Czechoslovak Mathematical Journal

Václav Fabian
A decision function

Czechoslovak Mathematical Journal, Vol. 6 (1956), No. 1, 31–45

Persistent URL: http://dml.cz/dmlcz/100177

Terms of use:

© Institute of Mathematics AS CR, 1956

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.

This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* http://dml.cz

A DECISION FUNCTION

VÁCLAV FABIAN, Praha.

(Received February 10, 1955.)

A family of sequential decision functions, which choose the more probable of two events, is given.

1. Summary

Let

$$X = X_1, X_2, X_3, \ldots$$

be a sequence of independent and identically distributed random variables with

$$P(X_i = 1) = p\,, \quad P(X_i = 0) = 1 - p\,, \tag{1.1}$$

where p is a unknown element of

$$\mathfrak{M} = \langle 0, \tfrac{1}{2}) \cup (\tfrac{1}{2}, 1\rangle\,. \tag{1.2}$$

A decision procedure with two possible decisions

$$d_1 : p < \tfrac{1}{2}\,, \quad d_2 : p > \tfrac{1}{2}\,, \tag{1.3}$$

satisfying following two conditions:

probability of an incorrect decision $\leq \alpha$ for all $p \in \mathfrak{M}$, (1.4)

with probability one (for each $p \in \mathfrak{M}$) the procedure will choose a decision after finite number of observations taken, (1.5)

can be defined as follows:

Let γ be a real number, $\gamma > 1$ and

$$A \geq \frac{1}{\alpha} \sum_{i=1}^{\infty} \frac{1}{i^\gamma}\,, \quad a = \mathrm{L}(A)\,, \tag{1.6}$$

(where $2^{\mathrm{L}(x)} = x$).

Denote by $]z[$ the smallest integer greater or equal to z.

Define the sequence $\{n_i\}$:

$$n_0 = 0\,, \quad n_1 =]a[\,, \tag{1.7}$$

$$n_i = \left]a + \gamma\, \mathrm{L}(i) + \mathrm{L}\left[\binom{n_{i-1}}{i-1}\right]\right[\,, \quad (i \geq 2) \tag{1.8}$$

and the function g by the relation

$$g(z) = i \quad \text{if} \quad n_i \leq z < n_{i+1}\,, \tag{1.9}$$

which is possible, because $n_i < n_{i+1}$ for all i.

Define

$$S_\nu = \sum_{j=1}^{\nu} X_j\,, \quad \nu = 1, 2, \ldots \tag{1.10}$$

and let $\{\nu_i\}$ be an increasing sequence of natural numbers.

Denote by **E** and **F** the following events:

E: There exists an index i, such that $S_{\nu_i} < g(\nu_i)$ and $S_{\nu_j} \leq \nu_j - g(\nu_j)$ for $j = 1, 2, \ldots, i-1$. (1.11)

F: There exists an index i, such that $S_{\nu_i} > \nu_i - g(\nu_i)$ and $S_{\nu_j} \geq g(\nu_j)$ for $j = 1, 2, \ldots, i-1$. (1.12)

The decision function, which accepts d_1 if **E** and d_2 if **F** occurs, will be denoted by $D(\alpha, \gamma, \{\nu_i\})$. We will prove that $D(\alpha, \gamma, \{\nu_i\})$ satisfies (1.4) and (1.5) in sections 2 and 3 respectively.

We point out the possibility of the choice of a sequence $\{\nu_i\}$, i. e., the possibility of grouping observations. Always it is superfluous to use the sequence $\nu_i = i$, for the sequence $\nu_i = n_i$ leads to the same decisions.

In section 4 two tables of the numbers n_i which are less than 502 are given for $\gamma = 1{,}5$, $\alpha = 0{,}01$ and $0{,}1$.

In section 5 the upper bound for the median of the number of observations required by the procedure is given.

In section 6 examples of application are given.

Obviously the decision function $D(\alpha, \gamma, \{\nu_i\})$ is not admissible. It would be possible to reach an improvement (i) using a better approximation than (2.17), (ii) defining $P = P(\mathbf{F})$ in (2.3), (iii) using an other summable sequence in the requirement $1 - P_m \leq \frac{1}{m^\gamma}$ $(m = 1, 2, \ldots)$.

2. Construction of $D(\alpha, \gamma, \{\nu_i\})$

$\{S_\nu\}$ is a sufficient and transitive sequence. Hence, $X_1, X_2, \ldots, X_n$ observed, we can base our decision only on S_n according to [1]. Then among all non-randomized decision functions the only admissible way is to

accept d_1 if $S_n < h(n)$,
accept d_2 if $S_n > k(n)$

continue sampling if $h(n) \leq S_n \leq k(n)$, where $h(n)$ and $k(n)$ are constant for fixed n.

For

$$0 \leq S_n - S_{n-1} \leq 1 \tag{2.1}$$

we can assume that also

$$0 \leq h(n) - h(n-1) \leq 1 \ . \tag{2.2}$$

Defining $h(n) = n - k(n) = g(n)$ we can prove (1.4) and (1.5). On the other hand we are not able to prove this choice of the functions h and k to be optimal.

