

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/309728666>

Implementation of Data Normality Testing as a Microsoft Excel® Library Function by Kolmogorov–Smirnov Goodness-of-fit Statistics Okeniyi et al (2014) Ibima Conf

Data · May 2014

CITATIONS
0

READS
496

3 authors:



Joshua Olusegun Okeniyi

Covenant University Ota Ogun State, Nigeria

130 PUBLICATIONS 1,716 CITATIONS

[SEE PROFILE](#)



Elizabeth Okeniyi

Covenant University Ota Ogun State, Nigeria

36 PUBLICATIONS 433 CITATIONS

[SEE PROFILE](#)



Aderemi A. Atayero

254 PUBLICATIONS 2,399 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Open Access and Open Education in the Global South [View project](#)



Developing an Entrepreneurial Ecosystem [View project](#)

Implementation of Data Normality Testing as a Microsoft Excel® Library Function by Kolmogorov-Smirnov Goodness-of-fit Statistics

OKENIYI, Joshua Olusegun^{1,*}, OKENIYI, Elizabeth Toyin^{2,**} and
ATAYERO, Aaron Aderemi A^{3,***}

¹Mechanical Engineering Department, Covenant University, P.M.B.1023, Ota, Nigeria

²Petroleum Engineering Department, Covenant University, P.M.B.1023, Ota, Nigeria

³Electrical and Information Engineering Department, Covenant University, P.M.B.1023, Ota, Nigeria

*Corresponding author: joshua.okeniyi@covenantuniversity.edu.ng; Tel: +2348069836502

** lizabeth.okeniyi@covenantuniversity.edu.ng; lizzoken@yahoo.com

*** atayero@covenantuniversity.edu.ng; lizzoken@yahoo.com

Abstract

This paper deliberates on the implementation of data Normality test as a library function in Microsoft Exel® spreadsheet software, in which researchers normally stores data for anlysis and processing, by Kolmogorov-Smirnov goodness-of-fit statistics. The implementation procedure followed algorithmic program development of the Normality Kolmogorov-Smirnov D statistics for the one-sided and the two-sided test criteria as a library function in the Microsoft Excel® environment. For this, the Visual Basic for Applications was employed for deploying macro embedment in the spreadsheet software. Successful implementation of the Normality K-S D statistics fosters the development of the Normality K-S p -value estimation procedure also as a library function in the Ms Excel® environment. Tests of these implementations bear potency of accurate, speedy and economical procedure for undertaking Normality testing in research, for data of up to sample size $n \leq 2000$.

Keywords: Kolmogorov-Smirnov sampling distributions; Kolmogorov-Smirnov probabilistic value; machine precision arithmetic; one-sided one-sample cumulative statistics; two-sided one-sample cumulative statistics; Microsoft Excel® library function

AMS Subject Classifications: 62A05, 62G10

1.0 Introduction

The Normal probability distribution function (pdf) is a well known statistical tool for describing and/or representing data obtained from samples or from population of interests in many disciplines of research including humanity, business, science and engineering (Okeniyi et al, 2013; Omotosho et al, 2012; Banholzer et al, 2006; Ho, 2006; Evans et al, 2000; Bury, 1999). This symmetrical, bell-shaped, continuous statistical distribution could be identified with sample space extending from minus infinity to plus infinity even as it is a standard to which other distributions are compared (Bury, 1999). The mathematical expression of the Normal probability density function is given by (Okeniyi et al, 2013; Iwueze, 2007; Krishnamoorthy, 2006; Soong, 2004; Montgomery and Runger, 2003; Lipson and Sheth, 1973):

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{x-\mu}{2\sigma^2}\right]; \quad \{\forall x: -\infty < x < \infty\}; \{\forall \mu: -\infty < \mu < \infty\}; \{\forall \sigma: \sigma > 0\}$$

(1)

Where: $x \equiv$ given univariate random variable; $\mu \equiv$ the location parameter of the Normal pdf; and $\sigma \equiv$ the scale parameter of the Normal pdf.

Thus Equation (1) is commonly depicted, in literature (Okeniyi and Okeniyi, 2012; Krishnamoorthy, 2006; Shao, 2003; Bilodeau and Brenner, 1999; Bury, 1999), especially for datasets that come from the Normal distribution function, using the shorthand form $N(\mu, \sigma^2)$. Also, the cumulative distribution function (cdf) of the Normal pdf is given by (Okeniyi et al, 2013; Iwueze, 2007; Krishnamoorthy, 2006; Soong, 2004; Montgomery and Runger, 2003, Lipson and Sheth, 1973):

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x \exp\left[-\frac{x-\mu}{2\sigma^2}\right] dx$$

(2)

The Normal pdf finds usefulness in data modelling due to the easy applications of its maximum likelihood estimations of location and scale parameters as the respective measure of central tendency and measure of variability of the dataset of random sample or population (Krishnamoorthy, 2006; Bury, 1999; Lipson and Sheth, 1973). For instance, unlike many other probability distribution functions, the maximum likelihood estimates of the location and of the scale parameters of the Normal

probability distribution function are well known and are easily obtained, respectively, from (Krishnamoorthy, 2006; Bury, 1999):

$$\mu \equiv \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \tag{3}$$

$$\sigma = S \cdot \sqrt{\frac{(n-1)}{n}} \tag{4}$$

The term S from Equation (4) could be obtained from:

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \tag{5}$$

Where \bar{x} is the sample mean, n is the number of data in the given dataset of the univariate random variable x and S is the sample standard deviation of the random variable x .

