mtcars.R

hp

2021-11-12

# read in the mtcars data set  
data(mtcars)  
mtcars

## mpg cyl disp hp drat wt qsec vs am gear carb  
## Mazda RX4 21.0 6 160.0 110 3.90 2.620 16.46 0 1 4 4  
## Mazda RX4 Wag 21.0 6 160.0 110 3.90 2.875 17.02 0 1 4 4  
## Datsun 710 22.8 4 108.0 93 3.85 2.320 18.61 1 1 4 1  
## Hornet 4 Drive 21.4 6 258.0 110 3.08 3.215 19.44 1 0 3 1  
## Hornet Sportabout 18.7 8 360.0 175 3.15 3.440 17.02 0 0 3 2  
## Valiant 18.1 6 225.0 105 2.76 3.460 20.22 1 0 3 1  
## Duster 360 14.3 8 360.0 245 3.21 3.570 15.84 0 0 3 4  
## Merc 240D 24.4 4 146.7 62 3.69 3.190 20.00 1 0 4 2  
## Merc 230 22.8 4 140.8 95 3.92 3.150 22.90 1 0 4 2  
## Merc 280 19.2 6 167.6 123 3.92 3.440 18.30 1 0 4 4  
## Merc 280C 17.8 6 167.6 123 3.92 3.440 18.90 1 0 4 4  
## Merc 450SE 16.4 8 275.8 180 3.07 4.070 17.40 0 0 3 3  
## Merc 450SL 17.3 8 275.8 180 3.07 3.730 17.60 0 0 3 3  
## Merc 450SLC 15.2 8 275.8 180 3.07 3.780 18.00 0 0 3 3  
## Cadillac Fleetwood 10.4 8 472.0 205 2.93 5.250 17.98 0 0 3 4  
## Lincoln Continental 10.4 8 460.0 215 3.00 5.424 17.82 0 0 3 4  
## Chrysler Imperial 14.7 8 440.0 230 3.23 5.345 17.42 0 0 3 4  
## Fiat 128 32.4 4 78.7 66 4.08 2.200 19.47 1 1 4 1  
## Honda Civic 30.4 4 75.7 52 4.93 1.615 18.52 1 1 4 2  
## Toyota Corolla 33.9 4 71.1 65 4.22 1.835 19.90 1 1 4 1  
## Toyota Corona 21.5 4 120.1 97 3.70 2.465 20.01 1 0 3 1  
## Dodge Challenger 15.5 8 318.0 150 2.76 3.520 16.87 0 0 3 2  
## AMC Javelin 15.2 8 304.0 150 3.15 3.435 17.30 0 0 3 2  
## Camaro Z28 13.3 8 350.0 245 3.73 3.840 15.41 0 0 3 4  
## Pontiac Firebird 19.2 8 400.0 175 3.08 3.845 17.05 0 0 3 2  
## Fiat X1-9 27.3 4 79.0 66 4.08 1.935 18.90 1 1 4 1  
## Porsche 914-2 26.0 4 120.3 91 4.43 2.140 16.70 0 1 5 2  
## Lotus Europa 30.4 4 95.1 113 3.77 1.513 16.90 1 1 5 2  
## Ford Pantera L 15.8 8 351.0 264 4.22 3.170 14.50 0 1 5 4  
## Ferrari Dino 19.7 6 145.0 175 3.62 2.770 15.50 0 1 5 6  
## Maserati Bora 15.0 8 301.0 335 3.54 3.570 14.60 0 1 5 8  
## Volvo 142E 21.4 4 121.0 109 4.11 2.780 18.60 1 1 4 2

am.glm <- glm(formula = am~hp+wt, data = mtcars, family = binomial)  
am.glm

##   
## Call: glm(formula = am ~ hp + wt, family = binomial, data = mtcars)  
##   
## Coefficients:  
## (Intercept) hp wt   
## 18.86630 0.03626 -8.08348   
##   
## Degrees of Freedom: 31 Total (i.e. Null); 29 Residual  
## Null Deviance: 43.23   
## Residual Deviance: 10.06 AIC: 16.06

# Hypotheses test for individual covariates  
summary(am.glm)

##   
## Call:  
## glm(formula = am ~ hp + wt, family = binomial, data = mtcars)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.2537 -0.1568 -0.0168 0.1543 1.3449   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 18.86630 7.44356 2.535 0.01126 \*   
## hp 0.03626 0.01773 2.044 0.04091 \*   
## wt -8.08348 3.06868 -2.634 0.00843 \*\*  
## ---  
## 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: 10.059 on 29 degrees of freedom  
## AIC: 16.059  
##   
## Number of Fisher Scoring iterations: 8

1-pchisq(43.230-10.059,31-29)

