Journal of the Korean Statistical Society (2003), **32**: *3, pp 299-309*

A SELECTION PROCEDURE FOR GOOD LOGISTICS POPULATIONS

PARMINDER SINGH[1] AND A. N. GILL[2]

ABSTRACT

Let $\pi_1, \ldots, \pi_k$ be $k(\geq 2)$ independent logistic populations such that the cumulative distribution function (*cdf*) of an observation from the population π_i is

$$F_i(x) = \frac{1}{1 + \exp\left\{-\pi\left(x - \mu_i\right)/(\sigma\sqrt{3})\right\}}, \ |x| < \infty,$$

where μ_i $(-\infty < \mu_i < \infty)$ is unknown location mean and σ^2 is known variance, $i = 1, \ldots, k$. Let $\mu_{[k]}$ be the largest of all μ's and the population π_i is defined to be 'good' if $\mu_i \geq \mu_{[k]} - \delta_1$, where $\delta_1 > 0$, $i = 1, \ldots, k$. A selection procedure based on sample median is proposed to select a subset of k logistic populations which includes all the good populations with probability at least P^*(a pre-assigned value). Simultaneous confidence intervals for the differences of location parameters, which can be derived with the help of proposed procedures, are discussed. If a population with location parameter $\mu_i < \mu_{[k]} - \delta_2$ $(\delta_2 > \delta_1)$, $i = 1, \ldots, k$ is considered 'bad', a selection procedure is proposed so that the probability of either selecting a bad population or omitting a good population is at most $1 - P^*$.

AMS 2000 subject classifications. Primary 62F07; Secondary 62F99.
Keywords. Subset selection, sample median, probability of correct selection, simultaneous confidence intervals.

1. INTRODUCTION

Consider $\pi_1, \ldots, \pi_k$ be $k(\geq 2)$ independent logistic populations such that population π_i is characterized by the unknown mean μ_i and common known variance σ^2, $i = 1, \ldots, k$. The cumulative distribution function (*cdf*) associated with

Received January 2003; accepted June 2003.
[1]Department of Mathematics, Guru Nanak Dev University, Amritsar, India (e-mail : pjasmeet@go.com)
[2]Department of Statistics, Panjab University, Chandigarh, India

$\pi_i\,(1 \leq i \leq k)$ is

$$F_i(x) = \frac{1}{1 + \exp\left\{-\pi\,(x - \mu_i)\,/(\sigma\sqrt{3})\right\}},\ |x| < \infty.$$

Without loss of generality, we can take that $\sigma = 1$. Let $\mu_{[1]} \leq \cdots \leq \mu_{[k]}$ be the unknown ordering of μ's. Define the set G of good populations as $G = \left\{i : \mu_i \geq \mu_{[k]} - \delta_1\right\}$ and the set B of the bad populations as $B = \left\{i : \mu_i < \mu_{[k]} - \delta_2\right\}$, where $\delta_2 > \delta_1 > 0$. The interest of the experimenter is to select a subset $S \subseteq (1, \ldots, k)$ which contains all the good populations with probability at least P^* (a pre-specified value), $k^{-1} < P^* < 1$.

Initially, ranking and selection problems have generally been studied by using either the indifference zone approach due to Bechhofer (1954) or subset selection approach of Gupta (1965). In the indifference approach the selected subset size is restricted to one and in the subset selection the selected subset size is random. Later, Desu (1970) restricted to the criteria of badness and defined a set of bad populations as $B_1 = \left\{i : \mu_i < \mu_{[k]} - \delta_2\right\}$. He proposed a subset selection procedure for which the correct selection is achieved if $S \subseteq B_1^c$, where B_1^c is complement of set B_1. Lam (1986) proposed a new procedure for selecting good normal populations in terms of location parameters. For subset selection procedures to select ϵ-best populations in the selected subset, we refer to Laan (1991, 1992). Chen and Dudewicz (1984) proposed subset selection procedures to select the best populations in the selected subset. Gill *et al.* (1993) extended Lam's (1986) approach to the scale parameters. For the first time, Verhulst (1945) used the logistic distribution function as growth function. After this, logistic distribution function finds many applications in the various fields, namely, population growth, bioassay, medical diagnosis, public health, logistic regression, *etc.* For detail applications, we refer Balakrishnan (1992).