The ease of computation of the parameters of the Normal distribution and most probably, the direct application of the parameters for drawing requisite conclusions concerning the random sample have led to the Normal pdf being the most applied among the statistical distributions. However, according to ASTM G16–95 R04, the indiscreet application of the Normal pdf modelling to data that distributed in other manner bear potency of grossly erroneous conclusions in the interpretation of prevalent condition represented by the scatter of the data. Avoiding this, therefore, makes it a fundamental requirement of theoretical and empirical research that the researchers test the Normality of the scatter of their data, i.e. ascertain that the data comes from or distributed like the Normal probability distribution function. For instance, Ho (2006) identified tests ascertaining Normality of data as preconditions for employing t -test, factorial analysis of variance and factor analysis among other types of statistical analyses of data. In that text, the use of non-parametric tests like the Mann–Whitney U, Kruskal–Wallis, Wilcoxon Signed Ranks tests were recommended only for data violating the Normality condition.

However, testing the Normality of data is not as simple as the estimation of the location and scale parameters of the Normal distribution. Normality testing requires non-parametric procedures that could be esoteric and which are known as the goodness-of-fit (GoF) testing. Among many methods that had been studied in literature the Kolmogorov-Smirnov GoF had been proffered as the test of choice by researcher for studying continuous distribution, of which the Normal pdf is a widely used example (Breton et al, 2008; Yazici and Yolacan, 2007). In addition, the Kolmogorov-Smirnov GoF statistics is supported by many studies for small sample size (Breton et al, 2008; Yazici and Yolacan, 2007; DeCoursey, 2003) which makes it suitable for engineering data where availability of large sample size may be prohibitive for economical reasons. However, the use of Kolmogorov-Smirnov goodness-of-fit testing for ascertaining that a set of data comes from or distributed like a particular distribution or otherwise requires analytical procedures that could be cumbersome for non-statisticians or non-mathematicians. Soong (2004) exemplified a procedure for a one-sided goodness-of-fit test but that culminated in the use of tables in the text before the Normality criteria of the dataset could eventually be decided. Such methods for large sample sizes, e.g. $n = 1500$, could culminate in drudgery that could dissuade Normality testing, human errors and consequent erroneous interpretation of the scatter of such data.

According to Drucker (2006), a critical principle of innovative culture demanded that an innovation has to be simple for it to be effective. Despite the fact that some authors from the present study had implemented Kolmogorov-Smirnov goodness-of-fit test as a Microsoft Excel® library function with requisite examples and software availability in report submitted elsewhere (Okeniyi and Okeniyi, 2012), the usage of that algorithmic software development remains limited. That implementation necessarily requires the estimation of the Kolmogorov-Smirnov D statistics given by (Okeniyi et al, 2013; Okeniyi and Okeniyi, 2012; Omotosho et al, 2012; Kutoyants, 2011; Weber et al, 2006; Gibbons and Chakraborti, 2003):

$$D_n^- = -\inf_x (F^*(x) - F(x)) \equiv \max_{1 \leq i \leq n} \left(F(X_{(i)}) - \frac{i-1}{n} \right) \quad (6)$$

Which is the one-sided test that $F^*(x)$ is smaller compared to $F(x)$, for sample size n . Also the one-sided test that $F^*(x)$ is larger compared to $F(x)$ need to be obtained from

$$D_n^+ = \sup_x (F^*(x) - F(x)) \equiv \max_{1 \leq i \leq n} \left(\frac{i}{n} - F(X_{(i)}) \right) \tag{7}$$

While the two-sided test which measures maximum absolute distance require the estimation of the statistics

$$D_n = \sup_{-\infty < x < \infty} |F^*(x) - F(x)| \equiv \max \{ D_n^-, D_n^+ \} \tag{8}$$

In that paper, Okeniyi and Okeniyi (2012) described procedure for accurately evaluating Kolmogorov-Smirnov p -values as a Microsoft Excel® library function using machine precision, based on IEEE 64-bit (8-byte) floating-point, arithmetic. This finds embedment as a macro (Simon, 2002), implemented in Visual Basic for Applications®, which could be invoked from the Microsoft Excel® environment, i.e. a common spreadsheet environment where research data are normally stored, processed and analysed.

It is importantly notable that the library functions exist in Excel for the maximum likelihood estimations of the Normal distribution parameters, i.e. Equation (3), Equation (4) and, as may be required, Equation (5). These exist respectively as:

“=AVERAGE(Number arguments corresponding to sample/population)”

“=STDEVP(Number arguments corresponding to population)”

“=STDEV(Sample arguments corresponding to sample)”

Also, Equation (1) and Equation (2) could be estimated in Microsoft Excel® through the usage of requisitely invoked condition in the library function given by:

NORMDIST(x, μ, σ , pdf or cdf condition)

In spite of these, however, there is no function in the highly important data storage, data processing and data analysing spreadsheet software for estimating the Kolmogorov-Smirnov statistics, i.e. Equation (6), Equation (7) and Equation (8).

The existence of such library function, especially based on the cumulative density function of the Normal distribution would have assisted the use of the procedures described in Okeniyi and Okeniyi (2012) for highly simplified and quicker tests of Normality of sample data. This is potent at facilitating focus of researcher on interpretation of their data with the added advantage of saving cost and time that could have been expended on purchasing and learning to use additional software in the stead of the cumbersome use of tables for Normality testing. Thus, this study deliberates on the implementation of data Normality test by the Kolmogorov-Smirnov goodness-of-fit statistics in the environment of Visual Basic for Applications® (VBA®) that could be invoked as a library function within the Microsoft Excel® software.

