The semnova Package for Latent Repeated Measures ANOVA

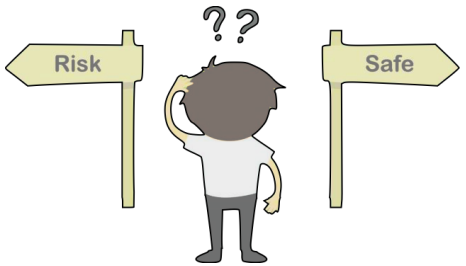
Benedikt Langenberg, RWTH Aachen University
Axel Mayer, RWTH Aachen University

RWTH AACHEN UNIVERSITY

Exemplary Research Question

Does noise affect risky decision making? (Syndicus et al., 2016)

no noise vs. speech vs. office noises



Measured variables

- The Choice Dilemma Questionnaire (12 items, percentages)
- The Risk Scenario Questionnaire (20 items, 10-point scale)

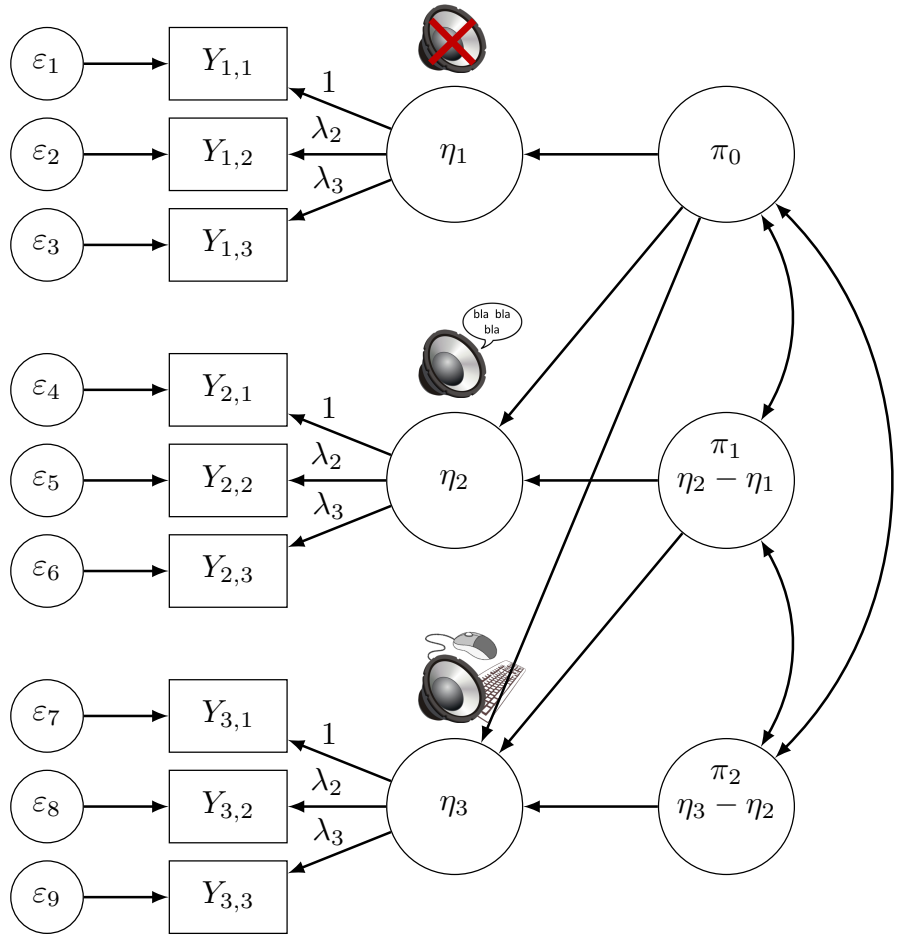
$\Rightarrow$ traditionally analyzed by repeated measures ANOVA (GLM) using averaged indicators (mean scores) as dependent variables or using separate analyses

Advantages of SEM over repeated measures ANOVA

- More power due to explicit error modeling
- More complex covariance structures allowed
  - Data does not have to satisfy sphericity
  - Covariance structure may differ among groups
  - Test for error structures available (e.g. compound symmetry, sphericity)
- Interindividual differences may be investigated
  - Exogenous variables may be included explaining differences among conditions
- Advanced methods of handling missing data and violations of normality available
- Model fit available
- Robust estimators available
- Test for measurement invariance

