International Journal of Statistical Sciences *ISSN 1683–5603*
Vol. 6 (Special Issue), 2007, pp 163-179

Testing the Center Using Median

Mezbahur Rahman and Sheikh Rumanur Rahman
Minnesota State University
Mankato, MN 56001, USA
Email: mezbahur.rahman@mnsu.edu and sheikh.rahman@gmail.com

[Received October 7, 2005; Revised March 20, 2007; Accepted July 7, 2007]

Abstract

In testing for the center of a distribution, the use of median is explored. A new test using the distribution of the median is proposed. The method is implemented for the normal (Gaussian) population. Power of the test is compared with standard parametric and nonparametric tests using the mean and the median as the center parameter.

Keywords and Phrases: Binomial probability; Sign test; T test; Wilcoxon signed rank test; Z test

AMS Classification: 62F03.

1 Introduction

The arithmetic mean has been used for such a long time and for many other reasons became most favorite statistical parameter to the statisticians. The most commonly identified rivals to the mean are the mode and the median. In this paper, we will examine the tests utilizing the mean and also will develop a test utilizing the median in testing for the center parameter. The mode does not possess essential qualities to be considered as a good candidate as the distributions of the mode are intractable. There can be more than one mode in one sample and the determination of the mode becomes uncertain when the sample size becomes smaller. The median can be determined just uniquely as the mean. The mean has been preferred over the median for the following reasons:

- "The median is not affected by the magnitude of the values so long as a change in the magnitude of any value does not alter its position with regard to the median, whereas the mean depends upon the individual values ", Savur (1937).

- "The mean is consistent and sufficient and is the most efficient of all statistics", Savur (1937).

The main reason mean is considered favorable against the median is that the mean takes into account the individual values in the sample, it is more representative of the samples and gives more information from the sample than the median. On the other hand we would like to explore the opportunity that in some instances the median can provide more powerful base which can lead us to more precise decision in testing significance involving a sample. We will discuss most commonly used parametric tests in statistics.

2 Parametric Tests Using the Mean

Let us consider that the sample is from a population with the center parameter M, then the possible hypotheses are:

$$\begin{array}{lll} TestA. & H_0 : M = M_0, & H_1 : M \neq M_0 \\ TestB. & H_0 : M \leq M_0, & H_1 : M > M_0 \\ TestC. & H_0 : M \geq M_0, & H_1 : M < M_0 \end{array}$$

When we assumed that the sample is from a normal population and the population variance (σ^2) is known, the test statistic is

$$Z = \frac{\bar{X} - M_0}{\sigma/\sqrt{n}}, \tag{1}$$

where Z has a standard normal distribution, when σ is not known, is replaced by it's maximum likelihood estimate, but when σ is not known, an alternative test statistic is used

$$T = \frac{\bar{X} - M_0}{S/\sqrt{n}}, \tag{2}$$

where S is the sample standard deviation and T follows t-distribution with $n-1$ degrees of freedom. In general, when the distribution of the population is unknown but sample size $n \geq 30$ then also in practice, the test statistic (2) is used as the Central Limit Theorem applies and T is equivalent to Z. Some modifications of T test are studied by Chen (1995) and the references there in.

For smaller samples and when the population distribution is not known, nonparametric procedures are used. In the following section we describe two commonly used nonparametric procedures.

3 Nonparametric Tests for the Center Parameter

Since populations do not always meet the assumptions of underlying parametric tests, we frequently need inferential procedures whose validity does not depend on rigid assumptions. Nonparametric statistical procedures meet this need in many instances. Nonparametric methods require minimal assumptions about the form of the distribution of the population. For instance, it might be assumed that the data are randomly selected from a population that has a continuous distribution, but no other assumptions are made. The simplicity of nonparametric methods, the widespread availability of such methods in statistical packages and the desirable statistical properties of such methods make them powerful tools in data analysis.

In general, two main types of statistical procedures are treated as nonparametric which are Truly Nonparametric methods and Distribution-free procedures. True nonparametric procedures are not concerned with population parameters. The distribution-free procedures do not depend on the functional form of the population from which the samples are drawn.

The following are some situations where nonparametric methods are appropriate.

1. The hypothesis to be tested does not involve a population parameter.
2. The data has been measured on a scale weaker than required for the parametric procedure that would otherwise be applied.
3. The assumptions necessary for validity of the parametric procedures are not met.

Here we will discuss nonparametric procedures regarding the center parameter of a distribution. In standard parametric tests, mean is considered as the location parameter, but in nonparametric tests, usually, the median is considered as the location parameter.

3.1 One Sample Sign Test

The sign test is designed to test a hypothesis about the location of a population distribution. It is most often used to test the hypothesis about a population median. The hypotheses can be formulated as given in Section 2 and will be referred throughout.

The procedure is, record the sign of the difference obtained by subtracting the hypothesized median M_0 from each sample value as $X_i - M_0$. So, there will be n signed differences $i = 1, 2, \ldots, n$.

If the null hypothesis is true - that is, if the population median is in fact equal to M_0 - we expect a random sample from the population to have about as many as plus signs as many as minus signs when the n differences $X_i - M_0$ have been computed. If we observe a sufficiently small number of either plus or minus signs, we reject null hypothesis A. If we observe a sufficiently small number of minus signs, we reject null hypothesis B, and if we observe a sufficiently small number of plus signs, we reject null hypothesis C. This test is very simple to administer as it requires Binomial probabilities to compute the P-values which are readily available in almost all elementary statistics textbooks and also easy to compute using a scientific calculator.

Decision Rule:

Test A: Reject H_0 at the α level of significance if the probability, when H_0 is true, of observing as few (or fewer) of the less frequently occurring sign in a random sample of size n is less than or equal to $\alpha/2$.

Test B: Reject H_0 at the α level of significance if the probability, when H_0 is true, of observing as few (or fewer) minus signs as are actually observed in a random sample of size n is less than or equal to α.

Test C: Reject H_0 at the α level of significance if the probability, when H_0 is true, of observing as few (or fewer) plus signs as are actually observed in a random sample of size n is less than or equal to α.

For small sample, this test is 95% efficient. The relative efficiency of sign test decreases as the population size increases. For further details and power comparisons the readers are referred to Daniel (1990) and the references therein.

3.2 Wilcoxon Signed Rank Test

The Wilcoxon Signed Rank test is designed to test a hypothesis about the location (median) of a population distribution. Wilcoxon test uses the magnitude of differences $X_i - M_0$ as given in Section 3.1. We need to know whether a sample measurement falls above or below the average.

To use the Wilcoxon procedure, we first rank the differences in order of absolute size. Then we assign the original signs of the differences to the ranks and compute two sums: the sum of the ranks with negative signs and the sum of the ranks with positive signs. Since the Wilcoxon signed-rank test uses more information than the sign test, it is often a more powerful test. The Wilcoxon signed-rank test assumes that the sampled population is symmetric. When the sampled population meets this assumption, then the conclusions about the population median also applies to the population mean. When the population is not symmetric, sign test is preferred over the Wilcoxon test.

To obtain the test statistic, we use the following procedure:

1. Subtract the hypothesized median from each observation; that is for each element, find $D_i = X_i - M_0$ If any D_i is equal to zero, that element is eliminated from calculation and sample size is reduced accordingly.

2. Rank the differences $|D_i|$ from the smallest to the largest.

3. Assign the sign to each rank according to the sign of the difference of which it is the rank.

4. Obtain the sum of the ranks with positive signs T^+ and obtain the sum of the ranks with negative signs T^-. We take the smaller of this two - T^+ and T^-, as the test statistic. Then we compare this value to the critical value from the Table A.3 given in Daniel (1990).

The table for the critical values for this test is also readily available in the textbooks where the method is introduced. A good number of elementary statistics textbooks, all nonparametric textbooks, and almost all applied statistics textbooks would introduce the method.

Decision rule:

Test A: We reject H_0 at the level of significance α if the calculated value of T is smaller than or equal to tabulated T for n and pre-selected $\alpha/2$.

Test B: Reject H_0 at the level of significance α if T^+ is less than or equal to tabulated T for n and pre-selected α.

Test C: Reject H_0 at the level of significance α if T^- is less than or equal to tabulated T for n and pre-selected α.

Wilcoxon signed rank test is very efficient. However, the efficiency depends on the distribution of D_i. The best efficiency occurs if D_i is normally distributed. For further details and power comparisons the readers are referred to Daniel (1990) and the references therein.