2.0 Program Implementation Methods

2.1 Algorithm for K-S D

The algorithmic procedure employed for implementing the K-S statistics in this study include modification of the descriptions itemised for the one-sided D^+ estimation by Soong (2004). By this the algorithmic procedure employed for the VBA® program implementation include:

- Capturing dataset of univariate data, x , from the Microsoft Excel® spreadsheet environment;
- Re-arrangement of captured data into an increasing ordered statistics of univariate data, $x_{(i)}$;
- Evaluating the maximum likelihood estimates of the location parameter, μ , and the scale parameter, σ , of the captured;
- Determination of the observed Normal cumulative distribution function, $F^*(x_{(i)})$, at the instances of each $x_{(i)}$ as required for D^- or D^+ and/or for D_n ;
- Determination of the theoretical Normal cumulative distribution function, $F(x_{(i)})$, at the instances of each $x_{(i)}$ from the ordered statistics of univariate data;

- Evaluation of D^- as per Equation (6) or D^+ as per Equation (7) and/or D_n as per Equation (8); Determination of the observed Normal cumulative distribution function, $F^*(x_{(i)})$, at the instances of each $x_{(i)}$ as required for D^- or D^+ and/or for D_n ;

2.2 Implementing K-S D in VBA

Implementation of K-S D statistics in VBA followed the flow chart in Figure 1. For a library function deployment, the data to be captured depend on the input of the user which is expected to be an array of univariate data. For this, the number of input data specified by the user was queried in this implementation through use of the “Application” object in conjunction with the “WorksheetFunction” property and the “Count” method by:

```
KountCells = Application.WorksheetFunction.Count(DatArray)
```

This find usefulness for re-arranging the input data by the user through requisitely assigned variable for storing the input data in the VBA program implementation, tagged “DatStor”. The re-arrangement procedure was implemented by the “bubble sorting” code (Gottfried, 2001) that invoked a data interchanging modular program “InterChangDat” in the form:

```
For i = 1 To KountCells - 1
    For k = i + 1 To KountCells
        If DatStor(i) > DatStor(k) Then
            Call InterChangeDat (DatStor(i), DatStor(k))
        End If
    Next k
Next i
```

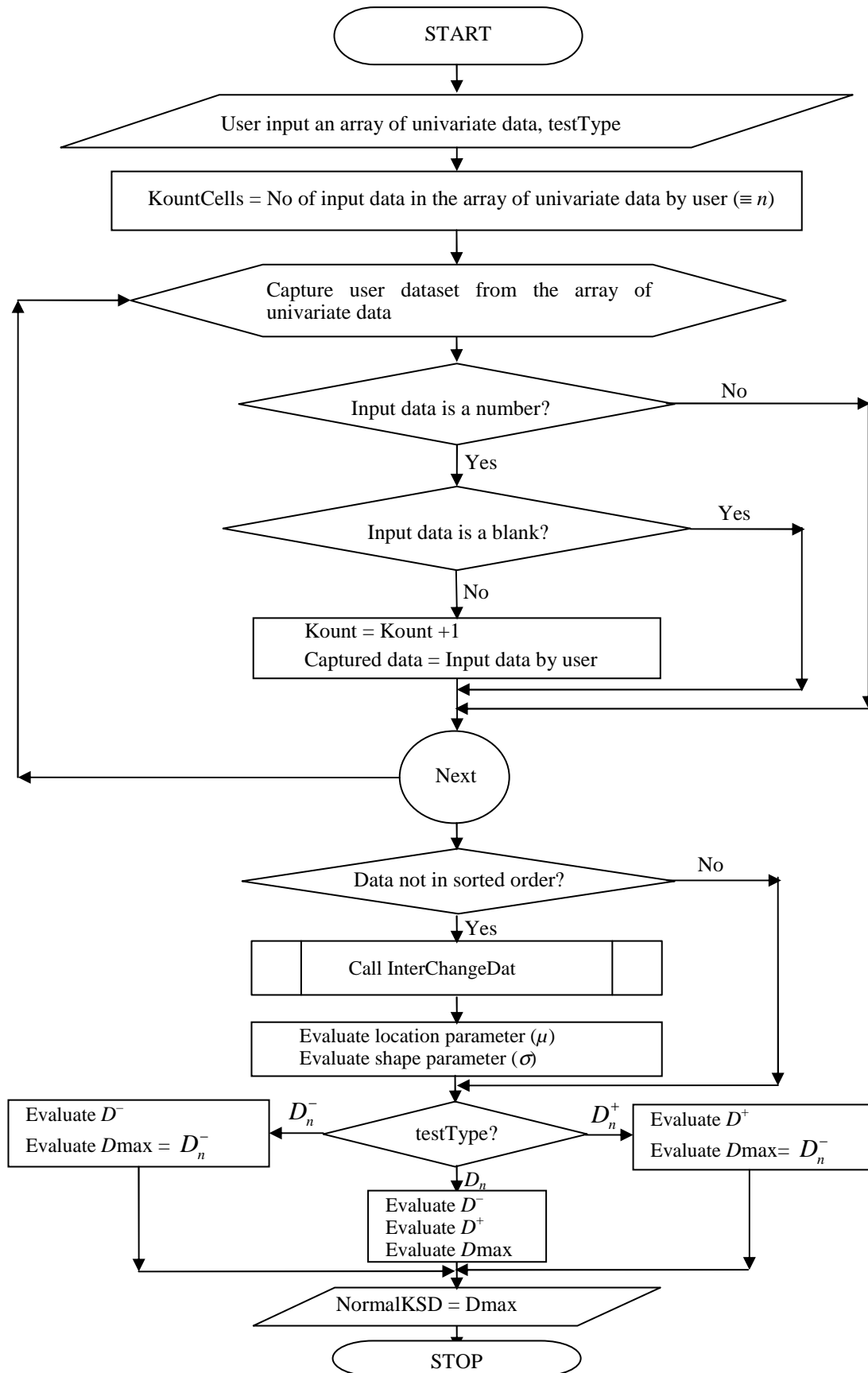



Figure 1: Program Flowchart for the Evaluation of Kolmogorov-Smirnov D statistics

Maximum likelihood estimate of the location parameter and the scale parameter were evaluated by the respective VBA statements that employed the “Average” and the “Stdev” methods, as appropriate, in the form of:

```
NormMu = Application.WorksheetFunction.Average(DatArray)
```

```
NormSig = (Application.WorksheetFunction.StDevP(DatArray))
```

