Chapter-VII

Study of Essential Hypertension among the Euglycemic Patients Having Ischemic Heart Disease

7.1 Introduction

The hypertension is one of the major health problems in present life and practice, which is the root of many diseases. Castelli (1983) was the first to review the cardiovascular diseases in terms of the risk factor parameters in the Framingham Heart Study over a period of 18 years. Kaplan (1992) declared that hypertension is a ‘Silent Killer’ because most of the times it is asymptomatic and goes undetected. Kaplan (1994) also came up with the facts that most people would develop hypertension during their life time. Kannel(1996) gave the prevention and treatment of blood pressure. Gupta (1997) performed a meta-analysis of all available epidemiological studies and concluded that there was a significant increase in the systolic blood pressure but not in the diastolic blood pressure in India. Dusten (1999) showed that hypertension is one of the major risks factors for various diseases like coronary artery disease, stroke, cardiac failure and renal damage. Weibert (2000) gave the fact that more than 90% of patients have essential hypertension and Carter and Saseen (2002) stated that fewer than 5% of people have secondary hypertension. Campanini (2002) stated that hypertension is the leading cause of death world wide and one of the world’s great public health problems. Recently, Bhavani et. al. (2005) worked on lipid profile and apolipoprotein epolymorphism in essential hypertension and Massie et al (2005) proposed current medical diagnosis and treatment relating to hypertension. Hajjar et
al (2006) showed that prior to 1990, population data suggest that hypertension prevalence was decreasing; however, recent data suggest that it is again on the rise. Hence, programs that improve hypertension control rates and prevent hypertension are urgently required. A person is said to be hypertensive if his/her systolic blood pressure is more than 139.99mm Hg and diastolic blood pressure is more than 89.99mm Hg. Body mass index is calculated as: Weight (Kg)/Height$^2$ (meter). If body mass index is less than 18.50 then the person is under weight and if it is more than 29.99 then obese. SPSS 14.0 software has been used to analyze the data.

### 7.2 Proposed Formulation and Results

To study the primary hypertension, it is necessary to measure the systolic blood pressure (SBP) and diastolic blood pressure (DBP). The SBP and DBP vary with the change in the value of many variables such as body mass index (BMI), total cholesterol (TC), high density lipoprotein-cholesterol (HDL), low density lipoprotein-cholesterol (LDL), tri-glyceride (TG), fasting blood sugar (FBS), insulin (I), blood urea (BU), serum creatinine (SC), serum glutamate oxaloacetate transeminase (SGOT), serum glutamate pyruvate transeminase (SGPT) etc. So, our main aim is to predict the values of SBP and DBP to identify the hypertension which are functions of independent risk factors such as BMI, TC, HDL, LDL, TG, FBS, I, BU, SC, SGOT, SGPT.

The regression is a statistical technique with the help of which we are in a position to predict the value of dependent variable for the given vales of independent variables. The regression of SBP and DBP can be written mathematically as:
SBP = \( f_1(BMI, TC, HDL, LDL, TG, FBS, I, BU, SC, SGOT, SGPT) \) \hspace{1cm} (7.2.1) \\
DBP = \( f_2(BMI, TC, HDL, LDL, TG, FBS, I, BU, SC, SGOT, SGPT) \) \hspace{1cm} (7.2.2)

In (7.2.1) and (7.2.2), we assume that the factors involved in the functions are linearly related. So, under this assumption we can write the functions \( f_1 \) and \( f_2 \) in general as:

\[
SBP = f_1(X_1, X_2, \ldots, X_n) = Y_1 = a_0 + a_1X_1 + a_2X_2 + \ldots + a_nX_n + \varepsilon_1 \tag{7.2.3}
\]

\[
DBP = f_2(X_1, X_2, \ldots, X_n) = Y_2 = b_0 + b_1X_1 + b_2X_2 + \ldots + b_nX_n + \varepsilon_2 \tag{7.2.4}
\]

