Asian Journal of Probability and Statistics

3(3): 1-8, 2019; Article no.AJPAS.47695

Comparative Study on the Degree of Randomness of Few Popular Random Number Tables

Pramit Pandit[1*] and Bishvajit Bakshi[2]

[1]*Department of Agricultural Statistics, Bidhan Chandra Krishi Viswavidyalaya, Mohanpur, Nadia, India.*
[2]*Department of Agricultural Statistics, Applied Mathematics and Computer Science, University of Agricultural Sciences, Bengaluru, India.*

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJPAS/2019/v3i330093
Editor(s):
(1) Dr. Mervat Mahdy Ramadan Mahdy, Associate Professor, Department of Statistics, Mathematics, & Insurance, College of Commerce, Benha University, Egypt.
Reviewers:
(1) Yahaya Shagaiya Daniel, Kaduna State University, Nigeria.
(2) Olumide adesina, Olabisi Onabanjo University, Nigeria.
Complete Peer review History: http://www.sdiarticle3.com/review-history/47695

Received: 27 December 2018
Accepted: 13 March 2019
Published: 19 March 2019

Review Article

Abstract

In the field of statistics as well as in the different branches of experimental sciences, random number tables have been playing a vital role for the purpose of selecting random samples. Among the existing different random number tables, four tables namely, Tippet's random number table, Fisher and Yates random number table, Kendall and Smith's random number table and random number table of RAND Corporation are of most frequent use. The current study aims at attempting to make a comparative review on the degree on randomness of these four most frequently used random number tables based on χ^2 test, run test and deviation test. From the findings based on χ^2 test, the highest degree of randomness has been observed in random number table due to RAND Corporation followed by due to Kendall and Smith, Tippet, Fisher and Yates, respectively. In case of run test, the highest degree of randomness has been noticed in random number table due to Fisher and Yates followed by due to Tippet, RAND Corporation, Kendall and Smith, respectively. However, from the findings based on the deviation test, the highest degree of randomness has been observed in random number table due to Kendall and Smith followed by due to Fisher and Yates, RAND Corporation, Tippet, respectively. It can observed that the findings obtained in the studies based on different tests are not alike. Consequently, there is necessity to search for the reasons of the difference between these findings. Moreover, it can also be concluded that attempts should be made by the researchers to construct new random numbers table with enhanced degree of randomness than that of the existing tables.

**Corresponding author: E-mail: pramitpandit@gmail.com;*

Keywords: Fisher and Yates; Kendall and Smith; RAND Corporation; randomness; random number table; Tippet.

1 Introduction

Random number tables have been playing a vital role in statistics as well as in the different branches of experimental sciences for the purpose of selecting random samples. Use of these tables are much more effective than selecting the random samples manually with dice, cards etc. Several random number tables have already been constructed by the renowned researchers. Those contributions are mainly due to Tippet [1], Fisher and Yates [2], Kendall and Smith [3,4], Mahalanobis [5], Quenouille [6], Rand Corporation [7], Snedecor and Cochran [8], Hald [9], Royo and Ferrer [10], Moses and Oakford [11], Rohlf and Sokal [12], Manfred [13], Rao, Mitra and Matthai [14] etc. Methods of drawing of random four-digit numbers, random five-digit numbers, random six-digit numbers and random seven-digit numbers from a combination of independent tables of random two-digit numbers and random three-digit numbers have also already been developed [15,16,17,18,19,20,21,22]. However, usage of computational random number generators have also been observed to be emerging. If carefully prepared, the process of filtering and testing can eliminate any noticeable bias or asymmetry from the numbers such that the tables provide the most 'reliable' random numbers available to the casual user. Among these different random number tables, four tables namely, Tippett's random number table, Fisher and Yates random number table, Kendall and Smith's random number table and random number table of RAND Corporation are of most frequent use [23]. The current study aims at attempting to make a comparative review on the degree on randomness of these four most frequently used random number tables.

2 Frequently Used Random Number Tables

2.1 Tippet's random number table

This table consists of 10,400 four-digit random numbers. Karl Pearson emphasized on testing statistical theories by sampling experiments. Tippet's random number could put to use for this purpose [24].

2.2 Fisher and Yates random number table

From the 10th to 19th digits of A.S. Thompson's 20-figure logarithmic tables, Fisher and Yates obtained the random numbers. In choosing from those digits, An element of randomness was introduced by using playing cards for the selection of half pages of the tables and of a column between 10^{th} to 19^{th} and finally for allotting these digits to the 50th place in a block [25].

