)
- \( H - 2 \) \( P(x,y) \) is measurable on \([a,b]^2\).
- \( H - 3 \) \( \int_a^b P(x,y)dy = 1 \)
Let $\Sigma(k)$ denote the collection of all $k$-term sequences of elements of $P$. For $\sigma(k) \in \Sigma(k)$,

$$\pi_{\sigma(k)}(x,y)$$

$$= \int_a^b \cdots \int_a^b P_k(x,z_{k-1}) P_{k-1}(z_{k-1},z_{k-2}) \cdots P_1(z_1,y) dz_1 dz_2 \cdots dz_{k-1}$$

Under the hypotheses above

$$\sup x \pi_{\sigma(k)}(x,y) - \inf x \pi_{\sigma(k)}(x,y) \leq [1 - (b-a)\delta]^{k-1}(\Delta-\delta) \quad (2.1.8)$$

**Proof:** By induction. For $k = 1$ the result follows directly from $H - 1$.

$$\sup x \pi_{\sigma(k+1)}(x,y) - \inf x \pi_{\sigma(k+1)}(x,y)$$

$$= \sup_{\alpha,\beta} \left\{ \pi_{\sigma(k+1)}(\alpha,y) - \pi_{\sigma(k+1)}(\beta,y) \right\}$$

$$= \sup_{\alpha,\beta} \left\{ \int_a^b [P_{k+1}(\alpha,z) - P_{k+1}(\beta,z)] \pi_{\sigma(k)}(z,y) dz \right\}$$

Let $S = \left\{ z \in [a,b] : P_{k+1}(\alpha,z) \geq P_{k+1}(\beta,z) \right\}$. If $\overline{S}$ denotes the complement of $S$ relative to $[a,b]$ and $\mu$ is Lebesgue measure,

$$\int_S [P_{k+1}(\alpha,z) - P_{k+1}(\beta,z)] dz + \int_{\overline{S}} [P_{k+1}(\alpha,z) - P_{k+1}(\beta,z)] dz$$

$$= 1 - 1 = 0$$
and

\[
\int_S \left[ P_{k+1}(a,z) - P_{k+1}(\beta, z) \right] \, dz \leq 1 - \delta \mu(S) - \delta \mu(S) = 1 - \delta(b-a)
\]

Then

\[
\sup_x \pi_{\sigma(k+1)}(x,y) - \inf_x \pi_{\sigma(k+1)}(x,y)
\]

\[
= \sup_{a,\beta} \left\{ \int_S \left[ P_{k+1}(a,z) - P_{k+1}(\beta, z) \right] \pi_{\sigma(k)}(z,y) \, dz \right. \\
+ \left. \int_S \left[ P_{k+1}(a,z) - P_{k+1}(\beta, z) \right] \pi_{\sigma(k)}(z,y) \, dz \right\}
\]

\[
\leq \sup_{a,\beta} \left\{ \sup_z \pi_{\sigma(k)}(z,y) \int_S \left[ P_{k+1}(a,z) - P_{k+1}(\beta, z) \right] \, dz \\
+ \inf_z \pi_{\sigma(k)}(z,y) \int_S \left[ P_{k+1}(a,z) - P_{k+1}(\beta, z) \right] \, dz \right\}
\]

\[
= \sup_{a,\beta} \left\{ \sup_z \pi_{\sigma(k)}(z,y) - \inf_z \pi_{\sigma(k)}(z,y) \right\} \int_S \left[ P_{k+1}(a,z) - P_{k+1}(\beta, z) \right] \, dz \\
\leq \left[ 1 - (b - a)\delta \right]^{k-1} (\Delta - \delta) \left[ 1 - (b - a)\delta \right]
\]

\[
= \left[ 1 - (b - a)\delta \right]^k (\Delta - \delta)
\]
It will be convenient to use the following notation to represent the n-fold convolution of n arbitrary kernels defined on the square \([a,b] \times [a,b]\).

\[
K_n K_{n-1} \cdots K_1(x,y) \equiv \int_a^b \cdots \int_a^b K_n(x,z_{n-1})K_{n-1}(z_{n-1},z_{n-2})\cdots K_1(z_1,y)dz_1dz_2\cdots dz_{n-1}
\]

**Lemma 2.1.2:** For each of the kernels \(M_n(x,y)\) in Equation 2.1.1, define

\[
P_n(x,y) = \frac{M_n(x,y)\phi_n(y)}{\lambda_n \phi_n(x)} \quad (2.1.9)
\]

Note that \(P_n(x,y)\) satisfies the hypotheses of Lemma 1 with

\[
\delta = \left(\frac{m}{M}\right)^2 \left(\frac{1}{b-a}\right) \quad \text{and} \quad \Delta = \left(\frac{M}{m}\right)^2 \left(\frac{1}{b-a}\right)
\]

If

\[
\int_a^b \left| \phi_n(x) - \phi_{n+1}(x) \right| dx \xrightarrow{n} 0,
\]

then for any finite K,

\[
\int_a^b \left| \sum_{n+k} M_{n+k-1} \cdots M_{n+k-1}(x,y)\phi_{n+k}(y) \right| dy \xrightarrow{n} 0
\]

**Proof:** By induction. For \(k = 1\), the result follows from Definition 2.1.9.
\[
\begin{align*}
&\int_a^b \frac{M_{n+1} \ldots M_{n+k+1}(x, z) \phi_{n+k+1}(z)}{\prod_{n+1} \lambda_i \phi_{n+k+1}(x)} \, dz \\
&= \int_a^b \frac{M_{n+k+1}(x, y)M_{n+k} \ldots M_{n+1}(y, z) \phi_{n+k+1}(z)}{\prod_{n+k} \lambda_i \phi_{n+k+1}(x)} \, dz \\
&= \int_a^b \int_a^b P_{n+k+1}(x, y) \frac{M_{n+k} \ldots M_{n+1}(y, z) \phi_{n+k+1}(z)}{\prod_{n+k} \lambda_i \phi_{n+k+1}(y)} \, dy \\
&- \int_a^b P_{n+k+1}(x, y) \frac{M_{n+k} \ldots M_{n+1}(y, z) \phi_{n+k}(z)}{\prod_{n+k} \lambda_i \phi_{n+k}(y)} \, dy \\
&+ \int_a^b P_{n+k+1}(x, y) \frac{M_{n+k} \ldots M_{n+1}(y, z) \phi_{n+k}(z)}{\prod_{n+k} \lambda_i \phi_{n+k}(y)} \, dy \\
&- \int_a^b P_{n+k+1}(x, y)P_{n+k} \ldots P_{n+1}(y, z) \, dy \, dz
\end{align*}
\]
Now

\[
\frac{\phi_{n+k+1}(z) - \phi_{n+k}(z)}{\phi_{n+k+1}(y) - \phi_{n+k}(y)} \leq \frac{1}{\phi_{n+k+1}(y)} \left| \phi_{n+k+1}(z) - \phi_{n+k}(z) \right|
\]

\[+ \frac{\phi_{n+k}(z)}{\phi_{n+k+1}(y) \phi_{n+k}(y)} \left| \phi_{n+k} - \phi_{n+k+1}(y) \right|\]

\[
\leq \frac{M(b-a)}{\tilde{M}} \left| \phi_{n+k+1}(z) - \phi_{n+k}(z) \right|
\]

\[+ \frac{M^3}{\tilde{M}^3} (b-a) \left| \phi_{n+k}(y) - \phi_{n+k+1}(y) \right|\]
where the last inequality follows from the Bounds 2.1.5 on the characteristic functions. Then

\[
\int_a^b \left| \frac{\phi_{n+k+1}(z)}{\phi_{n+k}(y)} - \frac{\phi_{n+k}(z)}{\phi_{n+k}(y)} \right| dy \, dz \leq \frac{M(b-a)^2}{n} \int_a^b \left| \frac{\phi_{n+k+1}(z)}{\phi_{n+k}(y)} - \frac{\phi_{n+k}(z)}{\phi_{n+k}(y)} \right| dz + \frac{M^2(b-a)^2}{m^3} \int_a^b \left| \frac{\phi_{n+k}(y)}{\phi_{n+k}(y)} - \frac{\phi_{n+k+1}(y)}{\phi_{n+k+1}(y)} \right| dy \tag{2.1.11}
\]

and these two integrals tend with \( n \) to zero by hypothesis.

That the right hand integral in Equation 2.1.10 tends also to zero with \( n \) follows from the induction hypothesis and the uniform boundedness of the integrand.

It is now possible to prove the following theorem.

**Theorem 2.1.1:** Let \( \{f_n(x)\} \) be defined as in Equation 2.1.1 with the measurable kernels \( K_n(x,y) \) restricted according to Equation 2.1.2. If \( \int_a^b \left| f_n(x) - f_{n+1}(x) \right| dx \to 0 \), then

\[
\left| f'_n(x) - f'_n(x) \right| \to 0.
\]

**Proof:** Choose \( \varepsilon > 0 \). Define \( P_n(x,y) \) as in Definition 2.1.9.

Let \( p^k_n(y) = \sup_x P_{n+k}P_{n+k-1} \cdots P_{n+1}(x,y) \)

It follows from Lemma 1 that

\[
\left| P_{n+k}P_{n+k-1} \cdots P_{n+1}(x,y) - p^n_k(y) \right| \leq [1 - (b-a)^2]^{k-1}(a-\delta)
\tag{2.1.12}
\]
and Inequality 2.1.12 does not depend on \( n \).

Choose \( k_o = \min \left\{ k : [1 - (b - a) \delta]^{k-1} (\Delta - \delta) < \frac{\varepsilon}{b - a} \right\} \) (2.1.13)

Then

\[
\int_a^b \left| P_{n+k_o} P_{n+k_o-1} \cdots P_{n+1}(x, y) - p_n(y)^k \right| dy < \varepsilon
\]

For \( k_o \) fixed as in Equation 2.1.13 and arbitrary \( n \)

\[
\frac{f_{n+k_o}(x)}{n+k_o} - \phi_{n+k_o}(x) \int_a^b \frac{p_n(y)f'_n(y)}{\phi_{n+k_o}(y)} dy
\]

\[
= \int_a^b \left[ \frac{\Pi^{M_{n+k_o}} M_{n+k_o-1} \cdots M_{n+1}(x, y)}{n+k_o} - \phi_{n+k_o}(x) \frac{p_n(y)}{\phi_{n+k_o}(y)} \right] f'_n(y) dy
\]

\[
+ \frac{\phi_{n+k_o}(k)}{\phi_{n+k_o}(y)} P_{n+k_o} P_{n+k_o-1} \cdots P_{n+1}(x, y) - \phi_{n+k_o}(x) \frac{p_n(y)}{\phi_{n+k_o}(y)} f'_n(y) dy
\]
\[
\sup_y f_n'(y) \leq \frac{\int_a^b \frac{\phi_n(x) n+k_0}{\phi_n(x+y) n+k_0} \left( \prod_{k=1}^n \lambda_k \phi_n(x) \right) dx}{\int_a^b M(y) f_n(y) dy} = \frac{\frac{M_n(y,z) f_n(z) dz}{\int_a^b f_n(y) dy}}{\int_a^b M_n(y,z) f_n(z) dz dy} \leq \frac{M}{m(b-a)} \quad (2.1.16)
\]
and

$$\frac{\hat{\phi}_n(x)}{\hat{\phi}_n(y)} \leq \frac{M}{m}^2 \quad (2.1.17)$$

Since Lemma 2.1.2 insures the existence of an $N(k_0, \varepsilon) > 0$, $n > N(k_0, \varepsilon)$ implies

$$\sum_{n+k_0}^{\infty} \sum_{n+k_0}^{\infty} \lambda_1 \phi_n(x) \leq P_{n+k_0} P_{n+k_0-l} \cdots P_{n+l} (x, y) \int dy < \varepsilon,$$

Inequality 2.1.15 implies

$$\left| \frac{f_{n+k_0}(x)}{n+k_0} - \phi_{n+k_0}(x) \right| \int_{a}^{b} \frac{p_n(y)f_n(y)}{\phi_{n+k_0}(y)} dy$$

\[ \leq \frac{|M|}{|m|} \left( \frac{1}{b-a} \right) (2\varepsilon) \]