Here \( X_i \); \( i = 1, 2, \ldots, n \) are the independent variables, \( a_i \); \( i = 0, 1, 2, \ldots, n \) and \( b_i \); \( i = 0, 1, 2, \ldots, n \) are real constant whereas \( \varepsilon_1 \) and \( \varepsilon_2 \) are assumed to be normally distributed with mean zero and constant variance. The values of constants involved is obtained by using the principal of least squares which implies that
\[ \hat{a}_0 = \bar{Y} - \hat{a}_1 \bar{X}_1 + \hat{a}_2 \bar{X}_2 + ... + \hat{a}_n \bar{X}_n \]

\[ \hat{a}_i = -\frac{\sigma_{y_i}}{\sigma_{x_i}} \frac{\omega_{i(l+i)}}{\omega_{11}} \quad ; \quad i = 1, 2, \ldots, n \quad (7.2.5) \]

\[ \hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}_1 + \hat{b}_2 \bar{X}_2 + ... + \hat{b}_n \bar{X}_n \]

\[ \hat{b}_i = -\frac{\sigma_{y_i}}{\sigma_{x_i}} \frac{\omega_{i(l+i)}}{\omega_{11}} \quad ; \quad i = 1, 2, \ldots, n \quad (7.2.6) \]

Here, \( \omega_{i(l+i)} \) and \( \omega_{11} \) is the co-factor of the element in the 1\textsuperscript{st} row and (i+1)\textsuperscript{th} column of \( \omega \) defined below:

\[
\omega = \begin{bmatrix}
1 & r_{x_1y} & r_{x_2y} & \ldots & r_{x_ny} \\
r_{x_1y} & 1 & r_{x_2x_1} & \ldots & r_{x_nx_1} \\
r_{x_1x_2} & r_{x_2x_1} & 1 & \ldots & r_{x_nx_2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
r_{x_1x_n} & r_{x_2x_n} & r_{x_3x_n} & \ldots & 1
\end{bmatrix}
\]

To estimate the values of constants involved in the multiple linear regression models (7.2.3) & (7.2.4), we took the data on random sample of 80 persons selected in the period 2006-2009 by the department of Medicine and Bio-chemistry at Govt. Rajindera Hospital, Patiala in which observations on systolic blood pressure, diastolic blood pressure, body mass index, fasting blood sugar, blood urea, serum creatinine, serum glutamate oxaloacetate transeminase, serum glutamate pyruvate transeminase, total cholesterol, high density lipoprotein-cholesterol, low
density lipoprotein-cholesterol, tri-glyceride and insulin on each person have been obtained (see appendix). To check which factors are significantly affecting the SBP and DBP, we use t-test on partial regression coefficients of SBP and DBP independently i.e. we calculate

\[ t_j = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \]

; \ j = 1, 2, \ldots, n

On the basis of the data, we have first obtained the values of statistic-t which are given in the following Table-7.2.1 and Table-7.2.2:

Table-7.2.1

<table>
<thead>
<tr>
<th>Variable</th>
<th>Partial Regression Coefficient of SBP</th>
<th>Standard Error of Estimate</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>110.186</td>
<td>34.915</td>
<td>3.156</td>
<td>.002</td>
</tr>
<tr>
<td>BMI</td>
<td>1.617</td>
<td>.324</td>
<td>.460</td>
<td>3.511</td>
</tr>
<tr>
<td>FBS</td>
<td>.139</td>
<td>.046</td>
<td>.286</td>
<td>.485</td>
</tr>
<tr>
<td>BU</td>
<td>-.950</td>
<td>-.139</td>
<td>.827</td>
<td>-1.149</td>
</tr>
<tr>
<td>SC</td>
<td>18.074</td>
<td>.140</td>
<td>15.630</td>
<td>1.156</td>
</tr>
<tr>
<td>SGOT</td>
<td>-.301</td>
<td>-.101</td>
<td>.451</td>
<td>-.668</td>
</tr>
<tr>
<td>SGPT</td>
<td>.146</td>
<td>.053</td>
<td>.406</td>
<td>.359</td>
</tr>
<tr>
<td>TC</td>
<td>.103</td>
<td>.147</td>
<td>.325</td>
<td>.317</td>
</tr>
<tr>
<td>HDL</td>
<td>-.879</td>
<td>-.255</td>
<td>.375</td>
<td>-2.344</td>
</tr>
<tr>
<td>TG</td>
<td>.084</td>
<td>.253</td>
<td>.058</td>
<td>1.450</td>
</tr>
<tr>
<td>LDL</td>
<td>-.177</td>
<td>-.202</td>
<td>.274</td>
<td>-1.646</td>
</tr>
<tr>
<td>Insulin</td>
<td>1.270</td>
<td>.380</td>
<td>.967</td>
<td>1.313</td>
</tr>
</tbody>
</table>
Table-7.2.2

Regression Coefficients of Model (7.2.4)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Partial Regression Coefficient of DBP</th>
<th>Standard Error of Estimate</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>74.156</td>
<td>21.299</td>
<td>3.482</td>
<td>.001</td>
</tr>
<tr>
<td>BMI</td>
<td>.982</td>
<td>.348</td>
<td>.281</td>
<td>3.497</td>
</tr>
<tr>
<td>FBS</td>
<td>-.017</td>
<td>-.010</td>
<td>.175</td>
<td>-.098</td>
</tr>
<tr>
<td>BU</td>
<td>.157</td>
<td>.041</td>
<td>.504</td>
<td>.312</td>
</tr>
<tr>
<td>SC</td>
<td>-1.308</td>
<td>-.018</td>
<td>9.535</td>
<td>-.137</td>
</tr>
<tr>
<td>SGOT</td>
<td>-.275</td>
<td>-.163</td>
<td>.275</td>
<td>-.999</td>
</tr>
<tr>
<td>SGPT</td>
<td>.170</td>
<td>.109</td>
<td>.247</td>
<td>.686</td>
</tr>
<tr>
<td>TC</td>
<td>.128</td>
<td>.323</td>
<td>.198</td>
<td>.646</td>
</tr>
<tr>
<td>HDL</td>
<td>-.560</td>
<td>-.288</td>
<td>.229</td>
<td>-2.449</td>
</tr>
<tr>
<td>TG</td>
<td>.007</td>
<td>.035</td>
<td>.035</td>
<td>.188</td>
</tr>
<tr>
<td>LDL</td>
<td>-.198</td>
<td>-.399</td>
<td>.167</td>
<td>-1.183</td>
</tr>
<tr>
<td>Insulin</td>
<td>.911</td>
<td>.483</td>
<td>.590</td>
<td>1.545</td>
</tr>
</tbody>
</table>

From the Table-7.2.1 and Table-7.2.2, we see that there are only BMI and HDL factors which are significantly affecting the SBP and DBP. The effect of other factors seemed to be insignificant. But actually, we have noted that non-diabetic persons having normal kidney and liver were selected. That is why the values of statistic-t in the Table-7.2.1 and Table-7.2.2 are showing insignificant effect of rest of the variables. Practically, HDL, TG and LDL are the part cholesterol, so we have consider BMI, HDL, TG and LDL factors in the models (7.2.3) and (7.2.4) for predicting SBP and DBP.

Taking

\[ Y_1 = SBP \quad Y_2 = DBP \]

\[ X_1 = BMI \quad X_2 = HDL \]
\[ X_3 = TG \quad X_4 = LDL \]

From the data, the values of partial regression coefficients and constants corresponding to models (7.2.3) & (7.2.4) are obtained and given in Table-7.2.3 and Table-7.2.4 respectively.