2.3 Kendall and Smith's random number table

In the year of 1939, a set of 100,000 digits were published by M.G. Kendall and B. Babington Smith. Those digits were produced by a specialized machine in conjunction with a human operator [26].

2.4 Random number table of RAND Corporation

In the mid-1940s, development of a large table of random number table was set about by the RAND Corporation with the Monte Carlo method. With the help of a hardware random number generator, 'A Million Random Digits' with 100,000 Normal Deviates were produced. The RAND table used electronic simulation of a roulette wheel attached to a computer, the results emanated from that were then filtered and tested with substantial care before being used to generate the table [27].

3 Tests Used for Checking Randomness

3.1 χ^2 test

Pearson's chi-square test has been used in order to test whether the occurrences of the numbers appeared in the table is random or not [28,29,30,31]. This is equivalent to test that equal numbers of 0s, 1s, 2s, 3s, … , 9s are present in the table or not.

Let N be the number of occurrences of the ten digits in the table and O_i = Observed frequency of the digit i, E_i = Expected frequency of the digit i (i = 0 , 1 , 2 , ... , 9) among the N occurrences. Then the χ^2 statistic for testing the null hypothesis, "the occurrences of the digits in the table is random" i.e. "each digit has the probability 0.1 to occur in any position", which is equivalent to testing "the discrepancy between the observed frequencies and the corresponding expected frequencies of the digits is insignificant" [23] is $\chi^2 = \sum_{i=1}^{9} \frac{(O_i - E_i)^2}{E_i}$, which follows χ^2 distribution with 9 degrees of freedom.

This statistic can be employed to examine the randomness of the whole table as well as of any part of the table provided that the test satisfies the necessary assumptions of simple random sample, sample size, expected cell count and independence [32,33,34]. The frequency test was applied to each 100^{th} occurrences.

3.2 Run test

The run test is a non-parametric test to test the randomness for a two valued data sequence [35]. A run test is based on the null hypothesis that from the same distribution, each element in the sequence has been drawn independently.

Let us consider the following hypothesis:

H_0: The occurrences of numbers in a table are in random manner.
H_1: The occurrences of the numbers in the table are not in random manner.

Let, U = Number of observed runs yielded by n successive numbers in a table. Then, U follows a binomial distribution with expectation E(U) and variance V(U) given by $\frac{(n+2)}{2}$ and $\frac{n(n-1)}{4(n-1)}$, respectively.

Then for large n, under H_0, the test statistic $Z = \frac{U-E(U)}{\sqrt{V(U)}} \sim N\,(0, 1)$

One has to accept or reject the null hypothesis H_0 on comparing the values of |Z| with the corresponding theoretical value of |Z| namely 1.96 (at 5% level of significance) and 2.58 (at 1% level of significance). The test was applied to each 200^{th} occurrences.

3.3 Deviation test

The statistic t can be considered as the ratio of the departure of the estimated value of a parameter from its hypothesized value to its standard error. The t-test is any statistical hypothesis test in which the test statistic follows a Student's t-distribution under the null hypothesis [36,37].

Let $d_i = d_i$ (N) be the deviation of the observed number of occurrences of the digit i from its theoretical number of occurrences among N occurrences of the 10 digits (i = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9). Then among the 10 deviations (d_0, d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8, d_9), independent values can be assumed by any nine. Now, if the occurrences of the 10 digits are random, then di = 0 in the ideal situation. However, due to chance error, d_i may assume non-zero value.

Thus, d_i's chance errors but not assignable error if the occurrences of the 10 digits in the set of the N occurrences. The chance variables are assumed to be independently & identically distributed.as N (0, σ^2). Testing of randomness of occurrences of the 10 digits is equivalent to testing the hypothesis H_0 that $E(d_i) = 0$ for i = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9.

Test statistic can be expressed as $t = \frac{\bar{d}}{s/\sqrt{n}} \sim t_{n-2}$,where $\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$ and $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2$

H_0 is rejected at the significance level α if the calculated value of t is found to be exceeding its corresponding theoretical value that corresponds to the level of significance α with (n-2) degrees of freedom. The test was applied to each 2000th occurrences.

