Ann. Inst. Statist. Math.
33 (1981), Part A, 471-474

ON THE INDIFFERENCE ZONE APPROACH TO SELECTION—A CONSISTENCY RESULT

PAUL D. FEIGIN AND ISHAY WEISSMAN

(Received Nov. 5, 1979; revised Apr. 10, 1981)

Summary

The problem of selecting the t best among k populations is considered. The concept of *ψ-correct selection* is defined and it is shown that the indifference zone approach is consistent in the following sense. The minimum probability of ψ-correct selection over the entire parameter space is no less than the minimum probability of correct selection over the preference zone.

1. Introduction

Consider the situation of k populations $\Pi_1, \cdots, \Pi_k$ and k random observations $X_1, \cdots, X_k$, where each X_i is taken from Π_i with distribution function $F(\cdot\,;\theta_i)$. The problem of selection is to find those t populations with the largest θ-values (to be referred to as the t best populations). Suppose the family $\{F(\cdot\,;\theta)\colon \theta\in\Theta\}$ is stochastically increasing, that is, $F(x;\,\cdot)$ is non-increasing for each fixed x (here Θ is a subset of the real line). A reasonable decision rule R, often used for this situation, is to select those t populations which give rise to the t largest X_i. Let the indices $[i]$ order the θ_i, that is

$$\theta_{[1]}\leqq\theta_{[2]}\leqq\cdots\leqq\theta_{[k]}\ .$$

Then a subset of t populations is a *correct selection* (CS) if it equals $\{\Pi_{[k-t+1]}, \cdots, \Pi_{[k]}\}$.

Let $\Omega=\{\boldsymbol{\theta}=(\theta_1, \cdots, \theta_k)\colon \theta_i\in\Theta\}$ be the parameter space. As $\boldsymbol{\theta}$ varies over Ω, the probability of a correct selection $\mathrm{P}_{\boldsymbol{\theta}}\{CS\,|\,R\}$ may be as low as $1\Big/\binom{k}{t}$. In the indifference zone approach of Bechhofer [2], the experimenter specifies a preference zone $M\subset\Omega$ and designs the experiment (e.g. if each X_i is a sample mean of size n, he may choose n) so that

$$p(M)=\inf_{\boldsymbol{\theta}\in M}\mathrm{P}_{\boldsymbol{\theta}}\{CS\,|\,R\}\geqq P^* \tag{1.1}$$

where $1\Big/\binom{k}{t}<P^*<1$. Following Barr and Rizvi [1], the preference zone is defined in terms of an increasing function ψ, $\psi(x)>x$ for all x, namely,

$$M\equiv\Omega_\psi=\{\boldsymbol{\theta}:\ \psi(\theta_{[k-t]})\leqq\theta_{[k-t+1]}\}\ .$$

(For the classical cases of location or scale parameters $\psi(x)=x+d$ $(d>0)$ or $\psi(x)=cx$ $(c>1)$ respectively.)

The rationale for the indifference zone approach is that we want to control the minimum probability of correctly selecting the t best populations whenever they are really to be preferred (i.e. the t best are sufficiently far apart from the other $k-t$ populations). For this approach to be *consistent* we argue that any selection of t populations Π_i, for which $\psi(\theta_i)>\theta_{[k-t+1]}$, should be considered a correct selection (to be denoted ψ-CS) and that

$$p_\psi(\Omega)=\inf_{\boldsymbol{\theta}\in\Omega}\mathrm{P}_{\boldsymbol{\theta}}\{\psi\text{-}CS\,|\,R\}\geqq p(\Omega_\psi)\ . \tag{1.2}$$

It is our purpose to prove the consistency of the indifference zone approach. Note however that CS always implies ψ-CS and on Ω_ψ both, CS and ψ-CS, are equivalent. Hence

$$p_\psi(\Omega)=\min\{p_\psi(\Omega_\psi^c),\ p(\Omega_\psi)\}\ , \tag{1.3}$$

where $p_\psi(A)=\inf_{\boldsymbol{\theta}\in A}\mathrm{P}_{\boldsymbol{\theta}}\{\psi\text{-}CS\,|\,R\}$. Thus consistency implies that the inequality in (1.2) is in fact equality. The consistency property will follow if we show that

$$\mathrm{P}_{\boldsymbol{\theta}}\{\psi\text{-}CS\,|\,R\}\geqq p(\Omega_\psi)\quad\text{for all}\quad \boldsymbol{\theta}\in\Omega_\psi^c\ . \tag{1.4}$$

This is proved in Section 2. In Section 3 we have some general remarks.

