KU LEUVEN

Workshop on Sample Size Planning for Intensive Longitudinal Studies

Ginette Lafit, Jordan Revol, Mihai A. Constantin, & Eva Ceulemans

June 8th, 2023

Society for Ambulatory Assessment Conference 2023

About us

Ginette Lafit (ginette.lafit@kuleuven.be)

- PostDoc at Research Group of Quantitative Psychology and Individual Differences and Center of Contextual Psychiatry, KU Leuven

Jordan Revol (jordan.revol@kuleuven.be)

- Ph.D. at Research Group of Quantitative Psychology and Individual Differences, KU Leuven

Mihai A. Constantin (mihai@mihaiconstantin.com)

- Ph.D. at Department of Methodology and Statistics, Tilburg University

Eva Ceulemans (eva.ceulemans@kuleuven.be)

- Professor at Research Group of Quantitative Psychology and Individual Differences, KU Leuven

Preliminaries

Goal of the workshop

The workshop provides a 'road map' on how to determine the sample size in intensive longitudinal (IL) designs

Materials

Slides of the workshop and materials are available at samplesize.help

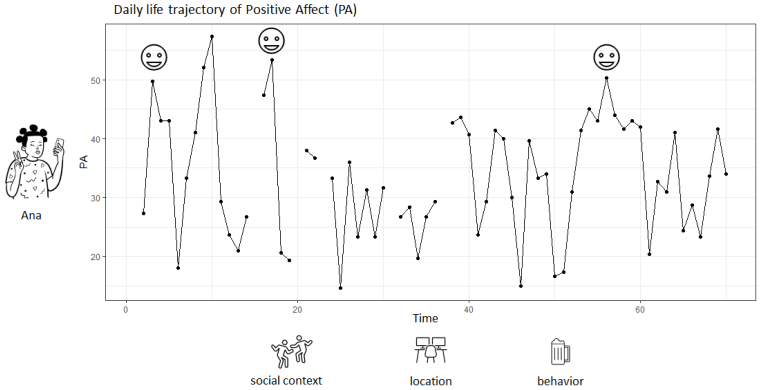
Feel free to ask questions anytime!

Overview

1. Intensive Longitudinal (IL) designs
2. Sample size planning in IL research
3. Research questions in IL research
4. Sample size planning for VAR(1) models in $N = 1$ designs
5. Sample size planning for multilevel models
6. Advanced methods for sample size planning

Investigate dynamics of daily life psychological processes

How do complex psychological processes evolve dynamically within a person's daily life?



Intensive longitudinal designs

- Intensive longitudinal (IL) designs: persons are repeatedly measured over time
- Methods to gather IL data: Experience Sampling Method (ESM)

Figure: Overview ESM. Taken from Olivia J. Kirtley

Structure of IL data

- Sample size: N is the number of persons and T is the number of repeated measurements
- **Time-varying variables**: repeated measurements within persons
- **Time-invariant variables**: person-level variables

example IL study: Leuven clinical study was conducted on persons diagnosed with depression

ID	Day	Beep	PA	NA	Anhedonia	Diagnosis	Depression
1	1	1	NA	NA	NA	1	12
1	1	2	27.33	30.40	26.00	1	12
1	1	3	49.67	23.80	25.00	1	12
2	1	1	73.67	11.00	20.00	0	4
2	1	2	64.33	10.80	18.00	0	4
2	1	3	69.67	11.20	10.00	0	4

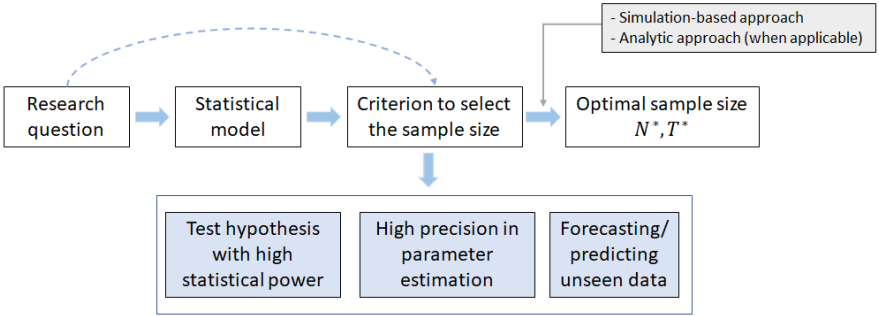
The data set is available at: https://emotedatabase.com/datasets/3

How we can determine the sample size in an IL study?

- The sample size of a study determines how much information is present in a data set to derive reliable conclusions
- If the goal of a study is to test a hypothesis: a criterion to select the sample size is statistical power
- However, other criteria to select the sample size of a study (e.g., precision in parameter estimation, predictive accuracy)

Sample size planning

Selecting the sample size involves the following steps:



Example. Sample size planning in $N = 1$ designs

Goal: investigate within-person dynamics of a variable for a single person

A key measure of emotional dynamics: emotional inertia

Emotional inertia refers to the degree to which emotional states are resistant to change (Kuppens et al., 2010)

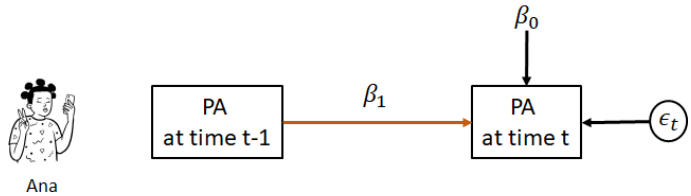
Operationalization of emotional inertia

Within-person emotion autoregressive effect: current values of the emotional variable are predicted by the value of the emotional variable at the previous time point

