SEQUENTIAL METHODS IN STATISTICS
BANACH CENTER PUBLICATIONS, VOLUME 16
PWN–POLISH SCIENTIFIC PUBLISHERS
WARSAW 1985

SEQUENTIAL PROBABILITY RATIO TESTS FOR STOCHASTIC PROCESSES: A REVIEW NOTE

NORBERT SCHMITZ

Institute of Mathematical Statistics, University of Münster,
Münster, G.F.R.

A survey is given on some properties of sequential probability ratio tests when these are applied for hypotheses on a stochastic process. In particular the termination property, the exponential boundedness of the sample size, and optimality properties are discussed for both the cases of simple and of composite hypotheses.

1. Introduction; Wald's SPRT for simple hypotheses in the i.i.d. case

The basic idea of sequential analysis consists of the following: Instead of choosing a fixed sample size for a statistical procedure one considers, at each stage of the investigation, the informations already received, and examines whether these are sufficient to terminate the experiment or whether another observation should be taken leading to additional costs. A. Wald was the first who systematically considered sequential decision procedures; in particular he developed the sequential probability ratio test (SPRT) which subsequently played a key-role in sequential analysis:

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed (i.i.d.) random variables. It is desired to decide between two simple hypotheses

$$H_i\colon\ P^{X_n} = Q_i; \qquad i = 1, 2, \qquad n \in N,$$

concerning the distribution of the X_n where, of course, $Q_1 \neq Q_2$. Denoting the Radon–Nikodym derivatives of the hypothetical distributions with respect to a σ-finite dominating measure by f_i, we define the *likelihood ratio* by

$$q(x_1, \ldots, x_n) := \frac{f_2(x_1, \ldots, x_n)}{f_1(x_1, \ldots, x_n)} = \prod_{\nu=1}^{n} \frac{f_2(x_\nu)}{f_1(x_\nu)}, \tag{1}$$

where $(x_1, \ldots, x_n)$ denotes the sample values.

This statistic is used to define a sequential test δ_{k_1,k_2} – Wald's SPRT – using the *stopping rule*

$$N = N_{k_1,k_2} := \inf\{n \in N\colon\ q(x_1, \ldots, x_n) \notin (k_1; k_2)\} \tag{2}$$

where k_1, k_2 are (suitable) constants

$$0 < k_1 < 1 < k_2 < \infty$$

$(\inf \emptyset := \infty)$, and by the *decision rule* $\varphi = (\varphi_n)_{n \in N}$ where

$$\varphi_n(x_1, \ldots, x_n) := \begin{cases} d_1 & \text{if } q(x_1, \ldots, x_n) \leqslant 1, \\ d_2 & \text{if } q(x_1, \ldots, x_n) > 1 \end{cases} \tag{3}$$

and d_i denotes the decision in favour of H_i, $i = 1, 2$ (it would suffice to define φ_n on $\{N_{k_1,k_2} = n\}$ resp.).

Since the sample size of SPRT is not fixed in advance but is, because of its dependence on the observations, a random variable the following questions arise:

(i) Will the test δ_{k_1,k_2} terminate, with probability 1, in a finite time, i.e.,(1)

$$P_i(N_{k_1,k_2} < \infty) = 1, \qquad i = 1, 2?$$

(Tests which give rise to positive probability for infinite sample sizes will rarely be suitable for practice.)

(ii) Do the expected sample sizes

$$E_i(N_{k_1,k_2}), \qquad i = 1, 2,$$

exist? (These values will, e.g., be of interest for comparisons with fixed sample size tests.)

(iii) Gains the SPRT δ_{k_1,k_2} advantages compared with other procedures, especially with classical tests?

For the i.i.d. case the answers are:

(1; i) ([36], p. 157) Every SPRT δ_{k_1,k_2} is *closed*, i.e., terminates with probability one in a finite time.

