Path Models II

Practical examples

Path Models II

Today's goal:

How to evaluate path models

Outline:

- Examples
- Model fit

Practical examples

Specifying and testing models

Practical examples

Knijnenburg et al. (2012): "Inspectability and Control in Social Recommenders", *RecSys'12*

The TasteWeights system uses the overlap between you and your friends' Facebook "likes" to give you music recommendations.

- Friends "weights" based on the overlap in likes w/ user
- Friends' other music likes—the ones that are not among the user's likes—are tallied by weight
- Top 10 is displayed to the user

Practical examples

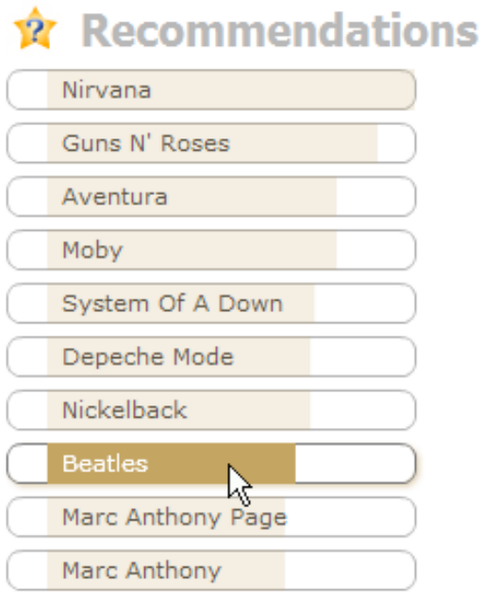
3 control conditions:

- No control (just use likes)
- Item control (weigh likes)
- Friend control (weigh friends)

Practical examples

2 inspectability conditions:

- List of recommendations vs. recommendation graph



Practical examples

twp.dat, variables:

- **cgraph**: inspectability manipulation (0: list, 1: graph)
- **citem-cfriend**: two dummies for the control manipulation (baseline: no control)
- **cig** (citem * cgraph) and **cfg** (cfriend * cgraph)
- **satisfaction** with the system (7 items)
- **quality** of the recommendations (6 items)
- **perceived_control** over the system (5 items)
- **understandability** of the system (5 items)

Practical examples

twp.dat, variables:

- **expertise**: user music expertise
- **trust**: propensity to trust
- **familiarity**: familiarity with recommenders
- average **rating** of, and number of **known** items in, the top 10
- **time** taken to inspect the recommendations