It follows from Equations 2.1.5 and 2.1.7 and the definition of $k_0$ that

$$\int_{a}^{b} \frac{p_n(y)f_n(y)}{\phi_{n+k_0}(y)} dy \geq \frac{|m|}{|M|}^2$$
Thus for $n \geq N(k_o, \epsilon)$

$$|\frac{f_{n+k_o}(x)}{n+k_o} - \frac{p_{n+k_o}^k(y)f_n(y)}{\frac{1}{\lambda_1}\int_a^{n+1} \frac{p_{n+k_o}^1(y)}{\phi_{n+k_o}(y)} \, dy} - \phi_{n+k_o}(x)| < C \epsilon \quad (2.1.18)$$

and

$$|\frac{\int_a^b f_{n+k_o}(x) \, dx}{n+k_o} - \frac{b}{n+k_o} \int\frac{p_{n+k_o}^k(y)f_n(y)}{\frac{1}{\lambda_1}\int_a^{n+1} \frac{p_{n+k_o}^1(y)}{\phi_{n+k_o}(y)} \, dy} | < C(b-a) \epsilon \quad (2.1.19)$$

where $C = \left[\frac{M^5}{m} \right] \frac{2}{(b-a)}$

Then

$$\left| f'_{n+k_o}(x) - \phi_{n+k_o}(x) \right| = \left| \frac{f_{n+k_o}(x)}{p_{n+k_o}^1(y)} - \phi_{n+k_o}(x) \right|$$

$$= \left| \frac{f_{n+k_o}(x)}{\int_a^{n+1} f_{n+k_o}(x) \, dx} \right| - \left| \frac{\int_a^b f_{n+k_o}(x) \, dx}{n+k_o} \right| - \left| \frac{b}{n+k_o} \int\frac{p_{n+k_o}^k(y)f_n(y)}{\frac{1}{\lambda_1}\int_a^{n+1} \frac{p_{n+k_o}^1(y)}{\phi_{n+k_o}(y)} \, dy} \right|$$
\[
\begin{align*}
&\left(\int_a^b f_{n+k_0}^\prime(x) dx \right) - f_{n+k_0}^\prime(x) \\
&= \frac{1}{n+k_0} \left[ \prod_{i=1}^{n+1} \int_a^b p_n(y) f_n(y) \frac{\phi_{n+k_0}^\prime(y)}{\phi_{n+k_0}(y)} dy \right] \\
&\quad + \frac{f_{n+k_0}^\prime(x)}{n+k_0} \left(\int_a^b p_n(y) f_n(y) \frac{\phi_{n+k_0}^\prime(y)}{\phi_{n+k_0}(y)} dy \right) - \phi_{n+k_0}^\prime(x)
\end{align*}
\]

\[\leq C \varepsilon \sup_{x} (b-a) \sup_{x} f_{n+k_0}^\prime(x) + 1 \]

\[\leq C \left( \frac{N}{M} + 1 \right) \varepsilon\]

That is, for \( n \geq N(k_0, \varepsilon) + k_0 \)

\[ \left| f_n^\prime(x) - \phi_n(x) \right| < C \left( \frac{N}{M} + 1 \right) \varepsilon \] (2.1.20)

where \( \varepsilon > 0 \) was arbitrary.

It should be noted that the condition \( \int_a^b \phi_n(x) - \phi_{n+1}(x) dx \to 0 \) would be implied by the pointwise restriction

\[ \phi_n(x) - \phi_{n+1}(x) \to 0, \] and thus that

\[ \phi_n(x) - \phi_{n+1}(x) \to 0 \Rightarrow f_n^\prime(x) - \phi_n(x) \to 0 \] (2.1.21)
It is also clear that the Conclusion 2.1.21 can be stated

\[ |\phi_n(x) - \phi_{n+1}(x)| \xrightarrow{n \to \infty} |f_n(x) - f_{n+1}(x)| \xrightarrow{n \to \infty} 0 \quad (2.1.22) \]

Theorem 2.1.2 is a converse to Equation 2.1.22. The following lemma will be useful in its proof.

**Lemma 2.1.3**: Let \( M(x,y) \) be a measurable kernel on \([a,b]^2\), and suppose that there exist positive numbers \( \underline{m} \) and \( \underline{M} \) such that \( \underline{m} \leq M(x,y) \leq \underline{M} \). Let \( \lambda \) denote the dominant root of \( M(x,y) \) with associated right and left characteristic functions \( \phi \) and \( \psi \) normed so that \( \int_a^b \phi = \int_a^b \psi = 1 \). Then if \( M^k(x,y) \) denotes the \( k \)th iterate of \( M(x,y) \),

\[ \left| \frac{M^k(x,y)}{\lambda^k} - \phi(x)\psi(y) \right| \leq C(\frac{\underline{m}}{\underline{M}}, b-a) \left[ 1 - \left( \frac{\underline{m}}{\underline{M}} \right)^2 \right]^{k-1} \]

where \( C(\frac{\underline{m}}{\underline{M}}, b-a) = \left( \frac{\underline{M}}{\underline{m}} \right)^2 - \left( \frac{\underline{m}}{\underline{M}} \right)^2 \left( \frac{1}{b-a} \right) \)

**Proof**: Let \( P(x,y) = \frac{M(x,y)\phi(y)}{\lambda^k\phi(x)} \) \hspace{1cm} (2.1.23)

Then \( P(x,y) \) satisfies the hypothesis of Lemma 1 with

\[ \delta = \left( \frac{\underline{m}}{\underline{M}} \right)^2 \frac{1}{b-a} \quad \text{and} \quad \Delta = \left( \frac{\underline{M}}{\underline{m}} \right)^2 \frac{1}{b-a} \quad \text{and} \]

\[ \sup_x P^k(x,y) - \inf_x P^k(x,y) \leq [1 - (b-a)\delta]^{k-1}(\Delta - \delta) \quad (2.1.24) \]
Furthermore
\[ \sup_x P^{k+1}(x,y) \leq \sup_x P^k(x,y) \] \hspace{1cm} (2.1.25)
and
\[ \inf_x P^{k+1}(x,y) \geq \inf_x P^k(x,y) \] \hspace{1cm} (2.1.26)

Therefore the sequences \( \left\{ \sup_x P^k(x,y) \right\} \) and \( \left\{ \inf_x P^k(x,y) \right\} \) have a common limit \( p(y) \), and

\[ P^k(x,y) - p(y) \leq (\Delta - \delta)[1 - (b - a)\delta]^{k-1} \] \hspace{1cm} (2.1.27)

That the function \( p \) is a left characteristic function of the kernel \( P(x,y) \) associated with the dominant root \( 1 \) follows from Equation 2.1.27 which states that \( \lim_{k \to \infty} P^k(x,y) = p(y) \)

Writing \( P^{k+1}(x,y) = \int_a^b P^k(x,z) P(z,y) \, dz \)

and taking limits with respect to \( k \) on both sides,

\[ p(y) = \int_a^b p(z) P(z,y) \, dz \]

Now \( M(x,y) = \frac{\lambda \hat{\phi}(x) P(x,y)}{\hat{\phi}(y)} \), and \( \hat{\phi} \) is a left characteristic function of \( M(x,y) \) associated with \( \lambda \).
Thus \( \psi(y) = \frac{1}{\lambda} \int_{a}^{b} M(x,y)\psi(x)dx = \int_{a}^{b} \frac{\psi(x)\phi(x)P(x,y)dx}{\phi(y)} \)

so that \( \psi(y)\phi(y) = \int_{a}^{b} \psi(x)\phi(x)P(x,y)dx \)

Since \( P(x,y) \) is positive, its dominant root 1 is simple.

Therefore \( p(y) = \psi(y)\phi(y) \) \( \text{(2.1.28)} \)

Combining Equations 2.1.23, 2.1.27 and 2.1.28,

\[
\left| \frac{M^k(x,y)\phi(y)}{\phi(x)} - \psi(y)\phi(y) \right| \leq (\Delta - \delta)[1 - (b-a)\delta]^{k-1}
\]

Thus \( \left| \frac{M^k(x,y)}{\phi(x)} - \phi(x)\psi(y) \right| \leq \frac{\phi(x)}{\phi(y)} (\Delta - \delta)[1 - (b-a)\delta]^{k-1} \)

\[
= C \left( \frac{\bar{m}}{\bar{M}}, b - a \right) \left[ 1 - \left( \frac{\bar{m}}{\bar{M}} \right)^2 \right]^{k-1} \text{ (2.1.29)}
\]

Result 2.1.27 is indicated by Doob (1953) and used in conjunction with Equation 2.1.23 by Harris (1964) to obtain a somewhat less specific form of Inequality 2.1.29. The additional calculations in Lemma 2.1.3 insure the uniformity
in the approach to its limit of any \( \frac{M_k(x, y)}{\lambda_n} \), \( \bar{m} \leq M_n(x, y) \leq \bar{M} \).

Equation 2.1.1 can be rewritten

\[
f_{n+1}^\prime(x) = \rho_{n+1} \int_a^b M_{n+1}(x, y)f_n(y)dy \quad n = 0, 1, 2 \ldots \quad (2.1.30)
\]

where \( f_n(x) \) is defined in Equation 2.1.7

\[
\rho_{n+1} = \frac{\int_a^b f_n(x)dx}{\int_a^b f_{n+1}(x)dx} \quad (2.1.31)
\]

Note that

\[
\frac{1}{\bar{M}(b-a)} \leq \rho_n \leq \frac{1}{\bar{m}(b-a)} \quad (2.1.32)
\]

**Theorem 2.1.2:** Let the kernels \( M_n(x, y) \) be restricted as in Equation 2.1.2. If \( \left| f_n'(x) - f_{n+1}'(x) \right| \xrightarrow{n} 0 \), then

\[
\left| f_n(x) - \hat{f}_n(x) \right| \xrightarrow{n} 0.
\]

It is convenient to begin by proving **Lemma 2.1.4:** Under the hypotheses of Theorem 2.1.2,

\[
\left| \rho_n \int_a^b M_n(x, y)f_n(y)dy - f_n(x) \right| \xrightarrow{n} 0 \text{ for any finite } k.
\]
Proof of Lemma 2.1.4: By induction. For $k = 1,$

$$\left| \int_{a}^{b} M_{n}(x,y) f_{n}'(y) dy - f_{n}'(x) \right| = \rho_{n} \int_{a}^{b} M_{n}(x,y) \left( f_{n}'(y) - f_{n-1}'(y) \right) dy$$

$$\leq \frac{M}{M(b-a)} \int_{a}^{b} \left| f_{n}'(y) - f_{n-1}'(y) \right| dy$$

Since the sequence $\left\{ |f_{n}'(y) - f_{n+1}'(y)| \right\}$ is bounded and converges pointwise to zero, it follows from the Lebesgue Dominated Convergence Theorem that $\int_{a}^{b} \left| f_{n}'(x) - f_{n+1}'(x) \right| dx \xrightarrow{n \to \infty} 0.$

$$\left| \int_{a}^{b} M_{n}^{k+1}(x,y) f_{n}'(y) dy - f_{n}'(x) \right|$$

$$= \rho_{n} \int_{a}^{b} M_{n}(x,y) \left[ \rho_{n} \int_{a}^{b} M_{n}(y,z) f_{n}'(z) dz - f_{n-1}'(y) \right] dy$$

$$= \rho_{n} \int_{a}^{b} M_{n}(x,y) \left[ \rho_{n} \int_{a}^{b} M_{n}(y,z) f_{n}'(z) dz - f_{n}'(y) + f_{n}'(y) - f_{n-1}'(y) \right] dy$$

$$\leq \rho_{n} \int_{a}^{b} M_{n}(x,y) \left[ \rho_{n} \int_{a}^{b} M_{n}(y,z) f_{n}'(z) dz - f_{n}'(y) \right] dy$$

$$+ \rho_{n} \int_{a}^{b} M_{n}(x,y) \left( f_{n}'(y) - f_{n-1}'(y) \right) dy$$
\[
\frac{M}{m(b-a)} \left[ \int_a^b \int_a^b \rho_{n,n}^{k,y,z}(y,z) f_n'(z) \, dz - f_n'(y) \, dy \right] \\
+ \frac{M}{m(b+a)} \int_a^b \left| f_n'(y) - f_{n+1}'(y) \right| \, dy
\]

That \[ \int_a^b \int_a^b \rho_{n,n}^{k,y,z}(y,z) f_n'(z) \, dz - f_n'(y) \, dy \xrightarrow{n \to \infty} 0 \] follows from the induction hypothesis and the Lebesgue Dominated Convergence Theorem.

Proof of Theorem 2.1.2: For arbitrary \( n \) and \( k \),

\[
\left| \frac{f_n'(x)}{(\rho_n \lambda_n)^k (\psi_n, f_n')} - \tilde{\phi}_n(x) \right| = \frac{1}{(\psi_n, f_n')} \frac{f_n'(x)}{(\rho_n \lambda_n)^k} - (\psi_n, f_n') \tilde{\phi}_n(x) \\
= \frac{1}{(\psi_n, f_n')} \left[ \frac{f_n'(x)}{(\rho_n \lambda_n)^k} - \frac{M_{n,n} f'(x)}{\lambda_n^k} + \frac{M_{n,n} f'(x)}{\lambda_n^k} - (\psi_n, f_n') \tilde{\phi}_n(x) \right] \\
\leq \frac{1}{(\psi_n, f_n')} \left[ \frac{f_n'(x)}{(\rho_n \lambda_n)^k} - \rho_{n,n}^{k,k} f_n'(x) \right] \\
+ \int_a^b \left| \frac{M_{n,x,y}}{\lambda_n^k} - \tilde{\phi}_n(x) \psi_n(y) f_n'(y) \right| \, dy
\]
Let \( \varepsilon > 0 \) be given.