---

(1) $P_i := Q_i^N$ denotes the distribution of the sequence $(X_n)_{n \in N}$ under H_i; the expectation under P_i is denoted by E_i, $i = 1, 2$.

(1; ii) ([32]) For every SPRT δ_{k_1,k_2} the sample size N_{k_1,k_2} is exponentially bounded, i.e., there exist moments

$$E_i\big((N_{k_1,k_2})^k\big), \qquad i = 1, 2,$$

of arbitrary order $k \in N$.

(1; iii) ([37], [41], see also [17], p. 98, [5], [19], [22], [8], p. 365, [10], p. 93, [31], [18]) Let δ_{k_1,k_2} be an SPRT with error probabilities

$$\alpha_i(\delta_{k_1,k_2}) := P_i(\varphi_{N_{k_1,k_2}} = d_{3-i}), \qquad i = 1, 2,$$

and $\delta' = (N', \varphi')$ another test with

$$\alpha_i(\delta') := P_i(\varphi' = d_{3-i}) \leqslant \alpha_i(\delta_{k_1,k_2}), \qquad i = 1, 2. \tag{4}$$

Then

$$E_i(N_{k_1,k_2}) \leqslant E_i(N'), \qquad i = 1, 2, \tag{5}$$

and both inequalities in (5) are strict if

$$\big(\alpha_1(\delta'), \alpha_2(\delta')\big) \neq \big(\alpha_1(\delta_{k_1,k_2}), \alpha_2(\delta_{k_1,k_2})\big)$$

thus the SPRT's are *(uniformly) optimal* with respect to the expected sample size.

Moreover, an important step in the proof of (1; iii) consists in showing a further optimality property of the SPRT's:

(1; iv) ([37], see also [17], p. 105) Let s_i be the loss due to the false decision d_{3-i}, $i = 1, 2$, let the observation costs increase linearly (with slope c), and let $(\pi, 1-\pi)$ be an a priori distribution on (H_1, H_2). Then the Bayes risk

$$\pi[s_1\alpha_1(\delta) + cE_1(N)] + (1-\pi)[s_2\alpha_2(\delta) + cE_2(N)]$$

is either minimized by an SPRT or by a test which decides without any observation (2).

2. SPRT's for simple hypotheses on a stochastic process

Especially in bio-sciences and in social-sciences (where people can learn) the assumptions of Section 1

$$\text{independent observations,} \tag{6}$$

$$\text{repetitions, i.e., same distribution in each step,} \tag{7}$$

$$\text{unlimited observability,} \tag{8}$$

(2) These tests may formally be subsumed under the SPRT's by defining $q_0 := 1$ and admitting constants $k_1 > 1$ and $k_2 < 1$ respectively.

are rather severe restrictions for the applicability of the model. One is therefore led to consider time-discrete stochastic processes

$$(\Omega, \mathscr{S}, P; X = (X_n)_{n\in N})$$

or even more general processes

$$(\Omega, \mathscr{S}, P; X = (X_t)_{t\in T}), \qquad T \text{ an ordered set;}$$

– to avoid technicalities we formulate, in the sequel, several propositions only for time-discrete stochastic processes though generalizations are possible.

Firstly we note (see [35], p. 130) that the SPRT may be defined also in the case of two simple hypotheses

$$H_i\colon\ P^X = P_i \qquad i = 1, 2$$

on a stochastic process in just the same way as before (see (2), (3)) – merely the representability of

$$q(x_1, \dots, x_n) = \frac{f_2(x_1, \dots, x_n)}{f_1(x_1, \dots, x_n)}$$

as product of the single likelihood-ratios is lost. But again the questions (i)–(iii) arise – and now the answers are not as gratifying as in the i.i.d. case:

Simple examples (see, e.g., [24]) show that even the closedness of the SPRT's is lost in general – and therefore also question (ii) has a negative answer. But it turns out that those testing problems where all SPRT's are closed may be characterized in a rather satisfactory way:

THEOREM 2.1 ([26]/[30])(3). *The following statements are equivalent*:

(a) *Every* SPRT δ_{k_1,k_2} *is closed.*

(b) *For every* $(\alpha_1, \alpha_2) > (0, 0)$ *there exists a closed* SPRT δ_{k_1,k_2} *such that*

$$\alpha_i(\delta_{k_1,k_2}) \leqslant \alpha_i, \qquad i = 1, 2.$$

(c) *For every* $(\alpha_1, \alpha_2) > (0, 0)$ *there exists a closed test* δ *such that* $\alpha_i(\delta) \leqslant \alpha_i$, $i = 1, 2$.