4 The Proposed Test for Center Using the Median

Let us consider $X_1, X_2, \ldots, X_n$ be a random sample of size n from a population with the density function $f(x)$ and the corresponding distribution faunction is $F(x)$. Let the ordered sample be $Y_1, Y_2, \ldots, Y_n$. Let Y_k be the median of the sample, then the density function of Y_k can be written as

$$g(y) = \frac{n!}{(k-1)!(n-k)!}(F(y))^{k-1} f(y)(1-F(y))^{n-k} \tag{3}$$

where $k = (n+1)/2$ for n is odd. For n even, densities for Y_k is obtained by obtaining for $k = n/2$ and for $k = n/2 + 1$ separately using (3) and then by finding the density for the average of the two corresponding random variables. In equation (3), a specified H_0 can be incorporated by replacing $f(x)$ and $F(x)$ by $f(x; M_0)$ and $F(x; M_0)$. Then the corresponding P-value is computed for the sample median m using $g(y)$ in (3) as $2P(Y_k \leq m)$, if $m < M_0$, $2P(Y_k \geq m)$, if $m > M_0$ for the two-tailed test A. Similary, for the one-tailed tests B and C, the P-values are computed as $P(Y_k \geq m)$ and $P(Y_k \leq m)$, respectively. In cases when $g(y)$ and $G(y)$ are not in simpler form, that is, the P-values could not be computed explicitly, a numerical integration would be used as can be seen for Normal parent populations, and are described in the following section.

4.1 Test For The Center Using The Median: Normal Case

Let us consider the sample is from a standard normal population. Then (3) can be written as

$$g(y) = \frac{n!}{(k-1)!(n-k)!}(\Phi(y))^{k-1}\phi(y)(1-\Phi(y))^{n-k} \tag{4}$$

where $\phi(y)$ is the standard normal pdf (probability density function) and $\Phi(y)$ is the standard normal cdf (cumulative distribution function). Then the P-values can be obtained using the procedure described in Section 4. When the data is from a normal population with mean μ and variance σ^2, the unbiased estimates are used and standardized to facilitate the use of (4). That is, equations (1) or (2) are used along with equation (4) in computing P-values.

In the following Section 5 we perform the Monte-Carlo simulation to study the level of significance and the power. In Section 5.1, in the level of significance study, the samples are from a standard normal population. In Section 5.2, samples are selected from non-normal population to compare the quantiles of the p-values. In Section 5.3 a power study is performed using simulation. Then in Section 6 we apply the process along with the standard procedures for two different data sets to show some real life applications and their feasibilities.

5 Simulation Study

To study the performance of the test when H_0 is true, we compute P-values when data is obtained from a standard normal distribution and from a non-normal distribution. Then we performed the power study to compare with the competitors.

5.1 Level of significance

Ten thousand samples are drawn from the standard normal distribution for sample sizes $n = 5, 9, 13, 17, 25, 31, 41, \text{and } 51$. Here we considered odd sample sizes only to simplify the computations. We computed the P-values for the Test A (two-tailed) with $M_0 = 0$ using the standard parametric Z test (1), the proposed median test (MED) is computed using (4), using the Sign Test (BIN) described in Section 3.1., and using the Wilcoxon Signed Rank Test (W) described in Section 3.2. All the computations are done using the MATLAB software.

We can see in Table 1 that in all the tests the P-values' quantiles are closely related to the corresponding percentages. For all sample sizes, Z and MED tests are closer than BIN and W tests. MED test has performed better than Z test except for sample size 31 (see Table 1a).

5.2 Quantiles for P-values for Non-normal Samples

In real life, the practitioners rarely investigate whether the data is from the assumed distribution or not. Here we investigate the robustness of the underlying distribution of the population from which the samples are by considering samples from a non-normal distribution.

Ten thousand samples are drawn from the standard exponential distribution for sample sizes $n = 9, 13, 17, 25, 31, 41,$ and 51. Quantiles for the P-values for the Test A (two-tailed) with $M_0 = 0.6931$ and $M_0 = 1$ for the four procedures as in Section 5.1. Here we considered $M_0 = 0.6931$ as the median of the standard exponential distribution is 0.6931 and $M_0 = 1$ is considered as the mean of the

Table 1: Pvalue quantiles when $M_0 = 0$ and the samples are from the standard normal distribution

Test	q_{01}	q_{05}	q_{10}	q_{50}	q_{90}	q_{95}	q_{99}
			$n = 5$				
Z	0.0078	0.0474	0.0995	0.5044	0.9064	0.9513	0.9904
MED	0.0092	0.0533	0.1047	0.5000	0.9008	0.9499	0.9903
BIN	0.0625	0.0625	0.3750	1.0000	1.0000	1.0000	1.0000
W	0.0625	0.0625	0.1250	0.6250	1.0000	1.0000	1.0000
			$n = 9$				
Z	0.0108	0.0523	0.1043	0.5010	0.8966	0.9467	0.9884
MED	0.0099	0.0519	0.1029	0.4995	0.9011	0.9519	0.9892
BIN	0.0391	0.1797	0.1797	0.5078	1.0000	1.0000	1.0000
W	0.0117	0.0547	0.1289	0.5703	0.9102	1.0000	1.0000
			$n = 13$				
Z	0.0082	0.0493	0.0985	0.4948	0.9033	0.9513	0.9915
MED	0.0096	0.0494	0.0973	0.4992	0.8962	0.9451	0.9895
BIN	0.0225	0.0923	0.2668	0.5811	1.0000	1.0000	1.0000
W	0.0105	0.0574	0.1099	0.5417	0.9460	1.0000	1.0000
			$n = 17$				
Z	0.0107	0.0514	0.0982	0.4969	0.9004	0.9515	0.9916
MED	0.0096	0.0494	0.0973	0.5034	0.9057	0.9534	0.9907
BIN	0.0127	0.0490	0.1435	0.6291	1.0000	1.0000	1.0000
W	0.0148	0.0495	0.0929	0.4925	0.9058	0.9434	0.9811
			$n = 25$				
Z	0.0096	0.0463	0.0977	0.4948	0.8973	0.9468	0.9889
MED	0.0081	0.0439	0.0977	0.4998	0.8958	0.9502	0.9887
BIN	0.0146	0.1078	0.1078	0.6900	1.0000	1.0000	1.0000
W	0.0119	0.0480	0.0980	0.4758	0.9036	0.9464	0.9893
			$n = 31$				
Z	0.0095	0.0504	0.1017	0.5030	0.9013	0.9487	0.9872
MED	0.0120	0.0525	0.0992	0.4975	0.8984	0.9482	0.9911
BIN	0.0107	0.0708	0.1496	0.7201	1.0000	1.0000	1.0000
W	0.0115	0.0500	0.0997	0.5052	0.8909	0.9531	0.9844
			$n = 41$				
Z	0.0118	0.0547	0.1086	0.5018	0.8957	0.9467	0.9891
MED	0.0113	0.0545	0.1040	0.5057	0.8998	0.9488	0.9897
BIN	0.0115	0.0596	0.1173	0.5327	1.0000	1.0000	1.0000
W	0.0124	0.0560	0.1067	0.5045	0.8969	0.9535	0.9948
			$n = 51$				
Z	0.0093	0.0537	0.1029	0.5009	0.9003	0.9521	0.9913
MED	0.0105	0.0507	0.1011	0.4939	0.9020	0.9521	0.9886
BIN	0.0110	0.0919	0.1608	0.5758	1.0000	1.0000	1.0000
W	0.0112	0.0523	0.0990	0.4997	0.8956	0.9477	0.9925

Table 1a: Rankings of absolute differences of P-values and the Quantiles

Test	q_{01}	q_{05}	q_{10}	q_{50}	q_{90}	q_{95}	q_{99}
	$n=5$						
Z	2	1	1	2	2	2	2
MED	1	2	2	1	1	1	1
BIN	3.5	3.5	4	4	3.5	3.5	3.5
W	3.5	3.5	3	3	3.5	3.5	3.5
	$n=9$						
Z	2	2	2	2	2	2	2
MED	1	1	1	1	1	1	1
BIN	4	4	4	3	4	3.5	3.5
W	3	3	3	4	3	3.5	3.5
	$n=13$						
Z	3	2	1	2	1	1	1
MED	1	1	2	1	2	2	2
BIN	4	4	4	4	4	3.5	3.5
W	2	3	3	3	3	3.5	3.5
	$n=17$						
Z	2	4	1	1	1	1	2
MED	1	2	2	2	2	2	1
BIN	3	3	4	4	4	4	4
W	4	1	3	3	3	3	3
	$n=25$						
Z	1	2	2.5	2	1	2	1
MED	2	3	2.5	1	3	1	2
BIN	4	4	4	4	4	4	4
W	3	1	1	3	2	3	3
	$n=31$						
Z	1	2	3	2	1	1	2
MED	4	3	2	1	3	2	1
BIN	2	4	4	4	4	4	4
W	3	1	1	3	2	3	3
	$n=41$						
Z	3	2	3	1	3	2	2
MED	1	1	1	3	1	1	1
BIN	2	4	4	4	4	4	4
W	4	3	2	2	2	3	3
	$n=51$						
Z	2	3	3	2	1	1.5	1
MED	1	1	2	3	2	1.5	3
BIN	3	4	4	4	4	4	4
W	4	2	1	1	3	3	2