We will prove (1.4) for our g. Apparently it suffices to prove that, supposing for a moment $p = \frac{1}{2}$,

$$P = P(S_n \geq g(n)\ , \quad n = 1\ , 2, \ldots) \geq 1 - \alpha\ . \tag{2.3}$$

Let us call by random walk (m) any finite sequence $(a_1, a_2, \ldots, a_m)$ with $0 \leq a_{i+1} - a_i \leq 1$ $(i = 1, \ldots, m-1)$, $0 \leq a_1 \leq 1$ and by an admissible random walk (m) any random walk (m), which satisfies

$$a_i \geq g(i) \quad i = 1, 2, \ldots, m\ .$$

Further we denote by V_m (R_m) the number of all different (admissible) random walks (m).

If

$$P_m = P(S_j \geq g(j)\ , \quad j = 1, 2, \ldots, m)\ , \tag{2.4}$$

then

$$P_m = \frac{R_m}{V_m} \tag{2.5}$$

and

$$P = \lim_{m \to \infty} P_m\ . \tag{2.6}$$

It is

$$R_m = 2R_{m-1} \quad \text{if} \quad m \text{ non } \epsilon \{n_i\} \tag{2.7}$$

and

$$R_m = 2R_{m-1} - A_{m-1} \quad \text{if} \quad m \epsilon \{n_i\}\ , \tag{2.8}$$

where $A_m(j)$ is the number of admissible random walks (m) $(a_1, \ldots, a_m)$ with $a_m = j$ and $A_m = A_m(g(m))$.

For

$$V_m = 2^m \tag{2.9}$$

it is

$$P_m = P_{m-1} \quad \text{if} \quad m \text{ non} \epsilon \{n_i\} \tag{2.10}$$

and

$$P_m = P_{m-1} - \frac{A_{m-1}}{2^m} \quad \text{if} \quad m \epsilon \{n_i\}\ . \tag{2.11}$$

Hence

$$P = 1 - \sum_{i=1}^{\infty} \frac{A_{n_{i-1}}}{2^{n_i}} \,. \tag{2.12}$$

The values P_m or A_m can be computed by a direct (and theoretically simple) way, e. g. by use of a modified Pascal triangle. However for large m the computations are extremaly labourious. Thus we must use approximations.

Obviously $j \geq g(m)$ implies

$$A_m(j) = A_{m-1}(j-1) + A_{m-1}(j) \,, \tag{2.13}$$

$$A_m(j) \leq \binom{m}{j}, \tag{2.14}$$

$$A_m(g(m) - 1) = 0 \,. \tag{2.15}$$

Hence

$$A_{n_i-1} = A_{n_{i-1}} \tag{2.16}$$

and from (2.14) we get

$$A_{n_i-1} \leq \binom{n_{i-1}}{i-1}. \tag{2.17}$$

Now, from (1.8) it follows

$$2^{n_i} \geq 2^a i^\gamma \binom{n_{i-1}}{i-1}; \tag{2.18}$$

from (2.17)

$$2^{n_i} \geq 2^a i^\gamma A_{n_i-1} \tag{2.19}$$

and with respect to (1.6)

$$\sum_{i=1}^{\infty} \frac{A_{n_i-1}}{2^{n_i}} \leq \frac{\sum_{i=1}^{\infty} \frac{1}{i^\gamma}}{A} \leq \alpha \tag{2.20}$$

and thus from (2.12)

$$P \geq 1 - \alpha \,. \tag{2.21}$$

From (2.3) it follows, that $P(\mathsf{E}) < 1 - P \leq \alpha$ for $p > \frac{1}{2}$; for $p < \frac{1}{2}$ $P(\mathsf{F}) < 1 - P \leq \alpha$.

Thus we have proved the following theorem:

Theorem 1. *For every decision function $D(\alpha, \gamma, \{\nu_i\})$ the probability of a wrong decision is less than α.*

3. Proof of (1.5)

Throughout this section let $0 < \alpha < 1$ and $1 < \gamma$ be given real numbers.

3.1 Lemma. *There exists a real number $\delta > 1$ such that, putting $\delta_i = i\delta$,*

$$\delta > a \,, \quad \delta_i \geq a + \gamma\, \mathrm{L}(i) + \mathrm{L}\left[\binom{\delta_{i-1}}{i-1}\right] \quad \textit{for} \quad i > 1 \,. \tag{3.2}$$

