ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
IX, 1 (1966)

J. ODERFELD (Warszawa)

A CONTRIBUTION TO THE SEQUENTIAL ANALYSIS

§ 1. Introduction. The sequential analysis introduced by A. Wald in 1947 was an improvement of the probability ratio test and since that date further improvements and refinements have been made by Wald himself and by other research-workers. Nowadays a more or less detailed description of the sequential analysis can be found in various textbooks, e.g. [1]. For the purpose of this paper it is sufficient to say that the sample size is not fixed in advance but depends on the observations, and that sequential tests possess an important optimum property as regards the expected sample size.

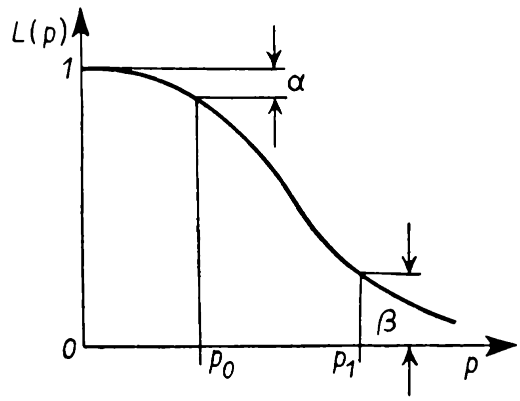

Fig. 1

This paper is restricted to Wald's procedure used for testing simple hypotheses on the value of p in a zero-one distribution where p stands for the unknown fraction of "ones" (sometimes called *successes*).

For the sake of convenience some information will be quoted without proof.

Let H_0: $p = p_0$ and H_1: $p = p_1$ $(0 < p_0 < p_1 < 1)$ denote a pair of simple hypotheses.

Denote by a and β the probabilities of errors of the first and of the second kind respectively. The points $(p_0, 1-a)$ and (p_1, β) lie on the graphical image of the operating-characteristic function $L = L(p)$, where L is the probability of accepting H_0 (Fig. 1).

Write further

$$A = \frac{1-\beta}{a}, \quad B = \frac{\beta}{1-a}, \quad R = p_1/p_0, \quad K = \frac{1-p_1}{1-p_0}$$

and let N be the sample size and T — the number of successes in the sample taken from a population with a fixed fraction p.

The decision rule is the following:

$$(1)\quad \begin{cases} \text{accept } H_0 \text{ if } T \leqslant a_N = \dfrac{\log B}{\log(R/K)} - N\dfrac{\log K}{\log(R/K)}, \\ \text{accept } H_1 \text{ if } T \geqslant r_N = \dfrac{\log A}{\log(R/K)} - N\dfrac{\log K}{\log(R/K)}, \\ \text{examine the item No. } T+1 \text{ if } a_N < T < r_N. \end{cases}$$

The operating-characteristic (O.C.) function is approximately given by

$$(2)\quad p = \frac{1-K^h}{R^h-K^h}, \quad L(p) = \frac{A^h-1}{A^h-B^h}.$$

The average sampling number (A.S.N.) is approximately equal to

$$(3)\quad \overline{N} = \frac{L(p)\log B + [1-L(p)]\log A}{p\log R + (1-p)\log K}.$$

For some special cases exact formulae replacing (2) and (3) are known and in some cases (2) and (3) are exact (cf. [1], pp. 100 and 118). The question of approximation is beyond the scope of this paper and all the formulae derived in the sequel will be strictly equivalent to (1), (2) and (3) respectively.

§ 2. Choosing parameters of a sequential rule. Usually the first step in fixing the rule (1) is the selection of four parameters p_0, p_1, α, β. This is a correct way of acting from the point of view of a mathematician but it means that the whole burden of selection is put on the practician who is invited first to fix his hypotheses, i.e. to select p_0 and p_1, and further to fix the related probabilities of errors α and β. The practician is often advised to take into account losses and to minimize them by a suitable choice of the four mentioned parameters but applying this advice would require some arbitrary probabilistic assumptions, a rather cumbersome computation and — what is the worst — an economic analysis involving data that are usually incomplete.

In practice the α and β values are chosen first: in 99 out of 100 cases some standard values, say $\alpha = 0.05$ and $\beta = 0.10$ are selected. The second step consists in fixing p_0 and p_1 according to some previous experience in similar conditions. Next, the expected sample size for $p = p_0$ and/or $p = p_1$ is checked; if one finds oneself to be not rich enough or able to spend more, one alters the p_1 value.

