Direct and Indirect Tests for Randomness

Chanoknath Sutanapong

About the author

★Chanoknath Sutanapong is an independent researcher. She may be reached by email at: chanoknath.sutanapong@gmail.com

ABSTRACT

The purpose of this paper is to explain the importance of randomness in data analysis. We point out the difference between random data and random selection. The scope of this paper is limited to randomness in discrete data. We explain the various tests used by NIST as the basis to understand randomness testing in discrete data. We also present an indirect test for randomness by relying on trend testing. Randomness is defined as lacking predictable pattern. Predictable pattern may be seen through trend. We argue that if there is a significant trend, the series is not random and vice versa. In this paper, we claim that trend test may be used an indirect test for randomness.

Keywords: Discrete data, randomness, sampling method, trend

CITATION:

Sutanapong, Chanoknath (2018). "Direct and Indirect Tests for Randomness." *Inter. J. Res. Methodol. Soc. Sci.*, Vol., 4, No. 1: pp. 63–83. (Jan. – Mar. 2018); ISSN: 2415-0371.

1.0 INTRODUCTION

1.1 Introduction to random data and random selection

Random selection deals with the sampling method. The misunderstanding here is that many researchers incorrectly believe that random selection produces random data. A sampling method does not produce the type of data nor can it change the nature of the data. If the data is random, no matter what kind of sampling method is used, the data would remain random. This randomness may be revealed through random testing. On the other hand, if the data is non-random, no matter what kind of sampling method is employed, the data would also remain non-random. For that reason, the use of "random sampling" method does not guarantee that the data would be random. The benefit of random sampling is to avoid selection bias. Selection bias occurs when each element of the population does not have equal chance of being selected. The statement "equal chance of being selection" coincidentally is a common definition of random selection; thus, to avoid the accusation of selection bias, researcher tends to claim to have used random selection.

In order to impose equal probability selection or random selection, it is necessary that the population size be known. This requirement is impractical because in most circumstances in social science, the population size is either unknown or if it is known, it is imprecise because population is dynamic, i.e. drop out, death, new birth and migration. Therefore, true random is not possible. The most that one can do is to "assume" that the population is non-dynamic, and the selection is made as if the population is finite. In such a case, the so-called random sampling is a mere presumption. The sampling made is no more than convenience sampling; it is only through self-fulfilling claim that it is said "randomly selected." Such a claim should not escape heighten ethical scrutiny. More attention should be given to randomness of the data, not how the data was selected.

Random data set means that the statistical test verifies that the data elements are random. The clam to randomness is substantiated through statistical test. The conclusion made is different from what had been said in the case of random selection. The claim by random test states that "the data came from a random process." It is this random process that is the holly grail in statistics because common statistical test requires that the data be random. "Be random" here means that the data comes from a random process, and random process means that the condition from which the data was sampled had been random. This is a completely different language from making a claim that the sampling method used was random selection. The focus on the randomness of the data under statistical tests does not give homage to the sampling method. As for selection bias, it would also show in the test as non-randomness in the data if there was indeed bias. Even if there is an alleged or suspicion of selection bias, if the data is indeed random, as verified by one or more of statistical tests, then such allegation and suspicion is vitiated.

1.2 Measurement of randomness

The objective of this section is concerned with the measurement of randomness. The methods used in the measurement of randomness are many. They can be categorized into data types, namely: (i) univariate, (ii) bivariate, and (iii) times series. In univariate data, the numbers come in one sequence or string. The test is imposed upon the string to verify whether the string or sequence is random; if so, it is said that the numbers come from a random process. Generally, in bivariate analysis, the sequence is dichotomize into (1, 0) where 1 is defined as category of interest or "success" and 0 is the category of non-interest of "failure." The dichotomy either comes from a true binary data in a form of (Yes | No) or it is made "dichotomous" through the application of mathematical expectation where the expected value or the mean is used as a separator dividing the sequence of numbers into two groups: below the mean and equal to and greater than the mean.

In the bivariate case, the data comes from two data array that may be categorized into array X and array Y. In such a case, the proof for randomness is verified through proving that X_i is comprised of independent elements $x_1, x_2, ... x_n$. Similarly, the same test is used to verify that the array Y_i also contain independent elements of $y_1, y_2, ... y_n$. Moreover, a third test in bivariate is to attest that X_i and Y_i are independent, i.e. the occurrence of Y_i does not depend on X_i so as to make the value of Y_i predictable with a given value of X_i.

A third for of number sequence used is randomness analysis is time series data. Time series are bivariate data consisting of X_i independent variable and Y_i dependent variable where the relationship between these two variable may be expressed by the regression model, thus: $X(t) = \beta_0 + \beta_1 X_1 + ... + B_n X_t + \varepsilon_i$.

In this model, it is no longer of interest as to the causality of Y_i, i.e. whatever explains the occurrence of Y_i is no longer the object of analysis. Whatever the causality or relation of X_i to Y_i is, it is taken for granted. Thus, whether the value of Y_i is produced, treat such series of events: $y_1, y_2, ..., y_n$ which occurs at time $t_1, t_2, ..., t_n$ respectively as $y_{t_1 = X(t_1)}, y_{t_2 = X(t_2)}, ..., y_{t_n = X(t_n)}$ and thus the event is listed simply as: $X(t_1), X(t_2), ..., X(t_n)$. This is called time series data array. In this scenario, the test for randomness is to look for the independents among the error term of each $X(t_i)$. If the series of $\varepsilon_i : \varepsilon_1, \varepsilon_2, ..., \varepsilon_n$ are statistically independent, it is said that the arrays of $X(t_i)$ are independent and, therefore, are random. Under these circumstances, autocorrelation test is used as a tool for verifying randomness among the time series $X(t_i)$.

2.0 Trend tests as indirect testing for randomness

If a common definition and characteristic of randomness is the lack of predictable trend or pattern, the most practical place to start the discussion of randomness is trend analysis.

One cautious note about trend test, the acceptance or rejection under the trend test is based on statistical significant test, i.e. the trend may exist but it is not considered statistically significant. Under this standard, the definition of randomness is relaxed because for cases where the trend test rejects, they may still manifest some kind of patterns, but that pattern does not pass the statistical significant test. This means that if the confidence interval of 0.95 is used, a pattern that would pass at 0.90 would be rejected at 0.95 as 'non-significant trend.' This does not necessary mean that pattern recognition of a data set of 0.90 trend display does not show a pattern. In fact, a pattern failing to reject a null hypothesis using 0.95 confidence interval may still be a "trend" or "pattern" but just does not pass 0.95 confidence interval test. For this reason, the various trend tests are not definitive test to define whether the data is random for failing to manifest trend pattern. Nevertheless, the introduction of trend test puts randomness in perspective to the extent that trend is defined as a recognizable pattern and that randomness is defined as the lack of recognizable pattern.

