HW7

Zeqiu.Yu

2022-11-22

Homework 7

14.7

```r
Data <- read.table("./BookDataSets/Chapter 14 Data Sets/CH14PR07.txt",header=FALSE)
colnames(Data) <- c("Y", "X")
```

(a.)

```r
glm.out = glm(Y~X, family=binomial(logit), data=Data)
summary(glm.out)
```

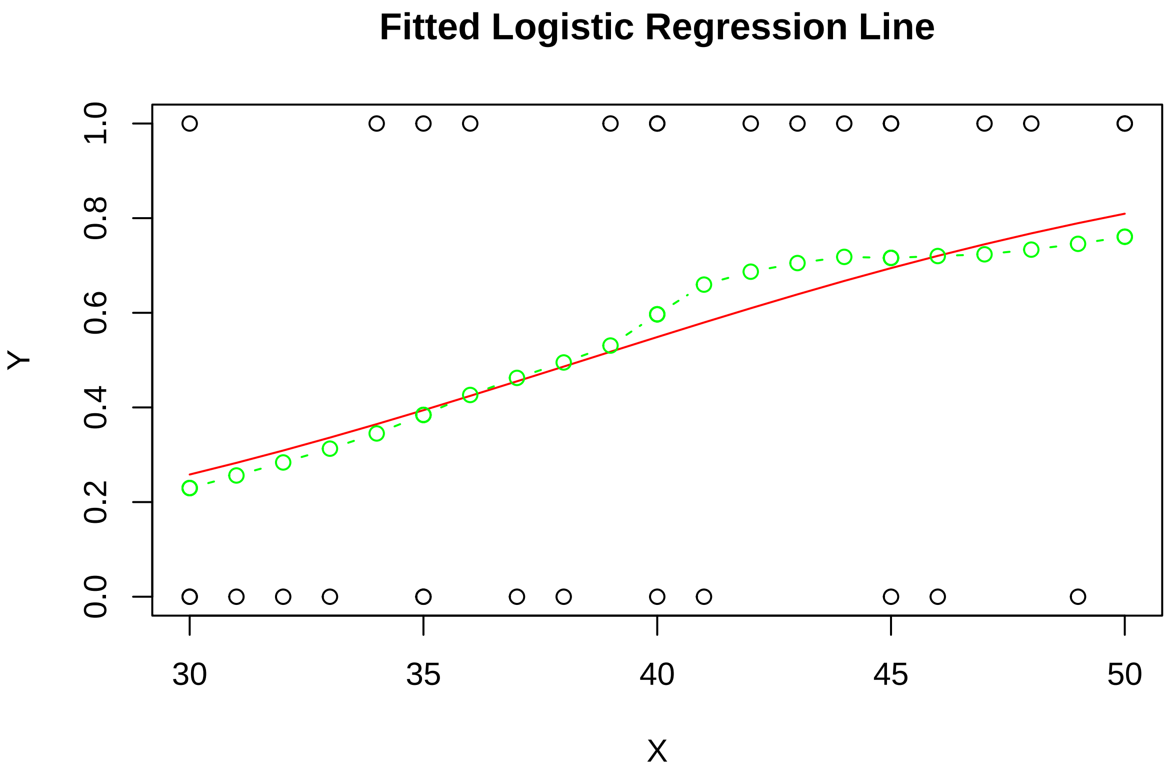
```
##
## Call:
## glm(formula = Y ~ X, family = binomial(logit), data = Data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.7651  -1.0012   0.6502   0.9828   1.6455
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.80751    2.65576  -1.810   0.0703 .
## X            0.12508    0.06676   1.874   0.0610 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 41.455  on 29  degrees of freedom
## Residual deviance: 37.465  on 28  degrees of freedom
## AIC: 41.465
##
## Number of Fisher Scoring iterations: 4
```

The maximum likelihood estimates: $\beta_0 = -4.80751$ and $\beta_1 = 0.12508$, and $\hat{\pi} = \frac{exp(-4.80751+0.12508X)}{1+exp(-4.80751+0.12508X)}$.

(b.)

```r
plot(Y~X, data=Data)
lines(Data$X[order(Data$X)], glm.out$fitted[order(Data$X)],
      type="l", col="red")
title(main="Fitted Logistic Regression Line")
```

```
Data.smooth <- predict(loess(Y~X, data=Data, span=0.75))
points(Data$X[order(Data$X)], Data.smooth[order(Data$X)],
      type="b",lty=2, col="green")
legend(5,0.9, c("logistic","loess smooth"), col=c("red", "green"), lty=c(1:2))
```



The fitted logistic response function fitts appears to fit well.

c.

```
exp_b1 <- exp(glm.out$coefficients[2])
exp_b1
```

```
##        X
## 1.133237
```

exp(b1) is 1.1332371. For an one-unit increase in the dollar increase in annual dues, we expect to see about 13.33% increase in the odds of that the membership will not be renewed.

d.

```
new <- data.frame(X=40, Y=1)
#y.hat <- predict(glm.out, new)      #predict log(p/(1-p))
#p.hat <- exp(y.hat)/(1+exp(y.hat))  #Transfer to probability
# or use option of "response" directly
p.hat <-predict(glm.out, new, type="response")
p.hat
```

```
##         1
## 0.5487487
```

The estimated probability that association members will not renew their membership if the dues are increased by $40 is 0.5487487.

e.

```
(log(0.75/0.25)+4.80751)/0.12508
```

```
## [1] 47.21876
```

Since $\hat{\pi} = 0.75$, $log(\frac{\hat{\pi}}{1-\hat{\pi}}) = log(\frac{0.75}{0.25}) = -4.80751 + 0.12508X$. Hence, X = 47.2187583.

14.15

(a.)

```
conf.beta1 <- confint(glm.out, "X", level = 0.90)
```

```
## Waiting for profiling to be done...
```

```
exp(conf.beta1)
```

```
##      5 %     95 %
## 1.021468 1.276781
```

An approximate 90 percent confidence interval for $exp(\beta_1)$ is [1.021468, 1.276781].
With 90 percent confidennce, we believe that, for an one-unit increase in the dollar increase in annual dues, the odds of that the membership will not be renewed will increase by between 2.1468% and 27.6781%.

b)

```
summary(glm.out)
```

```
##
## Call:
## glm(formula = Y ~ X, family = binomial(logit), data = Data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.7651  -1.0012   0.6502   0.9828   1.6455
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.80751    2.65576  -1.810   0.0703 .
## X            0.12508    0.06676   1.874   0.0610 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 41.455  on 29  degrees of freedom
## Residual deviance: 37.465  on 28  degrees of freedom
## AIC: 41.465
##
## Number of Fisher Scoring iterations: 4
```

```
qnorm(0.95, lower.tail=FALSE)
```

```
## [1] -1.644854
```

H_0: $\beta_1 = 0$
H_a: $\beta_1 \neq 0$.

Decision rule: $z^* = \frac{b_1}{s\{b_1\}} = \frac{0.12508}{0.06676} = 1.873577$, $z(1 - 0.1/2) = z(0.95) = 1.645$, if $|z*| \geq z(0.95)$, conclude Ha.
Conclusion: $|z*| = 1.873577 \geq z(0.95) = 1.645$. Hence, conclude H_a.