STAT 796 HW 3

Colleen Brents

February 9, 2019

# 1. The following questions ask about descriptive summaries of the cohort.

1. What proportion of the cohort had myopia within 5 years of study entry? The answer is 0.131

|  |  |  |  |
| --- | --- | --- | --- |
| myopic | Mom NOT myopic | Mom myopic |  |
| 0 (N) | 279 | 258 | 537 |
| 1 (Y) | 26 | 55 | 81 |
|  | 305 | 313 | 618 |

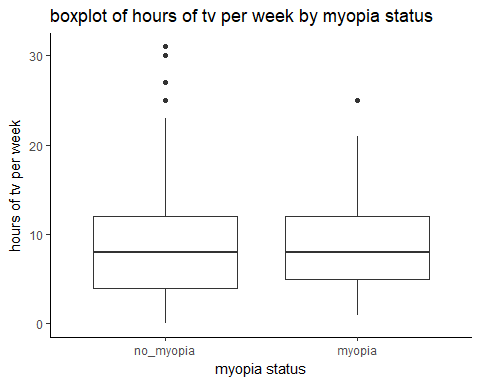
1. Report the proportion of children whose mom had myopia among
2. all children, The answer is 0.506
3. children without myopia The answer is 0.48
4. children with myopia. The answer is 0.679

|  |  |  |  |
| --- | --- | --- | --- |
| myopic | Dad NOT myopic | Dad myopic |  |
| 0 (N) | 285 | 252 | 537 |
| 1 (Y) | 25 | 56 | 81 |
|  | 310 | 308 | 618 |

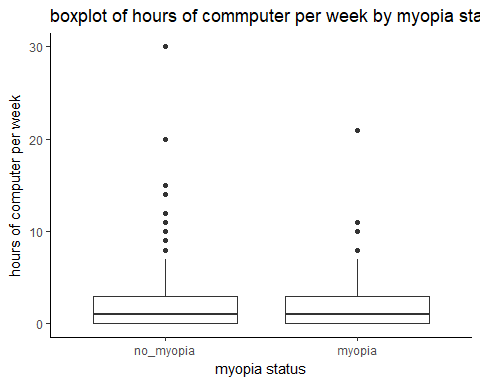
1. Report the proportion of children whose dad had myopia among
2. all children, The answer is 0.498
3. children without myopia The answer is 0.469
4. children with myopia. The answer is 0.691

|  |  |  |  |
| --- | --- | --- | --- |
| myopic | male | female |  |
| 0 (N) | 281 | 256 | 537 |
| 1 (Y) | 35 | 46 | 81 |
|  | 316 | 302 | 618 |

1. Report the proportion of children who were male among
2. all children, The answer is 0.511
3. children without myopia The answer is 0.523
4. children with myopia. The answer is 0.432
5. Make a boxplot of the hours of TV watched by children, stratified by myopia status.



1. Make a boxplot of the hours of computer time for the children, stratified by myopia status



## 2. For this question, fit a logistic regression model using myopia status as the outcome. Include age, indicator of sex, indicator of mom with myopia, indicator of dad with myopia, hours spent watching TV, and hours spent on the computer as the predictor variables.

1. For each of the three indicator predictor variables, report the coefficient estimate and interpret its meaning in a single sentence.

The estimated regression coefficient for the indicator variable of sex is 0.492. This is the average difference in log odds for myopia in females compared to males of the same age, same hours of tv and computer time, and same parental myopia status (mom or dad).

The estimated regression coefficient for the indicator variable of mom with myopia is 0.873. This is the average difference in log odds of myopia in children with moms with myopia compard to those with myopia-free of the same sex, age, dad myopia status, and hours of screen time (computer and tv)

The estiamted regression coefficient for the indicator predictor variable of dad with myopia is 0.971. This is the average difference in log odds of myopia in cchildren with dads with myopia compared to those with myopia-free dads of the same age, sex, mom myopic status, and hours and screen time (computer and tv)

1. For each of the three indicator predictor variables, report an estimated odds ratio and interpret its meaning in a single sentence.

The odds ratio for myopia comparing females and males, among those with same age, parental myopic status, and screen time (comuter and tv) is 1.635

The odds ratio for myopia comparing those with moms with and without myopia, among those with the same age, paternal myopic status, same sex, and same screen time (computer and tv) is 2.394

The odds ratio for myopia comparing those with dad with and without myopia, among those of the same age, sex, materal myopix status, and same screen time (computer and tv) is 2.641

1. How do the odds ratios you calculated in (b) compare to the data summaries in (b)-(d) of Question 1?

Data summaries in 1bcd show an increased proportion of children having myopia if they have a parent with myopia (vs mother or dather without myopia) or if the child is a female (vs male). The odds ratios in 2b agree with this statement that there is an associatin between myopia and sex and positive parental myopia characteristics, as the odds ratio in 1bcd exceed 1.  
Both 1bcd and 2b show an association between myopia in children and that child's sex (female vs male) and their parents' (mom or dad) myopic stuatus - particulatly if the child is female and/or has a mom and/or dad with myopia.