There are three trend tests commonly used in the field, namely (i) reverse Arrangement Test (RAT), (ii) Military Handbook Test (MHT), and (iii) La Place Trend test (LTT). The RAT method is given by:

$$Z = \frac{R - \left(\frac{r(r-1}{4}\right) + 0.50}{\sqrt{\frac{(2r+5)(r-1)r}{72}}} \tag{1}$$

where R = reversal counts, and r = repair time. The decision is governed by the following tables depending on whether we are looking for improving trend of degrading trend. The first type of trend is called an improving trend. The critical value of improving trend is given by the table below.

Table 1: One-sided test for improving trend

Repair: r	R(90%)	R(95%)	R(99%)
4	6	6	-
5	9	9	10
6	12	13	14
7	16	17	19
8	20	22	24
9	25	27	30
10	31	33	36
11	37	39	43
12	43	46	50

Source: http://www.itl.nist.gov/div898/handbook/apr/section2/apr234.htm

A second scenario of trend is a degrading trend where the pattern indicates that the process is deteriorating or "things are getting worse." The deteriorating trend uses the following critical values for testing its significance.

Table 1: One-sided test for degrading trend

Repair: r	R(90%)	R(95%)	R(99%)
4	0	0	-
5	1	1	0
6	3	2	1

7	5	4	2
8	8	6	4
9	11	9	6
10	14	12	9
11	18	16	12
12	23	20	16

Source: http://www.itl.nist.gov/div898/handbook/apr/section2/apr234.htm

The MHT approach is used when the values comes from a system that is described by Power Law. The Military handbook method is given by:

$$\chi^2_{2r} = 2\sum_{i=1}^{r} \ln\left(\frac{T_{end}}{T_i}\right) \tag{2}$$

where $T_i : T_1, T_2, ..., T_{end}$ are repair time counts. It is equivalent to *r* in the RAT equation above.

The third type of trend test is called the Laplace Trend Test. The LTT method is used when the data comes from an exponential system. LTT is given by:

$$Z = \frac{\sqrt{12r}\sum_{i=1}^{n}\left(T_i - \frac{T_{end}}{2}\right)}{rT_{end}} \tag{3}$$

If the data set shows that there is a significant trend under any of the three tests mentioned above, it is conclusive that the data is not random. The use of the trend test may be proach to verify if the data is random since “randomness” lacks predictable pattern.

3.0 NIST approach to randomness testing of discrete data

The national Institute of Standards and Technology (NIST) is an agency under the U.S. Department of Commerce (Ministry of Commerce). The mission of NIST is to “[p]romote U.S. innovation and industrial competitiveness by advancing measurement science, standards, and technology in ways that enhance economic security and improve our quality of life.” (Perry, 1953; p. 123). In random testing, NIST offers 15 different tests to verify whether a sequence of number is random.

The 15 tests offered by NIST is based on binary digits or categorical data test based on (1, 0) discrete data. Although these 15 tests are considered sufficient for binary data, they are not adequate to deal with continuous data and times series data. For this reason, continuous and tome series data are treated outside of NIST’s 15 tests. In this paper, we explain the first three of the fifteen tests in the NIST battery of test for randomness. The National Institute of Standards and Technology (NIST) offers a battery of 15 tests to verify randomness. These 15 tests are:

1. Frequency (monobit) test;
2. Frequency test within block;
3. Runs test;
4. Tests for the longest-run-of-ones in a block;
5. Binary matrix ranked test;
6. Discrete Fourier transform (spectral) test;
7. Non-overlapping template matching test;
8. Overlapping-template matching test;
9. Maurer’s “university statistical” test;
10. Linear complexity test;

11. Serial test;
12. Approximate entropy test;
13. Cumulative sum test;
14. Random excursion test; and
15. Random excursion variant test.

In order to use the random verification method advocated by NIST, the data has to be discrete or categorical comprising of (1, 0). If the data comes from a continuous scalar form, such as a survey instrument answer choice of: (0, 1, 2, 3), it is still possible to make the continuous scalar into a discrete form by separating the zero and non-zero into two categorical that could then be mapped into the form (1, 0) required by NIST. For instance, (0, 1, 2, 3) could be re-categorized into discrete form by equating {1, 2, 3}, non-zero values, to 1 = 1, 2 = 1 and 3 = 1. The data set now becomes (1, 1, 1, 0).

In the alternative, other methods for random testing that does not fall under the 15 methods proposed by NIST could accommodate continuous data set without having to go through discretization process described above. For instance, the adjacent test deals with non-categorical data set on "as is" basis with the assumption that there is an assumed cloud within which the randomness is defined. If the test shows that the observed score falls outside of the cloud, it is considered non-random. The "cloud" is defined by an interval consisting of the lower and upper values. It is said that, if the observed value under the adjacent test falls below the lower limit, it is conclusive that the data comes from a non-random process. However, if the values fall outside of the "upper limit," it is not inconclusive in making the conclusion whether the data comes from a non-random process.

In both cases, NIST's approach and other approaches, it is important to note that the study of randomness focuses on the "data", i.e. the sample, not the method by which the data had been selected, i.e. "sampling method." This distinction, and the point where the emphasis is put when deciding on randomization, underscores the ineffectual insistence on random selection. Even if the data had been produced by random selection, if the data fails the random test that uses the data as the testing value, random selection could produce a data set that is non-random. The only answer that random selection could provide is that the selection process is non-biased. This rationale for random selection is still weak. Assume that the data is random; no matter how the data is selected the sample would still show that it is random data. That is, if by its nature the data is random, even if a convenience sampling method is used to select the data, it would still produce a random data. The nature of the data does not change as the result of the sampling method.

Numerical examples will be given to illustrate each test. In cases where the original data is not binary (1, 0) or (Yes | Not), the data string will be transformed into (1,0) using the expected value as the reference point, i.e. if $0 \le x_i < \overline{X}$ then 0 and $\overline{X} \le x_i$, or simply:

$$p(X) = \begin{cases} \overline{X} \le x_i \\ 0 \quad \text{Otherwise} \end{cases} \tag{4}$$

In the second case where bivariate X and Y are involved, autocorrelation and correlation coefficient tests will be used.

