Vol.11.Issue.2.2023 (April-June)

BULLETIN OF MATHEMATICS AND STATISTICS RESEARCH

A Peer Reviewed International Research Journal

http://www.bomsr.com

Email:editorbomsr@gmail.com

INTERNATIONAL STANDARD SERIAL NUMBER 2348-0580

RESEARCH ARTICLE

The Validity of t-test, Mann-Whitney test and Z test for Testing Significant differences between two Sample Means When Sample size is 10 or below

Ramnath Takiar
Apartment no. 11, Building 9, 3 rd Floor,
Mangol Expoprt Town, 1st Khorro, Olympic Street
Sukhbaatar District, Ulaanbaatar, Mongolia.
Email: ramnathtakiar@gmail.com, ramnath_takiar@yahoo.co.in
&
Scientist G – (Retired)
National Centre for Disease Informatics and Research (NCDIR),
Indian Council of Medical Research (1978-2013)
Bangalore – 562110, Karnataka, India
DOI:10.33329/bomsr.11.2.1

Ramnath Takiar

ABSTRACT

In the present study, an attempt is made to evaluate the performance of t-test, Mann-Whitney test as compared to Z-test in testing the possible significant or non-significant differences between two sample means or between a sample mean and the population mean. The results presented are based on the 500 samples of size 10, 6 and 3 from predefined four Normal populations. Overall, the study covers 18000 comparisons between sample means and the respective population mean. It also covers an equal number of comparisons for testing the possible significant differences between two sample means by three selected significance tests.

It is surprising to note that for the samples of size 10, at α = 5%, t-test can pick up only 31.1% of the expected significant differences between two sample means which decreased to 11.4% for the sample size 3. This suggests that t-test is not valid when the sample size is 10 or below. In comparison, at α = 5%, for the sample size of 10, Mann-Whitney test showed the validity of 30.4% while Z test with estimated variance (Z-EV) showed the validity of

39.9%. At Sample size 3, the validity of Mann-Whitney test and Z-EV test is observed to be less and is 20.1% and 31.5%, respectively.

In view of very low validity observed, it is concluded that neither t-test nor Mann-Whitney test is suitable to be used when the sample size is 10 or below. In view of higher negative validity seen for Z-EV test as compared to t-test in the present study, as well as in my previous study for the sample size 9, 13 and 20, it is recommended that for sample size above 10 and below 30, Z-EV test can be used in the place of t-test, preferably with α = 10%.

KEY WORDS: t-test, Z-test, Mann-Whitney test, Normal samples, Validity, 10% α level.

INTRODUCTION

In most of the research studies, the interest usually lies in comparing the two sample means. If the sample is large enough, say, more then 30, than Z test is advocated. In case of small samples, below 30, t-test is recommended. The assumptions in application of both the tests are that i) The samples are independent and drawn from the Normal populations; ii) Both the sample populations, under considerations are having comparable variances (Snedecor and Cochran 1967, Gupta and Kapoor 2001, Gupta 2012).

In a recent study (Takiar 2021), it was shown that Z-test is better in picking up the valid significant differences as compared to t-test even when the sample size is small like 20,13 and 9. However, it is interesting to see what happened to Z-test and Mann-whiney test as compared to t-test when the sample size is chosen to be below 10. Hence, in the present study, an attempt is made to evaluate the validity of t-test as compared to Z-test and Mann-Whitney test when the sample size is below 10?

OBJECTIVES

In the present study, an attempt is made

1) To explore the validity of t-test when the sample size is 10 or below 10?
2) In comparison to t-test, how do the Z-test and Mann-Whitney U test perform when the sample size is 10 or below 10? and
3) To evaluate and compare the effect of decreasing sample size on the validity of the selected Significant tests?

MATERIAL AND METHODS

GENERATION OF NORMAL POPULATIONS

For the generation of a normal population, the function key, "Random Number Generation" provided in StatPlus 7.6.5 is used. This function key allows you to generate desired number of Normal samples of chosen size with the specified mean and SD. Using the same key, four sets of 100 normal samples of size 200 are generated with a different pre-specified mean and SD. From each set, the most suitable sample with the skewness close to zero and the kurtosis close to 3.0 is selected. The four samples of size 200 so chosen are named as P1, P2, P3 and P4. For the study purposes, they are termed as the Normal Populations of size 200. It is ensured that when we compare the distribution of P1 with

P2 on one hand and P3 and P4 on the other hand, the distributions are significantly different from each other. The major parameters of above four populations are provided in Table 1.

Table 1: Description of Parameters of Selected Populations with the result of Significance test

Parameter	NORMAL POPULATION			
	P1	P2	P3	P4
N	200	200	200	200
Mean	55.5	44.2	65.8	76.1
VAR-P	256.4	136.82	281.78	322.06
SD	16.01	11.7	16.79	17.94
Skewness	0.02	-0.11	0.06	0.002
Kurtosis	2.90	3.00	2.97	3.03
Z Value	8.03		5.95	
P-value	**< 0.001**		**< 0.001**	
Critical t value at 0.001	3.09		3.09	

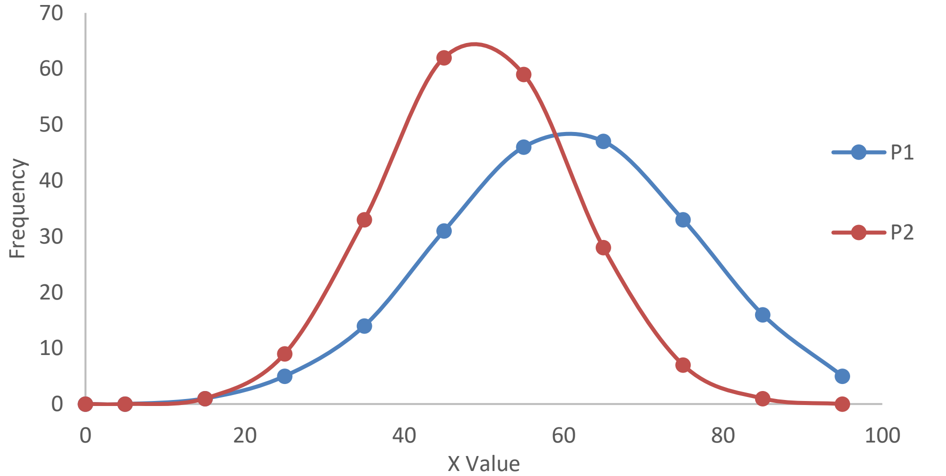
SMOOTHEN FREQUENCY DISTRIBUTION OF SELECTED POPULATIONS

To each population, normal distribution was fitted. The fitted distributions are found to be good. The smoothen frequency distribution of Population P1, P2 and P3, P4 are shown in Fig.1 and Fig. 2.