### Table-7.2.3

**Regression Coefficients of Model (5.2.3)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Partial Regression Coefficient of SBP</th>
<th>Standard Error of Estimate</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unstandardized Coefficients ( \hat{a}_i )</td>
<td>Standardized Coefficients ( \hat{a}_i^{\ast} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Constant)</td>
<td>93.530</td>
<td>17.796</td>
<td>5.256</td>
<td>.000</td>
</tr>
<tr>
<td>BMI</td>
<td>1.543</td>
<td>.309</td>
<td>.440</td>
<td>3.506</td>
</tr>
<tr>
<td>HDL-C</td>
<td>-.731</td>
<td>-.212</td>
<td>.305</td>
<td>-2.394</td>
</tr>
<tr>
<td>TG</td>
<td>.137</td>
<td>.413</td>
<td>.029</td>
<td>4.776</td>
</tr>
<tr>
<td>LDL-C</td>
<td>.194</td>
<td>.221</td>
<td>.078</td>
<td>2.476</td>
</tr>
</tbody>
</table>

### Table-7.2.4

**Regression Coefficients of Model (5.2.4)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Partial Regression Coefficient of SBP</th>
<th>Standard Error of Estimate</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unstandardized Coefficients ( \hat{b}_i )</td>
<td>Standardized Coefficients ( \hat{b}_i^{\ast} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Constant)</td>
<td>60.740</td>
<td>10.836</td>
<td>5.605</td>
<td>.000</td>
</tr>
<tr>
<td>BMI</td>
<td>.920</td>
<td>.326</td>
<td>.268</td>
<td>3.433</td>
</tr>
<tr>
<td>HDL-C</td>
<td>-.371</td>
<td>-.191</td>
<td>.186</td>
<td>-1.997</td>
</tr>
<tr>
<td>TG</td>
<td>.059</td>
<td>.315</td>
<td>.018</td>
<td>3.374</td>
</tr>
<tr>
<td>LDL-C</td>
<td>.114</td>
<td>.231</td>
<td>.048</td>
<td>2.399</td>
</tr>
</tbody>
</table>
Substituting the values from above tables, we obtain

\[ SBP = 93.530 + 1.543 \text{BMI} - 0.731 \text{HDL} + 0.137 \text{TG} + 0.194 \text{LDL} + \epsilon_1 \]  \hspace{1cm} (7.2.7)

\[ DBP = 60.740 + 0.920 \text{BMI} - 0.371 \text{HDL} + 0.059 \text{TG} + 0.114 \text{LDL} + \epsilon_2 \]  \hspace{1cm} (7.2.8)

To compare the rate of change in the values of SBP and DBP with the change in the values of different independent variables (factors) such as BMI, HDL, TG and LDL, we first convert variables in the standard form by using the transformation:

\[ Y^* = \frac{Y - E(Y)}{\sigma_Y} \]

Here, \( Y^*, Y, E(Y) \) and \( \sigma_Y \) denote the standardized variable, original variable, average value of the variable \( Y \) and standard deviation of variable \( Y \) respectively, because range of variation of independent variables differ widely. Similarly, all the variables have been converted.

On the basis of these standardized variables, the models (7.2.7) and (7.2.8) reduces to the following standardized models:

\[ SBP^* = 0.309 \text{BMI}^* - 0.212 \text{HDL}^* + 0.413 \text{TG}^* + 0.221 \text{LDL}^* + \epsilon_1^* \]  \hspace{1cm} (7.2.9)

\[ DBP^* = 0.326 \text{BMI}^* - 0.191 \text{HDL}^* + 0.315 \text{TG}^* + 0.231 \text{LDL}^* + \epsilon_2^* \]  \hspace{1cm} (7.2.10)

Here, \( \epsilon_1^* \) & \( \epsilon_2^* \) are assumed to be normally distributed with mean zero and variance one.
To check the assumption whether $\varepsilon_1^*$ & $\varepsilon_2^*$ are normally distributed with mean zero and variance one, we apply following tests:

i) Free Hand Curve Method

ii) Q-Q Plot Method

iii) Box Plot Method

i) Free Hand Curve Method:

The histogram is drawn for the grouped frequency distribution of the standardized residuals for both linear model (7.2.9) and model (7.2.10). A free hand curve is drawn on these histograms in figure 7.2.1 and figure 7.2.2 shows that both the errors almost follow normal distribution.

