STATISTICAL MODEL FOR BETTER NUTRITION ASSESSMENT USING ANTHROPOMETRIC METHODS

by

Wickramasinghe Jayasekara Lokuge Aruna Damayanthe

Declaration

The work described in this thesis was carried out by me under the supervision of Dr. P. Kalukottage and a report on this has not been submitted whole or in part to any university or any other institution for another Degree/Diploma.

..............................
W.J.L.A.Damayanthi
Reg.No: GS/M.Sc./Ap.St./2827/06
Index No: PGD/PS/121
Declaration

I/We certify that the above statement made by the candidate true and that thesis is suitable for submission to the university for the purpose of evaluation.

[Signature]

Dr. P. Kalukottage, (Supervisor)
Senior Lecturer,
Department of Statistics and Computer Science,
University of Sri Jayewardenepura.

2009. 11. 20
DEDICATION

Affectionately dedicated to my loving Parents
ACKNOWLEDGEMENT

First I would like to acknowledge sincere gratitude to Dr. S.G.Banneheka, course coordinator, M.Sc. in applied statistics, University of Sri Jayawardenepura Sri Lanka, for giving me an opportunity to undertake this research.

My earnest thanks is expressed to my internal supervisor Dr. P. Kalukottage, Senior Lecturer, Department of Applied statistics, University of Sri Jayawardenepura Sri Lanka, for much appreciated guidance and encouragement given to me throughout the period of the research. Her experience, wide knowledge about the Statistics and their real world applications and clear understanding about research work were very helpful to me properly focus on the subject and understand its principles.

My special thanks is expressed to my external supervisor Dr. K.D.R.R. Silva, Head of the department of Applied Nutrition, Wayamba University of Sri Lanka, for his supervision, invaluable guidance and great courtesy of sharing his valuable time and experience with me.

My heart felt gratitude is bestowed upon Dr. T.S.G. Peiris, senior lecturer, University of Moratuwa, who have gave me to invaluable guidance and subject knowledge for the completion of this study.

My gratitude also goes to Mr. K.V. Fernando, Mr. P. H. D. Taranga Ranasingha, Ms. Thilanka Ranatunga and for their untiring support and encouragement given to me at all times.

Finally, my heart felt gratitude goes out to each and everyone who extended their assistance in whichever way possible to accomplish this research.
### CONTENT

<table>
<thead>
<tr>
<th>Section</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dedication</td>
<td>I</td>
</tr>
<tr>
<td>Acknowledgment</td>
<td>II</td>
</tr>
<tr>
<td>Content</td>
<td>III</td>
</tr>
<tr>
<td>List of graphs</td>
<td>VI</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>VII</td>
</tr>
<tr>
<td>Abstract</td>
<td>VIII</td>
</tr>
</tbody>
</table>

**Chapter 1**

**INTRODUCTION**

1.1 Background and justification

1.1.1 Nutritional transition

1.1.2 Nutrition assessment methods

1.1.2.1 Anthropometric methods

1.1.3 Project context

1.2 Research questions

1.3 Objectives

1.3.1 General objectives

1.3.2 Specific objectives

1.4 Scope and limitations

1.4.1 Scope

1.4.2 Limitations

1.5 Organization of the report

**Chapter 2**

**LITERATURE REVIEW**

2.1 Introduction

2.2 Body composition

2.3 Body fat

2.4 Assessment of body composition

2.4.1 Body Mass Index (BMI)
2.4.2 Comparison of BMI and other anthropometric methods 10
2.4.3 Limitations of BMI as an indicator for fat mass 12
2.5 Different methods to assess body fat 12
  2.5.1 Under water weighing 12
  2.5.2 Computer tomography (CT) 13
  2.5.3 Body impedance analysis (BIA) 13
  2.5.4 Skin fold thickness 13
   2.5.4.1 Equations of assessing body fat by measuring skin fold thickness 14
2.6 Summary 14

Chapter 3
METHODOLOGY
3.1 Introduction 16
3.2 Study Design 16
3.3 Setting and time 16
3.4 Subjects and sample size 17
3.5 Sampling technique and selection of subjects 17
3.6 Data collection 17
   3.6.1 General information 17
   3.6.2 Anthropometrics 17
    3.6.2.1 Weight 17
    3.6.2.2 Height 18
    3.6.2.3 Hip 18
    3.6.2.4 Waist 18
    3.6.2.5 Mid Upper Arm Circumference 18
    3.6.2.6 Tight Circumference 18
    3.6.2.7 Skin folds thicknesses 18
     3.6.2.7.1 Triceps 19
     3.6.2.7.2 Abdomen 19
     3.6.2.7.3 Suprailiac 19
     3.6.2.7.4 Thigh skin fold 19
3.7 Data Analysis 20
Chapter 4
DATA ANALYSIS
4.1 Introduction 24
4.2 Data cleaning process 24
4.3 Descriptive analysis of response variables 24
4.4 Linear regression models 25
   4.4.1 Linear regression model with all regressor variables 25
      4.4.1.1 Model adequacy 26
   4.4.2 Developing regression model using best subset regression method 27
      4.4.2.1 Regression model according to the best sub set with Interaction 28
         4.4.2.1.1 Model adequacy 29
      4.4.2.2 Best fitted regression model 30
         4.4.2.2.1 Model adequacy 31
      4.4.2.3 Prediction interval for the observation 33
4.5 Accuracy comparison of interval estimation with commonly used anthropometric indices 34
   4.5.1 Accuracy comparison of interval estimation with anthropometric Indices 34
   4.5.2 Finding suitable cut off values for anthropometric indices suits with Sri Lankan context 35

Chapter 5
CONCLUSION 37
REFERENCES 38
Annex-1 41
## List of graphs

<table>
<thead>
<tr>
<th>Graph</th>
<th>Description</th>
<th>Page number</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Individual relationship between response variable (BF%) vs each regressor variable.</td>
<td>24</td>
</tr>
<tr>
<td>4.2</td>
<td>Normality of residuals</td>
<td>26</td>
</tr>
<tr>
<td>4.3</td>
<td>Constant variance of residuals</td>
<td>27</td>
</tr>
<tr>
<td>4.4</td>
<td>Normality of residuals</td>
<td>29</td>
</tr>
<tr>
<td>4.5</td>
<td>Constant variance of residuals</td>
<td>30</td>
</tr>
<tr>
<td>4.6</td>
<td>Normality of residuals</td>
<td>31</td>
</tr>
<tr>
<td>4.7</td>
<td>Constant variance of residuals</td>
<td>32</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------</td>
<td></td>
</tr>
<tr>
<td>CHD</td>
<td>Coronary Heart Disease</td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td>Body Mass Index</td>
<td></td>
</tr>
<tr>
<td>FFM</td>
<td>Fat Free Mass</td>
<td></td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organizations</td>
<td></td>
</tr>
<tr>
<td>WHR</td>
<td>Waist to Hip Ratio</td>
<td></td>
</tr>
<tr>
<td>WC</td>
<td>Waist Circumference</td>
<td></td>
</tr>
<tr>
<td>WHtR</td>
<td>Waist to Height Ratio</td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>Computer Tomography</td>
<td></td>
</tr>
<tr>
<td>BIA</td>
<td>Bio-electrical Impedance Analysis</td>
<td></td>
</tr>
<tr>
<td>MUAC</td>
<td>Mid Upper Arm Circumference</td>
<td></td>
</tr>
<tr>
<td>TC</td>
<td>Thigh Circumference</td>
<td></td>
</tr>
<tr>
<td>BF%</td>
<td>Body Fat Percentage</td>
<td></td>
</tr>
</tbody>
</table>
ABSTRACT

Increased body fat is the main risk factor for several nutrition related non-communicable diseases such as hypertension, cardio-vascular disease, diabetes etc. There are several methods to directly assess body fat percentage with higher accuracy levels but they need advanced equipments, higher cost and advanced technical knowledge. Anthropometric indices use several body measurements and are also widely used as simple methods to assess the risk of nutrition related non-communicable diseases.