A Minimal Example

Latent repeated measures ANOVA is based on the latent growth components approach (Mayer et al., 2012)

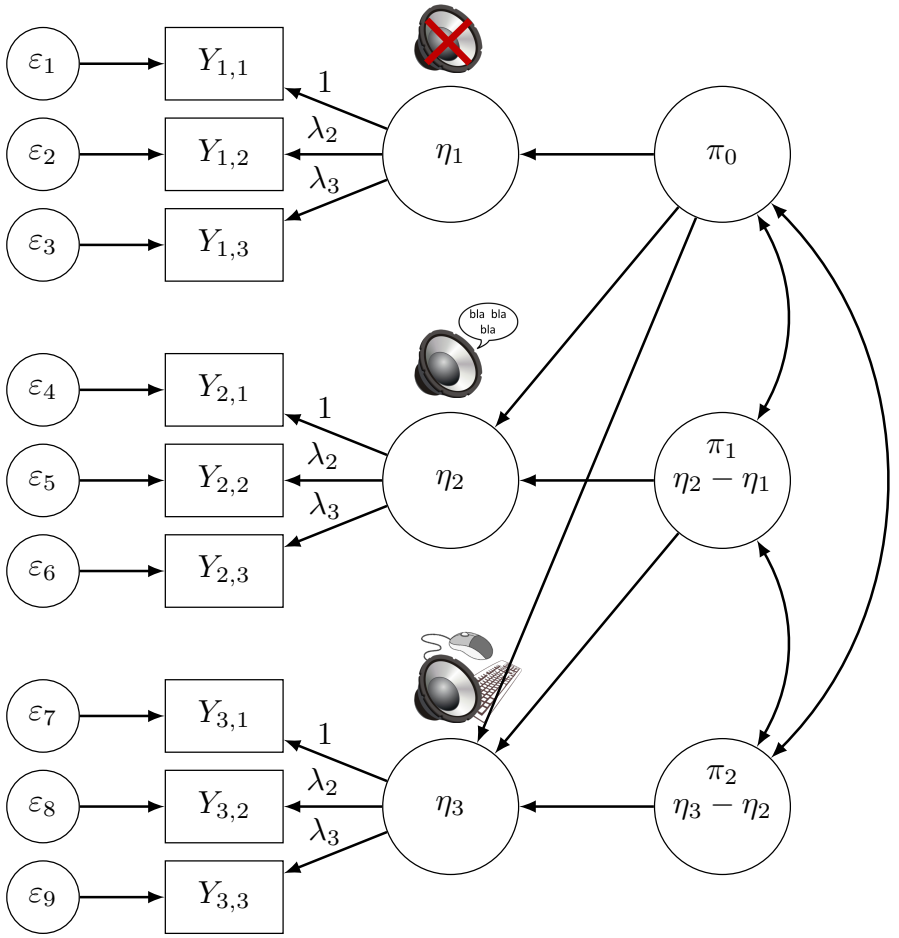


A Minimal Example

Latent repeated measures ANOVA is based on the latent growth components approach (Mayer et al., 2012)

In general, transform η into latent effect variables π:

$$\boldsymbol{\pi} = \overset{\mathbf{C}}{\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}} \boldsymbol{\eta}$$