The shape of logistic distribution, in many aspects, is similar to the normal distribution, however have a heavier tail than that of normal distribution. Hence, whenever there is suspicion of outliers, the logistic distribution is more suitable approximation to the normal. This motivated us to propose a selection procedures for selecting the good logistic populations. In the past, Lorenzen and McDonald (1981) proposed a selection rule, based on the sample median to select a subset of logistic populations. Han (1979) has studied the rule based on sample means and gives the approximated selection constants. Later, Laan (1989) studied the subset selection rule for the logistic distribution, based on the sample mean, for the case when sample size equal to one.

In Section 2, we discuss the proposed selection procedure and necessary constants are tabulated. Simultaneous confidence intervals for the differences $\mu_{[k]} - \mu_i$, $i = 1, \ldots, k$ and $\mu_{[i]} - \mu_{[j]}$, $i \neq j$, that can be derived with the help of the proposed procedures, are discussed in Section 3. In Section 4, a class of selection procedure is proposed so that the probability of omitting a good population or selecting a bad population is less than or equal to $1 - P^*$.

2. PROPOSED SELECTION PROCEDURE

Our goal is to select a random size subset, say S, of k populations which includes the set G. Thus a correct selection (CS) event will occur if $G \subseteq S$. For a given P^* ($k^{-1} < P^* < 1$) any subset selection procedure R is said to satisfy P^* condition if

$$P_{\underline{\mu}}\{CS|R\} \geq P^*, \quad \forall \underline{\mu} \in \Omega, \tag{2.1}$$

where $\Omega = \left\{\underline{\mu} : \underline{\mu} = (\mu_1, \ldots, \mu_k)', -\infty < \mu_i < \infty, i = 1, \ldots, k\right\}$. Condition (2.1) implies that selected subset must contain the set of good populations with probability at least P^*, pre-specified constant, regardless of the configuration of the parameters. Let $X_{i1}, \ldots, X_{in}$ be a random sample size n $(= 2m+1)$ from i^{th} population π_i and $X_{i:m}$ be the corresponding sample median from the i^{th} population π_i, $i = 1, \ldots, k$. Define the statistic W_m as

$$W_m = \max_{1 \leq i \leq k} (X_{i:m} - \mu_i) - \min_{1 \leq i \leq k} (X_{i:m} - \mu_i). \tag{2.2}$$

Let $F_m(\cdot)$ and $f_m(\cdot)$ denote the *cdf* and *pdf*, respectively of the sample median of a sample of size $n(= 2m + 1)$ from the standard logistic population, *i.e.*, from the logistic population with mean 0 and variance unity. It is easy to see that

$$f_m(x) = \frac{\Gamma(2m)}{\Gamma^2(m)} a\left(e^{-ax}\right)^m \left(1 + e^{-ax}\right)^{-2m}, \ |x| < \infty,$$

and

$$F_m(x) = \frac{\Gamma(2m)}{\Gamma^2(m)} \sum_{j=0}^{m-1} \binom{m-1}{j} (2m - j - 1)^{-1} (-1)^{m-1-j} \left(1 + e^{-ax}\right)^{j+1-2m},$$

where $a = \pi/\sqrt{3}$. The *cdf* of W_m will be

$$G_m(t) = k \int_{-\infty}^{\infty} \left\{F_m(x + t) - F_m(x)\right\}^{k-1} f_m(x) dx.$$