Linear regression

1. Import the dataset

2. Run the model

```
m <- lm(satisfaction ~ quality + perceived_control +
understandability, data = twp)
```

3. Get the model summary

```
summary(m)
```

Linear regression

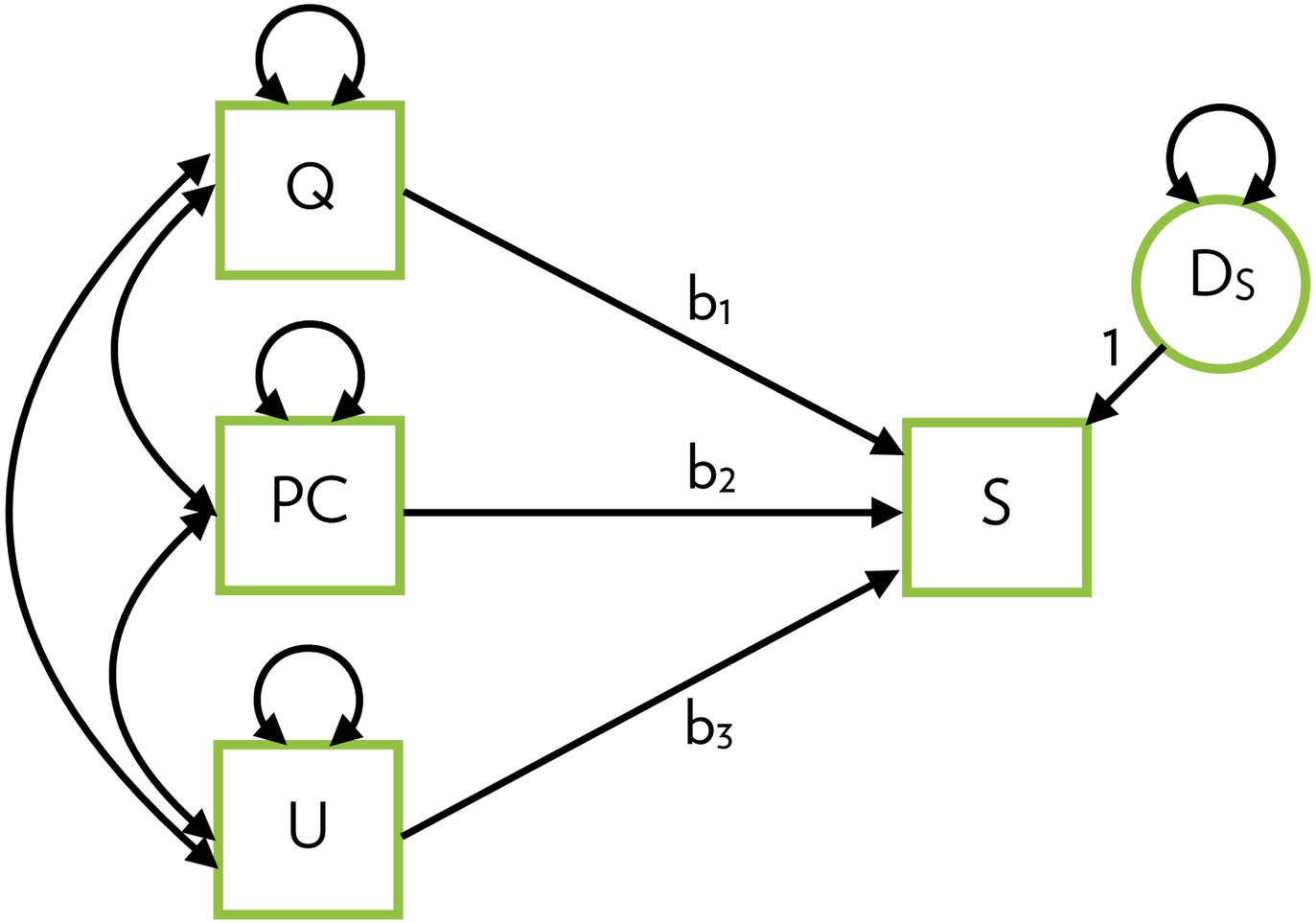
```
Residuals:
     Min       1Q   Median       3Q      Max 
-17.6760  -2.2287   0.3407   2.9143   9.6971 

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)    
(Intercept)          2.35401    0.50722   4.641 5.48e-06 ***
quality              0.40151    0.06054   6.632 1.87e-10 ***
perceived_control    0.74217    0.08400   8.836  < 2e-16 ***
understandability    0.11932    0.08136   1.467    0.144    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.225 on 263 degrees of freedom
Multiple R-squared:  0.5185, Adjusted R-squared:  0.513 
F-statistic: 94.42 on 3 and 263 DF,  p-value: < 2.2e-16
```

Same in lavaan

How many degrees of freedom?

How many observations?

How many parameters?



