The Generalized Discretized Sequential Probability Ratio Test and its Application in Insurance Mathematics

Body Math

by
Volker Mammitzsch
Department of Mathematics and Computer Science
Philipp University, D-35032 Marburg /Germany
E-mail: mammit@mathematik.uni-marburg.de

Abstract: WALD's famous sequential probability ratio test for comparing two simple
hypotheses P_1 and P_2 is extended to the case when instead of successive observations of i.i.d. random variables general onservations can be taken and final decisions are allowed at a discrete series of pre-assigned time points. It is shown that the following properties are equivalent:
(a) Each test of that type is closed.
(b) The errors of first and second kind of those tests can be made arbitrarily small.
(c) P_1 and P_2 are orthogonal probability measures.
This result is applied to the risk process in insurance mathematics and a uniqueness problem is discussed.

The discretized generalized sequential probability ratio test (DGSPRT) is a generalization of A. WALD's well-known sequential probability ratio test, cf. [8]. Therefore, we first give a short description of that testing procedure here, cf. also [1]

a. WALD's sequential probability ratio test

Let be $X_1, X_2, ...$ i.i.d. random variables with unknown density function f. The simple hypothesis "$f = f_1$" is to be tested against the simple alternative "$f = f_2$" ,where $f_1 \neq f_2$ are two known density functions.

1991 Mathematics Subject Classification: 62L10, 62M99

Key words and phrases: Closed sequenctial tests, error probabilities, orthogonal hypotheses, risk process, compound Poisson process, excess of loss.

Define the likelihood ratio

$$q^{(n)} := q^{(n)}(X_1, ..., X_n) := \frac{f_1(X_1)\cdots f_1(X_n)}{f_2(X_1)\cdots f_2(X_n)}, n = 1, 2, ...,$$

and, moreover, with $a, b \in \mathbb{R}$, $a < b$ the stopping rule

$$T := \inf \left\{ n : q^{(n)} \leq a \ \text{ or } \ q^{(n)} \geq b \right\}$$

and the (terminal) decision rule

$$\delta := \begin{cases} f_1 & \text{if } \ q^{(T)} \geq b \\ \\ f_2 & \text{if } \ q^{(T)} \leq a \end{cases} .$$

This means that we take observations as long as the likelihood ratio is strictly between a and b and stop taking observations as soon as the likelihood ratio leaves the open interval (a, b), and decide in favour of f_1 if the likelihood ratio in the moment of stopping is "big" and otherwise in favour of f_2.

The pair (T, δ) is called WALD's sequential probability ratio test. It has two remarkable properties.

1. $P(T < \infty \mid f_1) = P(T < \infty \mid f_2) = 1$

which says that the test with probability one comes to an end, no matter which one of the hypotheses f_1 or f_2 is true.

2. For each pair (α_1, α_2) with $0 < \alpha_i < 1$ $(i = 1, 2)$ there exist reals a, b with $a < b$ such that

$$P(\delta = f_2 \mid f_1) \leq \alpha_1 \ \text{ and } \ P(\delta = f_1 \mid f_2) \leq \alpha_2$$

which says that for appropriate a and b the error of the first kind and the error of the second kind can be made arbitrarily small.

b. The discretized generalized sequential probability ratio test

WALD's test has been modified in several ways in order to meet the needs for more general testing situations, see e.g. [2],. [7]. Here we will generalize the concept for the case when observations are possible in continuous time.

Definitions. Let be given

$(\mathcal{A}_t)_{t\in[0,\infty)}$ an increasing familiy of $\sigma-$algebras on Ω, $\mathcal{A}_t$ representing all possible observations until time t; $\mathcal{A}_\infty := \sigma\left(\bigcup_{t\in[0,\infty)} \mathcal{A}_t\right)$;

P_1 and P_2 probability measures on $\mathcal{A}_\infty$ called hypotheses; a sequence $0 \leq t_1 < t_2 < ...$ with $t_n \to \infty$ as $n \to \infty$ representing the time points in which a (terminal) decision is possible, $D := \{t_1, t_2, ...\}$.