Figure 7.2.1: Histogram with mean = 0 and Variance = 0.95 for model (7.2.9)
ii) Q-Q Plot Method:

The most useful tool for assessing normality is a Q-Q plot. This is a scatter plot with the quantiles of the scores on the horizontal axis and expected normal scores on the vertical axis. A plot of these scores against the expected normal scores should reveal a straight line. The expected normal scores are calculated by using the following formula:

$$\text{Expected } z\text{-score} = \left\{ z \mid P[Z \leq z] = \frac{I - 0.5}{N} \right\}$$

Here, Z denotes the standard normal variable and I is the rank of observation in the set of N observations.

The graph corresponding to both models between expected normal scores and observed values were drawn in figures 7.2.3 & figure 7.2.4 respectively. From figure 7.2.3 and figure 7.2.4, we see that all the points are nearly falling around straight line except two outliers which again signifies that both the errors follow normal distribution with mean zero and variance one.
iii) Box-Plot Method:

A box plot is a way of summarizing a set of data measured on an interval scale. It is a type of graph which is used to show the shape of the distribution, its central value, and variability. The box plot is constructed corresponding to the distribution based upon five points i.e. extreme values and quantiles. The skewness of the distribution is measured w.r.t. the position of box
between the extreme values and kurtosis is measured w.r.t. the size of box relative to whiskers. On the basis of the data, we have constructed box plot corresponding to values of errors in model (7.2.9) and model (7.2.10) in figure 7.2.5 and figure 7.2.6 respectively. The position of the boxes in both figures is almost lies in the centre of the extreme values which indicates that a distribution of residuals in both models is almost symmetric. On the other hand, the size of boxes in both figures is almost equal to corresponding distance between extreme values and boxes, indicating that the curve to be almost mesokurtic. Hence, we may conclude that distribution of residuals is almost normal.

Figure 7.2.5: Box Plot for Model (7.2.9)
7.3 Analysis of Fitted Models & Their Results:

On the basis of given data, we perform ANOVA to test the significance of the regression models (7.2.9) and (7.2.10) to determine if there is a linear relationship between the response variable and any of the regressor variables. This procedure is often thought of an overall or global test of model adequacy. The appropriate hypotheses for the regression coefficients of a multiple linear regression model are

\[ H_0 : \beta_1 = \beta_2 = \ldots = \beta_n = 0 \]

\[ H_1 : \beta_j \neq 0 \quad \text{for at least one j} \]

Rejection of this null hypothesis implies that at least one of the regressors contributes significantly to the model. In this testing procedure, the total sum of square (TSS) is partitioned into a sum of square due to regression (RSS) and error sum of square (ESS).

i.e.
\[ TSS = RSS + ESS \]

and under null hypothesis the test statistic-F given by

\[ F_0 = \frac{RSS/n}{ESS/(N-n-1)} \]

follows the \( F_{n,N-n-1} \) distribution.

In models (7.2.3) & (7.2.4) regression coefficients are \( \beta_j = a_j \) and \( \beta_j = b_j \) respectively.

On the basis of the data, we have obtained the results corresponding to the model (7.2.9) and model (7.2.10) which are given in the Table-7.3.1 and Table-7.3.2. Both the tables show that the p-value upto three decimal places is zero. Hence, we conclude that both the fitted models (7.2.9) and (7.2.10) are adequate models to study the SBP and DBP.

