Table of Contents

Chapter	Title	Page
စာရေးသူ၏အမှာစာ		V
1 Introduction to Multivariate Analysis		**1**
1.1	What is Multivariate Analysis (MA) ?	1
1.2	Measurement Scales and Selection Techniques	2
1.3	Graphical Examination of the Data	5
1.4	Searching for Missing Data	10
1.5	Outlier Detection Techniques	12
1.5.1	Z-Scores	13
1.5.2	Turkey's Method (Boxplot)	14
1.5.3	Mahalanobis (Squared) Distance	15
1.5.4	Cook's Distance	16
1.6	Statistical Assumption-Normality	16
1.6.1	Shapiro-Wilks test	18
1.6.2	Kolmogorov-Smirnov test	19
1.7	Statistical Assumption-Homoscedasticity	20
1.7.1	Bartlett test	21
1.7.2	Levene test	21
1.7.3	Boxs M test	21

1.8 Statistical Assumption-Linearity 22
1.9 Data Transformation 23
1.10 Packages for Multivariate Analysis 28

2. Multiple Linear Regression 30
2.1 Introduction 30
2.2 Multiple Linear Regression Model 30
2.3 Building a Linear Regression Model 35
2.3.1 Performance of Fitted Regression 36
2.3.2 Testing for Regression Coefficients 38
2.3.3 Graphical Diagnostic for Residuals 39
2.4 Checking Three Main Important Assumptions 43
2.4.1 Multicollinearity in Regression Analysis 43
2.4.2 Heteroscedasticity in Regression Analysis 44
2.4.3 Autocorrelation in Regression Model 45
2.5 Empirical Study of Multiple Linear Regression 46

3. Logistic Regression 64
3.1 Introduction 64
3.2 Binary or Binomial Logistic Regression 65
3.2.1 Binary Logistic Regression Model 65
3.2.2 Performance of Binary Logistic Regression 67
3.2.3 Empirical Study of Binary Logistic Regression 71
3.3 Multinomial Logistic Regression 79
3.3.1 Multinomial Logistic Regression Model 79
3.3.2 Empirical Study of Multinomial Logistic Regression 80
3.3.3 Performance of Fitted Multinomial Logistic Regression 85
3.4 Ordinal Logistic Regression 87
3.4.1 Ordinal Logistic Regression Model 87
3.4.2 Empirical Study of Ordinal Logistic regression Model 87
3.4.3 Performance of Fitted Ordinal Logistic Regression 94

4. Principal Component Analysis 95
4.1 Introduction 95
4.2 Principal Component Model 96
4.3 Calculating Principal Components 97
4.4 Empirical Study of Principal Component Analysis 100

5. Factor Analysis 114
5.1 Introduction 114
5.2 The Orthogonal Factor Model 116
5.3 Method of Estimation 117
5.3.1 Principal Component Method 117
5.3.2 Maximum Likelihood Method 119
5.4 Factor Rotation 120
5.4.1 Orthogonal Rotation 120
5.4.2 Oblique Rotation 120
5.5 Overall Measures of Intercorrelation 121
5.5.1 Bartlett Test of Sphericity 121
5.5.2 Measure of Sampling Adequacy 122
5.6 Determine the Number of Factors 122
5.7 Internal Consistency Reliability (Cronbach's Alpha Measurement) 123
5.8 Empirical Study of Factor Analysis (Principal Component) 124
5.9 Empirical Study of Factor Analysis (Maximum Likelihood) 138

6. Analysis of Variance(ANOVA) 140
6.1 Introduction 140
6.2 One-way ANOVA 141
6.3 Two-Way ANOVA 143
6.4 Repeated Measures ANOVA 145
6.5 ANOVA Models' Formula in R 146
6.6 Empirical Study of One Way ANOVA 146
6.7 Empirical Study of Two Way ANOVA 153
6.8 Empirical Study of Repeated Measures ANOVA:One-way 165
6.9 Empirical Study of Repeated Measures ANOVA:Two-way 171

7. Multiple Analysis of Variance (MANOVA) 175
7.1 Introduction 175
7.2 One-way MANOVA 176
7.3 Two-way MANOVA 178
7.4 Test Statistics for MANOVA 179
7.5 Empirical Study of MANOVA 181

