DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICAL SCIENCES
UNIVERSITY OF DELHI

PROJECT REPORT

ADVANCE STATISTICAL COMPUTING USING R

RANDOM NUMBER GENERATION FROM VARIOUS DISTRIBUTIONS

(AS PART OF THE COURSE PAPER 404: ADVANCED STATISTICAL COMPUTING AND DATA MINING)

UNDER THE GUIDANCE AND SUPERVISION OF

MR. ABHISHEK K. UMRAWAL
(ASSISTANT PROFESSOR)

SUBMITTED BY:

VISHAL KUMAR
M. SC. STATISTICS, SEMESTER – IV

DATE: FEB 28, 2016
PLACE: NEW DELHI

ABSTRACT

To understand the computational aspect of Data Mining, a practical approach using statistical software R was introduced in Statistics Department of University of Delhi. The main purpose of this approach is to understand how random numbers are generated, to validate that the numbers generated mimic the actual population under consideration. Then using that generated sample to estimate population parameters using method of Maximum Likelihood estimation and using some iterative techniques for those distribution involving two parameters.

This study has been divided into 4 parts:

1. In this part random numbers were generated using Linear Congruential Generator (LCG) for the choice of a priori integers due to Lehmer (1951). Then validity tests like histogram, Q-Q plot, Chi-square test for goodness of fit and Kolmogorov-Smirnov test were performed on the random numbers.

2. In this part random numbers were generated from Normal and Gamma distribution using Acceptance/Rejection method. Then MLE's were computed and validity tests like histogram, Q-Q plot, Chi-square test for goodness of fit and Kolmogorov-Smirnov test were performed on the random numbers.

3. In this part random numbers were generated from Bivariate Normal distribution and 3-D histogram was made. Then random numbers were generated from Marshall-Olkin Bivariate Exponential(MOBE) distribution and from Block & Basu Bivariate Exponential Distribution(BBBE).

4. In this part random numbers were generated from Weibull distribution, then MLE's were computed using Newton-Raphson method. Expectation-Maximization Algorithm was used to find MLE's from a mixture of two normal distribution.

ACKNOWLEDGEMENTS

I have given my best efforts to make this a good report. However, it would not have been possible without the kind support, guidance and constant supervision of our professor **MR. ABHISHEK K. UMRAWAL** and I would like to express my gratitude towards him for providing necessary help regarding the concepts and techniques used in this case study.

Lecture notes released by **MR. ABHISHEK K. UMRAWAL,** provided the necessary knowledge to tackle the given problem at hand and to understand the basic "why" concept behind any technique and algorithm used in this report.