Proof. $a + \gamma \mathrm{L}(i) + \mathrm{L}\left[\binom{\delta_{i-1}}{i-1}\right] = a + \gamma \mathrm{L}(i) +$

$$+ \sum_{j=1}^{i-1} [\mathrm{L}(\delta_{i-1} - i + 1 + j) - \mathrm{L}(j)] \leq a + \gamma \mathrm{L}(i) +$$

$$+ \frac{1}{\log 2} \int_1^{i-1} [\log(\delta_{i-1} - i + 1 + x) - \log x] \, \mathrm{d}x + \mathrm{L}(\delta_{i-1} - i + 2) =$$

$$= a + \gamma \mathrm{L}(i) + \delta_{i-1} \mathrm{L}(\delta_{i-1}) - (i-1)\mathrm{L}(i-1) - (\delta_{i-1} - i + 2) \, .$$

$$. \mathrm{L}(\delta_{i-1} - i + 2) + \mathrm{L}(\delta_{i-1} - i + 2) = a + \gamma \mathrm{L}(i) + \delta_{i-1}\mathrm{L}(\delta_{i-1}) -$$

$$- (\delta_{i-1} - i + 1)\mathrm{L}(\delta_{i-1}) + (\delta_{i-1} - i + 1)\mathrm{L}(\delta_{i-1}) -$$

$$- (\delta_{i-1} - i + 1)\mathrm{L}(\delta_{i-1} - i + 2) - (i-1)\mathrm{L}(i-1) \leq$$

$$\leq a + \gamma \mathrm{L}(i) + (i-1)\mathrm{L}(\delta_{i-1}) +$$

$$+ (\delta_{i-1} - i + 1)\frac{i-2}{(\delta_{i-1} - i + 1)\log 2} -$$

$$- (i-1)\mathrm{L}(i-1) \leq i[a + \gamma + 1 + \mathrm{L}(\delta)]$$

and it suffices to choose δ in such a way that

$$\delta \geq 1 + a + \gamma + \mathrm{L}(\delta)$$

and (3.2) holds. Thus the lemma is proved.

Unfortunately we have not succeeded to prove the asymptotic behaviour of n_i by the simple manner of the preceding proof. Thus in proving (1.5) we must continue in a way somewhat complicated.

3.3 Lemma. *If*

$$G(k) = k \, \mathrm{L}\left(1 + \frac{1}{k-1}\right) + \mathrm{L}(k-1) \, ,$$

then, for $k < 2$, $\frac{G(k)}{k}$ *is decreasing and*

$$G(k) < k \, . \tag{3.4}$$

Proof.

$$G(2) = 2 \quad \text{and} \quad \left(\frac{G(k)}{k}\right)' = - \frac{\mathrm{L}(k-1)}{k^2} < 0 \, .$$

3.5 Lemma. *Let B be a real number and*

$$F_i(a, k) = a \, \mathrm{L}\left(1 + \frac{1}{k-1}\right) + \mathrm{L}(k-1) + \frac{B}{i} \, . \tag{3.6}$$

Then

$$2 \leq a \leq k_1 < k_2 \Rightarrow F_i(a, k_1) < F_i(a, k_2) \, . \tag{3.7}$$

Further, there exists a sequence $\{\mu_i\}_{i=1}^{\infty}$ *such that* $\mu_i > \mu_{i+1}$, $\mu_i \to 2$ *and*

$$k > \mu_i \Rightarrow F_i(k, k) < k \, . \tag{3.8}$$

Finally, there exists a function $\eta(i, k) > 0$ *such that*

$$k > \mu_i,\ k - \mu(i, k) < a \leq k \Rightarrow F_i(a, k) < a \tag{3.9}$$

and

$$\mu_i < k_1 \leq k_2 \Rightarrow k_1 - \eta(i + 1, k_1) \leq k_2 - \eta_i(i, k_2)\ . \tag{3.10}$$

Proof. The existence of the sequence $\{\mu_i\}$ with the required properties follows at once from the preceding lemma, for

$$F_i(k, k) = G(k) + \frac{B}{i}\ .$$

(3.7) follows from

$$\frac{\partial}{\partial k} F_i(a, k) = \frac{1}{(k-1)\log 2}\left[1 - \frac{a}{k}\right] > 0$$

for $k > a \geq 2$.

The existence of $\eta(i, k)$ satisfying (3.9) follows from (3.8) for $F_i(a, k)$ is continuous in a. For every i, k we define $\eta(i, k)$ as the supremum of the set of numbers satisfying (3.9).

The right side of (3.10) is equivalent to

$$a > k_2 - \eta(i, k_2) \Rightarrow a > k_1 - \eta(i + 1, k_1)$$

and this is equivalent to

$$F_i(a, k_2) < a \Rightarrow F_{i+1}(a, k_1) < a\ .$$

But the last inequality follows from

$$F_i(a, k_2) > F_{i+1}(a, k_1)$$

which follows from (3.7).

3.11 Lemma. *There exists a sequence* $\{\beta_i\}$ *such that* $\beta_1 \geq a + 1$,

$$\beta_i \geq a + 1 + \gamma\, \mathrm{L}(i) + \mathrm{L}\left[\binom{\beta_{i-1}}{i-1}\right], \quad (i > 1) \tag{3.12}$$

and

$$\frac{\beta_i}{i} \to 2\ .$$

Proof. Let δ be the number of lemma 3.1 and $\{\mu_i\}$ the sequence of the preceding lemma, if we put $B = \dfrac{\delta + \gamma}{\log 2}$.