SAMPLE SELECTION

The Scheme of sample selection by different population and size is shown in Table 2.

Fig.1: Smoothen Frequency distribution of Population P1 and P2



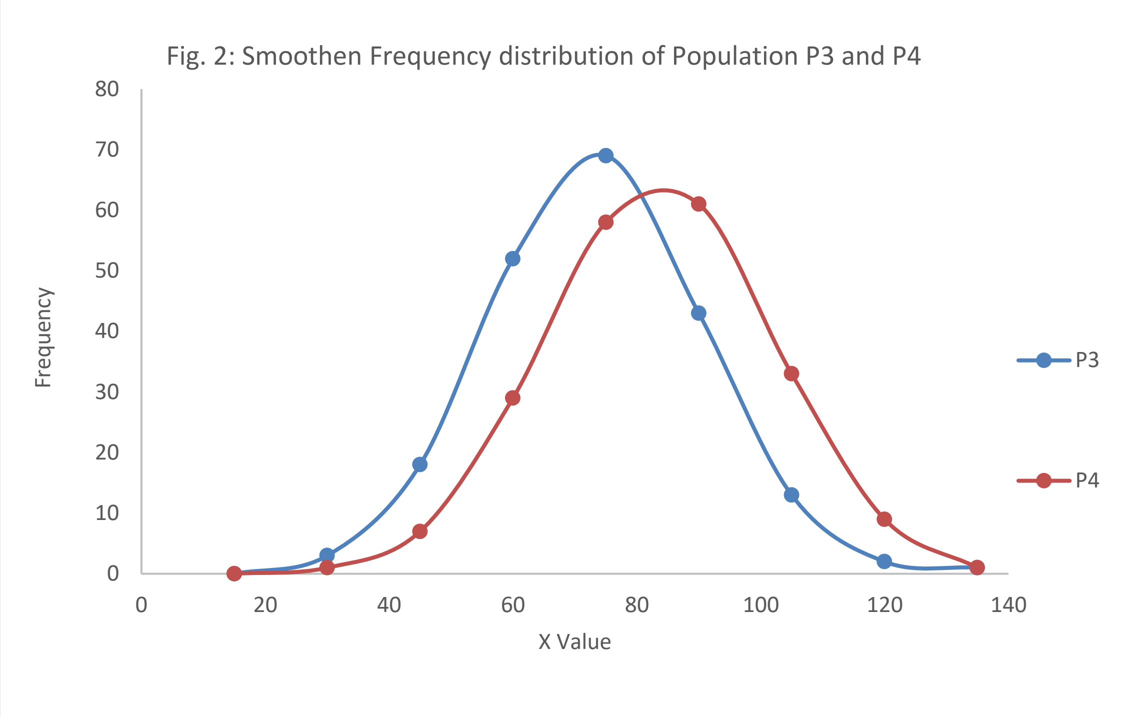


Table 2: Scheme of Sample Selection according to Population, Sample size and Number of Samples drawn

Population	Sample Size	Number of Samples Drawn
P1	10	500
	6	500
	3	500
P2	10	500
	6	500
	3	500
P3	10	500
	6	500
	3	500
P4	10	500
	6	500
	3	500
Total	Pooled	6000

As shown in Table 2, from each population, 500 Random samples of size 10, 6 and 3 are generated. Each sample mean allows us to be compared with the respective Population Mean, resulting in 500 t-tests and an equal number of Z-tests with Known Variance (Z-KV) and Z-test with Estimated Variance (Z-EV) for a given sample size. Thus, for all the four Populations (P1, P2, P3 and P4), in total, 6000 t-tests, Z-KV test, Z-EV test are carried out. This amounts to 18000 comparison of sample means with that of Population mean.

The formula used are t = $\frac{\bar{x}-\mu}{\frac{S}{\sqrt{n}}}$; Z = $\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$

It is to be noted that for t-statistic, variance estimate is done with (n-1) as the denominator (Snedecor and Cochran 1967, Gupta and Kapoor 2001, Gupta 2012) while for Z statistic, σ is used since variance is already known. Z test is also carried out with the sample estimate of variance with "n" as the denominator.

SCHEME OF COMPARISONS BETWEEN THE MEANS OF TWO DIFFERENT POPULATIONS

The number of possible mean comparisons according to each sample size and Population is shown in Table 3.

Table 3: Number of Mean Comparisons according to Each Sample size and the Population

Population P1	Population P2	Number of Mean comparisons	Population P3	Population P4	Number of Mean comparisons
Sample size	Sample size		Sample size	Sample size	
10	10 A	500	10	10 A	500
	10 B	500		10 B	500
6	6 A	500	6	6 A	500
	6 B	500		6 B	500
3	3 A	500	3	3 A	500
	3 B	500		3 B	500
Total		3000	Total		3000

A – Refers to original sequence of Samples**; B –** refers to changed sequence of samples

It is to be noted that 500 samples are drawn from each population. However, for a given sample size, for mean comparisons, sample means are compared between the populations of P1 and P2 on one hand and P3 and P4 on the other hand. By changing the sequence of samples in P2 and treating this as another set, 500 more comparisons are made between P1 and P2. Thus, in total, 1000 mean comparisons are attempted between P1 and P2. Proceeding in a similar way, again a total of 1000 mean comparisons are attempted between P3 and P4.

The formulae used are as follows: t = $\frac{\bar{x}_1-\bar{x}_2}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$; Z -EV= $\frac{\bar{x}_1-\bar{x}_2}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$

Where $S^2 = \frac{1}{(n-1)}\Sigma(x_i-\bar{x})^2$ $\quad s^2 = \frac{1}{n}\Sigma(x_i-\bar{x})^2$

SIGNIFICANCE TEST FOR TWO SAMPLE MEAN COMPARISONS

In the present study, the comparison between two sample means is carried out using four tests namely

1. t-test,

2. Z-test with known variance (For comparing sample means with that of population mean)

3. Z-test with estimated sample variance,

4. Mann-Whitney test

SIGNIFICANCE AND α LEVEL

In my earlier study (Takiar, 2021), the validity of t-test (< 50%) and Z-test (55.0%) was shown to be very low. Further with the rise in α levels, an improvement in the validity was seen. Accordingly, for the present study purposes, three α levels are chosen namely 1) 5% 2) 10% and 3) 15%.

