STAT 43000/STAT 53001 Applied Statistics Spring 2023

Lab 6

Q.N. 1) *mtcars* data in R 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).

a) Create a simple logistic regression model using weight as a predictor variable.

```
> model=glm(am~wt,family=binomial)
> model

Call:  glm(formula = am ~ wt, family = binomial)

Coefficients:
(Intercept)           wt
     12.040       -4.024

Degrees of Freedom: 31 Total (i.e. Null);  30 Residual
Null Deviance:      43.23
Residual Deviance: 19.18        AIC: 23.18
>
```

Π=[1+exp(-12.040+4.024wt)]

b) Perform the test of significance of the weight of the car.

```
> summary(model)

Call:
glm(formula = am ~ wt, family = binomial)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-2.11400  -0.53738  -0.08811   0.26055   2.19931

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   12.040      4.510   2.670  0.00759 **
wt            -4.024      1.436  -2.801  0.00509 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 43.230  on 31  degrees of freedom
Residual deviance: 19.176  on 30  degrees of freedom
AIC: 23.176

Number of Fisher Scoring iterations: 6
```

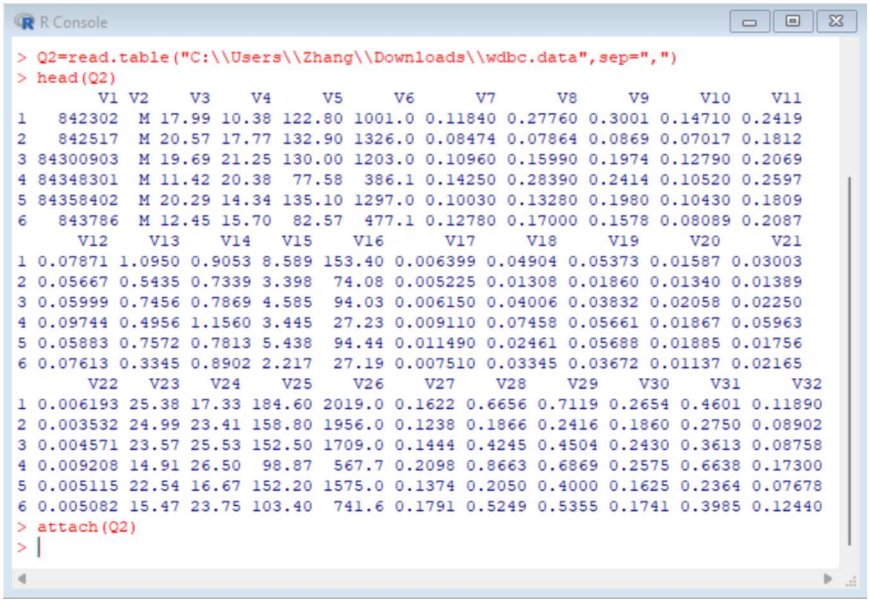
Pr(>|z|) value 0.00759 is too small. We reject H0(b1 is not important) and accept H1(b1 exists)

c) What is the probability that a 2000 lbs car is manual?

```
> predict(model,data.frame(wt=2),type="resp")
Error: unexpected input in "predict(model,data.frame(wt=2),type=""
> predict(model,data.frame(wt=2),type="response")
Error: unexpected input in "predict(model,data.frame(wt=2),type=""
> predict(model,data.frame(wt=2),type="resp")
        1
0.9818796
> predict(model,data.frame(wt=2),type="response")
        1
0.9818796
>
```

Q.N. 2) Breast Cancer Wisconsin (Diagnostic) Data Set contains 569 observations with 32 variables including 1) ID number 2) Diagnosis (M = malignant, B = benign) 3) Radius, …….

a) Import the data in R.



b) Fit a simple logistic regression model using Radius as a predictor variable.

```
> m2=glm(V2~V3,family="binomial")
> m2

Call:  glm(formula = V2 ~ V3, family = "binomial")

Coefficients:
(Intercept)           V3
    -15.246        1.034

Degrees of Freedom: 568 Total (i.e. Null);  567 Residual
Null Deviance:      751.4
Residual Deviance: 330  AIC: 334
>
```

$\Pi=[1+\exp(-15.246+1.034V3)]^{-1}$