SYNOPSIS
SYNOPSIS

The foundation of the subject, Quantitative Linguistics was laid down by Augustus, De Morgan, a professor of Mathematics at University of London, around 1872. T.C. Mendenhall, a physicist, latter on confirmed its existence through his work during the year 1887 to 1901. Subsequently the subject was developed by many scientists like G.K. Zipf, G.U. Yule, C.A. Williams, G. Herdan, C.E. Shannon. During the middle part of this century, a lot of studies on Quantitative Linguistics for English, Latin, Italian, German, Spanish and French languages were carried out. In India, some casual studies on Kannada, Tamil, Telugu, Bengali, Marathi, Assamese languages and Karnatic music have been made. A systematic investigation into verse and prose has been done in detail only for Bengali and Assamese language.

A systematic investigation into certain quantitative aspects of Sanskrit literature—variation in style over different literary eras, the disputed authorship, information theoretic treatment, assessment and reduction of non-sampling errors—forms the content of this thesis. The characteristics under consideration have been word-length, sentence-length,
type-token ratio, entropy and the statistical methods employed are essentially empirical fittings of distributions and testing of differences between different eras and various authors.

Chapter I gives a brief review of the previous work closely related to the problems considered in this dissertation and a summary of the present work.

In chapter II, we have made the general investigation into certain quantitative aspects by considering different literary eras. The whole study is planned on the sample survey basis and centres around the estimation of word-length, sentence-length and proportion of parts of speech and calculation of standard errors.

Chapter III deals with the analysis of word-count data for certain sampled pieces of Sanskrit prose and compares the results with those available for few foreign languages. This chapter also includes the mathematical analysis of the type-token relationship and fitting of the stochastic model to it.
Chapter IV, we have investigated the randomness of the series of word-length and that of sentence-length for the prose work of Sanskrit language. Statistical indices of auto-correlation coefficients between consecutive values in both the series have been computed and used for the test of randomness.

Chapter V examines the nature of curves for the word-length and sentence-length distribution. The values of test indices viz. $-D, K$ and $\chi^2$ statistics lead to an overall superiority of the fitting of compound Poisson distribution against the Negative binomial one.

In chapter VI, we have considered the problem of disputed authorship of the Kavyas Kumāra-sambhava (from 9th to 17th cantos) and Ritusamhār. The problem relates to a literary controversy around its conclusive authorship. Type-token ratio study, Yule's characteristic study, weight rate analysis and General statistical analysis of disputed pieces and its comparison with authentic Kavyas of Kalidasa form the basis of inference.

In chapter VII, information theoretic study of Sanskrit language and assessment of non-sampling errors
have been made. The entropy of Sanskrit language have been compared with the entropies of other languages - English, German, Russian and Assamese.