I am afraid that this picture will be a little disappointing to a theoretical mathematician but nevertheless it is fairly realistic.

What I propose here is to reduce the number of arbitrary parameters and to concentrate on such a set of parameters that has a direct interpretation in the practician's language.

Let us first examine the formula (1). It depends on 4 parameters $p_0, p_1, \alpha_0, \beta_1$ but it has only 3 degrees of freedom joined with 3 constants: $\log B/\log(R/K)$, $\log A/\log(R/K)$ and $\log K/\log(R/K)$. If we agree to add one constraint, e.g. $\alpha = \beta$ (discussion of this condition will be deferred to the end of this paragraph) there remain but 2 degrees of freedom.

Of course this does not mean that any two parameters will define unambiguously the rule (1). For instance, p_0 and p_1 and the condition $\alpha = \beta$ lead to an infinity of rules (1). On the other hand we will show (§ 3) that the rule (1) is determined if one takes as parameters $\bar{p}$ and $\bar{n}$, where $\bar{p}$ is the unique root of the equation $L(p)-1/2 = 0$ and $\bar{n}$ is the expected sample size if $p = \bar{p}$.

The notion of $\bar{p}$ is not new. In statistical quality control $\bar{p}$ is often called "indifference quality" and has been known at least since 1948 (cf. [2]). Speaking more generally, every test on a pair of hypotheses leads eventually to a decision in favour of one of two courses of action; if $p = \bar{p}$, there is no preference for either course. It is much easier to choose a single value for $\bar{p}$ than to select values for p_0 and p_1 which without α_0 and β have no interpretation at all.

At first glance, the notion of $\bar{n}$ seems to be somewhat artificial but happily enough in many situations encountered in practice $\bar{n}$ is very close to the upper bound of the expected sample size and as such it sets the upper bound for expected expenses.

It remains to justify the condition $\alpha = \beta$. The answer could be very short. Any condition involving α or β (e.g. $\alpha = 0.05$ and $\beta = 0.10$) or both means choosing a subclass of all possible sequential rules; the condition $\alpha = \beta$ implies that a_N and r_N (see (1)) differ in the sign of the free term only. This brings considerable simplification, e.g. formula (2) can be put in a distinct form.

But in addition, we can ask how this condition will affect the O.C. function. Let us fix p_0, p_1, α, β. The O.C. function is now defined by (2); let its graphical image be represented by the heavy line in Fig. 2. On this curve take a point whose ordinate is α and denote its abscissa by p_1'.

Now take four quantities $p_0, p_1', \alpha, \alpha$. They define a new O.C. curve (dotted line) which has two(1) common points with the original one: $(p_0, 1-\alpha)$ and (p_1', α). For the ordinate β the original abscissa is p_1 and the new one is p_1''. In most practical situations α and β are small and under this assumption $|p_1''-p_1|$ is much smaller than p_1. Even if $L = 1/2$, the

(1) There are also two points (0, 1) and (1, 0) that are common for all sequential rules regardless of the values of the parameters.

abscissae of both curves — the original one $\overline{p}$ and the new one $\overline{p}''$ — are alike.

For instance, taking for parameters of the test $p_0 = 0.05$, $p_1 = 0.10$, $a = 0.05$, $\beta = 0.10$ we have $p_1' = 0.1099$ and $\overline{p} = 0.0745$; taking for parameters $p_0 = 0.05$, $p_1' = 0.1099$, $a = \beta = 0.05$ we obtain $p_1'' = 0.10077$ and $\overline{p}'' = 0.0765$. Thus the difference $|p_1''-p_1|$ is negligible and the difference $|\overline{p}''-\overline{p}|$ is very small.