For each instances of ordered $x_{(i)}$ data, the absolute differences were evaluated through:

```
DMinus(i) = NormCDF(i) - (i - 1) / (KountCells)
```

```
Dplus(i) = (i / (KountCells)) - NormCDF(i)
```

Where “NormCDF(i)” represent the theoretical Normal cumulative density function evaluated by the “NormDist” method in accordance with:

```
NormCDF(i) = Application.WorksheetFunction.NormDist(DatStor(i), NormMu, NormSig, True)
```

The maximum absolute distance was obtained through an interchange program code implementation given, for D^+ (i.e. Dplus) as an instance, by:

```
DMax = Dplus(1)
```

```
For i = 2 To KountCells
```

```
  If DMax < Dplus(i) Then
```

```
    Temp = DMax
```

```
    DMax = Dplus(i)
```

```
    Dplus(i) = Temp
```

```
  End If
```

```
Next i
```

The “Dmax” obtained from this constitutes the K-S D , for this instance D^+ , that could be employed for evaluating K-S p -value from the procedure reported by Okeniyi and Okeniyi (2012).

2.3 Mimicking Ms Excel® function with array argument operand

It is of interest to mimic or imitate Ms Excel® function having array operand as argument. Examples of these include the “Sum” function and the “Average” function. These functions find capability at ignoring containing blanks and/or non-numeric values even as input data for these functions could span many columns. For effecting recognition and exclusion of non-numeric part of the input array by the user was implemented through the conditional statement:

If Not IsNumeric(DatArray.Cells(i, j).Value) Then kountIt = kountIt - 1

While the exclusion of the blank segment was effected through use of “ElseIf” program segment in the form of:

Elseif DatArray.Cells(i, j).Value <> "" Then DatStor(kountIt) = DatArray.Cells(i, j).Value

2.4 The Kolmogorov-Smirnov D library function for the Normal distribution

The foregoing was implemented as macro function of the Visual Basic for Applications® (VBA®) of Microsoft Excel through the declaration:

Public Function NormalKSD(DatArray As Variant, testType As Integer)

Where “DatArray” was conceived as the argument to capture user input of univariate data that could be optionally multi-column but single area or single arrayed. The argument “testType” could assume the any of three forms: testType = 0, for K-S D^- evaluation; testType = 1 for K-S D^+ evaluation; and testType = 2 for K-S D_n evaluation.

The function returns value of “Dmax” through the program coding:

NormalKSD = DMax

By these, the function “NormalKSD” could be invoked right from within Microsoft Excel® spreadsheet environment, in similitude to invoking an existing library function, by:

=NormalKSD(DatArray, testType)

2.5 Application of “NormalKSD” implementation for Normality K-S p -value

Ultimately, motivation for computing K-S D statistics was for estimating the K-S p -value on which the decision of normality would be based as per identified confidence interval/level of significance (Okeniyi and Okeniyi, 2012, Soong, 2004). Due to this, the implementation of the “NormalKSD” from this study was applied for directly estimating Kolmogorov-Smirnov p -value, for Normality testing of data, from dataset in the Microsoft Excel® spreadsheet environment. This find proposition as a library function that could be invoked as:

=NormalKSP(ArrayDat, ksdType)

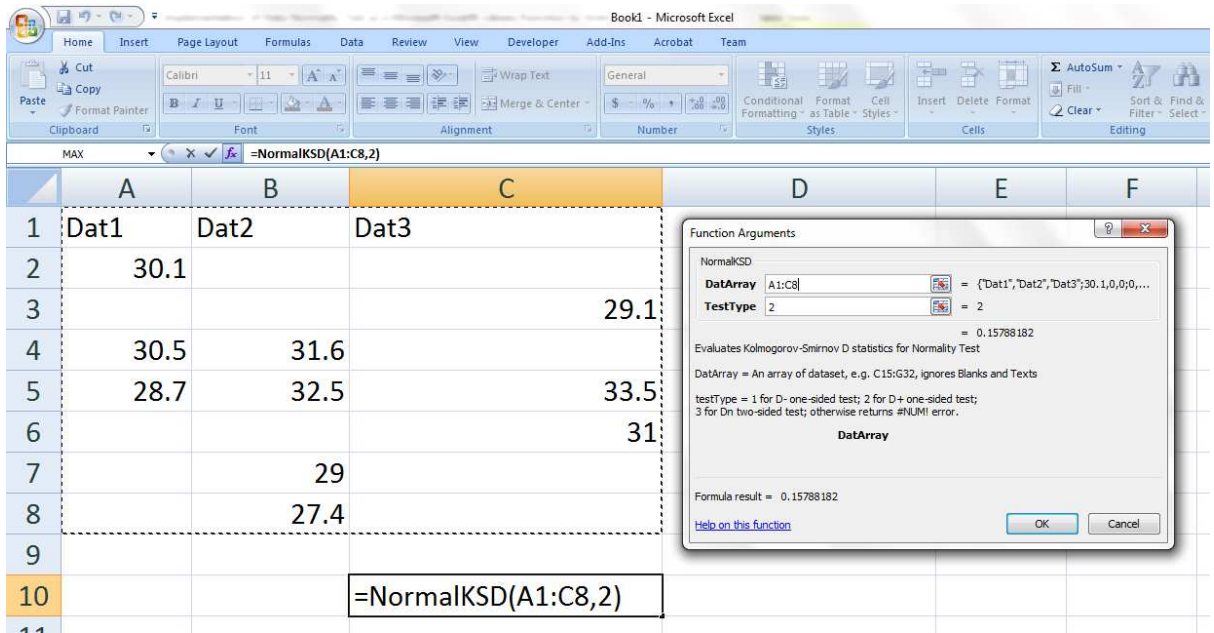
3.0 Results and Discussion of implemented program testing

The NormalKSD program implementation was applied from within Microsoft Excel® version 2007 spreadsheet environment. This was running on Hp Pavilion dv6 Notebook PC with Intel® Core™ i3 CPU M330 @ 2.13GHz, 2133MHz, 2 Core(s), 4 Logical Processor(s) and which employed Microsoft® Windows® 7 Home Premium as the operating system. The data used for this test-application were the same data employed by Soong (2004) for exemplifying computations of K-S D^+ (tagged d_2) by statistical tables and graphical methods.

