SEM and Longitudinal Data Analysis

PSY9140

Tilmann von Soest

Longitudinal data

Longitudinal study

Panel study

Follow-up study

Prospective study

2 versus more time points

Why longitudinal studies?

1) Examine individual developmental trajectories

2) Provide indications for causal relationships

Conditions for A causing B

1) Relationship: Relationship between A and B

2) Time precedence: A happens before B

3) Nonspuriousness: No C can explain relationship between A and B

Kenny, D. A. (1979). *Correlation and causality*. New York: Wiley.
Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
Kraemer, H. C., Kazdin, A. E., Offord, D. R., Kessler, R. C., Jensen, P. S., & Kupfer, D. J. (1997). Coming to terms with the terms of risk. *Archives of General Psychiatry, 54*, 337-343.

Analysis of Longitudinal Data

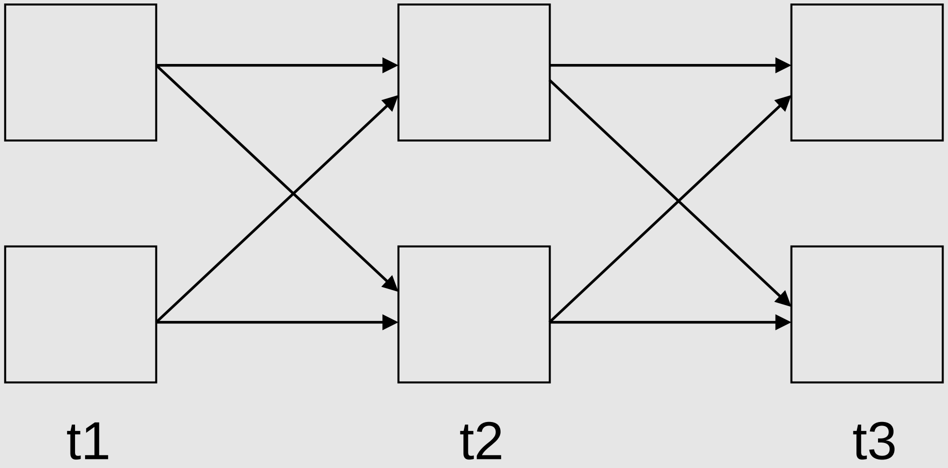
	continuous data	discrete/dichotoumos data
Two or few time points	autoregressive models, linear regression	autoregressive models, logistic regression
Several time points	growth curves	event history models (survival models)
Many time points	time series analysis	

Autoregressive Models

Cross-lagged Models

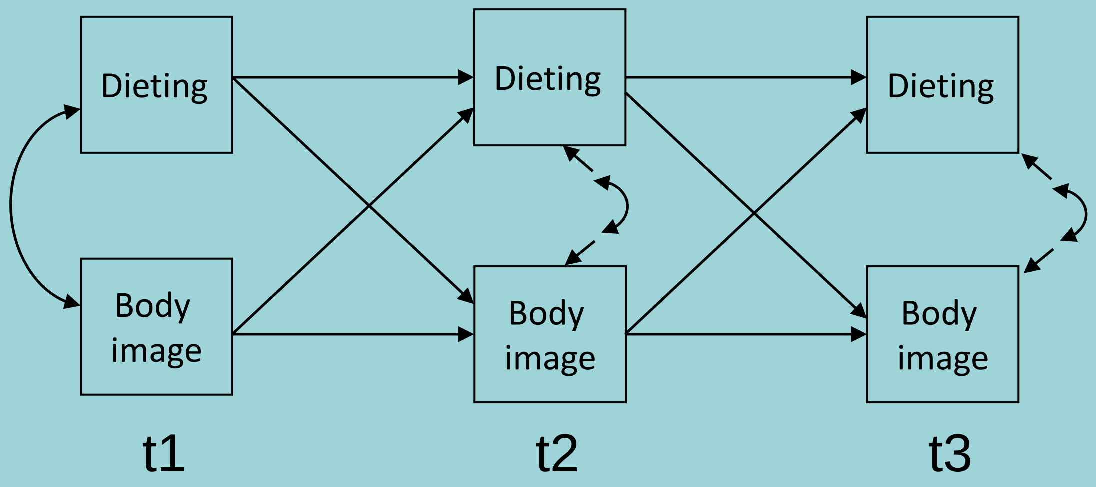
Autoregressive Models

(cross-lagged models)



EXERCISE 1

1) Use the script *day4.R, Exercise 1* and specify the following model (diet1 = dieting at T1, body1 = body image at T1 etc.)

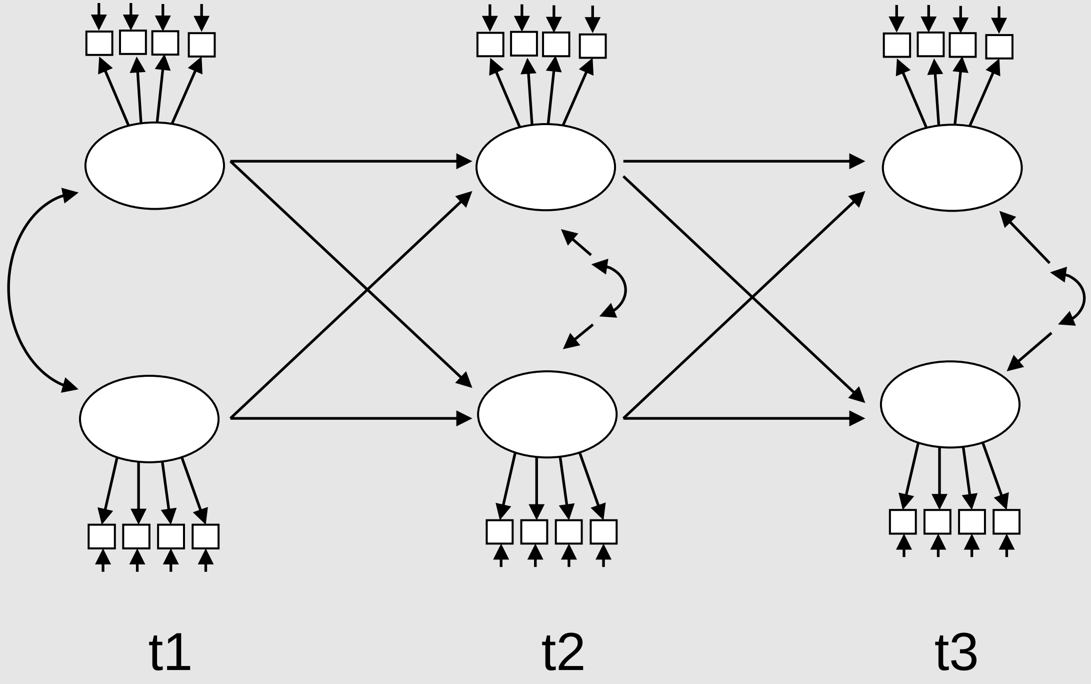


2) Check model fit and re-specify the model, according to modification indices and conceptual considerations.