A Minimal Example

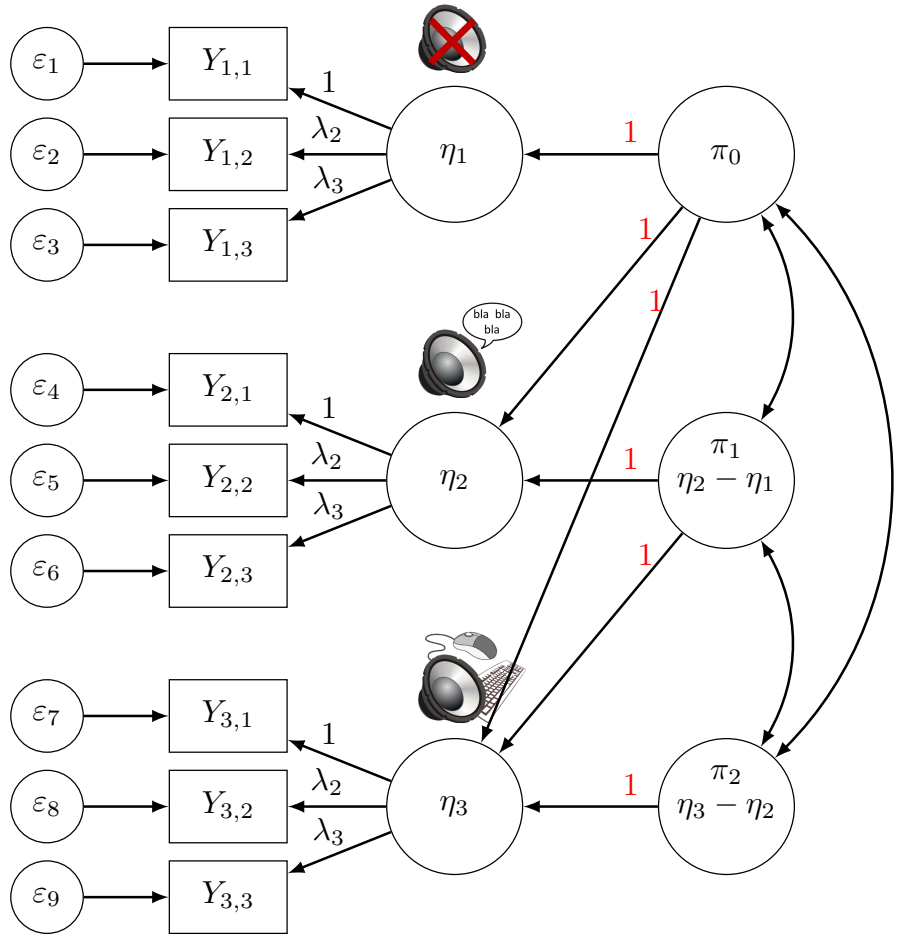
Latent repeated measures ANOVA is based on the latent growth components approach (Mayer et al., 2012)

In general, transform η into latent effect variables π:

$$\boldsymbol{\pi} = \overset{\mathbf{C}}{\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}} \boldsymbol{\eta}$$

Add row to make C invertible:

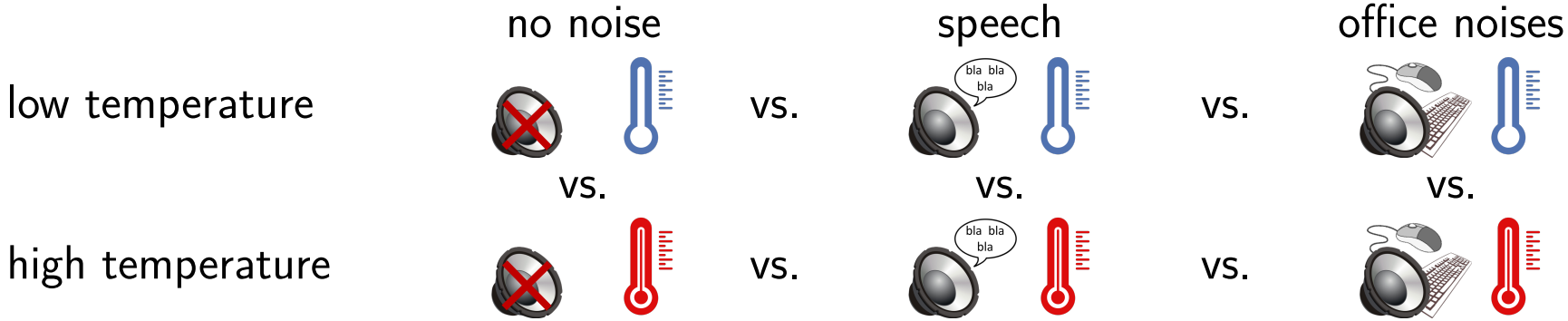
$$\boldsymbol{\pi} = \overset{\mathbf{C}}{\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}} \boldsymbol{\eta} \quad \Leftrightarrow \quad \boldsymbol{\eta} = \overset{\mathbf{B}^*}{\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}} \boldsymbol{\pi}$$

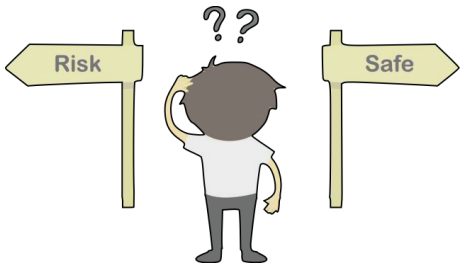


The semnova Package for Latent Repeated Measures ANOVA — Benedikt Langenberg (langenberg@psych.rwth-aachen.de); Axel Mayer — RWTH Aachen University — 28.02.2020