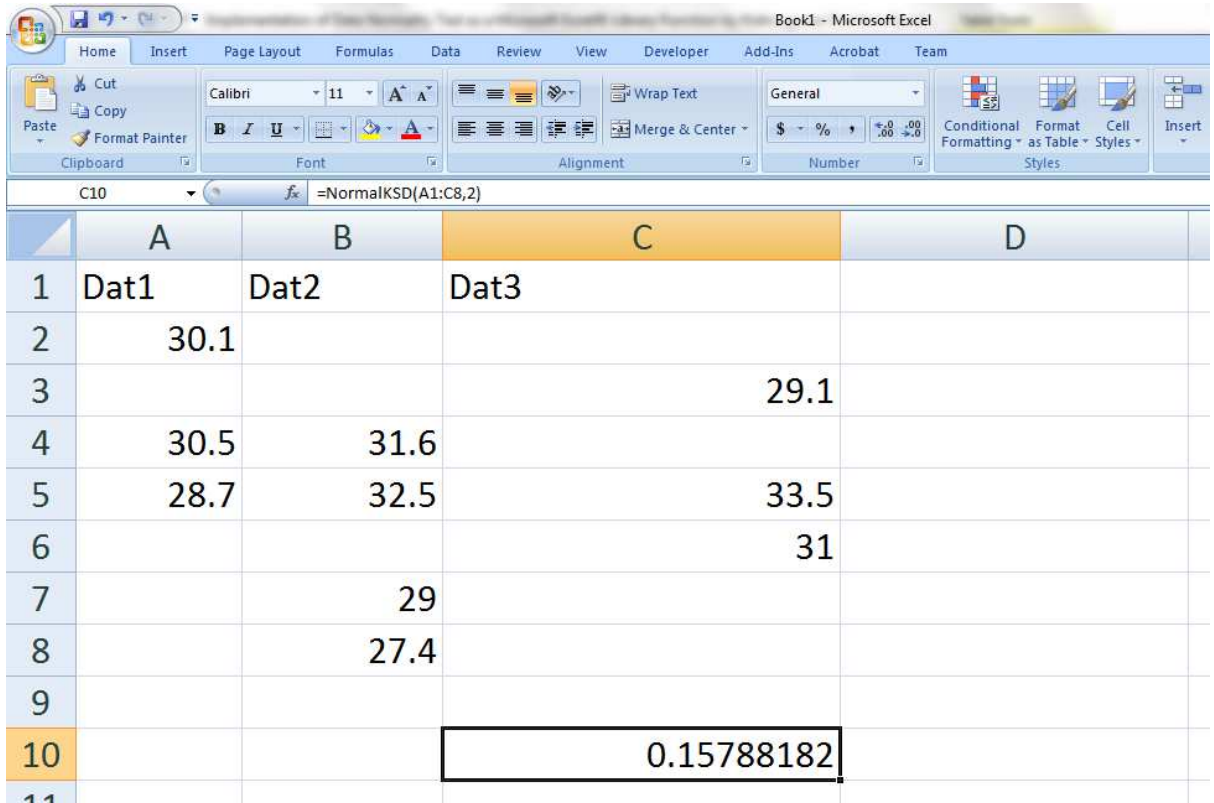
The direct computation of K-S D^+ of this dataset using the “NormalKSD” library function proposed in this study was as shown in Figure 2. In that Figure, the data were scattered for exemplifying multicolumn dataset, ignoring of cells with text(s) and blank(s) cells without data by the implemented library function, Figure 2(a). The result of the “NormalKSD” on the Ms Excel® was shown in Figure 2(b) as = 0.15788182 correct to 8 decimal places (8 d.p.). This result represents more accurate estimations compared to $d_2 = 0.1517$ that would have exhibited cumulative error due to the 4 d.p. obtained from the tables employed in Soong (2004).

Figure 3 showed the direct computation of K-S p -value in the Microsoft Excel® environment from the multicolumn dataset using the implemented “NormalKSP” library function. This direct computation gave the result of the Normality K-S p -value = 0.55271852 (correct to 8 d.p.). This result is > 0.05 significant level of α . By this conclusion bearing agreement with that from Soong (2004) on the data include

the direct deduction that it could not be rejected that the dataset distributed like or come from the Normal probability distribution function.

