TEST Engineering & Management

March-April 2020
ISSN: 0193-4120 Page No. 1230 - 1234

On Selecting the Best Population among Increasing Failure Rate Average Populations

[1]Keerthi Astagimath, [2]Shubhashree Joshi, [3]Parameshwar V. Pandit*

[1]Department of Statistics, Karnatak Arts College, Dharwad-580 001, India
[2,3]Department of Statistics, Bangalore University, Bangalore-560 056, India
[1]keerthimathad@gmail.com, [2]shubhashreejoshi13@gmail.com, *[3]panditpv12@gmail.com

Article Info
Volume 83
Page Number: 1230 - 1234
Publication Issue:
March - April 2020

Article History
Article Received: *24 July 2019*
Revised: *12 September 2019*
Accepted: *15 February 2020*
Publication: *14 March 2020*

Abstract

One of the important areas of research in Statistical Inference, namely, the selection and ranking of populations is well attended by many researchers. Here, a selection procedure is developed to select the population possessing stronger increasing failure rate average (IFRA) property among the several IFRA populations. A selection procedure is developed using a measure of departure from exponentiality towards IFRA. The strength of IFRA property of populations is determined by considering the measure given by Keerthi and Pandit [8]. The statistic proposed by Keerthi and Pandit [8] to test exponentiality against IFRA alternatives is based on a U-statistic with a kernel of degree four. Keerthi and Pandit [8] established the large sample properties of the statistic which is the basis for selection procedure. Also the performance of the selection procedure is evaluated using probability of correct selection.

Keywords: *Increasing Failure Rate Average (IFRA) class, Selection and ranking, U-statistic, Probability of correct selection.*

1. Introduction

A population among several populations is selected using the principle of ranking and selection with respect to some criteria. The population may be selected with reference to a measure of location or a measure of scale etc. A number of selection procedures are available in the literature for parametric families of distributions such as selecting the best normal, exponential, Weibull, Gamma etc. The populations are selected as the best based on ranks (Lehmann [9]) while Barlow and Proschan [2] developed selection procedure based on partial orderings of probability distributions. Further, Patel [12] studied such a technique based on means for selecting from class of increasing failure rate (IFR) distributions. For restricted families of distributions, the technique of selection is studied by Barlow and Gupta [1]. Gupta and Panchapakesan [6] gives an exhaustive review of ranking and selection problems. Some selection procedures available in the literature recently include Pandit and Math [11], Inginashetty and Pandit [7]. Pandit and Math [11] studied selection procedure for selecting best IFRA population whereas Inginashetty and Pandit [7] considered the problem of selecting best NBU population. Sharma et.al [13] developed selection procedure to select least IFRA population based on U-statistic. Recently, the problem of selecting the best DMRL population is studied by Pandit and Joshi [10]. Here, the problem of selecting the best increasing failure rate average (IFRA) class of distributions among populations possessing IFRA property is considered. We present below the definition of IFRA.

Definition: Let F be a absolutely distribution function, with $F(0)=0$. Then, F is an increasing failure rate average (IFRA) distribution function if

$\bar{F}(bx) \geq \{\bar{F}(x)\}^b \quad x>0,\ 0<b<1,$

Where $\bar{F} = 1-F$ (1.1)

The equality in (1.1) holds if and only if F is an exponential distribution.

It is easily seen that, if $\gamma(F) \geq \gamma(G)$, then the life distribution F is said to possess more IFRA -ness property than that of life distribution G.

Here, the criterion for selecting the best distribution possessing more IFRA–ness property among k IFRA populations is based on the value of $\gamma(F)$, assuming that the underlying distribution *F* is continuous.

March-April 2020
ISSN: 0193-4120 Page No. 1230 - 1234

The selection procedure here is based on the U-statistics $U_{n,b}(F_i)$ an unbiased estimate of $\gamma(F)$ which is defined as below:
Let

$$h_b(x_1,x_2,x_3,x_4)=\begin{cases}1, & \text{if } \mathrm{Min}(x_1,x_2)>b\mathrm{Max}(x_3,x_4)\\ 0, & otherwise\end{cases}$$

Where b is a fixed number such that $0 \le b \le 1$.