A Larger Example

Do noise or temperature affect risky decision making? (Syndicus et al., 2016)





Measured variables (again)

- The Choice Dilemma Questionnaire (12 items, percentages)
- The Risk Scenario Questionnaire (20 items, 10-point scale)

A Larger Example

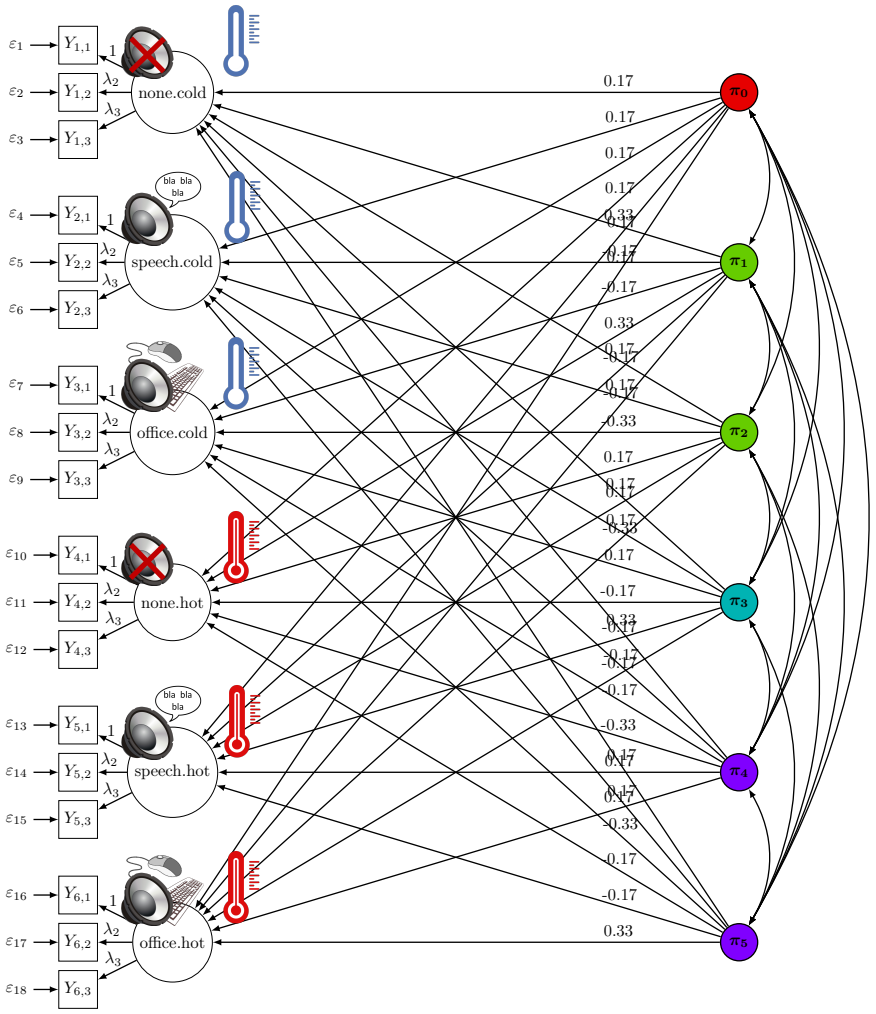
What the package does:

$$\mathbf{C} = \begin{array}{ll} & \text{Factor: noise / temp.} \\ & \begin{array}{cccccc} \text{none} & \text{low} & \text{high} & \text{none} & \text{low} & \text{high} \\ \text{cold} & \text{cold} & \text{cold} & \text{hot} & \text{hot} & \text{hot} \end{array} \\ \text{contrasts} \begin{array}{l} \text{intercept}\,(\pi_0) \\ \text{noise1}\,(\pi_1) \\ \text{noise2}\,(\pi_2) \\ \text{temp1}\,(\pi_3) \\ \text{noise1:temp1}\,(\pi_4) \\ \text{noise2:temp1}\,(\pi_5) \end{array} & \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 & 1 & -1 \\ 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 & -1 & 1 \end{pmatrix} \end{array}$$

