DEPARTMENT OF
STUDENTS' WELFARE
Discover. Learn. Empower.
CU
CHANDIGARH
UNIVERSITY

NAAC
GRADE A+
ACCREDITED UNIVERSITY

EXPERIMENT 5

CO3: To perform data collection and initial data handling by managing data structures and measurement levels.

Aim: To apply binary logistic regression on a medical dataset to predict the likelihood of a medical condition (e.g., presence or absence of a disease) based on patient characteristics (e.g., age, blood pressure, cholesterol).

Objective
To develop and interpret a binary logistic regression model using OriginPro 2024b that accurately predicts the presence or absence of a medical condition based on a set of patient-related predictor variables.

Step-by-Step Procedure

Step 1: Prepare the Data
1. **Create a New Workbook**:
   o Open OriginPro 2024b and create a new workbook.
2. **Input the Medical Data**:
   o Enter the following sample medical data into the workbook:

Sample Medical Data Table:

Patient	Age (Years)	Blood Pressure (mm Hg)	Cholesterol (mg/dL)	Condition (1 = Present, 0 = Absent)
1	45	140	200	1
2	50	160	220	1
3	35	120	180	0
4	40	130	190	0
5	55	170	240	1
6	60	180	250	1
7	30	110	170	0
8	50	150	210	1
9	40	140	200	0
10	35	130	185	0

Step 2: Open the Nonlinear Curve Fit Tool
1. **Go to the Analysis Menu**:
   o Navigate to Analysis > Fitting > Nonlinear Curve Fit > Open Dialog....
2. **Select a Logistic Model**:
   o In the "Function Selection" tab, you'll need to choose a logistic model. Since binary logistic regression isn't pre-built, you'll have to define it.
   o Use the equation for the logistic function:

$$P(Y = 1|X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n)}}$$

   o You may need to use a custom fitting function or find a logistic function that matches this form.
3. **Set the Dependent and Independent Variables**:
   o In the "Input Data" tab, select your binary dependent variable as the "Y" and your predictor variables as the "X".

Step 3: Perform the Fit
1. **Run the Nonlinear Fit**:
   - o After setting up the model and input data, click "Fit" to run the analysis.
   - o The software will estimate the coefficients $\beta_0, \beta_1, \beta_2, \ldots$ for the logistic regression model.
2. **Review the Results**:
   - o Check the output for the coefficients, standard errors, and other statistics related to the fit. You can interpret these as you would in any logistic regression analysis.

Step 4: Interpretation and Validation
1. **Examine the Fit Statistics**:
   - o Look at the goodness-of-fit measures to evaluate how well the model describes the data.
2. **Interpret the Coefficients**:
   - o Just like in any logistic regression analysis, interpret the coefficients in terms of odds ratios and their statistical significance.
3. **Residual Analysis (Optional)**:
   - o Perform residual analysis to check for outliers or other issues that might affect the model's performance.

OR

Using Generalized Linear Models (GLM)

If OriginPro has a "Generalized Linear Models" (GLM) tool:
1. **Open the GLM Tool**:
   - o Go to Statistics > General Linear Models > Generalized Linear Models.
2. **Specify Logistic Regression**:
   - o In the GLM dialog, specify that you want to use a logistic link function (logit) with a binary distribution for the dependent variable.
3. **Input Your Data**:
   - o Choose the dependent variable and the independent variables.
4. **Run the Analysis**:
   - o Click "OK" or "Compute" to run the logistic regression.