Same in lavaan

1. Import the dataset

2. Install and activate the lavaan package

3. Write a model specification

```
model <- "satisfaction ~ quality + perceived_control +
understandability"
```

4. Fit the model

```
fit <- sem(model, data="twp")
```

5. Get the model summary

```
summary(fit, rsquare=T)
```

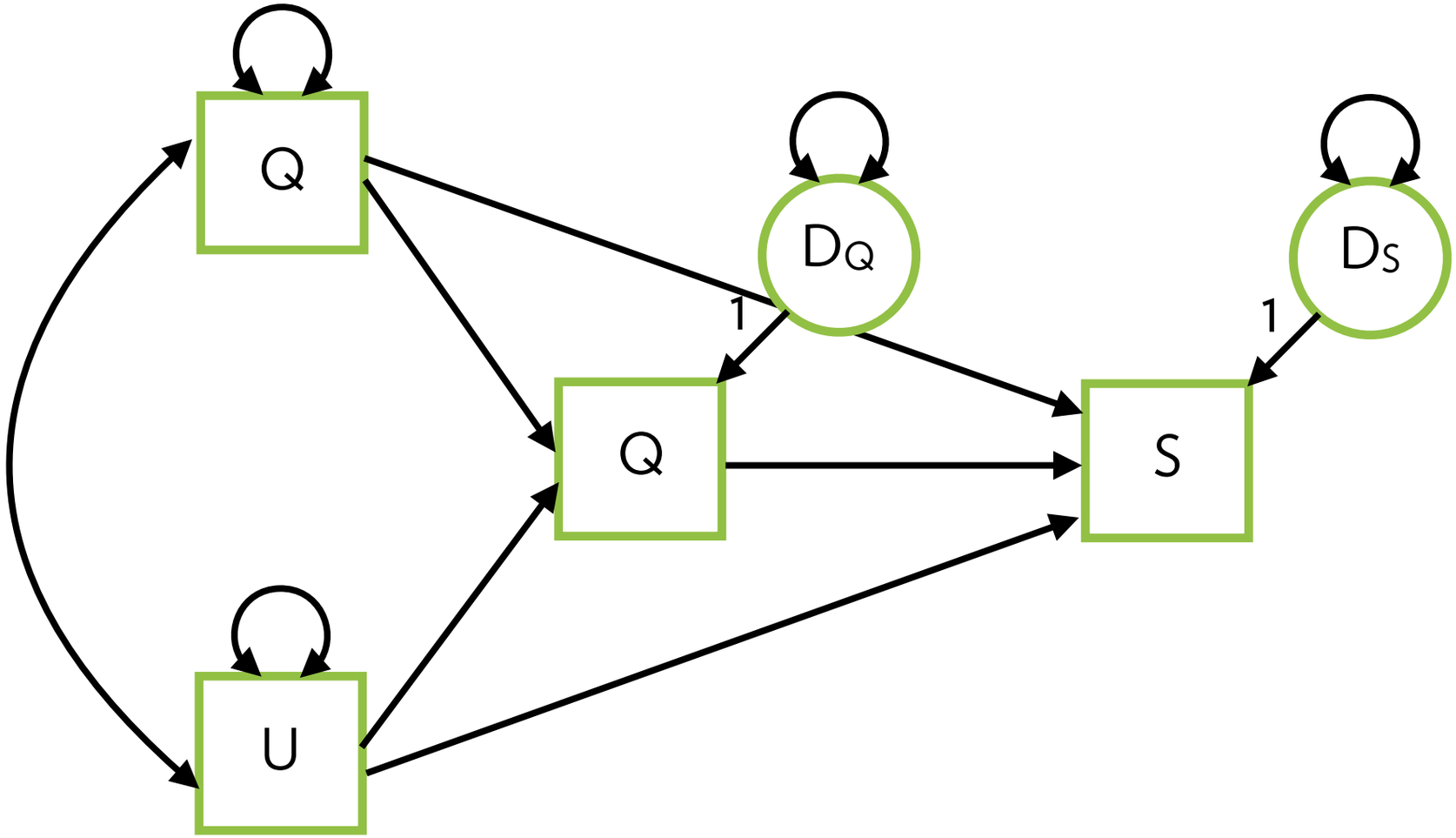
Same in lavaan

```
  Number of observations                           267

  Estimator                                         ML
  Minimum Function Test Statistic                0.000
  Degrees of freedom                                 0

Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)
  satisfaction ~
    quality           0.402    0.060    6.683    0.000
    perceivd_cntrl    0.742    0.083    8.903    0.000
    understandblty    0.119    0.081    1.478    0.139

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .satisfaction     17.582    1.522   11.554    0.000

R-Square:
                   Estimate
    satisfaction      0.519
```

Mediation

What if the effect of perceived_control and understandability are mediated by quality?

How many degrees of freedom?



Mediation

Write a model specification

```
mediated <- "satisfaction ~ quality + perceived_control +
understandability
quality ~ perceived_control + understandability"
```