Let $c = c\,(k,\, P^*,\, n)$ be the P^*-quantile of the random variable W_m, *i.e.*, c is the solution of the equation $G_m(c) = P^*$. Now,

$$\begin{aligned} G_m(c) &= k\int_{-\infty}^{\infty}\{F_m(x+c)-F_m(x)\}^{k-1} f_m(x)dx \\ &= \left(\frac{\Gamma(2m)}{\Gamma^2(m)}\right)^k \int_{\infty}^{-\infty}\left\{\sum_{j=0}^{m-1}\binom{m-1}{j}(2m-j-1)^{-1}(-1)^{m-1-j}\right\}^{k-1} \\ &\quad \times\left\{\left(1+e^{-a(x+c)}\right)^{j+1-2m}-\left(1+e^{-ax}\right)^{j+1-2m}\right\}^{k-1} \qquad (2.3)\\ &\quad \times\left\{a\left(e^{-ax}\right)^m\left(1+e^{-ax}\right)^{-2m}\right\}dx. \end{aligned}$$

Putting $y = (1 + e^{-ax})^{-1}$ and letting $c^1 = e^{-ac}$, the equation becomes

$$\begin{aligned} G_m(c) = \left(\frac{\Gamma(2m)}{\Gamma^2(m)}\right)^k k\int_{-\infty}^{\infty}&\left\{\sum_{j=0}^{m-1}\binom{m-1}{j}(2m-j-1)^{-1}(-1)^{m-1-j}\right\}^{k-1} \\ &\times\left\{\left(\frac{y}{(1-c^1)y+c^1}\right)^{-j-1+2m}-y^{2m-j-1}\right\}^{k-1}\left\{(1-y)^{m-1}(y)^{-m-1}\right\}dy. \end{aligned}$$

The integration on the right hand side of the above equation obtained by using numerical integration and $G_m(c)$ equated to P^*. Then values of c, obtained numerically for $k = 2(1)10$ and $m = 1(1)15$, are tabulated in Table 2.1, 2.2 and 2.3 for $P^* = 0.90$, 0.95 and 0.99, respectively.

Let $X_{[1]m} \leq \cdots \leq X_{[k]m}$ be the known ordering of sample median. The proposed selection procedures is

R : Include population π_i in the selected subset S if and only if

$$X_{i:m} \geq X_{[k]m} - \delta_1 - c.$$

That is, choose the set S as

$$S = \left\{i : X_{i:m} \geq X_{[k]m} - \delta_1 - c\right\}.$$

The following theorem shows that proposed selection procedure satisfies the P^*-condition given in (2.1).

THEOREM 2.1. *Let* $G = \left\{i : \mu_i > \mu_{[k]} - \delta_1\right\}$ *and* $c = c\,(k, P^*, n)$ *be the* P^*-*quantile of the distribution of the random variable* W_m. *Then* $P_{\underline{\mu}}\{G \subseteq S\} \geq P^*$ *for* $\mu \in \Omega$.

TABLE 2.1 *Value of c when $P^* = 0.90$*

n	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
1	2.3099	2.9317	3.3180	3.6013	3.8256	4.0115	4.1703	4.3089	4.4320
2	1.4507	1.8269	2.0553	2.2199	2.3488	2.4546	2.5442	2.6220	2.6907
3	1.1367	1.4271	1.6016	1.7265	1.8236	1.9030	1.9701	2.0280	2.0791
4	0.9642	1.2087	1.3547	1.4588	1.5396	1.6054	1.6608	1.7086	1.7507
5	0.8518	1.0667	1.1947	1.2857	1.3561	1.4134	1.4616	1.5031	1.5396
6	0.7711	0.9652	1.0804	1.1622	1.2254	1.2767	1.3199	1.3571	1.3897
7	0.7097	0.8879	0.9936	1.0684	1.1263	1.1732	1.2126	1.2466	1.2763
8	0.6610	0.8266	0.9247	0.9942	1.0478	1.0913	1.1278	1.1592	1.1867
9	0.6210	0.7765	0.8684	0.9335	0.9837	1.0244	1.0585	1.0879	1.1136
10	0.5876	0.7345	0.8213	0.8827	0.9300	0.9684	1.0006	1.0283	1.0525
11	0.5590	0.6986	0.7811	0.8394	0.8843	0.9207	0.9512	0.9775	1.0004
12	0.5342	0.6675	0.7462	0.8019	0.8447	0.8794	0.9085	0.9335	0.9554
13	0.5124	0.6402	0.7157	0.7690	0.8100	0.8432	0.8711	0.8950	0.9160
14	0.4931	0.6161	0.6886	0.7398	0.7792	0.8111	0.8379	0.8609	0.8810
15	0.4758	0.5944	0.6643	0.7137	0.7517	0.7825	0.8082	0.8304	0.8498

