I-2. Policy and structure: How to use this text

Statistics is a field of academic study in which methods to obtain natures or regularity of phenomena from obtainable data are studied. Data is plural form of datum and it is generally composed from more than two single datum. Each datum is expressed by single or plural factors or components. Statistics is used for various purposes. 1. Confirmation of difference among sub data population. 2. Separation of factors or component composing the data. 3. Expression of the relation among factors and variables. 4. Discrimination of sample population to sub sample populations. One possible structure of text book is to make chapters depending on the practical purpose. Such structure is convenient to search appropriate statistic methods, though we cannot avoid replicated explanation of principle theory, because same perceptions are used in different statistical method and different practical purposes. The author is targeting readers who has little knowledge of statistics and want to learn from basic theory. Structuring depending on practical purpose is unrealistic for this text book. The author broadly categorized statistical analysis to single variable analysis, multivariable analysis and the others. Here, the meaning of multivariable analysis is analyses include linear algebra (matrix calculation). Single variable analysis is analyses which can be explained without techniques of linear algebra. Because, most of targeting readers are beginners of advanced math and have little knowledge of linear algebra. Basic policy of this text book is "use easier explanation understandable without advance knowledge" and "provide every necessary basic knowledge for deeper understanding". For this reason, several mathematical explanations are provided in separated sections, and use many figures in the paragraphs to provide perception. The explanations may sometimes lack mathematical rigor, because the author is not professional in statistics and mathematics. The author asks the readers who have deep knowledge of statistics and mathematics to point out weakness of the explanations and inform the mistakes to the author. For readers who do not have any interest in such mathematical theories, you can skip those sections. For such readers, the author inserts tables showing calculation procedure of example dataset. As the result of author's policy and categorization, the author made 7 chapters after chapter I (Introduction). Chapter II is "Attitude of statistics and frequentism". The author wrote how the statistical judgement should be done. Chapter III is "Probability distributions". In this chapter, various probability distributions are introduced including how the distributions were made, namely binomial distribution, Poisson distribution, normal distribution, chai square distribution, student's t distribution and F distribution. In the explanations of process

to make distributions, the author used several mathematical techniques, namely Taylor expansion, Napier's constant, Jacobian, polar coordinate, multiple integral. The author learned such techniques in liberal art course in university. For readers who do not have such experiences, the author explains such techniques used in advanced math in separate chapters in second half of the chapter (III-3). Chapter IV is procedure to use probability distributions. The author compiled analysis of variance, simple linear regression and chi square test in chapter IV. Statistics so called multivariable analysis are compiled after chapter V. For understanding of multivariable analysis, readers need knowledge of linear algebra (at least matrix calculation). Chapter V is fundamental linear algebra. The author explains important matrix calculation techniques in Chapter V at first, and then various methods of multiple variable analysis such as multiple regression, partial correlation analysis, linear discriminant analysis, principle component analysis, factor analysis, multidimension scaling method are explained in Chapter VI. Chapter VII is for the other statistic methods. Fundamental Bayesian statistics, cluster analysis and conjoint analysis are included in chapter VII. This chapter is still under construction. The author is planning to add explanation of new statistic technologies.

Simple linear regression is a special case of multiple regression. The explanation of simple linear regression may have to be included in multiple linear regression. However, there may be readers who need to understand correlation simply without knowledge of linear algebra. The author put simple linear regression in IV and explains multiple regression in Chapter VI.

Practically. we need to use computer software or systems for calculation. The author will provide various R script for various statistical analyses as appendices. Fundamental explanation of Bayesian statistics and several simulation methods such as MCMC are also included in appendices.