4 Findings of the Randomness Tests

4.1 Findings of the χ^2 test

It is reported for Tippet's random number table that the highest observed chi-square value with 9 degrees of freedom is 15.814, whereas the theoretical value of chi-square with 9 degrees of freedom at 5% level of significance is 16.919. Thus, the lack of randomness of Tippet's random number table was found insignificant at 5% significance level. However, the observed chi-square value corresponds to its theoretical value at 7.5% level of significance. In other words, the lack of randomness of Tippet's random number table can be regarded as significant at the level of significance >7.5% and insignificant at the level of significance < 7.5% [23].

In case of Fisher and Yates random number table, it was observed that the highest observed chi-square value with 9 degrees of freedom is 26.118, which is higher than the theoretical value of chi-square with 9 degrees of freedom at both 5% and 1% significance level (16.919 and 21.666, respectively). Thus, the lack of randomness of Fisher & Yates random number table can be regarded as significant not only at 5% level of significance but also at 1% level. However, the observed chi-square value corresponds to its theoretical value at 0.055% level of significance. In other terms, the lack of randomness of Fisher and Yates random number table can be regarded as significant at the level of significance > 0.055% and insignificant at the level of significance < 0.055% [25].

It was mentioned for Kendall and Smith's random number table that the highest observed chi-square value with 9 degrees of freedom is 13.4, which is less than the corresponding theoretical value of chi-square at 5% level of significance. Thus, the lack of randomness of Kendall and Smith's random number table was found insignificant at 5% significance level. However, the observed chi-square value with 9 degrees of freedom namely 13.4 corresponds to the theoretical value of chi-square with 9 degrees of freedom at 18.1% level of significance. Thus, the lack of randomness of Kendall and Smith's random number table can be regarded as significant at the level of significance >18.1% and insignificant at the level of significance < 18.1% [23].

For the random number table due to Rand Corporation, it was observed that the highest observed chi-square value with 9 degrees of freedom is 12.518, which is less than the corresponding theoretical value of chi-square at 5% significance level. Thus, the lack of randomness of random number table due to Rand Corporation can be regarded as insignificant at 5% significance level. However, the observed chi-square value with 9 degrees of freedom namely 12.518 corresponds to the theoretical value of chi-square with 9 degrees of freedom at 24% significance level. In other words, the lack of randomness of random number table due to Rand Corporation can be regarded as significant at the level of significance >24% and insignificant at the level of significance < 24% [23].

From the findings based on χ^2 test, it is clear that highest degree of randomness is present in random number table due to RAND Corporation followed by due to Kendall and Smith, Tippet, Fisher and Yates, respectively.

4.2 Findings of the run test

On comparing the observed values with the corresponding theoretical Z values, the lack of randomness in the three parts containing 19^{th}, 21^{st} and 25^{th} 200 trials respectively in Tippet's random number table can be regarded as significant at 5% level but not at 1% level while the lack of randomness in the other parts can be treated as insignificant [38].

The lack of randomness in Fisher and Yates random number table was found to be non-significant at 5% level by comparing the observed values with the corresponding theoretical Z values [38].

The lack of randomness in the parts containing 1^{st}, 2^{nd}, 26^{th}, 33^{rd}, 36^{th}, 45^{th}, 46^{th}, 48^{th}, 65^{th}, 68^{th} and 70^{th} 200 trials respectively in Kendall and Smith's random number table was found significant at 5% significance level but not at 1%, while the lack of randomness in the other parts of the table can be treated as insignificant [26].

It was found on comparing the observed values with the corresponding theoretical Z values, that the lack of randomness in the four parts containing 35^{th}, 52^{nd}, 73^{rd}, 94^{th} 200 trials respectively in Rand Corporation random number table can be regarded as significant at 5% significance level but not at 1% level, while the lack of randomness in the other parts of the table can be considered as insignificant [38].

From the findings based on run test, it is clear that highest degree of randomness is present in random number table due to Fisher and Yates followed by due to Tippet, RAND Corporation, Kendall and Smith, respectively.

4.3 Findings of the deviation test

On comparing the observed values with the corresponding theoretical t values, that the lack of randomness of Tippet's random number table can be treated to be highly significant i.e. significant at both 5% and 1% significance level, except the four parts corresponding to the four sets of trials specifically 1^{st} 2000, 10^{th} 2000, 17^{th} 2000 and last 1600 trials. However, the lack of randomness in these four parts of the table was found significant at 5% level [39].