Fit the model

```
fit <- sem(mediated, data="twp")
```

Get the model summary

```
summary(fit, rsquare=T)
```

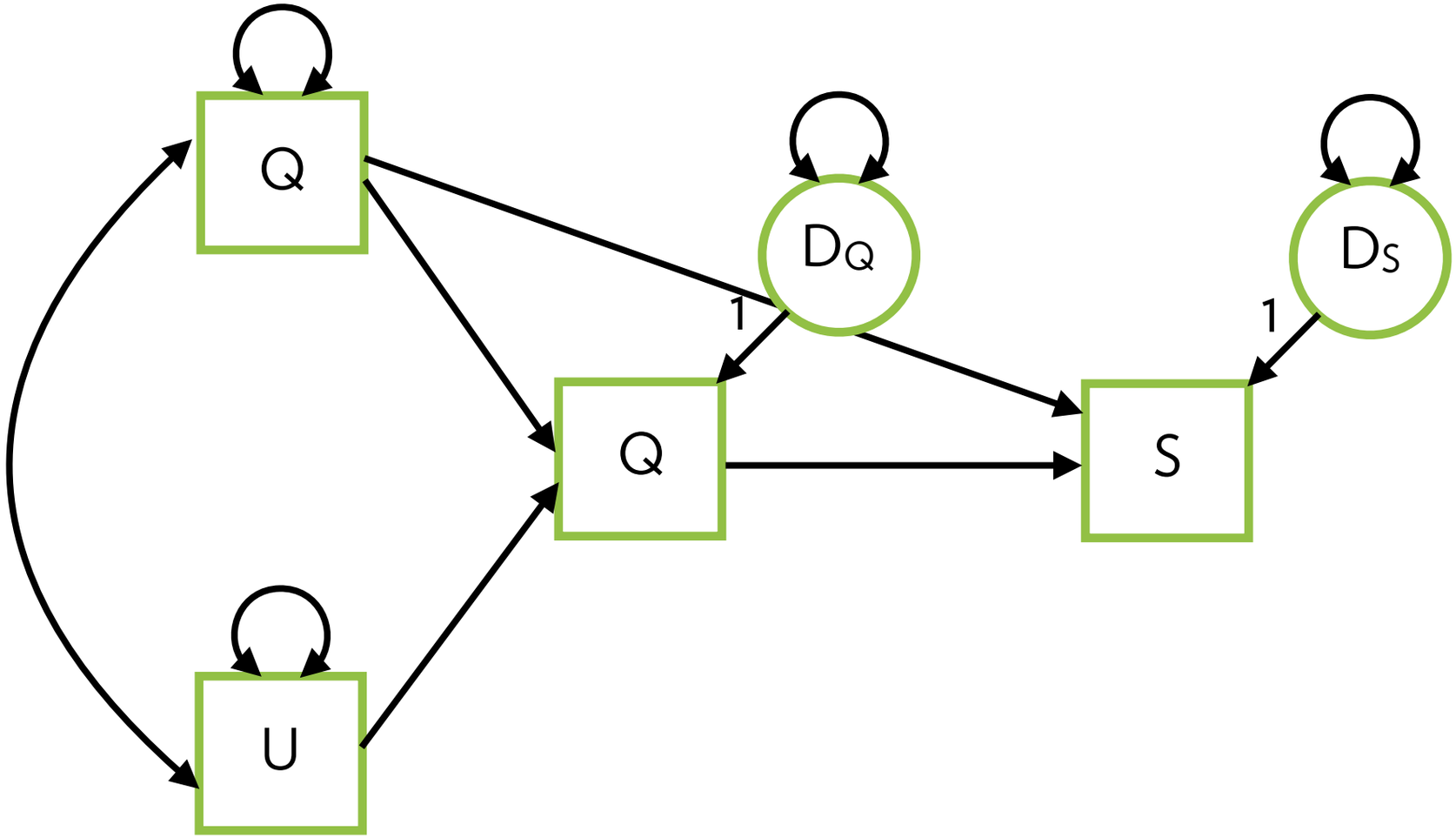
Mediation

```
Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)
  satisfaction ~
    quality           0.402    0.060    6.683    0.000
    perceivd_cntrl    0.742    0.083    8.903    0.000
    understandblty    0.119    0.081    1.478    0.139
  quality ~
    perceivd_cntrl    0.688    0.074    9.334    0.000
    understandblty    0.149    0.082    1.819    0.069

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .satisfaction     17.582    1.522   11.554    0.000
   .quality          18.241    1.579   11.554    0.000

R-Square:
                   Estimate
    satisfaction      0.519
    quality           0.292
```

Trimming

Remove understandability -> satisfaction

How many degrees of freedom?