Table 2: P-value Quantiles when $M_0 = 0.6931$ and samples are from the standard exponential distribution

Test	q_{01}	q_{05}	q_{10}	q_{50}	q_{90}	q_{95}	q_{99}
$n = 9$							
Z	0.0173	0.0461	0.0780	0.3417	0.8339	0.9138	0.9844
MED	0.0007	0.0239	0.0747	0.5402	0.9118	0.95199	0.9931
BIN	0.0391	0.1797	0.1797	0.7539	1.0000	1.0000	1.0000
W	0.0195	0.0547	0.0977	0.4961	0.9102	1.0000	1.0000
$n = 13$							
Z	0.0067	0.0246	0.0458	0.2486	0.7712	0.8742	0.9694
MED	0.0012	0.0314	0.0774	0.5381	0.9217	0.9614	0.9942
BIN	0.0225	0.0923	0.0923	0.5811	1.0000	1.0000	1.0000
W	0.0054	0.0327	0.0681	0.4143	0.8926	0.9460	1.0000
$n = 17$							
Z	0.0072	0.0249	0.0435	0.2146	0.7355	0.8481	0.9547
MED	0.0067	0.0536	0.1017	0.5402	0.9123	0.9587	0.9933
BIN	0.0490	0.1435	0.1435	0.6291	1.0000	1.0000	1.0000
W	0.0086	0.0312	0.0684	0.3812	0.8684	0.9434	0.9811
$n = 25$							
Z	0.0023	0.0079	0.0162	0.1338	0.6016	0.7825	0.9532
MED	0.0081	0.0665	0.1276	0.5907	0.9246	0.9630	0.9952
BIN	0.0146	0.1078	0.1078	0.6900	1.0000	1.0000	1.0000
W	0.0035	0.0128	0.0347	0.3674	0.8296	0.9250	0.9893
$n = 31$							
Z	0.0012	0.0051	0.0108	0.1045	0.5145	0.6839	0.9297
MED	0.0142	0.0683	0.1392	0.5696	0.9126	0.9518	0.9911
BIN	0.0294	0.0708	0.1496	0.7201	1.0000	1.0000	1.0000
W	0.0023	0.0125	0.0275	0.3272	0.8446	0.9219	0.9844
$n = 41$							
Z	0.0004	0.0021	0.0049	0.0571	0.3671	0.6305	0.9480
MED	0.0148	0.0664	0.1336	0.5841	0.9120	0.9570	0.9856
BIN	0.0115	0.0596	0.1173	0.5327	1.0000	1.0000	1.0000
W	0.0010	0.0076	0.0214	0.2679	0.8156	0.9123	0.9845
$n = 51$							
Z	0.0003	0.0010	0.0023	0.0298	0.2571	0.4224	0.8146
MED	0.0228	0.1023	0.1704	0.5631	0.9111	0.9475	0.9886
BIN	0.0110	0.0489	0.0919	0.5758	1.0000	1.0000	1.0000
W	0.0006	0.0048	0.0139	0.2091	0.7714	0.8956	0.9813

Table 2a: Rankings of absolute differences of P-values and the Quantiles

Test	q_{01}	q_{05}	q_{10}	q_{50}	q_{90}	q_{95}	q_{99}
	$n = 9$						
Z	3	1	2	3	3	2	2
MED	1	3	3	2	2	1	1
BIN	4	4	4	4	4	3.5	3.5
W	3	2	1	1	1	3.5	3.5
	$n = 13$						
Z	3	3	4	4	4	4	4
MED	1	2	2	1	2	2	1
BIN	4	4	1	2	3	3	2.5
W	2	1	3	3	1	1	2.5
	$n = 17$						
Z	2	3	4	4	4	4	4
MED	3	1	1	1	1	2	1
BIN	4	4	3	2	3	3	3
W	1	2	2	3	2	1	2
	$n = 25$						
Z	2	4	4	4	4	4	4
MED	1	1	2	1	1	1	1
BIN	4	3	1	3	3	3	3
W	3	2	3	2	2	2	2
	$n = 31$						
Z	3	4	4	4	4	4	4
MED	1	1	1	1	1	1	1
BIN	4	2	2	3	3	3	3
W	2	3	3	2	2	2	2
	$n = 41$						
Z	4	4	4	4	4	4	4
MED	2	2	2	2	1	1	1
BIN	1	1	1	1	3	3	3
W	3	3	3	3	2	2	2
	$n = 51$						
Z	4	3	4	4	4	4	4
MED	2	4	2	1	1	1	1
BIN	1	1	1	2	2	2	3
W	3	2	3	3	3	3	2

Table 3: P-value Quantiles when $M_0 = 1$ and samples are from the standard exponential distribution

Test	q_{01}	q_{05}	q_{10}	q_{50}	q_{90}	q_{95}	q_{99}
$n = 9$							
Z	0.0005	0.0126	0.0533	0.4678	0.8932	0.9460	0.9926
MED	0.0000	0.0010	0.0093	0.3508	0.8661	0.9411	0.9882
BIN	0.0039	0.0391	0.0391	0.5078	1.0000	1.0000	1.0000
W	0.0039	0.0273	0.0742	0.4961	0.9102	1.0000	1.0000
$n = 13$							
Z	0.0010	0.0263	0.0738	0.4860	0.8879	0.9489	0.9917
MED	0.0000	0.0019	0.0098	0.3065	0.8465	0.9264	0.9883
BIN	0.0129	0.0225	0.0923	0.5811	1.0000	1.0000	1.0000
W	0.0024	0.0266	0.0574	0.4548	0.8926	0.9460	1.0000
$n = 17$							
Z	0.0006	0.0160	0.0582	0.4674	0.8978	0.9439	0.9881
MED	0.0000	0.0012	0.0064	0.2247	0.7920	0.8977	0.9814
BIN	0.0023	0.0127	0.0490	0.3323	1.0000	1.0000	1.0000
W	0.0038	0.0168	0.0395	0.3686	0.8684	0.9434	0.9811
$n = 25$							
Z	0.0042	0.0334	0.0867	0.5156	0.9123	0.9566	0.9896
MED	0.0001	0.0019	0.0091	0.2097	0.7733	0.8863	0.9759
BIN	0.0009	0.0146	0.0146	0.2295	1.0000	1.0000	1.0000
W	0.0034	0.0144	0.0347	0.3533	0.8612	0.9250	0.9893
$n = 31$							
Z	0.0014	0.0242	0.0917	0.4900	0.9044	0.9560	0.9879
MED	0.0000	0.0009	0.0067	0.1567	0.7387	0.8684	0.9839
BIN	0.0002	0.0033	0.0107	0.2810	1.0000	1.0000	1.0000
W	0.0018	0.0084	0.0282	0.3369	0.8293	0.9219	0.9844
$n = 41$							
Z	0.0022	0.0241	0.0866	0.4821	0.8825	0.9497	0.9884
MED	0.0000	0.0006	0.0024	0.1068	0.5965	0.7794	0.9713
BIN	0.0001	0.0015	0.0043	0.1173	0.7552	1.0000	1.0000
W	0.0008	0.0072	0.0168	0.2679	0.8307	0.9123	0.9742
$n = 51$							
Z	0.0025	0.0392	0.0850	0.4955	0.8923	0.9521	0.9876
MED	0.0000	0.0004	0.0016	0.0658	0.5535	0.7572	0.9475
BIN	0.0001	0.0006	0.0018	0.0919	0.5758	0.7798	1.0000
W	0.0004	0.0055	0.0154	0.2266	0.7678	0.8771	0.9626