dependent variables

$$\mathbf{B}^* = \begin{array}{ll} & \text{Factor: noise / temp.} \\ & \begin{array}{cc} \text{none} & \text{cold} \\ \text{low} & \text{cold} \\ \text{high} & \text{cold} \\ \text{none} & \text{hot} \\ \text{low} & \text{hot} \\ \text{high} & \text{hot} \end{array} \\ \text{dep. variables} & \begin{pmatrix} 0.17 & 0.33 & 0.17 & 0.17 & 0.33 & 0.17 \\ 0.17 & -0.17 & 0.17 & 0.17 & -0.17 & 0.17 \\ 0.17 & -0.17 & -0.33 & 0.17 & -0.17 & -0.33 \\ 0.17 & 0.33 & 0.17 & -0.17 & -0.33 & -0.17 \\ 0.17 & -0.17 & 0.17 & -0.17 & 0.17 & -0.17 \\ 0.17 & -0.17 & -0.33 & -0.17 & 0.17 & 0.33 \end{pmatrix} \end{array}$$

contrasts: (π_0) intercept, (π_1) noise1, (π_2) noise2, (π_3) temp1, (π_4) noise1:temp1, (π_5) noise2:tmp1



A Larger Example

R package **semnova** implements latent repeated measures ANOVA using the SEM software package **lavaan** (Rosseel, 2012)

`semnova(...)`

- `data`: Data frame.
- `idata`: Matrix. Design matrix of the within-subject factors. Similar to the `idata` object in the `car` package.
- `mmodel`: List of character vectors. Each Element represents a latent dependent variable measured by the manifest indicators that are included in the character vector.
- (`icontrasts`: Character string. Default is "`contr.sum`". Specifies the type of contrasts to be used.)

A Larger Example

R package **semnova** implements latent repeated measures ANOVA using the SEM software package **lavaan** (Rosseel, 2012)

```
> head(data)

          Y11        Y12          Y13        Y21        Y22         Y23        Y31        Y32         Y33
1  0.2004584  1.7122335  1.86085243 -0.2483409  1.6641324  1.13194771 -0.1965220 -0.3876609  0.72844499
2  0.1862405  0.2657459 -0.09357174 -0.4528128 -0.1170475  1.19826603 -0.9518866 -0.9960841  2.29458413
3  4.2185414  4.1228080  0.72206631  1.5574117  0.2289177 -0.04011789  2.9190039  3.1094043  1.00288043
4  1.4312455  1.7345077  1.13627636  0.3325998  0.9038465  2.10896642  1.6668742  1.4398952  0.74878589
5  2.1724362  1.6230909  1.01891961  0.1978093 -0.6514590  0.80023643  0.2205186  2.4143326 -0.08174437
6  1.6229890  2.5948945 -0.01458020  3.0525912  1.7065496  1.38144415  3.6329593  2.2300305  1.78360290
          Y41        Y42        Y43        Y51        Y52         Y53        Y61        Y62        Y63
1 2.28418999  0.7364414  1.1718701  0.4309800  2.1110208 -0.04430411  1.0015881 -0.2578211  0.5504424
2 0.04583038 -0.4760048  0.8953298  0.1435606  0.9644196 -0.74461258  0.3374447  3.1675914  1.4721956
3 2.04458084  1.1012540  3.6971539  3.7982794  1.1863811  3.71389785  3.0867334  1.0604590  0.9689124
4 0.90092458  0.5537761  1.4479135  0.6998906  1.4130335  1.26029682  1.2081589  0.2769748 -0.9719528
5 1.94201956  1.7937876  2.1433103  0.1461332 -0.5443832  1.30563461  1.0690851  0.2793267  1.9604143
6 1.69692936  1.4636682  0.5518675  3.4503364  0.2924008  2.18199691  2.6190934  1.3106907  1.8708039
```

A Larger Example

R package **semnova** implements latent repeated measures ANOVA using the SEM software package **lavaan** (Rosseel, 2012)

```
> library(semnova)

> fit <- semnova(
      data   = data,
      idata  = expand.grid(
          noise       = c("none", "speech", "office"),
          temperature = c("cold", "hot")
      ),
      mmodel = list(
          none.cold = c("Y11", "Y12", "Y13"),
          low.cold  = c("Y21", "Y22", "Y23"),
          high.cold = c("Y31", "Y32", "Y33"),
          none.hot  = c("Y41", "Y42", "Y43"),
          low.hot   = c("Y51", "Y52", "Y53"),
          high.hot  = c("Y61", "Y62", "Y63")
      ),
  )

> summary(fit)