3) Both girls and boys are included in the sample. Specify a model with adjustment for gender.

Autoregressive Models based on CFA

1) Measurement invariance
2) Correlated residuals



DAY 4, EXERCISE 2

1) Use the input file *day4.R, Exercise 2*. Run the model and draw a path diagram of the model that is estimated. (*body1_1* is the first indicator of body image at T1, *body1_2* the second at T1 etc.)
2) The fit indices show that the model does not fit the data well. How could you improve the model in a way that is conceptually meaningful?
3) Re-specify the model to obtain measurement invariance across time.
4) As the model is specified now, the latent variables are correlated. Re-specify the model to obtain a cross-lagged, autoregressive model. Do the results provide information about the temporal associations between body image and dieting?
5) Include latent factors of the third measurement point in the model.

Measurement Invariance

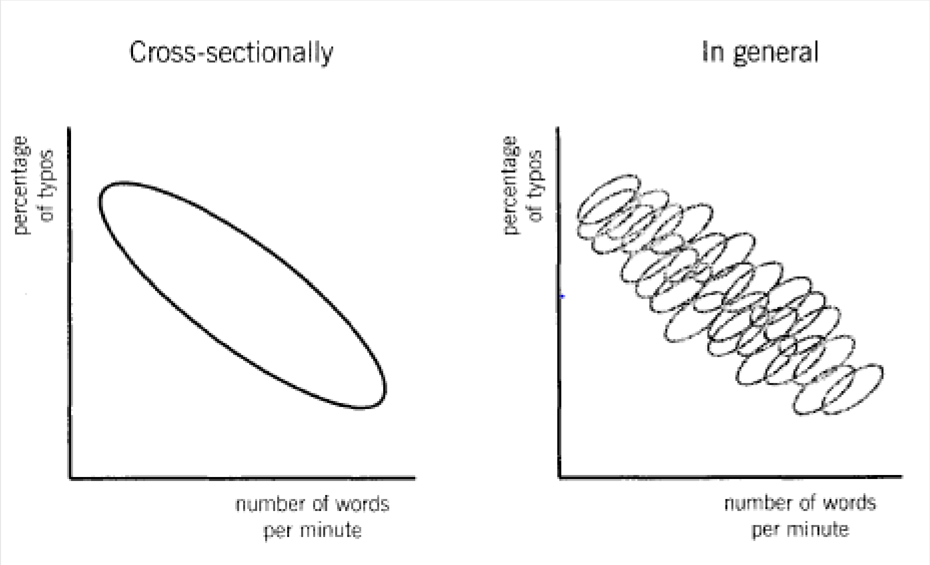
Type of invariance	Alternative term	Same factor model	Invariant factor loadings	Invariant intercepts	Invariant residual variance of indicators
Configural invariance	Configural invariance	X			
Weak factorial invariance	Metric invariance	X	X		
Strong factorial invariance	Scalar invariance	X	X	X	
Strict factorial invariance	Residual invariance	X	X	X	X

Pirralha, A. (2020). Testing for Measurement Invariance with Many Groups. https://bookdown.org/content/5737/

Ferrer, E., Balluerka, N., & Widaman, K. F. (2008). Factorial invariance and the specification of second-order latent growth models. Methodology, 4, 22-36. https://doi.org/10.1027/1614-2241.4.1.22

Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. Child Development Perspectives, 4(1), 10-18. https://doi.org/10.1111/j.1750-8606.2009.00110.x

Criticism of cross-lagged models



- Random Intercepts Cross-Lagged Panel Model (RI-CLPM, Hamaker et al., 2015)
- Autoregressive Latent Trajectory models (ALT, Curran et al., 2014)
- Cross-Lagged Panel Models With Fixed Effects (Allison et al., 2017)

Literature

Allison, P. D., Williams, R., & Moral-Benito, E. (2017). Maximum likelihood for cross-lagged panel models with fixed effects. *Socius, 3*, 1-17. doi:10.1177/2378023117710578

Berry, D., & Willoughby, M. T. (2017). On the practical interpretability of cross-lagged panel models: Rethinking a developmental workhorse. *Child Development, 88*, 1186-1206. doi:10.1111/cdev.12660

Curran, P. J., Howard, A. L., Bainter, S. A., Lane, S. T., & McGinley, J. S. (2014). The separation of between-person and within-person components of individual change over time: A latent curve model with structured residuals. *Journal of Consulting and Clinical Psychology, 82*, 879-894. doi:10.1037/a0035297

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods, 20*, 102-116. doi:10.1037/a0038889

Mulder, J. D., & Hamaker, E. L. (2021). Three extensions of the random intercept cross-lagged panel model. *Structural Equation Modeling*, 28, 638-648. doi:10.1080/10705511.2020.1784738

Usami, S., Murayama, K., & Hamaker, E. L. (2019). A unified framework of longitudinal models to examine reciprocal relations. *Psychological Methods*, 24, 637-657. doi:10.1037/met0000210

Newsom, J. T. (2015). *Longitudinal Structural Equation Modeling*. Routledge.

Growth curves

Random coefficient models

Linear mixed models

Multilevel models

Hierarchical regression models

(Latent) growth models

Latent trajectory models

Literature:

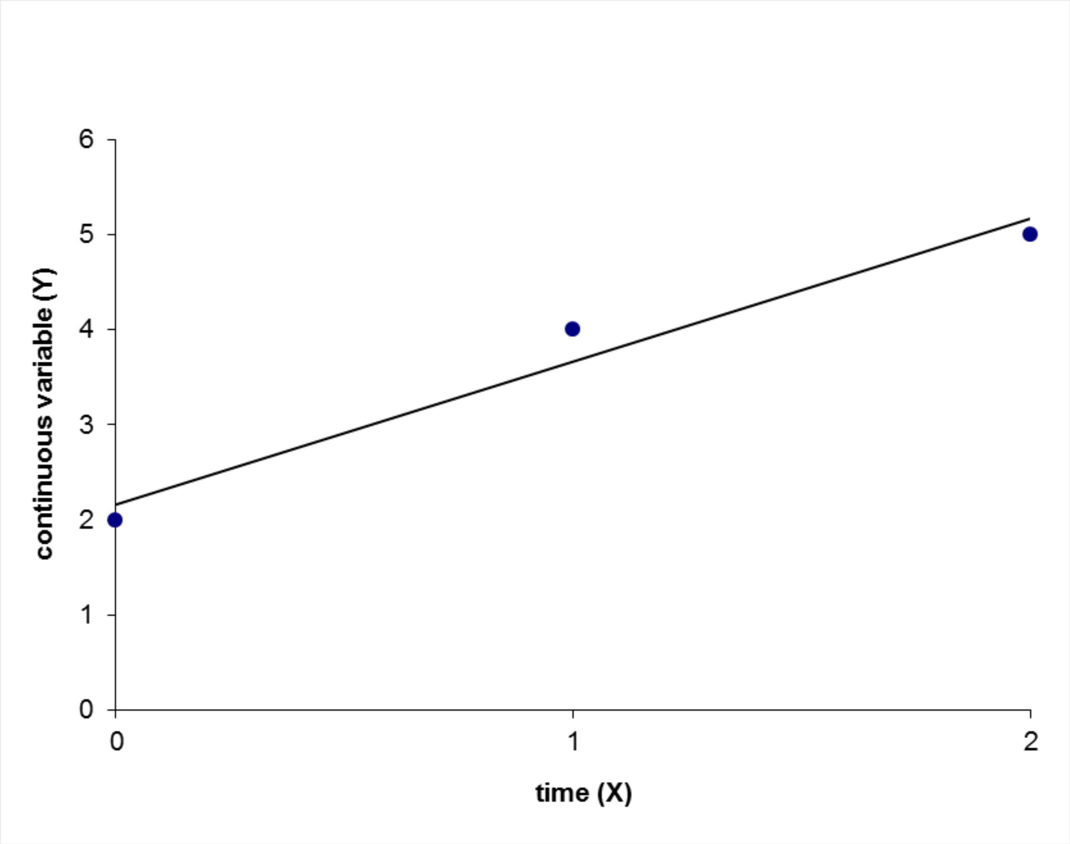
Bollen, K. A., & Curran, P. J. (2006). *Latent curve models. A structural equation perspective*. Hoboken, NJ: Wiley.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis. Modeling change and event occurrence*. Oxford: Oxford University Press.

Newsom, J. T. (2015). *Longitudinal Structural Equation Modeling*. Routledge.

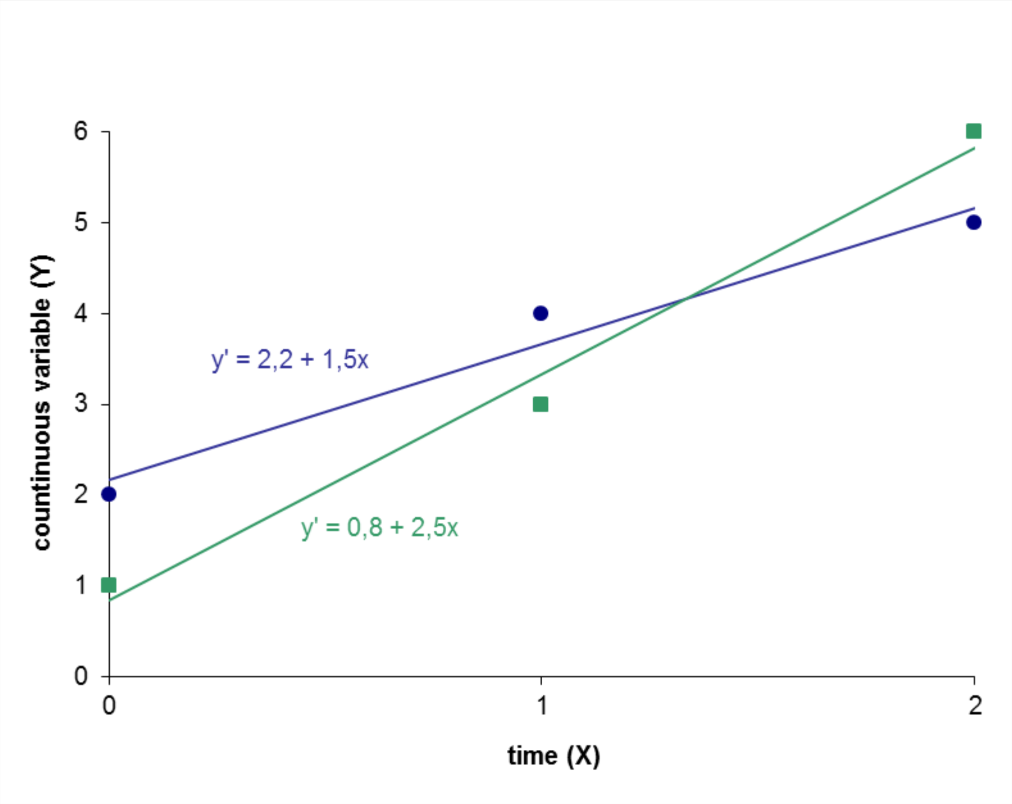
Example:

Individual development of body image from adolescence to adulthood



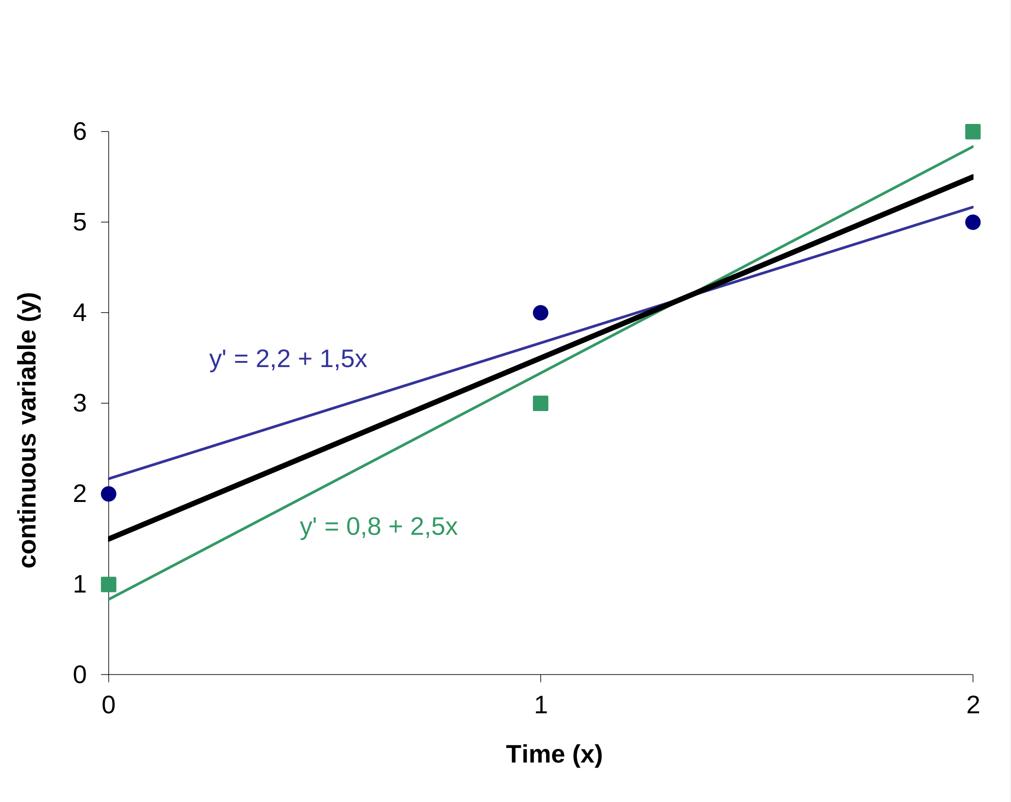
Y' = a + b*TIME

Y' = 2,2 + 1,5*TIME



$Y_i' = a_i + b_i*TIME$ $Y_{it}' = a_i + b_i*TIME$

$Y_i = a_i + b_i*TIME + e_i$ $Y_{it} = a_i + b_i*TIME + e_{it}$

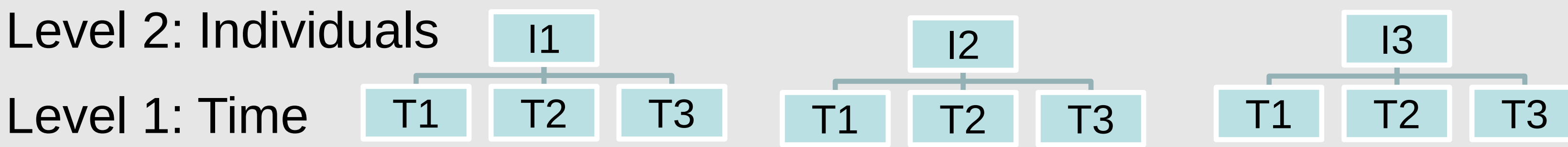


$Y_{it} = a_i + b_i$*TIME $+ e_{it}$ (trajectory equation/ level 1 equation)

$a_i = + \mu_a + e_{ia}$ (intercept equation/ level 2 equation)

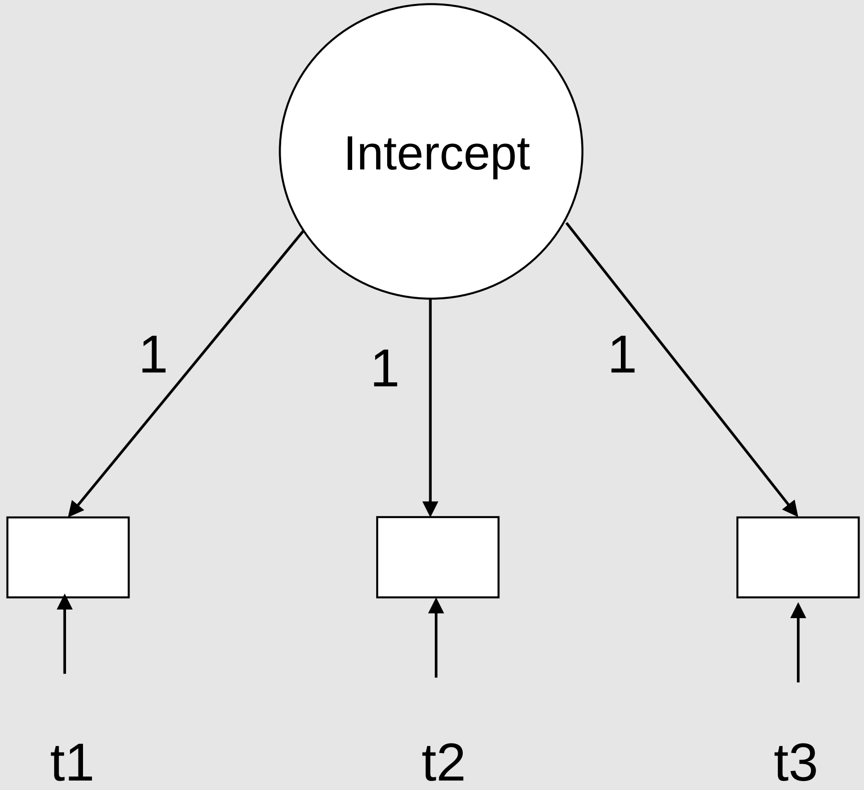
$b_i = + \mu_b + e_{ib}$ (slope equation/ level 2 equation)

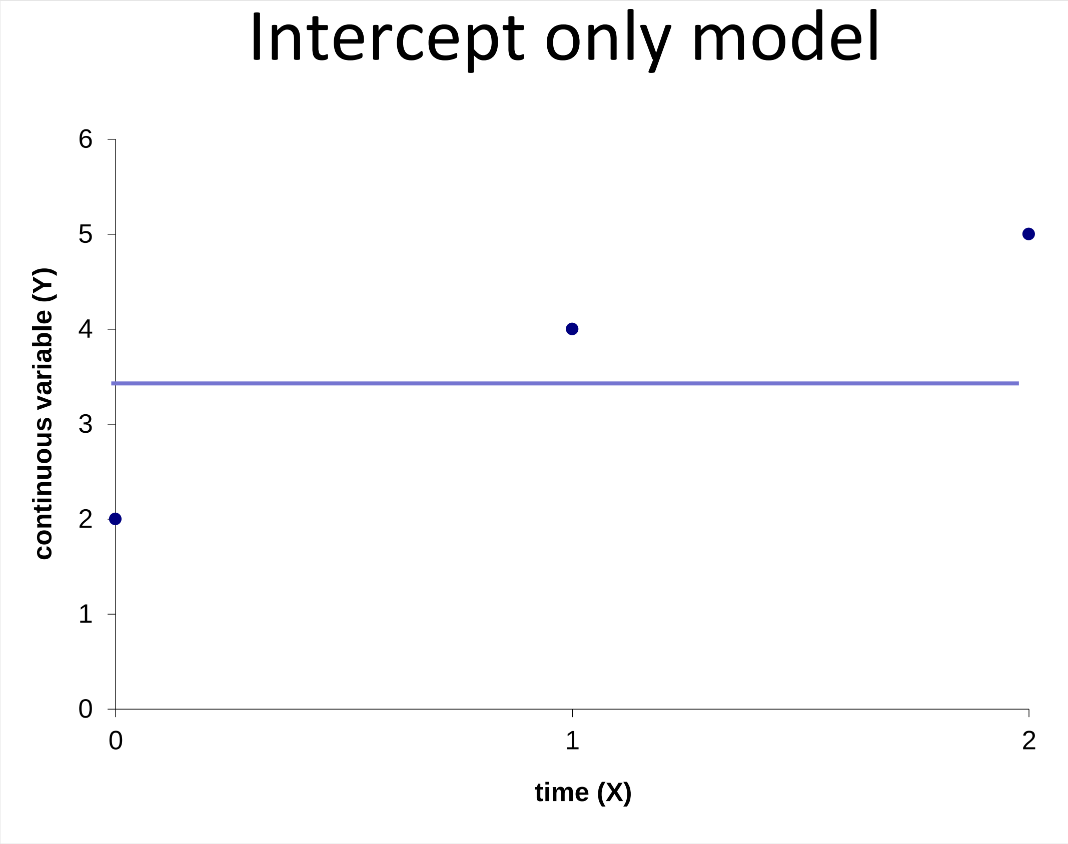
Longitudinal Multilevel Model



How do we specify growth curves in SEM?

