MULTIPURPOSE SAMPLE GENERATOR (MSG)

Waheed Babatunde Yahya (Ph.D.)

Mahmud Muhammed

USER'S MANUAL

MULTIPURPOSE SAMPLE GENERATOR (MSG)

VERSION 1.0.0

A Statistical Software For Generating Random Samples of Different Forms

Developed by:

Waheed Babatunde Yahya (Ph.D.)
Mahmud Muhammed

Department of Statistics, University of Ilorin, PMB 1515, Ilorin, Nigeria.
E-mail: **wbyahya@unilorin.edu.ng**; **mahmud4hadji@gmail.com**

USER'S MANUAL

PREAMBLE

This manual presents how to use Multipurpose Sample Generator (MSG) version 1.0.0 to generate random samples of different forms for further statistical analyses and usage.

SYSTEM REQUIREMENTS AND INSTALLATION

The MSG software can be installed on any 32 or 64 bits windows XP, windows Vista, windows 7 and 8 operating systems. A minimum of 512MB of RAM and at least 1GB hard disc space is required for optimal performance. In addition the system must have foxit-reader installed on it before the user can access the manual. The foxit reader setup can be downloaded free at: http://www.foxitsoftware.com/Secure_PDF_Reader/

To install MSG 1.0.0, simply double-click on the MSG setup file, a windows installer would pop up, click "**next**" to proceed, select users ("Just Me" or "Everyone") and choose installation folder, click "**next**" to finalize your installation. After the installation, the shortcut to MSG icon will be deposited on the desktop and in the user's program menu. Click on this icon to launch the MSG software.

The MSG 1.0.0 setup can be downloaded free at www.unilorin.edu.ng

CITATION INFORMATION

Any user of the MSG software for research purposes should reference the software using the following format:

Yahya, W.B. and Muhammed, M. (2014): MSG: A Multipurpose Sample Generator for generating random samples of different forms. *Department of Statistics, University of Ilorin, Nigeria.* http://www.unilorin.edu.ng/

Table of Contents

1.   INTRODUCTION

The detail of how to use the Multipurpose Sample Generator (MSG 1.0.0) software to generate random numbers and random samples of different forms is presented.

2.   HOME

The home page contains the title bar, the menu bars and a brief description of the software developers/authors. Apart from the home page, the menu bars provide the links to "`Random Number`", "`Simple Random Sampling`" and "`Stratified Random Sampling`" generation procedures which are the major tasks performed by the MSG package. Also, the menu bar has a link to view the user's manual and another one that provides detail address of the developers.

3.   RANDOM NUMBERS

Random numbers are numbers that evolve randomly by chance factor whose values can fall within a specified range of interest and their formation lacks any specific pattern.

3.1   *Random Numbers Generator*

To generate a set of random numbers using MSG software, the user is required to click "`Random Number`" on the menu bar and specify the range within which the random numbers should be generated using the "`From`" and "`To`" boxes. The user may specify a combination of numbers between 0 and 9 or a combination of characters from A to Z (either lower or upper case or both) in the "`prefix`" box that would precede and uniquely identify all the numbers to be generated.

The number of samples (n) to be generated should be specified in the "`Count`" box after which the user would click on the "`Generate`" button to view the random numbers generated. A dialogue box carrying the message "`Generated Successfully`" would appear immediately after the MSG has completed the process of generating the required random numbers and it would disappear by clicking the "`OK`" button on this box. The random numbers generated (results) would be displayed on the "`output`" box. A click on the "`Save`" button would automatically save this result after which the results would be re-displayed on the right panel of the MSG page. A click on the "`Print`" button would send this result to the selected printer. A click on the "`Refresh`" button would clear this result from the panel.

