Unit 11 Lab Report

Christopher Byrd

2024-12-10

Contents

Section 1: Hypothesis

The goal of this research paper is to determine how age, income, and regular attendance of church services affects prayer rates.

> Null: There is no significant correlation between age of those who pray regularly and their income, age, and church attendance.
> $Research_1$: There is a significant correlation between age of those who pray regularly and their income, age, and church attendance.

H_0: $\rho_{xy} = 0$
H_1: $r_{xy} \neq 0$

Section 2: Data

The data being utilized is from the Baylor Survey of Religious Life. Participants asked to respond to the question listed in the following code book:

Question or Description	Possible Answers
Age range of people who pray once or several times a day as reported in 2017.	Range: 17-95
By your best estimate, what was your total household income last year, before taxes?	1) $10,000 or less 2) $10,001 to $20,000 3) $20,001 to $35,000 4) $35,001 to $50,000 5) $50,001 to $100,000 6) $100,001 to $150,000 7) $150,001 or more
About how often do you spend time alone praying outside of religious services?	0) Never 1) Only on certain occasions 2) Once a week or less 3) A few times a week 4) Once a day 5) Several times a day
Attends a worship service about once a week or more.	1) No 2) Yes

Attendance was originally a nominal variable. To allow for a linear regression, it was recoded as a dummy variable.

Section 3: Method

To evaluate the data, the data set was first re-coded with shorter variable names, mainly because of my hate for long variable names. The new data set was then coerced to be a `data.frame` to make later actions easier (particularly with `ggplot2`, since data needs to be in tidy format).

```
# Re-coded with shorter names because idk I don't like long variable
# names. ¯\_( )_/¯

pf <- Module_11_Lab_Data$`Pray Frequency`
inc <- Module_11_Lab_Data$`Income Intervals`
age <- Module_11_Lab_Data$AGE
regat <- Module_11_Lab_Data$`Regular Attendance`

df <- data.frame(pf, inc, age, regat)
```

The next evaluation consisted of finding the median of income interval, age, and regular attendance. To cut down on the amount of code I needed to write, I made this a `for` loop. From here, the linear model (AKA regression) was calculated with prayer frequency regressed over income interval, age, and regular attendance of church services.

```
for (i in 2:4) {
  print(paste(median(df[[i]])))
}

lmresult <- lm(formula = pf ~ inc + age + regpr, df)
```

I did all this work below thinking I was working with Pearson's r, but I wasn't. I just put it here to display the work I put into this.

```
for (i in 1:length(lmresult$coefficients)) {
  if (lmresult$coefficients[i] > 0) {
    print(paste("The coefficient for", names(lmresult$coefficients[i]),
                "is", lmresult$coefficients[i], "and direct"))
  }   else {
    print(paste("The coefficient for", names(lmresult$coefficients[i]),
                "is", lmresult$coefficients[i], "and indirect"))
  }
  if (abs(lmresult$coefficients[i]) > 0.2) {
    print("There is a weak to strong relationship.")
  }   else {
    print("There is little to no relationship")
  }
}
```

Section 4: Results

Section 5: Discussion and Conclusion

Section 6: Graphics and Visuals

Figure 3: Matrix chart. r^2 of age and income interval.