```

A Larger Example

R package **semnova** implements latent repeated measures ANOVA using the SEM software package **lavaan** (Rosseel, 2012)

```
> library(semnova)

> fit <- semnova(
      data   = data,
      idata  = expand.grid(
          noise       = c("none", "speech", "office"),
          temperature = c("cold", "hot")
      ),
      mmodel = list(
          none.cold = c("Y11", "Y12", "Y13"),
          low.cold  = c("Y21", "Y22", "Y23"),
          high.cold = c("Y31", "Y32", "Y33"),
          none.hot  = c("Y41", "Y42", "Y43"),
          low.hot   = c("Y51", "Y52", "Y53"),
          high.hot  = c("Y61", "Y62", "Y63")
      ),
  )

> summary(fit)

```

A Larger Example

R package **semnova** implements latent repeated measures ANOVA using the SEM software package **lavaan** (Rosseel, 2012)

```
> library(semnova)

> fit <- semnova(
      data   = data,
      idata  = expand.grid(
          noise       = c("none", "speech", "office"),
          temperature = c("cold", "hot")
      ),
      mmodel = list(
          none.cold = c("Y11", "Y12", "Y13"),
          low.cold  = c("Y21", "Y22", "Y23"),
          high.cold = c("Y31", "Y32", "Y33"),
          none.hot  = c("Y41", "Y42", "Y43"),
          low.hot   = c("Y51", "Y52", "Y53"),
          high.hot  = c("Y61", "Y62", "Y63")
      ),
  )

> summary(fit)

```

A Larger Example

Output:

```
------------------------------------------

term: noise

Response transformation matrix:
            noise1   noise2
none.cold     -0.5  0.28868
speech.cold    0.0 -0.57735
office.cold    0.5  0.28868
none.hot      -0.5  0.28868
speech.hot     0.0 -0.57735
office.hot     0.5  0.28868

multiv. tests:
      Df test stat approx F num Df den Df  Pr(>F)
Wald   2    3.8139   1.9069      2    198 0.15125
Wilks  1    0.9047   5.1618      2     98 0.00739 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

univ. test:
       Sum Sq num Df Error SS den Df F value   Pr(>F)
F-test 3.8361      2   71.606    198  5.3037 0.005705 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

------------------------------------------
```

```
------------------------------------------

term: temperature

Response transformation matrix:
            temperature1
none.cold        -0.40825
speech.cold      -0.40825
office.cold      -0.40825
none.hot          0.40825
speech.hot        0.40825
office.hot        0.40825

multiv. tests:
      Df test stat  approx F num Df den Df Pr(>F)
Wald   1   0.00263 0.0026270      1     99 0.9592
Wilks  1   0.99995 0.0052905      1     99 0.9422

univ. test:
          Sum Sq num Df Error SS den Df F value Pr(>F)
F-test 0.0029077      1   54.412     99  0.0053 0.9422

------------------------------------------
```

A Larger Example

Output:

```
------------------------------------------

term: noise

Response transformation matrix:
             noise1   noise2
none.cold      -0.5  0.28868
speech.cold     0.0 -0.57735
office.cold     0.5  0.28868
none.hot       -0.5  0.28868
speech.hot      0.0 -0.57735
office.hot      0.5  0.28868

multiv. tests:
      Df test stat approx F num Df den Df  Pr(>F)
Wald   2    3.8139    1.9069      2    198 0.15125
Wilks  1    0.9047    5.1618      2     98 0.00739 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

univ. test:
       Sum Sq num Df Error SS den Df F value    Pr(>F)
F-test 3.8361      2   71.606    198  5.3037 0.005705 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

------------------------------------------
```

```
------------------------------------------

term: temperature

Response transformation matrix:
             temperature1
none.cold         -0.40825
speech.cold       -0.40825
office.cold       -0.40825
none.hot           0.40825
speech.hot         0.40825
office.hot         0.40825

multiv. tests:
      Df test stat  approx F num Df den Df Pr(>F)
Wald   1   0.00263 0.0026270      1     99 0.9592
Wilks  1   0.99995 0.0052905      1     99 0.9422

univ. test:
          Sum Sq num Df Error SS den Df F value Pr(>F)
F-test 0.0029077      1   54.412     99  0.0053 0.9422

------------------------------------------
```