For users of MS office 2013 version, a click on the "`Export`" button would export this result into "`My Document`" directory of the computer system as a Microsoft Excel file with Comma Separated Values (.csv) extension. For users of lower versions of MS office, the default printer of the computer system has to be changed to either "*OneNote*" or "*Microsoft XPS*" or "*Xml*" as the default printer. A click on the printer icon on the "`Print Preview`" dialogue box would

send the results to "*OneNote*" or "*Microsoft XPS*" or "*Xml*" as selected after which the results would be displayed for further usage.

The default printer of a computer system can be changed by checking on the "`view devices and printers`" in the control panel. Right click on the preferred printer to set it as default for usage by checking on the "`Set as default printer`" option.

4.   SIMPLE RANDOM SAMPLING

The simplest form of sampling is the simple random sampling (SRS). It is a sampling technique in which a subset of individuals or units (a sample) is randomly selected from a larger set (population) such that all members in the sample set has equal chance (probability) of being chosen into that set. As defined by Moore and McCabe (2006) and Moore (2007), a simple random sample of size n consists of n individuals from the population chosen in such a way that every set of n individuals has an equal chance to be in the sample actually selected.

4.1   *Simple Random Sample Generator*

Two forms of simple random samples can be generated using MSG software. One form is to generate only numeric SRS. Here, all members in the population from which the sample is to be selected have numerical representations. Here, all the elements in the population are assigned a sequence of numbers that uniquely identify them (with a difference of one unit between the successive number labels such as 1, 2, … , N, N being the total number of units in the population). A simple random sample of size n, $n \leq N$ can then be drawn from such a sequence of numbers while the sampling units carrying the selected number labels constitute the required sample from such population.

The second form is to generate SRS from a population whose elements have alpha-numeric or pure character representations.

i.   **Numeric SRS Generator:** To generate numeric simple random sample, the user is required to click on the "`Simple Random Sampling`" option on the menu bar. From the left panel labeled "`With Numbers`", specify the range within which the random samples should be selected using the "`From`" and "`To`" boxes. This range is the interval of the sequence of number-codes used to represent the original sampling units in the population. The number of samples (n) to be generated should be specified in the "`Count`" box after which the user would click on the "`Generate`" button to generate the required samples. Other steps on how to view and print the results generated as provided in Section 3.1 then follow from here.

ii. Character or Alpha-Numeric SRS Generator: For a data set that contains characters or combinations of characters and numbers (alpha-numeric), the user is required to input the data into the "`Input Data`" box on the "`With Characters`" interface of the MSG software. A click on the "`Add`" button would store the data inputted into to the "`Data`" panel after which the next data entry can be made again through the "`Input Data`" box. This process would continue until the last data point in the population is entered. All the data entry can be viewed by clicking on the scroll arrow on the "`Data`" panel. Specify the number of samples to be selected in the "`Count`" box and click on the "`Generate`" button to generate the required samples. Other steps on how to view and print the results generated as presented in Section 3.1 then follow from here.

However, it should be noted that this interface can as well be used to generate SRS from populations that contain only numeric sampling units of any form of sequence.

5. STRATIFIED RANDOM SAMPLING

Stratified random sampling is a variation of random sampling in which the population is partitioned into *h* homogenous sub-groups called strata and a (simple or systematic) random sample is selected from each stratum. The results from the strata are then aggregated to make inferences about the population.

In computational statistics, stratified sampling technique is employed to enhance the efficiency of the statistics computed from the sample data and reduce the sampling error. A major advantage of stratified random sampling is that of being the most representative of a population (Särndal et al, 2003).

5.1 *Stratified Random Sample Generator*

Like in the SRS, two forms of the stratified random samples can be generated using MSG software. These are numeric and alpha-numeric stratified random samples. To use the numeric stratified random samples interface of the MSG software, all members in the population from which the sample is to be selected within each stratum must have numerical representations in the form of sequence of number labels that uniquely identify them as described in Section 3.1 for SRS generator. The alpha-numeric stratified random samples generator interface generates stratified random samples from a population whose elements have alpha-numeric or pure character representations within each stratum.

i. Numeric Stratified Random Sample Generator: To generate numeric stratified random sample, the user is required to click on the "`Stratified Random Sampling`" option on the menu bar. From the left panel labeled "`With Numbers`", select the stratum from

which the random samples are to be generated among the five options "**Stratum 1**" to "**Stratum 5**" on the interface and specify the range within which the random samples should be selected using the "**From**" and "**To**" boxes. Click on "**Add**" button to send the numbers within the specified range into the data panel. Specify the number of samples (n) to be generated in the "**Count**" box and click on the "**Generate**" button to generate the required samples for that stratum. Repeat this process to generate samples from other strata.