We define two sequences $\{\beta_i\}$ and $\{\gamma_i\}$ as follows:

$$\gamma_i = a + 1 + \gamma\, \mathrm{L}(i) + \mathrm{L}\left[\binom{\beta_{i-1}}{i-1}\right], \tag{3.14}$$

$$\beta_i = i\delta \quad \text{for} \quad i = 1, 2, \ldots, i_1 - 1\ , \tag{3.15}$$

$$\beta_i = \beta_{i-1} + \mathrm{Max}\left(\gamma_i - \beta_{i-1}, \mu_i, \frac{\beta_{i-1}}{i-1} - \eta\left(i, \frac{\beta_{i-1}}{i-1}\right)\right) \qquad \text{for } i \geq i_1\,, \tag{3.16}$$

where i_1 is the first index for which $\mu_i < \delta$.

We shall now prove, that

$$\beta_i - \beta_{i-1} \leq \beta_{i-1} - \beta_{i-2}\,. \tag{3.17}$$

This is obvious for $i = 3, 4, \ldots, i_1 - 1$.

For $i \geq i_1$ (3.17) holds if and only if the following three conditions hold:

$$\gamma_i - \beta_{i-1} \leq \beta_{i-1} - \beta_{i-2}\,, \tag{3.17.1}$$

$$\mu_i \leq \beta_{i-1} - \beta_{i-2}\,, \tag{3.17.2}$$

$$\frac{\beta_{i-1}}{i-1} - \eta\left(i, \frac{\beta_{i-1}}{i-1}\right) \leq \beta_{i-1} - \beta_{i-2}\,. \tag{3.17.3}$$

For $i = i_1$ this reduces to the three inequalities $\gamma_{i_1} - \beta_{i_1-1} \leq \delta$, $\mu_{i_1} \leq \delta$, $\delta - \eta(i_1, \delta) \leq \delta$, which obviously are satisfied.

Let now (3.17) holds for $j = 1, 2, \ldots, i$.

Then

$$\gamma_{i+1} - \gamma_i = \mathrm{L}\left[\binom{\beta_i}{i}\right] - \mathrm{L}\left[\binom{\beta_{i-1}}{i-1}\right] + \gamma[\mathrm{L}(i+1) - \mathrm{L}(i)] =$$

$$= \sum_{\nu=1}^{i} [\mathrm{L}(\beta_i - i + \nu) - \mathrm{L}(\nu)] - \sum_{\nu=1}^{i-1} [\mathrm{L}(\beta_{i-1} - i + 1 + \nu) - \mathrm{L}(\nu)] +$$

$$+ \gamma[\mathrm{L}(i+1) - \mathrm{L}(i)] \leq$$

$$\leq \sum_{\nu=1}^{i} [\mathrm{L}(\beta_i - i + \nu) - \mathrm{L}(\beta_{i-1} - i + \nu)] + L(\beta_{i-1} - i + 1) -$$

$$- \mathrm{L}(i) + \gamma[\mathrm{L}(i+1) - \mathrm{L}(i)] \leq$$

$$\leq (\beta_i - \beta_{i-1}) \sum_{\nu=1}^{i} \mathrm{L}'(\beta_{i-1} - i + \nu) + \mathrm{L}(\beta_{i-1} - i + 1) - \mathrm{L}(i) + \frac{\gamma}{i \log 2}\,.$$

Using the inequality

$$\sum_{\nu=1}^{i} \mathrm{L}'(\beta_{i-1} - i + \nu) \leq \int_1^i \mathrm{L}'(\beta_{i-1} - i + t)\,dt + \mathrm{L}'(\beta_{i-1} - i + 1) \leq$$

$$\leq \mathrm{L}(\beta_{i-1}) - \mathrm{L}(\beta_{i-1} - i + 1) + \frac{1}{(\beta_{i-1} - i + 1)\log 2}$$

we get

$$\gamma_{i+1} - \gamma_i \leq (\beta_i - \beta_{i-1})\,\mathrm{L}\left(\frac{\beta_{i-1}}{\beta_{i-1} - i + 1}\right) + \mathrm{L}\left(\frac{\beta_{i-1} - i + 1}{i-1}\right) + \frac{\beta_i - \beta_{i-1} + \gamma}{i \log 2}$$

and because (according to the inductive assumption) $\beta_i - \beta_{i-1} \leq \beta_2 - \beta_1 \leq \delta$ and for $\beta_i \geq \gamma_i$,

$$\gamma_{i+1} - \beta_i \leq F_i\left(\beta_i - \beta_{i-1}, \frac{\beta_{i-1}}{i-1}\right). \tag{3.18}$$

From the inductive assumption and from (3.16)

$$\frac{\beta_{i-1}}{i-1} - \eta\left(i, \frac{\beta_{i-1}}{i-1}\right) \leq \beta_i - \beta_{i-1} \leq \frac{\beta_{i-1}}{i-1}\,.$$