From Lemma 2.1.3, choose

\[
k_o = \min \left\{ k : C \frac{b-a}{\lambda_n} \left[ 1 - \left( \frac{b}{M} \right)^2 \right]^{k-1} \leq \frac{\varepsilon}{2 \left( \frac{M}{m} \right)^2} \right\}
\]  

Then

\[
\int_a^b \left| \frac{M_n(x,y)}{\lambda_n} - \phi_n(x) \psi_n(y) \right| f_n(y) dy \leq \frac{\varepsilon}{2 \left( \frac{M}{m} \right)^2}
\]

for arbitrary \( n \).

Lemma 2.1.4 insures the existence of \( N(k_o,\varepsilon) \), \( n \geq N \implies \)

\[
\left| f_n(x) - \rho_n M^n f_n(x) \right| \leq \frac{\varepsilon}{k_o+2}
\]

Then for \( n \geq N(k_o,\varepsilon) \) it follows that

\[
\left| \frac{f_n(x)}{\rho_n \lambda_n} - \phi_n(x) \right| \leq \varepsilon
\]  

(2.1.34)

Since \( \int_a^b f_n(x) dx = \int_a^b \phi_n(x) dx = 1 \) for all \( n \), Equation 2.1.30 implies
Then
\[
\left| f_n'(x) - \delta_n(x) \right| = \left| f_n'(x) - \frac{f_n'(x)}{(\rho_n \lambda_n)^{k_0(\psi_n, f_n')}} + \frac{f_n'(x)}{(\rho_n \lambda_n)^{k_0(\psi_n, f_n')}} - \delta_n(x) \right|
\]
\[
\leq \left| 1 - \frac{1}{(\rho_n \lambda_n)^{k_0(\psi_n, f_n')}} \right| f_n'(x) + \frac{f_n'(x)}{(\rho_n \lambda_n)^{k_0(\psi_n, f_n')}} - \delta_n(x)
\]
\[
\leq \varepsilon (b - a) \frac{M}{m(b - a)} + \varepsilon
\]
\[
= \varepsilon \left( 1 + \frac{M}{m} \right)
\]  \hspace{1cm} (2.1.36)

The conclusion
\[
\left| f_n'(x) - f_{n+1}'(x) \right| \xrightarrow{n \to 0} 0 \implies \left| \delta_n(x) - \delta_{n+1}(x) \right| \xrightarrow{n \to 0} 0
\]
can be obtained immediately since
\[
\left| \delta_n(x) - \delta_{n+1}(x) \right| \leq \left| \delta_n(x) - f_n'(x) \right|
\]
\[
+ \left| f_n'(x) - f_{n+1}'(x) \right| + \left| f_{n+1}'(x) - \delta_{n+1}(x) \right|
\]
and the first and third terms tend to zero by Theorem 2.1.2.
It is useful now to summarize Theorems 2.1.1 and 2.1.2 as follows:

$$|\phi_n(x) - \phi_{n+1}(x)| \frac{\to}{n} 0 \iff |f'_n(x) - f'_{n+1}(x)| \frac{\to}{n} 0 \quad (2.1.37)$$

and either condition implies $$|f'_n(x) - \phi_n(x)| \frac{\to}{n} 0 \quad (2.1.38)$$

Then it is clear that

$$\{\phi_n(x)\} \text{ converges iff } \{f'_n(x)\} \text{ converges}$$

and in this case

$$\lim_{n} \phi_n(x) = \lim_{n} f'_n(x) \quad (2.1.39)$$

It is worthwhile to note that the results of this section apply as well to a sequence $$\{f_n\}$$ of p x 1 vectors generated according to

$$f_{n+1} = M_{n+1}f_n \quad n = 0, 1, 2 \ldots$$

where $$f_0$$ is a non-negative p - vector whose inner product with the p - vector $$(1, 1, \ldots, 1)' \equiv 1$$ is positive, and the p x p matrices $$M_n$$ are restricted by

$$0 < \underline{m} \leq m_{ij} \leq \overline{m} \quad (2.1.40)$$

As before, let $$\lambda_n$$ be the dominant characteristic root of $$M_n$$ with associated right and left characteristic vectors...
denoted respectively by \( \hat{\Psi}_n = (\hat{\Psi}_1^n, \hat{\Psi}_2^n, \ldots \hat{\Psi}_p^n) \) and
\( \Psi_n = (\Psi_1^n, \Psi_2^n, \ldots \Psi_p^n) \), where \( \hat{\Psi}_n \) and \( \Psi_n \) are chosen so that
\( (\hat{\Psi}_n, 1) = (\hat{\Psi}_n, \Psi_n) = 1 \). Define
\( f'_n = \frac{f_n}{(f_n, 1)} \).

Then
\[
|\hat{\Psi}_n - \hat{\Psi}_{n+1}| \xrightarrow{n \to 0} 0 \iff |f'_n - f'_{n+1}| \xrightarrow{n \to 0} 0
\]
and either condition implies \( |f'_n - \hat{\Psi}_n| \xrightarrow{n \to 0} 0 \).

Also \( \{\hat{\Psi}_n\} \) converges iff \( \{f'_n\} \) converges and
\[
\lim_{n \to \infty} f'_n = \lim_{n \to \infty} f'_n.
\]

(2.1.41)

The vector convergence indicated above is component-wise convergence. That is, if \( v_n = (v_1^n, v_2^n, \ldots v_p^n) \), then
\( v_n \xrightarrow{n \to 0} 0 \iff v_i^n \xrightarrow{n \to 0} 0, \ i = 1, 2, \ldots p \). The proofs are essentially the same as those given in the continuous case
with Transformation 2.1.9 restated as

\[
P = \frac{1}{\Lambda} \hat{\Psi}^{-1} M \hat{\Psi}
\]

where
\[
\hat{\Psi} = \begin{bmatrix}
\hat{\Psi}_1 & 0 & \cdots & 0 \\
\hat{\Psi}_2 & \hat{\Psi}_1 & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & \hat{\Psi}_p & \hat{\Psi}_p
\end{bmatrix}
\]

In a probabilistic setting the results of this section have two interpretations. Consider first a non-homogeneous
Markov chain with \( p \) states and transition matrices \( M_n \) whose
elements are bounded below by some positive $\Xi$. The $n$th stage probability distribution is given recursively by

$$f_n = M_n f_{n-1} \quad n = 1, 2, \ldots$$

where $f_0$ represents the initial probability vector.

In this case each $M_n$ is stochastic, that is

$$1' M_n = 1', \quad n = 1, 2, \ldots$$

1 is the dominant characteristic root of each $M_n$; $\hat{\psi}_n = 1$, $n = 1, 2, \ldots$, and the associated right characteristic vectors are denoted by $\hat{\phi}_n$.

Result 2.1.41 states that such a chain is strongly ergodic iff

$$\lim_{n \to \infty} f_n = \lim_{n \to \infty} \hat{\phi}_n.$$

Similarly, a non-stationary Markov chain with some interval $[a, b]$ as state space and transition probability kernels $M_n(x, y)$ subject to $0 < \underline{m} \leq M_n(x, y) \leq \overline{M}$, is, by Equation 2.1.39, strongly ergodic iff the sequence $\{\hat{\phi}_n(x)\}$ of right characteristic functions associated with the dominant root 1 converges, and $\lim_{n \to \infty} f_n(x) = \lim_{n \to \infty} \hat{\phi}_n(x)$.

If the sequence $\{\hat{\phi}_n\}$ is such that $|\hat{\phi}_n(x) - \hat{\phi}_{n+1}(x)| \to 0$, the same type of limiting behavior occurs in the sequence $\{f_n\}$ of probability vectors or densities. The converse is also true.
Alternatively the results can be interpreted in situations which differ from the above in that there exist absorbing states. Here attention is directed to the conditional asymptotic distribution, conditional upon non-absorption.

The presence of absorbing states is reflected in the fact that the transition matrices or kernels are no longer stochastic. In this case the conditions

\[
\sum_{i=1}^{p} m_{ij}^n = 1
\]

or

\[
\int_{a}^{b} M_n(x,y)dx = 1
\]

are replaced by

\[
\sum_{i=1}^{p} m_{ij}^n \leq 1
\]

or

\[
\int_{a}^{b} M_n(x,y)dx \leq 1.
\]

Such matrices and kernels are frequently termed substochastic.

If the transition matrices or kernels meet the Conditions 2.1.2 or 2.1.40 respectively, Results 2.1.39 and 2.1.41 can be applied to the sequence \( \{ f_n' \} \) of conditional distributions;
that is,
\[
\{f'_n\} \text{ converges iff } \{\phi_n\} \text{ converges and } \lim f_n = \lim \phi_n.
\]

B. Rates of Convergence and Error Bounds

It has been shown in Section 1 that for a sequence \( \{f'_n\} \) of functions generated iteratively according to
\[
f'_n(x) = \int_a^b M(x,y)f(y)dy
\]
and normed to integrate to one, the convergence to zero of \( \delta_n = \int_a^b \left| \phi_n(x) - \phi_{n+1}(x) \right| dx \) implies the convergence to zero of \( \left| f'_n(x) - \phi_n(x) \right| \). In this section the rate of this convergence is examined, and it is established that exponential convergence of the sequence \( \{\delta_n\} \) implies
\[
\left| f'_n(x) - \phi_n(x) \right| \rightarrow 0 \text{ exponentially.}
\]
Upper bounds on the difference \( \left| f'_n(x) - \phi_n(x) \right| \) are also given.

All constants \( \beta, \rho, \) and \( c_1 \) below may depend on the numbers \( M/M, (b-a), d \) and \( D \), but on nothing beyond these.

Lemma 2.2.1: Let \( \mathcal{M} \) be the set of all measurable kernels \( M(x,y) \) on \([a,b]^2\) with \( 0 < m \leq M(x,y) \leq M \). Let \( f \), non-negative, be such that \( 0 < d \leq f' \leq D \). Let \( M_1, i: 1,2, \ldots, k \) be in \( \mathcal{M} \), and let \( \lambda_i \) and \( \phi_i \) denote the dominant characteristic value and associated right characteristic function respectively of \( M_1 \).
Define
\[
P_i(x,y) = \frac{M_i(x,y) \phi_i(y)}{\lambda_i \phi_i(x)} \quad \text{and} \quad p^k(y) = \sup_x P_{k-1} \ldots P_1(x,y).
\]
Then
\[
B_k = \int_a^b \frac{p^k(y)f(y)dy}{\phi_k(y)} > c_1 > 0.
\]
where \( c_1 = \frac{\overline{m}(h-a)}{\overline{M}} \).