VALIDITY OF T-TEST AND Z-TEST

WHEN SAMPLES ARE DRAWN FROM THE SAME NORMAL POPULATION

The 500 samples of a given size (10, 6, 3) drawn from a given population (P1, P2, P3, P4) allow us to attempt 500 tests to see whether the selected sample is drawn from a pre-specified mean or not? In this case, the Null Hypothesis is that "The sample mean is not significantly different from the population mean." Since, the sample drawn is from the same normal population, it is logical not to reject the Null Hypothesis. Thus, A higher percentage of non-significant means obtained from 500 samples, will suggest that the t-test or Z-test is successful in picking up the desired non-significant differences of the sample means, correctly. A low percentage will indicate the lower validity of the t-test or that of Z-test. This type of validity for the study purposes is termed as the "Positive validity" of the test.

WHEN SAMPLES ARE DRAWN FROM TWO DIFFERENT NORMAL POPULATIONS

For the present study, we intend to compare 500 sample means of P1 with that of P2A, P2B, on one hand and sample means of P3 with that of P4A, P4B, on other hand. The chosen scheme of comparison allows us 1000 comparisons between two populations for each sample size (10, 6, 3). In this case, the formulated Null Hypothesis "The sample means are not significantly different from each other" must be rejected and the alternative Hypothesis that the sample means are significantly different from each other has to be accepted. A higher percentage of significant differences between the means, based on the predefined cut-off level, say above 80%, suggests the existence of the higher validity of the t-test or Z-test while a lower percentage than 80% will suggests that the validity of the test is lower. This type of validity for the study purposes is termed as the "Negative validity" of the test.

ANALYSIS OF THE DATA

For simultaneous comparisons of means of all the 500 samples, SPSS program, 2023 version, is utilized. For Z-test comparisons, with known variance and estimated variance, a program developed on Excel 2019 is utilized. The function keys available on Excel 2019, are utilized to arrive at the probability of the Z-statistic and t-statistic, calculated. The results obtained by t-test and Z-test, at the given α level, are also compared to see which test is better in picking up either the significant differences or the non-significant differences.

TYPICAL RANDOM SAMPLES

In order to given an idea to the readers, one typical random sample, of size 3, 6, 10 each, drawn from each of the population P1, P2, P3, P4, are shown in Table 4, 5, and 6.

Table 4: Typical one Random Sample from Each Population of Size 3

Number	P1	P2	P3	P4
1	71.11	37.38	70.00	102.80
2	62.95	46.82	84.00	81.20
3	59.42	71.31	68.10	57.20
MEAN	64.50	51.80	74.00	80.40
SD- t	6.00	17.51	8.68	22.81
SD-Z	4.90	14.30	7.09	18.62

Table 5: Typical one Random Sample from Each Population of Size 6

Number	P1	P2	P3	P4
1	71.11	41.11	62.80	72.70
2	62.95	65.70	65.70	72.20
3	59.42	47.75	66.60	74.10
4	60.04	54.40	99.60	60.90
5	46.72	61.23	63.70	60.70
6	60.93	71.59	67.90	74.00
MEAN	60.20	57.00	71.10	69.10
SD- t	7.87	11.41	14.11	6.47
SD-Z	7.18	10.42	12.88	5.91

Table 6: Typical one Random Sample from Each Population of Size 10

Number	P1	P2	P3	P4
1	71.90	49.59	51.80	77.10
2	64.27	9.17	39.10	34.90
3	51.29	39.39	84.40	92.00
4	40.53	28.43	38.20	62.60
5	27.05	47.81	66.10	52.40
6	49.50	39.46	37.30	76.60
7	76.57	48.52	76.70	27.80
8	59.50	34.16	63.70	82.40
9	59.58	41.11	49.20	87.00
10	58.56	47.78	54.50	65.90
MEAN	55.9	38.5	56.1	65.9
SD- t	14.6	12.41	16.38	21.72
SD-Z	13.85	11.77	15.54	20.6

From the data presented in above tables, it is evident that the SD fluctuate relatively more when the sample size is 3. In general, CV is around 30%.

COMPARISON OF SAMPLE MEANS WITH THAT OF POPULATION MEAN BY t-test AND Z-TEST WITH KNOWN VARIANCE AND WITH ESTIMATED VARIANCE

The Results of t-test and Z-test, for testing the expected non-significant differences between sample means and the population mean when drawn from the same population (P1, P2), by varying α levels are shown in Table 7.

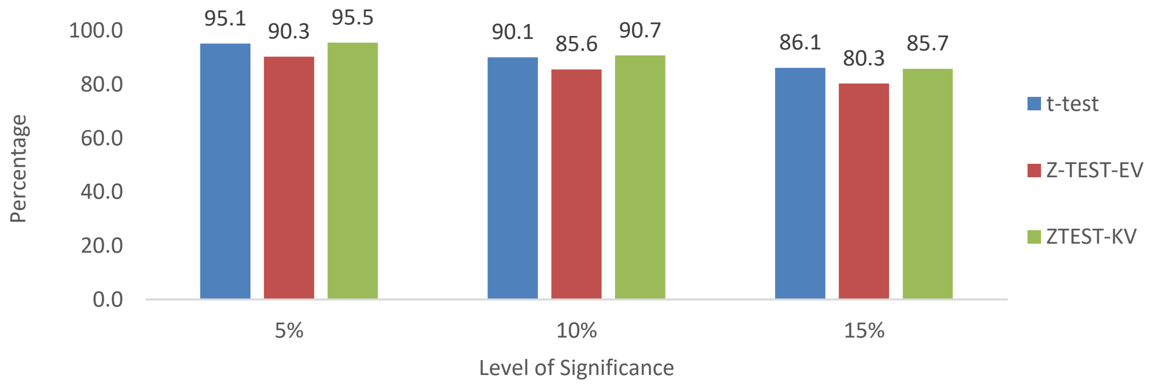
Table 7: The % of Sample means correctly classified as non-significant with the Population Mean by different test and different Alpha levels – Pooled for Population P1 and P2 At Sample size of 10

