International Journal of Statistika and Mathematika ISSN: 2277- 2790 E-ISSN: 2249-8605, Volume 8, Issue 3, 2014 pp 102-104

Selection of the Best New Better than used Population Based on Subsamples

Sujatha Inginshetty[1*], Parameshwar V. Pandit[2#]

[1]Department of Statistics, Gulbarga University, Gulbarga – 585106, Karnatka, INDIA.

[2]Department of Statistics, Bangalore University, Banglore – 560056, Karnatka, INDIA.

Corresponding Addresses:

[*]sujatasi_gug@rediffmail.com, [#]panditpv12@gmail.com

Research Article

Abstract: The present study considers the problem of selecting the 'Best' new better than used(NBU) population among the several NBU populations. The procedure to select the 'Best' NBU population is developed based on a measure of departure from exponentiality towards NBU, proposed by Pandit and Math(2009) for the problem of testing exponentiality against NBU alternatives in one sample setting. The selection procedure is based on large sample properties of the statistic proposed in Pandit and Math(2009).We also indicate some applications of the selection procedure.

Keywords and Phrases: New better than used(NBU), Selection and ranking, U-statistic, Probability of correct selection.

1. Introduction

The selection of the 'best population' among the several populations is based on the principle of ranking and selection developed in the literature. One of the criteria of selection is to select a population that belongs to some parametric family with reference to a parameter. Another approach is to select a population that belongs to Lehmann's (1963) procedure based on ranks, Barlow and Proschan's (1969) approach based on partial orderings of probability distributions. Further, a selection procedure based on means for selecting from class of increasing failure rate(IFR) distributions is given by Patel (1976). An extensive review of ranking and selection problems can be found in Gupta and Panchapakesan(1979). Here, we consider selection problem of selection for the NBU classes of distributions.

Pandit and Math (2009) gave $\gamma(F)=\int_0^\infty \bar{F}\left(\frac{x}{m}\right)dF(x)$ as a measure of NBU -ness and used it to develop a test for testing exponentiality against NBU distributions. When F is exponential, $\gamma(F)=\frac{1}{m+1}$ and if F is NBU $\gamma(F)\geq\frac{1}{m+1}$.

It is easy see that, if $\gamma(F)\geq\gamma(G)$, then a life distribution F is said to possess more NBU -ness property than that of another life distribution G,

Here, the criterion for selecting the best distribution possessing NBU –ness property among k NBU populations is based on the value of $\gamma(F)$, assuming that the underlying distribution F is continuous. The selection procedure here is based on the U- statistics $T_{n,}(F)$, an unbiased estimate of $\gamma(F)$ which is defined as below:

Let $X_1,X_2, \ldots ,X_n$ be a random sample from F. Define

$$T_n(F)=\frac{1}{\binom{n}{3}}\sum\sum_{i<j} h(X_i,X_j)$$

where $h(X_i,X_j)$ is the symmetrized version of $I\left(X_i>\frac{X_j}{k}\right)$ and $I(A)=\begin{cases}1 & if\ x\in A\\ 0 & otherwise\end{cases}$.

Here summation is taken over all $\binom{n}{2}$ combinations of integers (i,j) taken out of integers{1,2,3,…,n}. In this paper, the asymptotic normality of $T_n(F)$ (Refer Pandit and Math (2009)) is used for selecting the most NBU distribution among k NBU distributions.

2. Selection procedure

Let $F_1,F_2, \ldots, F_k$ be the distribution functions of k NBU distributions and hence the functional form of F_i and the NBU -ness measure $\gamma(F_i)$ of F_i are 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 goal is to select the population which has largest $\gamma_{[k]}$, where $\gamma_{[1]}\leq\gamma_{[2]}\leq\ldots\ldots\ldots\leq\gamma_{[k]}$ denote the ordered NBU -ness measure for k distributions. Selecting largest $\gamma_{[k]}$ refers to selecting the most NBU distribution.

Let $\underline{\gamma}=\left(\gamma_{[1]},\gamma_{[2]},\ldots..\gamma_{[k]}\right)$ and

$\underline{\Omega}=\left\{\underline{\gamma}:\frac{1}{m+1}\leq\gamma_{[1]}\leq\gamma_{[2]}\leq\ldots\ldots\ldots\leq\gamma_{[k]}\leq 1\right\}$ be the parameter space which is partitioned into a preference zone $\Omega(\delta^*)$

and an indifference zone $\Omega-\Omega(\delta^*)$, where $\Omega(\delta^*)$ is defined by $\Omega(\delta^*)=\left\{\underline{\gamma}:\gamma_{[k]}-\gamma_{[1]}\geq\delta^*\right\}$.