In the study of randomness, it is important to make a distinction between sampling method and data from the sample. The random test is the test of the data in order to answer the question of whether "the data came from a random process?" This is different from random sampling. Random sampling is a sampling method which is classified as *equal probability sampling*. It is given by:

$$P(sample) = \begin{cases} \binom{N}{n}^{-1} \text{ if } n(s) = n \\ 0 \quad \text{otherwise} \end{cases} \tag{5}$$

This is also called "simple random sampling." The assumption here is that if the sample is selected by a random sampling method, where all elements in the population have equal chance, data bias would not taint research. However, in real life, random sampling may not always be appropriate or available. The circumstances under which the data was collected, for instance a crime scene with limited sample availability, or in a situation where data accessibility is limited, random sampling may not be the best choice. In addition, one requirement of simple random sampling is that the population size must be known hence the term $\binom{N}{n}^{-1}$. However, if the population is dynamic, i.e. changing and the drop out of existing elements or addition of new elements changes and shifts population size, then simple random sample may not be possible. In most cases, *non-equal probability sampling* is the common option in sampling method of the Midzuno scheme.(Sampath, 2005; pp. 73-74; see also Midzuno, 1952; pp. 99-107) For instance, the non-equal probability sampling of the Midzuno scheme is given by:

$$P(sample) = \begin{cases} \dfrac{\hat{X}}{X} \dfrac{1}{\binom{N}{n}} \text{if } n(s) = n \\ 0 \quad \text{otherwise} \end{cases} \tag{6}$$

Where $\hat{X}$ = unbiased estimator of population total: $\hat{X} = \frac{N}{n} \sum_{i \in s} X_i$; X = population total of the size variable x in random sampling (Sampath, pp. 73-74).

It may be argued that even if the population size is unknown (non-finite), it may be capable of being estimated. In Sampath, the estimation of the population, under the Horvitz-Thompson estimator method, is explained thus:

> "Theorem 1.6 The Horvitz-Thompson estimator $\hat{Y}_{HT} = \sum_{i \in s} \frac{Y_i}{\pi_i}$ is unbiased for the population total and its variance is
>
> $$\sum_{i=1}^{N} Y_i^2 \left[\frac{1-\pi_i}{\pi_i} \right] + 2 \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i<j}}^{N} Y_i Y_j \left[\frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \right]$$
>
> *Proof* The estimator $\hat{Y}_{HT} = \sum_{i=1}^{N} \frac{Y_i}{\pi_i} I_i(s)$
>
> Taking the expectation of both sides we get
>
> $$E_P[\hat{Y}_{HT}] = \sum_{i=1}^{N} \frac{Y_i}{\pi_i} \pi_i = Y$$

Therefore
$$\hat{Y}_{HT} - Y = \sum_{i=1}^{N} \frac{Y_i}{\pi_i} I_i(s) - \sum_{i=1}^{N} Y_i$$
$$= \sum_{i=1}^{N} \frac{Y_i}{\pi_i} [I_i(s) - \pi_i]$$

Squaring both the sides and taking expectation we get

$$E_P[\hat{Y}_{HT} - Y]^2 = \sum_{i=1}^{N} \left[\frac{Y_i}{\pi_i}\right]^2 E_P[I_i(s) - \pi_i]^2$$
$$+ 2\sum_{i=1}^{N} \sum_{\substack{j=1 \\ i<j}}^{N} \frac{Y_i}{\pi_j} \frac{Y_j}{\pi_j} E_p[I_i(s) - \pi_j]$$
$$= \sum_{i=1}^{N} \left[\frac{Y_i}{\pi_i}\right]^2 V_P[I_i(s)] + 2\sum_{i=1}^{N} \sum_{\substack{j=1 \\ i<j}}^{N} \frac{Y_i}{\pi_j} \frac{Y_j}{\pi_j} \operatorname{cov}_P(I_i(s), I_j(s))$$
$$= \sum_{i=1}^{N} Y_i^2 \left[\frac{1-\pi_i}{\pi_i}\right] + 2\sum_{i=1}^{N} \sum_{\substack{j=1 \\ i<j}}^{N} Y_i Y_j \left[\frac{\pi_i - \pi_i \pi_j}{\pi_i \pi_j}\right]$$

Hence proof. □" (Samplath, pp. 14-19).

where the inclusion indicator was defined as (Sampath, p. 4):

$$I_i(s) = \begin{cases} I ... if ... s \in i, 1 \le i \le N \\ 0 otherwise \end{cases} \tag{7}$$

The inclusion probability is defined as:

$$\pi_i = \sum_{s \in i} P(s) \quad \text{and} \quad \pi_{ij} \sum_{s \in i, j} P(s) \tag{8}$$

The "approach", i.e. Horvitz-Thompson method (Hovitz, 1952; pp. 663-6) as discussed in Sampath, is not helpful in obtaining the total population N. Although it provides series of steps in mathematical proof, it presents little practical value. Therefore, it is necessary to look for other practical methods to estimate the population size. Without the complicated approach displayed by Sampath, the Horvitz-Thompson method could be explained simply as a method to estimate the mean and total of a superpopulation in a stratified sample. This point (stratified sampling) was not properly explained by Sampath when the Horvits-thompson was introduced. The Horvits-Thompson was introduced as "the most popular … estimator for population total …" when in fact it should have been qualified as an estimator for population total in stratified sampling (*see* Sampath, pp. 4-8). In so doing, the inverse weighting is used to account for the different proportion found in each strata of the population. Let $Y_i = 1,2,...,n$ be an independent sample from n taken from the sample space of $N \ge n$ where each n_i is a distinct stratum with the common mean of μ. Under this method, it is further assumed that there is a probability that each random item in n_i will be included in each stratum is called inclusion probability defined as π_i. Under this set up, the Horvits-Thompson estimator is given by:

$$\hat{Y}_{HT} = \sum_{i=1}^{n} \pi_i^{-1} Y_i \tag{9}$$

where $\hat{Y}$ = estimated superpopulation; the subscript HT = Horvitz-Thompson; π = inclusion probability; and Y_i = independent sample, i.e. strata, $Y_i = 1,2,...,n$ taken from the sample space of $\Omega : N \geq n$.

From (3.6), the estimated mean is given by:

$$\hat{\mu}_{HT} = N^{-1} \hat{Y}_{HT} \tag{10}$$

Which means that $\hat{Y}_{HT}$ is a portion of the total population or $\hat{\mu}_{HT} = \frac{\hat{Y}_{HT}}{N}$. By substituting (3.6), statement (3.7) becomes:

$$\hat{\mu}_{HT} = \frac{Y_i}{N\pi_i} \tag{11}$$

The process by which the estimated mean μ_{HT} may be thought of as an estimate obtained through series of bootstrap or Jackknife resampling technique (*viz.* Roderick and Rubin, 2002).