Proof: The proof follows from Equation 2.1.5 and the fact that
\[
\int_a^b p^k(y)dy \geq \int_a^b P_kP_{k-1} \cdots P_1(x,y)dy = 1.
\]

Define \( c = \frac{\overline{M}}{\overline{m}} \), \( \Delta = \left( \frac{\overline{M}}{\overline{m}} \right)^2 \frac{1}{b-a} \), \( \delta = \left( \frac{\overline{m}}{\overline{M}} \right)^2 \frac{1}{b-a} \), and
\[
\delta^*_i = \int_a^b \int_a^b \left| \mathfrak{s}_i(x) - \mathfrak{s}_{i+1}(x) \right| dx dy.
\]

It follows from Equations 2.1.4 and 2.1.5 that
\[
\delta \leq P_1(x,y) \leq \Delta.
\]

Lemma 2.2.2:
\[
\int_a^b \left| \frac{M_k \cdots M_1(x,y)\mathfrak{s}_k(y)}{\prod_{l=1}^k \lambda_l \mathfrak{s}_k(x)} - P_kP_{k-1} \cdots P_1(x,y)dy \right| \leq \frac{\Delta c}{b-a} \beta^k \left( \sup_{1 \leq i \leq k-1} \delta^*_i \right)
\]

Proof: First it will be shown by induction that for \( k \geq 2 \)
\[
\int_a^b \left| \frac{M_k \cdots M_1(x,y)\mathfrak{s}_k(y)}{\prod_{l=1}^k \lambda_l \mathfrak{s}_k(x)} - P_kP_{k-1} \cdots P_1(x,y)dy \right| \leq \frac{\Delta c}{b-a} \beta^k \left( \sup_{1 \leq i \leq k-1} \delta^*_i \right)
\]
\begin{equation}
\lambda \left( \sup_{1 \leq i \leq k-1} \delta_i^* \right) \frac{\Delta_{a}}{b-a} \sum_{j=0}^{k-2} \left( \frac{a}{b-a} \right)^j
\end{equation}

For \(k = 2\),

\[
\int_{a}^{b} \frac{M_{\lambda_{1}}(x,y) \phi_{2}(y)}{\lambda_{2} \lambda_{1} \phi_{2}(x)} - P_{2}P_{1}(x,y) \ dy
\]

\[
= \int_{a}^{b} \int_{a}^{b} \left[ \frac{M_{\lambda_{1}}(x,z) M_{\lambda_{1}}(z,y) \phi_{2}(y)}{\lambda_{2} \lambda_{1} \phi_{2}(x)} - P_{2}(x,z) P_{1}(z,y) \right] \ dz \ dy
\]

\[
= \int_{a}^{b} \int_{a}^{b} P_{2}(x,z) \left[ \frac{M_{\lambda_{1}}(z,y) \phi_{2}(y)}{\lambda_{1} \phi_{2}(z)} - P_{1}(z,y) \right] \ dz \ dy
\]

\[
= \int_{a}^{b} \int_{a}^{b} P_{2}(x,z) \frac{M_{\lambda_{1}}(z,y)}{\lambda_{1}} \left( \frac{\phi_{2}(y)}{\phi_{1}(z)} - \frac{\phi_{1}(y)}{\phi_{1}(z)} \right) \ dz \ dy \leq \frac{\Delta_{a}}{b-a} \delta_{1}^*
\]

Now, assuming Equation 2.2.1 for \(k-1\),

\[
\int_{a}^{b} \frac{M_{\lambda_{k-1}} \ldots M_{\lambda_{1}}(x,y) \phi_{k}(y)}{\prod_{l=1}^{k} \lambda_{l} \phi_{k}(x)} - P_{k}P_{k-1} \ldots P_{1}(x,y) \ dy
\]

\[
= \int_{a}^{b} \int_{a}^{b} P_{k}(x,z) \left[ \frac{M_{\lambda_{k-1}} \ldots M_{\lambda_{1}}(z,y)}{\prod_{l=1}^{k-1} \lambda_{l} \phi_{k}(z)} \left( \frac{\phi_{k}(y)}{\phi_{k}(z)} - \frac{\phi_{k-1}(y)}{\phi_{k-1}(z)} \right) \right] \ dz \ dy
\]
It follows from Equation 2.2.2 that

\[
\frac{M_k M_{k-1} \ldots M_1 (x, y) \phi_k (y)}{P_k P_{k-1} \ldots P_1 (x, y)} \leq \frac{\Delta \left( \frac{a}{b-a} \right)^{k-1} \delta_{k-1}^*}{\Delta \left( \frac{a}{b-a} \right)^{k-2} \sum_{j=0}^{k-3} \left( \frac{a}{b-a} \right)^j} \sup_{1 \leq i \leq k-1} \delta_i^*
\]

\[
\leq \frac{\Delta \left( \frac{a}{b-a} \right)^{k-1} \delta_{k-1}^*}{\Delta \left( \frac{a}{b-a} \right)^{k-2} \sum_{j=0}^{k-3} \left( \frac{a}{b-a} \right)^j} \sup_{1 \leq i \leq k-1} \delta_i^*
\]

\[
\leq \frac{\Delta a}{b-a} \left( \frac{a}{b-a} + 1 \right)^{k-2} \sup_{1 \leq i \leq k-1} \delta_i^*
\]

\[
\leq \frac{\Delta a}{b-a} (e)^k \sup_{1 \leq i \leq k-1} \delta_i^*
\]
where \( \delta = 1 + \frac{2}{b-a} \).

**Theorem 2.2.1:** Define \( \delta_1 = \int_a^b \left| \hat{\phi}_{i+1}(x) - \hat{\phi}_i(x) \right| \, dx \),

\[
f_k(x) = \int_a^b M_k M_{k-1} \cdots M_1(x,y) f(y) \, dy,
\]

and \( \lambda(\varepsilon,k) = (\beta)^k \varepsilon + \rho^k \)

where \( \rho = 1 - (b-a) \delta \). If \( k \) and \( \varepsilon \) are such that \( \sup_{1 \leq i \leq k-1} \delta_1 \leq \varepsilon \), then

\[
\left| f_k'(x) - \hat{\phi}_k(x) \right| \leq c_5 \lambda(\varepsilon,k)
\]

where \( c_5 = \frac{D}{d} \alpha^6 (\alpha^2 + 1) \).

**Proof:** It follows from 2.1.11 that \( \sup_{1 \leq i \leq k-1} \delta_1 \leq \varepsilon \) implies

\[
\sup_{1 \leq i \leq k-1} \delta_1^* \leq c_2 \varepsilon, \quad c_2 = \alpha(1+\alpha^2)(b-a)^2.
\]

Thus from

\[
\text{Lemma 2.2.2, } \int_a^b \left| \frac{M_k M_{k-1} \cdots M_1(x,y) \hat{\phi}_k(y)}{\prod_{l=1}^k \lambda_l \hat{\phi}_k(x)} - p_k p_{k-1} \cdots p_1(x,y) \right| \, dy
\]

\[
\leq c_3 (\beta)^k \varepsilon, \quad c_3 = \alpha^4 (1+\alpha^2)
\]
Define $A_k = \left[ \pi \lambda \int_a^b f(x) \, dx \right]^{-1}$.

Then

$$|A_k \phi_k(x) - B_k \phi_k(x)| = \int_a^b \frac{M_{k-1} \cdots M_1(x,y) f(y) \, dy}{\phi_k(x)^k - \phi_k(y)} - \phi_k(x) \int_a^b \frac{p(y) f(y) \, dy}{\phi_k(y)^k}$$

$$= \int_a^b \frac{M_{k-1} \cdots M_1(x,y) \phi_k(y) - \phi_k(x) p_k(y)}{\phi_k(y)^k} \, f'(y) \, dy$$

$$= \int_a^b \frac{\phi_k(x)}{\phi_k(y)} \left[ \frac{M_{k-1} \cdots M_1(x,y) \phi_k(y)}{\phi_k(y)^k} - p_k(y) \right] \, f'(y) \, dy$$

$$\leq D \alpha \int_a^b \frac{M_{k-1} \cdots M_1(x,y) \phi_k(y)}{\phi_k(y)^k} - p_k(y) \, dy$$

$$\leq D \alpha \int_a^b \frac{M_{k-1} \cdots M_1(x,y) \phi_k(y)}{\phi_k(y)^k} - p_k \cdots p_{k+1} \cdots p_1(x,y) \, dy$$

$$+ \int_a^b \frac{p_k \cdots p_{k-1} \cdots p_1(x,y) - p_k(y)}{\phi_k(y)^k} \, dy$$
where the last inequality follows from Lemmas 2.2.2 and 2.1.1. Thus

\[ |A_k f_k(x) - B_k \delta_k(x)| \leq D\alpha \left( c_3 b^k \varepsilon + \frac{(b-a)(\Delta-\delta)}{1-(b-a)\delta} \rho^k \right) \]

\[ \leq D\alpha \left( a^4 (1 + a^2) b^k \varepsilon + \frac{a^4 - 1}{a^2 (1 - (b-a)\delta)} \rho^k \right) \]

\[ \leq D\alpha \left( a^4 (1 + a^2) b^k \varepsilon + \frac{a^4 - 1}{a^2 (1 - \frac{1}{a^2})} \rho^k \right) \]

\[ \leq D\alpha \left( a^4 (1 + a^2) b^k \varepsilon + (a^2 + 1) \rho^k \right) \]

\[ \leq c_4 \lambda(\varepsilon,k), \]

where

\[ c_4 = D\alpha^5 (1 + a^2) \]
It follows from Lemma 2.2.1 and the above that
\[ \left| \frac{A_k f_k(x)}{B_k} - \phi_k(x) \right| \leq \frac{c_4}{c_1} \lambda(\epsilon, k) \]
and thus that
\[ \left| \frac{A_k}{B_k} \int_a^b f_k(x) - 1 \right| \leq (b-a) \frac{c_4}{c_1} \lambda(\epsilon, k) \]
Then,
\[ \left| f_k'(x) - \phi_k(x) \right| = \left| f_k'(x) \left( 1 - \frac{A_k}{B_k} \int_a^b f_k(x) \right) + \frac{A_k}{B_k} f_k(x) - \phi_k(x) \right| \]
\[ \leq \frac{c}{b-a} \left| 1 - \frac{A_k}{B_k} \int_a^b f_k(x) \right| + \frac{A_k}{B_k} \left| f_k(x) - \phi_k(x) \right| \]
\[ \leq \alpha \frac{c_4}{c_1} \lambda(\epsilon, k) + \frac{c_4}{c_1} \lambda(\epsilon, k) \]
\[ \leq c_5 \lambda(\epsilon, k) \]
where \( c_5 = \frac{c_4}{c_1} (\alpha + 1) \).
Corollary 2.2.1: Consider a sequence \( \{ \tilde{M}_n \} \) and a non-negative function \( f_0 \) such that \( a f_0 > 0 \). Let the sequence \( \{ f_n(x) \} \) be defined by \( f_n(x) = \int_a^b M_n(x,y) f_{n-1}(y) \, dy, \ n = 1,2,\ldots \)

Define
\[
\varepsilon_n = \sup_{j \geq n} \delta_j, \quad \varepsilon_n^* = \varepsilon_n \left[ \frac{n}{2} \right]
\]

\( v(n) = \min \left\{ \left[ \frac{n+1}{2} \right], n^* \right\} \)

where \( n^* = \max \left\{ \left[ \frac{\ln \frac{1}{\varepsilon_n^*}}{(\ln \frac{1}{\rho} + \ln \beta)} \right] - 1, 1 \right\} \)

Then
\[
\left| f'_n(x) - \phi_n(x) \right| \leq c_5 \lambda(\varepsilon_n^*, v(n))
\]

Proof: Identify quantities appearing in the lemmas, theorem and corollary as follows:

- \( k \) : \( v(n) \)
- \( f \) : \( f_n - v(n) \)
- \( \varepsilon \) : \( \varepsilon_n \)
- \( \varepsilon^* \) : \( \varepsilon_n^* \)

Note first that
\[
d = \frac{\bar{m}}{\bar{M}(b-a)} \leq f_n - v(n) \leq \frac{\bar{M}}{\bar{m}(b-a)} = D.
\]
Since
\[ \sup_{1 \leq i \leq v(n) - 1} \delta_{n-v(n)+1} \leq \varepsilon_{n-v(n)+1} \leq \varepsilon_{n} \leq \varepsilon_{n}^{*} = \varepsilon_{n}^{*}, \]
it follows from Theorem 2.2.1 that
\[ \left| f_{n}'(x) - \phi_{n}(x) \right| \leq \psi \lambda (\varepsilon_{n}^{*}, v(n)). \]

Corollary 2.2.2: If \( \delta_{n} \to 0 \), then
\[ \left| f_{n}'(x) - \phi_{n}(x) \right| \to 0. \]

Proof: It follows directly from the definitions of \( \varepsilon_{n} \) and \( \varepsilon_{n}^{*} \) that \( \delta_{n} \to 0 \) implies \( \varepsilon_{n} \to 0 \) which, in turn, implies \( \varepsilon_{n}^{*} \to 0 \).

Thus, for large \( n \),
\[ \lambda (\varepsilon_{n}^{*}, v(n)) = \beta v(n) \varepsilon_{n}^{*} + \rho v(n) \]
\[ \leq \beta \frac{1}{\varepsilon_{n}^{*} v(n)} \left( \ln \frac{1}{\rho} + \ln \beta \right) + \varepsilon_{n}^{*} + \rho v(n). \]
\[
\begin{align*}
& (\ln \beta) \left( \ln \frac{1}{\varepsilon_n} \right) / \left( \ln \frac{1}{\rho} + \ln \beta \right) \\
\leq & \quad e^{*} \left( \ln \frac{1}{\varepsilon_n} \right) / \left( \ln \frac{1}{\rho} + \ln \beta \right) + \rho \nu(n)
\end{align*}
\]

The tendency to zero of \( \varepsilon_n^* \) insures that \( \nu(n) \to \infty \), and

\[\rho = \left[ 1 - (b-a)\delta \right] < 1.\]

Corollary 2.2.3: Suppose there exists a positive \( \gamma < 1 \) such that \( \delta_n < \gamma^n \) for \( n \geq N \). Then, for large \( n \),

\[
\left| f'_n(x) - \phi_n(x) \right| \leq c_6 \max \left\{ \left( \gamma^2 \ln \frac{1}{\rho} \right)^n, \left( \frac{1}{\rho} \right)^n \right\}
\]

where \( c_6 = (1 + \frac{1}{\rho}) \).