The quantity $\frac{1}{m+1}\leq\delta^*$ and $\frac{1}{k}<p^*<1$ are pre-assigned by the experiment and selection procedure R is required to satisfy the condition

P(CS | R)≥P*,for all

$$\underline{\gamma}\in\Omega(\delta^*) \tag{A}$$

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 $T_n(F)$ and utilize the large sample properties of $T_n(F)$. The asymptotic distribution of $\sqrt{n}(T_n(F)-\gamma(F))$ is normal with mean zero and variance $4\xi_1(F)$ (see Pandit and Math(2009)). Since, the asymptotic distributions of each $\sqrt{n}(T_n(F))$ is normal with mean $\gamma(F)$ and variance $4\xi_1(F_i)$,the problem of selecting the more NBU population can be treated selection of the largest mean of the normal population.

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

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

$$\hat{h}_1(x_i)=\frac{1}{n-1}\sum_j h(x_i,x_j)\cdot$$

Now, the following two stage selection procedure R is proposed, assuming large samples.

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

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

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

Procedure R

The two stages of the selection procedure for the selection of the best NBU distribution is explained below: Take an initial sample of size n_o from the i-th population, i=1,2,…,k and compute $T_{n_0}(F_i)$ 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

$$\frac{1}{c}=4\hat{\xi}_{1n_0}(F_i)\left(\frac{1}{\delta^*}\right)^2.$$

The second stage sample size from π_i denote by n$_i$ is determined as follows:

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

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

$U_i=a_iT_{n_0}(F_i)+(1-a_i)T_{n_i}(F_i)$,

where0< a_i<1 is determined to satisfy

$$4\hat{\xi}_{1n_0}(F_i)\left[\frac{a_i^2}{n_0}+\frac{(1-a_i)^2}{n_i}\right]=\left(\frac{\delta^*}{t}\right)^2.$$

Here, $4\hat{\xi}_{1n_0}(F_i)\left[\frac{a_i^2}{n_0}+\frac{(1-a_i)^2}{n_i}\right]$ is strongly consistent estimator for variance of U_i. Note that no second sample is taken from the ith population, that is $n_i=0$, we define $(1-a_i)T_{n_i}(F)=0$ and $(1-a_i^2)n_i=0$.

Lemma

There exists a_i satisfying

$$4\hat{\xi}_{1n_0}(F_i)\left[\frac{a_i^2}{n_0}+\frac{(1-a_i)^2}{n_i}\right]=\left(\frac{\delta^*}{t}\right)^2.$$

Proof

Define
$$a_i=\begin{cases}1 & \text{if } \frac{1}{c}\leq n_0\\ \frac{1}{2}(1+\sqrt{(2n_0c-1)}) & \text{if } n_0<\frac{1}{c}\leq 2n_0\\ \frac{n_0}{n_1} & \text{if } \frac{1}{c}>2n_0\end{cases}.$$

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

$$4\xi_{1n_0}(F_i)\left[\frac{a_i^2}{n_0}+\frac{(1-a_i)^2}{n_i}\right]=\left(\frac{\delta^*}{t}\right)^2 \text{as}$$

$$4\xi_{1n_0}(F_i)\left(\frac{t}{\delta^*}\right)^2=\frac{1}{c}.$$

It is to be noted that a$_i$ can also be chosen to be $(1-(\sqrt{2n_0c-1}))/2$ and $(1+(\sqrt{2n_0c-1}))/2$. 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.

Theorem: For any p*, $p^*\in\left(\frac{1}{k},1\right)$ there exists n$_0$ large enough such that $\inf_{\Omega(\delta^*)}p(CS\,|\,R)\cong p^*$.