Table 3a: Rankings of absolute differences of P-values and the Quantiles

Test	q_{01}	q_{05}	q_{10}	q_{50}	q_{90}	q_{95}	q_{99}
	$n = 9$						
Z	2	3	2	3	1	1	2
MED	3	4	4	4	4	2	1
BIN	1.5	1	3	2	2	3.5	3.5
W	1.5	2	1	1	3	3.5	3.5
	$n = 13$						
Z	3	2	2	1	2	1	1.5
MED	4	4	4	4	3	3	1.5
BIN	1	3	1	3	4	4	2.5
W	2	1	3	2	1	2	2.5
	$n = 17$						
Z	3	2	1	1	1	1	1
MED	4	4	4	4	4	4	2
BIN	2	3	2	3	3	3	4
W	1	1	3	2	2	2	3
	$n = 25$						
Z	1	1	1	1	1	1	1
MED	4	4	4	4	4	4	4
BIN	3	2	3	3	3	3	3
W	2	3	2	2	2	2	2
	$n = 31$						
Z	2	1	1	1	1	1	1
MED	4	4	4	4	4	4	3
BIN	3	3	3	3	3	3	4
W	1	2	2	2	2	2	2
	$n = 41$						
Z	1	1	1	1	1	1	1
MED	4	4	4	4	4	4	4
BIN	3	3	3	3	3	3	2
W	2	2	2	2	2	2	3
	$n = 51$						
Z	1	1	1	1	1	1	1
MED	4	4	4	4	4	4	4
BIN	3	3	3	3	3	3	2
W	2	2	2	2	2	2	3

standard exponential distribution and a reasonable value to be considered as a center of the distribution. In Table 2, samples are generated from the standard exponential distribution and $M_0 = 0.6931$. Here we are expecting that the percentiles are not closely estimated in case of the Z test. Which is clearly seen for larger samples. For the MED test and the BIN test the quantiles are closer to the true percentiles compared to the W test. For the small samples, the MED test performed better than the BIN test. The rankings of the absolute differences for the P-values and quantile percentages are given in Table 2a.

It is to be noted here that if the median is the center parameter, performance of the MED test is comparable with the BIN test and even better for smaller samples though the samples are from a non-normal population. Also to be reminded that for a skewed data the BIN test performs better than the W test.

In Table 3, samples are again from the standard exponential population but $M_0 = 1$. Here we are expecting the Z test should show closer percentile values and which is true. But Table 3 is different than compared to Table 2 and Table 1 as the null hypothesis is no longer true if the median is the center parameter. In the Z test, the mean is the center parameter and hence the performance of Z as expected. The median is the center in the other three tests and are comparable. The worst performance in terms of the closeness to the true percentage is better as the test would be more powerful. Here we see that the MED test is showing higher deviation and hence better. The rankings of the deviations are given in Table 3a.

5.3 Power study

Here we generate random samples from a distribution which is not exactly specified by the null hypothesis. Then we count the number of rejections of H_0 using $\alpha = 0.01$ and $\alpha = 0.05$. The proportions of rejections are reported in Table 2. We consider different situations such as samples are from standard normal distribution and $M_0 = 1$, samples are from standard exponential distribution and $M_0 = 1, 0.69, 0.85$ as we know that the median of a standard exponential distribution is very close to 0.69. Then to represent a totally unrecognizable distribution, we consider two distributions which are the Claw Density and the Double Claw Density.

The Claw Density and the Double Claw Density are the following mixtures of the normal densities:

Claw Density

$$\frac{1}{2}N(0,1) + \sum_{l=0}^{4} \frac{1}{10} N\left(\frac{l}{2} - 1, \left(\frac{1}{10}\right)^2\right)$$

Double Claw Density

$$\frac{49}{100}N\left(-1, \left(\frac{2}{3}\right)^2\right) + \frac{49}{100}N\left(1, \left(\frac{2}{3}\right)^2\right) + \sum_{l=0}^{6} \frac{1}{350} N\left(\frac{l-3}{2}, \left(\frac{1}{100}\right)^2\right).$$

In Fig 1, the pictures of the Claw Density and the Double Claw Density are displayed as they not commonly used.

Figure 11: Claw and Double Claw Density

Table 4: Power Comparisons

α	Z	MED	BIN	W	α	Z	MED	BIN	W
Sample from $N(0,1)$; $H_0: M=1$; $H_1: M \neq 1$									
$n=13$									
0.01	0.703	1.000	0.360	0.645	0.05	0.919	1.000	0.680	0.899
$n=25$									
0.01	0.977	1.000	0.808	0.970	0.05	0.998	1.000	0.969	0.998
$n=51$									
0.01	1.000	1.000	0.999	1.000	0.05	1.000	1.000	1.000	1.000
Sample from $Exp(1)$; $H_0: M=1$; $H_1: M \neq 1$									
$n=13$									
0.01	0.038	1.000	0.017	0.024	0.05	0.092	1.000	0.091	0.108
$n=25$									
0.01	0.028	1.000	0.062	0.038	0.05	0.083	1.000	0.225	0.131
$n=51$									
0.01	0.020	1.000	0.160	0.070	0.05	0.061	1.000	0.452	0.204
Sample from $Exp(1)$; $H_0: M=0.69$; $H_1: M \neq 0.69$									
$n=13$									
0.01	0.012	1.000	0.005	0.010	0.05	0.0910	1.000	0.019	0.074
$n=25$									
0.01	0.056	1.000	0.006	0.037	0.05	0.024	1.000	0.052	0.124
$n=51$									
0.01	0.291	1.000	0.008	0.082	0.05	0.610	1.000	0.044	0.241
Sample from $Exp(1)$; $H_0: M=0.85$; $H_1: M \neq 0.85$									
$n=13$									
0.01	0.010	1.000	0.010	0.008	0.05	0.042	1.000	0.045	0.044
$n=25$									
0.01	0.011	1.000	0.014	0.008	0.05	0.059	1.000	0.012	0.049
$n=51$									
0.01	0.019	1.000	0.037	0.008	0.05	0.126	1.000	0.159	0.061
Sample from $Claw(0, 11/20)$; $H_0: M=1$; $H_1: M \neq 1$									
$n=13$									
0.01	0.831	1.000	0.494	0.781	0.05	0.963	1.000	0.778	0.955
$n=25$									
0.01	0.997	1.000	0.904	0.995	0.05	1.000	1.000	0.991	1.000
$n=51$									
0.01	1.000	1.000	1.000	1.000	0.05	1.000	1.000	1.000	1.000
Sample from $DoubleClaw(0, 34309/52500)$; $H_0: M=1$; $H_1: M \neq 1$									
$n=13$									
0.01	0.490	1.000	0.120	0.365	.05	0.790	1.000	0.321	0.717
$n=25$									
0.01	0.911	1.000	0.370	0.807	.05	0.987	1.000	0.726	0.959
$n=51$									
0.01	1.000	1.000	0.818	0.999	.05	1.000	1.000	0.960	1.000

Table 4a: Power Rankings

α	Z	MED	BIN	W	α	Z	MED	BIN	W
Sample from $N(0,1); H_0 : M = 1; H_1 : M \neq 1$									
$n = 13$									
0.01	3	1	4	2	0.05	2	1	4	3
$n = 25$									
0.01	2	1	4	3	0.05	2.5	1	4	2.5
$n = 51$									
0.01	2	2	4	2	0.05	2.5	2.5	2.5	2.5
Sample from $Exp(1); H_0 : M = 1; H_1 : M \neq 1$									
$n = 13$									
0.01	2	1	4	3	0.05	3	1	4	2
$n = 25$									
0.01	4	1	2	3	0.05	4	1	2	3
$n = 51$									
0.01	4	1	2	3	0.05	4	1	2	3
Sample from $Exp(1); H_0 : M = 0.69; H_1 : M \neq 0.69$									
$n = 13$									
0.01	2	1	4	3	0.05	2	1	4	3
$n = 25$									
0.01	2	1	4	3	0.05	4	1	3	2
$n = 51$									
0.01	2	1	4	3	0.05	2	1	4	3
Sample from $Exp(1); H_0 : M = 0.85; H_1 : M \neq 0.85$									
$n = 13$									
0.01	2.5	1	4	2.5	0.05	4	1	2	3
$n = 25$									
0.01	3	1	2	4	0.05	2	1	4	3
$n = 51$									
0.01	3	1	2	4	0.05	3	1	2	4
Sample from $Claw(0, 11/20);\ H_0 : M = 1; H_1 : M \neq 1$									
$n = 13$									
0.01	2	1	4	3	0.05	2	1	4	3
$n = 25$									
0.01	2	1	4	3	0.05	2	2	4	2
$n = 51$									
0.01	2.5	2.5	2.5	2.5	0.05	2.5	2.5	2.5	2.5
Sample from $DoubleClaw(0, 34309/52500);\ H_0 : M = 1; H_1 : M \neq 1$									
$n = 13$									
0.01	2	1	4	3	.05	2	1	4	3
$n = 25$									
0.01	2	1	4	3	.05	2	1	4	3
$n = 51$									
0.01	1.5	1.5	4	3	.05	2	2	4	2

In both cases we have used $M_0 = 1$. Here we have considered samples of sizes $n = 13, 25, 51$. As the samples are drawn from distributions with mean and standard deviation different that the standard normal distribution, data needed to be transformed. In the transformations we have used M_0 as the center and estimate of the standard deviation as the scale. In equation (1) σ is replaced by the maximum likelihood estimate and then MED test is computed and denoted as MED.