Let denote

$P_i^{(n)} := P_i \mid \mathcal{A}_{t_n}$ the restriction of P_i to $\mathcal{A}_{t_n}$, $i = 1,2,$ $n = 1,2,...$; $P_i^{(\infty)} := P_i$ $i = 1,2,$;

$P := \frac{1}{2}(P_1 + P_2)$ on $\mathcal{A}_\infty$ a dominating measure of the family $\{P_1, P_2\}$,

$P^{(n)} := P \mid \mathcal{A}_{t_n}$ the restriction of P to $\mathcal{A}_{t_n}$, $n = 1,2,...$.

By the RADON-NIKODYM Theorem, there exist densities

$$f_i^{(n)} := \frac{dP_i^{(n)}}{dP^{(n)}}, \quad \text{i=1,2; n=1,2,....}$$

Without loss of generality we assume

$$0 \leq f_i^{(n)} \leq 2 \;\text{ and }\; f_1^{(n)} + f_2^{(n)} = 2 \;\text{ for all }\; i = 1,2;\; n = 1,2,... .$$

Define the (generalized) likelihood ratio $q^{(n)}$ according to

$$q^{(n)} := \begin{cases} f_1^{(n)} \Big/ f_2^{(n)} & \text{if } f_2^{(n)} > 0 \\ \infty & \text{otherwise} \end{cases}$$

and for reals a, b with $0 < a < 1 < b < \infty$ the **stopping rule**

$$T_{a,b} := \inf \left\{ t_n : q^{(n)} \leq a \ \text{ or } \ q^{(n(} \geq b \right\}$$

and the **decision rule**

$$\delta := \begin{cases} P_1 & \text{if } q^{(T_{a,b})} \geq b \\ P_2 & \text{if } q^{(T_{a,b})} \leq a \end{cases} .$$

The pair $(T_{a,b}, \delta)$ is called **generalized discretized sequential probability ratio test**
(GDSPRT).

We say that a GDSPRT is **closed** if and only if $P_i(T_{a,b} < \infty) = 1$ for $i = 1, 2$ or,
equivalently, $P(T_{a,b} < \infty) = 1$ which means that the test with probability one will end in finite time, no matter which one of the hypotheses is true.

Let us recall that the probability measures are called **orthogonal** if and only if there exists an
$A \in A_\infty$ such that $P_1(A) = 0$ and $P_2(A) = 1$.

In what follows we need some lemmas

Lemma 1.

$$\sigma \left(\bigcup_{t \in [0,\infty)} \mathcal{A}_t \right) = \sigma \left(\bigcup_{t \in D} \mathcal{A}_t \right).$$

The proof is obvious.

Lemma 2. $q^{(n)} \to q^{(\infty)} \quad P$ a.e. as $n \to \infty$.

Proof: According to the martingale convergence theorem we have $f_i^{(n)} \to f_i^{(\infty)}$ P a.e. as $n \to \infty$, $i = 1, 2$. Observing $0 \le f_i^{(n)} \le 2$ and $f_1^{(n)} + f_2^{(n)} = 2$ for $i = 1, 2, \ldots$, $n = 1, 2, \ldots$, we immediately see that $q^{(n)} \to q^{(\infty)}$ P a.e. as $n \to \infty$.

Lemma 3. P_1 and P_2 orthogonal implies $P(0 < q^{(\infty)} < \infty) = 0$.

Proof: From P_1 and P_2 orthogonal it follows

$$P\left(\left\{f_1^{(\infty)} > 0\right\} \cap \left\{f_2^{(\infty)} > 0\right\}\right) = 0$$

which implies

$$P\left(0 < q^{(\infty)} < \infty\right) = P\left(\left\{f_1^{(\infty)} > 0\right\} \cap \left\{f_2^{(\infty)} > 0\right\}\right) = 0$$

Remark: The converse of Lemma 3 is true, too.

We are now ready to prove an equivalence theorem, a simple version of which for the discrete case can be found in [4].

Theorem. The following statements are equivalent:

(i) Each DGSPRT is closed.

(ii) For each pair $0 < \alpha_1, \alpha_2 < 1$ there exists a closed DGSPRT $(T_{a,b}, \delta)$ with $0 < a < 1 < b < \infty$ such that

$$P_i(\delta \ne P_i) < \alpha_i, \quad i = 1, 2.$$

(iii) P_1 and P_2 are orthogonal.