(d) *The measures* P_1, P_2 *are orthogonal,* i.e., *there exists a set* A *such that* $P_1(A) = 1$, $P_2(A) = 0$.

Criterion (d) often yields rather simple conditions/proofs for the closedness of SPRT's (e.g., 0–1-laws, laws of large numbers etc.) – see, e.g., Remark (2.4).

Generalizing a result of Savage and Savage ([23]) one can, moreover, derive sufficient conditions for the exponential boundedness of N_{k_1,k_2} – the

(3) This proposition may, in an obvious way, be generalized to multiple decision problems (see [29]).

most interesting of these leads to ensure the existence of $E_i(\exp N_{k_1,k_2} z)$ in a neighbourhood of $z = 0$ (see [26]). One obtains therefore also a partial result for question (ii). The answer to question (iii) differs essentially from that in the i.i.d. case; the following proposition which combines results of the papers [24] and [27] shows that the SPRT's will no longer have a uniform optimality property:

Remark 2.2. If just one of the assumptions (6), (7), (8) is violated then there does not exist, in general, any test which is uniformly optimal with respect to the expected sample size.(4)

On the other hand, this remark shows that the loss of the optimality property is no real objection against the SPRT's. Moreover, the following weak "optimality" property can be shown to hold for every SPRT:

Remark 2.3 ([7]). Every SPRT δ_{k_1,k_2} is weakly admissible, i.e. there does not exist any test $\delta' = (N', \varphi')$ such that

$$N' \leqslant N_{k_1,k_2} \quad P_i\text{-a.s.} \qquad \text{and} \qquad \alpha_i(\delta') \leqslant \alpha_i(\delta_{k_1,k_2}), \quad i = 1, 2,$$

with at least one of these inequalities (with positive probability) strict.

Beyond this, the next remark indicates that the SPRT's may be of practical interest also for certain stochastic processes:

Remark 2.4 ([24], [25]). Let $(X_n)_{n\in N}$ be a homogeneous Markov chain with finite state space. Then each SPRT δ_{k_1,k_2} is closed and N_{k_1,k_2} is exponentially bounded iff there does not exist any achievable subchain where both hypotheses coincide. In this case it may be expected that the SPRT's save about $70^0/_0$ of the sample size of corresponding fixed sample size tests.

An examination of Wald's original proof ([35], [36]) shows that his approximations of the error probabilities are also valid in the general case (see [35], p. 130–132, [38], p. 29):

LEMMA (2.5). (a) *If the* SPRT δ_{k_1,k_2} *is closed then*

$$\frac{\alpha_1(\delta_{k_1,k_2})}{1-\alpha_2(\delta_{k_1,k_2})} \leqslant k_1; \qquad \frac{\alpha_2(\delta_{k_1,k_2})}{1-\alpha_1(\delta_{k_1,k_2})} \leqslant 1/k_2.$$

(b) *Let for* $(\alpha_1, \alpha_2) > (0, 0)$ *an* SPRT *be defined by*

$$k_1 := \frac{\alpha_1}{1-\alpha_2}, \qquad k_2 := \frac{1-\alpha_1}{\alpha_2}.$$

If δ_{k_1,k_2} *is closed then*

$$\alpha_1(\delta_{k_1,k_2}) + \alpha_2(\delta_{k_1,k_2}) \leqslant \alpha_1 + \alpha_2.$$

(4) But it should be noted that the optimality property of the SPRT does not characterize the i.i.d. case (see [28]).