Thus from (3.9) it follows

$$\gamma_{i+1} - \beta_i \leq \beta_i - \beta_{i-1}\,. \tag{3.19}$$

Further

$$\beta_i - \beta_{i-1} \geq \mu_i > \mu_{i+1} \tag{3.20}$$

and from (3.16)

$$\beta_i - \beta_{i-1} \geq \frac{\beta_{i-1}}{i-1} - \eta\left(i, \frac{\beta_{i-1}}{i-1}\right).$$

But

$$\frac{\beta_{i-1}}{i-1} \geq \frac{\beta_i}{i} \geq \mu_i$$

and from (3.10)

$$\frac{\beta_{i-1}}{i-1} - \eta\left(i, \frac{\beta_{i-1}}{i-1}\right) \geq \frac{\beta_i}{i} - \eta\left(i+1, \frac{\beta_i}{i}\right)$$

and thus

$$\beta_i - \beta_{i-1} \geq \frac{\beta_i}{i} - \eta\left(i+1, \frac{\beta_i}{i}\right). \tag{3.21}$$

From (3.19), (3.20), (3.21) it follows that

$$\beta_{i+1} - \beta_i \leq \beta_i - \beta_{i-1}$$

and thus (3.17) holds for every i.

Apparently $\frac{\beta_i}{i}$ is a non-increasing sequence and thus there exists a limit of $\frac{\beta_i}{i}$, say $\varkappa$.

Suppose that $\varkappa > 2$. Then there exists an index i_2 such that

$$\mu_{i_2} < \varkappa\,.$$

Thus

$$F_{i_2}(\varkappa, \varkappa) < \varkappa\,.$$

From continuity it follows that there exist two positive numbers ε, $\eta < \varkappa - \mu_{i_2}$ such that

$$F_{i_2}(\varkappa + \varepsilon, \varkappa + \varepsilon) < \varkappa - \eta\,. \tag{3.22}$$

Obviously there exists an index i_3 such that $i_3 > i_2$,

$$\beta_{i_3} - \beta_{i_3-1} \leq \frac{\beta_{i_3-1}}{i_3 - 1} \leq \varkappa + \varepsilon\,.$$

Then according to (3.18) and (3.7)

$$\gamma_{i_3+1} - \beta_{i_3} \leq F_{i_3}(\varkappa + \varepsilon, \varkappa + \varepsilon) < \varkappa - \eta \,. \tag{3.23}$$

Obviously

$$\mu_{i_3+1} < \varkappa - \eta \tag{3.24}$$

and

$$\frac{\beta_{i3}}{i_3} - \eta\left(i_3, \frac{\beta_{i3}}{i_3}\right) < \varkappa - \eta \tag{3.25}$$

the latter inequality being a consequence of (3.22) for

$$\varkappa - \eta < a < \frac{\beta_{i3}}{i_3} \text{ implies } \left(\text{for } \frac{\beta_{i3}}{i_3} < \varkappa + \varepsilon\right)$$

$$F_{i3}\left(a, \frac{\beta_{i3}}{i_3}\right) < F_{i2}(\varkappa + \varepsilon, \varkappa + \varepsilon) < \varkappa - \eta < a \,.$$

From (3.23), (3.24) and (3.25) it follows

$$\beta_{i3+1} - \beta_{i3} \leq \varkappa - \eta$$

and

$$\lim \frac{\beta_i}{i} \leq \varkappa - \eta < \lim \frac{\beta_i}{i}\,, \quad \text{which is impossible.}$$

Thus (for $\beta_i - \beta_{i-1} \geq \mu_i > 2$)

$$\lim_{i\to\infty} \frac{\beta_i}{i} = 2 \tag{3.26}$$

and the lemma is proved.

Theorem II. *Let* n_i *be defined by* (1.7) *and* (1.8), *let* $\alpha < \frac{1}{2}$. *Then*

$$\lim_{i\to\infty} \frac{n_i}{i} = 2 \,.$$

Proof. From $\alpha < \frac{1}{2}$ and from theorem I it follows $\inf \frac{n_i}{i} \geq 2$. If n_i are defined by (1.7) and (1.8), then $n_i \leq \beta_i$ for all $i = 1, 2, \ldots$, where β_i are defined by (3.14), (3.15) and (3.16). For $\beta_1 \geq n_1$ and if $\beta_j \geq n_j$ for $j = 1, 2, \ldots, i - 1$, then

$$n_i \leq a + 1 + \gamma\, \mathrm{L}(i) + \mathrm{L}\left[\binom{n_{i-1}}{i-1}\right] \leq$$

$$\leq a + 1 + \gamma\, \mathrm{L}(i) + \mathrm{L}\left[\binom{\beta_{i-1}}{i-1}\right] = \gamma_i \leq \beta_i \,.$$

From $\frac{\beta_i}{i} \to 2$ it follows $\frac{n_i}{i} \to 2$.

Theorem III. *With probability one every test* $D(\alpha, \gamma, \{v_i\})$ *will decide after finite number of observations taken.*

More precisely, if $p \in \mathfrak{M}$, *then*

$$P(g(n) \leq S_n \leq n - g(n) \quad \textit{only for finitely many } n) = 1$$

Proof. The theorem is a direct consequence of the preceding theorem and of the strong law of large numbers.