PROOF. Define the event $A = W_m \leq c$ and let $\mu_{(i)}$ be the location parameter associated with $X_{[i]m}$, $i = 1, \ldots, k$. Then

$$\begin{aligned} A &= \left\{ \max_{1 \leq i \leq k} (X_{i:m} - \mu_i) - \min_{1 \leq i \leq k} (X_{i:m} - \mu_i) \leq c \right\} \\ &\subseteq \left\{ \min_{1 \leq i \leq k} (X_{i:m} - \mu_i) \geq X_{[k]m} - \mu_{(k)} - c \right\} \\ &\subset \left\{ \min_{1 \leq i \leq k} (X_{i:m} - \mu_i) \geq X_{[k]m} - \mu_{[k]} - c \right\} \\ &= \left\{ X_{i:m} \geq X_{[k]m} - \left(\mu_{[k]} - \mu_i \right) - c,\ i = 1, \ldots, k \right\}. \end{aligned} \tag{2.4}$$

Thus, $\mu_i \geq \mu_{[k]} - \delta_1$ and (2.4) together implies

$$\left\{ X_{i:m} \geq X_{[k]m} - \delta_1 - c, i = 1, \ldots, k \right\} = S.$$

Therefore,

$$P^* = P_{\underline{\mu}}(A) \leq P_{\underline{\mu}}(G \subseteq S), \quad \forall \underline{\mu} \in \Omega.$$

□

In the following section we discuss the simultaneous confidence intervals that can be derived from the proposed selection procedure.

TABLE 2.2 *Value of c when* $P^* = 0.95$

n	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
1	2.8204	3.4400	3.8257	4.1087	4.3328	4.5185	4.6771	4.8157	4.9386
2	1.7509	2.1174	2.3406	2.5019	2.6283	2.7321	2.8203	2.8968	2.9644
3	1.3659	1.6460	1.8147	1.9358	2.0301	2.1073	2.1725	2.2290	2.2787
4	1.1561	1.3906	1.5310	1.6313	1.7092	1.7727	1.8264	1.8727	1.9135
5	1.0200	1.2254	1.3479	1.4353	1.5029	1.5581	1.6045	1.6446	1.6798
6	0.9227	1.1076	1.2177	1.2960	1.3565	1.4059	1.4473	1.4831	1.5145
7	0.8487	1.0182	1.1190	1.1905	1.2458	1.2908	1.3286	1.3612	1.3898
8	0.7900	0.9475	1.0408	1.1071	1.1583	1.1999	1.2349	1.2650	1.2914
9	0.7420	0.8896	0.9771	1.0391	1.0869	1.1258	1.1585	1.1866	1.2113
10	0.7018	0.8412	0.9237	0.9822	1.0273	1.0639	1.0947	1.1212	1.1444
11	0.6675	0.7999	0.8782	0.9337	0.9765	1.0112	1.0403	1.0654	1.0874
12	0.6378	0.7641	0.8388	0.8917	0.9325	0.9656	0.9934	1.0173	1.0382
13	0.6117	0.7328	0.8043	0.8550	0.8940	0.9256	0.9522	0.9751	0.9951
14	0.5886	0.7050	0.7737	0.8224	0.8599	0.8903	0.9158	0.9377	0.9569
15	0.5679	0.6801	0.7464	0.7933	0.8294	0.8587	0.8832	0.9043	0.9229