By comparing the observed values with the corresponding theoretical t values, that the lack of randomness of Fisher and Yates random number table was found to be highly significant i.e. significant at both 5% and 1% significance level, except the two parts corresponding to the two sets of 11^{th} and 13^{th} 1000 trials. However, the lack of randomness in these two parts of the table was found at 5% level of significance [39].

The lack of randomness of Kendall and Smith's random number table can be treated to be highly significant i.e. significant at both 5% and 1% level of significance except the part corresponding to the set of 5^{th} 2000 trials, by comparing the observed values with the corresponding theoretical values of t. However, the lack of randomness in this particular part of the table is significant at 5% level [39].

On comparing the observed values with the corresponding theoretical t values, the lack of randomness of Rand Corporation random number table was found significant both at 5% and 1% level of significance except the five parts corresponding to the five sets viz. 3^{rd}, 7^{th}, 20^{th}, 23^{rd} and 25^{th} sets of 2000 trials. However, the lack of randomness in these five parts of the table was significant at 5% level [39].

From the findings based on the deviation test, it is clear that highest degree of randomness is present in random number table due to Kendall and Smith followed by due to Fisher and Yates, RAND Corporation, Tippet respectively.

5 Conclusion

Degree of randomness present in different random number tables (Tippet's random number table, Fisher and Yates random number table, Kendall and Smith's random number table and Random number table of RAND Corporation) based on χ^2 test, run test and deviation test are presented in Table 1.

Table 1. Ranks of the four random number tables

Random Number Table	**Rank with respect to the degree of presence of randomness based on χ^2 test**	**Rank with respect to the degree of presence of randomness based on run test**	**Rank with respect to the degree of presence of randomness based on deviation test**
Tippet's random number table	3	2	4
Fisher and Yates random number table	4	1	2
Kendall and Smith's random number table	2	4	1
Random number table due to RAND Corporation	1	3	3

It can observed that the findings obtained in the studies based on different tests are not alike. Consequently, there is necessity to search for the reasons of the difference between these findings. Chakrabarty's random number table has recently been developed, which has been observed to yield more randomness compared to the aforesaid tables on the basis of χ^2 test [40]. Moreover, it can also be concluded that attempts should be made by the researchers to construct new random numbers table with enhanced degree of randomness than that of the existing tables.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Tippett LHC. Random sampling numbers, tracts for computers no 15. Cambridge: Cambridge University Press; 1927.

[2] Fisher RA, Yates F. Statistical tables for biological, agricultural and medical research. 6th Edition. England: Longman Group Limited; 1982.

[3] Kendall MG, Smith BB. Randomness and random sampling numbers. J. R. Stat. Soc. 1938;101(1): 147–166.

[4] Kendall MG, Smith BB. A table of random sampling numbers, tracts for computers no 24. Cambridge: Cambridge University Press; 1939.

[5] Mahalanobis PC. Tables of random samples from a normal population. Sankhya. 1934;1:289–328.

[6] Quenouille MH. Tables of random observations from standard distributions. Biometrika. 1959;46: 178–202.

[7] Rand Corporation. A million random digits. Glenoe: Free Press; 1955.

[8] Snedecor GW, Cochran WG. Statistical methods. 6th Edition. Iowa: Iowa State University Press; 1967.

[9] Hald A. Statistical tables and formulas. Wiley; 1952.

[10] Royo J, Ferrer S. Tables of random numbers obtained from numbers in the Spanish National Lottery. Trab. Est. 1954;5:247–256.

[11] Moses EL, Oakford VR. Tables of random permutations. George Allen & Unwin; 1963.

[12] Rohlf FJ, Sokal RR. Ten thousand random digits. Freeman; 1969.

[13] Manfred M. Le Petit Livre de Nombres au Hasar. Paris: Édition d'artiste; 1971.

[14] Rao CR, Mitra SK, Matthai A. Random numbers and permutations. Calcutta: Statistical Publishing Society; 1966.

[15] Chakrabarty D. Drawing of random five-digit numbers from tables of random two-digit and three-digit numbers. Int. J. Adv. Res. Sci. Eng. Technol. 2016;3(7):2385–2306.

[16] Chakrabarty D. Drawing of random six-digit numbers from tables of random three-digit numbers. Int. J. Adv. Res. Sci. Eng. Technol. 2016;3(8):2507–2515.