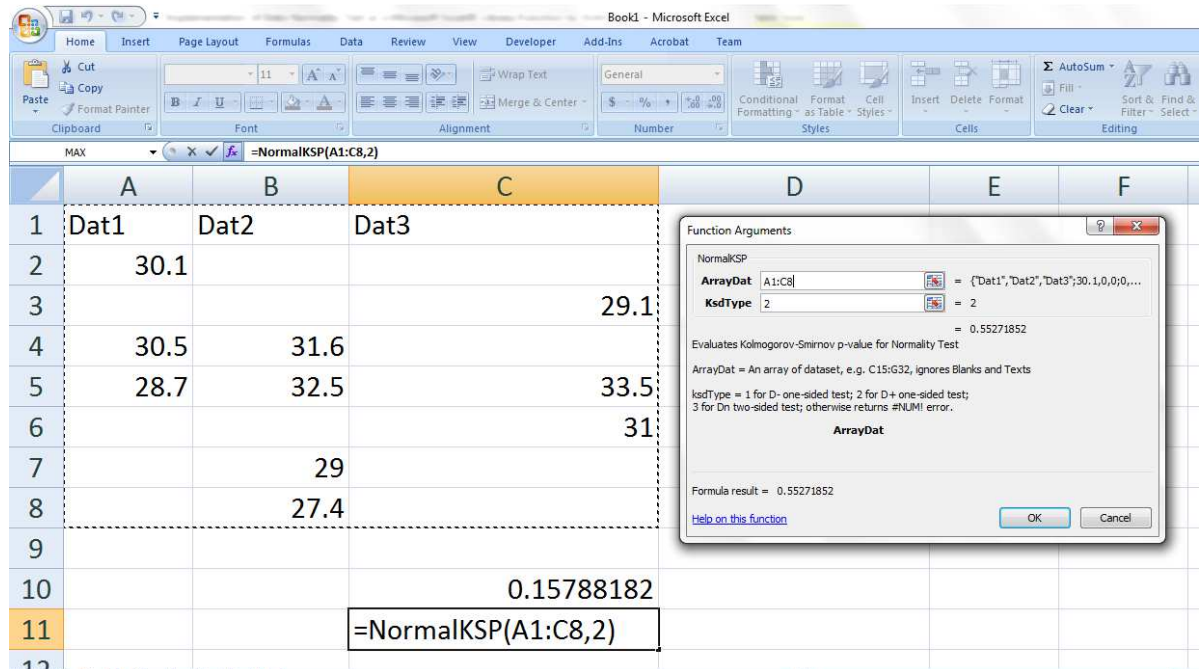


(a)

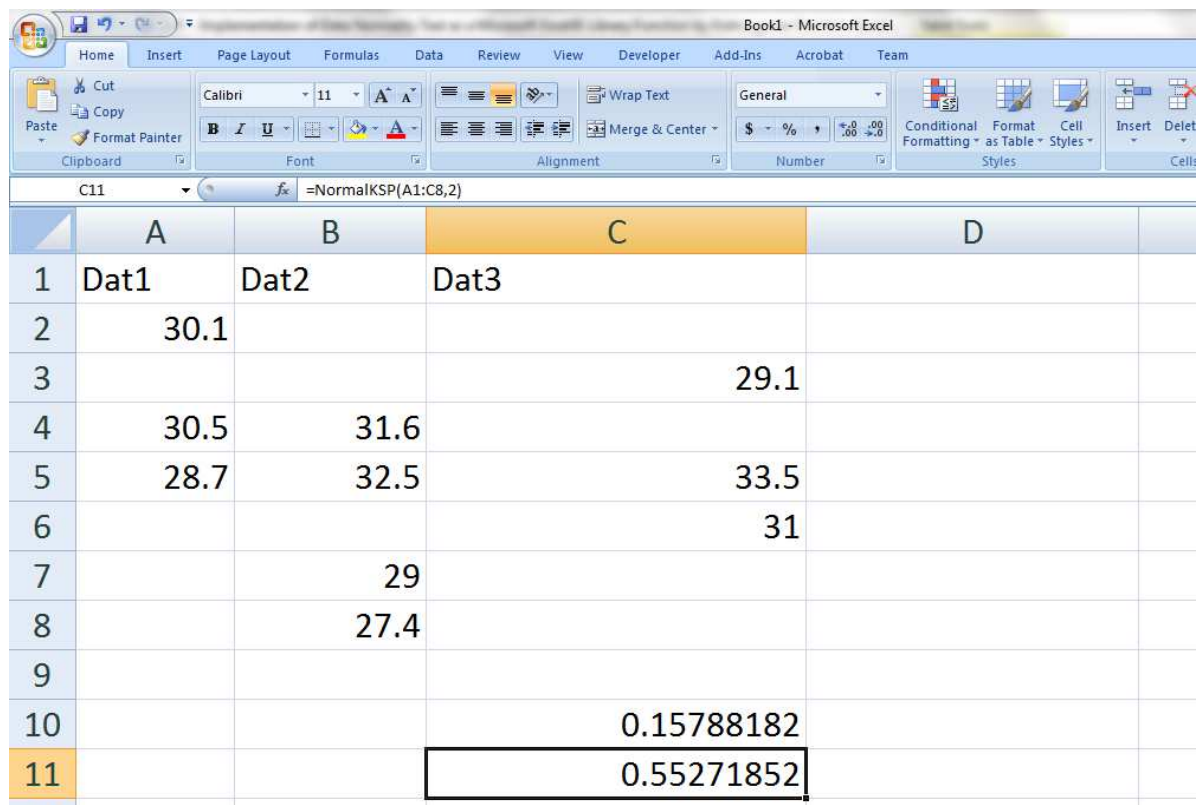


(b)

Figure 2: Evaluating Normality K-S D^+ (i.e. d_2) in Ms Excel® using the data from Soong 2004 in scattered form (a) selecting the array of multicolumn dataset with texts and blanks (b) evaluated Normality K-S D^+ by the implemented “NormalKSD” library function



(a)