Trimming

Write a model specification

```
trimmed <- "satisfaction ~ quality + perceived_control
quality ~ perceived_control + understandability"
```

Fit the model

```
fit <- sem(trimmed, data="twp")
```

Get the model summary

```
summary(fit, rsquare=T)
```

Trimming

```
Number of observations                        267

Estimator                                      ML
Minimum Function Test Statistic             2.175
Degrees of freedom                              1
P-value (Chi-square)                        0.140
```

```
Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)
  satisfaction ~
    quality           0.411    0.060    6.860    0.000
    perceivd_cntrl    0.767    0.082    9.353    0.000
  quality ~
    perceivd_cntrl    0.688    0.074    9.334    0.000
    understandblty    0.149    0.082    1.819    0.069

R-Square:
                   Estimate
    satisfaction      0.515
    quality           0.292
```

t-test

1. Run the model

```
m <- t.test(understandability ~ cgraph, data = twp)
```

2. Get the results

```
m

    Welch Two Sample t-test

data:  understandability by cgraph
t = -3.7022, df = 255.09, p-value = 0.000262
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.2819053 -0.6972146
sample estimates:
mean in group 0 mean in group 1
       1.083969        2.573529
```

As a regression

1. Run the model

```
m <- lm(understandability ~ cgraph, data = twp)
```

2. Get the model summary

```
summary(m)
```

As a regression

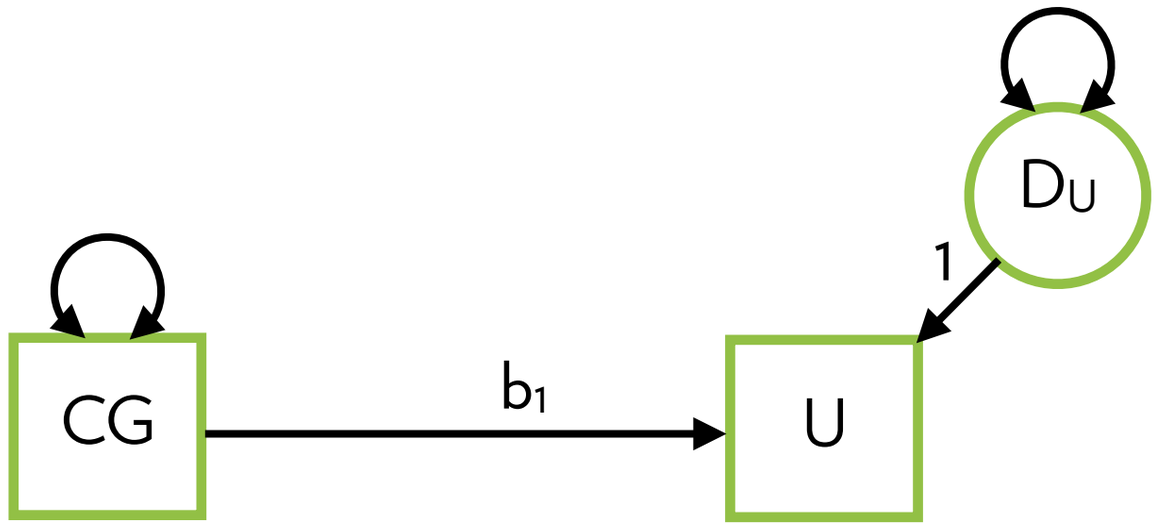
```
Residuals:
    Min      1Q  Median      3Q     Max
-7.0840 -2.0840  0.4265  2.4265  4.9160

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.0840     0.2863   3.786 0.000189 ***
cgraph        1.4896     0.4011   3.713 0.000249 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.277 on 265 degrees of freedom
Multiple R-squared:  0.04946,Adjusted R-squared:  0.04587
F-statistic: 13.79 on 1 and 265 DF,  p-value: 0.0002494
```

Same in lavaan

How many degrees of freedom?

How many observations?

How many parameters?