TABLE 2.3 *Value of c when* $P^* = 0.99$

n	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
1	3.9184	4.5342	4.9193	5.2020	5.4259	5.6110	5.7686	5.9058	6.0271
2	2.3728	2.7220	2.9366	3.0923	3.2147	3.3155	3.4013	3.4758	3.5417
3	1.8324	2.0933	2.2520	2.3665	2.4560	2.5295	2.5918	2.6459	2.6936
4	1.5429	1.7584	1.8887	1.9824	2.0553	2.1151	2.1657	2.2095	2.2481
5	1.3569	1.5442	1.6569	1.7377	1.8006	1.8520	1.8955	1.9330	1.9661
6	1.2248	1.3925	1.4932	1.5652	1.6211	1.6668	1.7054	1.7387	1.7680
7	1.1250	1.2781	1.3698	1.4353	1.4862	1.5277	1.5626	1.5928	1.6194
8	1.0460	1.1878	1.2725	1.3330	1.3799	1.4182	1.4504	1.4782	1.5027
9	0.9817	1.1142	1.1933	1.2498	1.2936	1.3292	1.3593	1.3852	1.4079
10	0.9279	1.0527	1.1273	1.1804	1.2216	1.2551	1.2833	1.3077	1.3290
11	0.8820	1.0005	1.0711	1.1214	1.1604	1.1921	1.2188	1.2418	1.2620
12	0.8423	0.9552	1.0225	1.0704	1.1075	1.1377	1.1631	1.1850	1.2042
13	0.8076	0.9156	0.9800	1.0258	1.0612	1.0901	1.1144	1.1353	1.1536
14	0.7768	0.8806	0.9424	0.9863	1.0203	1.0480	1.0713	1.0913	1.1089
15	0.7493	0.8492	0.9088	0.9511	0.9833	1.0104	1.0328	1.0521	1.0691

3. SIMULTANEOUS CONFIDENCE INTERVALS

Hsu (1981) initiated multiple comparisons with the best using the selection statements. Edward and Hsu (1983), Chen and Vanichbuncha (1989) and Gill *et al.* (1993) derived simultaneous confidence intervals in different setting. Simultaneous inference for the parameters $\mu_{[k]} - \mu_i$, $i = 1, \ldots, k$, referred as multiple comparison with best, can be derived from selection statement without decreasing the nominal confidence level of selection statement. The following theorems provide simultaneous confidence intervals for $\mu_{[k]} - \mu_i$, $i = 1, \ldots, k$ and $\mu_{[i]} - \mu_{[j]}$, $i \neq j$, $j = 1, \ldots, k$.

THEOREM 3.1. *Under the assumptions of Theorem 2.1, we have*

$$P_{\underline{\mu}}\Big\{G \subseteq S,\ \max(0,\ X_{[k]m} - X_{i:m} - c) \leq \mu_{[k]} - \mu_i \leq \max\Big(\max_{j\neq i}(X_{j:m} - X_{i:m}) + c,\ 0\Big),\ i = 1, \ldots, k\Big\} \geq P^*$$

PROOF. Consider the event A as defined in Theorem 2.1 and let $X_{(i)m}$ denote the random variable associated with $\mu_{[i]}$, $i = 1, \ldots, k$. Then by Theorem 2.1, we have

$$\begin{aligned} A &\subset \left\{\mu_{(k)} - \mu_i \geq X_{[k]m} - X_{i:m} - c,\ i = 1, \ldots, k\right\} \\ &\subset \left\{\mu_{[k]} - \mu_i \geq X_{[k]m} - X_{i:m} - c,\ i = 1, \ldots, k\right\} \\ &= \left\{\mu_{[k]} - \mu_i \geq \max\left(X_{[k]m} - X_{i:m} - c, 0\right)\right\}. \end{aligned} \tag{3.1}$$