4. Tables

Two tables are computed, the first contaning the values of n_i for $\gamma = 1{,}5$; $\alpha = 0{,}01$, the second for $\gamma = 1{,}5$; $\alpha = 0{,}1$. The values of $g(n_i) = i$ and $n_i - g(n_i)$ are added. (Tables 1. and 2. at the end of the paper.)

5. The number of observation required

For a given decision function $D(\alpha, \gamma, \{\nu_i\})$ and a given sequence $\{\nu_i\}$ the first member of $\{\nu_i\}$ for which we have not

$$g(\nu_n) \leq S_n \leq \nu_n - g(\nu_n)$$

is a random variable giving the number of observations required. We shall denote it by N.

Denote by M_p the median of N under assumption, that the X_i's are distributed according to $p \in \mathfrak{M}$. We are able to state only the following theorem:

Theorem IV. *Let $m \in \{\nu_n\}$ and let the inequality $g(m) \leq pm \leq m - g(m)$ not hold. Then*

$$M_p \leq m\,.$$

Proof. For $p < \frac{1}{2}$ it is well known that $P(S_m \leq mp) \geq \frac{1}{2}$. Hence, if $mp < g(m)$, then also $P(S_m < g(m)) \geq \frac{1}{2}$. For $p > \frac{1}{2}$ an analogous argument proves the theorem.

In the following table the upper bounds for the two tests $D_1 = D(0{,}01; 1{,}5; \{n_i\})$ and $D_2 = (0{,}1; 1{,}5; \{n_i\}$ are given.

Table 3

(The upper bounds for median of the number of observations required by D_i for various p.)

p	D_1	D_2
0,1 or 0,9	9	5
0,2 or 0,8	28	9
0,3 or 0,7	83	52
0,4 or 0,6	442	312

6. Applications

Remark 1. The decision function $D(\alpha, \gamma, \{\nu_i\})$ can be used for a modified problem in which the set $\mathfrak{N}$ of probability distributions of X has the following property:

If $F \in \mathfrak{N}$ then $X_1, X_2, \ldots$ are according to F mutually independent and X_i takes the values 1 resp. 0 with probability p_i resp. $1 - p_i$. Further there exists a constant $c_F > 0$ such that

$$p_i \leq \tfrac{1}{2} - c_F \quad \text{for all } i = 1, 2, \ldots,$$

or

$$p_i \geq \tfrac{1}{2} + c_F \quad \text{for all } i = 1, 2, \ldots$$

Remark 2. If we know that p non$\in (c_1, c_2)$, where (c_1, c_2) is an interval containing $\frac{1}{2}$, then the usual Wald's sequential test [3] will be obviously better then the present one. This is also the case, if we modify the condition (1.4) in this way: probability of an incorrect decision $\leq \alpha$ for all $p \in \langle 0, c_1 \rangle \cup \langle c_2, 1 \rangle$.

Example 1. Let $X_i = 1$ if the i-th random number in the first column in [2], p. 21, is 1 or 2. Let $X_i = 0$ otherwise. Then we have $S_5 = 0$ and the decision function D_2 gives the decision $p < \frac{1}{2}$ for $g(5) = 1$ according to Table 2.

Example 2. Suppose we will decide by a biological assay, if the toxicity of a drug (say P) is greater or equal to the toxicity of a standard (S), or if it is less (with regard to a given population and to given methods of application of the drugs).

To answer the question let us perform a sequence of experiments, each experiment consisting of drawing two subjects from the given population, applying P to the first and S to the second subject. We shall base our conclusions only on those experiments in which one of the two subjects dies and the other survives. We define $x_i = 1$ if in the i — the such experiment the first subject survies, $X_i = 0$ otherwise.

Thus, for example, if the sequence of observed values of the X_i's is 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, we can decide with the risk $\alpha \leq 0{,}1$ that $P(X_i = 0) > \frac{1}{2}$ i. e., that the toxicity of P is greater than that of S.

BIBLIOGRAPHY

[1] *R. R. Bahadur:* Sufficiency and Statistical Decision Functions. AMS *25*, No 3 (1954).

[2] *Hugo Steinhaus:* Tablica liczb przestasowanych czterocyfrowych. Rozprawy matematyczne, VI, Warszawa 1954, p. 21.

[3] *Abraham Wald:* Sequential Analysis. New York 1952, p. 91.