In Table 4, in all cases, the MED test shows the highest power. When the samples are from the exponential distribution, the MED test outperforms all the tests. In non-exponential samples, the Z test and the W test performed pretty well. The power rankings are given in Table 4a.

6 Application

Here we considered two examples to demonstate the application of the proposed method for both odd and even sample sizes. We also computed the P-values for the competing statistics mentioned in this paper. In Section 6.1. we give an example for odd sample size and in Section 6.2. we give an example for even sample size.

6.1 A Study of Myocardial Transit Times

In a study of myocardial transit times, Liedtke et al. (1973) measured appearance transit times in a series of subjects with angiographically normal right coronary arteries. The median appearance time for this group was 3.50 seconds. Another research team repeated the procedure on a sample of 11 patients with significantly occluded right coronary arteries and obtained the results shown in the following Table 5. The data is obtained from Daniel (1990, p.35). Could this research be concluded that the median of appreance transit time in the population from which its sample was drawn is different from 3.5 seconds?

Table 5: Appreance transit times for 11 patients with significantly occluded right coronary arteries

Sub.	1	2	3	4	5	6	7	8	9	10	11
Time	1.8	3.3	5.65	2.25	2.5	3.5	2.75	3.25	3.1	2.7	3.0

Here $M_0 = 3.5$, then using (1), (2), and (4) are used to compute P-values for Z, T (here, T test is used to satisfy the curiosity of a general audience), and MED tests as 0.1719, 0.1824, and 0.1719, respectively. While P-values for the Sign test and the Wilcoxon Signed Rank test are 0.0215 and 0.0840. We notice that using the standard 5% level of significance, only for the Sign test, the decision would be different. According to literature, for such continuous data, the Wilcoxon Signed Rank test is more reliable than the Sign test. We also notice that the P-values for all the parametric tests are similar.

6.2 A Study of Sleep Patterns

Agnew et al. (1967) in a study of sleep patterns, reported the data shown in Table 6 and obtained from Daniel (1990, p.47) shows the percentage of total sleep time spent in stage 0 sleep by 16 mentally and physically healthy males between the ages of 50 and 60. We want to test that the median percentage of the sleep time spent in stage 0 is 5 percent.

Table 6: Data showing percentage of sleep time spent in stage 0 by 16 mentally and physically healthy males between the ages of 50 and 60

Subject	1	2	3	4	5	6	7	8
Percentage	1.90	3.08	9.10	3.53	1.99	3.10	10.16	0.69
Subject	9	10	11	12	13	14	15	16
Percentage	1.74	2.41	4.01	3.71	8.11	8.23	0.07	3.07

Here $M_0 = 5$, then using (1), (2), and (4) are used to compute P-values for Z, T, and MED tests as 0.2083, 0.2419, and 0.0532, respectively. While P-values for the Sign test and the Wilcoxon Signed Rank test are 0.0768 and 0.3520. We notice that using the standard 5% level of significance, the decision would be same for all the tests. The P-value for the MED test is the closest to 5% but comparable with the value for the Sign test.

Acknowledgements

We thank the referee for constructive comments and suggestions which improved the earlier version of the paper.

References

Agnew, H. W., Jr., W. W. Webb, and R. L. Williams (1967). "Sleep Patterns in Late Middle Age Males: An EEG Study," *Electroencephalog. Clin. Neurophysiol.*, **23**, 168-171.

Chen, L. (1995). "Testing the Mean of Skewed Distributions," *Journal of the American Statistical Association*, **90**(430), 767-772.

Daniel, Wayne W. (1990). *Applied Nonparametric Statistics*, Duxbury, Thomson Learning, California, USA.

Liedtke, A. J., H. G. Kemp, D. M. Borkenhagen, and R. Gorlin (1973). "Myocardial Transit Times from Intra-Coronary Dye–Dilution Curves in Normal Subjects and Patients with Coronary Artery Disease," *American Journal of Cardiology*, **32**, 831-839.

Savur, S. R.(1937). "The Use of Median in Tests of Significance," *Proc. Indian Acad. Science*, **A5**, 564-576.

International Journal of Statistical Sciences *ISSN 1683–5603*
Vol. 6 (Special Issue), 2007, pp 181-192

On Fixed Critical Value for Preliminary Test Estimator

B. M. Golam Kibria
Dept. of Statistics
Florida International University
Miami, Florida 33199, USA
E-mail:kibriag@fiu.edu

Saralees Nadarajah
Dept. of Statistics
University of Nebraska
Lincoln, Nebraska 68583, USA
E-mail: snadaraj@unlserve.unl.edu

[Received October 14, 2005; Revised June 17, 2007; Accepted August 19, 2007]

Abstract

The performance of pre-test estimator depends on the level of significance. Since the existing max-min rule and minimax regret procedure are computer intensive, we propose one simpler alternative method for optimal level of significance. We perform a numerical comparison among these three methods. Numerical results suggest that the proposed and Brook methods are conservative for fixed sample size, whereas Han and Bancroft is flexible. If the researchers are very conservative about the minimum guaranteed efficiency, they might select our proposed or Brook's method. If they want to have the higher minimum guaranteed efficiency, they should select Han and Bancroft method. However, in later case the researchers have to accept the risk for the higher size of the test. The proposed method is easy to compute compared to Han and Bancroft and Brook methods.

Keywords and Phrases: Mean of F-distribution; Guaranteed Efficiency; Linear Regression; Preliminary Test; Optimal Level of Significance; Quadratic risk, Restricted Estimator, Unrestricted estimator.

AMS Classification: 62F10, 62F03.

1 Introduction

One common problem encountered with general linear regression models is to determine whether to place restrictions on the parameters or not. This leads to a choice of considering either restricted or unrestricted least squares estimator. For selecting either estimator, F-test statistic is used to make the decision. This encourages one to define a pre-test estimator. To describe the problem, consider the following linear regression model, $Y \sim N(X\beta, \sigma^2 I)$, where Y is an $n \times 1$ vector of observations on the dependent variable, which follows a normal distribution with fixed mean vector $X\beta$ and known variance $\sigma^2 I$, β is a $p \times 1$ vector of unknown regression parameters, and X is an $n \times p$ known design matrix of rank p $(n \geq p)$. We are interested to estimate the regression coefficients β when it is *a priori* suspected that β may be restricted to the subspace

$$H_0 : H\beta \quad = \quad h, \tag{1}$$

where H is a $q \times p$ known matrix of full rank $q(< p)$ and h is a $q \times 1$ vector of known constants. The choice of estimator for β whether restricted or unrestricted will depend on the outcome of the test. If we reject the null hypothesis, the unrestricted least squares estimator (URLSE) $\hat{\beta}^{UE} = C^{-1}X'Y$ will be used. Here, $C = X'X$ is called the information matrix. On the other hand, if the null hypothesis is true, the restricted least squares estimator (RLSE) $\hat{\beta}^{RE} = \hat{\beta}^{UE} - C^{-1}H'(HC^{-1}H')^{-1}(H\hat{\beta}^{UE} - h)$ will be used. As a result, one might combine the URLSE and RLSE to obtain a better performance of the estimator in presence of the uncertain prior information $H\beta = h$. This leads to the well known preliminary test least squares estimator (PTLSE) of β given by

$$\hat{\beta}_{\alpha}^{PT} = \hat{\beta}^{UE} - (\hat{\beta}^{UE} - \hat{\beta}^{RE})I(\mathcal{L} < \mathcal{L}_{\alpha}), \tag{2}$$

where $I(A)$ is the indicator function of the set A and $\mathcal{L}_{\alpha}$ is the upper $100\alpha\%$ point of the test statistic

$$\mathcal{L} = \frac{(H\hat{\beta}^{UE} - h)'(HC^{-1}H')^{-1}(H\hat{\beta}^{UE} - h)}{qs^2}, \tag{3}$$

where $s^2 = (n-p)^{-1}(y - X\hat{\beta}^{UE})'(y - X\hat{\beta}^{UE})$ is an unbiased estimator of σ^2. Under the null hypothesis, the test statistic $\mathcal{L}$ is distributed as central F distribution with q and $n - p$ degrees of freedoms. Under the non-null case, it has non-central F distribution with q and $n - p$ degrees of freedom and non-centrality parameter $\frac{1}{2}\Delta$, where

$$\Delta = \frac{\eta'(HC^{-1}H')^{-1}\eta}{\sigma^2},$$

and $\eta = H\beta - h$ is called the departure parameter.

