Indira Gandhi
National Open University
School of Sciences

MST-002
DESCRIPTIVE STATISTICS

Block

2

CORRELATION FOR BIVARIATE DATA

UNIT 5
Fitting of Curves **5**

UNIT 6
Correlation Coefficient **25**

UNIT 7
Rank Correlation **45**

UNIT 8
Intra-Class Correlation **61**

Curriculum and Course Design Committee

Prof. K. R. Srivathasan
Pro-Vice Chancellor
IGNOU, New Delhi

Prof. Parvin Sinclair
Pro-Vice Chancellor
IGNOU, New Delhi

Prof. Geeta Kaicker
Director, School of Sciences
IGNOU, New Delhi

Prof. Jagdish Prasad
Department of Statistics
University of Rajasthan, Jaipur

Prof. R. M. Pandey
Department of Bio-Statistics
All India Institute of Medical Sciences
New Delhi

Prof. Rahul Roy
Math. and Stat. Unit
Indian Statistical Institute, New Delhi

Dr. Diwakar Shukla
Department of Mathematics and Statistics
Dr. Hari Singh Gaur University, Sagar

Prof. Rakesh Srivastava
Department of Statistics
M. S. University of Baroda, Vadodara

Prof. G. N. Singh
Department of Applied Mathematics
I. S. M. Dhanbad

Dr. Gulshan Lal Taneja
Department of Mathematics
M. D. University, Rohtak

Faculty members of School of Sciences, IGNOU

Statistics
Dr. Neha Garg
Dr. Nitin Gupta
Mr. Rajesh Kaliraman
Dr. Manish Trivedi

Mathematics
Dr. Deepika Garg
Prof. Poornima Mital
Prof. Sujatha Varma
Dr. S. Venkataraman

Block Preparation Team

Content Editor
Dr. Meenakshi Srivastava
Department of Statistics
Institute of Social Sciences
Dr. B. R. Ambedkar University, Agra

Language Editor
Dr. Nandini Sahu
School of Humanities, IGNOU

Secretarial Support
Mr. Deepak Singh

Course Writer
Dr. Rajesh Tailor
School of Studies in Statistics
Vikram University, Ujjain

Formatted By
Dr. Manish Trivedi
Mr. Prabhat Kumar Sangal
School of Sciences, IGNOU

Programme and Course Coordinator: Dr. Manish Trivedi

Block Production

Mr. Y. N. Sharma, SO (P.)
School of Sciences, IGNOU

Acknowledgement: We gratefully acknowledge to Prof. Geeta Kaicker, Director, School of Sciences for her great support and guidance.

December, 2011

ISBN-978-81-266-

Further information on the Indira Gandhi National Open University may be obtained from University's Office at Maidan Garhi, New Delhi-110068 or visit University's website http://www.ignou.ac.in
Printed and published on behalf of the Indira Gandhi National Open University, New Delhi by the Director, School of Sciences.
Laser Typeset by: Tessa Media & Computers, C-206, A.F.E.-II, Okhla, New Delhi
Printed at:

CORRELATION FOR BIVARIATE DATA

In Block 1 of this course, you have studied the analysis of quantitative data mainly dealt with the quantitative techniques which describes the one or more variables e.g. height, weight, sales, income, etc. independently. Those units were broadly classified as measures of central tendency, measures of dispersion, moments, skewness and kurtosis. Often we come across the situation where information on two or more variables, together like height and weight, income and expenditure, literacy and poverty, etc. are available and our interest is to study the relationship between these two variables. The present block deals with the situations having information on two variables.

Unit 1 describes the fitting of various curves including straight line, second degree of parabola, power curves and exponential curves for the given set of data using principle of least squares. With the help of fitting of the curves one can estimate the dependent variable for given value of independent variable.

Unit 2 gives the concept of correlation which studies the linear association between two variables. The concept of correlation and correlation coefficient would be very helpful in regression analysis.

Unit 3 describes the rank correlation which handles the situation where study characteristics are not measureable but can be presented in the form of ranks according to merit of individuals. In this unit, you will study the rank correlation coefficient with its properties.

Unit 4 deals with two different types of situations. First in which no linear association exists between two variables but they may have some other type of curvilinear relationship. In this situation correlation coefficient fails to determine the intensity of relationship and we use correlation ratio. Another situation, when we are interested in studying the relationship among the members of a group or family, leads us to intraclass correlation coefficient. This unit describes the coefficient of determination, correlation ratio and intra-class correlation coefficient.

Suggested Readings:

- Ansari, M. A., Gupta, O. P. and Chaudhari S. S.; Applied Statistics, Kedar Nath Ram Nath & Co., Meerut 1979.
- Arora, S. and Bansi Lal; New Mathematical Statistics, Satya Prakashan, New Delhi, 1989.
- Chaturvedi, J. C.; Elementary Statistics, Prakash Brothers, Agra, 1963
- Elhance, D. N.; Fundamentals of Statistics, Kitab Mahal, Allahabad, 1987
- Goon, A. M., Gupta, M. K. and Das Gupta, B.; Fundamentals of Statistics-Vol-I; World Press Culcutta.
- Gupta, M. P. and Gupta, S. P.; Business Statistics; Sultan Chand & Sons Publications.
- Gupta S. C. and Kapoor, V. K.; Fundamentals of Mathematical Statistics, Sultan Chand & Sons Publications.

Notations and Symbols

$\frac{\partial}{\partial a}$	: Partial derivative with respect to a
U	: Sum of squares of errors
$\sum_{i=1}^{n}$	: Sum over i from 1 to n
log x	: Logarithm of x at the base 10
r = Corr (x, y)	: Correlation coefficient between X and Y
Cov (x, y)	: Covariance between X and Y
$V(x) = \sigma_x^2$	: Variance of X
σ_x	: Standard deviation of X
$\bar{x}$	: Mean of X
A	: Assumed mean
r_s	: Rank correlation coefficient
R_x	: Rank of X
d_i	: Difference between R_x and R_y
r_c	: Concurrent deviation
C	: Number of concurrent deviations
r^2	: Coefficient of determination
η	: Correlation ratio
r_{ic}	: Intra-class correlation coefficient
σ_m^2	: Variance of means