There are several tools used in population estimates; these methods can be classified into two types: (i) direct and (ii) indirect methods for estimating population size. The direct method is exemplified by the Release-Recapture method (RR). The RR method is practical for non-human population, i.e. capture, release, and recapture animals in the while. For human population, this method may be impractical because it is not easy to "tag" humans. The RR method is defined as:

$$\hat{N} = \frac{\text{(Total captured)(Number marked)}}{\text{Total caputred with mark}} \tag{12}$$

This method is known as *direct method* because it involves the use of actual counting of the population. The estimate of the population size is based on these trials of capture, release, and recapture. In social science, this capture-release-recapture method may not be feasible because it is not easy to "mark" or "tag" human beings. However, in specific cases, such as in marketing research or tracking customers, the method may be a useful tool for determining customer base or market size. For instance, tracking customers in a department store by telephone numbers or membership card is an equivalent of what zoologist does in the wild by capture-release-recapture subjects.

The second direct method of estimating population is to work with known population size from the prior period and adjust it for the number of new births, deaths and net migration. This second method is given as:

$$P_{t+n} = P_t + B_{t,t+n} - D_{t,t+n} + M_{t,t+n} \tag{13}$$

where

P_{t+n} = change of population from one period to the next; P_t = population size of the last period; $B_{t,t+n}$ = live birth that occurs during time between t and $t+n$; $D_{t,t+n}$ = death during the interval; and $M_{t,t+n}$ = net migration, i.e. $= M_{t,t+n} = I_{t,t+n} + O_{t,t+n}$

For an *indirect method* of population estimation is the census method, the formula for the census method is given as:

$$P_{est} = P_1 + \frac{n}{N}(P_2 + P_1) \tag{14}$$

where P_{est} = population estimate; n = number of months from P_1 census date of the estimate; N = number of months between census period; P_2 = population size in the last census; and P_1 = population in the second to last census.

From the two methods shown above, population size is not easily determined. Under the first method, capture-tag-release-recapture may not be feasible in human population study unless the population is in a controlled environment, such as club card scheme used in supermarket to track customers. The second method relies on prior population estimate. This method is feasible only if prior population size is known. In random sampling, it is necessary that the population size be known because random sampling is an equal probability sampling method. This sapling method requires that the population size be known in order to determine the probability of each element in the population. When the population size is not known or is uncertain, then equal probability sampling may not be feasible. Therefore, random sampling is equally not feasible in most circumstances. Moreover, random sampling serves one purposes: preventing selection biased. As for randomness testing, it is the data not the sampling method that matters.

The objective of the tests for randomness in this writing are not concerned with sampling methodology, but are concerned on the data itself. The query is "whether the data comes from a random process?" these two issues: random sample and random sampling in relations to data classification is summarized in the table below.

The difficulty presented in this second method is the use of the prior population size. When the population is dynamic, this number can only be an estimate. If time is allowed to lapse and the counting be affected, it is possible to count the population or determine the population size. However, without the benefit of a prior count, the population size may be difficult to determine. In some cases, researchers are faced with unknown population size (non-finite population).

Table 3: Data Produced by Random and non-Random Sampling Methods

	DATA SAMPLE		**Nota Bene**
SAMPLING	Random	Non-Random	*Random sampling does not guarantee a random data.*
Random	*RR*	*RN*	
Non-Random	*NR*	*NN*	

The ideal result of using random sampling method to obtained random data (*RR*) may not be achievable. Even if the sampling method is random, if the data itself comes from a population that is not random, then RN will be produced. For example, if the survey asks about a sensitive issue where members of the public hold different opinions on the issue, i.e. "Do you believe in capital punishment? In this case, the dispersion of the opinion may be wildly and is unpredictable; thus, the data is random (*RR*). Does it mean that the research is good? No, it means that the question solicits divisive responses.

Under the same sampling method: simple random sampling, the data may show non-randomness. For example, in public opinion survey, the selection method may be a simple random selection. However, if the population holds the same or similar opinion on a particular issue, i.e.

"Should there is a free and fair election?" The answer would most likely be {Yes}. In this case, the data is not random. Random by a common definition is lacking predictability. In this case, there is predictability. Does it mean that the research is bad? No, it just means that the data is not a random number set.

Other two scenarios: *NR* and *NN* may also occur. These different types of data results are not indicative of the quality of the research. Therefore, the use of sampling method as the lead indicator for testing randomness is not advisable. The focal point of the examination should be the data itself. Answer the query" "did the data come from a random process?" is more interesting than asking: "was simple random sampling method used?" The focus of this test is on the testing for "randomness in data" not in the sampling method.

4.0 UNIVARIATE DATA

4.1 Randomness in univariate case

In univariate case, the data would have to be codified into binary data form, i.e. 1 and 0 in a (1, 0) string. Univariate case draws on the battery of random tests offered by NIST. NIST battery of tests uses (1, 0) data string is the starting point to verify randomization. In cases where the data is in a form of categorical data, i.e. (Yes | No), the conversion into (1, 0) is a straight forward matter, i.e. (Yes = 1, No = 0). This type of binary data follows binomial distribution. The point of analysis may begin with the Laplace rule of Success to determine pointwise probability. The success is equated to Yes and Yes is equated to 1; the category No = 0. The probability of success is given by:

$$P(s) = \frac{s+1}{N+2} \tag{15}$$

where s = number of success or the counts of score of 1, and N = total number of observations. The probability of non-success or failure is given by:

$$P(f) = 1 - p \tag{16}$$

By convention, the probability of success is denoted as p and the probability of failure (non-success) is denoted as q. The pointwise probability forecast is given by:

$$P(X) = \frac{n!}{(n-X)!X!} p^n q^{n-X} \tag{17}$$

The test statistics follows the Z-equation, thus:

$$Z_{bin} = \frac{\frac{X}{n} - p}{\sqrt{\frac{pq}{n}}} \tag{18}$$

Assume that the data set is comprised of an answer choice in a scalar form: (0, 1, 2, 3). Generally, this form of answer score card is treated as a continuous data and, thus, continuous probability applies. However, in this case, the scale will be treated as dichotomous data (discrete) by dividing the score into zero and non-zero. The zero (0) will be counted as zero and the non-zero scores will be counted as 1. Thus, the scale is converted thus: $(0,1,2,3) \rightarrow (0,[1,1,1,])$. For example, if there are five questions where the answer were: (2, 3, 2, 0,0). The score of non-zeroes were