Proof: Since \( \varepsilon_n^* \to 0 \), it follows that

\[
\left[ \ln \frac{1}{\varepsilon_n^*} / \left( \ln \frac{1}{\rho} + \ln \beta \right) \right] > 1
\]

for large \( n \). Then
\[ v(n) = \min \left\{ [n+1], \left[ \frac{\ln \frac{1}{\varepsilon_n}}{(\ln \frac{1}{\rho} + \ln \beta)} \right] - 1 \right\} \]

Suppose first that

\[ v(n) = \frac{\ln \frac{1}{\varepsilon_n}}{(\ln \frac{1}{\rho} + \ln \beta)} - 1. \]

Then

\[ \rho v(n) \leq \frac{\ln \frac{1}{\varepsilon_n}}{(\ln \frac{1}{\rho} + \ln \beta)} \leq \frac{1}{\rho} e \]

and

\[ (\beta) v(n) \leq \frac{\ln \frac{1}{\rho}}{(\ln \frac{1}{\rho} + \ln \beta)} \]

so that, in view of the second corollary,

\[ \lambda \left( \varepsilon^*_n, v(n) \right) \leq c_6 \varepsilon^*_n \frac{\ln \frac{1}{\rho}}{(\ln \frac{1}{\rho} + \ln \beta)} \]

\[ \leq c_6 \left( \frac{1}{\gamma^2} \ln \frac{1}{\rho} \right) \left( \ln \frac{1}{\rho} + \ln \beta \right) \]

If, on the other hand, \[ v(n) = \left\lceil \frac{n+1}{2} \right\rceil \]

\[ \lambda \left( \varepsilon^*_n, v(n) \right) = (\beta) v(n) \varepsilon^*_n + \rho v(n) \]
\[\lim_{n \to \infty} \ln \left( \frac{1}{\varepsilon_n} \right) \leq \left( \ln \left( \frac{1}{\rho} \right) + \ln(\beta) \right) \]

\[\leq \ln(\beta) \cdot \frac{1}{\varepsilon_n} \leq \left( \ln \left( \frac{1}{\rho} \right) + \ln(\beta) \right) \]

\[\leq \left( \ln \left( \frac{1}{\rho} \right) + \ln(\beta) \right)^n + \left( \frac{1}{\rho^2} \right)^n\]

\[\leq c_6 \max \left[ \left( \ln \left( \frac{1}{\rho} \right) + \ln(\beta) \right)^n, \left( \frac{1}{\rho^2} \right)^n \right] \]

C. Asymptotic Tendency to Kernels of Rank One

Conditions which insure that a p-state non-stationary Markov chain is weakly ergodic are given by Sarymsakov (1953) and Hajnal (1956). Theorem 2.3.1 verifies that weak ergodicity also obtains when the state space is an interval \([a,b]\), provided that the transition kernels are uniformly positive, and gives a sufficient condition for asymptotic loss of initial information in the substochastic case.

Theorem 2.3.1: Consider a sequence \( \{f_n(x)\} \) of functions defined as in Equation 2.1.1 with kernels \( M_n(x,y) \) restricted by Equation 2.1.2. If \( \int_a^b \left| \psi_n(x) - \psi_{n+1}(x) \right| dx \to 0 \),
then there exists a sequence $q_n(x)$ of functions independent of $f_0(x)$ such that $\left| f_n'(x) - q_n(x) \right| \to 0$

Proof: Define $R_n(x, y) = \frac{M_n(x, y) \psi_n(x)}{\lambda_n \psi_n(y)}$

Then

$$0 \leq R_n(x, y) \leq \frac{F}{F}$$

where $F = \left( \frac{m}{M} \right)^3 \frac{1}{b-a}$ and $F = \left( \frac{M}{m} \right)^3 \frac{1}{b-a}$.

and $\int_a^b R_n(x, y) \, dx = 1$

Let $R_n'(y, x) = R_n(x, y)$. Since $R_n(y, x) \in P$ (see Lemma 2.1.1), there exists, for arbitrary $n$ and $k$, a function $P_n^k(x)$ such

that $\left| R_{n+1}R_{n+2}^{'} \cdots R_{n+k}^{'}(y, x) - P_n^k(x) \right| \leq (F - R)[1 - (b-a)r]^{k-1}$

(2.3.1)

Now $R_{n+1}^{'}R_{n+2}^{'} \cdots R_{n+k}^{'}(y, x) = (R_{n+k}^{'}R_{n+k-1}^{'} \cdots R_{n+1}^{'})(y, x)$,

so Equation 2.3.1 is equivalent to
It can be shown that if \( \int_a^b |\psi_n(x) - \psi_{n+1}(x)| \, dx \xrightarrow{n\to0} 0 \) then for any finite \( k \),

\[
\left| \int_a^b M_{n+k} M_{n+k-1} \cdots M_{n+1}(x,y) \psi_{n+k}(x) \, dy \right| \xrightarrow{n\to0} 0
\]  

The proof is essentially the same as that of Lemma 2.1.2.

Then, for arbitrary \( n \) and \( k \),

\[
\left| \frac{f_{n+k}(x)}{n+k} - \frac{p_n^{k}(x)}{\psi_{n+k}(x)} \psi_{n+k}^{'}(x) \right| \xrightarrow{n\to0} 0
\]
Let $\varepsilon > 0$ be specified. Choose $k_0$ so that

$$(\overline{R} - \overline{r})[1 - (b-a)\overline{r}]^{-k_0-1} < \varepsilon$$

Then for every $n$ it follows from Equation 2.3.2 that

$$\left| R_{n+k} \cdots R_{n+1}(x,y) - P_n^o(x) \right| < \varepsilon,$$

and

$$\frac{f_{n+k_0}(x)}{\psi_{n+k_0}(x)} \sum_{\lambda_i} \int_a^{b} f_n(z)dz - \frac{P_n^o(x)}{\psi_{n+k_0}(x)} \left( \psi_{n+k_0} f_n' \right)$$
In view of Equation 2.3.3 it is possible to choose \( N(k_o, \varepsilon) \) such that \( n > N \) implies

\[
\int_a^b \left( \frac{M_{n+k_o} M_{n+k_o-1} \cdots M_{n+1}(x,y) \psi_{n+k_o}(x)}{n+k_o \prod_{n+1}^{\infty} \lambda_1 \psi_{n+k_o}(y)} \right) dy + (b-a)\varepsilon
\]

Then for \( n \geq N(k_o, \varepsilon) \)

\[
\frac{f_{n+k_o}(x)}{n+k_o \prod_{n+1}^{\infty} \lambda_1 \psi_{n+k_o}(x)} - \frac{P_n(x)}{\psi_{n+k_o}(x)} \left( \psi_{n+k_o}, f_n' \right) < c_1 \varepsilon
\]

where \( c_1 = \frac{M}{\bar{M}} \left( \frac{b-a+1}{b-a} \right) \), and since \( \left( \psi_{n+k_o}, f_n' \right) \geq \frac{M}{\bar{M}} \),

it follows that

\[
\frac{f_{n+k_o}(x)}{n+k_o \prod_{n+1}^{\infty} \lambda_1 \psi_{n+k_o}(x)} - \frac{P_n(x)}{\psi_{n+k_o}(x)} \left( \psi_{n+k_o}, f_n' \right) < \left( \frac{b-a+1}{b-a} \right) \varepsilon
\]

\[ (2.3.4) \]
The Bounds 2.1.6 and the inequality

\[ \bar{r} \leq P_n^o(x) \leq \bar{R} \]

imply

\[
\left( \frac{m}{M} \right)^4 \leq \int_a^b P_n^o(x) \psi^{-1}_{n+k_0}(x)dx \leq \left( \frac{M}{m} \right)^4 \quad (2.3.5)
\]

It follows from Equation 2.3.4 that

\[
\left| \int_a^b f_{n+k_0}(x)dx \right| \left| \int_a^b P_n^o(x) dx \psi^{-1}_{n+k_0}(x)dx \right| < (b-a+1) \varepsilon
\]

and from Equation 2.3.5 that

\[
\left| \int_a^b f_{n+k_0}(x)dx \right| \left| \int_a^b P_n^o(x) dx \psi^{-1}_{n+k_0}(x)dx \right| \left| \int_a^b f_n'(x)dx \right| < (b-a+1) \left( \frac{M}{m} \right)^4 \varepsilon \quad (2.3.6)
\]

Then for \( n \geq N(k_0, \varepsilon) \),

\[
\left| f'_{n+k_0}(x) \right| \leq \frac{P_n^o(x) \psi^{-1}_{n+k_0}(x)}{P_n^o, \psi^{-1}_{n+k_0}}
\]
\[
I_\omega f_n + k_0(x) - \int_a^b f_{n+k_0}(x) \, dx \\
\left[ \left( \int_a^b f_{n+k_0}(x) \, dx \right) - \frac{\int_a^b f_{n+k_0}(x) \, dx}{\lambda_n \psi_{n+k_0}} \right] - \frac{f_{n+k_0}(x)}{\lambda_n \left( \psi_{n+k_0}, f_n \right) - \frac{\int_a^b f_{n+k_0}(x) \, dx}{\lambda_n \left( \psi_{n+k_0}, f_n \right)}}
\]

\[
\lambda \left( \left( \int_a^b f_{n+k_0}(x) \, dx \right) - \frac{\int_a^b f_{n+k_0}(x) \, dx}{\lambda_n \left( \psi_{n+k_0}, f_n \right) - \frac{f_{n+k_0}(x)}{\lambda_n \left( \psi_{n+k_0}, f_n \right)}} \right)
\]

\[
\leq \frac{\left( \frac{b-a+1}{b-a} \right)}{\frac{1}{\frac{M}{m}}} + \frac{1}{\left( \frac{\psi_{n+k_0}}{\lambda_n \left( \psi_{n+k_0}, f_n \right)} \right)}
\]

\[
\leq c_2 \varepsilon, \text{ where } c_2 = \left( \frac{\frac{M}{m}}{b-a} \right)^4 \frac{b-a+1}{b-a} \left( \frac{\frac{M}{m}}{b-a} + 1 \right)
\]
Since \( c \) is arbitrary the desired result is obtained

\[
q_{n+k_0}(x) = \frac{k_0(x) \psi_{n+k_0}(x)}{p_n(x) \psi_{n+k_0}(x)}
\]

That stochastic kernels \( M_n(x,y) \) satisfy the hypothesis of Theorem 2.3.1 is clear since in this case \( \psi_n(x) = 1 \) for all \( n \). In the substochastic situation the conditional \( n \)th stage probability distribution \( f_n'(x) \) will tend toward independence of \( f_0'(x) \), the initial distribution, even if \( \left\{ f_n'(x) \right\} \) does not converge, as long as the sequence \( \left\{ \psi_n(x) \right\} \) of left characteristic functions fulfills the hypothesis of Theorem 2.3.1.

D. Convergence Theorems for Characteristic Roots and Functions

Let \( \left\{ M_n(x,y) \right\} \) be a sequence of bounded, uniformly positive, measurable kernels on \([a, b]^2\), with dominant characteristic roots \( \left\{ \lambda_n \right\} \) and associated right and left characteristic functions \( \left\{ \hat{\gamma}_n \right\} \) and \( \left\{ \gamma_n \right\} \), respectively. It is the purpose of this section to establish that if successive differences between the kernels tend in an appropriate manner toward zero, the same behavior is exhibited by their dominant roots and associated characteristic functions.

Theorem 2.4.1: Let \( \left\{ M_n(x,y) \right\} \), \( 0 < \underline{M} \leq M_n(x,y) \leq \overline{M} \), be as described above with characteristic functions associated with
If \( \int_a^b \phi_n(x) \psi_n(x) \, dx = 1 \). If

\[
\int_a^b |M_n(x,y) - M_{n+1}(x,y)| \, dy \xrightarrow{n \to \infty} 0 \text{ uniformly in } x,
\]

then

\[
|\lambda_n - \lambda_{n+1}| \xrightarrow{n \to \infty} 0.
\]

Proof: The following characterization of the dominant root of \( M_n(x,y) \) is given by Harris (1964).

Let \( S_n = \{ \lambda_n > 0 : \text{ a bounded, non-negative function } f \}
\)

such that \( \int_a^b M_n(x,y) f(y) \, dy \geq \lambda_n f(x) \) but \( \int_a^b M_n(x,y) f(y) \neq \lambda_n f(x) \). Then \( \lambda_n = \text{l.u.b. } \lambda_n \).