Objective of the study is to develop a simple and reliable statistical model to predict body fat percentage using anthropometric measurements. This study was a cross sectional study which has covered total of 300 apparently healthy Sri Lankan office workers. Body fat percentage, anthropometric measurements and relevant general details were collected and analyzed. Best subset regression method used to build regression model to predict body fat percentage.

One sided prediction interval, based on the fitted regression model, was proposed to predict health risk of an individual. If the cutoff value lies within that interval, then the individual can be classified as having a health risk. This proposed method was tested with small sample, and found that their health risk is classified accurately.

Also the study suggested that the BMI cut off values should be adjusted as 25.3 for females and 26 for males to classify risk in the same context. In general, 25 can be considered as a better cut off value for BMI to classify risk for study population.
Chapter 1

Introduction
INTRODUCTION

1.1 Background and justification

1.1.1 Nutritional transition

With the rapid change in food habits and lifestyle patterns, diet related chronic diseases became a major public health problem in many countries. Coronary Heart Disease (CHD), diabetics, cancer, hypertension are related with personal nutritional status. Obesity (high fat percentage in the body) is the leading causative factor for above diseases. Nutrition transition, behavioral changes and sedentary lives are the major causes for obesity condition.

Coronary Heart Disease (CHD) is one of the leading causes of death, accounting for 27% of all deaths in 2000 in Sri Lanka. A key issue in the management of diet related chronic diseases is prevention. Many of these risk factors, including obesity, can be prevented or modified through appropriate lifestyle changes and dietary habits. However identify the risk in early stages, will allows for proper management of these diet related chronic diseases.

1.1.2 Nutrition assessment methods

Several assessment methodologies are adopted to determine the risk of diet related chronic diseases which classified into following main categories.

1. Anthropometric methods
2. Biochemical methods
3. Clinical methods
4. Dietary assessment methods

From these main types of assessment methods, dietary assessments are used to asses the food habits and food intake which enables to identify the possible risk factors. However these dietary assessments have higher workloads inherited with those assessment
methods. Those two methods take more time and experienced human resources to conduct detailed interviews with the subjects.

Measuring blood cholesterol and tri-glicerides levels can be considered as more accurate bio-chemical method of assessing risk of cardio vascular diseases. However cost of applying biochemical methods for nutritional assessments is higher when compared with other methods and also these techniques need higher expertise knowledge and technology.

Due to those limitations, anthropometric methods are widely used as a simple and feasible nutritional assessment method.

1.1.2.1 Anthropometric methods

Anthropometrics are the objective measurements of body muscle and fat. They are used to compare individuals, to compare growth in the young, and to assess weight loss or gain in the mature individual. Weight and height are the most frequently used anthropometric measurements, and skin fold measurement of several areas of the body is also taken.

Use of anthropometry requires two essential items: an anthropometric indicator and a cut-off point. The indicator, often called an anthropometric index, is a measurement or a combination of measurements made in the field, such as weight and height, or the combination of measurements with additional data, such as age. Common anthropometric indices are as follows.

1. BMI – Body mass index
2. Waist to hip ratio
3. Waist to height ratio
4. Mid upper arm circumference
5. Skin fold thickness - body fat percentage
To estimate body fat, skin fold measurement can be made using skin-fold calipers. Most frequently, triceps and sub scapular skin-folds are measured. Measurements can then be compared to reference data, and to previous measurements of the individual, if available. This can be considered as more reliable method of nutrition assessment as it gives an idea about the body fat level, which is known as a main risk factor of diet related chronic diseases.

Accuracy level of each indicator defers from one community to another based on the various demographic, cultural and behavioral factors. Therefore identification of simple, efficient and most accurate anthropometric index or combination of indices can be considered as the key aspect of early diagnosis and effective prevention from diet related chronic diseases.

Also the cutoff values of these indicators are set according to international standards and rarely adopted to more accurate regional and national standards. This leads to misinterpretation of nutritional status and eventually reduces the effectiveness of decisions made on the results of nutritional assessments. Identification of cutoff values with a suitable statistical analysis will also enhance the specificity and sensitivity.

1.1.3 Project context

Proposed project focused on developing a statistical model to identify simple and reliable anthropometric measurement/indices with the objective of enhancing the overall accuracy, and efficiency of nutritional assessments. At the same time the proposed project focused on enhancing sensitivity and specificity of anthropometric indices by defining more reliable cutoff values which suits to local conditions.
1.2 Research questions

Two main research questions were identified and the study design was carefully adopted to answer those identified questions. Those research questions are as follows.

1.2.1 What are the suitable cutoff values for common anthropometric indices, to be used for Sri Lankan population.

Body shape and size is measured by many anthropometric measurements. Most of these anthropometric measurements and indices are designed to identify the various patterns of fat depositions in various body parts. Several researches proved that there is a significant linear relationship between anthropometric measurements and body fat level. However there is a need for exploring suitable cutoff values for anthropometric indices.

1.2.2 What is the simple and highly significant anthropometric measurements, other relevant factors or their combination, to assess nutritional level instead of assessing body fat level directly?

Risk of diet related chronic diseases can be reduced by frequent nutrition assessments. Nutrition assessments can be done by assessing body fat levels or using other anthropometric assessment methods. However, assessment of body fat level is a difficult task as it requires advanced technical skills and equipments. Taking anthropometric measurements is relatively easy for general public. But each anthropometric measurement has its own limitations and accuracy to assess the risk of diet related chronic diseases. Therefore identifying suitable anthropometric measurement or combination which is highly significantly related with body fat levels will make nutrition assessments simpler and also accurate. Also the best anthropometric measurement or combination may differ from one community to another. Similar studies are rare for Sri Lankan population and it’s important to identify a model which suits local conditions.
1.3 Objectives

The study has set following objectives to find answers for the research questions identified.

1.3.1 General objectives

To determine accurate and efficient method to assess body fat percentage by developing a simple and reliable statistical model using either one anthropometric measurement/indices or combination.

1.3.2 Specific objectives

1. To identify the relationship between various anthropometric measurements and body fat level.

2. Develop a simple and reliable statistical model to assess body fat level which suits Sri Lankan context.

3. Determine and compare accuracy of developed model and anthropometric indices by cross checking with the assessed body fat levels.
1.4 Scope and limitations

1.4.1 Scope

The research was focused on apparently healthy adults worked on government and private sector institutions in Sri Lanka, and to assess body fat percentage.

