標準値がある場合の最良正規母集団の選択
– 等分散が仮定できない場合

熊本大・自然科学　高田　佳和

1　序

標準値 μ_0（既知）がある場合の母集団の選択について考察した．Π_i を正規母集団でその母平均を μ_i（未知）, その母分散を σ_i^2（既知，又は未知）, $i=1,\cdots,k(\geq 2)$ とする．$\mu_1,\cdots,\mu_k$ を大きさの順に並べ替えた値を $\mu_{[1]}\leq\cdots\leq\mu_{[k]}$ とし，$\mu_i=\mu_{[k]}$ のとき Π_i を最良母集団とする．標準値 μ_0 がある場合の正しい選択とは, $\mu_{[k]}>\mu_0$ ならば, 最良母集団を選択し, さもなければ, どの母集団も選択しないことである．

Π_i からの大きさ n_i の標本を $X_{i1},\cdots,X_{in_i}$，その標本平均を $\bar{X}_i=\sum_{j=1}^{n_i}X_{ij}/n_i$ とする $(i=1,\cdots,k)$．$\bar{X}_{[k]}$ を k 個の標本平均の最大値とする．選択方法は，$\bar{X}_{[k]}>\mu_0+c$ ならば，$\bar{X}_{[k]}$ を標本平均に持つ母集団を選択し，さもなければ，どの母集団も選択しないこととする．このとき, 選択基準として，Bechhofer (1954) の indifference-zone approach を適用した．与えられた $\delta_0^*,\delta_1^*,\delta_2^*,P_0^*,P_1^*$ $(0<\delta_1^*,\delta_2^*<\infty,-\delta_1^*<\delta_0^*<\infty,2^{-k}<P_0^*<1,(1-2^{-k})/k<P_1^*<1)$ に対して確率要求

$$\begin{aligned} P(\Pi_0) &\geq P_0^* \quad \text{whenever } \mu_{[k]}\leq\mu_0-\delta_0^*, \\ P(\Pi_{[k]}) &\geq P_1^* \quad \text{whenever } \mu_{[k]}\geq\mu_0+\delta_1^* \text{ and } \mu_{[k]}\geq\mu_{[k-1]}+\delta_2^* \end{aligned}$$

を満たすように，各母集団からの標本数と定数 c を決定する. ここで Π_0 は，どの母集団も選択しない事象を, $\Pi_{[k]}$ は, 最良母集団を選択する事象を表す.

Bechhofer and Turnbull (1978) は，母分散 σ_i^2 の値が既知で等しいとき，次の標本数と定数 c が確率要求を満たすことを示した.

$$n_i=\left[\frac{g^2\sigma^2}{\delta_2^{*2}}\right]+1,\quad i=1,\cdots,k,\qquad c=\frac{h}{g}\delta_2^*,$$

ここで，$[x]$ は x を超えない最大の整数を表し，σ^2 は共通の母分散の値である．定数 h，g は次の連立方程式の解である.

$$\begin{aligned} \Phi^k\left(h+\delta_0^*g/\delta_2^*\right) &= P_0^*, \\ \int_{h-\delta_1^*g/\delta_2^*}^{\infty}\Phi^{k-1}(y+g)d\Phi(y) &= P_1^*. \end{aligned}$$

更に，Bechhofer and Turnbull (1978) は，共通の母分散 σ^2 の値が未知の場合，確率要求を満たす二段階選択方法を提案した.

この発表では，母分散 σ_i^2 の値が異なる場合，確率要求を満たす標本数 n_i と定数 c を定める方法を提案し，その性質について考察した.

2　結果

先ず，母分散 σ_i^2 の値は異なるが既知である場合について考察した．この場合上記の母分散が等しい場合の方法から，次の標本数と定数 c が考えられる.

$$n_i=\left[\frac{g^2\sigma_i^2}{\delta_2^{*2}}\right]+1,\quad i=1,\cdots,k,\qquad c=\frac{h}{g}\delta_2^*.$$

このとき，定数 h，g をどのように定めれば確率要求が満たされるかについて考察した．Rinott (1978), Lam (1988) の方法を適用すれば，確率要求を満たす定数 h，g を定めることができることが分かった．定数 h，g を上記の Bechhofer and Turnbull (1978) の連立方程式の解である場合は，確率要求が満たされるかどうかは分からないが，$\delta_0^*=0, \delta_1^*=\delta_2^*$ の場合に，確率要求を満たすための定数 h, g の満たすべき条件が得られた．それによると通常の P_0^*, P_1^* の設定であれば確率要求は満たされることが分かった．Bechhofer and Turnbull (1978) の連立方程式の解 h，g から定まる標本数と定数 c が常に確率要求を満たすためには，標本平均ではなく，Dudewicz and Dalal (1975) による推定量を用いれば可能であることも分かった．次に，３つの方法による標本数の比較を行った．その結果, Bechhofer and Turnbull (1978) の方法が，Rinott (1978)，Lam (1988) の方法よりも標本数が少なくてすむことが分かった．Rinott (1978) と Lam (1988) の方法の比較では，数値的に調べると Lam (1988) の方法が標本数が少ないことが分かった．

次に，母分散 σ_i^2 の値が未知で異なる場合を考察した．確率要求を満たすために，Dudewicz and Dalal (1975), Rinott (1978), Lam (1988) の方法にもとづく二段階選択方法を提案した．更に，それらの標本数の比較を行った．初期標本数を等しくして比較すると，Rinott (1978), Lam (1988) の方法が第一段階で終われば Dudewicz and Dalal (1975) の標本数より少ないが，その差は 1 標本であり，第二段階へ進むならば，その標本数は Dudewicz and Dalal (1975) の標本数より多くなる．このことから，標本数だけの比較では，Dudewicz and Dalal (1975) の方法が優れているように思われる．より詳しく調べるために，$\delta_0^*=0, \delta_1^*=\delta_2^*$ の場合に期待標本数の漸近的比較を行った．その結果 Dudewicz and Dalal (1975) の方法が他の 2 つの方法に比べて，期待標本数が小さいことが分った．又，この問題に対してすでに提案されている Wilcox (1984) の二段階選択法についても検討した．その結果，Rinott (1978) の方法の方が標本数に関して優れていることが分った．

参考文献

Bechhofer, R.E. (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variances. *Ann. Math. Statist.*, **25**, 16–39.

Bechhofer, R.E. and Turnbull, B.W. (1978). Two (k+1)-decision selection procedures for comparing k normal means with a specified standard. *J. Am. Stat. Assoc.*, **73**, 385–392.

Dudewicz, E.L. and Dalal, S.R. (1975). Allocation of observations in ranking and selection with unknown variances. *Sankhā*, **B 37**, 28–78.

Lam, K. (1988). An improved two-stage selection procedure. *Commun. Statist. -Simula.*, **17**, 995–1006.

Rinott, Y. (1978). On two-stage selection procedures and related probability-inequalities. *Commun. Statist. -Theor. Meth.*, **7**, 799-811.

Wilcox, R.R. (1988) Selecting the best population, provided it is better than a standard: the unequal variance case. *J. Am. Stat. Assoc.*, **79**, 887–891.