A Larger Example

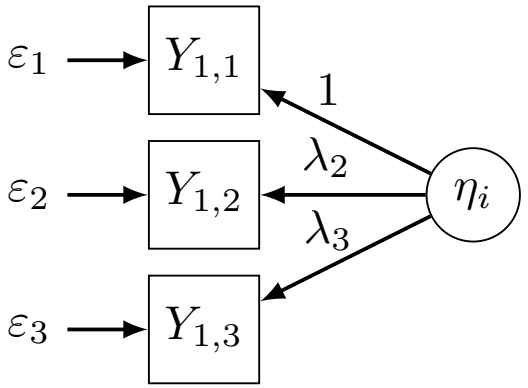
Output:

```
-----------------------------------------

term: noise

Response transformation matrix:
             noise1   noise2
none.cold      -0.5  0.28868
speech.cold     0.0 -0.57735
office.cold     0.5  0.28868
none.hot       -0.5  0.28868
speech.hot      0.0 -0.57735
office.hot      0.5  0.28868

multiv. tests:
      Df test stat approx F num Df den Df  Pr(>F)
Wald   2    3.8139   1.9069      2    198 0.15125
Wilks  1    0.9047   5.1618      2     98 0.00739 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

univ. test:
       Sum Sq num Df Error SS den Df F value    Pr(>F)
F-test 3.8361      2   71.606    198  5.3037 0.005705 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

-----------------------------------------
```

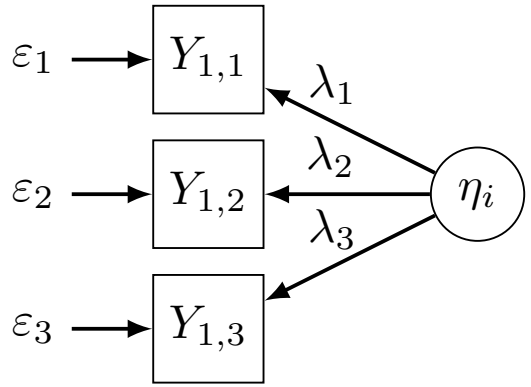
```
-----------------------------------------

term: temperature

Response transformation matrix:
             temperature1
none.cold         -0.40825
speech.cold       -0.40825
office.cold       -0.40825
none.hot           0.40825
speech.hot         0.40825
office.hot         0.40825

multiv. tests:
      Df test stat  approx F num Df den Df Pr(>F)
Wald   1   0.00263 0.0026270      1     99 0.9592
Wilks  1   0.99995 0.0052905      1     99 0.9422

univ. test:
          Sum Sq num Df Error SS den Df F value Pr(>F)
F-test 0.0029077      1   54.412     99  0.0053 0.9422

-----------------------------------------
```

Latent Variable Scaling

Marker-Variable Method



Effect-Coding Method



where

$$\frac{1}{K}\sum_{k=1}^{K}\lambda_k = 1$$

`semnova(..., ind_scaling = "first1", ...)`

- `ind_scaling`: Character String. Default is "`first1`" (marker-variable method). Can also be set to "`average1`" (effect-coding method).

Covariance Structures

Sphericity

$$\boldsymbol{\Sigma} = \begin{array}{l} \\ \text{intercept } (\pi_0) \\ \text{noise1 } (\pi_1) \\ \text{noise2 } (\pi_2) \\ \text{temp1 } (\pi_3) \\ \text{noise1:temp1 } (\pi_4) \\ \text{noise2:temp1 } (\pi_5) \end{array} \begin{array}{c} \begin{array}{cccccc} \text{intercept} & \text{noise1} & \text{noise2} & \text{temp1} & \text{noise1:temp1} & \text{noise2:tmp1} \end{array} \\ \begin{pmatrix} \sigma_{intercept} & & & & & \\ & \sigma_{noise} & 0 & & & \\ & 0 & \sigma_{noise} & & & \\ & & & \sigma_{temp} & & \\ & & & & \sigma_{noise:temp} & 0 \\ & & & & 0 & \sigma_{noise:temp} \end{pmatrix} \end{array}$$

```
semnova(..., sphericity = FALSE, ...)
```

- `sphericity`: Boolean. Default is `FALSE`. Imposes sphericity onto the model.