## [1] 6.266267e-08

# do anova  
am.glm0 <- glm(formula = am~1, data = mtcars, family = binomial)  
am.glm0

##   
## Call: glm(formula = am ~ 1, family = binomial, data = mtcars)  
##   
## Coefficients:  
## (Intercept)   
## -0.3795   
##   
## Degrees of Freedom: 31 Total (i.e. Null); 31 Residual  
## Null Deviance: 43.23   
## Residual Deviance: 43.23 AIC: 45.23

anova(am.glm0, am.glm, test = 'Chisq')

## Analysis of Deviance Table  
##   
## Model 1: am ~ 1  
## Model 2: am ~ hp + wt  
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)   
## 1 31 43.230   
## 2 29 10.059 2 33.171 6.267e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Prediction  
newdata <- data.frame(hp=120, wt=2.8)  
newdata

## hp wt  
## 1 120 2.8

predict(am.glm, newdata, type = 'response')

## 1   
## 0.6418125

# Probit Link  
myprobit <- glm(am~hp+wt, data = mtcars, family = binomial(link = "probit"))

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

myprobit

##   
## Call: glm(formula = am ~ hp + wt, family = binomial(link = "probit"),   
## data = mtcars)  
##   
## Coefficients:  
## (Intercept) hp wt   
## 10.40552 0.02126 -4.54219   
##   
## Degrees of Freedom: 31 Total (i.e. Null); 29 Residual  
## Null Deviance: 43.23   
## Residual Deviance: 9.861 AIC: 15.86

# Logit Link (Logistic Regression model)  
# fit log-linear model  
glm.Logit <- glm(formula = am~hp+wt, data = mtcars, family = poisson())  
glm.Logit

##   
## Call: glm(formula = am ~ hp + wt, family = poisson(), data = mtcars)  
##   
## Coefficients:  
## (Intercept) hp wt   
## 2.287139 0.008831 -1.580433   
##   
## Degrees of Freedom: 31 Total (i.e. Null); 29 Residual  
## Null Deviance: 23.42   
## Residual Deviance: 10.48 AIC: 42.48

# Prediction of am of hp=120 and wt=2.8  
newdata1 <- data.frame(hp=120, wt=2.8)  
newdata1

## hp wt  
## 1 120 2.8

predict(glm.Logit, newdata1, type = "response")

## 1   
## 0.3401468

# Using likelihood ratio to check hypotheses, test statistics and p-value  
summary(glm.Logit)

##   
## Call:  
## glm(formula = am ~ hp + wt, family = poisson(), data = mtcars)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.0494 -0.5398 -0.4261 0.3003 1.0600   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 2.287139 0.907534 2.520 0.01173 \*   
## hp 0.008831 0.005259 1.679 0.09309 .   
## wt -1.580433 0.494009 -3.199 0.00138 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for poisson family taken to be 1)  
##   
## Null deviance: 23.420 on 31 degrees of freedom  
## Residual deviance: 10.481 on 29 degrees of freedom  
## AIC: 42.481  
##   
## Number of Fisher Scoring iterations: 5

# Estimation: test model fitting  
glm.LRatio <- glm(am~1, data = mtcars, family = poisson())  
glm.LRatio

##   
## Call: glm(formula = am ~ 1, family = poisson(), data = mtcars)  
##   
## Coefficients:  
## (Intercept)   
## -0.9008   
##   
## Degrees of Freedom: 31 Total (i.e. Null); 31 Residual  
## Null Deviance: 23.42   
## Residual Deviance: 23.42 AIC: 51.42

anova(glm.LRatio, glm.Logit, test = 'Chisq')

## Analysis of Deviance Table  
##   
## Model 1: am ~ 1  
## Model 2: am ~ hp + wt  
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)   
## 1 31 23.421   
## 2 29 10.482 2 12.939 0.00155 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# plotting the probability for males and females  
head(subset(mtcars, select = 'am'))

## am  
## Mazda RX4 1  
## Mazda RX4 Wag 1  
## Datsun 710 1  
## Hornet 4 Drive 0  
## Hornet Sportabout 0  
## Valiant 0

factor(mtcars$am)

## [1] 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 1 1 1 1 1 1 1  
## Levels: 0 1

w = table(mtcars$am)  
w

##   
## 0 1   
## 19 13

t = as.data.frame(w)  
t

## Var1 Freq  
## 1 0 19  
## 2 1 13

# Red for male and yellow for female  
h<- hist(t$Freq, breaks=5, col=c("red","yellow"), xlab="am",  
 main="Histogram with Orvelay Line")  
xfit<-seq(min(t$Freq),max(t$Freq),length=40)  
yfit<-dnorm(xfit, mean = mean(t$Freq), sd=sd(t$Freq))  
yfit<-yfit\*diff(h$mids[1:2])\*length(t$Freq)  
lines(xfit, yfit, col="blue", lwd=2)