Proof: For any i,we can write

$$P\left(\frac{U_i-\gamma_i}{(\delta^*/t)}\le y\right)=P\left[\frac{U_i-\gamma_i}{\sqrt{\left(4\hat{\xi}_{1n_0}(F_i)\left[\frac{a_i^2}{n_0}+\frac{(1-a_i)^2}{n_i}\right]\right)}}\le y\right]$$

=

$$E_{\hat{\xi}_{1n_0}(F_i)}P\left[\frac{U_i-\gamma_i}{\sqrt{\left(4\hat{\xi}_{1n_0}(F_i)\left[\frac{a_i^2}{n_0}+\frac{(1-a_i)^2}{n_i'}\right]\right)}}\le y\frac{\sqrt{\hat{\xi}_{1n_0}(F_i)}}{\sqrt{\xi_{1n_0}(F_i)}}\ \Big|\hat{\xi}_{1n_0}(F_i)\right]$$

$$\cong E_{\hat{\xi}_{1n_0}(F_i)}\Phi\left[y\frac{\sqrt{\hat{\xi}_{1n_0}(F_i)}}{\sqrt{\xi_{1n_0}(F_i)}}\ \Big|\hat{\xi}_{1n_0}(F_i)\right]\cong\Phi(y) \qquad \text{(B)}$$

as $\left[\frac{\sqrt{\hat{\xi}_{1n_0}(F_i)}}{\sqrt{\xi_{1n_0}(F_i)}}\right]\to 1$.

Let $U_{(i)}$ denote the statistic corresponding to the population having parameter $\gamma_{[i]}$, i=1,2, …, k.

Then, we have

P(CS|R)=P[$U_{(i)}\le U_{(k)}$ *, i=1,2,…,k-1]*

=

$$P\left[\frac{U_{(i)}-\gamma_{[i]}}{\delta^*/t}\le\frac{U_{(k)}-\gamma_{(k)}}{\delta^*/t}+\frac{\gamma_{[k]}-\gamma_{[i]}}{\delta^*/t},\quad i=1,2,3,...,k-1\right]$$

$$\cong\int_{-\infty}^{\infty}\ \pi_{i=1}^{k-1}\ \Phi\left(z+\frac{\gamma_{[k]}-\gamma[i]}{\delta^*/t}\right)\ d\Phi(z) \qquad \text{(B')}$$

$$\ge\int_{0}^{\infty}\ \Phi^{k-1}\left(z+\frac{\delta^*}{\delta^*/t}\right)d\Phi(z) \qquad \text{(C)}$$

$$=\text{p*} \qquad \text{(D)}$$

Since $U_{(i)}$'s are independent.

The inequality (C) is true since the right hand side of (B') is minimized, when $\gamma_{[1]}=\gamma_{[2]}=........=\gamma_{[k-1]}=\gamma_{[k]}-\delta^*$ and (D) follows from

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

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 $\{\gamma:\gamma_{[k]}-\gamma_{[k-1]}\ge\delta^*\}$, $\quad \delta^*>\frac{1}{m+1}$.

The probability of correct selection is given by $P[U_{(k)}\ge U_{[i]}\ ;\ i=1,2,...,k-1]$.

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

Applications

A significant application of the procedure R is that, it can be used to select the "best" distribution according to its NBU -ness property to a number of interesting situations, for which no other selection procedures are available. It can be used to select among k given life distributions, with the same functional form but different shape parameters, the distribution with the largest mean life or the distribution with the largest mean residual life at a fixed time t_o> 0. The procedure can be used to select the best of k NBU distributions all of which have the same mean.

References

1. Barlow, R.E. and Gupta, S.S.(1969). Selection procedures for restricted families of distributions.Ann. Math. Statisti. 40, 905-917.
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. Dudewicz, E.J. (1971). Non-existence of a single-sample selection procedure whose P(CS) is independent of the variances, S.Afr.Staist.J., 5, 37-78.
4. Dudewicz, E.J. and Dalal, S.R.(!975). Allocation of observations in ranking and selection with unequal variances.Sankhya Ser. B, 37, 28-78.
5. Gupta, S.S. and Panchapakesan, S.(1979). Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations. Wiley New York.
6. Pandit, P.V and Math, C.M.. (2009). On testing exponentiality against new better than used alternatives. IAPQR Transactions, Vo.34, No.1, 1-17.
7. Patel, J.K. (1976). Ranking and Selection of IFR populations based on means, J. Amer. Statist. Assoc., 71, 143-146.
8. Park, D.H. and LalSexena, K.M. (1987). Nonparametric procedure for selecting the least NBU population: Large sample approach, Comm.Statist.-Theory Meth., 16(11), 3359-3377.
9. Tiwari,R.C. and Zalkikar, J.N.(1988). Testing whether F is 'more IFRA than is G', Microelectronics Reliability, Vol.28, No.5, 703-712.