The semnova Package for Latent Repeated Measures ANOVA — Benedikt Langenberg (langenberg@psych.rwth-aachen.de); Axel Mayer — RWTH Aachen University — 28.02.2020

Compound Symmetry

$$\boldsymbol{\Sigma} = \begin{array}{cc} & \begin{array}{cccccc} \text{none cold} & \text{low cold} & \text{high cold} & \text{none hot} & \text{low hot} & \text{high hot} \end{array} \\ \begin{array}{cc} \text{none} & \text{cold} \\ \text{low} & \text{cold} \\ \text{high} & \text{cold} \\ \text{none} & \text{hot} \\ \text{low} & \text{hot} \\ \text{high} & \text{hot} \end{array} & \begin{pmatrix} \sigma_a & \sigma_b & \sigma_b & \sigma_b & \sigma_b & \sigma_b \\ \sigma_b & \sigma_a & \sigma_b & \sigma_b & \sigma_b & \sigma_b \\ \sigma_b & \sigma_b & \sigma_a & \sigma_b & \sigma_b & \sigma_b \\ \sigma_b & \sigma_b & \sigma_b & \sigma_a & \sigma_b & \sigma_b \\ \sigma_b & \sigma_b & \sigma_b & \sigma_b & \sigma_a & \sigma_b \\ \sigma_b & \sigma_b & \sigma_b & \sigma_b & \sigma_b & \sigma_a \end{pmatrix} \end{array}$$

$$Var(Y_i) = Var(Y_j) \quad \forall i, j$$
$$Cov(Y_i, Y_j) = \rho \quad \forall i, j \text{ where } i \neq j$$

```
semnova(..., compound_symmetry = FALSE, ...)
```

- `compound_symmetry`: Boolean. Default is `FALSE`. Imposes compound symmetry onto the model.

Custom Contrasts / Hypotheses

$$\mathbf{C} = \begin{array}{ll} & \text{dependent variables} \\ \text{Factor} & \begin{array}{cccccc} \text{noise} & \text{none} & \text{low} & \text{high} & \text{none} & \text{low} & \text{high} \\ \text{temp.} & \text{cold} & \text{cold} & \text{cold} & \text{hot} & \text{hot} & \text{hot} \end{array} \\ \begin{array}{l} \text{contrast1} \\ \text{contrast1} \end{array} & \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \end{pmatrix} \end{array}$$

`lgc(...)`

- `C_matrix`: Constrast matrix. If not a square matrix, arbitrary orthogonal rows are added.
- `hypotheses`: List of integers vectors. Each element contains the row indices of the contrast matrix that are to be tested against zero.

Conclusion

Latent repeated measures ANOVA

- Extends the latent growth components approach
- Allows for latent variables in repeated measures analysis
- Allows for multi-factorial designs (of any size)
- Introduces SEM advantages to repeated measures ANOVA

Conclusion

Latent repeated measures ANOVA

- Extends the latent growth components approach
- Allows for latent variables in repeated measures analysis
- Allows for multi-factorial designs (of any size)
- Introduces SEM advantages to repeated measures ANOVA

R package `semnova`

- Implements latent repeated measures ANOVA
- Performs various tests
  - Multivariate: Wald, Wilks, permutation test
  - Univariate: F-test, permutation test
- Imposes different error structures
- Uses different measurement models
- Tests custom contrasts / hypotheses

Conclusion

Next steps:

- Test for sphericity
- Test for measurement invariance
- Implementation of `anova()` function
- Interindividual differences (latent covariates)
- Implementation of between-subject designs within latent repeated measures ANOVA (e.g., EffectLiteR approach, Mayer et al., 2016)
- Small sample sizes (permutation test, Bayesian extension)
- Enhance output / be user-friendly
- Documentation / test cases

Conclusion

Next steps:

- Test for sphericity
- Test for measurement invariance
- Implementation of `anova()` function
- Interindividual differences (latent covariates)
- Implementation of between-subject designs within latent repeated measures ANOVA (e.g., EffectLiteR approach, Mayer et al., 2016)
- Small sample sizes (permutation test, Bayesian extension)
- Enhance output / be user-friendly
- Documentation / test cases

Thank you for your attention!

`https://langenberg.github.io`

References

Mayer, A., Dietzfelbinger, L., Rosseel, Y., and Steyer, R. (2016). The EffectLiteR Approach for Analyzing Average and Conditional Effects. *Multivariate Behavioral Research*, 51(2-3):374–391.

Mayer, A., Steyer, R., and Mueller, H. (2012). A General Approach to Defining Latent Growth Components. *Structural Equation Modeling*, 19(4):513–533.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2):1–20.

Syndicus, M., Wiese, B. S., and van Treeck, C. (2016). In the Heat and Noise of the Moment: Effects on Risky Decision Making. *Environment and Behavior*, 50(1):3–27.