Moreover, Ghosh ([11]) obtained as an auxilliary result:

Remark (2.6). Every SPRT is unbiased.

Some further justification for the use of the SPRT in the nonindependent case is given by Ghosh ([10], p. 99).

On the other hand, Wald's approximations (see [36], p. 56) of the average sample number (ASN) will, in general, no longer be of real use – the reason is that they are based on Wald's equation (see [36], p. 53) which is proved under the i.i.d. assumption (but see also [40], [10], p. 70–82). Instead only rather crude bounds for the ASN are known (see, e.g., [12], [25]).

Moreover, also the Bayesian optimality property (1; iv) of the SPRT's is lost in general (see [24]); instead the stopping rule and the terminal decision have to take into account the "state" of the observed system ([24]). So the SPRT's do not play, in the general case, such a dominating role as in the i.i.d. case; nevertheless they are still of practical importance – in particular since a commonly accepted "scale" for comparing different tests seems to be missing.

Finally it should be noted that for the "canonical" extension of the i.i.d. case to time-continuous stochastic processes, the properties (1; i)–(1; iv) can also be extended:

Remark 2.7 ([6]), [13]). Let $T = [0; \infty)$ be the time set, $(\mathscr{A}_t)_{t\in T}$ a right-continuous filtration, and assume that the log-likelihood process (with respect to $\mathscr{A}_t$) is (under P_i) an integrable stochastic process with stationary independent increments which is continuous in P_i-probability. If $\delta_{k_1,k_2} = (\tau, \varphi)$ is an SPRT(5) with error probabilities $\alpha_i(\delta_{k_1,k_2})$ and $\delta' = (\tau', \varphi')$ is a (sequential) test such that

$$\alpha_i(\delta') \leqslant \alpha_i(\delta_{k_1,k_2}), \qquad i = 1, 2,$$

then

$$E_i(\tau') \geqslant E_i(\tau), \qquad i = 1, 2, \tag{9}$$

and both inequalities in (9) are strict if

$$\big(\alpha_1(\delta'), \alpha_2(\delta')\big) \neq \big(\alpha_1(\delta_{k_1,k_2}), \alpha_2(\delta_{k_1,k_2})\big).$$

In particular this optimality property of the SPRT's holds if the probability measures P_i belong to a family of the exponential class (see [9], [16], 2.2).

3. The SPRT for composite hypotheses in the i.i.d. case

The SPRT (2), (3) is constructed as a sequential test of one simple hypothesis against another. On the other hand, from a practical point of view the

(5) The stopping rule τ and the decision rule φ are defined analogously to (2), (3).

problem of testing two simple hypotheses seems to be rather artificial – in most real problems it is desired to decide between two composite hypotheses

$$H_i:\ P^X \in \mathscr{P}_i, \qquad i = 1, 2, \tag{10}$$

where $\mathscr{P}_i = \{P_\theta\colon \theta \in \Theta_i\}$ are disjoint families of possible distributions of the observations. But as in the "classical" theory (comp., e.g., the Neyman–Pearson lemma) one hopes that the consideration of simple hypotheses will be a first step for a successful treatment of the more complicated problem (10).

There is no unique way to construct an SPRT for composite hypotheses; the following three proposals seem to be the most important ones:

(a) Two special distributions $P_i \in \mathscr{P}_i$, $i = 1, 2$, are selected and used to define an SPRT according to (2), (3). Since this method can also be used in the general case we will discuss it in some detail.

(b) The composite hypotheses are reduced to simple ones by using the principle of invariance, if applicable, and then an invariant SPRT is defined in an obvious way. For the problems arising with this method we refer to the excellent review article [39] of Wijsman.