8. Analysis of Covariance (ANCOVA) 201
8.1 Introduction 201
8.2 ANCOVA Model 201
8.3 Assumptions of ANCOVA Model 203
8.4 Empirical Study of ANCOVA 204

Appendix 220
References and Resources 225

စာရေးသူ၏ အမှာစာ

စာရင်းအင်းပညာဘာသာရပ်၏ အသုံးဝင်မှုကို R နှင့် ဖြည့်စွက်ရင်း လူမှုပတ်ဝန်းကျင်များတွင် ဖြစ်ပေါ်နေသော နိစ္စဓူဝကိစ္စမှန်သမျှ ဆုံးဖြတ်ရန်အချက်များကို မှန်ကန်မြန်ဆန်စွာနဲ့ တွက်ချက်ဆုံးဖြတ်နိုင်ရန် အခြေခံစာရင်းအင်းပညာအတွက် R ကို ရေးသားထုတ်ဝေခဲ့သည်။ ကိန်းဂဏန်းအချက်အလက်များကို ကိုယ်စားပြုဖော်ပြသည့် ကိန်းရှင်တစ်ခုတည်း (univariate variable) ကို အသုံးပြုတွက်ချက် စိစစ်ဆုံးဖြတ် ကောက်ချက်ချရာမှသည် ကိန်းရှင်နှစ်ခု (bivariate variables) ကိုအသုံးပြု၍ အချင်းချင်းဆက်စပ်မှုများကို ဖော်ထုတ်ပြီး စာရင်းအင်း ပညာဆိုင်ရာ model တစ်ခုကိုဖန်တီး၍ နောင်အနာဂတ်တွင် ဖြစ်လာမည့်အခြေအနေကို ခန့်မှန်း တွက်ချက်နိုင်သည်အထိ အခြေခံဆိုင်ရာသိသင့်သည့် စာရင်းအင်းပညာရပ်ဆိုင်ရာနည်းလမ်းများ ပါဝင်ခဲ့သည်။ လေ့လာဆည်းပူးခဲ့သောပညာရပ်နှင့် ဆယ်စုနှစ်တစ်ခုစာ အုပ်ချုပ်မှုအတွေ့အကြုံ တို့ကိုပေါင်းစပ်လို့ လွယ်ကူစွာတွက်ချက်စိစစ်၍ ကောက်ချက်ချနိုင်ရန် ရေးသားထုတ်ဝေခဲ့ပြီး ဖြစ်ပါသည်။

ကောက်ချက်ချသည်ဆိုရာဝယ် အချက်အလက်များများနှင့်အကြောင်းအရာစုံလင်လေလေ ပိုမိုခိုင်မာမှန်ကန်သော ဆုံးဖြတ်ချက်ကို ရလေလေဆိုတာ ငြင်းဖွယ်ရာမရှိပြီ။ ဒီနေရာမှာ မဆိုင်သော အကြောင်းအရာထည့်သုံးမိခြင်းသည် ဆုံးဖြတ်ချက်ကို လွဲမှားစေနိုင်သောကြောင့် ကိန်းရှင်များ အကြားဆက်စပ်မှု၏ သဘောသဘာဝ ပမာဏများနှင့်အတူ ဆက်စပ်မှုရှိသော ကိန်းရှင်အများ (multiple variables) ကိုလေ့လာစိစစ်ဆုံးဖြတ်ခြင်းက ပိုမိုအရေးပါလာပါသည်။ တွက်ချက်မှုအထောက်အကူပြု မြန်နှုန်းမြင့်ကွန်ပျူတာများနှင့်အတူ လွယ်ကူစွာအသုံးပြု software များပေါ်ထွန်းလာခဲ့သော်လည်း အသုံးပြုရသည့် အဆင့်မြင့်စာရင်းအင်းပညာဆိုင်ရာ နည်းလမ်းများ၏ နက်နဲရှုပ်ထွေးမှုများက မဟာတံတိုင်းတစ်ခုသဖွယ် ဖြစ်တည်လာခဲ့သည်။ နည်းလမ်းပေါင်းစုံ သိရှိရင်ဖြင့် လွယ်ကူစွာ ဖြတ်သန်းကျော်ဖြတ်လို့ လိုရာခရီးကို ချောမောစွာ အမြန်ရောက်မည်သာဖြစ်သည်။ နက်နဲလှတဲ့ အဆင့်မြင့်စာရင်းအင်းပညာဆိုင်ရာ နည်းလမ်းပေါင်းစုံကို အလွယ်တကူသိရှိလေ့လာနိုင်အောင် ကွက်လပ်လေးတစ်ခု ဖြည့်ဆည်းပေးရန် လိုလားခဲ့ပါသည်။