Let \( \varepsilon > 0 \). Choose \( \eta < \varepsilon \left( \frac{\bar{m}}{\bar{M}(b-a)} \right) \leq \varepsilon \delta_n(y), n = 1,2, \ldots \)

Since \( \int_a^b |M_n(x,y) - M_{n+1}(x,y)| \, dy \xrightarrow{n \to \infty} 0 \) uniformly in \( x \), it is possible to choose \( N(\eta) \) such that \( n \geq N(\eta) \) implies

\[
\int_a^b |M_n(x,y) - M_{n+1}(x,y)| \, dy < \frac{\bar{m}(b-a)}{\bar{M}} \eta.
\]

Then

\[
\int_a^b M_n(x,y) \psi_{n+1}(y) \, dy - \int_a^b M_{n+1}(x,y) \psi_{n+1}(y) \, dy
\]
\[ \lambda \leq \int_a^b \left| M_n(x,y) - M_{n+1}(x,y) \right| \phi_{n+1}(y) dy \]

\[ \leq \frac{M}{M(b-a)} \int_a^b \left| M_n(x,y) - M_{n+1}(x,y) \right| dy < \eta \]

It follows that

\[ \int_a^b M_n(x,y) \phi_{n+1}(y) \geq \int_a^b M_{n+1}(x,y) \phi_{n+1}(y) dy - \eta \]

\[ \geq \lambda_{n+1} \phi_{n+1}(x) - \epsilon \phi_{n+1}(x) \]

\[ \geq (\lambda_{n+1} - \epsilon) \phi_{n+1}(x) \]

Therefore

\[ (\lambda_{n+1} - \epsilon) \epsilon \leq S_n \]

or equivalently

\[ \lambda_n \geq \lambda_{n+1} - \epsilon \quad \text{for } n \geq N. \quad (2.4.1) \]

Also

\[ \left| \int_a^b M_n(x,y) \phi_n(y) dy - \int_a^b M_{n+1}(x,y) \phi_n(y) dy \right| < \eta \text{ if } n \geq N, \text{ from which it follows that} \]
Thus \((\lambda_n - \epsilon) \in S_{n+1}\), so that

\[
\lambda_n - \epsilon \leq \lambda_{n+1} \quad \text{for} \quad n \geq N \tag{2.4.2}
\]

It follows from Equations 2.4.1 and 2.4.2 that \(n \geq N\) implies

\[
|\lambda_n - \lambda_{n+1}| < \epsilon.
\]

**Theorem 2.4.2:** If \(\int_a^b \left| M_n(x,y) - M_{n+1}(x,y) \right| \, dy \xrightarrow{n \to \infty} 0\) uniformly in \(x\), then

\[
\left| \phi_n(x) - \phi_{n+1}(x) \right| \xrightarrow{n \to \infty} 0.
\]

It is useful at this point to define \(K_n(x,y) = \frac{M_n(x,y)}{\lambda_n}\).

Note that the hypothesis of the theorem implies, in view of Theorem 2.4.1, that

\[
\int_a^b \left| K_n(x,y) - K_{n+1}(x,y) \right| \, dy \xrightarrow{n \to \infty} 0\] uniformly in \(x\).

A lemma will precede the proof of the theorem.

**Lemma:** Under the hypothesis of Theorem 2.4.2

\[
\left| \phi_n(x) - K_n^{(k)} \phi_n(x) \right| \xrightarrow{n \to \infty} 0 \quad \text{for every finite} \quad k.
\]
Proof of lemma: By induction.

For \( k = 1 \),

\[
\phi_n(x) - K_{n+1}(x, y) = \int_a^b (K_n(x, y) - K_{n+1}(x, y)) \phi_n(y) dy
\]

\[
\leq \frac{M}{m(b-a)} \int_a^b |K_n(x, y) - K_{n+1}(x, y)| dy \to 0.
\]

Then

\[
\phi_n(x) - K^{(k+1)}_{n+1}\phi_n(x) \leq \frac{M}{m(b-a)} \int_a^b |K^{(k+1)}_n(x, y) - K^{(k+1)}_{n+1}(x, y)| dy
\]

\[
\leq \frac{M}{m(b-a)} \int_a^b \int_a^b |K^{(k)}_n(x, z) K_n(z, y) dz - \int_a^b K^{(k)}_n(x, z) K_{n+1}(z, y) dz| dy
\]

\[
+ \int_a^b |K^{(k)}_n(x, z) K_{n+1}(z, y) dz - \int_a^b K^{(k)}_{n+1}(x, z) K_{n+1}(z, y) dz| dy
\]

\[
\leq \frac{M}{m(b-a)} \left\{ \int_a^b \int_a^b |K^{(k)}_n(x, z)| |K_n(z, y) - K_{n+1}(z, y)| dz dy
\]

\[
+ \int_a^b \int_a^b |K^{(k)}_n(x, z) - K^{(k)}_{n+1}(x, z)| K_{n+1}(z, y) dz dy \right\}
That the first and second terms in Equation 2.4.3 tend with \( n \) to zero follows from the hypothesis of the lemma and the induction hypothesis respectively.

Proof of Theorem 2.4.2: Let \( \epsilon > 0 \).

\[
|\hat{\phi}_n(x) - (\hat{\psi}_{n+1}, \hat{\psi}_n) \hat{\phi}_{n+1}(x)| 
\leq |\hat{\phi}_n(x) - K_{n+1}^k \hat{\phi}_n(x)| 
+ |K_{n+1}^k \hat{\phi}_n(x) - (\hat{\psi}_{n+1}, \hat{\psi}_n) \hat{\phi}_{n+1}(x)|
\]

It follows from Lemma 2.1.3 that

\[
|K_{n+1}^k \hat{\phi}_n(x) - (\hat{\psi}_{n+1}, \hat{\psi}_n) \hat{\phi}_{n+1}(x)| = 
\left| \int_a^b K_{n+1}^k(x, z) \hat{\phi}_n(z) \, dz - \int_a^b \hat{\phi}_n(x) \psi_{n+1}(z) \hat{\psi}_n(z) \, dz \right|
\]

\[
\leq \frac{\bar{M}}{\bar{m}(b-a)} \int_a^b \left| K_{n+1}^k(x, z) - \hat{\phi}_{n+1}(x) \psi_{n+1}(z) \right| \, dz
\]
\[ \leq C_2 \left( \frac{M}{\bar{M}}, b-a \right) \left( 1 - \frac{m}{\bar{M}} \right)^2 \left( \frac{M}{M} \right)^{k-1} \]

where
\[
C_2 \left( \frac{M}{\bar{M}}, b-a \right) = \left( \frac{1}{b-a} \right) \left( \frac{M}{\bar{M}} \right)^3 \left[ \left( \frac{M}{\bar{M}} \right)^2 - \left( \frac{M}{\bar{M}} \right)^3 \right] \]

If \( k_0 \) is chosen so that \( C_2 \left( \frac{M}{\bar{M}}, b-a \right) \left[ 1 - \left( \frac{M}{\bar{M}} \right)^2 \right]^{k_0-1} < \varepsilon/2 \)

and \( N(k_0, \varepsilon) \) so that \( n \geq N \) implies
\[
|\phi_n(x) - K^{(k_0)}_{n+1} \phi_n(x)| < \varepsilon/2, \]

it follows from Equation 2.4.2 that for \( n \geq N(k_0, \varepsilon) \)
\[
|\phi_n(x) - \langle \psi_{n+1}, \phi_n \rangle \phi_n(x)| < \varepsilon \]

Since
\[
|1 - \langle \psi_{n+1}, \phi_n \rangle| = \left| \int_a^b \left( \phi_n(x) - \langle \psi_{n+1}, \phi_n \rangle \phi_n(x) \right) dx \right| \leq \varepsilon(b-a),
\]

\[
|\phi_n(x) - \phi_{n+1}(x)| \leq \varepsilon \left( 1 + \frac{M}{\bar{M}} \right), \text{ where } \varepsilon > 0 \text{ was arbitrary.}
\]

Since the functions \( \{ \psi_n \} \) can be thought of as right characteristic functions for the sequence \( M_n(x,y) \) where \( M'_n(x,y) = M_n(y,x) \), it follows immediately that
Calculations essentially the same as those in the proofs of the two theorems of this section can be used to establish that if there exists a kernel \( M_o(x,y) \), satisfying the same conditions as those placed on the kernels \( M_n(x,y) \), such that

\[
\int_a^b |M_n(x,y) - M_o(x,y)| \, dy \xrightarrow{n \to 0} 0 \text{ uniformly in } x,
\]

then \( \lambda_n \xrightarrow{n \to 0} \lambda_0 \) and \( \psi_n \xrightarrow{n \to 0} \psi_0 \).

It should be noted that the results of Sections 1, 2 and 3 apply, by means of Section 4, to a sequence of functions generated iteratively according to

\[
f_{n+1}(x) = \int_a^b M_{n+1}(x,y) f_n(y) \, dy \quad n = 0, 1, 2, \ldots
\]

under the usual restrictions on the kernels in case

\[
\int_a^b |M_n(x,y) - M_{n+1}(x,y)| \xrightarrow{n \to 0} 0, \text{ uniformly in } x.
\]
III. APPLICATIONS TO STATISTICS

A. Sequential Probability Ratio Tests

Mengido (1963) has suggested approaching the analysis of truncated SPRT's by comparing a CAUOC (conditional asymptotic untruncated operating characteristic) function and a CATOC (conditional asymptotic truncated operating characteristic) function. The CAUOC function gives, as a function of the unknown parameter, the conditional probability of eventual acceptance, conditional on the event $E_n$: no decision yet at a large sample stage $n$. The CATOC function is the integral from the lower boundary to a truncation point $T$ of the asymptotic distribution of $n$-stage sample path end points, again conditional on $E_n$. These two functions make possible the approximate (asymptotically exact) matching, for one value of the unknown parameter, of the OC function of the untruncated SPRT with the OC of a suitably truncated procedure.

Mengido exploited this idea in the discrete case, specifically binomial sequential analysis. An important feature of the conditional distributional problems involved is that the conditional distributions underlying the CATOC functions are not degenerate, so that choosing a truncation point matching the two OC's at a suitable parametric value is feasible. An entirely analogous situation pertains in the continuous case.
Consider any continuous SPRT whose execution amounts to the construction and observation of a homogeneous continuous random walk between two absorbing barriers. If the density function of a step is given by $M_0(t)$, then it is clear that the conditional distribution of path end points for paths not absorbed at $a$ or $b$ by stage $n+1$ is given, for $a \leq x \leq b$, by

\[
\begin{align*}
\int_a^b \cdots \int_a^b M_0(x-z_n) \cdots M_0(z_{2-z_1})M_0(z_1)dz_1dz_2\cdots dz_n \\
\int_a^b \cdots \int_a^b M_0(x-z_n) \cdots M_0(z_{2-z_1})M_0(z_1)dz_1dz_2\cdots dz_n dx
\end{align*}
\]

\[
\int_a^b \int_a^b \frac{M_0^{(n)}(x,z_1)M_0(z_1)dz_1}{M_0^{(n)}(x,z_1)M_0(z_1)dz_1dx} = \frac{f_{\theta,n+1}(x)}{\int_a^b f_{\theta,n+1}(x)dx}
\]  

Hence, for truncation of the corresponding SPRT at some point $T$, $a < T < b$, the probability of acceptance, given evolution of the test to stage $n$, is
\[
\int_a^T f'_{\theta,n}(x)dx. \quad (3.1.2)
\]

If \( M_\theta(t) \) is positive for \( a-b < t < b-a \) then it is a corollary of Theorem 2.1.1 and indeed follows from a theorem of Doob (1953) that

\[
f'_{\theta,n}(x) \xrightarrow{n} \phi_\theta(x),
\]

where \( \phi_\theta(x) \) is the right characteristic function, taken to integrate to one, associated with the dominant characteristic value of the kernel \( M_\theta(x-y), a \leq x, y \leq b \). It then follows that the CATOC function in the sense of Mengido is given by

\[
\text{CATOC} (\theta) = \int_a^T \phi_\theta(x)dx \quad (3.1.3)
\]

The continuous version of the CAUOC is obtained as follows through an argument entirely analogous to that used in the discrete case.