Table 1

($\gamma = 1.5$. Probability of an incorrect decision < 0.01. Accept the hypothesis $p < \frac{1}{2}$ if for some i $S_{n_i} < i$, accept the hypothesis $p > \frac{1}{2}$ if for some i $S_{n_i} > n_i - i$.)

n_i	i	$n_i - i$	n_i	i	$n_i - i$	n_i	i	$n_i - i$	n_i	i	$n_i - i$
9	1	8	154	53	101	278	105	173	397	157	240
13	2	11	156	54	102	280	106	174	399	158	241
17	3	14	159	55	104	282	107	175	402	159	243
21	4	17	161	56	105	285	108	177	404	160	244
25	5	20	164	57	107	287	109	178	406	161	245
28	6	22	166	58	108	289	110	179	408	162	246
31	7	24	169	59	110	292	111	181	411	163	248
35	8	27	171	60	111	294	112	182	413	164	249
38	9	29	174	61	113	296	113	183	415	165	250
41	10	31	176	62	114	299	114	185	417	166	251
44	11	33	178	63	115	301	115	186	420	167	253
47	12	35	181	64	117	303	116	187	422	168	254
50	13	37	183	65	118	306	117	189	424	169	255
53	14	39	186	66	120	308	118	190	427	170	257
56	15	41	188	67	121	310	119	191	429	171	258
59	16	43	190	68	122	312	120	192	431	172	259
62	17	45	193	69	124	315	121	194	133	173	260
64	18	46	195	70	125	317	122	195	436	174	262
67	19	48	198	71	127	319	123	196	438	175	263
70	20	50	200	72	128	322	124	198	440	176	264
72	21	51	202	73	129	324	125	199	442	177	265
75	22	53	205	74	131	326	126	200	445	178	267
78	23	55	207	75	132	329	127	202	447	179	268
81	24	57	210	76	134	331	128	203	449	180	269
83	25	58	212	77	135	333	129	204	451	181	270
86	26	60	214	78	136	335	130	205	454	182	272
88	27	61	217	79	138	338	131	207	456	183	273
91	28	63	219	80	139	340	133	208	458	184	274
94	29	65	221	81	140	342	133	209	460	185	275
96	30	66	224	82	142	245	134	211	463	186	277
99	31	68	226	83	143	347	135	212	465	187	278
101	32	69	229	84	145	349	136	213	467	188	279
104	33	71	231	85	146	352	137	215	469	189	280
107	34	73	233	86	147	354	138	216	472	190	282
109	35	74	236	87	149	356	139	217	474	191	283
112	36	76	238	88	150	358	140	218	476	192	284
114	37	77	240	89	151	361	141	220	478	193	285
117	38	79	243	90	153	363	142	221	481	194	287
119	39	80	245	91	154	365	143	222	483	195	288
122	40	82	247	92	155	368	144	224	485	196	289
124	41	83	250	93	157	370	145	225	487	197	290
127	42	85	252	94	158	372	146	226	490	198	292
129	43	86	254	95	159	374	147	227	492	199	293
132	44	88	257	96	161	377	148	229	494	200	294
134	45	89	259	97	162	379	149	230	496	201	295
137	46	91	261	98	163	381	150	231	499	202	297
139	47	92	264	99	165	383	151	232	501	203	298
142	48	94	266	100	166	386	152	234	503	204	299
144	49	95	268	101	167	388	153	235	505	205	300
147	50	97	271	102	169	390	154	236	507	206	301
149	51	98	273	103	170	393	155	238			
152	52	100	275	104	171	395	156	239			

Table 2

($\gamma = 1.5$. Probability of an incorrect decision < 0.1. The decision procedure as for Table 1.)

n_i	i	$n_i - i$	n_i	i	$n_i - i$	n_i	i	$n_i - i$	n_i	i	$n_i - i$
5	1	4	148	54	94	271	107	164	391	160	231
9	2	7	150	55	95	273	108	165	393	161	232
13	3	10	152	56	96	276	109	167	396	162	234
17	4	13	155	57	98	278	110	168	398	163	235
20	5	15	157	58	99	280	111	169	400	164	236
23	6	17	159	59	100	283	112	171	402	165	237
26	7	19	162	60	102	285	113	172	404	166	238
29	8	21	164	61	103	287	114	173	407	167	240
32	9	23	167	62	105	289	115	174	409	168	241
35	10	25	169	63	106	292	116	176	411	169	242
38	11	27	171	64	107	294	117	177	413	170	243
41	12	29	174	65	109	296	118	178	416	171	244
44	13	31	176	66	110	299	119	179	418	172	246
47	14	33	178	67	111	301	120	181	420	173	247
50	15	35	181	68	113	303	121	182	422	174	248
52	16	36	183	69	114	305	122	183	424	175	249
55	17	38	185	70	115	308	123	185	427	176	251
58	18	40	188	71	117	310	124	186	429	177	252
60	19	41	190	72	118	312	125	187	431	178	253
63	20	43	193	73	120	315	126	189	433	179	254
66	21	45	195	74	121	317	127	190	436	180	256
68	22	46	197	75	122	319	128	191	438	181	257
71	23	48	200	76	124	321	129	192	440	182	258
73	24	49	202	77	125	324	130	194	442	183	259
76	25	51	204	78	127	326	131	195	445	184	261
79	26	53	206	79	128	328	132	196	447	185	262
81	27	54	209	80	129	330	133	197	449	186	263
84	28	56	211	81	130	333	134	199	451	187	264
86	29	57	214	82	132	335	135	200	453	188	265
89	30	59	216	83	133	337	136	201	456	189	267
91	31	60	218	84	134	339	137	202	458	190	268
94	32	62	221	85	136	342	138	204	460	191	269
96	33	63	223	86	137	344	139	205	462	192	270
99	34	65	225	87	138	346	140	206	465	193	272
101	35	66	227	88	139	348	141	207	467	194	273
104	36	68	230	89	141	351	142	209	469	195	274
106	37	69	232	90	142	353	143	210	471	296	275
109	38	71	234	91	143	355	144	211	473	197	276
111	39	72	237	92	145	357	145	212	476	198	278
114	40	74	239	93	146	360	146	214	478	199	279
116	41	75	241	94	147	362	147	215	480	200	280
119	42	77	244	95	149	364	148	216	482	101	281
121	43	78	246	96	150	366	149	217	485	202	283
123	44	79	248	97	151	369	150	219	487	203	284
126	45	81	251	98	153	371	151	220	489	204	285
128	46	82	253	99	154	373	152	221	491	205	286
131	47	84	255	100	155	375	153	222	493	206	287
133	48	85	257	101	156	378	154	224	496	207	289
136	49	87	260	102	158	380	155	225	498	208	290
138	50	88	262	103	159	382	156	226	500	209	291
140	51	89	264	104	160	384	157	227	502	210	292
143	52	91	267	105	162	387	158	229			
145	53	92	269	106	163	389	159	230			

