The Annals of Mathematical Statistics
1969, Vol. 40, No. 4, 1300–1324

ON PARTITIONING A SET OF NORMAL POPULATIONS BY THEIR LOCATIONS WITH RESPECT TO A CONTROL[1]

BY YUNG LIANG TONG

University of Nebraska

0. The problem and the approaches. This paper is concerned with a problem of partitioning a set of normal populations into two subsets according to their locations with respect to a control population, based on indifference zone formulation. Let $\Pi_0, \Pi_1, \cdots, \Pi_k$ be $(k+1)$ normal populations with means $\mu_0, \mu_1, \cdots, \mu_k$ and a common variance σ^2; and let Π_0 denote the standard or control population. For arbitrary but fixed constants δ_1^* and δ_2^* such that $\delta_1^* < \delta_2^*$, we define three disjoint and exhaustive subsets Ω_B, Ω_I and Ω_G of the set

$$\Omega = (\Pi_1, \Pi_2, \cdots, \Pi_k) \tag{0.1}$$

by

$$\begin{aligned} \Omega_B &= (\Pi_i : \mu_i \leqq \mu_0 + \delta_1^*) \\ \Omega_I &= (\Pi_i : \mu_0 + \delta_1^* < \mu_i < \mu_0 + \delta_2^*) \\ \Omega_G &= (\Pi_i : \mu_i \geqq \mu_0 + \delta_2^*). \end{aligned} \tag{0.2}$$

After observations have been taken, the set Ω is partitioned into two disjoint subsets S_B and S_G.

DEFINITION 0.1. A decision is a correct decision (CD) if $\Omega_B \subset S_B$ and $\Omega_G \subset S_G$.

An equivalent definition to Definition 0.1 is that $S_B \subset (\Omega_B \cup \Omega_I)$ and $S_G \subset (\Omega_G \cup \Omega_I)$. It is noted that the open interval $(\mu_0 + \delta_1^*, \mu_0 + \delta_2^*)$ is considered as the indifference zone and a correct decision puts no restrictions on those populations in the set Ω_I. With this consideration, it will be consistent to give the following

DEFINITION 0.2. A population $\Pi_i \varepsilon \Omega$ is misclassified if $\Pi_i \varepsilon (\Omega_B \cap S_G) \cup (\Omega_G \cap S_B)$.

Let P^* be an arbitrary but preassigned constant such that $2^{-k} < P^* < 1$. The statistical problem is to find a procedure R which consists of a sampling procedure and a terminal decision rule such that the appropriate probability requirement below is satisfied.

(1) When σ^2 is known,

$$P[CD \mid \boldsymbol{\mu}, \sigma^2; R] \geqq P^* \quad \text{for every mean vector } \boldsymbol{\mu}. \tag{0.3}$$

(2) When σ^2 is unknown,

$$P[CD \mid \boldsymbol{\mu}, \sigma^2; R] \geqq P^* \quad \text{for every } \boldsymbol{\mu} \text{ and every } \sigma^2 > 0. \tag{0.4}$$

Received 18 December 1967; revised 3 January 1969.

[1] This paper forms a part of the author's doctoral thesis submitted to the University of Minnesota. The research was partially supported by the National Science Foundation under grant No. NSF-GP-3813.