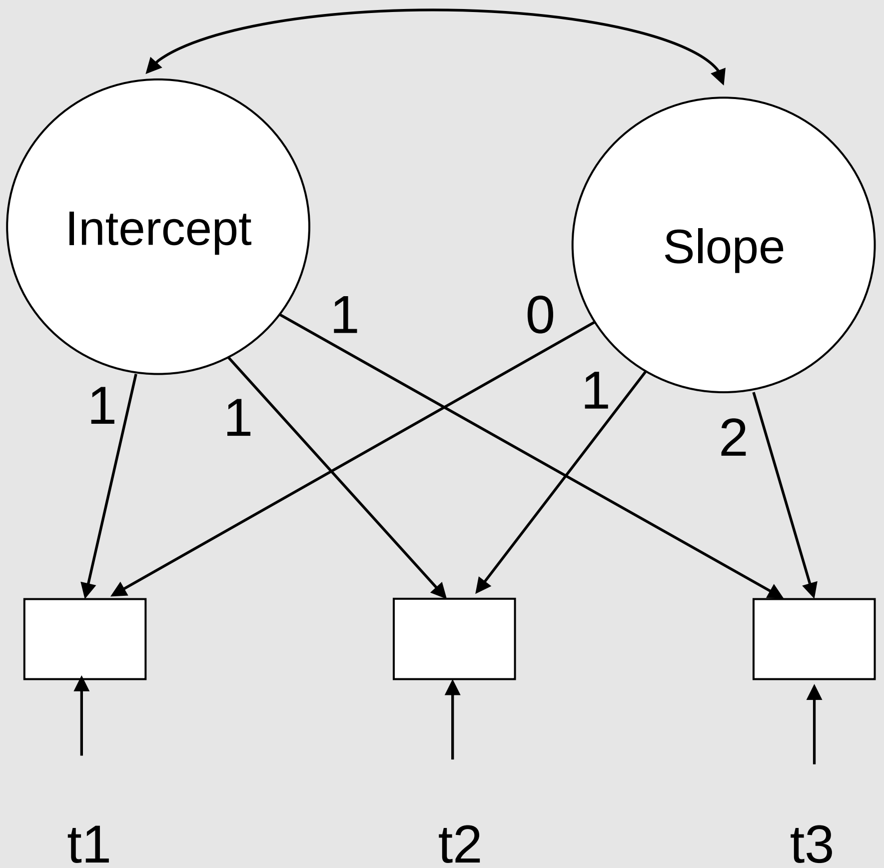
Intercept only model

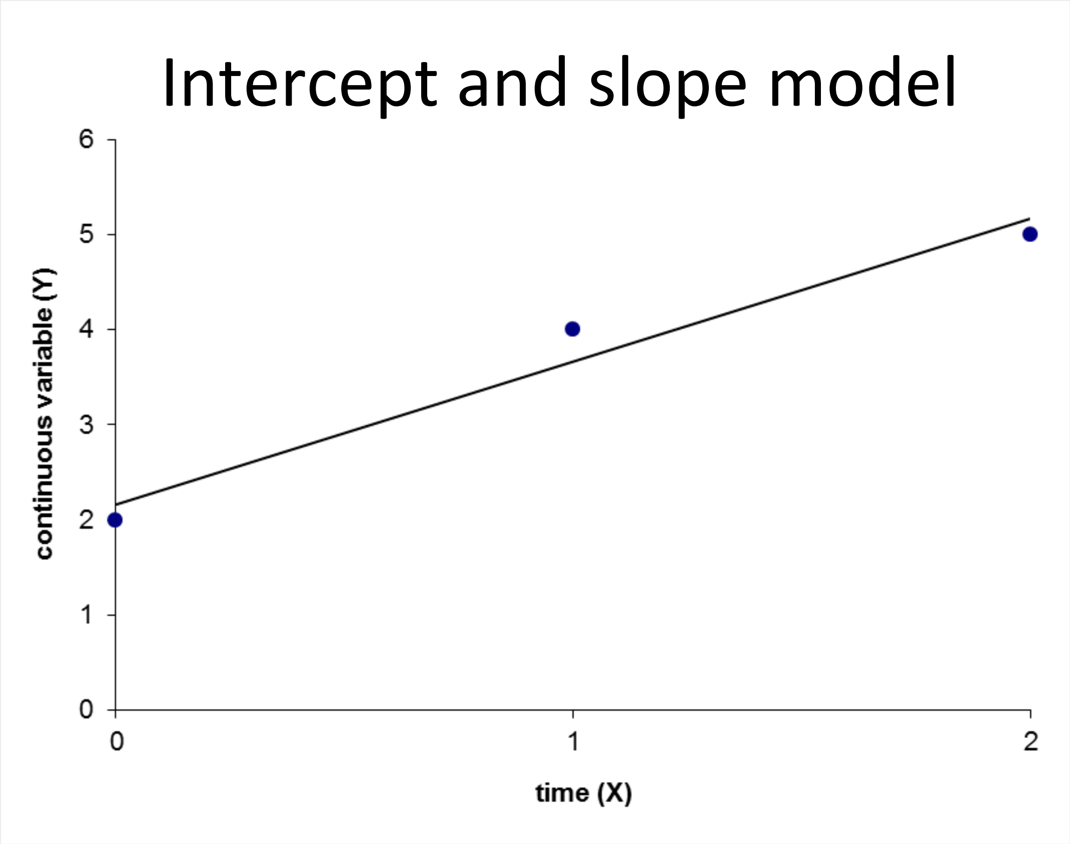



Intercept only model
continuous variable (Y)
time (X)
0
1
2
3
4
5
6
0
1
2

$$Y' = a$$

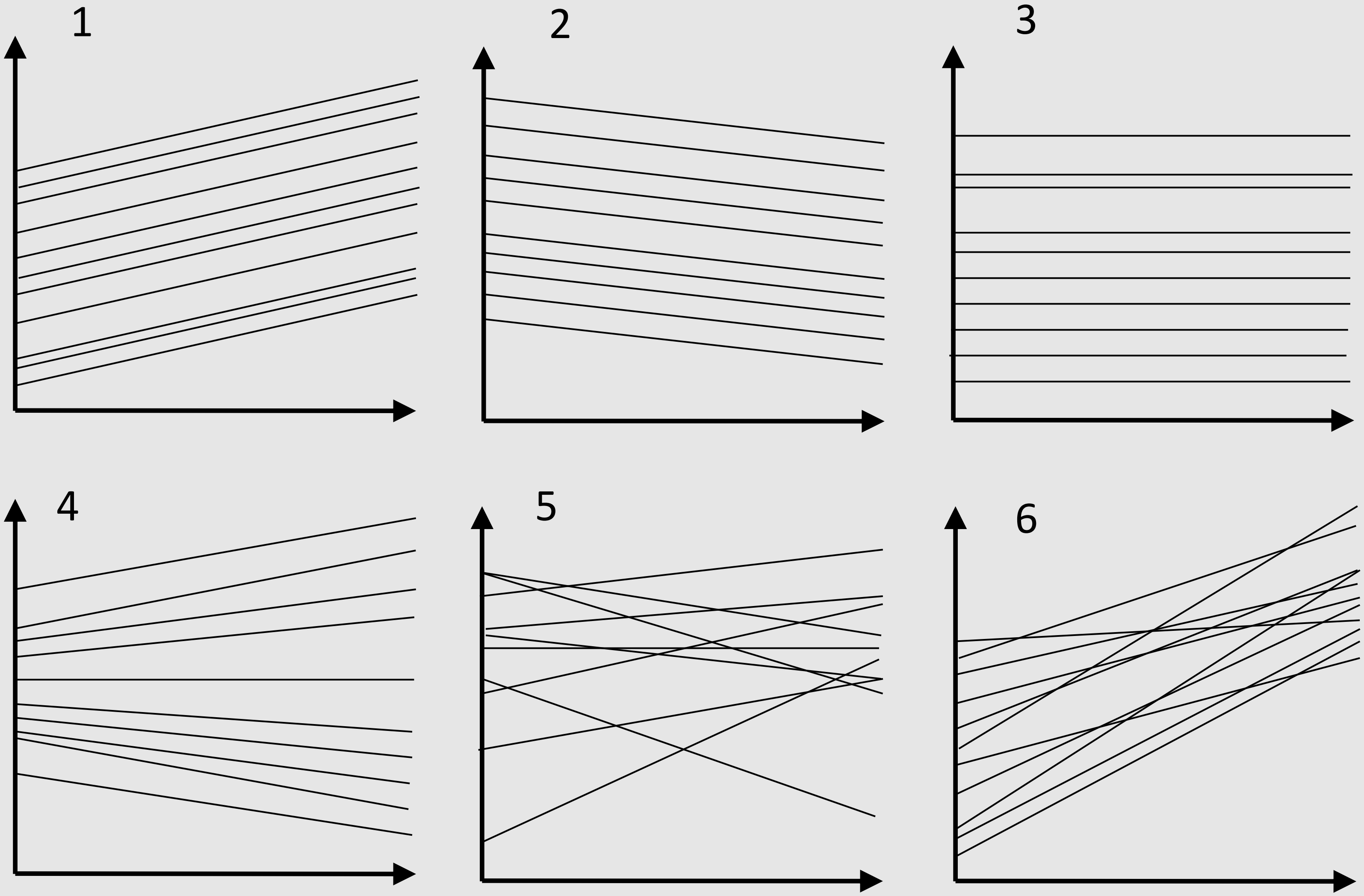
Intercept and slope model



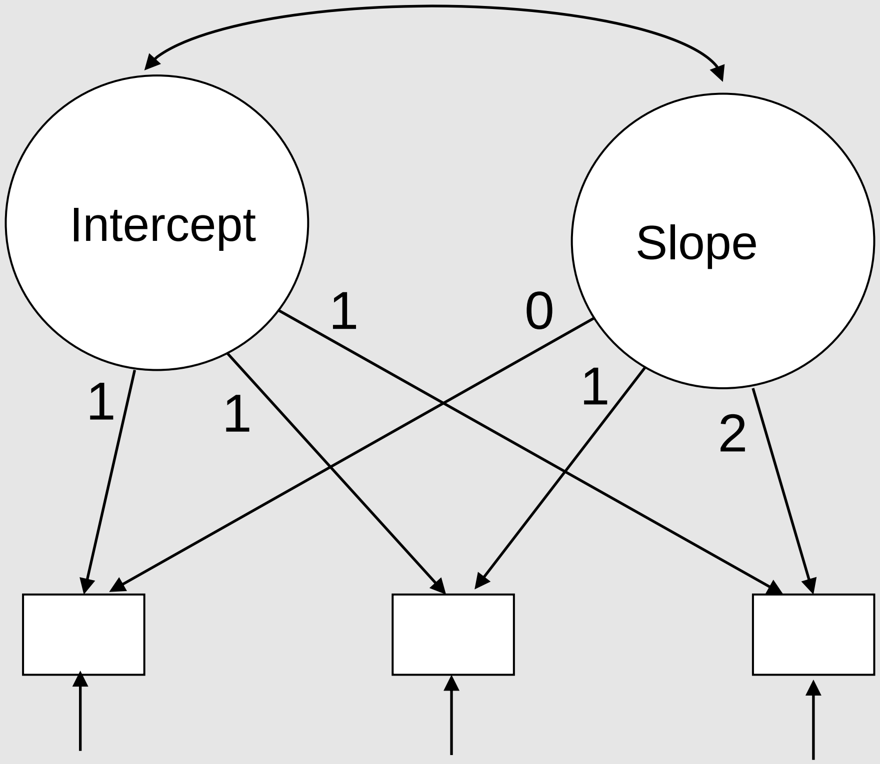

Intercept and slope model
continuous variable (Y)
time (X)
0
1
2
3
4
5
6
0
1
2

$Y' = a + b*TIME$

Intercept and slope mean and variance



Identification



2 variances

2 means

1 correlation

3 error variances

----

8 parameters

lavaan notation for growth curves

```
### CFA NOTATION FOR GROWTH CURVES ###

model <- "# LGC specification
            i =~ 1*body1 + 1*body2 + 1*body3
            s =~ 0*body1 + 1*body2 + 2*body3
          # Covariance between i and s
            i ~~ s
          # fixing intercept of indictors to 0
          # and estimate growth parameter means freely
            body1 ~ 0*1
            body2 ~ 0*1
            body3 ~ 0*1
          # alternative notation: body1 + body2 + body3 ~ 0*1
            i ~ 1
            s ~ 1
          "

fit <- sem(model, missing = "ML", data = df)



### SIMIPLIFIED NOTATION FOR GROWTH CURVES ###

model <- "# LGC specification
            i =~ 1*body1 + 1*body2 + 1*body3
            s =~ 0*body1 + 1*body2 + 2*body3
          # Covariance between i and s
            i ~~ s
        "

# Running the model
fit <- growth(model, missing = "ML", data = df)
```

EXERCISE 3

1) Use the script *day4.R, Exercise 3*. Estimate a growth curve based on the mean scores of body image (variables *body1* to *body3*) at three time points.
2) What kind of information do the results provide about the mean and variance of the intercept and the slope of body image?
3) The second data collection was two years after the first data collection. The third data collection was seven years after the first data collection. Estimate a new growth curve model that takes into account the uneven intervals between time points. How can we interpret the mean of the slope?