The probability \( \pi_\theta(n) \) of eventual acceptance under the untruncated SPRT, given no decision by sampling stage \( n \), is

\[
\pi_\theta(n) = \int_a^b \int_a^b M_\theta(x-y) f'_{\theta,n}(y)dydx
\]

\[
+ \sum_{i=1}^\infty \int_a^b \int_a^b \int_a^b M_\theta(x-z)M_\theta^{(i)}(z,y)f'_{\theta,n}(y)dydzdx
\]
Hence, if \( \lambda_a \) is the dominant characteristic root of

\[
M_\theta(x-y), \ a \leq x, y \leq b,
\]

\[
\pi_\theta(n) = \frac{1}{1-\lambda_\theta} \int_a^b \int_a^b M_\theta(x-y) \phi_\theta(y)dydx
\]

\[
\leq \int_a^b \int_a^b M_\theta(x-y) f'_\theta,n(y)dydx - \int_a^b \int_a^b M_\theta(x-y) \phi_\theta(y)dydx
\]

\[
+ \sum_{i=1}^{X} \int_a^b \int_a^b \int_a^b M_\theta(x-z) M_\theta^{(i)}(z,y) f'_\theta,n(y)dydzdx
\]

\[
- \lambda_i \int_a^b \int_a^b M_\theta(x-y) \phi_\theta(y)dydx
\]

\[
+ \sum_{i=X+1}^{\infty} \int_a^b \int_a^b \int_a^b M_\theta(x-z) M_\theta^{(i)}(z,y) f'_\theta,n(y)dydzdx
\]

\[
+ \frac{\lambda_{X+1}}{1-\lambda_\theta} \int_a^b \int_a^b M_\theta(x-y) \phi_\theta(y)dydx \quad (3.1.4)
\]

It will be assumed, in accordance with Chapter II, that

\( M_\theta(t) \) is positive for \( a-b < t < b-a \) and that there exists an

\( \eta > 0 \) such that
The third term in (3.1.4) is, in effect, the probability that with a starting point distribution \( f_{n}(y) \) a random walk with step distribution \( M_{\theta}(t) \) violates the bounds no earlier than at step \( K+2 \) and then at \( a \). This probability is, in turn, no greater than the probability that all of the first \( K+1 \) steps lie between \( a-b \) and \( b-a \), i.e. no greater than

\[
\sum_{i=K+1}^{\infty} \int_{a}^{b} \int_{a}^{b} M^{(i)}(x,y) f_{n}(y)dydx \leq \sum_{i=K+1}^{\infty} (1-\eta)^{i} = \frac{(1-\eta)^{K+1}}{\eta}
\]

There is, then, a \( K \) independent of \( n \) which makes the third term arbitrarily small. The same assumption (3.1.5) insures that \( \lambda_{\theta} < 1 \), and thus the fourth term can be made arbitrarily small, independently of \( n \), for large \( K \).

Finally, the convergence of \( f_{\theta,n} \) to \( \phi_{\theta} \) with \( n \) guarantees, through the dominated convergence theorem, for fixed \( K \) that the first and second terms can be made arbitrarily small for large enough \( n \).

Thus it follows that the CAUOC, in the sense of Mengido (1963), is given by

\[
\text{CAUOC} (\theta) = \frac{1}{1-\lambda_{\theta}} \int_{a}^{b} \int_{a}^{b} M_{\theta}(x-y) \phi_{\theta}(y)dydx
\]
Consider, for example, the SPRT for testing

\[ H_0 : f(x) = \frac{1}{2} e^{-|x-\mu_0|}, \quad -\infty < x < \infty \]

vs

\[ H_1 : f(x) = \frac{1}{2} e^{-|x-\mu_1|}, \quad -\infty < x < \infty \]

This SPRT will be based on the successive inequalities

\[ A < \sum_{1}^{n} |x_i - \mu_0| - \sum_{1}^{n} |x_i - \mu_1| < B \]

where A and B are determined by the desired error rates of the first and second kind.

It is easily verified that if \( \mu_1 - \mu_0 \geq B - A \), then this SPRT is in fact based on the successive inequalities

\[ a = \frac{A}{2} < \sum_{1}^{n} (x_i - \bar{\mu}) < \frac{B}{2} = b \]

where \( \bar{\mu} = \frac{\mu_1 + \mu_2}{2} \),

in other words on the random walk between absorbing barriers a and b with step size distribution

\[ K_\mu (t) = \frac{1}{2} e^{-|t - \mu + \bar{\mu}|} \]

For the kernel \( K_{\mu} (y-x) = \frac{1}{2} e^{-|y-x|} \), \( a \leq x, y \leq b \),

Kac (1945) gives the dominant characteristic root
\[
\frac{1}{\mu} = \frac{1}{\beta^2 + 1} \tag{3.1.6}
\]

with associated characteristic function

\[
\Phi_\mu(x) = C_1 \cos \beta x + C_2 \sin \beta x \tag{3.1.7}
\]

where

\[
C_1 = \frac{3(\cos \beta b + \sin \beta b)}{1 + \beta \sin (b-a)\beta - \cos (b-a)\beta}
\]

\[
C_2 = \frac{3(\sin \beta b - \cos \beta b)}{1 + \beta \sin (b-a)\beta - \cos (b-a)\beta}
\]

and \(\beta\) is the smallest positive root of the transcendental equation

\[
\tan (b-a)\beta = \frac{2\beta}{1 + \beta^2}
\]

It follows that for the above SPRT, satisfying the condition \(\mu_1 - \mu_0 \geq B - A\), truncating at a point \(T\) which satisfies

\[
\int_a^T \Phi_\mu(x)dx = \frac{1}{1-\lambda_\mu} \int_a^b K_{\mu}(x-y) \Phi_\mu(y)dydx \tag{3.1.8}
\]

approximately equates the truncated and untruncated OC functions for \(\mu = \bar{\mu}\).
As a second example, the SPRT for testing

\[ H_0: f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu_0)^2} \quad -\infty < x < \infty \]

vs

\[ H_1: f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu_1)^2} \quad -\infty < x < \infty \]

is based on the successive inequalities

\[ A < \sum_{i=1}^{n} (x_i - \mu_0)^2 - \sum_{i=1}^{n} (x_i - \mu_1)^2 < B \]

where, again, A and B are determined by the desired error rates. It is easily seen that this SPRT is equivalent, in case \( \mu_1 > \mu_0 \), to one based on the inequalities

\[ a < \sum_{i=1}^{n} (x_i - \bar{\mu}) < b \]

where \( a = \frac{A}{2(\mu_1-\mu_0)} \) and \( b = \frac{B}{2(\mu_1-\mu_0)} \)

and is thus equivalent to the random walk between absorbing barriers a and b with step size distribution
\[ M_\mu(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t - \mu + \overline{\mu})^2} \]

In this case, approximations for \( \lambda_\mu \) and \( \phi_\mu \) must be obtained before truncation rule (3.1.8) is applied.

The selection of \( \overline{\mu} \) as the value of \( \mu \) at which to equate the truncated and untruncated OC's is suggested by the fact that \( \lambda_\mu \) and \( \phi_\mu \) are then the dominant characteristic root and associated characteristic function of a symmetric kernel.

Since characteristic roots and functions of non-symmetric kernels are generally more difficult to obtain than those of symmetric kernels, it is worthwhile to note that for the non-symmetric kernels generated by normal densities, the problem reduces quite simply to the symmetric case.

Suppose, for example, that it is desired to equate the truncated and untruncated OC's for the normal SPRT considered above at a parameter value \( \mu \neq \overline{\mu} \). It is required then to obtain the dominant root \( \lambda_\mu \) and associated right characteristic function \( \phi_\mu \) of the kernel

\[ M_\mu(x,y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-y-\mu+\overline{\mu})^2} \quad a \leq x, y \leq b. \]

It is clear that \( \lambda_\mu \) and \( \phi_\mu \) solve the equation
\[ \lambda_{\mu} \psi_{\mu}(x) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-y-\mu-\bar{\mu})^2} \psi(y) dy \]

if and only if

\[ \lambda_{\mu} e^{-\frac{1}{2}(\mu-\bar{\mu})^2} \phi_{\mu}(x) e^{-(\mu-\bar{\mu})x} = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-y)^2} e^{-(\mu-\bar{\mu})y} \phi(y) dy \]

which implies the relations

\[ \lambda_{\mu} = \lambda_{\bar{\mu}} \quad -(\mu-\bar{\mu}) = \frac{1}{2}(\mu-\bar{\mu})^2 \]

(3.1.9)

and

\[ \phi_{\bar{\mu}}(x) = \phi_{\mu}(x) e^{-\mu-\bar{\mu})x} \]

(3.1.10)

Equations 3.1.9 and 3.1.10 are easily solved for \( \lambda_{\mu} \) and \( \phi_{\mu} \) in terms of \( \lambda_{\bar{\mu}} \) and \( \phi_{\bar{\mu}} \) so that the non-symmetric problem does, in fact, reduce to a symmetric one.

As a third example, consider the family of densities given by

\[ f_\theta(x) = e^{a(\theta) + b(\theta)x + h(x)} \]

(3.1.11)
where, for some $\theta_0$, $\theta_1$, $\theta_2$,

$$h(x) = -b(\theta_0)x - |x + \delta(\theta_1,\theta_2)|$$

The SPRT of $\theta_1$ vs. $\theta_2$ is based on successive inequalities

$$A < \frac{[b(\theta_2) - b(\theta_1)]}{b(\theta_2)} \sum_{i=1}^{n} (x_i + \delta) < B$$

where $A$ and $B$ depend on error rates and

$$\delta = \frac{a(\theta_2) - a(\theta_1)}{b(\theta_2) - b(\theta_1)}$$

The SPRT is equivalent, in case $b(\theta_2) > b(\theta_1)$, to a random walk with absorbing barriers at $a = \frac{A}{b(\theta_2) - b(\theta_1)}$ and

$$b = \frac{B}{b(\theta_2) - b(\theta_1)}$$

whose step size density is

$$M_{\theta}(t) = e^{a(\theta) + [b(\theta) - b(\theta_0)](t - \delta) - |t|}$$

Then a truncation point $T$ such that

$$\int_{a}^{T} \psi_{\theta}(x)dx = \frac{1}{1-\lambda_{\theta}} \int_{a}^{b} \int_{a}^{b} M_{\theta}(x-y) \psi_{\theta}(y) dy dx$$
where $\lambda_\theta$ and $\phi_\theta$ are the dominant root and associated right characteristic function of

$$M_\theta(x-y) = e^{a(\theta) + [b(\theta)-b(\theta_0)][x-y-\delta] - |x-y|}$$

approximately equates CAUOC ($\theta$) and CATOC ($\theta$).

To obtain $\lambda_\theta$ and $\phi_\theta$, note that $M_{\theta_0}(t) = e^{a(\theta_0)-|x-y|}$ is the symmetric kernel studied by Kac (1945) whose dominant root $\lambda_{\theta_0}$ and associated characteristic function $\phi_{\theta_0}$ are given by Equations 3.1.6 and 3.1.7 respectively. The relations between $\lambda_\theta$, $\phi_\theta$ and $\lambda_{\theta_0}$, $\phi_{\theta_0}$ may be established as follows.

If $\lambda_\theta$ and $\phi_\theta$ solve the integral equation

$$\phi_\theta(x) = \lambda_\theta \int_a^b M_\theta(x-y) \phi_\theta(y)dy$$

then

$$\phi_\theta(x) = \lambda_\theta \int_a^b e^{a(\theta) + [b(\theta)-b(\theta_0)][x-y-\delta] - |x-y|} \phi_\theta(y)dy$$

or

$$\frac{\phi_\theta(x)}{e^{b(\theta)-b(\theta_0)|x}}$$

$$= \lambda_\theta \int_a^b e^{a(\theta)-a(\theta_0)} \int_a^b e^{a(\theta_0)-|x-y|} \frac{\phi_\theta(y)}{e^{b(\theta)-b(\theta_0)|y}} dy$$
Thus \( \lambda_\theta = \frac{e^{[b(\theta) - b(\theta_0)]}}{e^{a(\theta) - a(\theta_0)}} \lambda_{\theta_0} \) \hspace{1cm} (3.1.13)

and

\[ \phi_\theta(x) = e^{[b(\theta) - b(\theta_0)]x} \phi_{\theta_0}(x) \] \hspace{1cm} (3.1.14)

**B. Generalized Sequential Probability Ratio Tests**

Weiss (1953) gives the following definition for a "generalized sequential probability ratio test," abbreviated GSPRT, for testing one simple hypothesis \( H_0 \) against another simple hypothesis \( H_1 \). There exist two sequences \( \{A_n\} \) and \( \{B_n\} \) of predetermined non-negative constants such that \( A_n \leq B_n \) for all \( n \). Sampling is continued as long as

\[ A_n < \frac{f_1(x)}{f_0(x)} < B_n \] \hspace{1cm} (3.2.1)

\( H_0 \) is accepted if the first violation occurs at the lower bound, \( H_1 \) if it occurs at the upper bound. The GSPRT differs, according to Weiss, from the familiar Wald (1947) SPRT only in that the bounds \( A_n \) and \( B_n \) are not necessarily the same at each stage of the sampling.
Properties of GSPRT's are discussed by Weiss (1953) and by Kiefer and Weiss (1957). This section is concerned with demonstrating a relationship between certain GSPRT's and non-stationary random walks which makes possible the application of the methods of the previous section to the truncation of these GSPRT's.