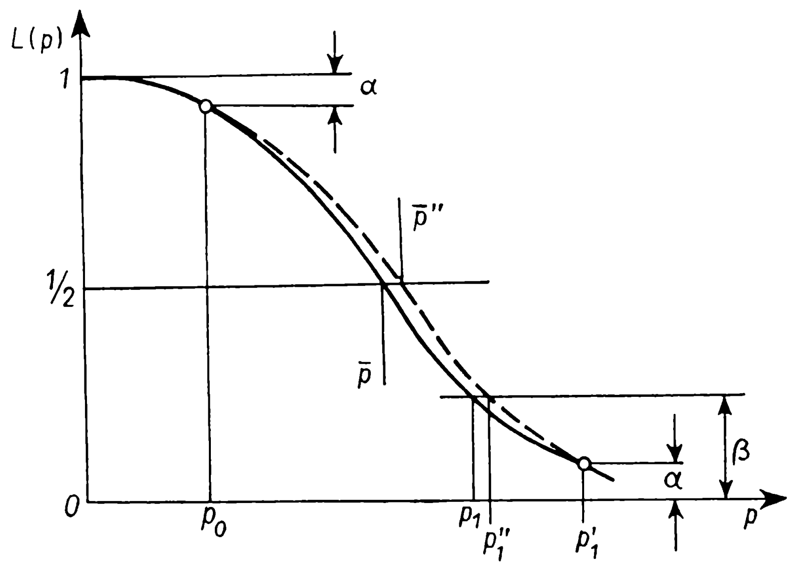

Fig. 2

The computed differences are partly due to the fact that formula (2) is not exact but the essential reason is that the sequential ratio test is not uniformly most powerful [1, p. 101].

§ 3. Introducing parameters $\overline{p}$ and $\overline{n}$. Denote the common value of a and β by γ and put $(1-\gamma)/\gamma = C$. Hence in (1), (2), (3) $A = C$, $B = C^{-1}$ and (2) can be written in the form

$$(4) \qquad p = \frac{1-K^h}{R^h-K^h}, \qquad L = \frac{C^h-1}{C^h-C^{-h}}.$$

If $h = 0$ the expressions (4) are undefined; de l'Hôspital's rule gives the true value $L = 1/2$ and

$$(5) \qquad \overline{p} = \frac{\log K}{\log K - \log R}.$$

Since $L(p)$ is a decreasing function $\overline{p}$ is the unique root of equation $L(p)-1/2 = 0$.

From (4) we obtain dL/dh and dp/dh and eventually

$$\frac{dL}{dp} = \log C \frac{(C^h+C^{-h}-2)(R^h-K^h)^2}{(C^h-C^{-h})^2[(RK)^h(\log R-\log K)+K^h\log K-R^h\log R]}.$$

This expression is undefined for $p = \overline{p}$, $h = 0$. Using de l'Hôspital's rule one finds

$$(6) \qquad -\frac{dL}{dp}\Big|_{\overline{p}} = \tfrac{1}{2}\log C\frac{\log K-\log R}{\log K\log R} = m$$

which together with (5) gives

$$(7) \qquad m = \frac{1}{2\overline{p}}\frac{\log C}{\log R}.$$

From (5) and (7)

$$(8) \qquad \log R = \frac{\log C}{2m\overline{p}}, \quad \log K = -\frac{\log C}{2m(1-\overline{p})}.$$

Substituting (8) to the right sides of a_N and r_N (see (1)), and replacing A by C and B by C^{-1} we have

$$(9) \qquad a_N = -2m\overline{p}(1-\overline{p})+\overline{p}N, \quad r_N = 2m\overline{p}(1-\overline{p})+\overline{p}N.$$

From (3) and (8)

$$(10) \qquad \overline{N} = 2m\overline{p}(1-\overline{p})\frac{1-2L(p)}{p-\overline{p}}.$$

Substitute now in the right side of (10) p by $\overline{p}$, $L(p)$ by $1/2$, and in the left side $\overline{N}$ by $\overline{n}$. Using once more de l'Hôspital's rule we get $\overline{n} = 4m^2\overline{p}(1-\overline{p})$, whence

$$(11) \qquad 2m = \sqrt{\frac{\overline{n}}{\overline{p}(1-\overline{p})}}.$$

From (9) and (11)

$$(12) \qquad a_N = -\sqrt{\overline{p}(1-\overline{p})\overline{n}}+\overline{p}N, \quad r_N = \sqrt{\overline{p}(1-\overline{p})\overline{n}}+\overline{p}N.$$

Now denote two constants

$$(13) \qquad a = \sqrt{\frac{\overline{p}}{(1-\overline{p})\overline{n}}}, \quad b = \sqrt{\frac{1-\overline{p}}{\overline{p}\overline{n}}}$$

and a new variable

$$(14) \qquad x = \frac{L}{1-L}.$$

From (4), (8), (13) and (14)

$$p = \frac{x^a - 1}{x^{a+b} - 1}. \tag{15}$$

From (10), (11) and (14)

$$\overline{N} = \sqrt{\overline{p}(1-\overline{p})\overline{n}}\,\frac{1-x}{1+x}\,\frac{1}{p-\overline{p}}. \tag{16}$$

The formulae (12), (15) and (16) are the sought solution. They depend on two parameters $\overline{p}$ and $\overline{n}$ only.