Same in lavaan

1. Write a model specification

```
tmodel <- “understandability ~ cgraph"
```

2. Fit the model

```
fit <- sem(tmodel, data="twp")
```

3. Get the model summary

```
summary(fit, rsquare=T)
```

Same in lavaan

```
Number of observations                           267

Estimator                                         ML
Minimum Function Test Statistic                0.000
Degrees of freedom                                 0
Minimum Function Value                 0.0000000000000

Parameter Estimates:

  Information                                 Expected
  Standard Errors                             Standard

Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)
  understandability ~
    cgraph            1.490    0.400    3.727    0.000

Variances:
                 Estimate  Std.Err  z-value  P(>|z|)
   .understandblty   10.657    0.922   11.554    0.000
```

ANOVA

1. Import the dataset (use the original tw here!)

2. Run the model

```
m <- aov(understandability ~ control, data = tw)
```

3. Get the model summary

```
summary(m)
```

```
             Df Sum Sq Mean Sq F value Pr(>F)
control       2   93.3   46.65   4.246 0.0153 *
Residuals   264 2900.1   10.99
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

As a regression

1. Get the lm summary

```
summary.lm(m)
```

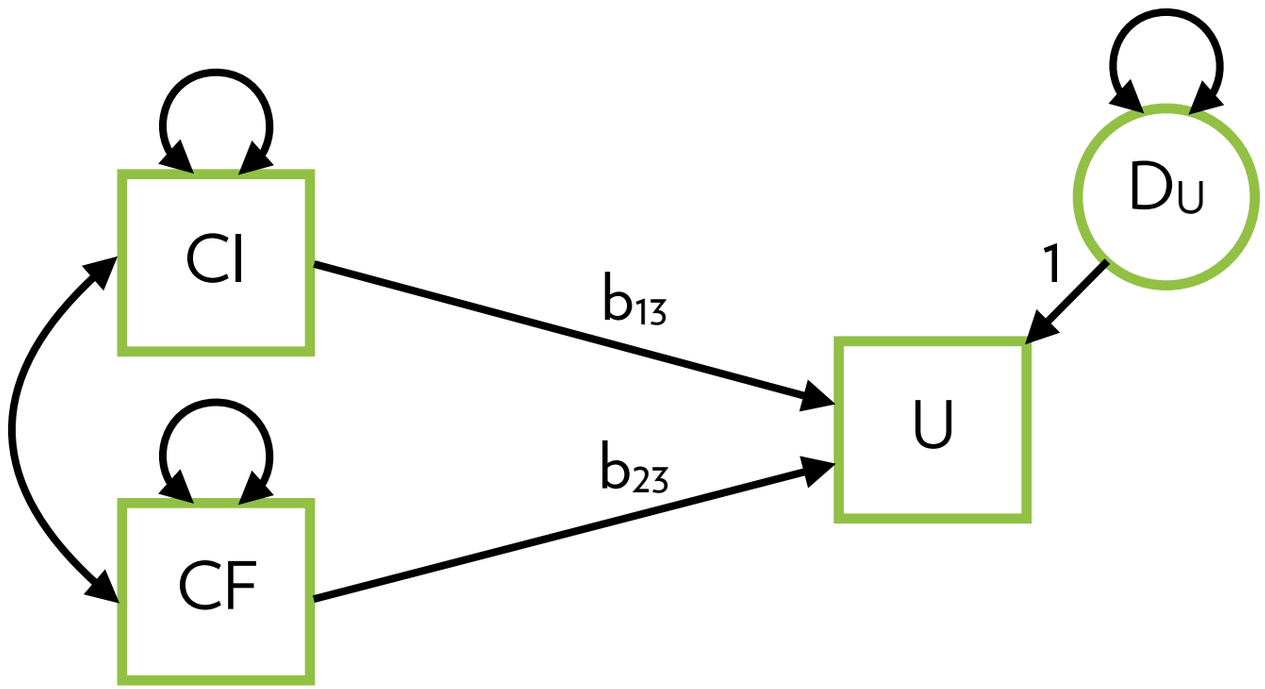
```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)    1.0435     0.3455   3.020  0.00278 **
controlfriend  1.3610     0.4928   2.762  0.00615 **
controlitem    1.0728     0.4971   2.158  0.03183 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.314 on 264 degrees of freedom
Multiple R-squared:  0.03117,Adjusted R-squared:  0.02383
F-statistic: 4.246 on 2 and 264 DF,  p-value: 0.01531
```

Same in lavaan

How many degrees of freedom?

How many observations?

How many parameters?



Same in lavaan

1. Write a model specification