Proof:

(i)⇒(ii): As $(T_{a,b},\delta)$ is closed, for suitable a we have

$$\begin{aligned} P_1(\delta \neq P_1) &= \sum_{k=1}^{\infty} \int\limits_{\{T_{a,b}=t_k,\ q^{(k)}\leq a\}} f_1^{(k)} dP \\ &\leq \sum_{k=1}^{\infty} \int\limits_{\{T_{a,b}=t_k\}} a\ f_2^{(k)} dP = a < \alpha. \end{aligned}$$

The proof for $P_2(\delta \neq P_2) \leq \frac{1}{b} < \alpha_2$ is similar.

(ii)⇒(iii): Choose a sequence of DGSPRT $(T^{(n)},\delta^{(n)})$ such that $P_i(\delta^{(n)} \neq P_i) < \frac{1}{2^n}$,

$i = 1,2.$

Moreover, define

$$A := \limsup_{n\to\infty} \left\{\delta^{(n)} \neq P_1\right\} = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} \left\{\delta^{(n)} \neq P_1\right\},$$

then

$$P_1(A) = \lim_{k\to\infty} P_1 \left(\bigcup_{n=k}^{\infty} \left\{\delta^{(n)} \neq P_1\right\} \right) \leq \lim_{k\to\infty} \sum_{n=k}^{\infty} \frac{1}{2^n} = 0$$

and

$$\begin{aligned} P_2(A) &= \lim_{k\to\infty} P_2 \left(\bigcup_{n=k}^{\infty} \{\delta^{(n)} \neq P_1\} \right) \\ &\geq \lim_{k\to\infty} P_2 \left(\delta^{(n)} \neq P_1\right) = \lim_{k\to\infty} \left(1 - P_2(\delta^{(k)} \neq P_2)\right) \\ &= 1 - \lim_{k\to\infty} \tfrac{1}{2^k} = 1. \end{aligned}$$

(iii)⇒(i): Let P_1 and P_2 be orthogonal and suppose that there exist reals a, b with

$0 < a < 1 < b < \infty$ such that

$$0 < P \left(\bigcap_{n=1}^{\infty} \left\{ a < q^{(n)} < b \right\} \right)$$

then from Lemma 2 we conclude

$$0 < P(a \leq q^{(\infty)} \leq b) \leq P(0 < q^{(\infty)} < b)$$

which contradicts Lemma 3.

c. An application in insurance mathematics

In collective risk theory we consider the so-called **risk process**

$$S_t = \sum_{j=1}^{N_t} X_j, \quad 0 \leq t < \infty,$$

where N_t denotes the number of claims in a portfolio up to time t with $(N_t)_{t\in[0,\infty)}$ supposed to be a Poisson process with parameter $\lambda > 0$,
$X_j \geq 0$ denoting the amount of the j-th claim, where $X_1, X_2, ...$ are i.i.d. with distribution
function F, $(N_t)_{t\in[0,\infty)}, X_1, X_2...$ independent.
S_t is the accumulated claim up to time $t \in [0, \infty)$ and is compound Poisson.
We want to test the hypothesis H_1 against the alternative H_2 given by

$$\begin{aligned} H_1 &: \lambda = \lambda_1, \quad F = F_1 \\ H_2 &: \lambda = \lambda_2, \quad F = F_2 \end{aligned}$$

where λ_i are known reals and F_i known distribution functions, $i = 1, 2$, with $(\lambda_1, F_1) \neq (\lambda_2, F_2)$.
Problem: Given arbitrarily small $\alpha_i > 0, i = 1, 2,$ is it possible to distinguish H_1 from H_2 by means of a DGSPRT with error probabilities of the first and second kind no larger than α_1 and α_2, respectively?
Answer: Yes!
A rough sketch of the proof can be found in [5]. A further generalization to

other stochastic processes are dealt with in [3] .