Estimating emotional inertia for positive affect

Emotional inertia of positive affect (PA)

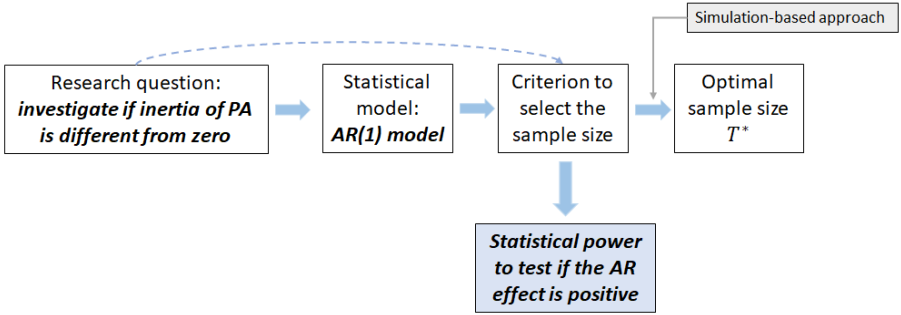
Within-person autocorrelation: current values of PA are predicted by the value of PA at the previous time point



where β_1 denotes the autoregressive effect (i.e., inertia)

Example. Sample size planning to investigate if the emotional of PA is positive

Selecting the sample size for the $N = 1$ study involves the following steps:



Power-based sample size planning

Power-based sample size planning

The goal is to design a study to yield sufficient statistical power to test specific hypotheses concerning parameters in the statistical model

Power analysis

Goal: select the sample size to reach high statistical power (e.g., 90%)

- Power is the probability of correctly rejecting the null hypothesis when there is an effect of a certain size
- Example: in a $N = 1$ design the goal is to test if the autoregressive effect of PA is different from zero

$$H_0 : \beta_1 = 0$$
$$H_1 : \beta_1 \neq 0$$

Null Hypothesis and Outcome of a Test

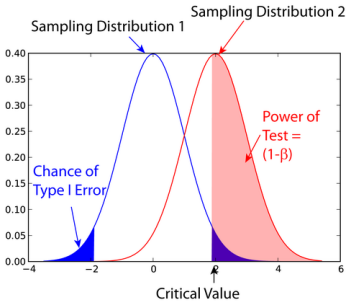
	H_0 is true	H_1 is true
H_0 is rejected	α=Pr(Type I error)	$1 - \beta$=power
H_0 is not rejected		β=Pr(Type II error)

Factors influencing statistical power

- Test statistics: t-test defined as $T = \frac{\hat{\beta}_1}{\text{SE}(\hat{\beta}_1)}$
- We reject H_0 if $|T| \geq \mathcal{T}^{-1}(1-\alpha)$, where $\mathcal{T}^{-1}(1-\alpha)$ is the critical value

Factors influencing statistical power

- size of the true effect size (β_1^*) is positively related to power
- Type I error rate (α) is inversely related to power
- sample size (T) is positively related to power: higher $T \rightarrow$ lower standard error $\text{SE}(\hat{\beta}_1)$

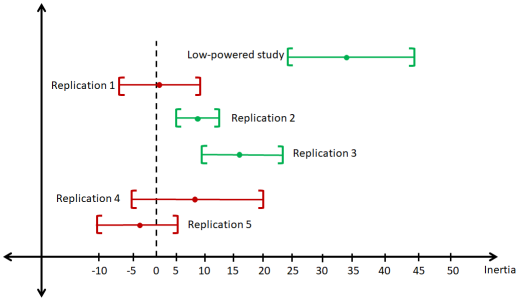


Power of a test. (2022, October 18). In Wikipedia. https://en.wikipedia.org/wiki/Power_of_a_test

The risk of low-powered studies

Common consequences of studies with low power

- The research findings of a low-powered study can differ considerably from the findings in the subsequent studies, in other words, they might not be replicated
- Low-powered studies also increase the chance of overestimation of the true effect



Methods for conducting power analysis

Analytic approach

Computes statistical power using formulas for the standard errors of the effect of interest. For example, G*Power uses analytic approximations to calculate statistical power.

$\rightarrow$ obtaining these formulas is not straightforward for complex models

Simulation-based approach

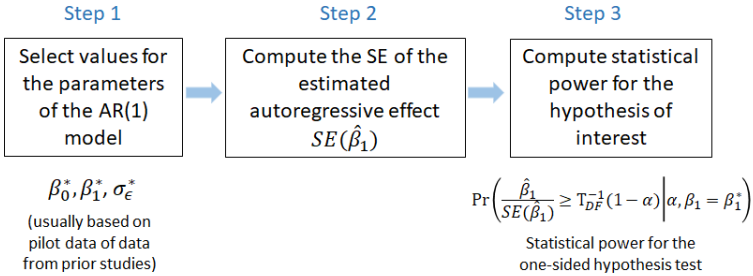
Computes statistical power using Monte Carlo simulations to generate data based on a statistical model and predefined values for the model parameters

$\rightarrow$ it can be applied to a wide variety of models

Steps of the analytic approach

- Example: select the number of measurement occasions T to test if the autoregressive effect of PA is positive