ရှုပ်ထွေးလှသော နေ့စဉ်လူမှုဘဝပြဿနာပေါင်းစုံကို ကိန်းရှင်နှစ်ခုထက် ပိုမိုသော ကိန်းရှင်အများကိုသုံး၍ လေ့လာစိစစ်သည့်နည်းလမ်း (multivariate analysis techniques) များကိုသုံး၍ ဖြေရှင်းနိုင်စေရန်အလို့ငှာ အလွယ်တကူလေ့လာတတ်မြောက်ရေးအတွက် လိုအပ် နေတဲ့ကွက်လပ်ကလေး ဖြည့်ဆည်းရန် "အဆင့်မြင့်စာရင်းအင်းပညာအတွက် R (R for Advanced Statistics)"ကိုမြန်မာလို ထပ်မံမွေးဖွားခဲ့သည်။ အဆင့်မြင့်စာရင်းအင်းပညာအတွက် တင်ပြရမည့်အကြောင်းအရာ နည်းလမ်းများ အများအပြားရှိသောကြောင့် နှစ်ပိုင်းခွဲပြီး၊ အတွဲ(၁) ကို ထုတ်ဝေလိုက်ပါသည်။

ဤစာအုပ်၏ အခန်း(၁)သည် multivariate analysis ၏ အခြေခံသဘောတရားများကို တင်ပြထားပါသည်။ အခန်း(၂)သည် အသုံးများသော multiple linear regression ကို အသေးစိတ် တင်ပြထားပါသည်။ အခန်း(၃)သည် logistic regression ကို binary logistic, multinomial logistic နှင့် ordinal logistic စသည့် အမျိုးအစား(၃)မျိုးဖြင့် တင်ပြထားပါသည်။ အခန်း(၄) နှင့် (၅) သည် principle component analysis နှင့် factor analysis တို့ကိုဖော်ပြထားပါသည်။ အခန်း(၆)သည် စစ်ဆေးကိန်းတစ်ခုဖြစ်သော analysis of variances (ANOVA) ကို ရှင်းပြထားပါသည်။ အဆိုပါစစ်ဆေးကိန်းနှင့်ဆက်စပ်လျက် multivariate analysis of variances (MANOVA) ကိုအခန်း(၇)တွင် တင်ပြထားပါသည်။ နောက်ဆုံး အခန်းဖြစ်သောအခန်း(၈)သည် analysis of covariances (ANCOVA) ကို တင်ပြထားပါသည်။ ဤစာအုပ်ပါ အချို့သောအခန်းများသည် multivariate analysis အတွက် နည်းလမ်းများမဟုတ်သော်လည်း နည်းပညာအရ ဆက်စပ်နေသောကြောင့် ထည့်သွင်းဖော်ပြထားပါသည်။ ဥပမာအားဖြင့် ANOVA ကို MANOVA ဖြင့် ဆက်စပ်လျက် ဖော်ပြထားပါသည်။

ယခုထုတ်ဝေလိုက်သော "အဆင့်မြင့်စာရင်းအင်းပညာအတွက် R - အတွဲ (၁)" ကို ဖတ်ရှု၍ နက်နဲသော အဆင့်မြင့်စာရင်းအင်းပညာဘာသာရပ်များကို R program နှင့်တွဲပြီး အလွယ်တကူ သင်ယူလေ့လာကာ ပညာရပ်ဝန်းကျင်ကောင်းတစ်ခု ဖန်တီးနိုင်ကြမည်ဟု ယုံကြည်မျှော်လင့်မိပါသည်။

ဒေါက်တာစိုးဝင်း, PhD
စာရင်းအင်းပညာ

ရက်စွဲ။ ။ ၁၉-၄-၂၀၁၉