Correlation between intercept and slope

- correlation is typically negative
  - regression to mean
  - «fact of life» (Rogosa et al., 1982)
- positive correlation: Matthew effect
  - "the rich get richer and the poor get poorer”
  - Matthew’s gospel

Rogosa, D. R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin, 92*, 726-748.

Predicting the intercept and slope

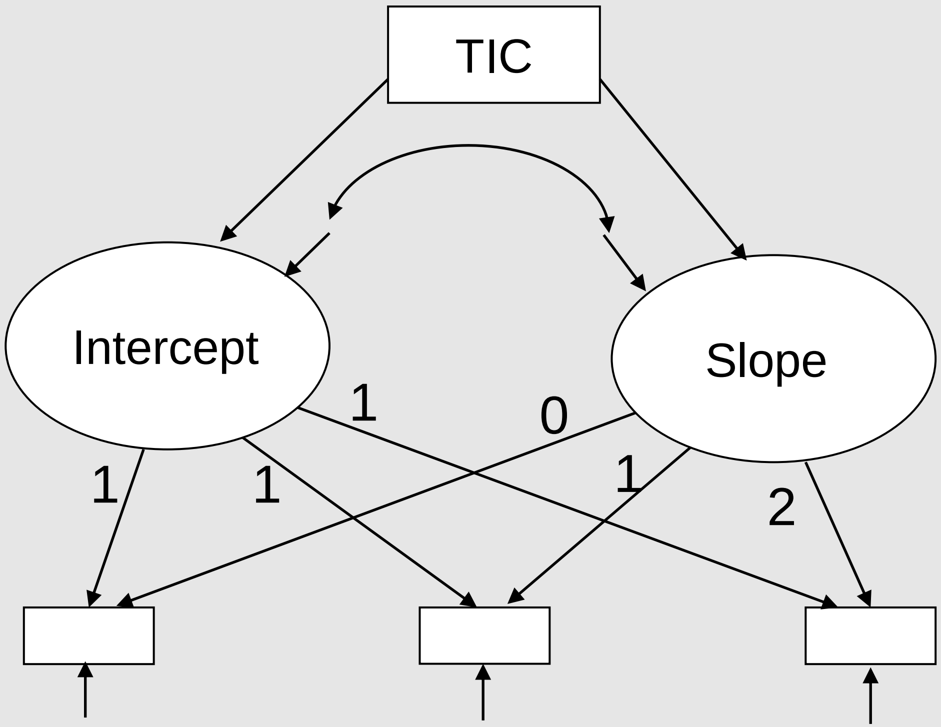
Growth curve with covariate (time invariant)

$Y_{it} = a_i + b_i$*TIME $+ e_{it}$ (level 1)

$a_i = + \mu_a + \sigma_a$*COVARIATE $+ e_{ia}$ (intercept)

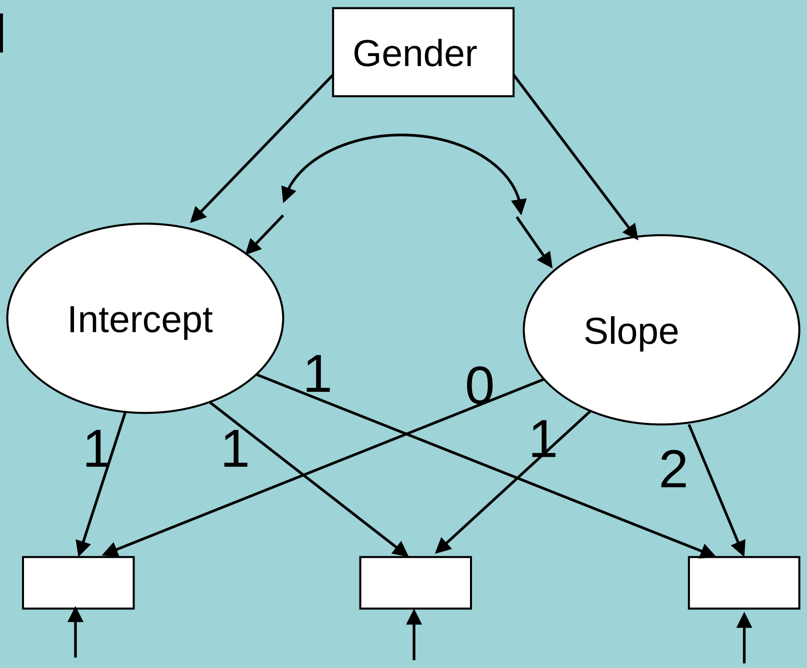
$b_i = + \mu_b + \sigma_b$*COVARIATE $+ e_{ib}$ (slope)

Growth curve with covariate (time invariant)



EXERCISE 4

1) Run the code in *day4.R, Exercise 4*.
2) How is gender related to the intercept and slope of body image, and what do such relationships tell us about how girls' and boys' body image is developing over time?
3) Model a growth curve for dieting and interpret the estimates of the slope and the intercept.
4) Include gender as time invariant covariate, predicting the intercept and the slope of dieting. Is there any problem with this model?



```
Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  i =~
    diet1             1.000                                0.623    0.886
    diet2             1.000                                0.623    0.891
    diet3             1.000                                0.623    0.896
  s =~
    diet1             0.000                                0.000    0.000
    diet2             2.000                                0.136    0.195
    diet3             7.000                                0.476    0.685

Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  i ~
    girl              0.521    0.026   19.688    0.000    0.836    0.412
  s ~
    girl              0.011    0.004    2.716    0.007    0.162    0.080

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
 .i ~~
   .s                -0.020    0.002  -11.542    0.000   -0.530   -0.530

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
   .diet1             0.000                                0.000    0.000
   .diet2             0.000                                0.000    0.000
   .diet3             0.000                                0.000    0.000
   .i                 1.495    0.020   74.210    0.000    2.400    2.400
   .s                 0.013    0.003    4.094    0.000    0.186    0.186

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
   .diet1             0.106    0.009   11.437    0.000    0.106    0.214
   .diet2             0.159    0.007   22.420    0.000    0.159    0.324
   .diet3             0.134    0.022    6.140    0.000    0.134    0.278
   .i                 0.322    0.013   25.493    0.000    0.830    0.830
   .s                 0.005    0.001    7.256    0.000    0.994    0.994
```