The preliminary test estimation has application in applied econometric analysis. It has been pioneered by Bancroft (1944), followed by Bancroft (1964), Han and Bancroft (1968), Judge and Bock (1978), Benda (1996), Chiou and Han (1999), Han (2002), Kibria and Saleh (2003) and very recently Kibria and Saleh (2005, 2006). A detailed review of the preliminary test estimation procedures is given by Han at al. (1988) and Gilies and Gilies (1993).

It follows from above that the performance of PTLSE depends on the unknown parameter Δ and the size of the test α. Indeed, the choice of α or critical value for F test is an important issue for the users of the PT estimator. Han and Bancroft (1968) proposed the max-min rule based on relative efficiency and Brook (1976) proposed the minimax regret procedure based on risk to determine the optimal significance level for the usual pre-test estimator. Since both procedures are computer intensive, a fixed critical value for PTLSE is proposed in this paper. A numerical comparison among these three procedure are given and discussed their relative merits.

The plan of the paper is as follows. In Section 2, we provide the risk functions of the estimators. The determination of optimal significance level is discussed in section 3. A summary of the paper is added in Section 4.

2 The Risk Analysis

2.1 The Risk Functions

Here, we present the quadratic risk functions of the estimators. Suppose $\hat{\beta}$ denotes an estimator of β, then for a given positive semi definite matrix M, the loss function of the estimator $\hat{\beta}$ is defined as

$$L(\hat{\beta}; M) = (\hat{\beta} - \beta)' M (\hat{\beta} - \beta)$$

and the corresponding risk function of the estimator $\hat{\beta}$ is

$$R(\hat{\beta}; M) = E(\hat{\beta} - \beta)' M (\hat{\beta} - \beta) = tr(U),$$

where U is the mean-squared error matrix of the estimator $\hat{\beta}$. The quadratic risk functions of the proposed estimators are (see Judge and Bock (1978)):

$$\begin{aligned}
R(\hat{\beta}^{UE}; M) &= \sigma^2 tr(C^{-1}M), \\
R(\hat{\beta}^{RE}; M) &= \sigma^2 tr(C^{-1}M) - \sigma^2 tr(A) + \eta' D \eta, \\
R(\hat{\beta}^{PT}; M) &= \sigma^2 tr(C^{-1}M) - \sigma^2 tr(A) G_{q+2,n-p}(l_1; \Delta) \\
&+ \eta' D \eta \{2G_{q+2,n-p}(l_1; \Delta) - G_{q+4,n-p}(l_2; \Delta)\},
\end{aligned} \tag{4}$$

where $D = (HC^{-1}H')^{-1}A$, $A = HC^{-1}MC^{-1}H'(HC^{-1}H')^{-1}$, $l_1 = \frac{q}{q+2}F_{\alpha,q,n-p}$, $l_2 = \frac{q}{q+4}F_{\alpha,q,n-p}$ and $G_{a,b}(*; \Delta)$ is the cdf of non-central F distribution with a and b degrees of freedom and non-centrality parameter Δ.

Figure 12: Risk plots for $n = 10$ and different values of p, q and α.

Figure 13: Risk plots for $p = 4$, $q = 2$ and different n and α.

3.1 Han and Bancroft's Method

Here we describe the maximum and minimum (Max & Min) rule proposed by Han and Bancroft (1968) for the optimal choice of the level of significance of the PTLSE for testing the null hypothesis (1). For fixed values of p and q, the relative efficiency of the PTLSE ($\hat{\beta}^{PT}$) compared to the URLSE is a function of α and Δ. Let us denote this relative efficiency by

$$\begin{aligned} E(\alpha, \Delta) &= \frac{R(\hat{\beta}^{UE}, C\sigma^{-2})}{R(\hat{\beta}^{PT}, C\sigma^{-2})} \\ &= \left[1 - \frac{1}{p}\left\{qG_{q+2,n-p}(l_1;\Delta) - \Delta(2G_{q+2,n-p}(l_1;\Delta) - G_{q+4,n-p}(l_2;\Delta))\right\}\right]^{-1} \end{aligned} \quad (5)$$

For known p and q, the relative efficiency is a function of α and Δ. For a given α, $E(\alpha, \Delta)$ is a decreasing function of Δ in the interval $[0, \Delta_{Min}(\alpha)]$ and an increasing function of Δ in the interval $[\Delta_{Min}(\alpha), \infty]$ and $E(\alpha, \Delta) \to 1$ as $\Delta \to \infty$. For $\alpha \neq 0$, it has maximum at $\Delta = 0$ with the value

$$\begin{aligned} E_{Max}(\alpha, 0) &= \left[1 - \frac{q}{p} G_{q+2,n-p}(l_1; \Delta)\right]^{-1} (\geq 1) \\ &= \left[1 - G_{q+2,n-p}(l_1; \Delta)\right]^{-1} (\geq 1) \quad \text{if } p = q. \end{aligned} \tag{6}$$

If we consider the value of $E(\alpha, \Delta)$ at $\alpha = 0$, we have $E(0, \Delta) = [1 - \frac{q}{p} + \Delta]^{-1}$ and $E(0, \Delta) = 1$ when $\Delta = \frac{q}{p}$. Thus, the efficiency is maximum for $0 \leq \Delta \leq \frac{q}{p}$ and selects $\hat{\beta}^{RE}$ as the PTLSE of β.

From Figures 1 and 2, we observed that the PTLSE is not uniformly best compared to URLSE or RLSE. Moreover, if Δ is unknown then one follows the *minimum guaranteed efficiency* procedures proposed by Han and Bancroft (1968) which in turn determine the optimal level of significance for given minimum guaranteed efficiency say E_{Min}. One looks for a suitable α from the set $S_\alpha = \{\alpha | E(\alpha, \Delta) \geq E_{Min}\}$. The PTLSE is chosen for which $E(\alpha, \Delta)$ is maximized over all $\alpha \in S_\alpha$ and Δ. Thus, one solves the equation

$$\min_{\Delta} E(\alpha, \Delta_{Min}(\alpha)) = E_{Min}. \tag{7}$$

From (7), we obtain the optimal significance level α^* for the PTLSE with minimum guaranteed efficiency E_{Min}.

3.2 Brook's Optimal Critical Values

This section discusses the Brook (1976) regret criterion based on quadratic risk function to obtain the optimal critical value, which is also available in Kibria and Saleh (2005). For a given critical value c_α, the risk function of $\hat{\beta}^{PT}$ with $M = \sigma^{-2}C$ is obtained as

$$R(\hat{\beta}^{PT}; \sigma^{-2}C) = p - qG_{q+2,n-p}(c_\alpha; \Delta) + \Delta\{2G_{q+2,n-p}(c_\alpha; \Delta) - G_{q+4,n-p}(c_\alpha; \Delta)\} \tag{8}$$

Notice that

$$R(\hat{\beta}^{UE}; \sigma^{-2}C) = p, \quad \text{and} \quad R(\hat{\beta}^{RE}; \sigma^{-2}C) = p - q + \Delta.$$

Then $\hat{\beta}^{RE}$ is better than $\hat{\beta}^{UE}$ if $\Delta < q$ and $\hat{\beta}^{UE}$ is better than $\hat{\beta}^{RE}$ if $\Delta \geq q$. Since Δ is unknown, we want to have an optimal value of c_α, which provides a reasonable value for the risk function for all Δ.