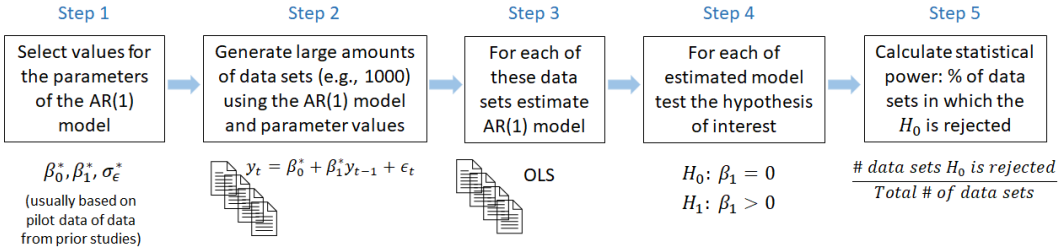
Given T, Hypothesis of interest (e.g., $H_0 : \beta_1 = 0$ vs. $H_1 : \beta_1 > 0$), and α



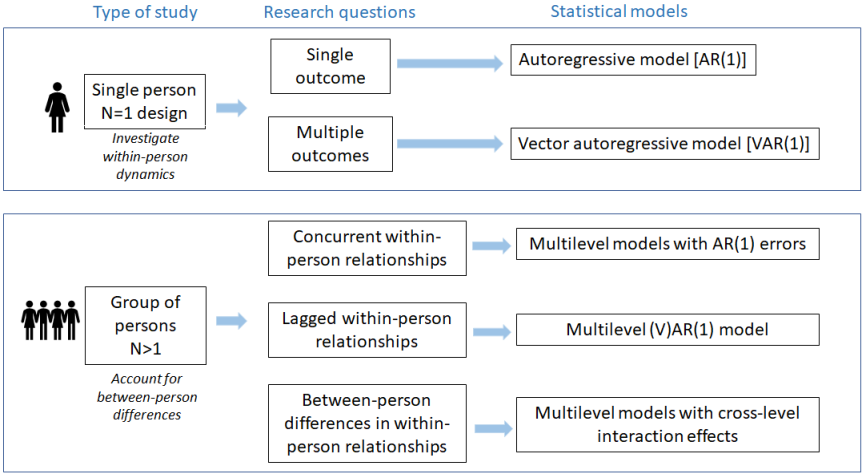
Steps of the simulation-based approach

- Example: select the number of measurement occasions T to test if the autoregressive effect of PA is positive

Given T, Hypothesis of interest (e.g., $H_0 : \beta_1 = 0$ vs. $H_1 : \beta_1 > 0$), and α



What type of research questions we can investigate with IL designs?



In this workshop we will focus on sample size planning for the following research questions

Single person $N = 1$ designs

- AR(1) model
- VAR(1) model

Group of persons $N > 1$ designs

- Multilevel models to investigate concurrent associations (i.e., at the same time point)
- Multilevel AR(1) models

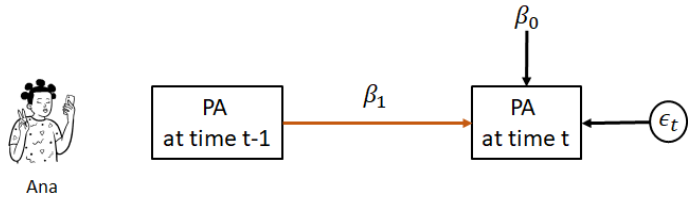
Sample size planning for VAR(1) models in $N = 1$ designs

Goal: select the number of repeated measurements T for VAR(1) models

- Power analysis in AR(1) and VAR(1) models
- Predictive accuracy analysis: a new criterion for selecting T in VAR(1) models when the goal of a study is to predict unseen data

Estimating inertia of PA: Autoregressive model

Autoregressive or AR(1) model for PA



AR(1) model as a linear model:

$$\text{PA}_t = \beta_0 + \beta_1 \text{PA}_{t-1} + \epsilon_t$$

where β_0 is the intercept, β_1 is the autoregressive effect, and ϵ_i denotes the within-person errors which are independent and identically distributed $N(0, \sigma_\epsilon^2)$

Example: AR(1) model estimation

AR(1) model can be estimated using ordinary least squares (OLS)

OLS estimation using R

```
fit.AR.PA = lm(PA ~ 1 + PA.lag, data = data)
summary(fit.AR.PA)
```

Estimate the standard deviation of the errors (σ_ϵ)

```
sd(residuals(fit.AR.PA))
[1] 9.566865
```

Estimate of the autoregressive effect: $\hat{\beta}_1 = 0.41$
Estimate of the standard deviation of the errors: $\hat{\sigma}_\epsilon = 9.57$

Observations	54 (16 missing obs. deleted)
Dependent variable	PA
Type	OLS linear regression

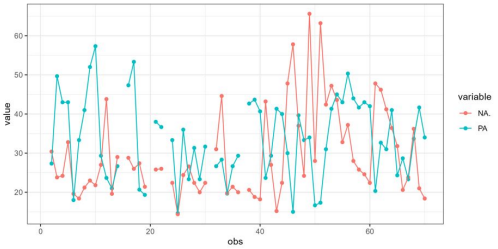
F(1,52)	9.80
R^2	0.16
Adj. R^2	0.14

	Est.	S.E.	t val.	p
(Intercept)	20.33	4.58	4.44	0.00
PA.lag	0.41	0.13	3.13	0.00

Standard errors: OLS

What if we are interested in the within-person dynamics of two variables?

Returning to Ana's example, how does her affect system (PA and NA) evolve in daily life?