(c) Already Wald ([35]) proposed the so-called method of weight functions: For each $\mathscr{P}_i$ a prior distribution is chosen and the families $\mathscr{P}_i$ are "replaced" by the according "mixture" of the distributions in $\mathscr{P}_i$, $i = 1, 2$. Thus the composite hypotheses are reduced to simple ones and an SPRT can be defined in the previous way (2), (3). For the choice of the prior "mixtures" certain guidelines were suggested by Wald.

For a short discussion of proposal (a) let us firstly remark that in this case analogous questions as (i)–(iii) arise – but now more complicated since one has to take into account all possible P_θ, $\theta \in \Theta := \Theta_1 \cup \Theta_2$.

Again the answers to questions (i) and (ii) are satisfactory:

THEOREM 3.1 ([36], [32], see also [10], p. 118). *Suppose that for all* $\theta \in \Theta$

$$P_\theta\left(\frac{f_2(X_1)}{f_1(X_1)} = 1\right) < 1; \tag{11}$$

then every SPRT δ_{k_1,k_2} *satisfies*

(i) δ_{k_1,k_2} *is closed*, i.e.,

$$\lim_{n\to\infty} P_\theta(N_{k_1,k_2} > n) = 0 \qquad \textit{for all} \quad \theta \in \Theta;$$

(ii) N_{k_1,k_2} *is exponentially bounded*, i.e.,

$$E_\theta\big((N_{k_1,k_2})^k\big) < \infty \qquad \textit{for all} \quad k \in N \ \textit{and} \ \theta \in \Theta.$$

Condition (11) may easily be checked; it is, e.g., fulfilled if $f_2(X_1)/f_1(X_1)$ has positive variance.

But the answer to question (iii) yields at the same time one of the main objections against the SPRT: Though every SPRT minimizes the ASN for the special distributions P_i, $i = 1, 2$, it may have rather unsatisfactory properties for "intermediate" distributions P_θ. For the time-continuous case of a Wiener process with known variance σ, hypotheses

$$H_1\colon\ \mu \leqslant \mu_1; \qquad H_2\colon\ \mu \geqslant \mu_2 := \mu_1 + \sigma/5,$$

and error bounds

$$\alpha_1 = \alpha_2 = 0.01,$$

Ghosh ([10], p. 238) gives, e.g., the average sample time for seven rival tests at μ_i and $\mu_3 := (\mu_1 + \mu_2)/2$; he obtains in particular:

Procedure	$E_1(\tau) = E_2(\tau)$	$E_3(\tau)$
Fixed sample	541.2	541.2
SPRT	225.2	527.9
Special truncated test	242.8	437.7
Special "Anderson" ([1]) test	249.4	402.2

A *uniformly* optimal test will therefore not exist in general. It is even possible that an SPRT needs, at an intermediate P_θ, an ASN which exceeds that of a competitive fixed sample size test (see [1], [10], p. 141).

On the other hand for special one-parameter families of distributions a restricted optimality property of the SPRT's can be proven (see [10], p. 105).

Therefore another "scale" for comparing different (sequential) tests is needed. Two main proposals, beside the decision-theoretical concepts (Bayes and minimax procedures), consist of the following:

(i) Construct a (sequential) test which minimizes $\sup_{\theta \in \Theta} E_\theta(N)$ out of all closed tests of strength (α_1, α_2) (see [14]) – this concept leads, for certain families of distributions, to truncated generalized SPRT's.

(ii) Construct a (sequential) test which has a local optimality property – e.g., maximizes the slope of the OC function at a "boundary" point θ_0.

4. Remarks on the SPRT for composite hypotheses on a stochastic process

Also in the case of composite hypotheses

$$H_i\colon\ P^X \in \mathscr{P}_i, \qquad i = 1, 2,$$

on a stochastic process it makes sense to select two special distributions $P_i \in \mathscr{P}_i$, $i = 1, 2$, and to use these to define an SPRT.