1.4.2 Limitations

1. Subjects were selected for age ranged from 25-45 years.

2. Sample size was limited to 300 and apparently healthy adults were selected.

3. Geographical area was restricted only to Colombo and suburbs.

4. Anthropometric methods were considered as simple and feasible nutritional assessment methods due to the financial limitation.
1.5 Organization of the report

The report has 6 chapters and introduction was included in 1st chapter which covered basic background information, justification, research questions, objectives, scope and limitations. Then chapter 2 has covered relevant facts from local and foreign researches. Those were reviewed to provide better understanding and background knowledge for the study. Methodology of the research was included in the chapter 3 and covered the details such as geographical area, sample size, sampling techniques, data collection, equations used, cut off values etc. Results were analyzed and discussed in Chapter 4 and revealed new knowledge explored in this study. Chapter 5 was dedicated to discuss conclusions made with results and discussions. All reference materials used for the study was listed and included as Chapter 6. In addition to these main chapters, other relevant documents related with the research were attached as annexes. Under that the data collection questioner was attached as annex 1.
Chapter 2

Literature Review
LITERATURE REVIEW

2.1 Introduction

Researches are being conducted in many countries covering several aspects of nutrition and in the process of generating new knowledge. This study has focuses on developing a statistical model for body fat levels and focused on developing simple and most accurate nutritional assessment methods to represent body fat levels.

Literature review was done to gather more information on basic terminology in nutrition and to explore the up to date knowledge on relevant topics. Following literature review used extracts from relevant researches and studies conducted in various countries. First, it has focused on body composition and further information of body fat (Section 2.3). Then discussed about the body composition assessment (Section 2.4) and further explained about the Body Mass Index (Section 2.4.1) as one of a major method for body composition analysis. Also the literature review has compared BMI with other anthropometric assessment methods (Section 2.4.2) to provide better understanding. Then it has discussed the limitations of BMI as an indicator for body composition analysis (Section 2.4.3). Next it was focused discussing other methods to assess body fat levels (Section 2.5) and skin fold thickness (Section 2.5.1), under water weighing (Section 2.5.2), computer tomography (Section 2.5.3) and body impedance analysis (Section 2.5.4) were covered under this discussion. Finally the literature review has focused on common equations for body fat percentage calculations (Section 2.6) in skin fold thickness method, which is the most common and reasonably accurate method for assessment of body fat.

2.2 Body composition

Most anthropometric methods used to assess body composition are based on two chemically distinct compartments of the body, namely fat and fat-free mass. Anthropometric techniques can indirectly assess fat and the fat-free mass, and variations in their amount and proportion can be used to assess nutritional status. For example, fat is the main storage form of energy in the body and is sensitive to acute
malnutrition. Thus, alterations in body fat content provide indirect estimates of changes in energy balance (Rosalind S., Giban, 2005)

2.3 Body fat

The body fat content is the most variable component of the body, differing among individuals of the same sex, height, and weight. On average, the fat content of women is higher than that of men, representing 32% of their total body weight compared with 25% for men (Lee RD, David C. and Boston N., 1996). Above cutoff values are internationally accepted as standards to determine the risk. Excesses or depletions of fat and fat free mass are associated with an increased risk of chronic diseases. The amount of FFM is considered to be directly correlated with health and longevity (Shephard R.J., 1994).

2.4 Assessment of body composition

There are many methods, ranging from simple to very complex, available for estimating body composition. Anthropometry, with measurements of skin fold thicknesses, body circumferences, body impedance analysis are the most widely used field methods for the estimating human body composition. In population studies, body fat is often assessed by anthropometry. Skin fold thickness is frequently used to estimate the percentage body fat (Rosalind S., Giban, 2005).

2.4.1 Body Mass Index (BMI)

Body Mass Index (BMI) is also considered as an indicator for body fat level of individuals (Rosalind S., Giban, 2005). Currently, assessment and screening of overweight and obesity frequently rely on the use of anthropometry in the form of the BMI.

BMI has been used routinely to classify subjects as obese or nonobese. The World Health Organization (WHO) and the International Obesity Task Force recommend the BMI cutoff point of 25 kg/m² for overweight and 30 kg/m² for obesity. This cutoff point has been derived mainly from mortality statistics from European and American populations. Several studies carried out mainly among Asian populations have

Also, studies have shown that Asian populations have high risks of type 2 diabetes, cardiovascular disease, and mortality from other causes at relatively lower BMI. This is attributable to the higher proportion of body fat in Asian populations (Norgan, N. G. 1994; Deurenberg-Yap, M., G. Schmidt, W. A. van Staveren, and P. Deurenberg 2000; Dudeja, V., A. Misra, R. M. Pandey, G. Devina, G. Kumar, and N. K. Vikram 2001; Deurenberg, P., M. Deurenberg-Yap, L. F. Foo, G. Schmidt, and J. Wang 2003; Deurenberg-Yap, M., S. K. Chew, and P. Deurenberg 2002). Therefore, it has been suggested that lower BMI cutoff points should be adopted for Asians. In 2004, the WHO Expert Consultation, after an analysis of population data from more than 10 countries, noted that the proportion of Asian people with a high risk of type II diabetes and cardiovascular disease is substantial at BMIs lower than the existing WHO cutoff point for overweight (≥25 kg/m²). The consultation also agreed that the WHO BMI cutoff points should be retained as international classifications and suggested lower BMI action points between 23 and 27.5 kg/m², which individual countries could use to define the cutoff points for increased risk for their populations (WHO Expert Consultation, 2004).

2.4.2 Comparison of BMI and other anthropometric methods

Recently the efficacy of the BMI was questioned and waist to hip ratio (WHR) or waist circumference (WC) is considered as more indicative measures for obesity. With this assumption, researches assessed the associations of BMI, WC, hip circumference, WHR and waist to height ratio (WHtR) to cardiovascular risk by calculating the area under the receiver-operating characteristic curve and adjusted odds ratios for metabolic syndrome, dyslipidemia, and type 2 diabetes. It was concluded that WHtR or WC may predict prevalence of cardiovascular risk better than BMI or WHR, even though the differences are small (Schneider et al, 2007). Lower risk is associated with fat
placement in hip and thighs. Also recommend that the WHR be less than 0.9 and less than 0.8 for adult males and females, respectively. When WHR is greater than these cut off points, then the risk for diseases rises steeply.

Another study concluded that the best predictor of hypertension among other anthropometric indices was WHtR for Australian Aboriginal people and waist circumference for Torres Strait Islanders. WHR was the best predictor for both diabetes and dyslipidemia in both populations. In multivariate regression analyses, WHR and BMI were independently associated with the probability of coronary heart disease (CHD) for Torres Strait Islanders. However, overall WHR appeared to be the best predictor of the estimated CHD risk for both populations (Wang Z, Rowley K, Wang Z, Piers L, O'Dea K., 2007).

Also researches focused on importance of identifying simple and more reliable tools for assessment of fat level or obesity level. Various anthropometric indices were calculated for type II diabetic subjects and compared them with the metabolic status which was evaluated based on other bio-chemical and clinical methods such as lipidograms which has proven relationship with complications of type II diabetics and CHD. WHtR was correlated with lipidograms in females whereas BMI was the best indicator in males (Katona-Dureković A, Stokić E., 2006).