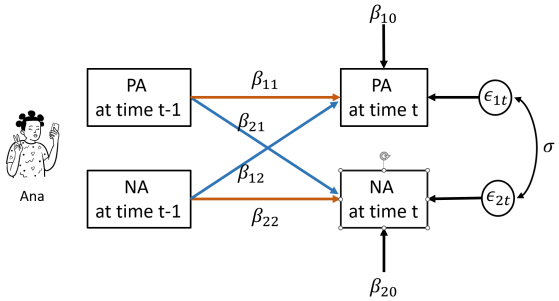


Vector autoregressive model [VAR(1)]: investigate temporal associations between variables

- Effect of a variable at time $t-1$ on the same variable at time t
- Effect of a variable at time $t-1$ on the other variable at time t

VAR(1) models for PA and NA

Goal: investigate within-person dynamics of a system of two variables



where β_{11} and β_{22} are the auto-regressive effects, β_{12} and β_{21} are the cross-regressive effects and ε are the error term that follows a multivariate normal distribution $N(0, \Sigma_\epsilon^2)$

VAR(1) models as a linear model

$$\mathsf{PA}_t = \beta_{10} + \beta_{11}\mathsf{PA}_{t-1} + \beta_{12}\mathsf{NA}_{t-1} + \epsilon_{1t}$$

$$\mathsf{NA}_t = \beta_{20} + \beta_{22}\mathsf{NA}_{t-1} + \beta_{21}\mathsf{PA}_{t-1} + \epsilon_{2t}$$

With:

$$\epsilon \sim \quad N\left[\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{12} & \sigma_{22}\end{bmatrix}\right]$$

Example: VAR(1) model estimation

VAR(1) model can be estimated using ordinary least squares (OLS)

OLS estimation using R

```
fit.PA = lm(PA ~ 1 + PA.lag + NA.lag, data = data)
summary(fit.PA)
fit.NA = lm(NA. ~ 1 + PA.lag + NA.lag, data = data)
summary(fit.NA)
```

Estimation variance-covariance matrix

```
res = cbind(residuals(fit.PA),residuals(fit.NA))
cov(res)
```

PA outcome:

	Est.	S.E.	t val.	p
(Intercept)	23.46	6.28	3.74	0.00
PA.lag	0.39	0.13	2.94	0.00
NA.lag	-0.08	0.11	-0.73	0.47

Standard errors: OLS

NA outcome:

	Est.	S.E.	t val.	p
(Intercept)	17.88	6.63	2.70	0.01
PA.lag	-0.02	0.14	-0.14	0.89
NA.lag	0.38	0.12	3.14	0.00

Standard errors: OLS

Variance-covariance:

```
##             [,1]       [,2]
## [1,] 90.572865   4.295723
## [2,]  4.295723 101.177796
```

Power analysis VAR(1) models

In a new study of a person with similar characteristics to Ana, we want to investigate if:

- The auto-regressive effect for NA is different from zero ($H_0 : \beta_{22} = 0$ vs. $H_1 : \beta_{22} \neq 0$)
- A negative effect of NA on PA ($H_0 : \beta_{21} = 0$ vs. $H_1 : \beta_{21} \neq 0$)

Using Ana's data, we set the values of the model parameters of the VAR(1) model:

$$PA_t = 23.46 + .39 * PA_{t-1} + (-.08) * NA_{t-1} + \varepsilon_{1t}$$

$$NA_t = 17.88 + .38 * NA_{t-1} + (-.02) * PA_{t-1} + \varepsilon_{2t}$$

$$\epsilon \sim \quad N\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 90.6 & 4.30 \\ 4.30 & 101.2 \end{bmatrix} \right]$$

Power for VAR(1)

Goal: select the number of repeated measurements (T) to reach high statistical power (e.g., 90%) with a simulation-based approach

Solution1: R script (see power_analysis_var1.Rmd)

Solution2: A shiny app (link)

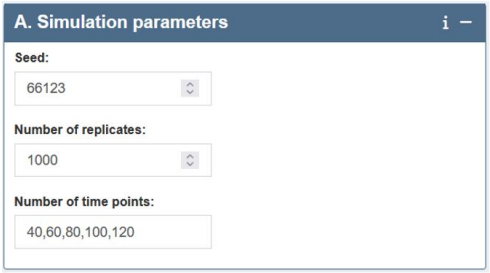
Script to run the shiny app

```
remotes::install_gitlab("ppw-okpiv/researchers/u0148925/shinyapp-paa_var_n1",
host="https://gitlab.kuleuven.be", force=TRUE)
library(paavar1)
run_paa_var1()
```

Power for VAR(1): using the shiny app

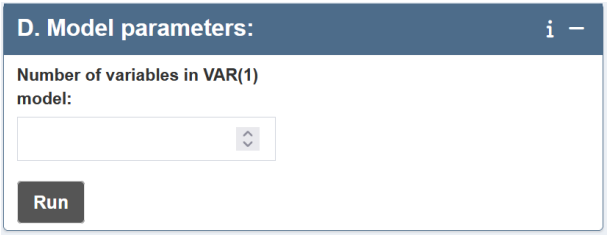
A: Simulation parameters: seed, #replicates and #time points (T)

C: Power settings: alpha and power target



Power for VAR(1): using the shiny app

Parameters: select the number of variables in the VAR(1) model



Power for VAR(1): using the shiny app

Parameters: set the values of the model parameters of the VAR(1) model

D. Model parameters: i –

Number of variables in VAR(1) model:

2

Y1 PA Y2 NA.