Proof: Define

$$\begin{aligned}
\mathcal{A}_t &:= \sigma(S_u : u \le t), \quad t \in (0,\infty] \\
\mathcal{A}_\infty &:= \sigma(S_u : u < \infty) = \sigma\left(\bigcup_{t\in(0,\infty]} \mathcal{A}_t\right) \\
D &:= \{t_1, t_2, ...\} \quad \text{as in part b.}
\end{aligned}$$

(i) If $F_1 \neq F_2$, then there exists an $x \in \mathbb{R}_+$ such that $F_1(x) \neq F_2(x)$. Consider the random variables

$$Z_j := 1_{[0,x]}(X_j), \quad j = 1, 2, ... \;.$$

Obviously, Z_j is measurable with respect to $\mathcal{A}_\infty$, and $Z_1, Z_2, ...$ are i.i.d. with $E(Z_j \mid H_i) = F_i(X)$, $i = 1, 2$. Moreover, by the strong law of large numbers, for

$$A := \left\{ \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} Z_j = F_2(x) \right\} \in \mathcal{A}_\infty$$

there holds $P(A \mid H_1) = 0$ and $P(A \mid H_2) = 1$. Thus, $P(\cdot \mid H_1)$ and $P(\cdot \mid H_2)$ are orthogonal, and our Theorem finishes the proof.

(ii) If $\lambda_1 \neq \lambda_2$, then choose a subsequence $t_j^* \in D$ with $t_{j+1}^* - t_j^* \geq 1$ for all $j = 1, 2, ...,$

and consider the random variables

$$Z_j := \frac{1}{t_{j+1}^* - t_j^*}(N_{t_{j+1}^*} - N_{t_j^*}), \quad j = 1, 2, ... \;.$$

Again, Z_j is measurable with respect to $\mathcal{A}_\infty$ for all $j = 1, 2, ...,$ and $Z_1, Z_2, ...$ are independent with

$$E(Z_j \mid H_i) = \lambda_i,$$
$$Var(Z_j \mid H_i) \leq \lambda_i, \quad i = 1,2; \quad j = 1,2,\dots \ .$$

As $\sum\limits_{j=1}^{\infty} \frac{1}{j^2} Var(Z_j \mid H_i) < \infty \ (i = 1,2),$ from Kolmogorov's convergence theorem using the same argument as in (i) we find that $P(\cdot \mid H_1)$ and $P(\cdot \mid H_2)$ are orthogonal, which as above finishes the proof.

d. A uniqueness problem

Usually, in part c we cannot observe λ and F directly but can only track the process

$(S_t)_{t\in[0,\infty)}$. This raises the following question, which was put to me by G. SIEGEL

(oral communication).

Problem: Is (λ, F) uniquely determined by the distribution of

$$(S_t)_{t\in[0,\infty)} \ ?$$

Answer: (1) Yes, if $F(0) = 0$, which means $P(X > 0) = 1$

(2) No, if $F(0) > 0$, which means $p := P(X > 0) < 1$, $q := P(X = 0) > 0$.

Proof: (1) For all $n = 1,2,\dots, x \geq 0$ there holds P a.e.

$$\begin{aligned} \{N_t = n\} &= \{(S_t) \text{ has } n \text{ jumps in } (0,t]\,, X_j > 0,\ j = 1,\dots,n\} \\ &= \{(S_t) \text{ has } n \text{ jumps in } (0,t]\}\,, \end{aligned}$$

thus λ is uniquely defined by $(S_t)_{t\in[0,\infty)}$. Furthermore, P a.e. holds $\{n\text{-th jump of } (S_t) > x\} = \{X_n > x,\ X_j > 0,\ j = 1,\dots,n-1\} = \{X_n > x\}$

thus F is uniquely defined by (S_t).

(2) Consider the thinned process (M_t) defined by

$$M_t = \text{number of jumps of } (N_t) \text{ with } X_{N_t} > 0,\ t \geq 0.$$

(M_t) is a Poisson process with parameter $\lambda_M = p\lambda$, c.f. [6] , Section 6.3. Moreover, define

$$G(x) = \begin{cases} \frac{1}{p}(F(x) - q) & \text{if } x \geq 0 \\ & \\ 0 & \text{if } x < 0 \end{cases} .$$

Clearly, G is a distribution function with $G \neq F$ and $G(0) = 0$. Let $Y_1, Y_2, ...$ be i.i.d. random variables with distribution function G and let be $(M_t), Y_1, Y_2, ...$ independent.

