Indira Gandhi
National Open University
School of Sciences

# MST-002
# DESCRIPTIVE STATISTICS

Block

3

## REGRESSION AND MULTIPLE CORRELATION

# 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
Dr. Neha Garg
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:

# REGRESSION AND MULTIPLE CORRELATION

In Block 2 of this course, you have studied different methods including curve fitting using principle of least-squares, correlation, correlation coefficient, rank correlation coefficient and intra-class correlation coefficient.

Unit 9 of this block discusses the linear regression which considers two variables, one as a dependent variable and another one as an independent variable. It describes the two regression lines, regression coefficient and draws distinction between correlation and regression. Angle between two regression lines is also discussed.

If the informations on three or more variables are available and someone is interested to study the relation between them, then one has to use the concepts of plane of regression. Unit 10 provides the basic concepts about the plane of regression considering only three variables. In this unit, you will also learn Yule's notation and concept as well as the properties and variance of residuals.

Unit 11 describes the multiple correlation coefficients with its properties. Correlation coefficient measures the linear strength of association between two variables whereas multiple correlation coefficients measure the strength of association between a dependent variable and joint effect of two or more independent variables.

Sometimes one considers the correlation between dependent variable and one independent variable while ignoring the effect of others when we are considering two or more variables. This type of situation is handled by partial correlation. Unit 12 is mainly focused on partial correlation coefficient. Multiple correlation coefficients are also explained in terms of partial and total correlation coefficient in this unit.

## Suggested Readings:


- 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.
- Ray, N. and Sharma, H. S.; Mathematical Statistics, Ram Prasad & Sons, Agra, 7th edn., 1983
- Sancheti, D. C. and Kapoor, V. K.; Statistics, Sultanchand & Sons, New Delhi, 7th edn., 1991
- Varshney, R. P.; Advanced Statistics, Jawahar Publication, Agra, 28th edn., 2003-2004

# Notations and Symbols

| | |
|---|---|
| $y_i$ | : Observed values of y $(i=1,2,3,...,n)$ |
| $Y_i$ | : Expected values of y $(i=1,2,3,...,n)$ |
| U | : Sum of squares of errors |
| $\frac{\partial U}{\partial a}$ | : Partial derivatives of U with respect to 'a |
| $Cov(x,y)$ | : Covariance between x and y |
| $V(x)=\sigma_x^2$ | : Variance of x |
| r = Corr (x, y) | : Correlation coefficient between x and y |
| $b_{yx}$ | : Regression coefficient of y on x |
| $b_{xy}$ | : Regression coefficient of x on y $\theta$ |
| $\theta_1$ | : Acute angle |
| $\theta_2$ | : Obtuse angle |
| $b_{12.3}$ | : Partial regression coefficient of $x_1$ on $x_2$ |
| $b_{13.2}$ | : Partial regression coefficient of $x_1$ on $x_3$ |
| $x_{1.23}$ | : Estimate of $x_1$ |
| $e_{1.23}$ | : Error of estimate or residual |
| $r_{ij}$ | : Correlation coefficient between $x_i$ and $x_j$ |
| $\sigma_i^2$ | : Variance of $x_i$ |
| Cov $(x_i, x_j)$ | : Covariance between $x_i$ and $x_j$ |
| $W_{ij}$ | : Co-factor of the element in the $i^{th}$ row and $j^{th}$ column of matrix W |
| $\sigma_{1.23}^2$ | : Variance of residual $e_{1.23}$ |
| $R_{1.23}$ | : Multiple correlation coefficient $x_1$ on $x_2$ and $x_3$ |
| $r_{12.3}$ | : Partial correlation coefficient between $x_1$ and $x_2$ |
| $e_{1.3}$ | : Residual for $x_1$ and $x_3$ |
| $e_{2.3}$ | : Residual for $x_2$ and $x_3$ |
| $V(e_{1.3})=\sigma_{1.3}^2$ | : Variance of residual $e_{1.3}$ |
| $Cov(e_{1.3}, e_{2.3})$ | : Covariance between residual $e_{1.3}$ and $e_{2.3}$ |