Transition matrix formula

Intercept matrix

	PA	NA.
	0	0

Transition matrix:

	PA	NA.
PA	0	0
NA.	0	0

Sigma matrix:

	PA	NA.
PA	1	0
NA.	0	1

Run

Power for VAR(1): using the shiny app

Intercepts:

$$\text{PA}_t = \beta_{10} + \beta_{11}\text{PA}_{t-1} + \beta_{12}\text{NA}_{t-1} + \epsilon_{1t}$$

$$\text{NA}_t = \beta_{20} + \beta_{22}\text{NA}_{t-1} + \beta_{21}\text{PA}_{t-1} + \epsilon_{2t}$$

Intercept matrix:

	PA	NA
	β_{10}	β_{20}

Power for VAR(1): using the shiny app

Intercepts:

$$PA_t = 23.46 + .39 * PA_{t-1} + (-.08) * NA_{t-1} + \varepsilon_{1t}$$

$$NA_t = 17.88 + .38 * NA_{t-1} + (-.02) * PA_{t-1} + \varepsilon_{2t}$$

Intercept input:

Intercept matrix

	PA	NA.
	23.46	17.88

Power for VAR(1): using the shiny app

Coefficients of the transition matrix:

$$\text{PA}_t = \beta_{10} + \beta_{11}\text{PA}_{t-1} + \beta_{12}\text{NA}_{t-1} + \epsilon_{1t}$$

$$\text{NA}_t = \beta_{20} + \beta_{22}\text{NA}_{t-1} + \beta_{21}\text{PA}_{t-1} + \epsilon_{2t}$$

Transition matrix:

	PA	NA
PA	β_{11}	β_{12}
NA	β_{21}	β_{22}

Power for VAR(1): using the shiny app

Coefficients of the transition matrix:

$$PA_t = 23.46 + .39 * PA_{t-1} + (-.08) * NA_{t-1} + \varepsilon_{1t}$$

$$NA_t = 17.88 + .38 * NA_{t-1} + (-.02) * PA_{t-1} + \varepsilon_{2t}$$

Input the values of the parameters of the transition matrix:

Transition matrix:

	PA	NA.
PA	0.39	-0.08
NA.	-0.02	0.38

Power for VAR(1): using the shiny app

Variance-covariance matrix of the within-person errors:

$$\epsilon \sim \quad N\left[\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma_{00} & \sigma_{01}\\ \sigma_{10} & \sigma_{11}\end{bmatrix}\right]$$

Variance-covariance matrix of the within-person errors:

	PA	NA
PA	σ_{11}	σ_{12}
NA	σ_{12}	σ_{22}

Power for VAR(1): using the shiny app

Variance-covariance matrix of the within-person errors:

$$\epsilon \sim \quad N\left[\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}90.6 & 4.30\\4.30 & 101.2\end{bmatrix}\right]$$

Input the values of the parameters of the variance-covariance matrix of the within-person errors:

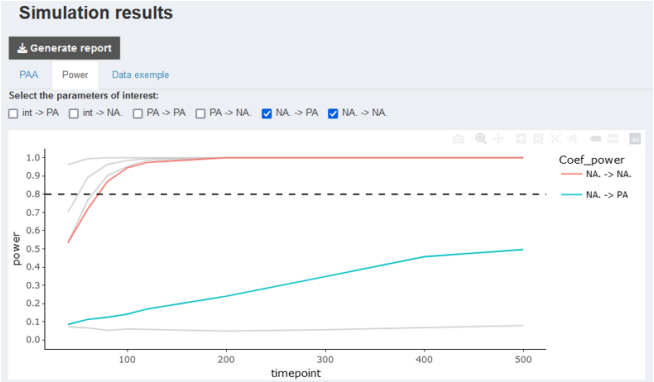
Sigma matrix:

	PA	NA.
PA	90.6	4.3
NA.	4.3	101.2

Run

Power for VAR(1): using the shiny app

- Power curve



- Sample size recommendation: 60 and 500+

Power for VAR(1): using the shiny app

- Summary table

Summary table

For each number of measurement occasions, displays the estimated power associated to each model parameter.

Coefficients	40	60	80	100	120	200	300	400	500
int -> PA	0.96	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
int -> NA	0.70	0.89	0.96	0.98	0.99	1.00	1.00	1.00	1.00
PAlag -> PA	0.53	0.77	0.90	0.95	0.98	1.00	1.00	1.00	1.00
PAlag -> NA	0.07	0.07	0.05	0.06	0.06	0.05	0.06	0.07	0.08
NA.lag -> PA	0.08	0.11	0.12	0.14	0.17	0.24	0.35	0.46	0.50
NA.lag -> NA	0.53	0.72	0.87	0.94	0.97	1.00	1.00	1.00	1.00

Simulation duration: 20 mins

Remark: power-based sample size recommendations differ according to the effect of interest!

Exercise

Goal: Select the sample size for a VAR(1) model with 3 variables: PA, NA, and anhedonia to design a new study for a person with similar characteristics to Ana.

