

*Abstract from the Proceedings of the International Mathematical Congress
Amsterdam, Sept. 1954.*

SEQUENTIAL TEST WITH THREE POSSIBLE DECISIONS FOR THE COMPARISON OF TWO UNKNOWN PROBABILITIES

CONSTANCE VAN EEDEN

We consider two series of independent trials, e.g. two processes, each trial resulting in a success or a failure with probabilities p , $1-p$ and p' , $1-p'$ respectively for the two processes.

Executing the trials in pairs consisting of one trial for each process, a sequential test with two possible decisions, developed by Wald [1] (p. 106), may be used for the comparison of p and p' .

For groups of trials of both processes a sequential test with two possible decisions has been described in [2]. This test is carried out as follows:

Suppose the group of trials constituting the i^{th} step of the test consists of n_i trials for the first process and m_i trials for the second one. If the number of successes for the two processes are a_i and b_i respectively, then

$$x_i = 2 \left\{ \arcsin \sqrt{\frac{a_i}{n_i}} - \arcsin \sqrt{\frac{b_i}{m_i}} \right\}$$

is for large n_i and m_i approximately normally distributed with mean

$$(1) \quad \mu = \mathcal{E} x_i = 2 \{ \arcsin \sqrt{p} - \arcsin \sqrt{p'} \} = 2 \arcsin \{ \sqrt{pq'} - \sqrt{p'q} \} \quad \begin{matrix} q' = 1-p' \\ q = 1-p \end{matrix}$$

and variance (2) $\sigma_i^2 = \sigma^2(x_i) = (n_i + 1)n_i^{-2} + (m_i + 1)m_i^{-2}$ (see e.g. [3]).

The ordinary sequential test with two possible decisions for the mean of a normal distribution with known variance, given by Wald [1] (p. 117), may then be applied to the x_i .

Both abovementioned tests for comparing two unknown probabilities can be generalized to tests with three possible decisions. For the first one this generalization may be based directly on a method described by de Boer [4].

The test for groups of trials can be generalized by applying the test with three possible decisions for the mean μ of a normal distribution with known variance, developed by Sobel and Wald [5] to the abovementioned random variables x_i .

For this test two values ξ and η and four values μ_1, μ_2, μ_3 and μ_4 of μ must be chosen, with (3) $\mu_1 < \xi < \mu_2 < \mu_3 < \eta < \mu_4$ the three possible decisions being

$$(4) \quad 1. \quad \mu < \xi, \quad 2. \quad \mu > \eta, \quad 3. \quad \xi \leq \mu \leq \eta,$$

the intervals (μ_1, μ_2) and (μ_3, μ_4) being indifference regions.

To translate this into terms of p and p' , let

$$(5) \quad \sqrt{pq'} - \sqrt{p'q} = \delta$$

then (see (1)) we have $\mu = 2 \arcsin \delta$.

The functional relation between p and p' for a given value of δ^2 is given by those two parts of the ellipse

$$(6) \quad p^2 + p'^2 - 2pp'(1 - 2\delta^2) - 2\delta^2(p + p') + \delta^4 = 0$$

for which $\delta^2 \leq p + p' \leq 2 - \delta^2$; this set of points (p, p') consists of a part $\mathcal{C}$ for which $p > p'$ and a part $\mathcal{C}'$ for which $p < p'$.

Choosing two values ε and ζ and four values $\delta_1, \delta_2, \delta_3$ and δ_4 of δ , with

$$(7) \quad \delta_1 < \varepsilon < \delta_2 < 0 < \delta_3 < \zeta < \delta_4,$$

(δ_1, δ_2) and (δ_3, δ_4) constituting the indifference regions, the three decisions (4) are equivalent with:

$$(8) \quad 1. \quad \delta_1 < \varepsilon, \quad 2. \quad \delta_2 > \zeta, \quad 3. \quad \varepsilon \leq \delta \leq \zeta$$

and hence with the following decisions for p and p' :

- (9) 1. the point (p, p') lies outside the part $\mathcal{C}'$ of the ellipse (6) corresponding to $\delta = \varepsilon$,
 2. the point (p, p') lies outside the part $\mathcal{C}$ corresponding to $\delta = \zeta$,
 3. the point (p, p') lies between or on the parts $\mathcal{C}'$ and $\mathcal{C}$ corresponding to $\delta = \varepsilon$ and $\delta = \zeta$.

The probabilities of errors are determined as described in [5].

REMARK. It is not necessary that p and p' are constant throughout the experiment. We only need a constant δ .

REFERENCES

- [1] A. WALD, Sequential analysis, New York 1944.
- [2] Statistical Research Group of the Columbia University, Sequential analysis of statistical data, applications, section 3, New York 1945.
- [3] R. A. FISHER, Proc. Roy. Soc. Edinburgh **42** (1922), 321—341.
- [4] J. DE BOER, Appl. Sci. Res. **3** (1953), 249—259.
- [5] M. SOBEL and A. WALD, Ann. Math. Stat. **20** (1949), 502—522.

VAN EEGHENSTRAAT 47.
 AMSTERDAM.