In addition to above matters, there is a dialogue in international community, about the reliability of cutoff values of anthropometric indices. A WHO expert consultation addressed the debate about interpretation of recommended BMI cut-off points for determining overweight and obesity in Asian populations, and considered whether population-specific cut-off points for BMI are necessary. They reviewed scientific evidence that suggests that Asian populations have different associations between BMI, percentage of body fat, and health risks than to European populations. The consultation concluded that the proportion of Asian people with a high risk of type 2 diabetes and cardiovascular disease is substantial at BMI lower than the existing WHO cut-off point for overweight, >25 kg/m² ( WHO Expert Consultation, 2004).

For many Asian populations, additional trigger points for public health action were identified as 23 kg/m² or higher, representing increased risk, and 27.5 kg/m² or higher as representing high risk. The suggested categories are as follows: less than 18.5 kg/m²
underweight; 18.5–23 kg/m² increasing but acceptable risk; 23–27.5 kg/m² increased risk; and 27.5 kg/m² or higher high risk. In addition to following these standards, identification of specific public health action points will be an important step towards better prevention of nutritional problems in each country (WHO Expert Consultation 2004).

2.4.3 Limitations of BMI as an indicator for fat mass

Use of BMI as indicator for fat mass of the people is controversial. One limitation of BMI is the reduced ability to differentiate levels of fatness and leanness among individuals. On the other hand assumption of BMI guidelines is that body mass, adjusted for stature squared, is closely associated with body fatness and consequent morbidity and mortality (Bray GA., 1996, Cole TJ., 1991). However, some individuals who are overweight are not over fat (eg, bodybuilders). Others have BMIs within the normal range and yet have a high percentage of their body weight as fat. Although these misclassified persons are uncommon relative to the population as a whole (US Department of Health and Human Services 1998), the question arises as to how they might be evaluated correctly according to body fatness. Unfortunately, there is no consensus on how body fat is linked with morbidity and mortality because of the absence of appropriate prospective studies. Specifically, no accepted published body fat ranges exist; those reported based on empirically set limits, population percentiles, and z scores have serious limitation (Frisancho AR, 1990).

2.5 Deferent methods to asses of body fat

2.5.1 Under water weighing

The most widely used technique of determining body fat is underwater weighing. The technique is based on that the volume of an object submerged in water equals the volume of water the object displaces. Thus, if the mass and the volume of a body are known, the density of the body can be calculated. Using another formula, percent body fat can be calculated from body density (Lee RD, David C. and Boston N., 1996).
2.5.2 Computer Tomography (CT)

Computer Tomography (CT) is an imaging technique producing highly detailed cross-sectional images of the body resulting from differences in the transmission of an X-ray beam through body tissues of differing density. CT has been particularly useful in studying the relative deposition of subcutaneous and intra-abdominal fat. The potential for using CT in assessing body composition and nutritional status has been limited by problems of radiation exposure and the high cost and limited availability of the instrument.

2.5.3 Body impedance analysis (BIA)

An alternative method for body-composition assessment is bioelectrical impedance analysis (BIA). This method has practical features similar to anthropometry (eg, safety, cost-effectiveness, convenience for the patient, and ease of use), and it has been used in large-scale studies of body composition and assessment of body fluid status (Chumlea WC, Guo S., 1994). BIA measures of resistance and impedance are proportional to body water volume, if body electrolyte status is normal, and to the length of the conductor or stature (eg, stature^2/resistance). BIA uses 2-component model (fat and fat-free) as the reference method and identified as a limitation. Methods such as hydrometry (Kushner R, Schoeller D., 1986) and hydro densitometry (Lukaski H, Johnson P, Bolonchuk W, Lykken G., 1985) assume a constant composition of the fat-free body and thus are limited in discriminating differences in body composition when factors such as physical activity, illness, and aging affect a person (Withers RT, Laforgia J, Heymsfield SB, 1999).

2.5.4 Skin fold thickness

Skin fold thickness measurements provide an estimate of the size of the subcutaneous fat depot, which, in turn, provides an estimate of total body fat. Such estimates are based on two assumptions: (a) the thickness of the subcutaneous adipose tissue reflects a constant proportion of the total body fat, and (b) the skin fold sites selected for measurement, either singly or in combination, represents the average thickness of the entire subcutaneous adipose tissue (Rosalind S., Giban, 2005). Assessing skin fold thickness also has own limitations such as high cost of equipment (skin fold caliper)
and requirement of highly skilled person to obtain the measurement. But skin fold measurement is a valid and reliable indicator of body composition.

2.5.4.1 Equations of assessing body fat by measuring skin fold thickness

Predicting percentage of body fat from skin fold measurements requires regression equations. These formulas have been developed by comparing a variety of skinfold and body density measurements. Multiple regression analysis is used and specific equations were derived for groups sharing certain characteristics such as gender and age.

Following generalized skin fold equations were developed by Durnin and Womersley (Deurenberg, P., M. Yap, and W. A. van Staveren, 1998) for males and females. The primary advantage of this approach is that one generalized equation can replace several population specific equations with no loss in prediction accuracy. Since fat placement differs between males and females, separate two equations are developed for each sex.

Generalized Body Composition Equations for Male and Female (Lee RD, David C. and Boston N., 1996)

Males

Percent body fat = 0.29288 \( (X_2) - 0.00050 (X_2)^2 + 0.15845 (X_8) - 5.76377 \)

Females

Percent body fat = 0.29699 \( (X_2) - 0.00043 (X_2)^2 + 0.02963 (X_8) + 1.4072 \)

Where,

\[ X_2 = \text{sum of abdomen, suprailiac, triceps, and thigh skinfolds} \]

\[ X_8 = \text{age in years} \]

2.6 Summery

The literature review has covered basics on body composition, body fat, and also about various methods to assess the body composition. BMI was identified as a commonly used anthropometric method while comparing it with other anthropometric methods.
Also skin fold thickness was identified as reasonably valid and reliable indicator to
assess body fat percentage. Finally literature review has discussed about two regression
equation which is used to derive fat composition from skin fold thickness values. This
knowledgebase will provide a better background for this study. It has explored the
possibility of developing new statistical model using anthropometric methods for better
representation of body fat percentage.
Chapter 3

Methodology
METHODOLOGY

3.1 Introduction

Philosophy of this research is positivism because,

- This research is conducted according to a highly structured methodology.
- Builds a model using quantifiable observations.

According to the previous researches conducted, it was accepted the theories that there is a linear relationship between body fat and anthropometric measurements. With accepting that theory, this study will focus on building a model to find body fat level of Sri Lankan population, using suitable anthropometric measurements. Therefore this research follows a deductive approach and conducted using survey strategy. This research can be categorized as a cross sectional study as data is collected at one point of time.

Remaining topics of this chapter discuss about the methodology which was followed in this research and section 3.2 covers the facts about the adopted study design while section 3.3 discusses about the time frame and the study setting. Section 3.4 has information about the subjects and sample sizes and sampling techniques were discussed in section 3.5. Sections 3.6 describe the detailed facts about data collection and the methods of taking each anthropometric measurement was discussed in this chapter. Finally, section 3.7 has covered about the data analysis methods used in the study.