It is clear that

$$\inf_{c_\alpha} R(\hat{\beta}^{PT}; \sigma^{-2}C) = R(\hat{\beta}^{RE}; \sigma^{-2}C) \quad \text{if} \quad \Delta < q$$

$$= \quad R(\hat{\beta}^{UE}; \sigma^{-2}C) \quad \text{if} \quad \Delta \geq q$$

Now, consider the regret function

$$\begin{aligned} Reg(\Delta, c_\alpha) &= R(\hat{\beta}^{PT}; \sigma^{-2}C) - \inf_{c_\alpha} R(\hat{\beta}^{PT}; \sigma^{-2}C) \\ &= q(1 - G_{q+2,n-p}(c_\alpha; \Delta)) \\ &- \Delta(1 - 2G_{q+2,n-p}(c_\alpha; \Delta) + G_{q+4,n-p}(c_\alpha; \Delta)) \quad \text{if} \quad \Delta < q \\ &= -qG_{q+2,n-p}(c_\alpha; \Delta) \\ &+ \Delta(2G_{q+2,n-p}(c_\alpha; \Delta) - G_{q+4,n-p}(c_\alpha; \Delta)) \quad \text{if} \quad \Delta \geq q. \end{aligned} \tag{9}$$

We find

$$\begin{aligned} \sup_{0<\Delta<q} Reg(\delta, c_\alpha) &= Reg(\Delta_L, c_\alpha), \\ \sup_{\Delta \geq q} Reg(\delta, c_\alpha) &= Reg(\Delta_U, c_\alpha). \end{aligned}$$

Then solve for c_α^* for which

$$Reg(\Delta_L, c_\alpha^*) = Reg(\Delta_U, c_\alpha^*), \tag{10}$$

where Δ_L and Δ_U are the values of Δ for which $Reg(\Delta, c_\alpha)$ is the maximum for $\Delta < q$ and $Reg(\Delta, c_\alpha)$ is the maximum for $\Delta \geq q$, respectively. The relative efficiency corresponding to the optimal c_α^* can be determined from Section 3.1. The optimal critical value c_α^* for different numerator and denominator degrees of freedoms have been tabulated by Brook (1976). We define the PTLSE based on optimal critical value c_α^* as

$$\hat{\beta}^{PT}_{Brook} = \hat{\beta}^{UE} - (\hat{\beta}^{UE} - \hat{\beta}^{RE}) I(\mathcal{L} < c_\alpha^*). \tag{11}$$

3.3 Fixed Critical Value

As we have seen both Han and Bancroft and Brook methods are computer intensive to obtain an optimal size of the test. Now we define the PT estimator based on fixed critical value, which is easy to calculate. Note that the test statistic

$$\mathcal{L} = \frac{(H\hat{\beta}^{UE} - h)'(HC^{-1}H')^{-1}(H\hat{\beta}^{UE} - h)}{qs^2}$$

has a central F distribution with q and $n - p$ degrees of freedoms under the null hypothesis. The mean of the F distribution is $\mu_F = E(F) = \frac{n-p}{n-p-2}$. We would like to use this mean as a fixed critical value of the test statistic. If we consider $\mu_F = \frac{n-p}{n-p-2}$ as the critical value of the test statistic (2) we obtain conservative size of the test (see Table 1) that works well compared to critical values suggested by Han and Bancroft

Table 1: Probability of F for quantiles μ_F and for different degrees of freedom

n	p	q									
		1	2	3	4	5	8	10	15	20	30
10	2	0.282	0.316	0.330	0.337	0.341	0.347	0.349	0.351	0.351	0.352
	3	0.275	0.308	0.320	0.327	0.330	0.335	0.337	0.338	0.339	0.340
	4	0.267	0.296	0.307	0.313	0.316	0.320	0.321	0.322	0.323	0.323
	5	0.253	0.279	0.288	0.292	0.294	0.297	0.298	0.299	0.300	0.300
	6	0.230	0.250	0.256	0.259	0.261	0.263	0.263	0.264	0.264	0.264
	7	0.182	0.192	0.196	0.197	0.197	0.198	0.198	0.199	0.199	0.199
20	2	0.303	0.346	0.365	0.376	0.383	0.393	0.397	0.401	0.403	0.405
	3	0.302	0.345	0.364	0.374	0.381	0.391	0.394	0.399	0.401	0.403
	4	0.301	0.344	0.362	0.372	0.378	0.388	0.392	0.396	0.398	0.399
	5	0.300	0.342	0.360	0.370	0.376	0.385	0.389	0.393	0.394	0.396
	8	0.295	0.335	0.352	0.360	0.366	0.374	0.377	0.380	0.382	0.383
	10	0.290	0.328	0.343	0.351	0.356	0.363	0.366	0.368	0.369	0.370
	12	0.282	0.316	0.330	0.337	0.341	0.347	0.349	0.351	0.351	0.352
	15	0.253	0.279	0.288	0.292	0.294	0.297	0.298	0.299	0.300	0.300
30	2	0.308	0.354	0.375	0.387	0.394	0.407	0.411	0.417	0.421	0.423
	3	0.308	0.354	0.374	0.386	0.394	0.406	0.410	0.416	0.419	0.422
	5	0.307	0.353	0.373	0.384	0.392	0.404	0.408	0.414	0.417	0.419
	8	0.306	0.350	0.370	0.381	0.389	0.400	0.404	0.409	0.412	0.414
	10	0.304	0.349	0.368	0.379	0.386	0.397	0.401	0.406	0.408	0.410
	15	0.300	0.342	0.360	0.370	0.376	0.385	0.389	0.393	0.394	0.396
	20	0.290	0.328	0.343	0.351	0.356	0.363	0.366	0.368	0.369	0.370
	25	0.253	0.279	0.288	0.292	0.294	0.297	0.298	0.299	0.300	0.300
40	2	0.311	0.358	0.379	0.392	0.400	0.414	0.419	0.426	0.429	0.433
	3	0.311	0.358	0.379	0.391	0.400	0.413	0.418	0.425	0.429	0.432
	5	0.310	0.357	0.378	0.391	0.399	0.412	0.417	0.424	0.427	0.430
	8	0.309	0.356	0.377	0.389	0.397	0.410	0.415	0.421	0.425	0.428
	10	0.309	0.355	0.376	0.388	0.396	0.409	0.413	0.420	0.423	0.426
	15	0.307	0.353	0.373	0.384	0.392	0.404	0.408	0.414	0.417	0.419
	20	0.304	0.349	0.368	0.379	0.386	0.397	0.401	0.406	0.408	0.410
	25	0.300	0.342	0.360	0.370	0.376	0.385	0.389	0.393	0.394	0.396
	30	0.290	0.328	0.343	0.351	0.356	0.363	0.366	0.368	0.369	0.370
	35	0.253	0.279	0.288	0.292	0.294	0.297	0.298	0.299	0.300	0.300
100	1	0.315	0.364	0.387	0.401	0.410	0.426	0.432	0.441	0.446	0.452
	2	0.315	0.364	0.387	0.400	0.410	0.426	0.432	0.441	0.446	0.451
	5	0.315	0.364	0.387	0.400	0.409	0.425	0.431	0.441	0.446	0.451
	10	0.315	0.364	0.386	0.400	0.409	0.425	0.431	0.440	0.445	0.450
	20	0.314	0.363	0.386	0.399	0.408	0.424	0.430	0.439	0.443	0.449
	30	0.314	0.363	0.385	0.398	0.407	0.423	0.428	0.437	0.441	0.446
	40	0.313	0.362	0.384	0.397	0.406	0.421	0.426	0.435	0.439	0.444
	50	0.312	0.360	0.382	0.395	0.404	0.418	0.424	0.431	0.436	0.440