```
amodel <- "understandability ~ ci*citem + cf*cfriend"
```

2. Fit the model

```
fit <- sem(amodel, data="twp")
```

3. Conduct the omnibus test

```
lavTestWald(fit, "ci==0; cf==0”)
```

4. Get the model summary

```
summary(fit, rsquare=T)
```

Same in lavaan

```
$stat
[1] 8.588976

$df
[1] 2

$p.value
[1] 0.01364356

$se
[1] “standard"
```

Chi-square test with 2 degrees of freedom

$$\chi^2(2) = 8.59,\ p = .0136$$

Same in lavaan

```
Number of observations                           267

Estimator                                         ML
Minimum Function Test Statistic                0.000
Degrees of freedom                                 0

Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)
  understandability ~
    citem     (ci)    1.073    0.494    2.170    0.030
    cfriend   (cf)    1.361    0.490    2.778    0.005

Variances:
                 Estimate  Std.Err  z-value  P(>|z|)
   .understandblty   10.862    0.940   11.554    0.000

R-Square:
                 Estimate
    understandblty    0.031
```

Factorial ANOVA

1. Run the model

```
m <- aov(understandability ~ inspectability * control, data =
tw)
```

2. Get the anova results with type III Sum of Squares (in the ‘car’ package)

```
Anova(m, type=3)
```

Factorial ANOVA

```
Anova Table (Type III tests)

Response: understandability
                         Sum Sq  Df F value   Pr(>F)
(Intercept)              172.20   1 16.3518 6.93e-05 ***
inspectability            73.09   1  6.9404 0.008931 **
control                   29.61   2  1.4060 0.246983
inspectability:control     3.88   2  0.1841 0.831962
Residuals               2748.50 261
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

As a regression

1. Get the lm summary

```
summary.lm(m)
```

```
Coefficients:
                                          Estimate Std. Error t value Pr(>|t|)
(Intercept)                                0.1522     0.4785   0.318  0.75070
inspectabilitygraphview                    1.7826     0.6766   2.634  0.00893 **
controlfriend                              1.6739     0.6766   2.474  0.01400 *
controlitem                                1.1555     0.7064   1.636  0.10307
inspectabilitygraphview:controlfriend     -0.5854     0.9652  -0.607  0.54469
inspectabilitygraphview:controlitem       -0.3031     0.9757  -0.311  0.75633
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.245 on 261 degrees of freedom
Multiple R-squared:  0.08181,  Adjusted R-squared:  0.06422
F-statistic: 4.651 on 5 and 261 DF,  p-value: 0.0004388
```

Same in lavaan

1. Write a model specification

```
fmodel <- "understandability ~ ci*citem + cf*cfriend + cgraph
+ cig*cig + cfg*cfg"
```

2. Fit the model

```
fit <- sem(fmodel, data="twp")
```

3. Conduct the omnibus tests

```
lavTestWald(fit, "cig==0; cfg==0")
lavTestWald(fit, "ci==0; cf==0")
```

4. Get the model summary

```
summary(fit, rsquare=T)
```

Same in lavaan

```
$stat
[1] 0.3766611

$df
[1] 2

$p.value
[1] 0.8283409

$se
[1] "standard"
```

Interaction effect:

$\chi^2(2) = 0.38, p = .828$

Same in lavaan

```
$stat
[1] 6.530504

$df
[1] 2

$p.value
[1] 0.03818731

$se
[1] "standard"
```

Main effect:

$X^2(2) = 6.53, p = .0382$

Meaningless, though... why?

Same in lavaan

```
Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)
  understandability ~
    citem     (ci)     1.156    0.698    1.655    0.098
    cfriend   (cf)     1.674    0.669    2.502    0.012
    cgraph             1.783    0.669    2.665    0.008
    cig       (cig)   -0.303    0.965   -0.314    0.753
    cfg       (cfg)   -0.585    0.954   -0.613    0.540

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .understandblty   10.294    0.891   11.554    0.000

R-Square:
                   Estimate
    understandblty    0.082
```

Robust estimator

1. Write a model specification

```
trimmed <- "satisfaction ~ quality + perceived_control
quality ~ perceived_control + understandability"
```

2. Fit the model **with a robust estimator**

```
fit <- sem(trimmed, data="twp", estimator="mlr")
```

3. Get the model summary

```
summary(fit, rsquare=T)
```

Non-robust

```
Number of observations                                267

Estimator                                              ML
Minimum Function Test Statistic                     2.175
Degrees of freedom                                      1
P-value (Chi-square)                                0.140

Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)
  satisfaction ~
    quality           0.411    0.060    6.860    0.000
    perceivd_cntrl    0.767    0.082    9.353    0.000
  quality ~
    perceivd_cntrl    0.688    0.074    9.334    0.000
    understandblty    0.149    0.082    1.819    0.069

R-Square:
                   Estimate
    satisfaction      0.515
    quality           0.292
```