But in order to show what may happen here in connection with our question (i) we consider a simple example which essentially goes back to [33]:

EXAMPLE 4.1. Let two distributions concerning a homogeneous Markov chain with 3 states be given by the initial distributions

$$^1p = \begin{pmatrix} 10/31 \\ 9/31 \\ 12/31 \end{pmatrix}, \qquad ^2p = \begin{pmatrix} 87/248 \\ 81/248 \\ 80/248 \end{pmatrix}$$

and the transition matrices

$$^1\boldsymbol{P} = \begin{pmatrix} 3/10 & 3/10 & 2/5 \\ 1/3 & 1/6 & 1/2 \\ 1/3 & 3/8 & 7/24 \end{pmatrix}, \qquad ^2\boldsymbol{P} = \begin{pmatrix} 1/6 & 1/2 & 1/3 \\ 1/2 & 1/6 & 1/3 \\ 2/5 & 3/10 & 3/10 \end{pmatrix}$$

– the initial distributions ^ip are just the stationary distributions corresponding to the $^i\boldsymbol{P}$. The $^i\boldsymbol{P}$ are irreducible and aperiodic; the P_i are orthogonal (see (2.1)). But let now the "true" distribution P be given by the transition matrix

$$\boldsymbol{P} = \begin{pmatrix} 0 & 0 & 1 \\ 3/4 & 1/4 & 0 \\ 1/3 & 2/3 & 0 \end{pmatrix}$$

– this $\boldsymbol{P}$ is irreducible and aperiodic, too; P is orthogonal to both P_i.

On the other hand, one easily verifies by direct calculations that for

$$\begin{aligned} x_0 = 1: & \\ x_0 = 2: & \quad \prod_{j=1}^{n} \frac{^2p_{x_{j-1},x_j}}{^1p_{x_{j-1},x_j}} \\ x_0 = 3: & \end{aligned} \begin{aligned} &\in \{5/6,\ 1,\ 2/3\} \\ &\in \{3/2,\ 1,\ 5/4\} \\ &\in \{6/5,\ 4/5,\ 1\} \end{aligned} \qquad \text{for all } (x_1, \ldots, x_n),\ n \in \boldsymbol{N},\ P\text{-a.s.}$$

This yields that every SPRT δ_{k_1,k_2} with

$$k_1 < 2/3, \qquad k_2 > 27/16$$

will never terminate (with probability 1). Therefore, to ensure even the termination property of the SPRT's rather strong assumptions have, in general, to be made.

Remarks 4.2. (a) Example 4.1 turns out to be a counterexample as well against Lemma 2 (the reference to [34] seems to be erroneous) as against Lemmas 3 and 4 of [21].

(b) Condition V.2 of [15] avoids the difficulties arising in Example 4.1; but to verify that condition it seems necessary to know the "true" distribution (then in fact no longer a statistical problem exists).

(c) Using a concept of degeneracy due to [20] it is moreover, possible to rectify the results of Phatarfod ([21]), but then the same difficulties arise as in the case treated by Küchler ([15]).

Observing that a nonincreasing sequence $(a_n)_{n\in N}$ with

$$0 < a_n \leqslant 1, \quad \forall n \in N,$$

fulfills

$$\lim_{n\to\infty} a_n = 0 \qquad \text{iff} \qquad \sum_{n=1}^{\infty} (1 - a_{n+1}/a_n) = \infty,$$

one obtains a simple characterization of the termination property (but which seems not to be of any practical importance):

Remark 4.3 (see [10], p. 161). Let $\delta = (N, \varphi)$ be a sequential test with $P_\theta(N > n) > 0$ for all(6) $n \in N$. Then δ is closed iff for each $\theta \in \Theta$

$$\sum_{n=1}^{\infty} [1 - P_\theta(N > n+1)/P_\theta(N > n)] = \infty.$$

The fact that

$$\lim_{n\to\infty} P_\theta\big(k_1 < q(X_1, \ldots, X_n) < k_2\big) = 0, \ \forall \theta \in \Theta,$$

is a sufficient condition for the termination property of the SPRT δ_{k_1,k_2}, is used by Ghosh ([10], p. 121, 161) to derive several sufficient conditions for SPRT's to be closed. But it seems, from a practical point of view, to be more fruitful to consider special classes of stochastic processes and to prove conditions for the termination property by using their special structure:

Remarks 4.4. (a) Stochastic processes whose log-likelihood process has stationary independent increments allow conditions analogous to those of Theorem 3.1.