Define $U_{n,b}(F_i)$ as U-statistic based on the kernel $h_b(x_1,x_2,x_3,x_4)$, which is given by

$$U_{n,b}(F_i)=\binom{n}{4}^{-1}\sum h_b^*(x_{i_1},x_{i_2},x_{i_3},x_{i_4})$$

Where summation is taken over all combinations of integers (i_1,i_2,i_3,i_4) chosen out of integers $(1,2,\ldots,n)$ and $h^*{}_b(x_1,x_2,x_3,x_4)$ is the symmetrized version of $h_b(x_1,x_2,x_3,x_4)$.

In this paper, the asymptotic normality of $U_{n,b}(F_i)$ (Refer Keerthi and Pandit [8]) is used for selecting the best IFRA distribution among k IFRA distributions.

2. Selection Procedure

Let $F_1, F_2, \ldots, F_k$ denote distribution functions of k IFRA populations, which are absolutely continuous and let the IFRA -ness measure $\gamma(F_i)$ of F_i be unknown, but it is assumed that F_i are continuous. For the sake of convenience, let $\gamma(F_i)$ be denoted by γ_i. The aim is to select the population possessing more IFRA property. The criterion that is used here is to select the population with largest γ_i, that is $\gamma_{[k]}$, where $\gamma_{[1]} \le \gamma_{[2]} \le \ldots\ldots\ldots. \le \gamma_{[k]}$ denote the ordered IFRA -ness measure for k distributions. Selecting largest $\gamma_{[k]}$ is to select the most IFRA distribution.

Let $\underline{\gamma}=(\gamma_{[1]},\gamma_{[2]},\ldots\ldots\gamma_{[k]})$ be the parameter and

$$\underline{\Omega}=\left\{\underline{\gamma}:\frac{1}{(b+1)(2b+1)}\le\gamma[1]\le\gamma[2]\le\ldots\ldots\ldots\le\gamma[k]\le 1\right\}$$

be the parameter space which is partitioned into a preference zone (PZ), $\Omega(\delta^*)$ and an indifference zone (IZ), $\Omega-\Omega(\delta^*)$, where $\Omega(\delta^*)$ is defined by $\Omega(\delta^*)=\left\{\underline{\gamma}:\gamma[k]-\gamma[k-1]\ge\delta^*\right\}$.

The quantity $\frac{1}{(b+1)(2b+1)}\le\delta^*$ and $\frac{1}{k}<p^*<1$ are pre-assigned by the experimenter and selection procedure R is required to satisfy the condition

P(CS | R) $\ge$ P*, for all $\underline{\gamma}\in\Omega(\delta^*)$ (A)

Selection of any population with $\gamma_{[k]}$ is regarded as the correct selection [CS] and condition (A) is referred to as P* condition.

The selection procedure here is based on the U-Statistics $U_{n,b}(F_i)$ and utilize the large sample properties of $U_{n,b}(F_i)$. The asymptotic distribution of $\sqrt{n}\left(U_{n,b}(F_i)-\gamma(F)\right)$ is normal with mean zero and variance $16\xi_1(F)$ (see Keerthi and Pandit [8]). Since, the asymptotic distributions of each $\sqrt{n}\left(U_{n,b}(F_i)\right)$ is normal with mean $\gamma(F)$ and variance $16\xi_1(F)$, the problem of selecting the more IFRA population can be treated as selection of the largest mean of the normal population (see Dudewicz [4], and, Dudewicz and Dalal [5]).