(2,3,2). These non-zeroes are converted to (1,1,1). Now the new data is in the form of (1,0) where there are through counts of 1 and two counts of 0 or (1,1,1,0,0). With this data form, the value p may be determined, thus:

$$P(s) = \frac{s+1}{N+2} = \frac{3+1}{5+2} = \frac{4}{7} = 0.57 \quad \text{and} \quad P(f) = 1 - p = 1 - 0.57 = 43 \text{ or}$$

$$p = 0.57 \text{ and } q = 0.43 .$$

With known p and q, the test statistics for any give number of X may be determined. Assume that one wants to predict the probability of $X = 3$, the probability be determined, thus:

$$P(3) = \frac{n!}{(n-X)!X!} p^n q^{n-X} = \frac{3!}{(5-3)!3!} 0.57^5 0.43^{5-3} = \frac{6}{2(6)} 0.57^5 0.43^2$$

$$= \frac{6}{12} 0.57^5 0.43^2 = 0.50(0.06)(0.18)$$

$$= 0.50(0.011)$$

$$= 0.0055$$

The test for statistical significance follows:

$$Z_{bin} = \frac{\frac{X}{n} - p}{\sqrt{\frac{pq}{n}}} = \frac{\frac{3}{5} - 0.57}{\sqrt{\frac{0.57(0.43)}{5}}} = \frac{0.60 - 0.57}{\sqrt{\frac{0.2451}{5}}} = \frac{0.03}{\sqrt{0.4902}} = \frac{0.03}{0.2214} = 0.1355$$

where $Z(0.1355)$ is equal to 0.556 or 55.60% chance of happening. The question then remains: Is this data string (1,1,1,0,0) random number or does the data set (1,1,1,0,0) come from a random process? Or does the number come from a random selection?

The question deals with two completely different things. "A number or data coming from a random process" focuses on the data itself. That the data comes from a random process is an inference made after having completed a random test, i.e. NIST or other forms of tests to verify whether the data are random numbers. The null hypothesis is generally stated that the number comes from a random process, if there is a statistical significance finding, it is said that the number did not come from a random process. The rationale for this posture of the null hypothesis is that the term "randomness" has a functional equivalence of equal probability of elements distributed within the sampling space. Thus, if all elements of the population in the sampling space have equal probability of being selected, then it is normally distributed within that space. Within such a space, if the confidence interval is set at, say 0.95 or 0.99, any elements of the data found outside of this predetermined C.I., it may then be concluded that they are non-random because if they are random, they would have been fallen within the C.I. region under the distribution curve.

The second part of the question asks whether the data "comes from a random selection." This second part of the question may not have a definite answer because a selection method or sampling method does not determine the type of data being selected. For instance, if the process under which the sample has been draw was random, i.e. the data was randomly distributed a sampling method will not change this nature of the data. That is to say if the data is randomly distributed, a non-random sampling method would still produce a random data and a random sampling method would also produce random data. The sampling method is not a determining factor. It is the nature of the data *res ipsa* that defines whether the data is random.

4.2 Monost bit test

Monobit test is the first of the 15 binary test for randomness offered by NISt. It is based on frequency count. In this test, there is one sequence or one block where $M = 1$ consisting of data string of 1 and 0. The objective is to compare the string of 1 to the string of 0 and determine whether there is a sensible pattern. If there is a pattern, it means that the sequence is not random. Since the sequence is comprised of 1 and 0, it can be model by a one-dimension random walk. Assume that p is the probability of 1 and $q = 1 - p$. The total probability is given by $p + q = 1$. Asume further that the taking of steps walking to the right is n_1 and the steps taken to the left is n_2. Thus, the total steps taken would be $n_1 + n_2 = N$. Therefore, the probability of taking n_1 steps to the right is given by:

The monobit test is a test based on binomial distribution because it deals with binary data: (1, 0) which falls within the category of Bernoulli random variables. The notation used for even occurrence is X_i; the string of these event which is called data string may be written as: $X_i = \varepsilon$ where $X = 2\varepsilon - 1$. The sequence is given as $S_n = X_1, X_2, ..., X_n = 2(\varepsilon_1, \varepsilon_2, ..., \varepsilon_n) - n$. In binary data analysis, it is necessary to designate the event of interest or category of interest. In the monobit test, the category is interest is 1. The probability of 1 in any time sequence is ½. If the number of observations is sufficient large, according to De Moivre-Laplace theorem, (Feller, 1968; vol. 1, sect. VII.3) the distribution of the binomial sum will approximate normal distribution. Recall that the Moivre-Laplace theorem is given by:

$$\lim_{n \to \infty} P_n \left[a \le \frac{S_n - np}{\sqrt{np(1-p)}} \le b \right] = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^1/2} dx \tag{19}$$

where $a, b \in R \cup \{\pm\infty\}$ and $a < b$ because a is the beginning point and b is he ending point in the integration. The convergence in a and b is uniform.

If $y \in \square, n \in \square$ and the probability of the event lies between 0 and 1: $0 < p < 1$; this condition defines:

$$k(y) = \left\lfloor np + y\sqrt{np(1-p)} \right\rfloor \tag{20}$$

Therefore,

$$\lim_{x \to \infty} \sum_{j=0}^{k(y)} \binom{n}{j} p^j (1-p)^{n-j} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-c} dx \tag{21}$$

where …

$a \quad = -\infty$

$b \quad = np + y\sqrt{np(1-p)}$

$c = \dfrac{(z - np)^2}{2np(1-p)}$

The monobit test is based on the Central Limit Theorem for a random walk: $S_n = X_1, X_2, ..., X_n$. The CLT is given by:

$$\lim_{n\to\infty} P\left(\frac{S_n}{\sqrt{n}} \le z\right) = \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\Delta}^{z} e^{-u^2/2} du \tag{22}$$

If (z) is positive, the probability of the random walk is given by:

$$P\left(\frac{|S_n|}{\sqrt{n}} \le z\right) = 2\Phi(z) - 1 \tag{23}$$

4.3 Frequency (Monobit) test

The frequency test attempts to verify whether the data string is random or came from a random process by looking at the ratio of 1 and 0 to the entire data string.

$$S_n = \sum_{i=1}^{n} X_i \tag{24}$$

where $X_i = (a+b)$ and that $a = \sum 1_i$, $b = \sum -1_i$ and $-1 = 0$. In the example, the string consists of (1,1,1,0,0); thus, $a = \sum 1_i = 1+1+1 = 3$ and $-1 = 0 \to -1, -1$ because there are two counts of zero. Therefore, $b = \sum -1_i = -2$. Now, the value for S_n may be easily determined thus: $S_n = 3 + (-2) = 1$. With known S_n, the observed ratio of 1 may now be determined: S_{obs} which given by:

$$|S_{obs}| = \frac{S_n}{\sqrt{n}} \tag{25}$$

The value for S_{obs} is $S_{obs} = 1/\sqrt{5} = 1/2.24 = 0.4464$.