3.2 Study Design

This study was a cross sectional study conducted to assess the body fat percentage by anthropometric methods to develop prediction equation among office workers.

3.3 Setting and time

The study has targeted both government and private sector workers in Colombo area and the study was carried out from April to August 2008.
3.4 Subjects and sample size

Total of 300 apparently healthy Sri Lankan office workers were selected from both male and females, age ranged from 25-45 years.

3.5 Sampling technique and selection of subjects

Assessment team visited government and private offices in Colombo area and invited for volunteers to take part in the assessment. Sample of them were selected who meet above criteria in sample selection.

3.6 Data collection

Anthropometric measurements and general details were collected and recorded in questioners. (See annex 1) Well trained persons were involved for taking anthropometric measurements to minimize the human error.

3.6.1 General information

General details such as age, sex, marital status and occupation were obtained. Face to face Interviews were used to collect the personal information.

3.6.2 Anthropometrics

Standing height, weight, hip, waist, mid-upper arm circumference, thigh circumference and skin fold thickness of triceps, abdomen, thigh and suprailiac were measured from both males and females. Measurers were well trained to obtain the anthropometric measurements before getting the measurements from study participants.

3.6.2.1 Weight

Body weight was measured using an electrical scale to the nearest 100g. The scale was placed on a level ground. Slippers and any other heavy items were removed. Subjects were still on the center of the platform of the scale to distribute the body weight on both feet evenly without touching any thing else.
3.6.2.2 Height

Height was measured to the nearest 0.1 cm, the barefoot subjects were requested to stand straight in front of the stadiometer. Subject was asked to look straight ahead and to stand with straight legs and heels together. Arms were to the side of the body in natural manner and shoulders were relaxed. Heels, buttock, shoulders’ blade and back of the head touched the pole. Using the above data, body mass index was calculated as weight (kg)/height (m2).

3.6.2.3 Hip

Widest circumference around the buttock was measured as hip circumference using flexible non-stretchable tape.

3.6.2.4 Waist

Lowest circumference below last rib bone and above umbilicus was measured as waist circumference using flexible non-stretchable tape.

3.6.2.5 Mid Upper Arm Circumference

Mid-upper arm circumference (MUAC) was measured using a non stretchable tape, taken from the left arm. Forearm was flexed at 90 degree angle and mid point was marked using the distance between the acromial process of the shoulder and the olecranon process of the elbow. The arm hung loosely by the side of the body. The upper arm was wrapped with tape at the mid point without compressing the soft tissue. Measurements were taken to the nearest 0.1 cm.

3.6.2.6 Thigh circumference

Study participants were asked to sit straight on the chair. Length between inguinal crease and end of the femur on either side of the patella was measured. Middle point of the above measured length was marked. Circumference around that mid point was measured using flexible non-stretchable tape.

3.6.2.7 Skin folds thicknesses

Skin fold thickness was measured using a caliper, at the abdomen, thigh, triceps and suprailliac on the left side of the body. Each skin fold measurement was replicated three
times. Measurement was recorded to the nearest 0.1 mm with minimum interval of 5 seconds of each measurement. This study has followed above mentioned standard methods carefully to reach the most accurate results.

3.6.2.7.1 Triceps

Using a non-stretchable flexible tape, mid-point of the left arm was marked. Subject was asked to keep the hand loosely by the side with palm of the hand facing anteriorly. Posterior midlines were marked at the level of the mid point. Measurer stood behind the subject to measure the triceps. The skin fold was grasped by the thumb and index finger of the left hand about 1 cm or \( \frac{1}{2} \) in. proximal to the marked site. The caliper was held in the right hand perpendicular to the long axis of the skin fold and dial facing up as easily readable. The caliper was placed on the site about 1 cm to the finger holding the skin fold and took measurement to the nearest 0.1 mm.

3.6.2.7.2 Abdomen

The subject stood erectly with the booby weight evenly distributed on both feet, abdominal muscles relaxed, and breathing quietly. A horizontal skin fold 3cm to the right of and 1cm below the midpoint of the umbilicus was measured.

3.6.2.7.3 Suprailiac

Suprailiac skin fold was measured just above the iliac crest at the midaxillary line. The long axis of the skin fold was parallel to the natural cleavage line of the skin. Subject stood erectly with feet together. Measurer stood to the left side of the subject. The skin fold was grasped by the thumb and index finger of the left hand about 1 cm anterior to the midaxillary line. The caliper was perpendicular to long axis of the skin fold and held by the right hand. The tips of the caliper were placed at the midaxillary line with calipers’ dial facing up and easily readable.

3.6.2.7.4 Thigh skin fold

This site was a vertical skin fold along the midline of the anterior aspect of the thigh midway between the junction of the midline and the inguinal crease and the proximal border of the patella, or knee cap. The subject was shifts the weight to the left foot and
relaxes the lay being measured by slightly flexing the knee with the foot flat on the floor.

The caliper was held in the right hand perpendicular to the long axis of the skin fold and dial facing up as easily readable. The caliper was placed on the site about 1cm to the finger holding the skin fold and took measurement to the nearest 0.1mm.

3.7 Data Analysis

3.7.1 Anthropometric indices and cut off values

Above mentioned anthropometric measurements were taken and relevant anthropometric indices were calculated using following equations. Cutoff values used in each anthropometric index were also discussed in following section. Study subjects can be classified based on the risk of non-communicable diseases, using these indices and relevant cutoff values.

BMI of all subjects was calculated using the following equation.

\[
\text{BMI} = \frac{\text{weight (Kg)}}{\text{(Height (m))}^2} \rightarrow (1)
\]

Based on BMI values study participants were grouped into following four categories.

- Underweight < 18.5 kg/m²;
- Normal 18.5–23 kg/m²
- Over weight- 23–27.5 kg/m²
- Obese - > 27.5 kg/m²

Waist to hip ratio (WHR) was calculated with following equation.

\[
\text{Waist: Hip} = \frac{\text{Waist}}{\text{Hip}} \rightarrow (2)
\]

Based on WHR values study participants were grouped into following two categories for males and females separately.

For males

- Acceptable < 0.9
- Risk ≥ 0.9
For females  

Acceptable <0.8  

Risk ≥ 0.8

Waist to height ratio (WHTR) will be calculated by,

\[
\text{Waist: Height} = \frac{\text{Waist}}{\text{Height}} \rightarrow (3)
\]

WHTR values were grouped into two categories for both male and female.

Normal < 0.5  

At risk ≥ 0.5

Above results were compared with the existing standard cutoff values and categorized based on the specific risk category.

Body fat percentage determined using triceps, suprailliac, abdomen, thigh skinfold thickness from the previously developed equation. Subjects were classified into risk groups based on the results, and the classification was used as the reference for validating the other anthropometric measurements/indices.

Following equations were used to calculate body fat percentage using above measurements taken.