Significance test	Sample size	Number of Comparisons	5%	10%	15%
t-test	10	1000	947	903	867
		%	**94.7**	**90.3**	**86.7**
Z-EV TEST (With Estimated Variance)	10	1000	908	862	813
		%	**90.8**	**86.2**	**81.3**
Z-KV TEST (With known Variance)	10	1000	957	915	863
		%	**95.7**	**91.5**	**86.3**

At 5% α level, with the sample size 10, the t-test showed 94.7% accuracy in picking up the expected non-significant differences between sample means and the population mean, while by the traditional Z-test with known variance (Z-KV) the accuracy was 95.7%, a bit more than the t-test. With the rise in α level, the accuracy comes down to around 86% for both t-test and Z-test. However, Z-test with estimated variance (Z-EV), irrespective of α levels, showed around 5% less accuracy in picking up the non-significant differences between sample means and the population mean. The rise in α levels from 5% to 15%, resulted in 10% fall in accuracy in picking up correctly the non-significant differences. Similar tabulations are done for the Population P3 and P4. In order to keep the number of tables in optimum limit, hereafter, it is thought logical to present the results, pooled for P1, P2, P3 and P4 in a single table, instead of presenting the results in 4 tables, every time.

The % of non-significant means correctly shown comparable to population mean, by the α levels and the Test of Significance, pooled for the Sample size 10, is shown in Fig. 3.

Fig. 3: The % of Non-Significant Sample Means, correctly classified by the Alpha Level and the Test of Significance - Pooled for P1,P2, P3,P4 - Sample size 10

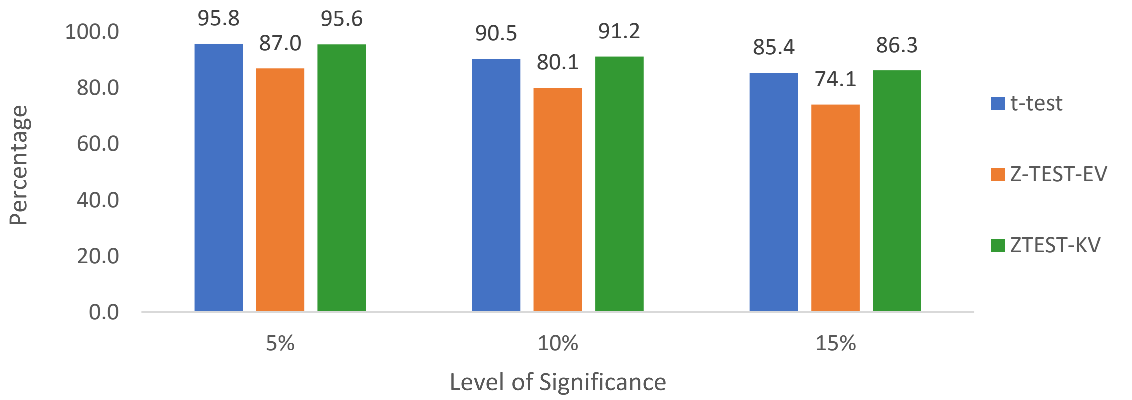


At 5% α level, with the sample size 10, the t-test showed 95.1% accuracy in picking up the expected non-significant differences between sample means and the population mean. For traditional Z-test with known variance (Z-KV), the accuracy is 95.5%, bit more than the t-test. With the rise in α level to 15%, the accuracy comes down to around 86% for both the t-test and Z-test. However, Z-test with estimated variance (Z-EV), irrespective of α levels chosen, showed around 5% less accuracy in picking up the non-significant differences between sample means and the population mean. The change in α level from 5% to 15%, resulted in around 10% fall in the accuracy of picking up correctly the non-significant differences.

The % of Non-Significant means correctly shown comparable to population mean, by the α levels and the Test of Significance, Pooled for P1, P2, P3, P4 for the Sample size 6, is shown in Fig. 4.

At 5% α level, with the sample size 6, the t-test showed 95.8% accuracy in picking up the expected non-significant differences between sample means and the population mean. For traditional Z-test (Z-KV), the accuracy is 95.6%, almost similar to what seen in the case of t-test. With the rise in α level to 15%, the accuracy comes down to around 86% for both t-test and Z-test. However, Z-test (Z-EV) exhibited around 10% fall in accuracy of picking up correctly the non-significant differences, like seen in the case of sample size 10.

Fig.4: The % of Non-Significant Sample Means, correctly classified by the Alpha Level and the Test of Significance - Pooled for P1,P2, P3,P4 - Sample size 6



The % of Non-Significant means correctly shown comparable to population mean, by the α levels and the Test of Significance, Pooled for P1, P2, P3, P4 for the Sample size 3, is shown in Fig. 5.

At 5% α level, with the sample size 3, the t-test showed 96.3% accuracy in picking up the expected non-significant differences between sample means and the population mean, By the traditional Z-test (Z-KV) the accuracy is 95.4%, almost like what seen in the case of t-test. With the rise in α level to 15%, for t-test and Z-test, the accuracy comes down to around 86.8% and 84.1%, respectively. However, Z-test (Z-EV), resulted in fall of around 13% in accuracy.

COMPARISON OF TWO SAMPLE MEANS BY t-test, MANN-WHITNEY TEST AND Z-TEST WITH ESTIMATED VARIANCE

The results of Significance tests for comparison between the sample means of populations P1 and P2 of Sample size 10 is shown in Table 8.