1. For each of the three other (non-indicator) predictor variables, report an estimated odds ratio and interpret its meaning in a single sentence.

The odds ratio for myopia for a one year (one unit) difference in age among children with the same sex, parental myopic status, and screen time (computer and tv) is 1.197

The odds ratio for myopia for a one hour difference in weekly tv-watching among children with the same age, sex, parental myopic status, and same computer time is 0.995

The odds ratio for myopia for a one hour difference in weekly computer-usage among children with the same age, sex, parental myopic status, and tv-watching time is 1.037

1. What is the estimated adjusted odds ratio for myopia corresponding to a difference of 5 hours in amount of TV watched each week? The answer is 0.978
2. What is the estimated adjusted odds ratio for myopia corresponding to a difference of 25 hours in amount of computer time each week? The answer is 2.483
3. What is the estimated adjusted odds ratio for myopia corresponding to a difference of 5 years in age at first visit? The answer is 2.461
4. Do the quantities calculated in (e)-(g) make sense in the context of this cohort? [This is a subjective question, but consider the range of these predictor variables in the cohort

A five hour difference in tv time, given the range was 0 to 31 with a mean 8.9, is reasonable within the context of this cohort. There are children that have a difference of tv t ime. The odds ratio (2.483) suggests there is essentaily no association between tv screen time and myopia. The odds ratio for 1 hour difference in tv time (0.995) also supports that lack of association. Perhaps the distance from the tv reduces the amount of eye strain? A 25 hour difference in computer time, given teh range was 0 to 30, with a mean 2.1, is not very reasonable in the context of this cohort as it is very extreme. The odds ratio 2.483 suggests an association between computer time and myopia (adjusting for other variables)  
A 5 years difference in age, given that the subject population spanned 5 to 9 years old, is less reasonable given that the range of participant population is 4 years. The odds ratio 2.461 suggests an increase in myopia risk as age increases (adjusting for other varaibles). My general impressions of myopia are that age and screen time are associated with an increased risk of developing myopia. Whil ethe quantities suggested in 2h are sometimes excessive, the results agree with my original impressions (wit hthe excecption of tv screen time). I think the PI should redo the experiement with smartphones (as that can replace both a tv and computer, especially for little kids).

## Appendix

knitr::opts\_chunk$set(echo = FALSE, message = FALSE, warning = FALSE)  
options(scipen=1, digits=3)  
library(ggplot2)  
library(dplyr)  
library(tidyr)  
library(readr)  
library(knitr)  
url <- paste0("https://raw.githubusercontent.com/cbrents/stat796logreg/master/hw3/myopia.csv")  
myopia<-read\_csv(url)  
#myopic within first 5 years of follow up 0 = no 1 = yes   
myopia %>%   
 count(myopic)  
onea<- 81/(537+81)  
myopia %>%   
 count(myopic, mom\_myopic)  
bi<- 313/618  
bii <- 258/537  
biii <- 55/81  
myopia %>%   
 count(myopic, dad\_myopic)  
ci<- 308/618  
cii <- 252/537  
ciii <- 56/81  
myopia %>%   
 count(myopic, female)  
di<- 316/618  
dii <- 281/537  
diii <- 35/81  
myopia <- myopia %>%   
 mutate(myopic = factor(myopic, levels = 0:1,   
 labels = c("no\_myopia", "myopia")))  
ggplot(myopia) +  
 theme\_classic()+  
 geom\_boxplot( aes(x=myopic, y=hr\_tv)) +  
 xlab("myopia status") +  
 ylab("hours of tv per week") +  
 ggtitle("boxplot of hours of tv per week by myopia status")  
ggplot(myopia) +  
 theme\_classic()+  
 geom\_boxplot( aes(x=myopic, y=hr\_comp)) +  
 xlab("myopia status") +  
 ylab("hours of computer per week") +  
 ggtitle("boxplot of hours of commputer per week by myopia status")  
myopia\_glm <- glm(myopic~age+female+mom\_myopic+dad\_myopic+hr\_tv+hr\_comp, family=binomial, data=myopia)   
  
coef(myopia\_glm)  
int<- coef(myopia\_glm)[1]  
age <- coef(myopia\_glm)[2]  
ind\_female <- coef(myopia\_glm)[3]  
ind\_mom\_myopic<-coef(myopia\_glm)[4]  
ind\_dad\_myopic<- coef(myopia\_glm)[5]  
hr\_tv <-coef(myopia\_glm)[6]  
hr\_comp <- coef(myopia\_glm)[7]  
  
or\_female<- exp(ind\_female)  
or\_mom <- exp(ind\_mom\_myopic)  
or\_dad <- exp(ind\_dad\_myopic)  
or\_age <- exp(age)  
or\_tv<- exp(hr\_tv)  
or\_comp <- exp(hr\_comp)  
  
twoe<- exp(5\*hr\_tv)  
twof<- exp(25\*hr\_comp)  
twog <- exp(5\*age)  
myopia %>%   
 select(age, hr\_tv, hr\_comp) %>%   
 summary()  
exp(hr\_tv)  
range(myopia$age)  
#should print all code