Statistics One

Lecture 12

ANOVA in R

Two segments

- Mediation analysis in R
- Moderation analysis in R

Lecture 12
Segment 1

Mediation analysis in R

Goal

- Write a script in R to test for mediation
  - Three regression analyses
    - Outcome = Predictor
    - Predictor = Mediator
    - Outcome = Predictor + Mediator

Goal

- Write a script in R to test for mediation
  - Three regression analyses
    - lm(Y ~ X)
    - lm(X ~ M)
    - lm(Y ~ X + M)

Example

- Fictional data
  - Outcome (Y)
    - Happiness
  - Predictors (X, M)
    - Extraversion (X)
    - Diversity of life experience (M)

Mediation example

- Data are available in the following file:
  - STATS1.EX.05.txt

Write a script

- First line(s) of code should be comments

```
# Statistics One, Lecture 12, example script
# Mediation analysis
# X is extraversion
# Y is happiness
# M is diversity of life experience
```

Write a script

- Read data into a dataframe called “med”

  med <- read.table(“STATS1.EX.05.txt”, header = T)

Write a script

- Print descriptive statistics and histograms to test univariate normal assumptions

```
describe(med)
hist(med$happy)
hist(med$extra)
hist(med$diverse)
```

Write a script

- Print scatterplots to test linear and homoscedasticity assumptions

```
plot(med$happy ~ med$extra)
  abline(lm(med$happy ~ med$extra))
plot(med$diverse ~ med$extra)
  abline(lm(med$diverse ~ med$extra))
plot(med$happy ~ med$diverse)
  abline(lm(med$happy ~ med$diverse))
```

Write a script

- Conduct three regression analyses

```
model1 = lm(med$happy ~ med$extra)
  summary(model1)
model2 = lm(med$diverse ~ med$extra)
  summary(model2)
model3 = lm(med$happy ~ med$extra + med$diverse)
  summary(model3)
```

happy = 2.19 + 0.275(extra)

```
> model1 = lm(med$happy ~ med$extra)
> summary(model1)

Call:
lm(formula = med$happy ~ med$extra)

Residuals:
     Min       1Q   Median       3Q      Max
-2.15357 -0.49763 -0.08476  0.77762  2.05286

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.1902     0.3723   5.882 1.84e-08 ***
med$extra     0.2752     0.1022   2.693  0.00773 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9467 on 186 degrees of freedom
Multiple R-squared: 0.03753,  Adjusted R-squared: 0.03236
F-statistic: 7.253 on 1 and 186 DF,  p-value: 0.007726
```

diverse = 1.63 + 0.284(extra)

```
> model2 = lm(med$diverse ~ med$extra)
> summary(model2)

Call:
lm(formula = med$diverse ~ med$extra)

Residuals:
     Min       1Q   Median       3Q      Max
-1.97565 -0.69191  0.05636  0.52089  2.23715

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.62789    0.35343   4.606  7.6e-06 ***
med$extra    0.28374    0.09702   2.925  0.00388 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.8987 on 186 degrees of freedom
Multiple R-squared: 0.04396,  Adjusted R-squared: 0.03882
F-statistic: 8.553 on 1 and 186 DF,  p-value: 0.003878
```

happy = 1.886 + 0.222(extra) + 0.1868(diverse)

```
> model3 = lm(med$happy ~ med$extra + med$diverse)
> summary(model3)

Call:
lm(formula = med$happy ~ med$extra + med$diverse)

Residuals:
     Min       1Q   Median       3Q      Max
-2.22437 -0.56277 -0.06535  0.77563  1.94231

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.88612    0.38780   4.864 2.46e-06 ***
med$extra    0.22224    0.10315   2.155   0.0325 *
med$diverse  0.18680    0.07623   2.451   0.0152 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9342 on 185 degrees of freedom
Multiple R-squared: 0.06779,  Adjusted R-squared: 0.05771
F-statistic: 6.727 on 2 and 185 DF,  p-value: 0.001513
```

Sobel z = 1.88

```
$Indirect.Effect
[1] 0.05300156

$SE
[1] 0.02821695

$z.value
[1] 1.878359

$N
[1] 188
```

Final script

```
# Statistics One, Lecture 12, example script
# Mediation analysis
# X: Extraversion
# Y: Happiness
# M: Diversity of life experiences

# Create object med, which is a dataframe and contains the data for the mediation analysis
med <- read.table("STATS1.EX.05.txt", header = T)

# Load libraries
# The multilevel package is being used so we can conduct the Sobel test
library(psych)
library(multilevel)

# Print descriptive statistics (skew and kurtosis values validate univariate normal assumptions)
describe(med)
```

Final script

```
# Histograms (shape of distributions validate univariate normal assumptions)
layout(matrix(c(1,2,3,4), 2, 2, byrow = TRUE))
hist(med$happy, breaks = 6)
hist(med$extra, breaks = 6)
hist(med$diverse, breaks = 6)

# Scatterplots (plots validate linear and homoskedastic assumptions)
layout(matrix(c(1,2,3,4), 2, 2, byrow = TRUE))
plot(med$happy ~ med$extra)
abline(lm(med$happy ~ med$extra))
plot(med$diverse ~ med$extra)
abline(lm(med$diverse ~ med$extra))
plot(med$happy ~ med$diverse)
abline(lm(med$happy ~ med$diverse))
```

Final script

```
# Mediation analysis
model1 = lm(med$happy ~ med$extra)
summary(model1)
model2 = lm(med$diverse ~ med$extra)
summary(model2)
model3 = lm(med$happy ~ med$extra + med$diverse)
summary(model3)

# Sobel test (is the indirect path statistically significant?)
indirect = sobel(med$extra, med$diverse, med$happy)
indirect
```