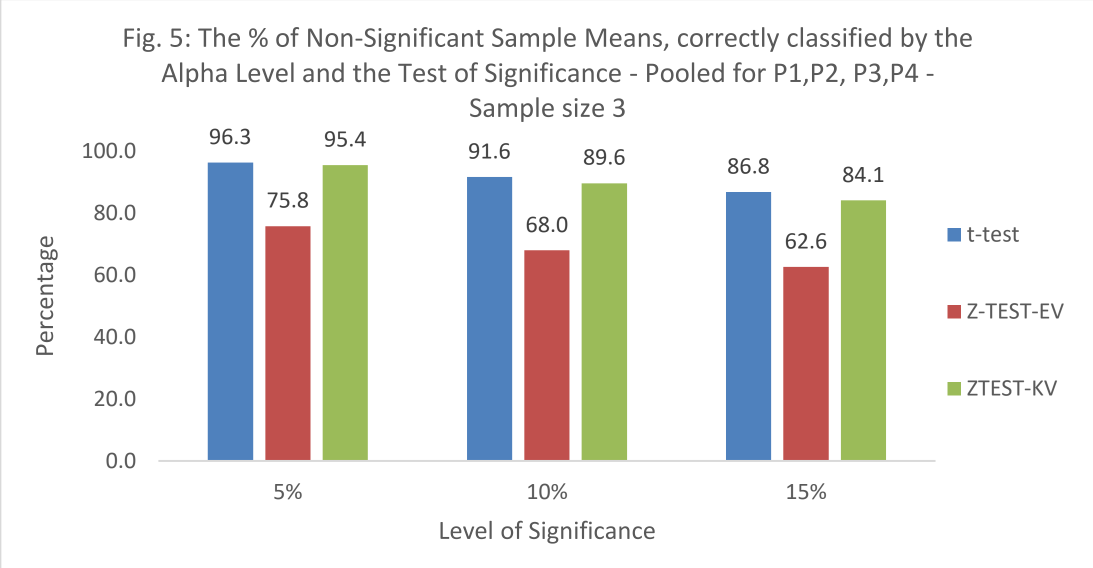
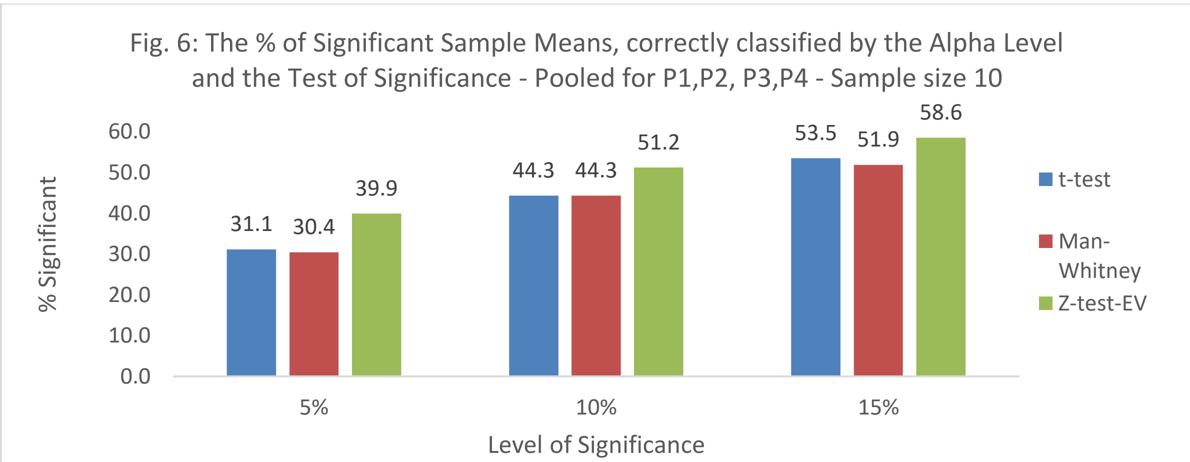


Table 8: The % of Significant Sample Means, correctly classified by the Alpha Level and the Test of Significance For Population P1 and P2 - Sample size 10

Significance test	Sample size	Number of Comparisons	Significant differences at a Level		
			5%	10%	15%
t-test	10	1000	375	521	627
		%	**37.5**	**52.1**	**62.7**
Mann-Whitney	10	1000	367	514	599
		%	**36.7**	**51.4**	**59.9**
Z test -EV	10	1000	476	600	683
		%	**47.6**	**60.0**	**68.3**

At the sample size of 10 and at 5% α level, t-test and Mann-Whitney test picked up around 37% of the expected significant mean differences, correctly. In comparison, Z-test picked up 10% more i.e., 47.6% of the expected mean differences, correctly. At 15% α level, Z-test picked up around 68%, almost 20% more as compared to that seen at 5% α level. In case of t-test and Mann-Whitney test, the corresponding rise was 25.2% and 23.2%, respectively.

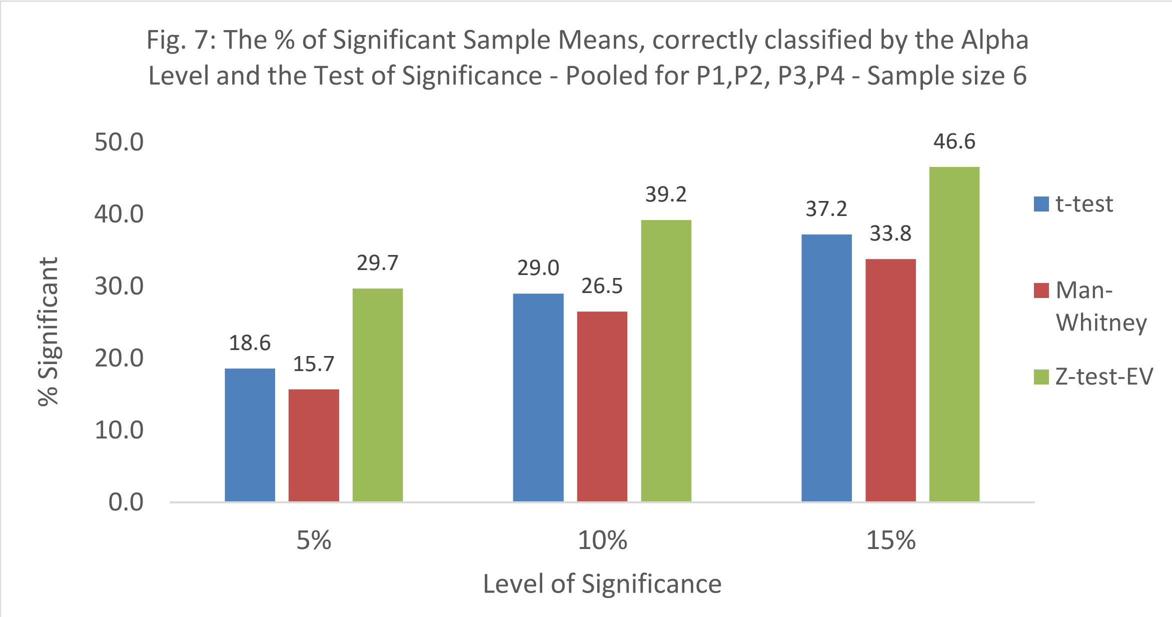
At sample size equal to 10, the results of significant differences in sample means found by the selected three tests when pooled for P1P2 and P3P4 are shown in Fig. 6.



