Chapter 7

SUMMARY AND SCOPE FOR FURTHER WORK

We have applied some known standard statistical tools to Linguistic data for studying writing style of the famous English author Thomas Hardy.

7.1 Summary Of Results.

i) Measures of Apparent Richness of Vocabulary of each of the three novels do not show heterogeneity.

ii) The large sample tests reveal that,

   a. Thomas Hardy did not maintain consistency in Vocabulary Richness in the novel “Desperate Remedies” and “Tess Of D’Urbervillies”.

   b. The novel “Two On A Tower”, is homogeneous in Vocabulary Richness.

iii) Kolmogorov-Smirnov test, when applied to case of Sentence Length data, showed that,

   a. every novel is homogeneous within itself, and
b. all three novels are homogeneous between themselves.

iv) Writing style of Thomas Hardy regarding Sentence Length is homogeneous over a period of twenty years.

v) Test of Homogeneity of Parallel samples, when applied to six samples from the three novels, reveals that, the samples are not homogeneous in case of marker word “The”.

vi) When the homogeneity test is applied to each of the three novels separately, the only novel that is homogeneous in case of marker word “The” is, “Two On A Tower”. The other two novels “Desperate Remedies” and “Tess Of D’Urbervillies” do not maintain uniformity.

vii) Application of Katz’s test indicates that Negative Binomial Probability distribution is suitable to all the six samples.

viii) Further Negative Binomial distribution fits well to all the six samples.

ix) Large sample test for difference between means of Characteristic $K^*$ is applied to two data sets from each of the three novels. The results are as follows:

a. There is significant difference between means of two samples from “Desperate Remedies”.

b. The difference between two means is insignificant for the novels “Two On A Tower” and “Tess Of D’Urbervillies”. Hence these two novels are consistent within themselves.
c. Further comparison of the two novels “Two On A Tower” and “Tess Of D’Urbervillies” shows consistency regarding $K^*$ among themselves.

x) A frequency distribution of Yule’s K is obtained for 260 samples. It showed that the distribution is positively skewed and leptokurtic.

xi) Lognormal probability distribution does not fit to the data.

xii) To find independence of sample size and Yules Characteristic $K$, Karl Pearson’s coefficient of correlation is obtained. Its value is $-0.238206$. This shows that sample size and Yule’s $K$ are not independent.

xiii) Graph 6.3 shows that for different sample sizes ranging from 500 to 1900 (below 2000) Yule’s $K$ is not constant.

xiv) Graph 6.4 shows that for different sample sizes ranging from 2100 to 2900 (above 2000) very little dispersion or values of Yule’s $K$ are nearly constant.

xv) Karl Pearson’s coefficient of correlation between sample size above 2100 and Yule’s $K$ show linear independence.

xvi) It is observed that with increase in sample size Yule’s $K$ comes closer.

xvii) Writing style of Thomas Hardy is more homogeneous in the novel “Two On A Tower” than in “Desperate Remedies” and “Tess Of D’Urbervillies” regarding stylistic parameters studied in this dissertation.
Style of an author is constituted by a number of different stylistic parameters like use of adjectives, phrases, nouns, pronouns, different marker words etc. Apparently, statistical examination of each of these parameters should throw light on his style.

— However, exhaustive study of all the parameters is not an easy task.

7.2 Results Compatible To Objectives

Regarding fulfilment of the objectives of this study, we could state as follows:

i) Writing style of Thomas Hardy is more homogeneous in the novel “Two On A Tower” as compared to “Desperate Remedies” and “Tess Of D’Urbervillies” with respect to stylistic parameters studied in this dissertation.

ii) a. \( E(\hat{\phi}) = \frac{1}{\mu} \)

b. \( V(\hat{\phi}) = \phi^4 V(\overline{\phi}) \)

where \( \phi \Rightarrow \) Measure of Apparent Richness of Vocabulary.

iii) a. \( E(K^*) = \frac{1}{N} \mu^2 \)

b. \( V(K^*) = \frac{4}{N^3} V^4 \left\{ \frac{\mu_4 - \mu_2^2}{4\mu_2^2} + \frac{\mu_2^2}{\mu_1^2} - \frac{\mu_3}{\mu_2^2} \right\} \)

where

\[ V = \frac{\sqrt{\mu_2}}{\mu_1'} , \mu_1' > 0 \]
iv) Negative Binomial distribution fits good to the data regarding marker word “The”.

v) Lognormal distribution does not suit to the data of Yule’s K.

vi) Yule’s K is constant for samples of sizes above 2000 words.

7.3 Scope For Further Work

The dissertation has its own limitations. Further, there is always scope for more and more research in the same or allied areas.

Some of the topics in which the research work can be extended are,

i) Comparative study of writing style of Thomas Hardy and some contemporary author.

ii) Comparative study of writing style in “Two On A Tower” and “Monoryawar Doghe ” which is Marathi translation of “Two On A Tower”.

iii) Study of style of Thomas Hardy as a poet.

iv) Further study of Characteristic K.

v) Study of writing style of Thomas Hardy with respect to other stylistic parameters which are not used in this dissertation.

vi) Study of writing style by using different statistical tools.