Table-7.3.1

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degree of Freedom</th>
<th>Mean Square</th>
<th>F-Ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>23850.998</td>
<td>4</td>
<td>5962.749</td>
<td>17.327</td>
<td>.000</td>
</tr>
<tr>
<td>Residual</td>
<td>25810.390</td>
<td>75</td>
<td>344.139</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>49661.388</td>
<td>79</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Predictors: (Constant), LDL-C, TG, BMI, HDL-C

Table-7.3.2

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degree of Freedom</th>
<th>Mean Square</th>
<th>F-Ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>6276.535</td>
<td>4</td>
<td>1569.134</td>
<td>12.298</td>
<td>.000</td>
</tr>
<tr>
<td>Residual</td>
<td>9569.415</td>
<td>75</td>
<td>127.592</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>15845.950</td>
<td>79</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Predictors: (Constant), LDL-C, TG, BMI, HDL-C
From both Table-7.3.1 and Table-7.3.2, we see that the p-value upto three decimal places is zero, indicating the rejection of null hypothesis in both the models.

In order to test, which model in the fitted models (7.2.9) and (7.2.10) is better and explain more variation in dependent variable, we obtain the multiple correlation coefficients (R) and adjusted coefficient of multiple determinations (adjusted R²) from the ANOVA defined in section 7.3 as:

\[ R^2 = \frac{RSS}{TSS} \quad \text{&} \quad \text{adjusted } R^2 = 1 - \frac{ESS/N - n}{TSS/N - 1} \]

In linear regression model, R² gives the percentage of explained variation in the dependent variable by the independent variables. In general, R² always increases when a regressor is added to the model, regardless of the value of the contribution of that variable. Therefore, it is difficult to judge whether an increase in R² is really telling us anything important. In order to remove this problem, we use adjusted R² statistic which will only increase on adding a variable to the model if the addition of the variable reduces the residual mean square. That is why, Adjusted R² statistic is very useful in comparing two or more than two regression models.

The values of R, R² and R²(adjusted) and standard error of estimate corresponding to both models (7.2.9) & (7.2.10) obtained by using the data given in appendix are given in Table-7.3.3.
From Table-7.3.3, we see that the values of the adjusted $R^2$ for models (7.2.9) and (7.2.10) are 45% and 36% respectively which signifies that former model explains more variation in SBP as compared to the later model in DBP. Therefore, model (7.2.9) can be considered as more adequate and reliable as compared to model (7.2.10).

To study the hypertension which generally depends on SBP and DBP, we find that BMI, TG, LDL and HDL are the factors affecting the SBP and DBP more significantly. It has been shown that the relationship of SBP and DBP with factors BMI, TG, LDL and HDL is linear. The effect of these factors on SBP and DBP has been ranked w.r.t. their magnitude i.e.

\[ TG > BMI > LDL > HDL \]

It is shown that TG is the most important risk factor of hypertension & heart diseases and models proposed are adequate for the prediction of SBP and DBP. To verify the goodness of fit of the proposed models, we draw graphs corresponding to the estimated values and the actual (real) values for SBP and DBP separately.

**Graph-7.3.1**
From both the graphs, we see that lines corresponding to estimated and actual values almost coincide each other which show that both fitted models are adequate for predicting the values of SBP and DBP.

Summary

Two separate multiple linear regression models have been used by taking SBP & DBP as dependent variables and body mass index, tri-glyceride, low density lipoprotein-cholesterol and high density lipoprotein-cholesterol as independent variables (independent risk factors) to study the essential hypertension among the euglycemic patients of ischemic heart disease. Adequacy of the models has been verified using analysis of variance (F-test), histogram, Q-Q plot and box whiskers plot. It has been shown that body mass index; tri-glyceride, low density lipoprotein-cholesterol and high density lipoprotein-cholesterol are the factors affecting the SBP and DBP significantly. It has also been shown that TG and BMI are more affecting the blood pressure as compared to other factors. The graphs have also been plotted corresponding to fitted models and original data which show that fitted models are best models to obtain the values of SBP and DBP. An effort has been made to analyze the data and interpret it with the help of statistical software package SPSS 14.0.