If more than five strata are desired, simply click on "**clear**" option beside the data panel of strata 1 to clear the initial data and repeat the above process beginning from this stratum to make-up to the number of strata desired. Other steps on how to view and print the results generated as provided in Section 3.1 then follow from here.

ii. **Character or Alpha-Numeric Stratified Random Sample Generator:** For data set that contains characters or combinations of characters and numbers (alpha-numeric), select the stratum from which the random samples are to be generated among the five options "**Stratum 1**" to "**Stratum 5**" on the "**With Characters**" interface and enter the data through the "**Data**" panel. A click on the "**Add**" button would store the data inputted into to the "**Data**" panel after which the next data entry can be made again through the "**Data**" panel. This process would continue until the last data point in the population is entered. All the data entry can be viewed by clicking on the scroll arrow on the "**Data**" panel. Specify the number of samples to be selected in the "**Count**" box and click on the "**Generate**" button to generate the required samples for that stratum. Repeat this process to generate samples from other strata.

If more than five strata are desired, simply click on "**clear**" option beside the data panel of strata 1 to clear the initial data entered and repeat the above process beginning from this stratum to make-up to the number of strata desired. Other steps on how to view and print the results generated as provided in Section 3.1 then follow from here.

6.   CONCLUSION

The current version of the Multipurpose Sample Generator software (MSG 1.0.0) is designed to generate random samples of different forms (numeric or alpha-numeric). Two basic sampling techniques of Simple Random Sample and Stratified Random Sample with replacement were implemented by MSG 1.0.0. Thus, the next version of the software (MSG 1.0.1) would implement these sampling schemes under the sampling with and without replacements.

Other sampling schemes that would be incorporated in future are the double sampling, cluster sampling, systematic sampling and so on. All these shall constitute part of the improvement that would be implemented by the next version of the software, MSG 1.0.1.

Users of the current version, MSG 1.0.0 are therefore enjoined to forward their suggestions and comments that would improve the subsequent versions of the software to the developers using the e-mail addresses provided.

Like the Super QC software (Yahya and Adeniyi, 2012), the MSG software can be freely accessed and downloaded at www.unilorin.edu.ng .

REFERENCES

[1]   Hunt, N. and Tyrrell, S. (2001): Stratified Sampling. Coventry University (Accessed on 12th July 2012).

[2]   Microsoft Visual Studio 2010 (Microsoft cooperation) www.microsoft.com

[3]   Moore, D.S. (2007): *The Basic Practice of Statistics*, 4th Edition, Freeman.

[4]   Moore, D.S. and McCabe, G.P. (2006): *Introduction to the Practice of Statistics*, 5th Edition, Freeman.

[5]   Särndal, C., Swensson, B and Wretman, J. (2003): Stratified Sampling. In: *Model Assisted Survey Sampling*. New York: Springer. pp. 100–109. ISBN 0-387-40620-4.

[6]   Walker, J. (2009): *HotBits: Genuine Random Numbers*. Retrieved 2009-06-27.

[7]   Yates, D.S., Moore, D.S. and Starnes, D.S. (2008). *The Practice of Statistics*, 3rd Ed. Freeman. ISBN 978-0-7167-7309.

[8]   Yahya W.B. and Adeniyi I. A. (2012): Super QC-Pack: A novel Statistical Software for analysing Quality Control Data. *Annals. Computer Science Series,* 10(1):59-72. *http://anale-informatica.tibiscus.ro/*