2. The main result

For the proof of (1.4) we need a general result concerning stochastic monotonicity. Let $A\subseteqq\{1,\cdots,k\}$ and $A^c=\{1,\cdots,k\}\backslash A$ and let $N(\boldsymbol{\xi})=\#\{i\in A:\ \xi_i>\bigvee_{j\in A^c}\xi_j\}$, $\boldsymbol{\xi}=(\xi_1,\cdots,\xi_k)\in R^k$.

LEMMA 1. *If $Y=(Y_1,\cdots,Y_k)$ is a random vector with independent components, then $N(Y)$ is stochastically increasing in Y_i, $i\in A$ and stochastically decreasing in Y_j, $j\in A^c$.*

PROOF. We have to show that if $Y^*\in R^k$ is another random vector with independent components such that $Y_i^*\geqq_{\mathrm{st}}Y_i$, $i\in A$ and $Y_j^*\leqq_{\mathrm{st}}Y_j$, $j\in A^c$ then $N(Y)\leqq_{\mathrm{st}}N(Y^*)$. Clearly $N(\boldsymbol{\xi})$ (for non-random $\boldsymbol{\xi}$) is a non-decreasing function of ξ_i, $i \in A$ and non-increasing function of ξ_j, $j \in A^C$. The result follows from the fact that non-decreasing transformations preserve the stochastic ordering.

Now, assuming that the family $\{F(\cdot;\theta): \theta \in \Theta\}$ is stochastically increasing, for correct selection it is well known that

$$\begin{aligned}(2.1)\qquad &\inf_{\boldsymbol{\theta} \in \Omega_\psi} \mathrm{P}_{\boldsymbol{\theta}}\{\bigwedge_{i=k-t+1}^{k} X_{[i]} > \bigvee_{j=1}^{k-t} X_{[j]}\} \\ &= \inf_{\theta \in \Theta} t\int_{-\infty}^{\infty} F^{k-t}(y;\theta)\{1-F(y;\psi(\theta)\}dF(y;\psi(\theta)) .\end{aligned}$$

Let $A_u=\{k-u+1, \cdots, k\}$ $(t \leqq u \leqq k)$, $\boldsymbol{X}=(X_{[1]}, \cdots, X_{[k]})$ and let N_u and A_u play the role of N and A as above. Define the one-parameter subset $B=\{\boldsymbol{\theta}=\boldsymbol{\theta}(\theta)\colon \theta_{[i]}=\theta,\ i \leqq k-t;\ \theta_{[j]}=\psi(\theta),\ j>k-t;\ \theta \in \Theta\} \subset \Omega$. Then (2.1) can be written as

$$(2.2)\qquad p(\Omega_\psi)=\inf_{\boldsymbol{\theta} \in B} \mathrm{P}_{\boldsymbol{\theta}}\{N_t(\boldsymbol{X}) \geqq t\}$$

(note $0 \leqq N_t \leqq t$).