Table 2: Maximum & Minimum Guaranteed Efficiency of $PTLSEs$

n	p	q		α								Brook	Fixed
				2.5%	5%	10%	15%	20%	30%	40%	50%	23.8%	23.0%
6	2	1	E_{Max}	1.81	1.68	1.49	1.37	1.28	1.16	1.09	1.05	1.23	1.24
			E_{Min}	0.29	0.40	0.54	0.64	0.72	0.82	0.89	0.94	0.76	0.75
			Δ_{Min}	12.69	8.96	6.28	5.11	4.45	3.64	3.22	2.95	4.09	4.15
				2.5%	5%	10%	15%	20%	30%	40%	50%	20.5%	28.2%
10	2	1	E_{Max}	1.77	1.63	1.44	1.33	1.25	1.14	1.08	1.04	1.24	1.16
			E_{Min}	0.39	0.49	0.61	0.70	0.76	0.85	0.91	0.95	0.76	0.84
			Δ_{Min}	8.03	6.31	4.93	4.24	3.82	3.31	3.01	2.83	3.79	3.40
				2.5%	5%	10%	15%	20%	30%	40%	50%	18.8%	30.1%
18	2	1	E_{Max}	1.74	1.60	1.42	1.31	1.23	1.13	1.07	1.04	1.25	1.13
			E_{Min}	0.45	0.54	0.65	0.72	0.78	0.86	0.92	0.95	0.77	0.86
			Δ_{Min}	6.55	5.41	4.42	3.91	3.58	3.16	2.92	2.77	3.64	3.16
				2.5%	5%	10%	15%	20%	30%	40%	50%	20.4%	31.6%
12	4	2	E_{Max}	1.84	1.73	1.57	1.45	1.37	1.24	1.16	1.10	1.36	1.23
			E_{Min}	0.48	0.58	0.69	0.76	0.81	0.87	0.92	0.95	0.81	0.88
			Δ_{Min}	11.69	9.35	7.36	6.37	5.71	4.90	4.36	4.00	5.68	4.81
				2.5%	5%	10%	15%	20%	30%	40%	50%	17.9%	34.4%
20	4	2	E_{Max}	1.82	1.70	1.54	1.43	1.34	1.22	1.15	1.09	1.37	1.18
			E_{Min}	0.56	0.64	0.73	0.79	0.83	0.89	0.93	0.95	0.82	0.91
			Δ_{Min}	9.05	7.61	6.31	5.59	5.14	4.51	4.12	3.82	5.32	4.33
				2.5%	5%	10%	15%	20%	30%	40%	50%	16.8%	35.2%
28	4	2	E_{Max}	1.81	1.69	1.53	1.42	1.33	1.22	1.14	1.09	1.38	1.17
			E_{Min}	0.58	0.66	0.74	0.80	0.84	0.89	0.93	0.96	0.82	0.92
			Δ_{Min}	8.33	7.12	5.98	5.35	4.96	4.39	4.03	3.76	5.20	4.18
				2.5%	5%	10%	15%	20%	30%	40%	50%	15.8%	36.2%
64	4	2	E_{Max}	1.80	1.68	1.51	1.40	1.32	1.21	1.14	1.09	1.39	1.16
			E_{Min}	0.61	0.68	0.76	0.81	0.85	0.90	0.93	0.96	0.82	0.92
			Δ_{Min}	7.58	6.58	5.62	5.11	4.75	4.24	3.91	3.67	5.02	4.03
				2.5%	5%	10%	15%	20%	30%	40%	50%	13.6%	38.3%
30	6	4	E_{Max}	2.62	2.38	2.05	1.83	1.67	1.45	1.30	1.20	1.88	1.32
			E_{Min}	0.63	0.70	0.78	0.83	0.86	0.91	0.94	0.96	0.82	0.94
			Δ_{Min}	11.93	10.37	8.93	8.12	7.55	6.76	6.22	5.77	8.30	6.28
				2.5%	5%	10%	15%	20%	30%	40%	50%	11.7%	39.7%
66	6	4	E_{Max}	2.59	2.34	2.01	1.80	1.64	1.43	1.29	1.19	1.93	1.30
			E_{Min}	0.67	0.74	0.81	0.85	0.88	0.92	0.95	0.96	0.82	0.94
			Δ_{Min}	10.52	9.35	8.21	7.55	7.06	6.40	5.95	5.56	7.94	5.95
				2.5%	5%	10%	15%	20%	30%	40%	50%	11.0%	40.1%
126	6	4	E_{Max}	2.58	2.33	2.00	1.79	1.63	1.42	1.29	1.19	1.95	1.28
			E_{Min}	0.69	0.75	0.81	0.85	0.88	0.92	0.95	0.97	0.82	0.95
			Δ_{Min}	10.10	9.02	7.97	7.33	6.91	6.31	5.86	5.50	7.82	5.86

(1968) and Brook (1976). The size of the test for different degrees of freedom and quantiles are presented in Table 1.

We propose the PTLSE based on critical value μ_F as

$$\hat{\beta}^{PT}_{Fixed} = \hat{\beta}^{UE} - (\hat{\beta}^{UE} - \hat{\beta}^{RE})I(\mathcal{L} < \mu_F). \quad (12)$$

The proposed estimator is to change the way that researchers undertake the preliminary test.

Table 2 compares $\hat{\beta}^{PT}_{\alpha}$, $\hat{\beta}^{PT}_{Brook}$ and $\hat{\beta}^{PT}_{Fixed}$ under the quadratic risk function and minimum and maximum guaranteed efficiency criteria. The last two columns of the table provide the maximum and minimum guaranteed relative efficiencies for optimal critical values provided by Brook (1976) and fixed critical value, respectively. For given q and $n-p$, one enters the table and looks for the smallest relative efficiency E_{Min} he/she wishes to accept. For example, suppose $q = 2$, $n-p = 12$ and the experimenter wishes to have an estimator with a minimum guaranteed efficiency of 0.75. From the table, we recommend him/her to select $\alpha = 0.15$, corresponding to $\hat{\beta}^{PT}$, because such a choice of α would yield an estimator with a minimum efficiency of 0.76 and a maximum efficiency of 1.45. Note that with this condition the minimum guaranteed efficiency of $\hat{\beta}^{PT}$ using Brook's optimal critical value is 0.81 with a maximum efficiency of 1.36. By fixed critical value, the minimum guaranteed efficiency is 0.88 with a maximum efficiency of 1.23.

4 Summary

In this paper, we have compared the methods of Han Bancroft (1968) and Brook (1976) along with a proposed fixed critical value for obtaining an optimal significance level to formulate a PTLSE. To determine the Han Bancroft's level one has to specify a value which is the smallest relative efficiency the investigator is willing to accept. However, the Brook's level balances the loss and gain to determine a level based on a regret function. Since, a theoretical comparison among these three methods is hard to make, a numerical comparison has been performed.

From Table 2, it is observed that the minimum guaranteed efficiency by Brook's method vary between 0.76 and 0.82 and the maximum efficiency vary between 1.23 and 195.0 for α $(0.11 < \alpha < 0.24)$. The minimum guaranteed efficiency by Fixed critical method vary between 0.75 and 0.95 and the maximum efficiency vary between 1.24 and 1.32 for α $(0.23 < \alpha < 0.40)$. The corresponding minimum and maximum guaranteed efficiencies by Han Bancroft method's are between 0.29 and 0.97 and 1.05 and 2.59, respectively, for $0.025 < \alpha < 0.50$. Both Fixed and Brook methods are conservative for fixed sample size, whereas Han Bancroft method is flexible because a higher minimum guaranteed efficiency can be chosen to determine the significance level. In either method, the researchers have to take some risk. If the researchers are concerned or very conservative about the minimum guaranteed efficiency, they might select Fixed or Brook's method. However, if they are willing to accept higher size of the test but want to have higher guaranteed minimum efficiency, they should select Han Bancroft's method. The proposed method is easy to compute compared to Han and Bancroft or Brook's method.

Acknowledgments

The authors are thankful to the referee whose various suggestions have improved the presentation of the paper.

References

Bancroft, T. A. (1944). On biases in estimation due to use of preliminary tests of significance. *Annals of Mathematics and Statistics*, 15, 190-204.

Bancroft, T. A. (1964). Analysis and inference for incompletely specified models involving the use of preliminary test(s) of significance. *Biometrics*, 20, 427-442.

Benda, N. (1996). Pre-test estimation and design in the linear model. *Journal of Statistical Planning and Inference*, 52, 225-240.

Brook, R. J. (1976). On the use of a regret function to set significance points in prior tests of estimation. *Journal of the American Statistical Association*, 71, 126-131.

Chiou, P. and Han, C-P. (1999). Conditional interval estimation of the ratio of various components following rejection of a pre-test. *Journal of Statistical Computation and Simulation*, 63, 105-119.

Han, C-P. (2002). Influential observations in a preliminary test estimation of the mean. *Pakistan Journal of Statistics*, 18, 321-333.

Han, C-P, and Bancroft, T. A. (1968). On pooling means when variance is unknown. *Journal of the American Statistical Association*, 63, 1333-1342.

Han, C-P, Rao, C. V. and Ravichandran, J. (1988). Inference based on conditional specification: a second bibliography. *Communications in Statistics-Theory and Methods*, 17, 1945-1964.

Gilies, J. A. and Gilies, D. E. A. (1993). Pre-test estimation in econometrics: recent developments. *Journal of Economic Survey*, 7, 145-197.

Judge, G.G. and Bock, M.E. (1978). *The Statistical Implications of Pre-test and Stein-rule Estimators in Econometrics*, North-Holland Publishing Company, Amsterdam.

Kibria, B. M. G. and Saleh, A. K. Md. E. (2003). Estimation of the mean vector of a multivariate normal distribution under various test statistics. *Journal of Probability and Statistical Science*, 1, 141-155,

Kibria, B. M. G. and Saleh, A. K. Md. E. (2005). Comparison between Han Bancroft and Brook methods to determine the optimal significance level for pre-test estimator. *Journal of Probability and Statistical Science*, 3, 293-303.

Kibria, B. M. G. and Saleh, A. K. Md. E. (2006). Optimum critical value for pre-test estimators. *Communications in Statistics-Simulation and Computation*, 35 (2), 309-319.