Moreover,

$$\begin{aligned} A &\subseteq \left\{X_{i:m} - \mu_i \leq X_{(k)m} - \mu_{[k]} + c,\ i = 1, \ldots, k\right\} \\ &\subset \left\{\mu_{[k]} - \mu_i \leq X_{(k)m} - X_{i:m} + c,\ i = 1, \ldots, k\right\} \\ &\subset \Big\{\mu_{[k]} - \mu_i \leq \max_{j\neq i}(X_{j:m} - X_{i:m}) + c,\ i = 1, \ldots, k\Big\} \\ &= \Big\{\mu_{[k]} - \mu_i \leq \max\big(0,\ \max_{j\neq i}(X_{j:m} - X_{i:m}) + c\big),\ i = 1, \ldots, k\Big\}. \end{aligned} \tag{3.2}$$

But in Theorem 2.1 we see that

$$A \subset \{G \subseteq S\}. \tag{3.3}$$

Therefore, from (3.1), (3.2) and (3.3),

$$\begin{aligned} P^* &= P_{\underline{\mu}}(A) \\ &\le P_{\underline{\mu}}\Big\{G \subseteq S, \max(0, X_{[k]m} - X_{i:m} - c) \le \mu_{[k]} - \mu_i \\ &\qquad \le \max\Big(\max_{j\neq i}(X_{j:m} - X_{i:m}) + c,\ 0\Big),\ i = 1, \dots, k\Big\}. \end{aligned}$$

Hence result follows. □

For deriving simultaneous confidence intervals for the difference $\mu_{[i]} - \mu_{[j]}$, $i \neq j$, $j = 1, \dots, k$ of the ranked μ-values we will use some results stated in the following lemmas.

LEMMA 3.2. *For any i between 1 and k, we have*

$$\min_{1\le j\le i}\left(X_{[j]m} - \mu_{(j)}\right) \le X_{[i]m} - \mu_{[i]} \le \max_{i\le j\le k}\left(X_{[j]m} - \mu_{(j)}\right),$$

where $\mu_{(i)}$ is the parameter associated with $X_{[i]m}$, $i = 1, \dots, k$.

PROOF.

$$\begin{aligned} \min_{1\le j\le i}\left(X_{[j]m} - \mu_{(j)}\right) &\le \min_{1\le j\le i}\left(X_{[i]m} - \mu_{(j)}\right) \\ &= \left(X_{[i]m} - \max_{1\le j\le i}\mu_{(j)}\right) \\ &\le \left(X_{[i]m} - \mu_{[i]}\right). \end{aligned} \tag{3.4}$$

Similarly, we can show that

$$\max_{i\le j\le k}\left(X_{[j]m} - \mu_{(j)}\right) \ge X_{[i]m} - \mu_{[j]}. \tag{3.5}$$

Thus the lemma follows from (3.4) and (3.5). □

LEMMA 3.3. *We have the inequalities*

$$\min_{1\le i\le k}\left(X_{[i]m} - \mu_{(i)}\right) \le \min_{1\le i\le k}\left(X_{[i]m} - \mu_{[i]}\right),$$

$$\max_{1\le i\le k}\left(X_{[i]m} - \mu_{[i]}\right) \le \max_{1\le i\le k}\left(X_{[i]m} - \mu_{(i)}\right).$$

The proof of this lemma follows from Lemma 3.1 and hence omitted.

THEOREM 3.4. *Assume that assumptions of Theorem 2.1 hold. Then*

$$P\Big\{X_{[i]m}-X_{[j]m}-c \le \mu_{[i]}-\mu_{[j]} \le X_{[i]m}-X_{[j]m}+ c,\ i \ne j,\ i,\ j = 1,\dots,k\Big\} \ge P^*.$$