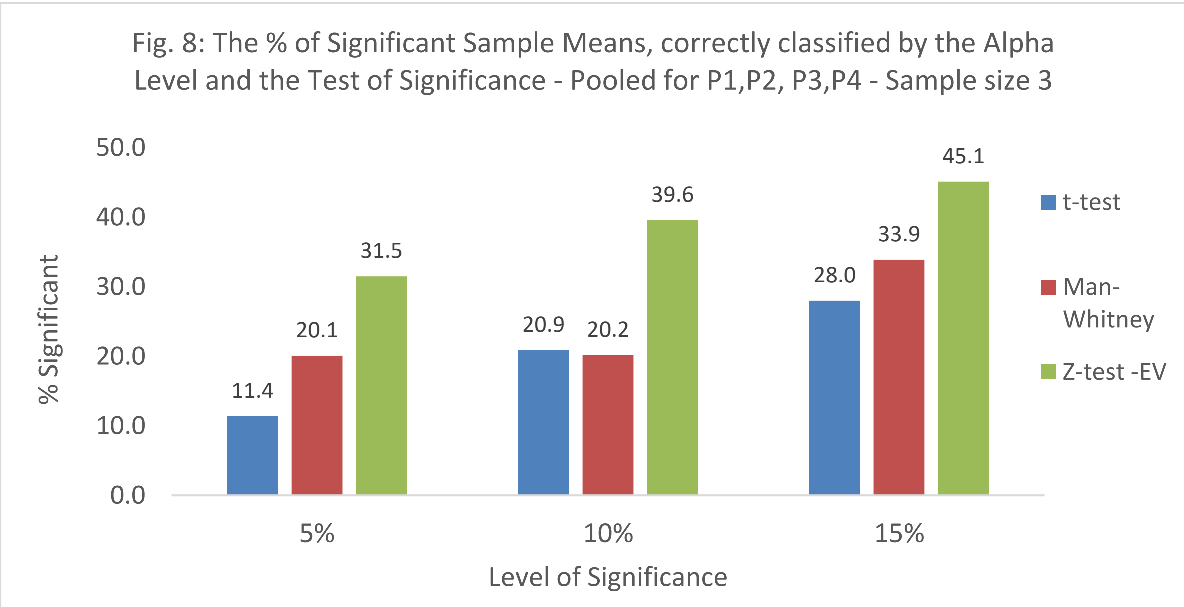
At the sample size 10 and at 5% α level, t-test and Mann-Whitney test picked up around 30% of the expected significant mean differences, correctly. In comparison, Z-EV test picked up around 40% of the expected mean differences, correctly. At selected three α levels, Z-EV test performed better compare to t-test and Mann-Whitney test. At 10% and 15% α level, Z-EV test picked up around 51% and 59% of the expected significant differences, correctly.

At sample size equal to 6, the results of significant differences in sample means found by the selected three tests when pooled for P1P2 and P3P4 are shown in Fig. 7.

At 5% α level, t-test could pick up 18.6% sample mean differences, correctly while at 10% level, it rose to 29% and at 15% α level, it rose to 37.2%. The performance of Mann-Whitney test was slightly less 3 to 4%. At 10% α level, Z-test showed 39.2% mean differences, correctly. At 15% α level, Z-test could pick up 46.6% significant differences in means, correctly.



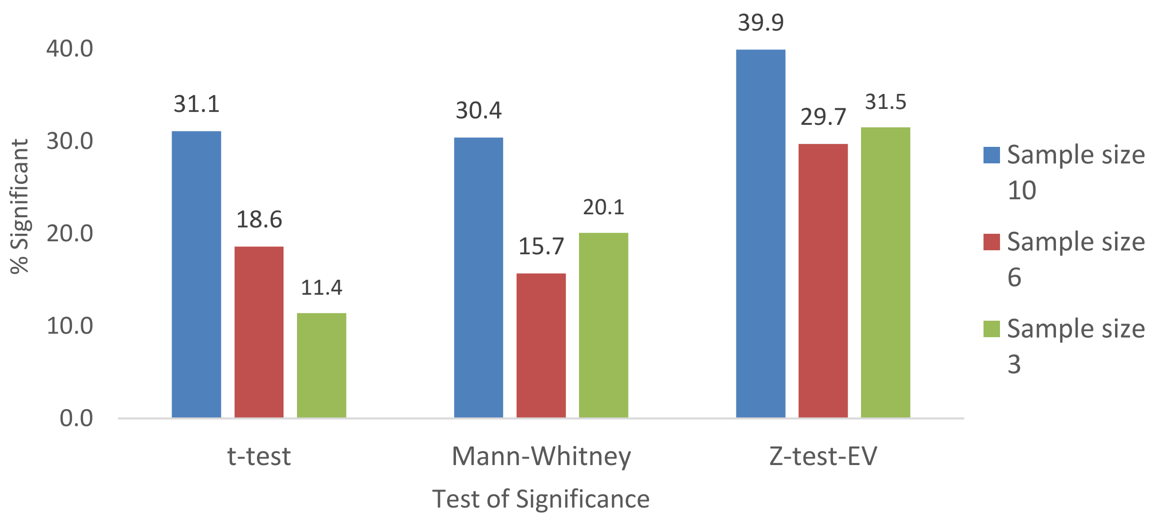
At the sample size 3, the results of significant differences in sample means found by the selected three tests when pooled for P1P2 and P3P4 are shown in Fig. 8.



At 5% α level, t-test could pick up 11.4% sample mean differences, correctly while at 10% level, it rose to 21% and at 15% rose to 28.0%. The performance of Mann-Whitney test is better as compared to t-test at 5% α level. At 15%, also, Mann-Whitney test showed around 5% higher percentage of picking up the significance differences. Z-test performance was far better as compared to t-test at all the three selected α levels. At 15% α level, Z-test picked up 45.1% mean differences, correctly.

An attempt is made in the present paper to find out the effect of sample size on the validity of t-test, Mann-Whitney test, and Z-test. The variation in the percentage significance by the sample size and 5% α level is shown in Fig. 9.