Define

$$R_t := \sum_{j=1}^{M_t} Y_j$$

and let denote $\varphi_t, \psi_t, \varphi_X$, and φ_Y the characteristic functions of S_t, R_t, X_t, and Y_t, respectively, then it holds for all real z

$$\begin{aligned} \varphi_x(z) &= \int_{[0,\infty)} e^{izx} dF(x) \\ &= e^{iz\cdot 0} \cdot q + \int_{(0,\infty)} e^{izx} dF(x) \\ &= q + \int_{(0,\infty)} e^{izy} p \, dG(y) = q + p \cdot \varphi_Y(z) \end{aligned}$$

and finally

$$\begin{aligned} \varphi_t(z) &= e^{it\lambda(\varphi_X(z)-1)} &&= e^{it\lambda(q+p\cdot\varphi_Y(z)-1)} \\ &= e^{it\lambda p(\varphi_Y(z)-1)} &&= \psi_t(z) \end{aligned} .$$

Part (2) of the answer has a nice application in reinsurance.

Consider an excess of loss contract which means that the reinsurer pays $Y_i = (X_i - \varepsilon)^+$ if a claim X_i occurs. Assume that there are two portfolios with claim distribution functions F_1 and F_2 and retentions ε_1 and ε_2 , respectively, which yield the same risk process to the reinsurer if and only if

$$(*) \quad \frac{1}{1-F_1(\varepsilon_1)}(F_1(x+\varepsilon_1)-F_1(\varepsilon_1)) = \frac{1}{1-F_2(\varepsilon_2)}(F_2(x+\varepsilon_2)-F_2(\varepsilon_2)) \text{ for all } x > 0.$$

This may be seen by putting $p_1 := 1/(1-F_1(\varepsilon_1))$ and $p_2 := 1/(1-F_2(\varepsilon_2))$. Because of

$$\begin{aligned} \tfrac{1}{1-F_i(\varepsilon_i)}(F_i(x+\varepsilon_i)-F_i(\varepsilon_i)) &= \tfrac{P_i(\varepsilon_i<X\leq x+\varepsilon_i)}{P(\varepsilon_i<X)} = \tfrac{P_i(0<X-\varepsilon_i\leq x)}{P(0<X-\varepsilon_i)} \\ &= P_i(X-\varepsilon_i \leq x \mid X-\varepsilon_i > 0),\ i=1,2\ , \end{aligned}$$

$(*)$ is equivalent with

$(**)$ $P_1(X-\varepsilon_1 \leq x \mid X-\varepsilon_1 > 0) = P_2(X-\varepsilon_2 \leq x \mid X.\varepsilon_2 > 0)$.

Example: If X is exponentially distributed, for each pair $\varepsilon_1, \varepsilon_2 > 0$ we have $P(X-\varepsilon_1 \leq x \mid X-\varepsilon_1 > 0) = P(X-\varepsilon_2 \leq x \mid X-\varepsilon_2 > 0)$.
Thus it is impossible to determine ε_1 and ε_2 by observing (S_t) only.

References

[1] FERGUSON, T.S.: Mathematical Statistics. A Decision Theoretic Approach.
Academic Press (1967)

[2] GOVINDARAJULU, Z.: A Survey of Sequential Statistical Analysis.
Sequential Methods of Statistics, Banach Center Publications, Vol. 16,
Polish Scientific Publishers, Warsaw (1985)

[3] KLEIN, A.: Zur Sequentiellen Analysis von Stochastischen Prozessen.

Doctoral Thesis, Marburg (1999)

[4] MAMMITZSCH, V.: Sequential Probability Ratio Tests and Orthogonal Hypotheses.

Prace Naukowe Akademii Ekonomicznej we Wroclawiu Nr. 442 (1988), p. 41-47

[5] MAMMITZSCH, V.: Sequentielle Tests zur Unterscheidung von Gesamtschaden-

verteilungen. Münchner Blätter zur Versicherungsmathematik, Heft 22 (Dezember 1992)

[6] SCHMIDT, K.D.: Lectures on Risk Theory, B.G. Teubner, Stuttgart (1996)

[7] SCHMITZ, N.: Sequential Probability Ratio Tests for Stochastic Processes:

a Review Note. Sequential Methods of Statistics, Banach Center Publications, Vol. 16,

Polish Scientific Publisher, Warsaw (1985)

[8] WALD, A.: Sequential Tests of Statistical Hypotheses. The Annals of Math. Stat.

Vol. 16, (1945), pp 117-186