Now the derivation of the formulae can be forgotten together with the hypotheses and with the test. There remain two courses of action: the first one has to be followed if $T \leqslant a_N$ and the second one — if $T \geqslant r_N$, where a_N and r_N are defined by (12); if neither of these two inequalities holds the sequential procedure has to be continued. Formula (15) gives the probability that eventually the first course of action will be followed and formula (16) the expected sample size.

§ 4. Miscellaneous remarks. Extremal values of the function $\overline{N} = \overline{N}(p)$ defined by (16) can be found from the necessary condition $d\overline{N}/dp = 0$ which leads to the equation

$$\frac{dp}{dx} - \frac{2(\overline{p}-p)}{1-x^2} = 0, \tag{17}$$

where

$$p = \frac{x^a - 1}{x^{a+b} - 1}$$

and

$$\frac{dp}{dx} = \frac{x^{2a-1}}{(x^{a+b}-1)^2}\big((a+b)x^{b-a} - ax^{-a} - bx^b\big).$$

Solving the equation (17) is in general a very timeconsuming operation but this is not necessary in many cases encountered in practice. Namely if $\overline{p}$ is small (e.g. not exceeding a few per cent) and $\overline{n}$ is great (e.g. not less than a few hundred) then the function $\overline{N}(p)$ has exactly one maximum and the ratio $\overline{N}_{\max}/\overline{n}$ is very near to 1 so that $\overline{n}$ can be interpreted as a good approximation of $\overline{N}_{\max}$. Some examples are brought together in Table 1.

TABLE 1

$\overline{n}$	$\overline{p}$	$\overline{N}_{\max}/\overline{n}$
200	0.02	1.0210
400	0.04	1.0047
1000	0.04	1.0018

Selection of parameters $\overline{p}$ and $\overline{n}$ has to be based on non-mathematical reasons (see § 2). However, the discrimination power of the sequential procedure could also be helpful. As a simple measure of that power one may take the ratio $(1/2m)/\bar{p}$ (see Fig. 3). From (11) and (13) it is easily seen that this ratio equals b. It seems that b should be of the order of 1/4.

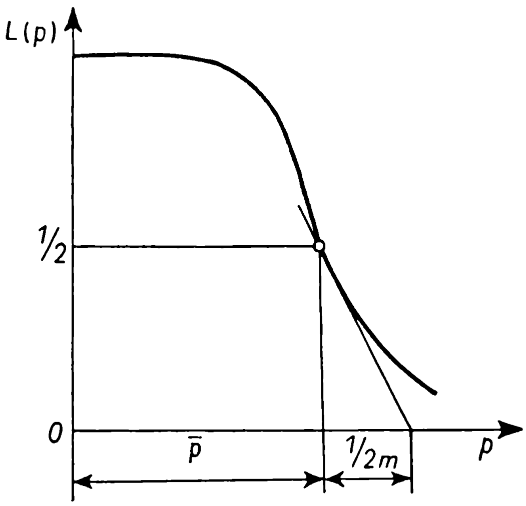

Fig. 3

It is worth mentioning that the smallest sample size permitting to take the first course of action is $[1/a+1]$; no item in the sample should then be a "success". The smallest sample size permitting to take the second course of action is $[1/b+1]$; every item in the sample should then be a "success".

This is easily found from (1), (12) and (13).

§ 5. **An example.** This paper is a by-product of a study on optimization of machines in the non-linear case with a random feasible range. At a stage of this study I had to check the probability p that a random range does not include a certain fixed point and this procedure had to be repeated for a few hundred points. By means of a digital computer the realizations of the random range were generated and the number T of "successes" (the range did not include the point) in N consecutive trials recorded. This procedure was continued until the accumulated record T, N permitted to accept the point in question as a feasible solution or to reject it.

Technical analysis led to $\bar{p} = 0.04$, and consideration of involved costs of running the computer suggested taking $\bar{n} = 400$.

From (13) $a = 0.010208$ and $b = 0.245$ were found and the discrimination power characterized by the value b was considered satisfactory. Further, the formulae (12), (14), (15) and (16) were used to find

$$a_N = -3.919+0.04\,N, \qquad r_N = 3.919+0.04\,N,$$

$$x = \frac{L}{1-L},$$

$$p = \frac{x^{0.010208}-1}{x^{0.255}-1},$$

$$\bar{N} = 3.919\frac{1-x}{1+x}\,\frac{1}{p-0{,}04}.$$

The smallest sample size giving right to recognize the point as a feasible solution was found to be 98 (only successes), and to reject it — 5 (no successes at all). The upper bound of the expected sample size was found to be 402 (see Table 1).