The test statistic of the frequency or monobit test is given by:

$$Z = \frac{S_{obs}}{\sqrt{2}} \tag{26}$$

From known S_{obs} the calculation for $Z_{monobit}$ follows: $Z = \frac{S_{obs}}{\sqrt{2}} = \frac{0.4464}{1.4142} = 0.3157$. Looking at the Z-table, find the corresponding p-value for $Z(0.3157)$, the value is 0.626. Under 0.95 confidence interval, this number is within the 0.95 confidence interval; there is no statistical significance. Recall that the hypothesis formulation under the frequency or monobit test is:

$$H_0 : (1-\alpha) < 1.65\} \to Random$$
$$H_A : (1-\alpha) > 1.65\} \to Non-Random$$

The null hypothesis is based on: $s = |S_n|/n$. This is compared to the observed value: $|s(obs)| = |= X_1, X_2, +...+ X_n|.\sqrt{n}$. The p-value or the area under the curve, bounded by the lower and upper bounds of the confidence interval (CI), is given by:

$$pValue = 2[1 - \Phi(|s(obs)|)] = erfc\left(|s(ob)| / \sqrt{n}\right) \tag{27}$$

The notation *erfc* stands for the *complementary error function* which is given by:

$$erfc(z) = \frac{2}{\sqrt{\pi}} \int_{z}^{\infty} e^{-u^2} du \tag{28}$$

In this case, $1 - \alpha = 0.626$ which is less than 1.65. Therefore, H_0 (null hypothesis or the assumption of randomness) cannot be rejected or the number string (1,1,1,0,0) came from a random process. Note that the position of the null hypothesis is assuming that the data comes from a random process. The alternative hypothesis is to prove otherwise. This is *counter intuitive* if one wants to prove randomness and the assumption of the alternative hypothesis is "non-randomness." NIST recommends that the sample size for this test is $n = 100$. In the example above, only five observations were used.

4.4 Frequency test within block

In the first case of frequency test, the entire string of data is treated as one block (M); thus, $n = M$. In the second case of frequency test: frequency test within block, the data string is segmented into *M* blocks; each *M* has peculiar patterns which is illustrated by the *Aperiodic Templates* below for a small value of pattern *m* where $2 \le m \le 5$.

$m = 2$	$m = 3$	$m = 4$	$m = 5$
01	001	0001	00001
10	011	0011	00011
	100	0111	00101
	110	1000	01011
		1100	00111
		1110	01111
			11100
			11010
			10100
			11000
			11110

Source: Barbour, and Holst, L., and Janson, S. (1992). Poisson Approximation. Oxford: Clarendon Press. Sects. 8.4 and 10.4. *Cited in* Rukhin A., Soto J., Nechvata J., Smid M., Barker E., Leigh S., Levenson M., Vangel M., Banks D., Heckert A., Drat J. and Vo S. (2010). *A Statistical Test Suite for Random and Pseudo-Randomg Number Generators for Cryptographic Applications*. National Institute of Standards and Technology (NIST); technology Administration, U.S. Department of Commerce. Special Publication 800-22, rev. 1a by Lawrence E. Bassham III, (April 2010). p.3-10.

The "frequency test within block" is still dealing with binary digit (bit). A bit is defined as a block of data that conveys information. Generally, the block is comprised of values of 1 and 0. A block of (1, 0) combination is known as a bit. Eight bits equals one byte. Hence, kilobyte (1,000 x 8 = 8,000 bits), Megabytes (1,000,000 8 = 8,000,000 bits), and gigabytes (100,000,000 x 8 = 8,000,000,000 bits) are referring to the information in binary form of a unit length of eight bits.

The first frequency test under NIST scheme is called monobit test because it involves only one block of bit. Each block is called M-bit block. In the frequency test, the block is a single block or the block size is $M = 1$. In "frequency test within block," the M-bit block is more than 1, i.e. $M > 1$. The purpose of the frequency test within block is to verify whether the frequency of ones within the M-bit block is approximately $M/2$ because the assumption of randomness assumes that the frequency of ones is approximately $M/2$. The following definitions are provided:

ε = bit sequence generated by the random number generator (RNG) or pseudo-random number generator (PRNG);
M = bit block; and
n = length of the bit string.

The test statistic used for verification is the chi-square test. The chi-square measures the ratio of the frequency of ones observed and comparing it to the expected ratio---which is approximately $M/2$ or half of the bit block. The chi-square test statistic is given by:

$$\chi^2 = \sum_{i=0}^{K} \frac{(v_i - N\pi_i)^2}{N\pi_i} \tag{29}$$

Alternatively, equation (2.1) is simplified to:

$$\chi^2 = 4M \sum_{i=1}^{N} \left(\pi - \frac{1}{2} \right)^2 \tag{30}$$

where the term π_i is defined as:

$$\pi_i = \frac{\sum_{j=1}^{M} \varepsilon_{(i-1)M+j}}{M} \tag{31}$$

The p-value for the test is determined by:

$$Pvalue = \int_{\chi^2_{obs}}^{\infty} \frac{e^{-u/2} u^{N/2-1}}{\Gamma(N/2) 2^{N/2}} du = \int_{\chi^2_{obs/2}}^{\infty} \frac{e^{-u} u^{N/2-1} du9}{\Gamma(N/2)} = igamc\left(\frac{K}{2}, \frac{\chi^2_{obs}}{2} \right) \tag{32}$$

This is known as *incomplete gamma function*.

Assume that the following string is given: $\varepsilon = 0110011010$. The length of the string is 10; therefore, $n = 10$. The *first step* is to divide the string into non-overlapping block is called N. The non-overlapping block is defined as $N = \left\lfloor \frac{n}{M} \right\rfloor$. Therefore, the string $\varepsilon = 0110011010$ can be divided into blocks thus:

011 = first N
001 = second N
101 = third N
0 = discarded.

The objective is to divide the data string into blocks. Each block (M) must contain at least one 1 in relations to relative strong of zeroes. How many ones should be in a block? Notice that in the string $\varepsilon = 0110011010$, the separating point for the block is the ending of 1 that would make the shortest block pattern in the string; as for the first data of the next block, it may be 1 o 0.
Therefore, $N = 3$. Note that each block is 3 or $M = 3$; thus, the number of block is $N = \left\lfloor \frac{n}{M} \right\rfloor = \left\lfloor \frac{10}{3} \right\rfloor = \lfloor 3.33 \rfloor = 3$. The 0.33 is disregarded as excess incomplete block. This first step is called *partition input sequence*.

Second step is to determine the proportion of ones in each M-bit block. This proportion is denotes as π_i which is given by equation (5.3):

$$\pi_i = \frac{\sum_{j=1}^{M} \varepsilon_{(i-1)M+j}}{M} \text{ for } 1 \le i \le N \qquad \text{which reads } \pi_i = \frac{\sum_{j=1}^{M} \varepsilon_{(i-1)M+j}}{M} = \frac{Ones}{\text{bit block}}, \text{ from the blocks:}$$

011	=	first N	has	2 ones
001	=	second N	has	1 one
101	=	third N	has	2 ones
0	=	discarded.		