Fig. 9: Effect of Sample size on the Negative Validity at 5 % alpha Level - Pooled for P1P2 and P3P4

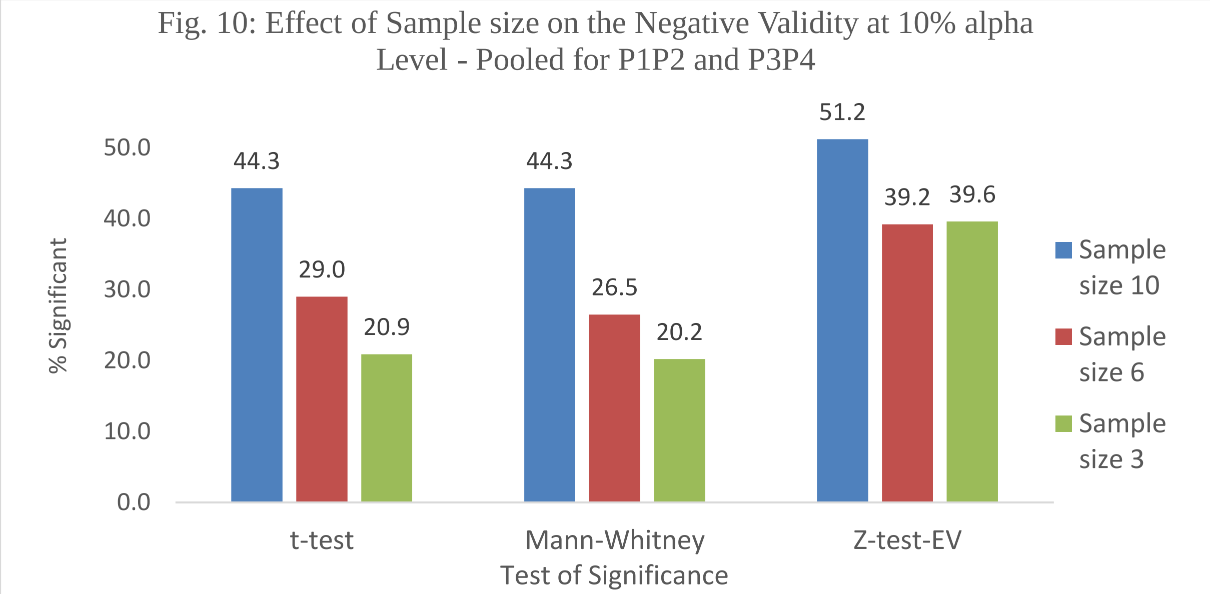


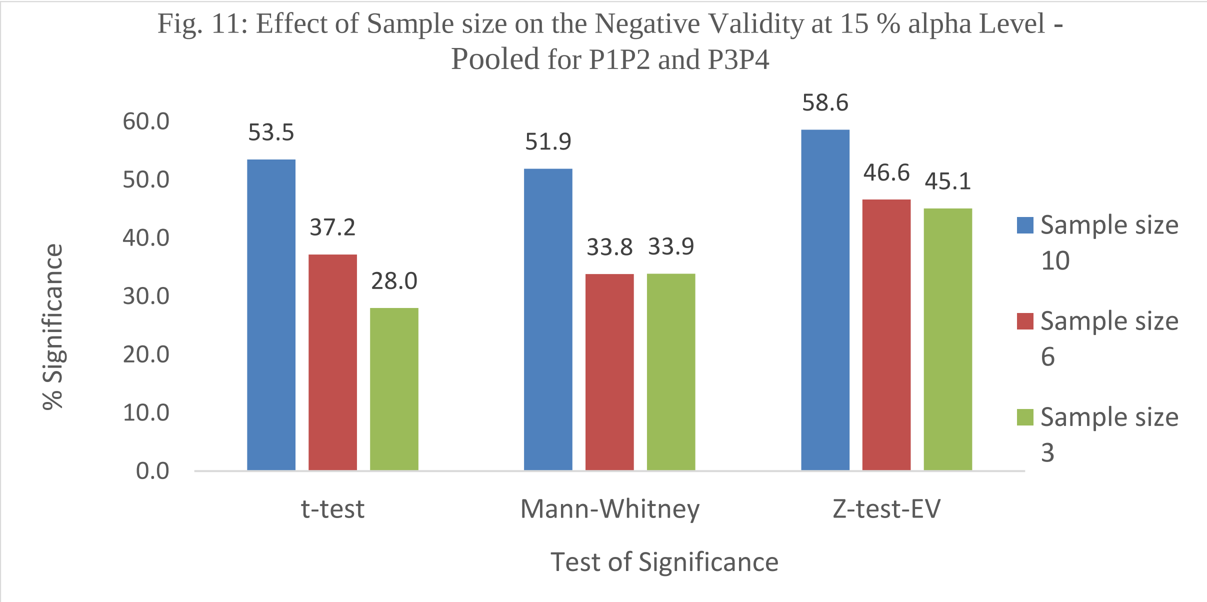
In case of t-test, at 5% α level, the negative validity changes from 31.1% for the sample size 10 to 11.4% for the sample size 3. Similarly, in case of Mann-Whitney test, the negative validity changes from 30.4% for the sample size 10 to 20.1% for the sample size 3. The Z-test performance is the best among all the tests, showing the negative validity to be 39.9% for the sample size 10 to 31.5% for the sample size 3.

The effect of sample size on the validity of t-test, Mann-Whitney test, and Z-test, at 10% α level is shown in Fig. 10.

In case of t-test, at 10% α level, the negative validity (the % significance) changes from 44.3% for the sample size 10 to 20.9% for the sample size 3. Similarly, in case of Mann-Whitney test, the negative validity changes from 44.3% for the sample size 10 to 20.2% for the sample size 3. The Z-test performance is the best among all the tests, showing the negative validity to be 51.2% for the sample size 10 to 39.6% for the sample size 3.

The effect of sample size on the validity of t-test, Mann-Whitney test, and Z-test, at 15% α level is shown in Fig. 11.





In case of t-test, at 15% α level, the negative validity changes from 53.5% for the sample size 10 to 28.0% for the sample size 3. Similarly, in case of Mann-Whitney test, the negative validity changes from 51.9% for the sample size 10 to 33.9% for the sample size 3. The Z-test performance is the best among all the tests, showing the negative validity to be 58.6% for the sample size 10 to 45.1% for the sample size 3.