Here,

$$\xi_1(F)=Cov\left[h_b^*(X_1,X_2,X_3,X_4)h_b^*(X_1,X_5,X_6,X_7)\right]$$

A strongly consistent estimator of $\xi_1(F)$ is given by

$$\hat{\xi}_{1n_0}(F)=\frac{1}{n-1}\sum_{i=1}^{n}(\hat{h}_1(x_i)-T_n(F))^2,$$

TEST Engineering & Management
March-April 2020
ISSN: 0193-4120 Page No. 1230 - 1234

Where

$$\hat{h}_1(x_i) = \frac{6}{(n-1)(n-2)(n-3)} \sum_{j<k} \sum_{k<l} \sum_{i \neq j,k,l} h(x_i, x_j, x_k, x_l).$$

The asymptotic normality of $U_{n,b}(F_i)$,i=1,2,…,k is used to develop two-stage selection procedure R to select the population possessing more IFRA-ness property among k IFRA populations, assuming large samples.

Let $t = t(k,p^*) > 0$ be the unique solution of the equation

$$\int_{-\infty}^{\infty} \Phi^{k-1}(z+t)\, d\Phi(z) = p^*,$$

Where $\Phi(.)$ is the distribution function of standard normal random variable.

Procedure R:

The selection procedure to select the best IFRA distribution involves two stages, that is as explained below:

Take an initial sample of size n_o from the population π_{i1} i=1,2,…k and compute $U_{n_0,b}(F_i)$ and $\hat{\xi}_{1n_0}(F_i)$, define $n_i = \max(2n_0, [1/c])$ where $[x]$ denote the smallest integer which is greater than or equal to x and $\dfrac{1}{c} = 16\hat{\xi}_{1n_0}(F_i)\left(\dfrac{1}{\delta^*}\right)^2$.

The second stage sample size from π_i denote by n_i' is determined as follows:

$$n_i' = \begin{cases} 0 & \text{if } \left[\dfrac{1}{c}\right] \le n_0 \\ n_i - n_0 & \text{if } \left[\dfrac{1}{c}\right] > n_0 \end{cases}.$$

Compute $U_{n_i',b}(F_i)$ based on n_i', the additional sample taken from π_i and define

$$T_i = a_i U_{n_0,b}(F_i) + (1-a_i) U_{n_i',b}(F_i),$$

Where $0 < a_i < 1$ is determined to satisfy

$$16\hat{\xi}_{1n_0}(F_i)\left[\frac{a_i^2}{n_0} + \frac{(1-a_i)^2}{n_i'}\right] = \left(\frac{\delta^*}{t}\right)^2.$$

Here, $16\hat{\xi}_{1n_0}(F_i)\left[\dfrac{a_i^2}{n_0} + \dfrac{(1-a_i)^2}{n_i'}\right]$ is strongly consistent estimator for variance of T_i. Note that no second sample is taken from the i^{th} population, that is $n_i' = 0$, we define $(1-a_i)U_{n_i',b}(F_i) = 0$ and $(1-a_i^2)n_i' = 0$.

Lemma:

There exists a_i satisfying

$$16\hat{\xi}_{1n_0}(F_i)\left[\frac{a_i^2}{n_0} + \frac{(1-a_i)^2}{n_i'}\right] = \left(\frac{\delta^*}{t}\right)^2.$$

Proof:
Define

$$a_i = \begin{cases} 1 & \text{if } \dfrac{1}{c} \le n_0 \\ \dfrac{1}{2}\left(1 + \sqrt{(2n_0 c - 1)}\right) & \text{if } n_0 < \dfrac{1}{c} \le 2n_0 \\ \dfrac{n_0}{n_i} & \text{if } \dfrac{1}{c} > 2n_0 \end{cases}.$$

Then, it is straight forward to show that a_i defined above satisfies

$$16\xi_{1n_0}(F_i)\left[\frac{a_i^2}{n_0} + \frac{(1-a_i)^2}{n_i'}\right] = \left(\frac{\delta^*}{t}\right)^2 \text{ as}$$

$$16\xi_{1n_0}(F_i)\left(\frac{t}{\delta^*}\right)^2 = \frac{1}{c}.$$

It is to be noted that a_i can be chosen as $(1-(\sqrt{2n_0 c - 1}))/2$ and $(1+(\sqrt{2n_0 c - 1}))/2$. In such a case, the initial sample size n_0 is equal to the additional sample size n_i and the coefficients a_i and $(1-a_i)$ become interchangeable. The following theorem gives the procedure R that it attains p*, the pre-fixed probability of correct selection.