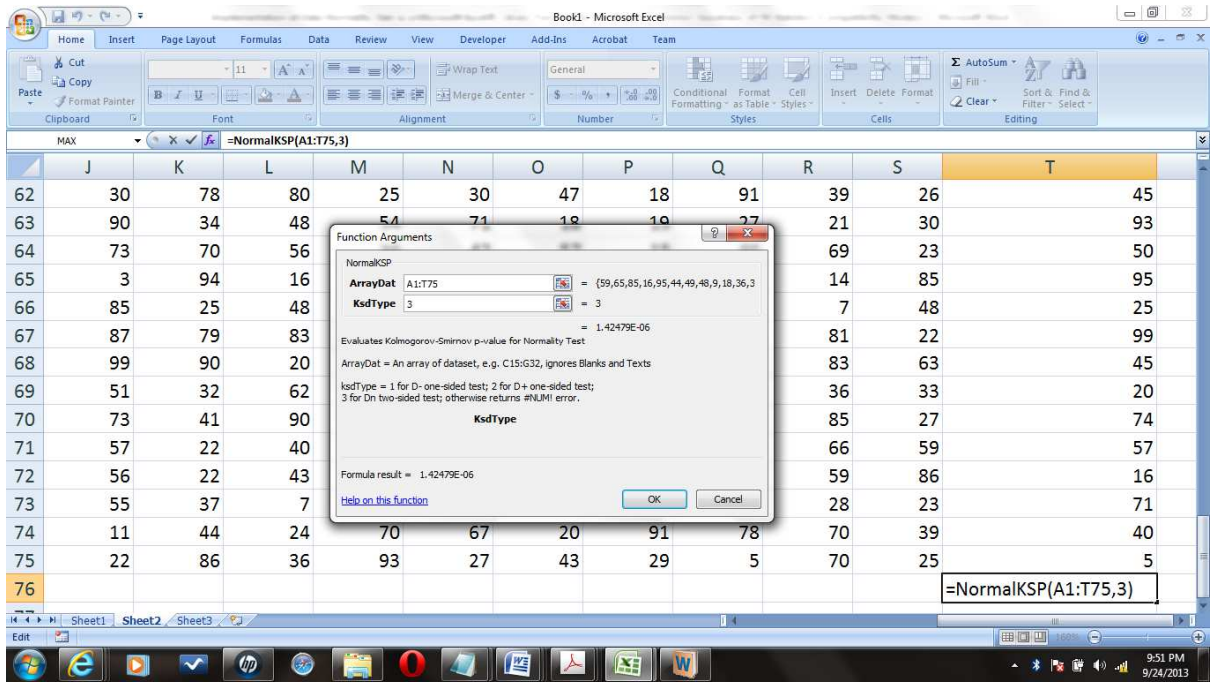


(b)

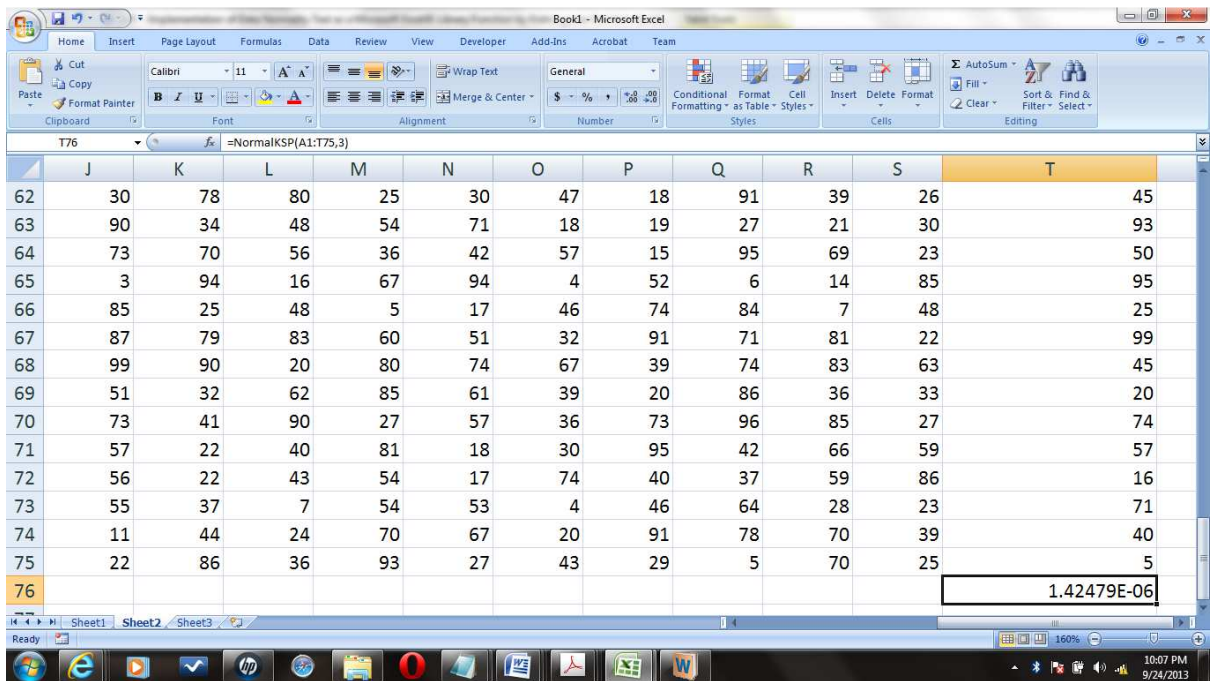
Figure 3: Evaluating Normality K-S p -value in Ms Excel® using the data from Soong 2004 in scattered form (a) selecting the array of multicolumn dataset with texts and blanks (b) evaluated Normality K-S p -value by the implemented “NormalKSP” library function

However, benefit of this library function implementations could be better understood from the application of the quick ascertaining of the random data between 1 and 99, using the “RandBetween(1,99)” library function but in 20 columns \times 75 rows = 1500 data. Figure 4(a) showed the “NormalKSP” application to this dataset, at “A1:T75” of the Excel sheet, while the result was as shown in Figure 4(b) as Normality K-S p -value = 1.42479×10^6 . This is < 0.05 such that it could be quickly deduced that the scatter of the data did not distributed like or come from the Normal probability distribution function.