Therefore, π_i includes:
$\pi_1 = 2/3$
$\pi_2 = 1/3$
$\pi_3 = 2/3$

Third step, using equation (5.2.), the value of χ^2_{obs} my be determined, thus:

$$\chi^2 = 4M \sum_{i=1}^{N} \left(\pi - \frac{1}{2} \right)^2$$
$$= 4(3) \left[\left(\frac{2}{3} - \frac{1}{2} \right)^2 + \left(\frac{1}{3} - \frac{1}{2} \right)^2 + \left(\frac{2}{3} - \frac{1}{2} \right)^2 \right]$$
$$= 12[(0.66 - 0.50)^2 + (0.33 - 0.50)^2 + (0.66 - 0.50)^2]$$
$$= 12[(-0.16)^2 + (-0.17)^2 + (-0.16)^2]$$
$$= 12[(0.0256) + (0.0289) + (0.0256)]$$
$$= 12(0.0801) = 0.9612$$

Fourth step, find the critical value for chi-square in the chi-square table at a given degree of freedom and percentage confidence level. The degrees of freedom in the chi-square equation is $df = N = 3$.

For 0.99 confidence interval, the chi-square critical value is 11.30 and for 0.95 confidence interval, the critical value is 7.80. The value for the observation is $\chi^2_{obs} = 0.9621$ which is less than both 7.80 for CI of 95% and less than 11.30 for CI of 99%. Recall that the hypothesis formulation for the random test is:

$$H_0 : \chi^2_{obs} < \chi^2_{0.95} = random$$
$$H_A : \chi^2_{obs} > \chi^2_{0.95} = non - random$$

In this case, it is not possible to reject the null hypothesis. Therefore, it is concluded that the data string $\varepsilon = 0110011010$ comes from a random process.

Equation (30) and equation (31) use K and M degrees of freedom respectively. The following Tables provides the corresponding classes of data string with possible combination of probabilities.

Table 4: *K = 3; M = 8*

Classes	**Probabilities**
$v \le 1$	$\pi_0 = 0.2148$
$v = 2$	$\pi_1 = 0.3672$
$v = 3$	$\pi_2 = 0.2305$
$v \ge 4$	$\pi_3 = 0.1875$

Table 5: *K = 5; M = 128*

Classes	**Probabilities**
$v \le 4$	$\pi_0 = 0.1174$
$v = 5$	$\pi_1 = 0.2430$
$v = 6$	$\pi_2 = 0.2493$
$v = 7$	$\pi_3 = 0.1752$
$v = 8$	$\pi_4 = 0.1027$
$v \ge 9$	$\pi_5 = 0.1124$

Table 6: *K = 5; M = 512*

Classes	**Probabilities**
$v \le 6$	$\pi_0 = 0.1170$
$v = 7$	$\pi_1 = 0.0.2460$
$v = 8$	$\pi_2 = 0.2523$
$v = 9$	$\pi_3 = 0.1755$
$v = 10$	$\pi_4 = 0.1027$
$v \ge 11$	$\pi_5 = 0.1124$

Table 7: *K = 5; M = 100*

Classes	Probabilities
$v \le 7$	$\pi_0 = 0.1307$
$v = 8$	$\pi_1 = 0.2437$
$v = 9$	$\pi_2 = 0.2452$
$v = 10$	$\pi_3 = 0.1714$
$v = 11$	$\pi_4 = 0.1002$
$v \ge 12$	$\pi_5 = 0.1088$

Table 8: *K = 6; M = 10,000*

Classes	Probabilities
$v \le 10$	$\pi_0 = 0.0882$
$v = 11$	$\pi_1 = 0.2092$
$v = 12$	$\pi_2 = 0.2483$
Les 1 - $v = 13$	$\pi_3 = 0.1933$
$v = 14$	$\pi_4 = 0.1208$
$v = 15$	$\pi_5 = 0.0675$
$v \ge 16$	$\pi_6 = 0.0727$

Source: Tables 1 – 5 *see* Pal Revesz (1990). Random Walk in Random and Non-Random Environments, Singapore: World Scientifc. Pp. 55; *cited in* Rukhin et al. p. 3-5.

4.5 Run test for randomness

The Runs Test is considered non-parametric because it does not depend on any parameter as in a polynomial equation. The test examines the substrings of consecutive 1's and 0's. These strings are considered homogeneous. The objective is to look at the oscillation of the substring and verify whether the oscillation is too fast or too slow. The following terms are defined:

V_n = number of runs;

π = fixed proportion, i.e. $\pi = \sum_j \varepsilon_j / n$

The reference point is 0.50; the objective of the test is to determine whether the substring is close to 0.50 or $\left|\pi - \frac{1}{2}\right| \le \frac{2}{\sqrt{n}}$ which is defined by the probability of the runs as:

$$\lim_{x \to \infty} P\left(\frac{V_n - 2n\pi(1-\pi)}{2\sqrt{n}\left(\pi(1-\pi)\right)}\right) \tag{33}$$

In order to evaluate the run (V_n), define that $k = 1, 2, ..., n-1$ and $r(k) = 0$ if the present string is equal to the next string, i.e. $\varepsilon_k = \varepsilon_{k+1}$ and $r(k) = 1$ if $\varepsilon_k \ne \varepsilon_{k+1}$. Then the number of runs is given by:

$$V_n = \sum_{k=1}^{n-1} r(k-1) \tag{34}$$

The pValue or the area under the curve within which is considered non-significance region or the region bounded by the lower and upper boundaries of confidence interval is given by:

$$pValue = erfc\left(\frac{|V_n(obs) - 2n\pi(1-\pi)|}{2\sqrt{2}\left(\pi(1-\pi)\right)}\right) \tag{35}$$

The value of the runs V_n explains the characteristic of the oscillation. If V_n is too large, it means that the oscillation is fast and if the value of V_n is small, it means that the oscillation us slow. The significance tests centers around the observation and test of this oscillation.

The run test (also known as Wald-Wolferwitz test) (Wald and Wolfowitz, 1940; pp. 147-162; Mendenhall *et al.*, 1986) can be used to verify if the data comes from a random process (Bradley, 1068; chap. 12). A "run' is series of increasing values or series of decreasing values; a run is a ranked data set. A ranked data set is a data set arranged in ascending values or descending values. The number of data values in the set is called the "length" of the run. In random data test, the probability that $(l+1)$ is larger or smaller than l^{th} follows binomial distribution---the basis of the runs test. For example, the following is a data set that has been categorized into positive (+) and negative (-) signs:

+ + + + + - - - - - - - + + + - - - - + + + + - - - - - - -
1 2 3 4 5 6

There are 6 runs or $R = 6$ and the length of the data string is 30 counts. The run is counted by counting the times that the sign change from positive (+) to negative (-). In this case, there are 6 alternations. The purpose of the test is to verify that each data point is mutually independent of the others in the set. This mutual independence is an evidence of randomness. If the data in the set are not mutually independent, it means that one event depends on the other or that one event may be used to predict the occurrence of the other, such a condition is said *not* mutually independent and, therefore, is not random. The argument of the null hypothesis is that the data set is random; the alternative hypothesis is that the data is not random.