PROOF. From the definition of event A in Theorem 2.1 and from Lemma 3.2, we have

$$\begin{aligned} A &= \Big\{\max_{1\le j\le k} (X_{j:m}-\mu_i) - \min_{1\le i\le k} (X_{i:m}-\mu_i) \le c\Big\} \\ &\subseteq \Big\{\max_{1\le j\le k} (X_{[j]m}-\mu_{(j)}) - \min_{1\le i\le k} (X_{[i]m}-\mu_{(i)}) \le c\Big\} \\ &\subseteq \Big\{\max_{1\le j\le k} (X_{[j]m}-\mu_{[j]}) - \min_{1\le i\le k} (X_{[i]m}-\mu_{[i]}) \le c\Big\} \\ &= \Big\{\mu_{[i]}-\mu_{[j]} \le X_{[i]m}-X_{[j]m}+c,\ i\ne j,\ i,\ j=1,\dots,k\Big\}. \end{aligned} \tag{3.6}$$

Also from Lemma 3.2, we have

$$\begin{aligned} A &= \Big\{\min_{1\le j\le k} (X_{[j]m}-\mu_{(j)}) \ge \max_{1\le i\le k} (X_{[i]m}-\mu_{(i)}) - c\Big\} \\ &\subseteq \Big\{\min_{1\le j\le k} (X_{[j]m}-\mu_{[j]}) \ge \max_{1\le i\le k} (X_{[i]m}-\mu_{[i]}) - c\Big\} \\ &= \Big\{\mu_{[i]}-\mu_{[j]} \ge X_{[i]m}-X_{[j]m}-c,\ i\ne j,\ i,\ j=1,\dots,k\Big\}. \end{aligned} \tag{3.7}$$

On combining (3.6) and (3.7), we get

$$A \subseteq \Big\{X_{[i]m}-X_{[j]m}-c \le \mu_{[i]}-\mu_{[j]} \le X_{[i]m}-X_{[j]m}+c,\ i\ne j,\ i,\ j=1,\dots,k\Big\}.$$

Therefore,

$$\begin{aligned} P\Big\{X_{[i]m} &- X_{[j]m}-c \le \mu_{[i]}-\mu_{[j]} \le X_{[i]m}-X_{[j]m}+c,\ i\ne j,\ i,\ j=1,\dots,k\Big\} \\ &\ge P(A) = P^*. \end{aligned}$$

□

4. CONTROLLING BOTH TYPES OF ERRORS

In the selection problems an experimenter can commit two type of errors namely: (i) omitting a 'good' population and (ii) selecting a 'bad' population. Let the population π_i be considered good if $\mu_i \ge \mu_{[k]} - \delta_1$, $\delta_1 > 0$ and bad if $\mu_i < \mu_{[k]} - \delta_2$, where $\delta_2 > \delta_1$. Consider the interest of the experimenter is to control both types of errors. He also wants the selected subset to satisfy $G \subseteq S \subseteq B^c$ with high probability, where B^c is compliment of the set B.

THEOREM 4.1. *Let $\pi_1, \ldots, \pi_k$ be $k(\geq 2)$ be independent logistic populations with unknown location parameters $\mu_1, \ldots, \mu_k$, respectively and with common scale known variance σ^2 (assume $\sigma^2 = 1$). Let n be chosen such that $\delta_2 - \delta_1 > 2c = 2c(k, P^*, n)$ and let t be a constant satisfying $\delta_2 - c > t > \delta_1 + c$. Choose the set $S = \{i : X_{i:m} \geq X_{[k]m} - t\}$, then $P\{G \subseteq S \subseteq B^c\} \geq P^*$.*

PROOF. Let A be defined as in Theorem 2.1 with $\delta = \delta_1$, then

$$A \subset \left\{ \max_{1\leq i\leq k} (X_{i:m} - \mu_i) - \min_{1\leq i\leq k} (X_{i:m} - \mu_i) \leq t - \delta_1 \right\}. \tag{4.1}$$