Consider, for example, a GSPRT for testing

\[ H_0: f_0(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu_0)^2} \quad -\infty < x < \infty \]

vs.

\[ H_1: f_1(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu_1)^2} \quad -\infty < x < \infty \]

This GSPRT is based, in case \( \mu_1 > \mu_0 \), on the successive inequalities

\[ a_n < \sum_{l=1}^{n} (x_i - \bar{\mu}) < b_n \quad (3.2.2) \]

where \( a_n = \frac{\ln A_n}{\mu_1 - \mu_0} \), \( b_n = \frac{\ln B_n}{\mu_1 - \mu_0} \) and \( \bar{\mu} = \frac{\mu_1 + \mu_2}{2} \),

and is, thus, equivalent to a random walk with nth stage absorbing barriers at \( a_n \) and \( b_n \) and step size density function
If the sequences \( \{a_n\} \) and \( \{b_n\} \) converge to \( a \) and \( b \) respectively, and if \( b_n - a_n = c \) for all \( n \), the random walk with \( n \)th stage barriers at \( a_n \) and \( b_n \) and step size density given by Equation 3.2.3 is equivalent to one with absorbing barriers at \( a \) and \( b \) whose \( n \)th stage step density \( M_n(t) \) can be determined as follows.

Let \( Y_n = \sum_{i=1}^{n} (x_i - \mu) \), and let \( c_n = b_n - b = a_n - a \).

Then \( a_n < Y_n < b_n \) if and only if \( a = a_n - c_n < Y_n - c_n < b_n - c_n = b \). The \( n \)th stage step size is given by

\[
T = (Y_n - c_n) - (Y_{n-1} - c_{n-1}) = X_n - \mu - (c_n - c_{n-1})
\]

from which it follows that

\[
M_{\mu,n}(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t + \bar{\mu} - [\mu - (c_n - c_{n-1})])^2} \quad (3.2.4)
\]

The derivation of CATOC and CAUOC functions analogous to those of the previous section will proceed in a more general context.

Consider a continuous random walk between absorbing barriers \( a \) and \( b \) with density function for the \( n \)th stage step given by \( M_n(t) \). The conditional density of path end points
for paths not absorbed by stage $n$ is given, for $a < x < b$, by

$$f_n'(x) = \frac{1}{b} \int_a^b \prod_{i=1}^{n-1} m_i(x, z_{i-1}) m_{n-1}(z_{n-2}) \cdots m_1(z_1) dz_1 dz_2 \cdots dz_{n-1} \int_a^b f_n(x) dx$$

If the sequence $\{M_n(t)\}$ converges to a density $M_0(t)$ and if there exist positive numbers $\underline{m}$ and $\bar{M}$ such that $\underline{m} \leq M_n(x-y) \leq \bar{M}$, $a \leq x, y \leq b$, then it follows from Theorem 2.4.1 and Theorem 2.1.1 that

$$\left| f_n'(x) - \hat{\phi}_0(x) \right| \rightarrow 0 \text{ uniformly on } [a, b] \quad (3.2.5)$$

where $\hat{\phi}_0(x)$ is the right characteristic function associated with the kernel $M_0(x-y)$. Thus, for a truncation point $T$, $a < T < b$, the analogue of Equation 3.1.2 is given by

$$\text{CATOC } (M_0) = \int_a^T \hat{\phi}_0(x) dx \quad (3.2.6)$$

To derive an analogue for the CAUOC function, it will be assumed, as in the stationary case, that there exists an $\eta > 0$ such that
The probability $\pi(n)$ of eventual absorption at the barrier $a$, given evolution of the walk to stage $n$, is

$$\pi(n) = \int_a^b \int_a^b M_{n+1}(x-y)f_n'(y)dydx$$

$$+ \sum_{i=2}^{\infty} \int_a^b \int_a^b \int_a^b M_{n+i}(x-y)M_{n+i-1}(y,z)f_n'(z)dzdydx.$$
\[
\begin{align*}
&\leq \left| \int_{-\infty}^{a} \int_{a}^{b} M_{n+1}(x-y)f_{n}'(y)dydx - \int_{-\infty}^{a} \int_{a}^{b} M_0(x-y)\varphi_0(y)dydx \right| \\
&+ \sum_{i=2}^{K} \left[ \int_{-\infty}^{a} \int_{a}^{b} \int_{a}^{b} M_{n+1}(x-y)M_{n+1-1}\cdots M_{n+1}(y,z)f_{n}'(z)dzdydx \\
&- \lambda_{i-1}^{i-1} \int_{-\infty}^{a} \int_{a}^{b} M_0(x-y)\varphi_0(y)dydx \right] \\
&+ \sum_{i=K+1}^{\infty} \int_{-\infty}^{a} \int_{a}^{b} \int_{a}^{b} M_{n+1}(x-y)M_{n+1-1}\cdots M_{n+1}(y,z)f_{n}'(z)dzdydx \\
&+ \frac{\lambda_0^K}{1-\lambda_0} \int_{-\infty}^{a} \int_{a}^{b} M_0(x-y)\varphi_0(y)dydx
\end{align*}
\]

It follows from Equation 3.2.7 that \( \lambda_0 < 1 \). Thus the fourth term above can be made arbitrarily small for large enough \( K \) independently of \( n \). That an appropriate choice of \( K \) also makes the third term arbitrarily small can be argued as follows. The third term is, as in the stationary case, the probability that a random walk with initial density \( f_{n}'(y) \) is absorbed no sooner than stage \( K+1 \) and then at \( a \), and this probability is no greater than the probability that the walk
proceeds through $K$ stages. The convergence to $M_o(t)$ of the densities $M_n(t)$ implies, due to Equation 3.2.7 that there exists an $N_0$ such that for $n \geq N_0$,

$$\int_{a}^{b} M_n(t) dt \leq 1 - \eta/2$$

So that the third term is, for $n \geq N_0$, no greater than $(1 - \eta/2)^K/\eta/2$.

It remains to show, then, that the first two terms tend to zero with $n$.

The first term is bounded above by

$$\int_{a}^{b} \left| f_n'(y) - \tilde{f}_n(y) \right| dy + \int_{a}^{b} \int_{-\infty}^{\infty} \left( M_n(x-y) - M_o(x-y) \right) dx \left| \tilde{f}_o(y) \right| dy$$

The expression on the left tends to zero with $n$ by Theorems 2.4.1 and 2.1.1. If

$$g_n(y) = \int_{-\infty}^{a} (M_n(x-y) - M_o(x-y)) dx$$

it follows from a theorem due to Scheffé (1947) that $g_n(y) \xrightarrow{n} 0$. $g_n(y) \leq 1$ and $\tilde{f}_o(y)$ is bounded. Thus by the Lebesque dominated convergence theorem the expression on the right tends also to zero with $n$. 
Finally, to show that the second term tends to zero with \( n \), consider the \( r \)th, \( 2 \leq r \leq K-2 \), of the \( K-2 \) addends.

\[
\left| \int_{-\infty}^{a} \int_{a}^{b} M_{n+r}(x-y)M_{n+r-1} \cdots M_{n+1}(y,z)f_n'(z)dzdydx \right|
\]

\[
- \lambda_{r-1}^{r-1} \int_{-\infty}^{a} \int_{a}^{b} M_{o}(x-y)\phi_{o}(y)dydx
\]

\[
= \left| \int_{-\infty}^{a} \int_{a}^{b} M_{n+r}(x-y)G_{r,n}(y)dydx - \int_{a}^{b} \int_{-\infty}^{a} M_{o}(x-y)H_{r,n}(y)dydx \right|
\]

where

\[
G_{r,n}(y) = \int_{a}^{b} M_{n+r-1} \cdots M_{n+1}(y,z)f_n'(z)dz
\]

and

\[
H_{r,n}(y) = \int_{a}^{b} M_{o}^{(r-1)}(y,z)\phi_{o}(z)dz
\]

Now

\[
\left| G_{r,n}(n) - H_{r,n}(y) \right| = \int_{a}^{b} \left( M_{n+r-1} \cdots M_{n+1}(y,z) - M_{o}^{(r-1)}(y,z)f_n'(z)dz \right)
\]
\[ \int_a^b M_0^{(r-1)}(y,z)(f'_n(z) - \phi'(z))dz \]

\[ \sup_z f'_n(z) \int_a^b \left| M_{n+r-1} \cdots M_{n+1}(y,z) - M_0^{(r-1)}(y,z) \right| dz \]

\[ + \int_a^b \left| f'_n(z) - \phi'(z) \right| dz \]

and for fixed \( r \), \( \lim_{n \to \infty} G_{r,n}(y) - H_{r,n}(y) = 0 \).

The second term can be written

\[ \int_a^b \int_a^b M_{n+r}(x-y)G_{r,n}(y)dydx - \int_a^b \int_a^b M_0(x-y)G_{r,n}(y)dydx \]

\[ + \int_a^b \int_a^b M_0(x-y)H_{r,n}(y)dydx - \int_a^b \int_a^b M_0(x-y)H_{r,n}(y)dydx \]

\[ \leq \int_a^b \int_a^b (M_{n+r}(x-y) - M_0(x-y))dx \left| G_{r,n}(y) \right| dy \]

\[ + \int_a^b \left| G_{r,n}(y) - H_{r,n}(y) \right| dy \]
The sequences $G_{r,n}$ and $H_{r,n}$ are bounded for every $r$. By Scheffé's theorem and the dominated convergence theorem, the first expression tends with $n$ to zero, and the latter theorem, with the boundedness of the sequences, assures the convergence to zero of the second expression.

Thus for a sequence of densities $M_{n}(t)$ converging to a density $M_{o}(t)$, the CAUOC function is given by

$$\text{CAUOC} (M_{o}) = \frac{1}{1-\lambda_{o}} \int_{a}^{b} \int_{-\infty}^{\infty} M_{o}(x-y)\phi_{o}(y)dydx$$

(3.2.9)

It is clear, for the normal GSPRT with step density Equation 3.2.4, that the limiting density is

$$M_{\mu}(x-y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-y+\bar{\mu}-\mu)^2}$$

and a truncation point $T$ which approximately equates the CATOC and CAUOC functions is that for which

$$\int_{a}^{T} \hat{\phi}_{\mu}(x)\hat{\phi}_{\mu} = \frac{1}{1-\lambda_{\mu}} \int_{a}^{b} \int_{-\infty}^{\infty} M_{\mu}(x-y)\phi_{\mu}(y)dydx.$$ 

An additional example is provided by a GSPRT for testing $\theta_{1}$ vs. $\theta_{2}$ for the family of densities given by Equation 3.1.11. The GSPRT is based on successive inequalities.
\[ A_n < \left( b(\theta_2) - b(\theta_1) \right) \sum_{1}^{n} (x_i + \delta) < B_n \]

or equivalently, in case \( b(\theta_2) > b(\theta_1) \), on

\[ a_n = \frac{A_n}{b(\theta_2) - b(\theta_1)} < \sum_{1}^{n} (x_i + \delta) < \frac{B_n}{b(\theta_2) - b(\theta_1)} = b_n \]

where \( \delta \) is given by Equation 3.1.12.

If the sequences \( \{a_n\} \) and \( \{b_n\} \) converge to \( a \) and \( b \) respectively, and if \( b_n - a_n = c \) for all \( n \), this walk is equivalent to one with barriers at \( a \) and \( b \) with nth stage step size density given by

\[ M_n(t) = e^{a(t) + [b(t) - b(t_o)][t-\delta+c_n-c_{n-1}] - |t+c_n-c_{n-1}|} \]

where

\[ c_n = b_n - b = a_n - a \]

For this case the limiting density is

\[ M_\infty(t) = e^{a(t) + [b(t) - b(t_o)][t-\delta] - |t|} \]

and the truncation point \( T \) which approximately equates the CATOC and CAUOC functions is that for which
\[ \int_a \delta_0(x)dx = \frac{1}{1-\lambda_0} \int_a \int_{-\infty}^{\frac{1}{2}} M_{\theta}(x-y)\phi_0(y)dydx \]

where \( \lambda_0 \) and \( \phi_0 \) are found using Equations 3.1.6 and 3.1.7 in conjunction with Equations 3.1.13 and 3.1.14.


This research was supported by the National Institutes of Health and the National Science Foundation.

The problem was suggested and the research directed by Prof. Herbert T. David, who has contributed substantially to the results obtained and to whom the author is grateful on many counts.

The author recalls with gratitude the kindly encouragement given her by Prof. Theodore A. Bancroft.

When no statement of appreciation is adequate, perhaps the simplest is least inappropriate. The author wishes to thank William Lyon.