**Males**

\[
\text{Percent body fat} = 0.29288 \ (X_2) - 0.00050 \ (X_2)^2 + 0.15845 \ (X_8) - 5.76377 \rightarrow (4a)
\]

**Females**

\[
\text{Percent body fat} = 0.29699 \ (X_2) - 0.00043 \ (X_2)^2 + 0.02963 \ (X_8) + 1.4072 \rightarrow (4b)
\]

Where,

\[
X_2 = \text{sum of abdomen, suprailliac, triceps, and thigh skinfolds}
\]

\[
X_8 = \text{age in years}
\]

Internationally accepted following cut off values for both men and women were provided to categorize subjects into groups based on the risk of non-communicable chronic diseases.
Acceptable body fat percentage for male  -  < 25%
Risk body fat percentage for male  -  >= 25%
Acceptable body fat percentage for female  -  < 32%
Risk body fat percentage for female  -  >= 32%

3.7.2 Statistical approach

Various statistical methods will be used such as best subset regression models, prediction intervals, cross tables etc. to analyze the data set using SPSS and MINITAB statistical software packages.

3.7.2.1 Variables used in analysis

Response variable

➢ Body fat percentage

Body fat level was identified as a main risk factor for non-communicable chronic diseases and it was repeatedly proved by several studies in various countries. Also the slandered method used in this study has measured the skin fold thickness which has close relationship with the thickness of subcutaneous fat layer. Therefore this can be considered as more accurate and reliable method to determine the risk levels. Also the standard equations and cutoff values used are internationally accepted and used in various countries therefore valid for Sri Lankan population.

Independent variables

➢ Age
➢ Height
➢ Weight
➢ Hip
➢ Waist
Thigh circumference
Mid upper arm circumference

All of these anthropometric measurements are used individually or as combinations, to determine the risk of non-communicable diseases. Even though these anthropometric measurements do not directly measure the body fat, they have shown close relationships with the risk of non-communicable diseases, especially when used as anthropometric indices which combine two or more measurements.

Categorical Variables
  ➢ Sex
  ➢ Marital Status
  ➢ Occupation

3.7.2.2 Missing values

Missing value estimation was not done due to the impact for overall accuracy. However the study has paid close attention to maintain the accuracy and removed whole row for any subject which records missing values. Data collected from more than 300 subjects as some of subjects had to be removed due to missing values.

3.7.2.3 Statistical analysis

Data collected from apparently healthy adults from private and government sector institutes as described in sections 3.4, 3.5 and 3.6. These collected data were cross checked and accuracy was assessed. Then data fed into SPSS statistical package and data were cleaned.

Then anthropometric indices were calculated using equation 1, 2, and 3 discussed in section 3.7.1. Relationship between body fat percentage and each regressor variable was checked. Then, it was checked the correlation of regressor variables. A linear regression model was built with all regressors. Then a simple and reliable model was developed using best subset regression method. Further, prediction interval for the new observation was identified to increase accuracy of predicting health risk.
Chapter 4

Data Analysis
DATA ANALYSIS

4.1 Introduction

Data analysis can be considered as one of the important sections of the research as results are generated as main outcome which help to reach a conclusion. First, this chapter describes the data cleaning process which focused on correcting data entry errors. Then basic descriptive analysis was included which relevant to the response variable. In addition to above mentioned primary analysis, some regression models were built using regression model building techniques. Model adequacy was checked and compared between models to reach better conclusions. Finally the one best model was compared with the anthropometric indices.

4.2 Data cleaning process

Data cleaning is important as it helps to minimize the errors and wrong results and conclusions based on incompatible data. In addition, categorical variables were checked for data entry errors.

4.3 Descriptive analysis of response variables

Graph 4.1: Individual relationship between response variable (BF%) vs each regressor variable.
According to above scatter plot there is a linear relationship between response variable and each regressor variables. Positive linear relationships were identified between response variable (BF%) and all other regressor variables except height. Negative linear relationship was identified between BF% and height. In addition, these scatter plots shows gender is significant variable.

4.4 Linear regression models

4.4.1 Linear regression model with all regressor variables

The regression equation is

\[
BF% = 20.2 - 6.97 \text{ Gender} - 0.0169 \text{ age} - 0.192 \text{ Married} \\
+ 0.144 \text{ Weight} - 0.175 \text{ Height} - 0.0709 \text{ Hip} \\
+ 0.335 \text{ Waist} - 0.0785 \text{ MUAC} + 0.124 \text{ TC} \\
+ 1.26 \text{ Occupation}
\]

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>SE Coef</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>20.195</td>
<td>4.908</td>
<td>4.12</td>
<td>0.000</td>
</tr>
<tr>
<td>Gender</td>
<td>-6.9745</td>
<td>0.3651</td>
<td>-19.10</td>
<td>0.000</td>
</tr>
<tr>
<td>age</td>
<td>-0.01688</td>
<td>0.03835</td>
<td>-0.44</td>
<td>0.660</td>
</tr>
<tr>
<td>Married</td>
<td>-0.1919</td>
<td>0.3101</td>
<td>-0.62</td>
<td>0.536</td>
</tr>
<tr>
<td>Weight</td>
<td>0.14444</td>
<td>0.02609</td>
<td>5.54</td>
<td>0.000</td>
</tr>
<tr>
<td>Height</td>
<td>-0.17507</td>
<td>0.02693</td>
<td>-6.50</td>
<td>0.000</td>
</tr>
<tr>
<td>Hip</td>
<td>-0.07087</td>
<td>0.03582</td>
<td>-1.98</td>
<td>0.049</td>
</tr>
<tr>
<td>Waist</td>
<td>0.33510</td>
<td>0.03541</td>
<td>9.46</td>
<td>0.000</td>
</tr>
<tr>
<td>MUAC</td>
<td>-0.07848</td>
<td>0.07292</td>
<td>-1.08</td>
<td>0.283</td>
</tr>
<tr>
<td>TC</td>
<td>0.12416</td>
<td>0.03535</td>
<td>3.51</td>
<td>0.001</td>
</tr>
<tr>
<td>Occupation</td>
<td>1.2619</td>
<td>0.3620</td>
<td>3.49</td>
<td>0.001</td>
</tr>
</tbody>
</table>

\[ S = 2.08326 \quad R-Sq = 88.8% \quad R-Sq(adj) = 88.4% \]

Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>10</td>
<td>9915.82</td>
<td>991.58</td>
<td>228.48</td>
<td>0.000</td>
</tr>
<tr>
<td>Residual Error</td>
<td>289</td>
<td>1254.25</td>
<td>4.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>299</td>
<td>11170.07</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.4.1.1 Model adequacy

Normality of the residuals

Graph 4.2: Normality of residuals

Normality assumption of residuals was confirmed as the above result shows the P-value (0.63) which is greater than the significant level 0.05.

Constant variance of residuals

Graph 4.3: Constant variance of residuals
Above graph shows that the residuals are approximately spreaded as horizontal band and also plotted symmetrically around zero. This shows that the residuals have approximately constant variance. Also above findings implies that the resulted model is linear.

According to the P-values of above t-test age, marital status and mid upper arm circumference (MUAC) were not significant at the significance level of 0.05.

4.4.2 Developing regression model using best subset regression method

Study aims to develop a simple and reliable model to predict body fat percentage; therefore best subset regression method was used only for significant variables in above model and result is given below.