Following the lines of proof of the Theorem 2.1 it is easy to show that the event in the right hand side of (4.1) is contained in the event $G \subseteq S$. Therefore,

$$A \subseteq (G \subseteq S). \tag{4.2}$$

Also

$$\begin{aligned}
\{S \subseteq B^c\} &= \{B \subseteq S^c\} \\
&= \{X_{i:m} < X_{[k]m} - t \text{ , for all } i\text{'s such that } \mu_i < \mu_{[k]} - \delta_2\} \\
&\supseteq \{X_{(k)m} - X_{i:m} > t \text{ , for all } i\text{'s such that } \mu_i < \mu_{[k]} - \delta_2\} \\
&\supseteq \{X_{(k)m} - \mu_{[k]} - X_{i:m} - \mu_i > t - \delta_2,\ i = 1, \ldots, k\} \qquad (4.3) \\
&\supseteq \{(X_{(k)m} - \mu_{[k]}) - (X_{i:m} - \mu_i) > -c,\ i = 1, \ldots, k\} \\
&= \{(X_{i:m} - \mu_i) - (X_{(k)m} - \mu_{[k]}) < c,\ i = 1, \ldots, k,\ i = 1, \ldots, k\} \\
&\supseteq A.
\end{aligned}$$

From (4.2) and (4.3), we get

$$A \subseteq \{G \subseteq S \subseteq B^c\}.$$

Therefore,

$$P^* = P(A) \leq P_{\underline{\mu}}\{G \subseteq S \subseteq B^c\}, \quad \forall \underline{\mu} \in \Omega.$$

Hence the result follows. □

ACKNOWLEDGEMENTS

The authors are thankful to the referees for their valuable suggestions and improvements.

REFERENCES

BALAKRISHNAN, N. (1992). *Handbook of the Logistic Distribution,* Marcel Dekker, New York.

BECHHOFER, R. E. (1954). "A single-sample multiple decision procedure for ranking means of normal population with known variances", *The Annals of Mathematical Statistics,* **25**, 16–39.

CHEN, H. J. AND DUDEWICZ, E. J. (1984). "A new subset selection theory for all best populations and its applications", In *Developments in Statistics and Its Applications* (A. M. Abouammoh, E. A. El-Neweihi, E.-E. A. Aly and M. A. Al-Osh, eds.), 63–87, King Saud University Press, Riyadh.

CHEN, H. J. AND VANICJBUNCHA, K. (1989). "Simultaneous upper confidence bounds for distances from the best two parameter exponential distribution", *Communications in Statistics-Theory and Methods,* **18**, 3019–3031.

DESU, M. M. (1970). "A selection problem", *The Annals of Mathematical Statistics,* **41**, 1596–1603.

EDWARD, D. G. AND HSU, J. C. (1983). "Multiple comparisons with the best treatment", *Journal of The American Statistical Association,* **78**, 965–971.

GILL, SHARMA AND MISHRA (1993). "Selection of good populations : The scale parameter case", *Journal of South African Statistics,* **27**, 11–21.

GUPTA, S. S. (1965). "On some multiple decision (selection and ranking) and ranking rules", *Technometrics,* **7**, 225–245.

HAN, S. (1979). *Contributions to Selection and Ranking Theory with Special Reference to Logistic Population.* Ph.D. Thesis (also Technical report No. 87–138), Dept of Statistics, Prudue University, West Lafayette, Indiana, Wiley.

HSU, J. C. (1981). "Simultaneous confidence intervals for testing for all distances from the best", *The Annals of Statistics,* **9**, 1026–1034.

LAM, K. (1986). "A new procedure for selecting good populations", *Biometrika,* **73**, 201–206.

LAAN, P. VANDER (1989). "Selection from logistic populations", *Statistica Neerlandica,* **43**, 169–174.

LAAN, P. VANDER (1991). "Subset selection for an ϵ-best population : Efficiency results", COSOR-Memoranda, **19**, Eindhoven University of Technology, Eindhoven.

LAAN, P. VANDER (1992). "Subset selection of an almost best treatment", *Biometrical Journals,* **34**, 647–656.

LORENZEN, T. J. AND MCDONALD, G. C. (1981). "Selecting logistics populations using the sample median", *Communications in Statistics-Theory and Methods,* **10**, 101–124.

VERHULST, P. J. (1945). "Recherches mathématiques sur la loi d'accroissement de la population", *Académie de Bruxelles,* **18**, 1–38.