H_0 : the sequence was produced in random manner
H_1 : the sequence was NOT produced in a random manner

The test statistic for the runs test is given by:

$$Z = \frac{R - \overline{R}}{S_R} \tag{36}$$

where R = number of observed number of runs; $\overline{R}$ = expected number of runs:

$$\overline{R} = \frac{2n_1 n_2}{n_1 + n_2} + 1 \tag{37}$$

S_R = standard deviation of the runs where …

$$S_R^2 = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1+n_2)^2(n_1+n_2-1)} \tag{38}$$

In the alternative, the variance may be written as:

$$\sigma^2 = \frac{(\mu-1)(\mu-2)}{N-1} \tag{39}$$

Recall that $\mu = \bar{R}$ and $S_R^2 = \sigma^2$ and n_1 = number of + and n_2 = number of – event.

The significance level is *alpha*: α. The runs test rejects the null hypothesis if:

$$|Z| > Z_{1-\alpha/2} \tag{40}$$

For large sample is defined as $n_1 > 10$ and $n_2 > 10$. For a large sample, use $Z_{0.95} = 1.65$. The runs test can answer the following question: *Were the sample data generated from a random process?*

For a numerical illustration, the following data set is provided in a pre-classified (=) and (-) form:

+ + + + + - - - - - - - + + + - - - - + + + + - - - - - - -
1 2 3 4 5 6

The objective is to use the Runs Test to verify whether the data string comes from a random process. In this string, there are 6 runs and the size of the data set or the length of the data string is 30 counts or $n = 30$. The counts are divided into two categories: (-) and (-). The next step is to determine the correct value for $n_1 + n_2 = 30$. Assume that $n_1 = +$ counts and $n_1 = -$ counts. Segments 1, 3 and 5 carry (+) signs. There are a total of 12 counts (+) and 18 counts of (-). Therefore, $n_1 = 12$ and $n_2 = 18$.

The second step is to determine the value for $\bar{R}$, thus:

$$\bar{R} = \frac{2n_1n_2}{n_1+n_2} + 1 = \frac{2(12)(18)}{12+18} + 1 = \frac{216}{30} + 1 = 7.2 + 1$$

$$\bar{R} = 8.20$$

The third step is to calculate the standard deviation, by using the variance formula of the Runs test, thus:

$$S_R^2 = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1+n_2)^2(n_1+n_2-1)} = \frac{2(12(18)((2(12)(18)) - 12 - 18)}{(12+18)^2(12+18+1)} = \frac{432(432-12-18)}{900(31)}$$

$$= \frac{432(402)}{27900}$$

$$= \frac{173,664}{27900}$$

$$= 6.2245$$

Recall that standard deviation is the square root of the variance; therefore, the standard deviation for this Runs test is: $S_R = \sqrt{S_R^2} = 2.50$. The test statistic may now be calculated, thus:

$$Z = \frac{|R - \bar{R}|}{S_R} = \frac{6 - 8.20}{2.50} = \frac{2.20}{2.50} = 0.88$$

With the critical value of 0.88, the percentage probability may be looked up in the Z-table. For $Z(0.88)$, the p-value is $1 - \alpha = 0.811$ which is less than 0.95. Recall that the hypothesis statements were:

$H_0 : Z_{obs} < Z(0.95) \rightarrow Random$

$H_A : Z_{obs} \geq Z(0.95) \rightarrow Non - Random$

The critical value for $Z(0.95)$ is 1.65. In this case, the critical value for $Z(0.88)$ is 0.811. Since $Z(0.88) < Z(0.95)$ or $0.811 < 1.65$. The null hypothesis cannot be rejected. Therefore, it is concluded that the data came from a random process.

Run-sequence plots are easy way to graphically summarize a univariate data set. Common assumptions of univariate data set is that they behave like: (i) random drawing; (ii) come from a fixed distribution; (iii) with common location; and (iv)with common scale. The run-sequence allows us to see the shift in location and scale. It also allows us to detect the outliers. The run-sequence is formed by the vertical axis (response variable Y_i) and the horizontal axis ($X_i(1,2,...)$). The questions answered by the run-sequence are: (1) Are there any shifts in location? (2) Are there any shifts in variation? And (3) Are there any outliers?

IV. CONCLUSION

The concept of randomness is the foundation of statistics since statistical test deals with distribution and percentage probability of event occurrence within the distribution. In both types of probability: discrete and continuous, each event in the distribution space is considered random event, hence, equal probability distribution in normal distribution cases. Therefore, since randomness plays an indispensable role in the discussion of statistical tests, it is necessary for researcher to know how to test for randomness in the data. This paper presents three direct tests for randomness in discrete data. In addition, we also present three indirect tests fort randomness through trend analysis. The three direct tests were under NIST approach. The proposed three indirect tests were a mixed of (i) reverse Arrangement Test (RAT), (ii) Military Handbook Test (MHT), and (iii) La Place Trend test (LTT).

REFERENCES

Bradley (1968). Distribution-Free Statistical tests, Chapter. 12.

Feller, W. (1968). An Introduction to Probability Theory and Its Applications (Volume 1). Wiley. ISBN 0-471-25708-7. Section VII. 3.

Horvitz, D. G.; Thompson, D. J. (1952). "A generalization of sampling without replacement from a finite universe." Journal of the American Statistical Association, 47, 663–685.

Perry, John (1953). The Story of Standards, Funk and Wagnalls, Library of Congress Cat. No. 55-11094, p. 123.

Mendenhall, Scheaffer, and Wackerly (1986), Mathematical Statistics with Applications, 3rd Ed., Duxbury Press, CA.

Midzuno, H. (1952)."On the sampling system with probability proportionate to sum of sizes," Ann. Inst. Statist. Math., 3 (1952), pp. 99–107.

Roderick J.A. Little, Donald B. Rubin (2002). Statistical Analysis With Missing Data, 2nd ed., Wiley. ISBN 0-471-18386-5.

Sampath, S. (2005). Sampling Theory and Methods. 2nd ed. Alpha Science International. Harrow: UK. ISBN 1-84265-214-1., pp. 73-74.

Wald, A. and Wolfowitz, J. (1940). "On a test whether two samples are from the same population. Ann. Math Statist. 11, 147-162.