Theorem: For any p*, $p^* \in \left(\dfrac{1}{k}, 1\right)$ there exists n_0 large enough such that $\inf\limits_{\Omega(\delta^*)} p(CS \mid R) \cong p^*$.

To select the most IFRA distribution, we select the distribution which yields $T_{[k]}$. In this case the preference zone is defined for fixed δ^* as $\{\gamma : \gamma_{[2]} - \gamma_{[1]} \ge \delta^*\}$, $\delta^* > 0$.

The probability of correct selection is given by

$$P\left[T_{(i)} \ge T_{[1]}\ ;\ i = 2, \ldots\ldots, k\right].$$

This probability of Correct Selection is minimized when $\gamma_{[1]} = \gamma_{[2]} = \ldots\ldots = \gamma_{[k-1]} = \gamma_{[k]}$ and minimum value is the assigned p* that is used to compute t using $\int_{-\infty}^{\infty} \Phi^{k-1}(z+t)d\Phi(z) = p^*$.

3. Simulation Study

Simulation study is conducted to evaluate the performance of the procedure R using linear failure rate (LFR), Makeham and Weibull distributions. The distribution functions of the above three life distributions considered for simulation are given below

1. Linear failure rate (LFR) distribution:

$$F(x) = 1 - \exp\left\{-\left(x + \frac{\theta x^2}{2}\right)\right\}, \qquad x \ge 0, \theta \ge 0.$$

2. Makeham Distribution:

$$F(x) = 1 - \exp\{-(x + \theta(x + \exp\{-x\} - 1))\}, \qquad x \ge 0, \theta \ge 0$$

3. Weibull Distribution:

$$F(x) = 1 - \exp\{-x^{\theta}\}, \qquad x \ge 0, \theta > 0.$$

In table 1, we present the number of populations, values of parameters and initial sample size considered for simulation. A simulation study is conducted to estimate P(CS) and expected sample size. The selection procedure is repeated 1000 times and the proportion of correct selection are computed. The estimate of P(CS) is the proportion of correct selection in 1000 repetitions and the estimate of expected sample size is the average of the total sample sizes needed to reach the decision. The estimates of probability of correct decision P(CS) and expected sample size for LFR, Makeham and Weibull distributions are tabulated in table 2, 3, and 4 respectively. The value of t corresponding to k=3 and k=4 are read from Bechhofer [3] and the values of t are 1.6524 and 1.8932 respectively when P*=0.80.

Table 1: Combination of k, θ and n_0 considered for simulation

Life Distributions	K	Θ	Initial sample size n_0
LFR	3 4	1,2,3, 1,2,3,4	15,20,25,30,40, 50 15,20,25,30,40, 50
Makeham	3 4	1,2,3 1,2,3,4	15,20,25,30,40, 50 15,20,25,30,40, 50
Weibull	3 4	2,3,4 2,3,4,5	15,20,25,30,40, 50 15,20,25,30,40, 50

Table 2: Expected Sample Size Estimates and P(CS) among k LFR distributions with P*=0.8

LFR	θ	n_0=15		n_0=20		n_0=25		n_0=30		n_0=40		n_0=50	
		EN	P(CS)	EN	P(CS)	EN	P(CS)	EN	P(CS)	EN	P(CS)	EN	P(CS)
K=3 δ*=0.0005	1	38	0.86	40	0.92	49	0.92	52	0.94	50	0.95	45	0.95
	2	42		22		46		52		50		45	
	3	28		22		48		53		51		45	
k=4 δ*=0.0005	1	34	0.86	42	0.93	65	0.94	51	0.945	51	0.95	80	0.95
	2	63		74		26		51		49		47	
	3	58		38		51		52		49		47	
	4	41		51		52		51		50		78	

Table 3: Expected Sample Size Estimates and P(CS) among k Makeham distributions with P*=0.8