(b) For independent but not identically distributed X_i the condition

$$\inf_{i\in N} \operatorname{Var}_\theta\big(f_2(X_i)/f_1(X_i)\big) > 0 \qquad \text{for each} \quad \theta \in \Theta$$

is sufficient for the termination property of the SPRT's (see [10], p. 161).

(c) Conditions for homogeneous Markov chains with finite state space were already mentioned in Remark 4.2.

(d) The termination property of SPRT's for timecontinuous Markov processes was considered by Andrieu ([2]).

Moreover, for some special cases also conditions for the exponential boundedness of N_{k_1,k_2} are derived (see e.g. [15]). But even a partial answer to question (iii) seems, as far as we know, to be missing – Ghosh ([10], 146–148) gives some arguments that the SPRT's in the general case will be expected not to have any optimality property (see also 2.2).

(6) $P_\theta(N > n_0) = 0$ yields no problem of closedness.

References

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

[2] C. Andrieu, *Une formulation, et la finitude, du test séquentiel du rapport des vraisemblances dans le cas Markovien à temps continu*, Z. Wahrsch. Verw. Gebiete **32** (1975), 333–344.

[3] I. Basawa and B. Prakasa Rao, *Statistical Inference for Stochastic Processes*, Academic Press, London 1980.

[4] B. Bhat, *Bayes solution of sequential decision problems for Markov dependent observations*, Ann. Math. Statist. **35** (1964), 1656–1662.

[5] D. Burkholder and R. Wijsman, *Optimum properties and admissibility of sequential tests*, Ann. Math. Statist. **34** (1963), 1–17.

[6] A. Dvoretzky, J. Kiefer and J. Wolfowitz, *Sequential decision problems for processes with continuous time parameter. Testing hypotheses*, Ann. Math. Statist. **24** (1953), 254–264; (**30** (1959), 1265).

[7] B. Eisenberg, B. Ghosh and G. Simons, *Properties of generalized sequential probability ratio tests*, Ann. Statist. **4** (1976), 237–251.

[8] T. Ferguson, *Mathematical Statistics. A Decision Theoretic Approach*, Academic Press, New York 1967.

[9] J. Franz and W. Winkler, *Über Stoppzeiten bei statistischen Problemen für homogene Prozesse mit unabhängigen Zuwächsen*, Math. Nachr. **70** (1976), 37–53.

[10] B. Ghosh, *Sequential Tests of Statistical Hypotheses*, Addison–Wesley, Reading 1970.

[11] J. Ghosh, *On the monotonicity of the OC of a class of sequential probability ratio tests*, Calcutta Statist. Assoc. Bull. **9** (1960), 139–144.

[12] W. Hall, *Embedding submartingales in Wiener processes with drift, with applications to sequential analysis*, J. Appl. Probab. **6** (1969), 612–632.

[13] A. Irle and N. Schmitz, *On the optimality of the SPRT for processes with continuous time parameter*, Math. Operationsforsch. Statist. Ser. Statist. **15** (1984), 91–104.

[14] J. Kiefer and L. Weiss, *Some properties of generalized sequential probability ratio tests*, Ann. Math. Statist. **28** (1957), 57–70, 14–17, 72–74.

[15] I. Küchler, *Der sequentielle Quotiententest bei irreduziblen homogenen Markovschen Ketten mit endlichem Zustandsraum*, Math. Operationsforsch. Statist. Ser. Statist, **9** (1978), 227–239.

[16] —, *On the exponential class of processes with independent increments-analytical treatment and sequential probability ratio tests*, Akademie der Wissenschaften der DDR, WSS-04/79, Berlin 1979.