DISCUSSION

The results of the study are based on the comparisons of 18000 sample means with that of population mean. An equal number of comparisons are made to test the significance differences or otherwise between two sample means when samples are known to be drawn from different normal populations. It is to be noted that irrespective of the sample size (10, 6 or3), when the variance is known, the t-test and Z-test gave almost comparable results and can pick up more than 95% expected non-significant differences (Positive validity), correctly. With the change in α level to 15%, the ability to pick up the expected non-significant differences falls by 10% that is around 85% for t-test and Z-KV test. Assuming 80% as the expected cut off level for positive validity, Z-EV test appears to be good enough when alpha is up to 10% and the sample size is 6 . Based on these observations, it can be said that all the three tests are good enough to pick up the expected non-significant differences between sample and population mean when the sample size is 6 or 10 and α level is 10%. The scenario changes, when an attempt is made to evaluate the significant differences between two sample means especially when samples are known to be drawn from two different normal populations. At 5% α level, it is seen that the negative validity of the t-test is 31.1%, 18.6% and 11.4% for the sample size of 10, 6 and 3, respectively. I do not think that the negative validity of 31.1% would be acceptable to any researcher. This points out that t-test is not at all suitable when the sample size is 10 or below. However, if we change the α = 10%, the negative validity of t-test goes up to 44.3% and to 53.5% when α is chosen to be 15%. The possible reason for high positive validity (above 95%) and low negative validity (below 35%) is the way SD is defined with (n-1) as the denominator instead of n. This leads to high variance especially when n is low like below 10 (Takiar 2022), resulting in higher critical differences then expected as compared to when SD is calculated with n as the denominator. This, in turn results in a very high percentage of accepting the Ho and conversely it results in significantly lower percentage of rejecting the Ho when it is supposed to have rejected.

The negative validity, at all the three selected α levels, is slightly lower in the case of Mann-Whitney test. In comparison, Z-EV-test performs better as compared to t-test when the sample size is 10, giving the negative validity to be 39.9%, 51.2% and 58.3% for α =5%, 10% and 15%, respectively. This raises a doubt in use of t-test for small samples particularly when it was opted in place of Z-test, believing that it is a better choice of test to examine the possible significant or non-significant differences between sample means. In view of Z-EV test scoring better in negative validity as compared to t-test, Z-EV test can be opted instead of t-test even for testing the significance differences between 2 sample means when n is above 10 and below 30.

CONCLUSIONS

- t-test is not suitable when the sample size is 10 or below.
- t-test has shown a very high, above 95%, positive validity for finding the expected non-significant differences, correctly.
- However, t-test has a very low negative validity, below 32% when α = 5%.
- The negative validity rises to 44.0% with the rise in α = 10%.
- The possible reason for accepting Ho more often than rejecting it is due to the way SD is defined with (n-1) as the denominator.
- Mann-Whitney test showed little lower validity as compared to t-test at all the α levels and at all the three sample sizes namely 10, 6 and 3.
- This also suggests that even Mann-Whitney test is not suitable when the sample size is 10 or below 10.
- Z-EV test like t-test, has shown a very high, above 95%, positive validity for finding the expected non-significant differences correctly.
- Z-EV test showed the negative validity to be 39.9% when α = 5%.
- The negative validity for Z-EV test improved to 51.2% when α = 10%.

- For the sample size above 10 and below 30, Z-EV test can safely be used instead of t-test assuming α to be 10%.
- In case, t-test has to be used with sample size below 10, it is advised to use α =10% or 15%.

REFERENCES

[1]. Gupta SC 2012: Fundamental of Statistics, Seventh Edition, Himalaya Publishing House; Page 19.1-19.2, 19.12-19.14.

[2]. Gupta SC, Kapoor VK 2001: Fundamentals of Mathematical Statistics, Sultan Chand & Sons; Tenth revised Edition, Page 14.1-14.3 and 14.24-14.26.

[3]. IBM Corp. (2015). IBM SPSS Statistics for Windows -Version 23.0, IBM Corp.

[4]. Microsoft Corporation, 2019. Microsoft Excel, Available at: https://office.microsoft.com/excel.

[5]. Snedecor GW, Cochran WG 1967; Statistical Methods, Sixth Edition; Oxford & IBH Publishing Co., Page 59-60, 100-105.

[6]. StatPlus 7.6.5.0 2021, AnalystSoft Inc. - Statistical analysis program. Version v7. See https://www.analystsoft.com/en/

[7]. Takiar R (2021): The Validity of t-test and Z-test for Small One sample and Small Two Sample test, Bulletin of Mathematics and Statistics Research, Vol. 9(4), Page 42-57.

[8]. Takiar R (2022): Sample Variance – Is it Really an Unbiased Estimate of the Population Variance? Bulletin of Mathematics and Statistics Research, Vol. 10(1), Page 21-30.

A Brief Biography of Corresponding Author

Dr. Ramnath Takiar: I am a Post graduate in Statistics from Osmania University, Hyderabad. I did my Ph.D. from Jai Narain Vyas University of Jodhpur, Jodhpur, while in service, as an external candidate. I worked as a research scientist (Statistician) for Indian Council of Medical Research from 1978 to 2013 and retired from the service as Scientist G (Director Grade Scientist). I am quite experienced in large scale data handling, data analysis and report writing. I have 65 research publications in national and International Journals related to various fields like Nutrition, Occupational Health, Fertility and Cancer epidemiology. During the tenure of my service, I attended three International conferences namely in Goiana (Brazil-2006), Sydney (Australia-2008) and Yokohoma (Japan-2010) and presented a paper in each. I also attended the Summer School related to Cancer Epidemiology (Modul I and Module II) conducted by International Agency for Research in Cancer (IARC), Lyon, France from 19th to 30th June 2007. After my retirement, I joined my son at Ulaanbaatar, Mongolia. I worked in Ulaanbaatar as a Professor and Consultant from 2013-2018 and was responsible for teaching and guiding the Ph.D. students. I also taught Mathematics to undergraduates and Econometrics to MBA students. During my service there, I also acted as the Executive Editor for the in-house Journal "International Journal of Management". I am still active in research and have published 5 research papers during 2021-23.