Makeham	θ	n_0=15		n_0=20		n_0=25		n_0=30		n_0=40		n_0=45	
		EN	P(CS)	EN	P(CS)	EN	P(CS)	EN	P(CS)	EN	P(CS)	EN	P(CS)
K=3 δ*=0.0007	1	39	0.82	36	0.90	45	0.90	49	0.94	49	0.92	47	0.96
	2	46		63		44		49		49		61	
	3	58		38		45		48		50		47	
k=4 δ*=0.0007	1	34	0.84	61	0.91	65	0.91	52	0.95	50	0.94	68	0.96
	2	53		43		36		53		46		60	
	3	54		43		44		52		48		47	
	4	43		35		45		49		49		68	

Table 4: Expected Sample Size Estimates and P(CS) among k Weibull distributions with P*=0.8

Weibull	θ	n_0=15		n_0=20		n_0=25		n_0=30		n_0=40		n_0=45	
		EN	P(CS)	EN	P(CS)	EN	P(CS)	EN	P(CS)	EN	P(CS)	EN	P(CS)
k=3	2	57		81		56		85		81		86	

March-April 2020
ISSN: 0193-4120 Page No. 1230 - 1234

δ*=0.0005	3	29	0.87	63	0.92	55	0.93	60	0.94	81	0.91	82	0.95
	4	50		57		96		61		81		97	
k=4 δ*=0.0005	2	87	0.92	62	0.93	94	0.96	138	0.96	86	0.96	87	0.97
	3	65		85		90		74		86		51	
	4	68		94		78		60		90		103	
	5	52		44		50		70		87		67	

References

[1] Barlow, R.E. and Gupta, S.S, “Selection procedures for restricted families of distributions”, Ann. Math. Statistics. 40, (1969), pp. 905-917.

[2] Barlow, R.E. and F. Proschan, “A note on tests for monotone failure rate”, Ann. Math. Statistics. 40, (1969), pp595-600.

[3] Bechhofer, R.E, “A single sample multiple decision procedure for ranking means of normal populations with known variances”, Ann. Math. Statist. 25, (1954), pp. 16-39.

[4] Dudewicz, E.J, “Non-existence of a single-sample selection procedure whose P(CS) is independent of the variances”, S.Afr.Staist.J., 5, (1971), pp. 37-78.

[5] Dudewicz, E.J. and Dalal, S.R, “Allocation of observations in ranking and selection with unequal variances” Sankhya Ser. B, 37, (!975), pp. 28-78.

[6] Gupta, S.S. and Panchapakesan, S, “Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations” Wiley New York, (1979)

[7] Inginshetty, Sujatha and Pandit, P.V, “Selection of the Best New Better than used Population based on Subsamples” International Journal of Statistika and Mathematika, Vo.08, No.3, (2014), pp. 102-104.

[8] Keerthi Astagimath and Parameshwar V. Pandit, “A New Test For Testing Exponentiality Against Increasin Failure Rate Average Alternatives”, Aryabhatta Journal of Mathematics & Informatics Vol.10, No.2, July-Dec, 2018.

[9] Lehmann, E. L, “A class of selection procedures based on ranks” *Math. Ann.* (1963), pp. 268–275.

[10] Pandit, P. V. and Joshi Shubhashree, “A Selection Procedure to Select Best Population Based on Decreasing Mean Residual Life” , International Journal of Mathematics and Statistics (2017), pp. 106-115.

[11] Pandit, P.V and Math, C.M., “On Selecting best IFRA population: A Large Sample Approach” Indian J. of Math. And Math. Sci., Vol.06, No.2, (2010),pp. 233-239.

[12] Patel, J.K, “Ranking and Selection of IFR populations based on means”, J. Amer. Statist. Assoc., 71, (1976), pp. 143-146.

[13] Sharma, S. K., Kanchan Jain and Neetu Singla, “A Selection Procedure to Select Least Increasing Failure Rate Distribution, Comm. Statist.-Theor. Meth. (2015) pp. 1786-1796.