**Best Subsets Regression: BF% versus Gender, Weight, ...**

| Vars | R-Sq | R-Sq(adj) | Mallows | C-p | S     | Occ | Cup | GWH | eee | W | pai | nii | a | t | dgg | Hi | i | ehh | his | To | rtt | pt | Cn |
|------|------|----------|---------|------|-------|-----|-----|-----|-----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1    | 47.4 | 47.3     | 1062.9  | 4.4389 | X     |     |     |     |     |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 1    | 46.5 | 46.3     | 1087.1  | 4.4782 | X     |     |     |     |     |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 2    | 85.5 | 85.4     | 82.0    | 2.3389 | X     | X    |     |     |     |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 2    | 73.8 | 73.6     | 384.5   | 3.1418 | X     | X    |     |     |     |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 3    | 86.2 | 86.1     | 63.5    | 2.2781 | X     | X    | X   |     |     |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 3    | 86.0 | 85.8     | 70.2    | 2.2994 | X     | X    | X   |     |     |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 4    | 87.7 | 87.5     | 27.8    | 2.1574 | X     | X    | X   | X   |     |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 4    | 86.8 | 86.6     | 52.5    | 2.2397 | X     | X    | X   | X   |     |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 5    | 88.2 | 88.0     | 16.3    | 2.1148 | X     | X    | X   | X   |     |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 5    | 88.0 | 87.8     | 21.7    | 2.1334 | X     | X    | X   | X   |     |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 6    | 88.5 | 88.3     | 10.3    | 2.0903 | X     | X    | X   | X   | X   |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 6    | 88.3 | 88.0     | 17.2    | 2.1146 | X     | X    | X   | X   | X   |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| 7    | 88.7 | 88.4     | 8.0     | 2.0787 | X     | X    | X   | X   | X   |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
Highlighted models, of above results, were fitted and then checked for model adequacy and accuracy. The 5th model was selected as the best model. Gender, waist and occupation were the regressors contained in that model. Reason for this selection was that, it satisfied model assumptions better, and also its accuracy was higher.

4.4.2.1 Regression model according to the best sub set with interaction

The regression equation is

\[ BF\% = -6.55 - 23.1 \text{ Gender} + 0.389 \text{ Waist} + 10.9 \text{ Occupation} + 0.170 \text{ waist*gender} - 0.0940 \text{ waist*occu} \]

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>SE Coef</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-6.553</td>
<td>1.220</td>
<td>-5.37</td>
<td>0.000</td>
</tr>
<tr>
<td>Gender</td>
<td>-23.130</td>
<td>2.264</td>
<td>-10.22</td>
<td>0.000</td>
</tr>
<tr>
<td>Waist</td>
<td>0.38881</td>
<td>0.01407</td>
<td>27.63</td>
<td>0.000</td>
</tr>
<tr>
<td>Occupation</td>
<td>10.939</td>
<td>4.605</td>
<td>2.38</td>
<td>0.018</td>
</tr>
<tr>
<td>waist*gender</td>
<td>0.17031</td>
<td>0.02457</td>
<td>6.93</td>
<td>0.000</td>
</tr>
<tr>
<td>waist*occu</td>
<td>-0.09405</td>
<td>0.04742</td>
<td>-1.98</td>
<td>0.048</td>
</tr>
</tbody>
</table>

\[ S = 2.10633 \quad R\text{-Sq} = 88.3\% \quad R\text{-Sq(adj)} = 88.1\% \]

Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>5</td>
<td>9865.7</td>
<td>1973.1</td>
<td>444.74</td>
<td>0.000</td>
</tr>
<tr>
<td>Residual Error</td>
<td>294</td>
<td>1304.4</td>
<td>4.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>299</td>
<td>11170.1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.4.2.1.1 Model adequacy

Normality of the residuals

Graph 4.4: Normality of residuals

Constant variance of residuals

Graph 4.5: Constant variance of residuals
P-values of above t-test shows that all regressors were significant at the 0.05 significance level. Also 88.1% of variability of body fat percentage is explained by this model. But normality assumption of residuals was violated at the 0.05 level of significance. Interaction between waist and occupation was removed by considering P-values of t-test as that P-value was relatively higher and almost reached to the 0.05 significance level. Then regression model was fitted with other regressors except above mentioned interaction between waist and occupation.

**4.4.2.2 Best fitted regression model**

The regression equation is

\[
BF\% = -6.06 - 23.2 \text{ Gender} + 0.383 \text{ Waist} + 1.84 \text{ Occupation} + 0.171 \text{ waist*gender}
\]

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>SE Coef</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-6.06</td>
<td>1.200</td>
<td>-5.05</td>
<td>0.000</td>
</tr>
<tr>
<td>Gender</td>
<td>-23.17</td>
<td>2.275</td>
<td>-10.19</td>
<td>0.000</td>
</tr>
<tr>
<td>Waist</td>
<td>0.383</td>
<td>0.01386</td>
<td>27.66</td>
<td>0.000</td>
</tr>
<tr>
<td>Occupation</td>
<td>1.835</td>
<td>0.3645</td>
<td>5.04</td>
<td>0.000</td>
</tr>
<tr>
<td>waist*gender</td>
<td>0.170</td>
<td>0.02469</td>
<td>6.92</td>
<td>0.000</td>
</tr>
</tbody>
</table>

S = 2.11677     R-Sq = 88.2%  R-Sq(adj) = 88.0%

**Analysis of Variance**

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>4</td>
<td>9848.3</td>
<td>2462.1</td>
<td>549.48</td>
<td>0.000</td>
</tr>
<tr>
<td>Residual Error</td>
<td>295</td>
<td>1321.8</td>
<td>4.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>299</td>
<td>11170.1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.4.2.2.1 Model adequacy

Normality of the residuals

Graph 4.6: Normality of residuals

Normality assumption of residuals was confirmed as the above result shows the P-value (0.223) which is greater than the significant level 0.05.

Constant variance of residuals

Graph 4.7: Constant variance of residuals
Above graph shows that the residuals are approximately spreaded as horizontal band and also plotted symmetrically around zero. This shows that the residuals have approximately constant variance. Also above findings implies that the resulted model is linear.

According to the P-values of t-test, all regressors were significantly contributed to the model at the 0.05 level of significance. Also this model explains 88% of variability in body fat percentage.

Following output was obtained to check the Cp value of the fitted model discussed above.