I am indebted to Eva Šetinová and Carmen Marková for carrying out the computations.

Резюме

ОБ ОДНОЙ РЕШАЮЩЕЙ ФУНКЦИИ

ВАЦЛАВ ФАБИАН (VÁCLAV FABIAN), Прага.

(Поступило в редакцию 10/II 1955 г.)

Пусть

$$X_1, X_2, X_3, \ldots$$

— последовательность независимых и тождественно распределенных случайных переменных,

$$P(X_i = 1) = p\,,$$

$$P(X_i = 0) = 1 - p\,,$$

где p — неизвестный элемент множества

$$\mathfrak{M} = \langle 0, \tfrac{1}{2}) \cup (\tfrac{1}{2}, 1\rangle\,.$$

Построим решающую функцию с двумя возможными решениями

$$d_1\colon \quad p < \tfrac{1}{2}\,,$$

$$d_2\colon \quad p > \tfrac{1}{2}\,.$$

Пусть α — некоторое число интервала (0, 1), а γ — произвольное число, большее единицы. Пусть будет

$$A \geq \frac{1}{\alpha} \sum_{i=1}^{\infty} \frac{1}{i^{\gamma}}\,, \quad a = \mathrm{L}(A)\,,$$

где $2^{\mathrm{L}(X)} = X$.

Обозначим через $]z[$ наименьшее целое число, которое или больше или равно числу z.

Определим последовательность $\{n_i\}$:

$$n_0 = 0\,, \quad n_1 =]a[\,,$$

$$n_i = \left]a + \gamma \mathrm{L}(i) + \mathrm{L}\left[\binom{n_{i-1}}{i-1}\right]\right[$$

и функцию g соотношением

$$g(z) = i \quad \text{для} \quad n_i \leq z < n_{i+1}\,,$$

что вполне возможно, так как $n_i < n_{i+1}$ для всех i.

Далее, определим

$$S_\nu = \sum_{j=1}^{\nu} X_j\,, \quad \nu = 1, 2, \ldots$$

и $\{\nu_i\}$ пусть будет возрастающей последовательностью, каждый член которой является натуральным числом.

Через **E** и **F** обозначим следующие случайные явления:

E: Существует индекс i такой, что $S_{\nu_i} < g(\nu_i)$ и $S_{\nu_j} \leq \nu_j - g(\nu_j)$ для $j = 1, 2, \ldots$

F: Существует индекс i такой, что $S_{\nu_i} > \nu_i - g(\nu_i)$ и $S_{\nu_j} \geq g(\nu_j)$ для $j = 1, 2, \ldots$

Решающая функция $D(\alpha, \gamma, \{\nu_i\})$, которая принимает решение d_1, как только наступает явление **E**, и решение d_2, если наступает явление **F**, обладает следующими свойствами:

1. Вероятность ошибочного решения меньше или равна α для всех $p \in \mathfrak{M}$.

2. С вероятностью единица (для каждого $p \in \mathfrak{M}$,) выбирает она решение после конечного числа наблюдений.

Оценка медианы числа наблюдений, необходимых для достижения решения, дана в 5-том параграфе.

В 4-том параграфе приведены таблицы значений n_i, меньших 502 для $\gamma = 1{,}5$, $\alpha = 0{,}01$ и $\alpha = 0{,}1$.

В 6-том параграфе приводятся примеры.

Решающая функция $D(\alpha, \gamma, \{\nu_i\})$, очевидно, не допустимая. Улучшения можно было бы добиться, во-первых, применением лучшей аппроксимации, чем 2.17, во-вторых, определением $P = P(\mathbf{F})$ в соотношении (2.3) и, наконец, использованием иного ряда в условии $1 - P_m \leq \dfrac{1}{m^\gamma}$ $(m = 1, 2, \ldots)$.