

**SAVEETHA SCHOOL OF ENGINEERING**  
**SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES**  
**ITA 0443 - STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEM**  
**DAY 4– LAB MANUAL**

**Reg No:192124086**

**Name:L.Uthra**

**LINEAR REGRESSION ANALYSIS IN R**

**Exercise**

1. Using linear regression analysis establish a relationship between height and weight of a person using the input vector given below.

# Values of height

151, 174, 138, 186, 128, 136, 179, 163, 152, 131

# Values of weight.

63, 81, 56, 91, 47, 57, 76, 72, 62, 48

Predict the weight of a person with height 170. Visualize the regression graphically.

**Program:**

```
height <- c(151, 174, 138, 186, 128, 136, 179, 163, 152, 131)
```

```
weight <- c(63, 81, 56, 91, 47, 57, 76, 72, 62, 48)
```

```
abline(lm(weight~height))
```

```
line<-lm(weight~height)
```

```
A<-data.frame(height=170)
```

```
pre<-predict(line,A)
```

```
print(pre)
```

**Output:**

```
> print(pre)
```

```
1
```

```
76.22869
```

2. Download the Dataset "water" From Rdataset Link. Find out whether there is a linear relation between attributes "mortality" and "hardness" by plot function. Fit the Data into the Linear Regression model. Predict the mortality for the hardness=88

**Program:**

```
dataset<-read.csv("water.csv")  
  
relation<-lm(mortality~hardness,data = dataset)  
  
A<-data.frame(hardness=88)  
  
pre<-predict(relation,A)  
  
print(pre)
```

**Output:**

```
> print(pre)
```

```
1
```

```
1392.46
```

## MULTIPLE REGRESSION ANALYSIS IN R

**Exercise:**

1. Generate a multiple regression model using the built in dataset mtcars. It gives a comparison between different car models in terms of mileage per gallon (mpg), cylinder displacement("disp"), horse power("hp"), weight of the car("wt") and some more parameters.

Establish the relationship between "mpg" as a response variable with "disp", "hp" and "wt" as predictor variables. Predict the mileage of the car with dsp=221, hp=102 and wt=2.91.

**Program:**

```
data(mtcars)  
  
model<-lm(mpg~disp+hp+wt,data=mtcars)  
  
pred<-data.frame(disp=221, hp=102, wt=2.91)  
  
predict(model, pred)
```

**Output:**

```
> print(pre)
      1
1392.46
```

2. Consider the data set "delivery" available in the R environment. It gives a deliverytime ("delTime") of production materials (number of productions "n.prod") with the given distance ("distance") to reach the destination place.

a) Create the model to establish the relationship between "delTime" as a response variable with "n.prod" and "distance" as predictor variables.

b) Predict the delTime for the given number of production ("n.prod")=9 and distance ("distance")=450

## LOGISTIC REGRESSION ANALYSIS IN R

**Exercise**

1. Create a logistic regression model using the "mtcars" data set with the information given below.

The in-built data set "mtcars" describes different models of a car with their various engine specifications. In "mtcars" data set, the transmission mode (automatic or manual) is described by the column am which is a binary value (0 or 1). Create a logistic regression model between the columns "am" and 3 other columns - hp, wt and cyl.

**Program:**

```
model <- glm(am ~ hp + wt + cyl, data = mtcars, family = binomial)

summary(model)
```

**Output:**

```
glm(formula = am ~ hp + wt + cyl, family = binomial, data = mtcars)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.17272	-0.14907	-0.01464	0.14116	1.27641

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	19.70288	8.11637	2.428	0.0152 *
hp	0.03259	0.01886	1.728	0.0840 .
wt	-9.14947	4.15332	-2.203	0.0276 *
cyl	0.48760	1.07162	0.455	0.6491

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 43.2297 on 31 degrees of freedom  
Residual deviance: 9.8415 on 28 degrees of freedom  
AIC: 17.841

Number of Fisher Scoring iterations: 8

## POISSON REGRESSION ANALYSIS IN R

### Exercise :

1. Create a Poisson regression model using the in-built data set "warpbreaks" with information given below.

In-built data set "warpbreaks" describes the effect of wool type (A or B) and tension (low, medium or high) on the number of warp breaks per loom. Consider "breaks" as the response variable which is a count of number of breaks. The wool "type" and "tension" are taken as predictor variables.

### Program:

```
model <- glm(breaks ~ wool + tension, data = warpbreaks, family = poisson)

summary(model)
```

### Output:

```
glm(formula = breaks ~ wool + tension, family = poisson, data = warpbreaks)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-----	----	--------	----	-----

-3.6871 -1.6503 -0.4269 1.1902 4.2616

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	3.69196	0.04541	81.302	< 2e-16 ***
woolB	-0.20599	0.05157	-3.994	6.49e-05 ***
tensionM	-0.32132	0.06027	-5.332	9.73e-08 ***
tensionH	-0.51849	0.06396	-8.107	5.21e-16 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 297.37 on 53 degrees of freedom

Residual deviance: 210.39 on 50 degrees of freedom

AIC: 493.06

Number of Fisher Scoring iterations: 4