Lecture 9
Segment 2

MR analysis, standardized

Goal

- Write a script in R
  - Multiple regression analysis with standardized regression coefficients
  - Compare models

Example

- Fictive data
  - Outcome (Y)
    - Physical endurance (endurance)
  - Predictors (X1, X2)
    - Age (age)
    - Years engaged in active exercise (activeyears)

Write a script

- To get standardized regression coefficients
  - Simply use the “scale” function

Write a script

- To compare models
  - Simply use the “anova” function

Additional code to script

```
# Regression analyses (standardized)
model1.z = lm(scale(endur$endurance)~scale(endur$age))
summary(model1.z)
model2.z = lm(scale(endur$endurance)~scale(endur$activeyears))
summary(model2.z)
model3.z = lm(scale(endur$endurance)~scale(endur$age) + scale(endur$activeyears))
summary(model3.z)

# Model comparisons
comp1 = anova(model1.z, model3.z)
comp1
comp2 = anova(model2.z, model3.z)
comp2
```

endurance = 0.00 + -0.13(age)

```
Call:
lm(formula = scale(endur$endurance) ~ scale(endur$age))

Residuals:
     Min       1Q   Median       3Q      Max
-2.31751 -0.70552  0.00901  0.62584  2.85325

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)       2.837e-17  6.351e-02   0.000   1.0000
scale(endur$age) -1.259e-01  6.364e-02  -1.978   0.0491 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9941 on 243 degrees of freedom
Multiple R-squared: 0.01584,  Adjusted R-squared: 0.01179
F-statistic: 3.911 on 1 and 243 DF,  p-value: 0.04911
```

endurance = 0.00 + 0.34(activeyears)

```
Call:
lm(formula = scale(endur$endurance) ~ scale(endur$activeyears))

Residuals:
     Min       1Q   Median       3Q      Max
-2.19330 -0.65320  0.05157  0.53104  2.87296

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)               -1.340e-17  6.028e-02   0.000        1
scale(endur$activeyears)  3.365e-01  6.041e-02   5.571  6.7e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9436 on 243 degrees of freedom
Multiple R-squared: 0.1133, Adjusted R-squared: 0.1096
F-statistic: 31.04 on 1 and 243 DF,  p-value: 6.697e-08
```

endurance = 0.00 + -0.24(age) + 0.40(activeyears)

```
Call:
lm(formula = scale(endur$endurance) ~ scale(endur$age) + scale(endur$activeyears))

Residuals:
     Min       1Q   Median       3Q      Max
-2.01489 -0.63813  0.05269  0.52061  2.51664

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)              -2.183e-17  5.857e-02   0.000 1.000000
scale(endur$age)         -2.402e-01  6.119e-02  -3.925 0.000113 ***
scale(endur$activeyears)  4.044e-01  6.119e-02   6.610 2.44e-10 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9168 on 242 degrees of freedom
Multiple R-squared: 0.1663, Adjusted R-squared: 0.1594
F-statistic: 24.14 on 2 and 242 DF,  p-value: 2.754e-10
```

Model comparison

```
Analysis of Variance Table

Model 1: scale(endur$endurance) ~ scale(endur$age)
Model 2: scale(endur$endurance) ~ scale(endur$age) + scale(endur$activeyears)
  Res.Df    RSS Df Sum of Sq      F    Pr(>F)
1    243 240.13
2    242 203.41  1    36.721 43.687 2.438e-10 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Model comparison

```
Analysis of Variance Table

Model 1: scale(endur$endurance) ~ scale(endur$activeyears)
Model 2: scale(endur$endurance) ~ scale(endur$age) + scale(endur$activeyears)
  Res.Df    RSS Df Sum of Sq      F    Pr(>F)
1    243 216.36
2    242 203.41  1     12.95 15.406 0.0001131 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Compare to correlations

	age	activeyears	endurance
age	1.0000000	0.2826635	-0.1258528
activeyears	0.2826635	1.0000000	0.3365433
endurance	-0.1258528	0.3365433	1.0000000