THEOREM 1. *If* $\{F(\cdot;\theta): \theta \in \Theta\}$ *is stochastically increasing then*

$$\mathrm{P}_{\boldsymbol{\theta}}\{\psi\text{-}CS\,|\,R\} \geqq p(\Omega_\psi) \quad \textit{for all} \quad \boldsymbol{\theta} \in \Omega_\psi^C .$$

PROOF. Suppose $\boldsymbol{\theta} \in \Omega_\psi^C$. Then there exist a $\theta_0 \in \Theta$ and an integer u $(t<u \leqq k)$ such that

$$\begin{aligned}(2.3)\qquad &\theta_{[1]} \leqq \cdots \leqq \theta_{[k-u]} \leqq \theta_0 < \theta_{[k-u+1]} \leqq \cdots \leqq \theta_{[k-t]} \leqq \psi(\theta_0) \\ &\psi(\theta_0)=\theta_{[k-t+1]} \leqq \cdots \leqq \theta_{[k]} .\end{aligned}$$

For this $\boldsymbol{\theta}$, a ψ-CS is any subset of size t of the populations $\{\pi_{[k-u+1]}, \cdots, \pi_{[k]}\}$, hence

$$\mathrm{P}_{\boldsymbol{\theta}}\{\psi\text{-}CS\,|\,R\}=\mathrm{P}_{\boldsymbol{\theta}}\{N_u(\boldsymbol{X}) \geqq t\} .$$

It follows by Lemma 1 that for $\boldsymbol{\theta}$ satisfying (2.3) we have

$$\begin{aligned}&\mathrm{P}_{\boldsymbol{\theta}}\{N_u(\boldsymbol{X}) \geqq t\} \geqq \mathrm{P}_{\boldsymbol{\theta}(\theta_0)}\{N_u(\boldsymbol{X}) \geqq t\} \\ &\geqq \mathrm{P}_{\boldsymbol{\theta}(\theta_0)}\{N_t(\boldsymbol{X}) \geqq t\} \geqq \inf_{\boldsymbol{\theta}(\theta) \in B} \mathrm{P}_{\boldsymbol{\theta}(\theta)}\{N_t(\boldsymbol{X}) \geqq t\}=p(\Omega_\psi) ;\end{aligned}$$

the second inequality follows from the fact that $N_t \leqq N_u$ $(t<u)$. This completes the proof.

3. Concluding remarks

Being "ψ-correct" in the *subset selection* approach is called *eliminating the non t-best* by Carroll, Gupta and Huang [3] (see also Gupta and Panchapakesan [5]). If the experimenter is concerned with being "ψ-correct" in classifying the $k-t$ worst populations as well as the t best ones, he is led to the following definition, based on Fabian [4]. A selection of t populations is F-correct (F-CS) if $\psi(\theta_j)>\theta_i$ for every selected Π_j and nonselected Π_i. It is then easy to verify that

$$\begin{aligned} &CS \Rightarrow F\text{-}CS \Rightarrow \psi\text{-}CS && \text{on } \Omega \quad (1\leqq t<k)\,, \\ &CS \Leftrightarrow F\text{-}CS \Leftrightarrow \psi\text{-}CS && \text{on } \Omega_\psi \quad (1\leqq t<k)\,, \\ &F\text{-}CS \Leftrightarrow \psi\text{-}CS && \text{on } \Omega \quad (t=1)\,. \end{aligned}$$

Thus Fabian's result

$$\inf_{\theta\in\Omega} \mathrm{P}_\theta\{F\text{-}CS\,|\,R\}=p(\Omega_\psi) \quad (t=1) \tag{3.1}$$

is a special case of (1.2). It is still an open question whether (3.1) is true for general t.

Acknowledgements

Thanks are due to Tom Santner and a referee for their valuable suggestions which improved the paper.

TECHNION-ISRAEL INSTITUTE OF TECHNOLOGY

REFERENCES

[1] Barr, D. R. and Rizvi, M. H. (1966). An introduction to ranking and selection procedures, *J. Amer. Statist. Ass.*, **61**, 640–646.

[2] 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.

[3] Carroll, R. J., Gupta, S. S. and Huang, D. Y. (1975). Selection procedures for the t best populations, *Commun. Statist.*, **4**, 987–1008.

[4] Fabian, V. (1962). On multiple decision methods for ranking population means, *Ann. Math. Statist.*, **33**, 248–254.

[5] Gupta, S. S. and Panchapakesan (1979). *Multiple Decision Procedures*, John Wiley and Sons, New York.