[17] E. Lehmann, *Testing Statistical Hypotheses*, Wiley, New York 1959.

[18] G. Lorden, *Structure of sequential tests minimizing an expected sample size*, Z. Wahrsch. Verw. Gebiete **51** (1980), 291–302.

[19] T. Matthes, *On the optimality of sequential probability ratio tests*, Ann. Math. Statist. **34** (1963), 18–21.

[20] H. Miller, *A convexity property in the theory of random variables, defined on a finite Markov chain*, Ann. Math. Statist. **32** (1961) 1261–1270.

[21] R. Phatarfod, *Sequential analysis of dependent observations, I*, Biometrika **52** (1965), 157–165.

[22] H. Richter, *Optimale Sequenzwiederholungstests*, Z. Wahrsch. Verw. Gebiete **5** (1966), 156–166.

[23] I. Savage and L. Savage, *Finite stopping time and finite expected stopping time*, J. Roy. Statist. Soc. Ser. B **27** (1965), 284–289.

[24] N. Schmitz, *Likelihoodquotienten-Sequenztests bei homogenen Markoffschen Ketten*, Biometrische Z. **10** (1968), 231–247.

[25] –, *Stichprobenumfänge des LQST bei homogenen Markoff-Ketten*, Oper. Res. Verf. **VIII** (1970), 271–279.

[26] –, *Existenz von sequentiellen Tests zu vorgegebenen Niveaus*, Arch. Math. (Basel) **XXI** (1970), 617–628.

[27] –, *A note on the optimality of k-stage tests*, Metrika **19** (1972), 72–75.

[28] –, *Zur Optimalität des Likelihoodquotienten-Sequenztests*, Transactions Sixth Prague Conference 1971, 781–791, 1973.

[29] –, *Sequentielle Mehrentscheidungsverfahren mit vorgeschriebenem Irrtumsvektor*, Oper. Res. Verf. **XVII** (1973), 317–340.

[30] –, *Corrigendum zu "Sequentielle Mehrentscheidungsverfahren mit vorgeschriebenem Irrtumsvektor"* Oper. Res. Verf. **XXIV** (1977), 176–177.

[31] G. Simons, *An improved statement of optimality for sequential probability ratio tests*, Ann. Statist. **4** (1976), 1240–1243.

[32] C. Stein, *A note on cumulative sums*, Ann. Math. Statist. **17** (1946), 498–499.

[33] R. Strehl, *Bemerkungen zur exponentiellen Beschränktheit des Likelihoodquotienten-Sequenztests bei Vorliegen von Abhängigkeiten*, Unpublished Diplomarbeit, Münster 1978.

[34] M. Thornett, *Some results on the first latent root of a matrix* $(p_{ij}e^{th_{ij}})$ *where* (p_{ij}) *is a stochastic matrix*, University of Western Australia, Research Report, 1961.

[35] A. Wald, *Sequential tests of statistical hypotheses*, Ann. Math. Statist. **16** (1945), 117–186.

[36] –, *Sequential Analysis*, Wiley, New York 1947.

[37] A. Wald and J. Wolfowitz, *Optimum character of the sequential probability ratio test*, Ann. Math. Statist. **19** (1948), 326–339.

[38] B. Wetherill, *Sequential Methods in Statistics*, Chapman and Hall, London 1966.

[39] R. Wijsman, *Stopping time of invariant sequential probability ratio tests*, in: P. Krishnaiah (ed.), *Developments in Statistics*, vol. 2, 1979, 235–314.

[40] J. Wolfowitz, *The efficiency of sequential estimates and Wald's equation for sequential processes*, Ann. Math. Statist. **18** (1947), 215–230.

[41] –, *Remark on the optimum character of the sequential probability ratio test*, Ann. Math. Statist. **37** (1966), 726–727.

Presented to the semester
Sequential Methods in Statistics
September 7–December 11, 1981