<table>
<thead>
<tr>
<th>Vars</th>
<th>R-Sq</th>
<th>R-Sq(adj)</th>
<th>Mallows C-p</th>
<th>Mallows S</th>
<th>rtnt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>46.5</td>
<td>46.3</td>
<td>1037.7</td>
<td>4.4782</td>
<td>X</td>
</tr>
<tr>
<td>1</td>
<td>20.0</td>
<td>19.7</td>
<td>1698.7</td>
<td>5.4765</td>
<td>X</td>
</tr>
<tr>
<td>2</td>
<td>85.5</td>
<td>85.4</td>
<td>68.6</td>
<td>2.3389</td>
<td>X X</td>
</tr>
<tr>
<td>2</td>
<td>83.2</td>
<td>83.1</td>
<td>125.2</td>
<td>2.5147</td>
<td>X X</td>
</tr>
<tr>
<td>3</td>
<td>87.1</td>
<td>87.0</td>
<td>28.4</td>
<td>2.2021</td>
<td>X X X</td>
</tr>
<tr>
<td>3</td>
<td>86.2</td>
<td>86.1</td>
<td>50.8</td>
<td>2.2781</td>
<td>X X X</td>
</tr>
<tr>
<td>4</td>
<td>88.2</td>
<td>88.0</td>
<td>5.0</td>
<td>2.1168</td>
<td>X X X X</td>
</tr>
</tbody>
</table>

Above results showed that the fitted model has lower Cp value and residual mean square with higher R-sq (adj). According to the above results, the fitted model can be considered as best one to predict body fat percentage.
Best fitted model is

\[ BF\% = -6.06 - 23.2 \text{ Gender} + 0.383 \text{ Waist} + 1.84 \text{ Occupation} + 0.171 \text{ waist*gender} \]

Where,

Gender is categorized as follows.

1- Male
0 - Female

Fitted linear regression model to predict body fat percentage of males

\[ BF\% = -29.26 + 0.554 \text{ Waist} + 1.84 \text{ Occupation} \]

Fitted linear regression model to predict body fat percentage of females

\[ BF\% = -6.06 + 0.383 \text{ Waist} + 1.84 \text{ Occupation} \]

Where,

Occupation is a categorical variable which take value 1 for those with less physical activity level, while 0 for those who are physically active.

4.4.2.3 Prediction interval for the observation

Health risk of each individual can be accurately predicted using, individual interval estimation, than mean estimation given by the fitted model. Therefore one sided prediction interval is proposed based on the fitted model and it is as follows.

According to the above findings, 95 % prediction interval for the new observation \( y_0 \) at \( x = x_0 \) is as follows.

\[ 0 \leq y_0 \leq \hat{y}_0 + t_{a,n-2} [\text{MSE}(1+1/n+(x_0-\bar{x})^2/S_{xx})]^{1/2} \]

Following are the main steps of assessing the risk of an individual, using the proposed prediction interval which is based on the fitted regression model.

First, find the prediction interval for an individual based on the fitted regression model.

Then, if the cutoff value for specific gender of an individual lies within the interval, person can be classified as risk, with 95% confidence.
Cut off values of body fat percentage to identify health risk, is as follows.

Risk body fat percentage for male - >= 25%
Risk body fat percentage for female - >= 32%

4.5 Accuracy comparison of interval estimation with commonly used anthropometric indices

Accuracy of interval estimation was checked and compared with the accuracy of most common anthropometric indices such as BMI, WHR and WHtR.

4.5.1 Accuracy comparison of interval estimation with anthropometric indices

Following results shows the accuracy comparison of interval estimation and other anthropometric indices.

<table>
<thead>
<tr>
<th>Classified Group</th>
<th>True Group</th>
<th>Risk</th>
<th>Not Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk</td>
<td>124</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>Not Risk</td>
<td>0</td>
<td>125</td>
<td></td>
</tr>
</tbody>
</table>

Accuracy of interval estimation = \( \frac{(124+125) \times 100}{300} \) = 83%

<table>
<thead>
<tr>
<th>Classified Group</th>
<th>True Group</th>
<th>Risk</th>
<th>Not Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk</td>
<td>112</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>Not Risk</td>
<td>12</td>
<td>124</td>
<td></td>
</tr>
</tbody>
</table>

Accuracy of BMI = \( \frac{(112+124) \times 100}{300} \) = 78.7%
The interval estimation has clearly shown a higher accuracy, when it was compared with the accuracy of other anthropometric indices.

### 4.5.2 Finding suitable cut off values for anthropometric indices suits with Sri Lankan context

The study has focused on other anthropometric indices and the suitability of their cutoff values, to be used in the Sri Lankan context. BMI has shown 4% of more serious error which implies that the cutoff values should be adjusted to suit the local context. Therefore suitable cutoff values were found as follows.

Risk groups were separated from the study sample and their BMI values were checked separately for Male and Female groups. It this process, some of the risk individuals has reported smaller BMI values than the standard cutoff value (27.5) and the observed minimum BMI values were deferent for Males (26) and Females (25.3). As per these results, the study suggests 25 as the suitable cutoff value for the Sri Lankan context.

Accuracy was checked for the adjusted BMI cutoff value (25) and results are as follows.
According to the above results, overall accuracy of the BMI was 72.3 % after adjustment, while keeping more serious error in 0 level.

Waist to Hip ratio and Waist to Height ratio has also undergone through the same process and found that the existing standard cutoff values has accurately classified risk group. Therefore those cutoff values can be considered as suitable for Sri Lankan context.
Chapter 5

Conclusion
CONCLUSION

Following two models can be considered as convenient methods to predict body fat percentages in Sri Lankan office workers.

1. Linear regression model to predict body fat percentage of male

\[ BF\% = -29.26 + 0.554 \text{ Waist} + 1.84 \text{ Occupation} \]

2. Linear regression model to predict body fat percentage of female

\[ BF\% = -6.06 + 0.383 \text{ Waist} + 1.84 \text{ Occupation} \]

Where,

Occupation is a categorical variable which take value 1 for those with less physical activity level, while 0 for those who are physically active.

One sided prediction interval, based on the above regression model, was proposed to predict the health risk.

95 % prediction interval for the new observation \( y_0 \) at \( x = x_0 \), which is related with above model, is as follows.

\[ 0 \leq y_0 \leq \hat{y}_0 + t_{\alpha,n-2} [\text{MSE}(1+1/n+(x_0-\bar{x})^2/S_{x_0})]^{1/2} \]

This is the one of the most important finding of this study and prediction interval can be suggested as the best method of assessing health risk.

Anthropometric indices such as Body Mass Index, waist to height ratio and waist to hip ratio are widely used at present to assess the risks related to body fat percentage. Above interval estimation was proved that it has higher accuracy to predict health risk of an individual when compared to those anthropometric indices.

This study proposed a reliable method to predict health risk, however the accuracy and reliability can be further increase by extending the same study to a larger sample.
REFERENCES


38


Annex-1
Department of Applied Science, Faculty of Graduate Studies
University of Sri Jayawardenepura

Re. No _____________________________

General Questionnaire to Assess Statistical Model for Better Nutrition Assessment
Using Anthropometric Methods

1) General Information
1.1 Gender: Male-1 _____ Female-2 _____
1.2 Birth Day: D _____ M _____ Y _____ Age: _____
1.3 Occupation _____________________________
1.4 What is your marital Status?
   Married-1 _____ Unmarried-2 _____
1.5 Appearance
   Obesity-1 _____ Non-obesity-2 _____

2) Measurements
2.1 Weight: _____________________________
2.2 Height: _____________________________
2.3 Hip: _____________________________
2.4 Waist: _____________________________
2.5 MUAC: 1 2 3 _____________________________
2.6 TC: 1 2 3 _____________________________
2.7 Skin fold thickness:
   Subscapular 1 2 3 _____________________________
   Suprailiac 1 2 3 _____________________________
   Triceps 1 2 3 _____________________________
   Thigh Skinfold 1 2 3 _____________________________