This quick approach at ascertaining Normality of dataset makes the implementation of the “NormalKSD” statistics and the subsequent “NormalKSP” to constitute improvement over the methods described in Okeniyi and Okeniyi for evaluating Normality K-S p -value. For instance, the procedure for Normality K-S p -value estimations described in Okeniyi and Okeniyi (2012) requires the number of univariate data, “ n ”, computation of K-S D statistics, “ d ”, and the sidedness of the K-S GoF test, “num_sided” as function arguments. By the implementation in this study, the rigorous estimation of Normality K-S D statistics is avoided with the added advantage of accurate K-S p -value estimation for facilitating faster decision of Kolmogorov-Smirnov goodness-of-fit criteria. Also, that the method from Okeniyi and Okeniyi (2012) on which the estimation of the Normality K-S p -value were based could be applied to dataset of up to $n \leq 2000$ data constitute further advantage of modelling comparably large univariate data.



(a)



(b)

Figure 4: Application of the “NormalKSP” library function for quickly ascertaining Normality of $n = 20 \times 72 \equiv 1500$ data (a) selecting the array of 20 columns and 75 rows (b) evaluated Normality K-S p -value for the large univariate data

4.0 Conclusion

Implementation of Data Normality test as a Microsoft Excel® library function by the Kolmogorov-Smirnov goodness-fit-statistics had been deliberated on in this study. From this the following conclusion can be drawn:

- Normality test of data is required for many statistical analyses as well as for describing data of samples from many disciplines;
- Computation of K-S D statistics for Normality test of data using conventional methods could be prone to errors, drudgery and may be outright esoteric such that it could dissuade non statistical/non-mathematical researchers for conducting the highly important Normality test;
- Implementation of the Normality K-S D statistics through the proposed “NormalKSD” library function in Microsoft Excel® is potent with accurate, speedy and economical estimation of the important absolute maximum difference for K-S GoF statistics;
- The “NormalKSD” implementation fosters the “NormalKSP” library function deployment that facilitate quick and accurate estimation of Normality K-S p -value from within the Microsoft Excel spreadsheet environment, where researchers normally store data for analysis and further processing without, need for purchasing or learning how to use other types of software for ascertaining data Normality.

References

- Banholzer B., Brockmann, T. and Brameshuber W. 2006. Material and bonding characteristics for dimensioning and modelling of textile reinforced concrete (TRC) elements. *Materials and Structures* **39**: 749–763.
- Bilodeau, M. Brenner, D. 1999. *Theory of Multivariate Statistics*. Springer-Verlag, New York.

- Breton, M. D., Devore, M. D. and Brown, D. E. 2008. A tool for systematically comparing the power of tests for normality, *Journal of Statistical Computation and Simulation*, **78**(7): 623–638.
- Bury, K. 1999. *Statistical distributions in engineering*. Cambridge University Press.
- Evans, M., Hastings, N., and Peacock, B. 2000. *Statistical distributions – 3rd Edition*. John Wiley and Sons, Inc., New York.
- Gibbons J. D. & Chakraborti S. 2003. *Nonparametric Statistical Inference, Fourth Edition, Revised and Expanded*, Marcel Dekker Inc., New York.
- Gottfried, B. S. 2001. *Schaum's Outline of Theory and Problems of Programming with Visual Basic*, McGraw-Hill, Inc.
- Ho, R. 2006. *Handbook of univariate and multivariate data analysis and interpretation with SPSS*. Taylor and Francis Group, LLC, Boca Raton, FL.
- Krishnamoorthy, K. 2006. *Handbook of statistical distributions with applications*. Taylor and Francis Group, LLC, Boca Raton, FL.
- Kutoyants, Y. A. 2011. Goodness-of-fit tests for perturbed dynamical systems, *Journal of Statistical Planning and Inference*, **141**: 1655–1666.
- Lipson, C. and Sheth, N. 1973. *Statistical design and analysis of Engineering Experiments*. McGraw-Hill, Inc., New York.
- Montgomery, D. C. and Runger, G. C. 2003. *Applied Statistics and Probability for Engineers, 3rd edition*, John Wiley & Sons, Inc., New York.
- Okeniyi, J. O. and Okeniyi E. T. 2012. Implementation of Kolmogorov-Smirnov *p*-value computation in Visual Basic®: implication for Microsoft Excel® library function. *Journal of Statistical Computation and Simulation*, **82**: 1727-1741.

- Okeniyi, J. O., Obiajulu, U. E., Ogunsanwo, A. O., Odiase, N. W. and Okeniyi, E. T. 2013. CH₄ emission model from the waste of *Sus Domesticus* and *Gallus Domesticus* in Nigerian local farms: environmental implications and prospects. *Mitigation and Adaptation Strategies for Global Change* **18(3)**: 325–335.
- Okeniyi, J. O., Obiajulu, U. E., Ogunsanwo, A. O., Odiase, N. W. and Okeniyi, E. T. 2013. CH₄ emission model from the waste of *Sus Domesticus* and *Gallus Domesticus* in Nigerian local farms: environmental implications and prospects. *Mitigation and Adaptation Strategies for Global Change* **18(3)**: 325–335.
- Omotosho, O. A., Okeniyi, J. O., Ajayi, O. O. and Loto, C. A. 2012. Effect of synergies of K₂Cr₂O₇, K₂CrO₄, NaNO₂ and aniline inhibitors on the corrosion potential response of steel reinforced concrete in saline medium. *International Journal of Environmental Sciences*, **2(4)**: 2346–2359.
- Shao, J. 2003. *Mathematical statistics, 2nd Edition*, Springer Science+Business Media, LLC.
- Simon, J. 2002. *Excel Programming: Your visual blueprint for creating interactive spreadsheets*, Hungry Minds, Inc.
- Soong, T. T. 2004. *Fundamentals of Probability and Statistics for Engineers*, John Wiley & Sons, Ltd.
- Weber, M. D., Leemis, L. M. & Kincaid, R. K. 2006. Minimum Kolmogorov-Smirnov test statistic parameter estimates, *Journal of Statistical Computation and Simulation*, **76(3)**: 195–206.
- Yazici, B. and Yolacan, S. 2007. A comparison of various tests of normality, *Journal of Statistical Computation and Simulation*, **77(2)**: 175–183.