[17] Chakrabarty D. Drawing of random six-digit numbers from tables of random two-digit numbers. Int. J. Adv. Res. Sci. Eng. Technol. 2016;3(9):2643–2655.

[18] Chakrabarty D. Drawing of random six-digit numbers from a single table of random two-digit numbers. Int. J. Adv. Res. Sci. Eng. Technol. 2016;3(10):2743–2753.

[19] Chakrabarty D. Drawing of random six-digit numbers from a single table of random three-digit numbers. Int. J. Adv. Res. Sci. Eng. Technol. 2016;3(11):2905–2914.

[20] Chakrabarty D. Drawing of random seven-digit numbers from tables of random two-digit numbers and of three-digit numbers. Int. J. Adv. Res. Sci. Eng. Technol. 2016;3(12):3003–3012.

[21] Chakrabarty D. Drawing of random four-digit numbers from independent tables of random two-digit numbers in selection of random sample. Biom. Biostat. Int. J. 2016;4(7):1-6.

[22] Chakrabarty D. Drawing of random four-digit numbers from a single table of random two-digit numbers. Int. J. Adv. Res. Sci. Eng. Technol. 2017;4(2):3877–3887.

[23] Chakrabarty D, Sarma BK. Comparison of degree of randomness of the tables of random numbers due to Tippet, Fisher & Yates, Kendall & Smith and RAND Corporation. J. Reliab. Stat. Stud. 2017; 10(1):27-42.

[24] Nair KR. On Tippetts ‘Random Sampling Numbers’. Sankhya. 1938;4(1):65-72.

[25] Sarmah BK, Chakrabarty D. Testing of randomness of the number generated by fisher and Yates. Int. J. Eng. Sci.Res. Technol. 2014;3(11):632-636.

[26] Sarmah BK and Chakrabarty D. Examination of proper randomness of numbers of M. G. Kendall and B. Babington Smith’s Random numbers table: Run test. Int. J. Multidiscip. Res. Mod. Educ. 2015; 1(2):223-226.

[27] Moore PG. Reviews: A million random digits with 100,000 normal deviates. Biometrika. 1955; 42(3-4):543–543.

[28] Yates F. Contingency table involving small numbers and the chi square test. J. R. Stat. Soc. 1934; 1(2):217–235.

[29] Corder GW, Foreman DI. Nonparametric statistics: A step-by-step approach. New York: Wiley; 2014.

[30] Greenwood PE, Nikulin MS. A guide to chi-squared testing. New York: Wiley; 1996.

[31] Chernoff H, Lehmann EL. The use of maximum likelihood estimates in χ^2 tests for goodness of fit. Ann. Math. Stat. 1954;25(3):579–586.

[32] Gupta SC, Kapoor VK. Fundamentals of mathematical statistics. 10th Ed. New Delhi: Sultan Chand & Sons; 2014.

[33] Zar JH. Biostatistical analysis. 5th Ed. Upper Saddle River, N.J.: Prentice-Hall/Pearson; 2010.

[34] Rangaswamy R. A textbook of agricultural statistics. 2nd Ed. New Delhi: New Age International Publishers; 2010.

[35] Bradely JV. Distribution free statistical tests. 1st Ed. USA: Prentice Hall; 1968.

[36] Sahu PK, Pal SS, Das AK. Estimation and inferential statistics. New Delhi: Springer (India) Pvt. Ltd.; 2015.

[37] Sahu PK. Applied statistics for agriculture, veterinary, fishery, dairy and allied fields. Springer Nature: New Delhi; 2016.

[38] Chakrabarty D. Random numbers tables due to Tippet, Fisher & Yates, Kendall & Smith and Rand Corporation: Comparison of Degree of Randomness by Run Test. J. Biostat. Biom. App. 2018;3(1): 1-8.

[39] Chakrabarty D. Deviation test: Comparison of degree of randomness of the tables of random numbers due to Tippet, Fisher & Yates, Kendall & Smith and Rand Corporation. SM J. Biom. Biostat. 2017; 2(3):1-6.

[40] Chakrabarty D. Chakrabarty's tables of random numbers: A comparison of degree of randomness with that of some existing random numbers tables. Amer. J. Biom. Biostat. 2018;2(1):3-10.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://www.sdiarticle3.com/review-history/47695