Robust

```
Number of observations                                267

Estimator                                          ML      Robust
  Minimum Function Test Statistic                2.175       1.476
  Degrees of freedom                                 1           1
  P-value (Chi-square)                           0.140       0.224
  Scaling correction factor                                  1.473
    for the Yuan-Bentler correction

Parameter Estimates:

  Information                                 Observed
  Standard Errors                    Robust.huber.white
```

Robust

```
Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)
  satisfaction ~
    quality           0.411    0.071    5.759    0.000
    perceivd_cntrl    0.767    0.084    9.085    0.000
  quality ~
    perceivd_cntrl    0.688    0.075    9.237    0.000
    understandblty    0.149    0.098    1.516    0.130

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)
   .satisfaction      2.534    0.576    4.398    0.000
   .quality           6.582    0.315   20.903    0.000

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .satisfaction     17.726    1.943    9.121    0.000
   .quality          18.241    1.903    9.588    0.000

R-Square:
                   Estimate
    satisfaction      0.515
    quality           0.292
```

Model fit

how to evaluate a path model

Model chi-square

Chi-square test of model fit:

- Tests whether there any significant misfit between estimated and observed correlation matrix
- The chi-square value is compared against the model df
- The test does **not** work for a just-identified model
- When the model df is large, the model is usually significant ($p < .05$)... parsimonious models are rarely perfect!
- Alternative metric: chi-squared / df < 3 (good fit) or < 2 (great fit)

Baseline test

Chi-square test against baseline:

- Tests whether the model that considers all variables independent (unrelated)
- df of this test is # parameters – # variables
- This is like the F-test in linear regression
- In other words, it is almost always significant, and thus rather useless

Relative fit

Relative fit indices: how much improvement compared to the baseline?

- Ranging from 0.00 to 1.00

CFI: $((\chi^2_b - df_b) - (\chi^2_m - df_m)) / (\chi^2_b - df_b)$

- Bentler suggests that it should be > 0.96

TLI (or NNFI): $((\chi^2_b/df_b) - (\chi^2_m/df_m)) / (\chi^2_b/df_b - 1)$

- Suggested that it should be > 0.95

Approximate fit

RMSEA (root mean square error of approximation):

- $\sqrt{(\chi^2_m - df_m)} / \sqrt{(df_m * (N-1))}$
- < 0.01 (excellent fit), < 0.05 (good fit), < 0.08 (mediocre fit), > 0.10 (poor fit)
- May be too high with low N and low df
- 90% confidence interval of RMSEA should fall below 0.10
- P-value RMSEA <= 0.05 should be > 0.05 (meaning the RMSEA is not significantly larger than 0.05)

Model fit in R

Getting additional fit statistics in R:

```
summary(fit, rsquare=T, fit.measures=T)
```

```
  Number of observations                           267

  Estimator                                         ML      Robust
  Minimum Function Test Statistic                2.175       1.476
  Degrees of freedom                                 1           1
  P-value (Chi-square)                           0.140       0.224
  Scaling correction factor                                  1.473
    for the Yuan-Bentler correction

Model test baseline model:

  Minimum Function Test Statistic              287.397     209.526
  Degrees of freedom                                 5           5
  P-value                                        0.000       0.000
```

Model fit in R

```
User model versus baseline model:

  Comparative Fit Index (CFI)                    0.996       0.998
  Tucker-Lewis Index (TLI)                       0.979       0.988

  Robust Comparative Fit Index (CFI)                         0.997
  Robust Tucker-Lewis Index (TLI)                            0.987


Root Mean Square Error of Approximation:

  RMSEA                                          0.066       0.042
  90 Percent Confidence Interval          0.000  0.191       0.000  0.151
  P-value RMSEA <= 0.05                          0.266       0.405

  Robust RMSEA                                               0.051
  90 Percent Confidence Interval                             0.000  0.212
```

“It is the mark of a truly intelligent person
to be moved by statistics.”

T H A N K S !

George Bernard Shaw