Of course, only the formulae for a_N and r_N had to be put into the computer programme.

References

[1] E. L. Lehmann, *Testing statistical hypotheses*, New York 1959. Wiley & Sons.

[2] H. A. Freeman et al., *Sampling inspection control*, New York 1949, Mc Graw-Hill.

THE MATHEMATICAL INSTITUTE OF THE POLISH ACADEMY OF SCIENCES

Received on 7. 12. 1965

J. ODERFELD (Warszawa)

PRZYCZYNEK DO ANALIZY SEKWENCYJNEJ

STRESZCZENIE

Praca dotyczy analizy sekwencyjnej Walda, stosowanej do weryfikowania hipotez prostych o wartości frakcji p jedynek w rozkładzie zero-jedynkowym.

Zwykle ustala się najpierw hipotezy H_0: $p = p_0$ i H_1: $p = p_1$ oraz prawdopodobieństwa α i β błędów pierwszego i drugiego rodzaju. Ta czwórka liczb determinuje regułę postępowania (1), funkcję operacyjno-charakterystyczną (2) i funkcję oczekiwanej długości badania (3). W pracy przytoczono (1), (2) i (3) w postaci znanych aproksymacji.

Następnie wyjaśniono powody, dla których w praktyce wybór czwórki parametrów p_0, p_1, α, β jest kłopotliwy. Zwrócono uwagę, że (1), (2), (3) mają tylko 3 stopnie swobody, mimo że występują tam 4 parametry. Umotywowano dodatkowy więz $\alpha = \beta$ i wprowadzono 2 nowe parametry $\bar{p}, \bar{n}$, których sens jest dla praktyka wyraźny. Parametr $\bar{p}$ oznacza wartość obojętną frakcji p, przy której funkcja operacyjno-charakterystyczna przyjmuje wartość 1/2. Parametr $\bar{n}$ oznacza oczekiwaną długość badania, gdy $p = \bar{p}$. Pokazano, że w pospolitych warunkach $\bar{n}$ prawie nie różni się od górnego kresu oczekiwanej długości badania.

Wprowadzono bardzo proste wzory (12), (15) i (16), które zawierają tylko 2 parametry $\bar{p}$ i $\bar{n}$ i które całkowicie opisują regułę postępowania i jej konsekwencje matematyczne. W interpretacji nowych wzorów nie ma mowy o testowaniu hipotez; używa się tylko znanych pojęć z teorii decyzji.

Podano przykład zastosowania do optymizacji maszyn w przypadku nieliniowym z obszarem losowym.

Я. ОДЕРФЕЛЬД (Варшава)

К ВОПРОСУ О ПОСЛЕДОВАТЕЛЬНОМ АНАЛИЗЕ

РЕЗЮМЕ

Настоящая работа касается последовательного анализа Вальда, применяемого при проверке простых гипотез о частоте p появления единицы в распределении типа нуль-один.

Обычно сначала устанавливаются гипотезы H_0: $p = p_0$ и H_1: $p = p_1$ а также вероятности α и β погрешностей первого и второго рода. Эта четвёрка чисел, определяет правило поведения (1), оперативную характеристику (2) и функцию ожидаемой продолжительности исследования (3). В работе формулы (1), (2) и (3) даны в виде известных аппроксимаций.

Излагаются причины, по которым применение в практике четвёрки параметров p_0, p_1, α, β затруднительно. Обращается внимание на то, что (1), (2), (3) имеют только 3 степени свободы, несмотря на то, что выступают там 4 параметра. Даётся мотивировка добавочной связи $\alpha = \beta$ и вводятся 2 новых параметра $\bar{p}, \bar{n}$, смысл которых для практика ясен. Параметр $\bar{n}$ обозначает нейтральное значение относительного количества p, при котором оперативная характеристика принимает значение $\frac{1}{2}$. Параметр $\bar{n}$ обозначает продолжительность исследования при $p = \bar{p}$. Показано, что в обычных условиях $\bar{n}$ почти не отличается от верхнего предела ожидаемой продолжительности исследования.

Вводятся очень простые формулы (12), (15), (16), которые содержат только два параметра $\bar{p}$ и $\bar{n}$ и которые полностью описывают правило поведения и его математические последствия. В интерпретации новых формул нет речи о тестировании гипотез; употребляется только известные понятия из теории решений.

Даётся пример приложения к оптимизации машин в нелинейном случае со случайной областью.