EPIDEMIOLOGY TAKEAWAYCAT

KILONZO WAVITI

2023-04-02

#Reading of the mtcars  
# Load the mtcars dataset  
data(mtcars)  
  
#Number of observations and variables  
# Load the mtcars dataset  
data(mtcars)  
  
# Check the dimensions of the dataset  
dim(mtcars)

## [1] 32 11

#To create the univariate logistic regression with mpg,disp and cyl  
  
# Load the mtcars dataset  
data(mtcars)  
  
# Convert am to a binary variable  
mtcars$am <- as.factor(mtcars$am - 1)  
  
#Creation of the univariate logistcs  
# Load the mtcars dataset  
data(mtcars)  
  
# Convert am to a binary variable  
# Univariate logistic regression model with mpg as independent variable  
mpg\_model <- glm(am ~ mpg, data = mtcars, family = binomial())  
  
# Univariate logistic regression model with cyl as independent variable  
cyl\_model <- glm(am ~ cyl, data = mtcars, family = binomial())  
  
# Univariate logistic regression model with disp as independent variable  
disp\_model <- glm(am ~ disp, data = mtcars, family = binomial())  
  
#Examining the significance of the data  
# Summary of the mpg model  
summary(mpg\_model)

##   
## Call:  
## glm(formula = am ~ mpg, family = binomial(), data = mtcars)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.5701 -0.7531 -0.4245 0.5866 2.0617   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -6.6035 2.3514 -2.808 0.00498 \*\*  
## mpg 0.3070 0.1148 2.673 0.00751 \*\*  
## ---  
## 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: 29.675 on 30 degrees of freedom  
## AIC: 33.675  
##   
## Number of Fisher Scoring iterations: 5

# Summary of the cyl model  
summary(cyl\_model)

##   
## Call:  
## glm(formula = am ~ cyl, family = binomial(), data = mtcars)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.6265 -0.5656 -0.5656 0.7871 1.9554   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 3.7777 1.5456 2.444 0.01452 \*   
## cyl -0.6912 0.2536 -2.725 0.00642 \*\*  
## ---  
## 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: 33.951 on 30 degrees of freedom  
## AIC: 37.951  
##   
## Number of Fisher Scoring iterations: 4

# Summary of the disp model  
summary(disp\_model)

##   
## Call:  
## glm(formula = am ~ disp, family = binomial(), data = mtcars)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.5651 -0.6648 -0.2460 0.7276 2.2691   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 2.630849 1.050170 2.505 0.01224 \*   
## disp -0.014604 0.005168 -2.826 0.00471 \*\*  
## ---  
## 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: 29.732 on 30 degrees of freedom  
## AIC: 33.732  
##   
## Number of Fisher Scoring iterations: 5

#Creation of the multivariate logistic regression  
# Multivariable logistic regression model  
multivar\_model <- glm(am ~ mpg + cyl + disp, data = mtcars, family = binomial())  
  
# Summary of the multivariable model  
summary(multivar\_model)

##   
## Call:  
## glm(formula = am ~ mpg + cyl + disp, family = binomial(), data = mtcars)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.2497 -0.7439 -0.1683 0.4698 2.3450   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -11.42678 8.77554 -1.302 0.1929   
## mpg 0.38166 0.27454 1.390 0.1645   
## cyl 1.36320 0.87338 1.561 0.1186   
## disp -0.02334 0.01354 -1.724 0.0848 .  
## ---  
## 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: 25.709 on 28 degrees of freedom  
## AIC: 33.709  
##   
## Number of Fisher Scoring iterations: 6

#Creation of a table  
# Create a table with the results  
results\_table <- data.frame(  
 Variable = c("mpg", "cyl", "disp", "Intercept"),  
 Coefficient = c(round(coefficients(multivar\_model), 4)),  
 OR = c(round(exp(coefficients(multivar\_model)), 4)),  
 CI\_95 = c(round(confint(multivar\_model), 4)),  
 p\_value = c(round(summary(multivar\_model)$coefficients[,4], 4))  
)

## Waiting for profiling to be done...

## Warning in data.frame(Variable = c("mpg", "cyl", "disp", "Intercept"),  
## Coefficient = c(round(coefficients(multivar\_model), : row names were found from  
## a short variable and have been discarded

# Display the table  
results\_table

## Variable Coefficient OR CI\_95 p\_value  
## 1 mpg -11.4268 0.0000 -32.1968 0.1929  
## 2 cyl 0.3817 1.4647 -0.0485 0.1645  
## 3 disp 1.3632 3.9087 -0.1922 0.1186  
## 4 Intercept -0.0233 0.9769 -0.0550 0.0848  
## 5 mpg -11.4268 0.0000 3.0494 0.1929  
## 6 cyl 0.3817 1.4647 1.0492 0.1645  
## 7 disp 1.3632 3.9087 3.3472 0.1186  
## 8 Intercept -0.0233 0.9769 0.0000 0.0848