Intercept and slope equations

$a_i = 1.50 + 0.52*GIRL + e_{ia}$ $(\sigma_a \neq 0^{**})$

$b_i = 0.01 + 0.01*GIRLS + e_{ib}$ $(\sigma_b \neq 0^{**})$

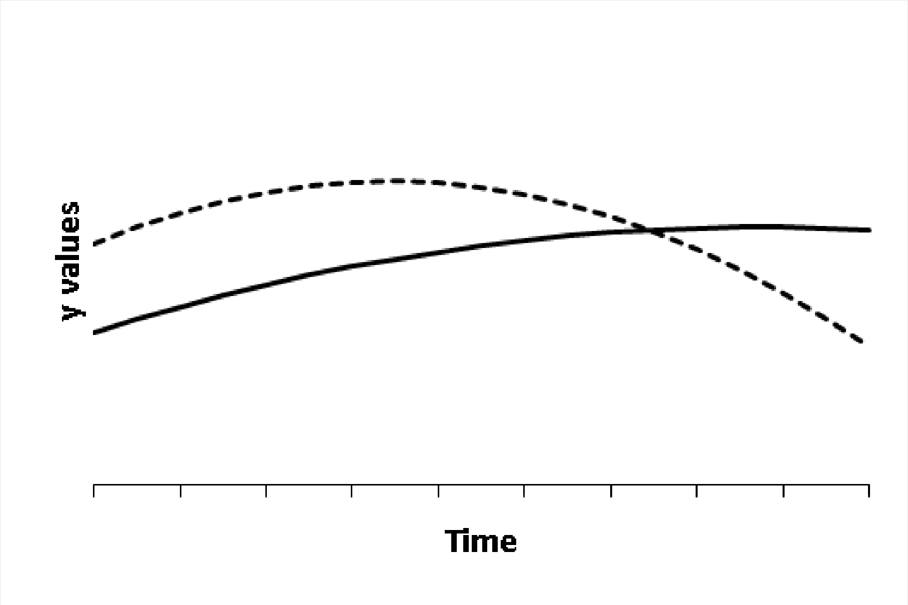
→ Girls have a higher intercept and a higher slope

Parallel growth curves

EXERCISE 5

1) Run the code in *day4.R Exercise 5*. Two growth curves are estimated in one model: One for dieting and one for body image. Examine how intercepts and slopes are correlated with each other.
2) Re-specify the model such that the intercept of dieting predicts the slope of body image. The intercept of body image should predict the slope of dieting.
3) Draw a diagram of the model you have just estimated.

Non-linear growth curves



$$Y' = a + b*TIME + c*TIME^2$$

EXERCISE 6

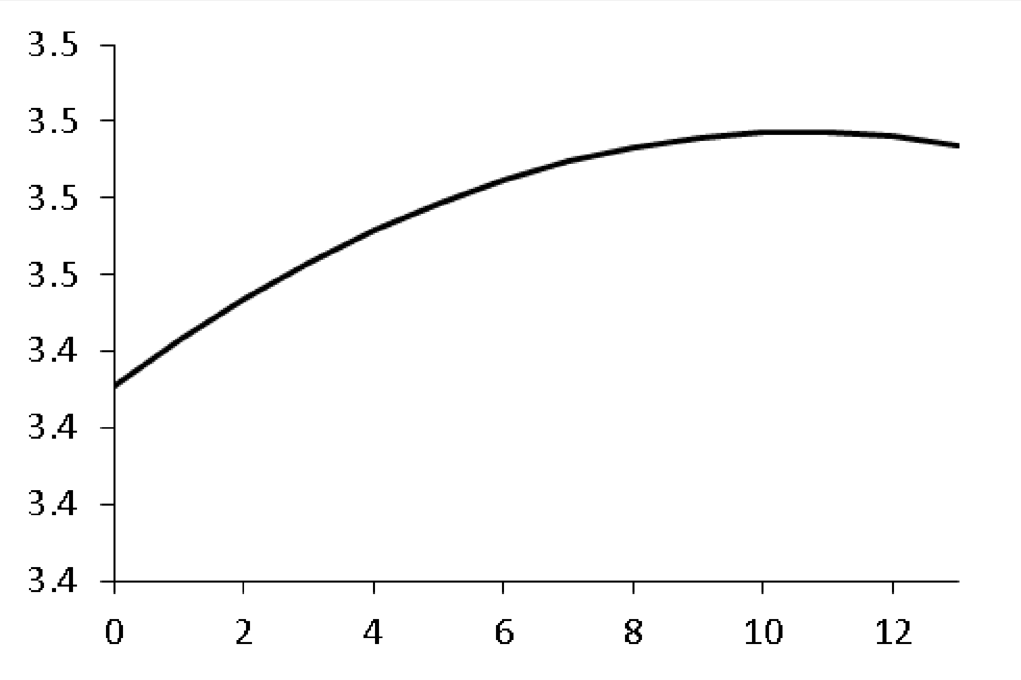
1) Run the code in *day4*.R, Exercise 6 and interpret the means of the intercept and slope.
2) Include a quadratic slope in the model by changing the growth curve command to

   *i =~ 1*body1 + 1*body2 + 1*body3 + 1*body4*

   *s =~ 0*body1 + .2*body2 + .7*body3 + 1.3*body4*

   *q =~ 0*body1 + .04*body2 + .49*body3 + 1.69*body4*
3) Try to draw a graph representing the development of body image according to the new, non-linear trajectory.

Graphical representation of growth curve



$$Y' = 3.43 + 0.13*TIME - 0.06*TIME^2$$

Literature

- Model identification

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models. A structural equation perspective*. Hoboken, NJ: Wiley.

- Non-linear growth curves

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models. A structural equation perspective*. Hoboken, NJ: Wiley.

- Missing data

Allison, P. D. (2003). Missing data techniques for structural equation modeling. *Journal of Abnormal Psychology, 112*, 545-557.

Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*, 147-177.

- Reliability of change scores

Lund, T. (2001). *Måling av forandring. En innføring*. Oslo: Unipub.

Rogosa, D. R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin, 92*, 726-748.

- Correlation between intercept og slope

Rogosa, D. R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin, 92*, 726-748.