We want high power (.8) for:

- auto-regressive effect of PA
- cross-regressive effect of NA on PA
- cross-regressive effect of NA on anhedonia

Follow the steps:

1. Give as inputs: the model, #time points (T = 100, 150)
2. Run the simulation and interpret the results
3. Select sample size

PA outcome:

	Est.	S.E.	t val.	p
(Intercept)	22.60	6.51	3.47	0.00
PA.lag	0.39	0.13	2.91	0.01
NA.lag	-0.13	0.14	-0.91	0.37
anhedonia.lag	0.06	0.10	0.55	0.59

Standard errors: OLS

NA outcome:

	Est.	S.E.	t val.	p
(Intercept)	14.84	6.66	2.23	0.03
PA.lag	-0.03	0.14	-0.18	0.85
NA.lag	0.22	0.14	1.54	0.13
anhedonia.lag	0.20	0.10	1.91	0.06

Standard errors: OLS

Anhedonia outcome:

	Est.	S.E.	t val.	p
(Intercept)	14.95	8.92	1.68	0.10
PA.lag	0.04	0.18	0.24	0.81
NA.lag	0.33	0.19	1.71	0.09
anhedonia.lag	0.31	0.14	2.20	0.03

Standard errors: OLS

Variance-covariance matrix:

```
##           [,1]      [,2]      [,3]
## [1,] 90.034929  2.375884  16.02124
## [2,]  2.375884 94.326082  37.21901
## [3,] 16.021244 37.219012 169.22933
```

Exercise: solution

PA outcome:

	Est.	S.E.	t val.	p
(Intercept)	22.60	6.51	3.47	0.00
PA.lag	0.39	0.13	2.91	0.01
NA.lag	-0.13	0.14	-0.91	0.37
anhedonia.lag	0.06	0.10	0.55	0.59

Standard errors: OLS

NA outcome:

	Est.	S.E.	t val.	p
(Intercept)	14.84	6.66	2.23	0.03
PA.lag	-0.03	0.14	-0.18	0.85
NA.lag	0.22	0.14	1.54	0.13
anhedonia.lag	0.20	0.10	1.91	0.06

Standard errors: OLS

Anhedonia outcome:

	Est.	S.E.	t val.	p
(Intercept)	14.95	8.92	1.68	0.10
PA.lag	0.04	0.18	0.24	0.81
NA.lag	0.33	0.19	1.71	0.09
anhedonia.lag	0.31	0.14	2.20	0.03

Standard errors: OLS

Variance-covariance matrix:

```
##             [,1]      [,2]      [,3]
## [1,] 90.034929  2.375884  16.02124
## [2,]  2.375884 94.326082  37.21901
## [3,] 16.021244 37.219012 169.22933
```

Intercepts

PA	NA	Anhed.
22.6	14.84	14.95

Transition

	PA	NA	Anhed.
PA	.39	-.13	.06
NA	-.03	.22	.2
Anhed.	.04	.33	.31

Variance-covariance

	PA	NA	Anhed.
PA	90	2	16
NA	2	94	37
Anhed.	16	37	169

Exercise: solution

D. Model parameters:

Number of variables in VAR(1) model:

3

Y1 PA **Y2** NA. **Y3** Anhed.

Transition matrix formula

Intercept matrix

	PA	NA.	Anhed.
	22.6	14.84	14.95

Transition matrix:

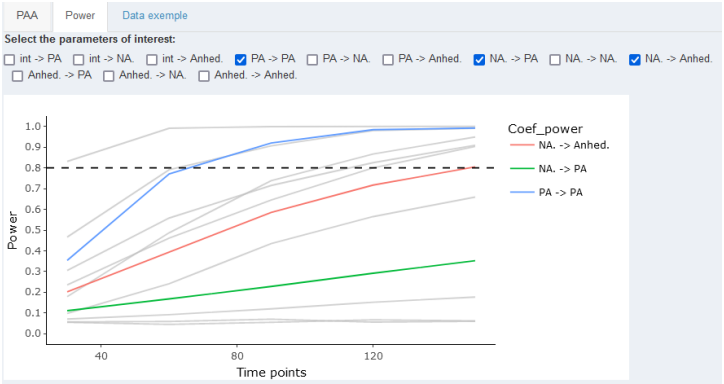
	PA	NA.	Anhed.
PA	0.39	-0.13	0.06
NA.	-0.03	0.22	0.2
Anhed.	0.04	0.33	0.31

Sigma matrix:

	PA	NA.	Anhed.
PA	90	2	16
NA.	2	94	37
Anhed.	16	37	169

Run

Exercise: solution



- What sample size recommendation? 70, 150 or more?

Exercise: sensitivity analysis

Goal: Run sensitivity analysis to explore uncertainty around the recommended sample sizes.

Explore how the sample size recommendations change depending on:

- Hypothesized parameters' values:
  - Auto-regressive PA: .39 to .8
  - Variance PA: 90 to 180
- Number of replicates (R=100)
- Select lower and upper bounds of CI of the auto-regressive PA

PA outcome:

	2.5 %	97.5 %
(Intercept)	9.5296021	35.6778667
PA.lag	0.1208077	0.6605412
NA.lag	-0.4063407	0.1530965
anhedonia.lag	-0.1488169	0.2600881

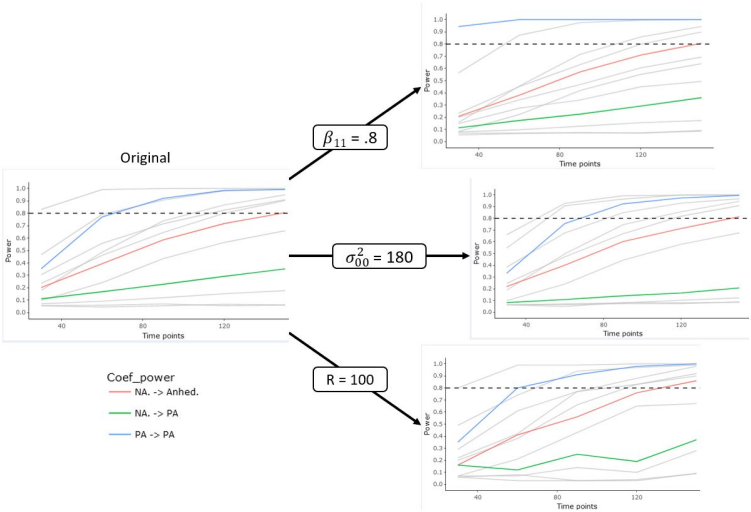
NA outcome:

	2.5 %	97.5 %
(Intercept)	1.46062438	28.2247622
PA.lag	-0.30162831	0.2508175
NA.lag	-0.06736614	0.5052475
anhedonia.lag	-0.01071022	0.4078257

Anhedonia outcome:

	2.5 %	97.5 %
(Intercept)	-2.97591249	32.8729333
PA.lag	-0.32611856	0.4138472
NA.lag	-0.05709339	0.7098859
anhedonia.lag	0.02741213	0.5880142

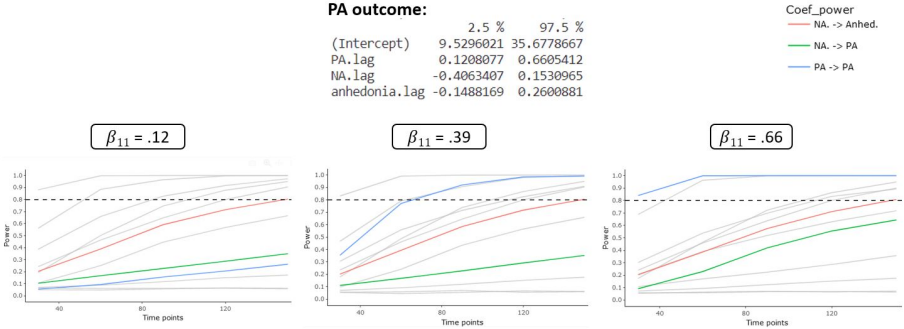
Exercise: sensitivity analysis



Exercise: sensitivity analysis

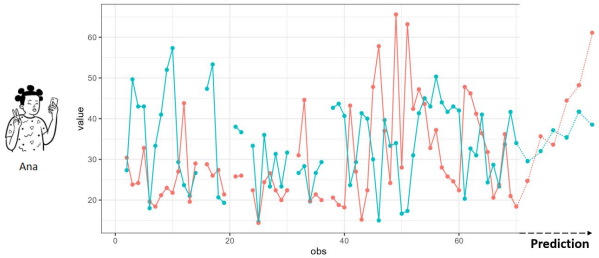
PA outcome:

	2.5 %	97.5 %
(Intercept)	9.5296021	35.6778667
PA.lag	0.1208077	0.6605412
NA.lag	-0.4063407	0.1530965
anhedonia.lag	-0.1488169	0.2600881



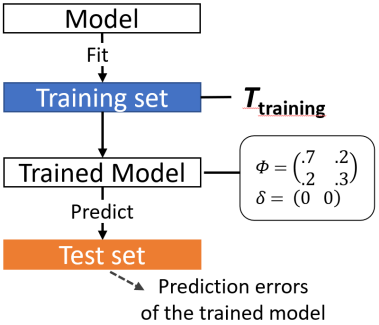
- Sample size uncertainty: from 30 to 150+

Predictive accuracy: what is the goal?



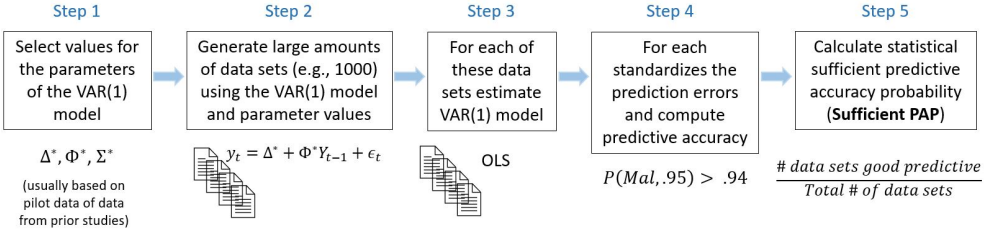
- Predictive accuracy = performance of the whole model on unseen data
- Predict at t+1 (forecasting) and generality (overfitting)

How to compute predictive accuracy?



Predictive accuracy analysis (PAA)

PAA: Optimize the number of measurement occasions to have an **high probability** to achieve a **good predictive accuracy** using a simulation-based approach.



PAA: an illustration

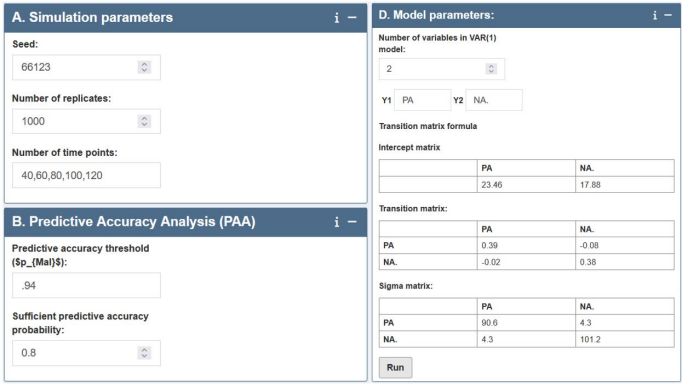
For a participant similar to Ana, we want a model of the affect system (PA and NA) that:

- Predict well the next values (t+1)
- Can generalize well to unseen data (prevents overfitting)

Goal: Determine the number of measurement occasions (T) for a-priori high predictive accuracy and a good probability to reach it

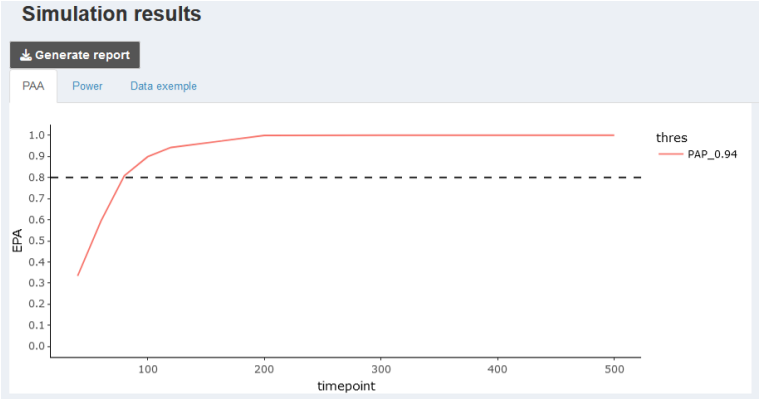
PAA: using shiny app

Hypothesized model: VAR(1) model for PA and NA (see previous slides)



PAA: using shiny app

- Sufficient predictive accuracy probability curve



PAA: using shiny app

- Summary table

Summary table

For each threshold and number of measurement occasions, displays the Sufficient Predictive Accuracy Probability estimated using the simulation.

	40	60	80	100	120	200	300	400	500
PAP_0.94	0.33	0.59	0.81	0.90	0.94	1.00	1.00	1.00	1.00

Based on simulation's parameters:

- For PAA with P_Mal = 0.94 and Sufficient PAP = 0.8 : 80 time points required.

- Sample size recommended: T=80

Exercise

Goal: select the sample size for predictive accuracy for a VAR(1) model with 3 variables: PA, NA, and anhedonia to design a new study for a person with similar characteristics to Ana.

Using Ana's data, we estimated the parameters of the VAR(1) model and we already ran the simulation for power.

Follow the steps:

1. Select sample size
2. Does it differ from the previous model?
3. What if we change the expected values (sensitivity analysis)?

PA outcome:

	Est.	S.E.	t val.	p
(Intercept)	22.60	6.51	3.47	0.00
PA.lag	0.39	0.13	2.91	0.01
NA.lag	-0.13	0.14	-0.91	0.37
anhedonia.lag	0.06	0.10	0.55	0.59

Standard errors: OLS

NA outcome:

	Est.	S.E.	t val.	p
(Intercept)	14.84	6.66	2.23	0.03
PA.lag	-0.03	0.14	-0.18	0.85
NA.lag	0.22	0.14	1.54	0.13
anhedonia.lag	0.20	0.10	1.91	0.06

Standard errors: OLS

Anhedonia outcome:

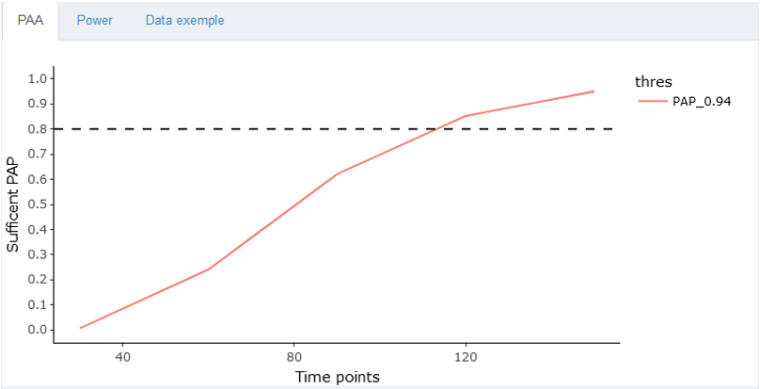
	Est.	S.E.	t val.	p
(Intercept)	14.95	8.92	1.68	0.10
PA.lag	0.04	0.18	0.24	0.81
NA.lag	0.33	0.19	1.71	0.09
anhedonia.lag	0.31	0.14	2.20	0.03

Standard errors: OLS

Variance-covariance matrix:

```
##              [,1]      [,2]      [,3]
## [1,] 90.034929  2.375884  16.02124
## [2,]  2.375884 94.326082  37.21901
## [3,] 16.021244 37.219012 169.22933
```

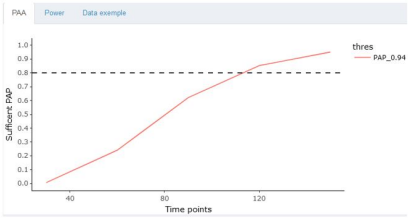
Exercise: Solution



- Sample size recommendation: 110

Exercise: Sensitivity analysis

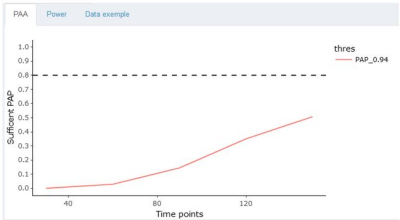
Original



	PA	NA	Anhed.
PA	.39	-.13	.06
NA	-.03	.22	.2
Anhed.	.04	.33	.31

Transition

Sensitivity

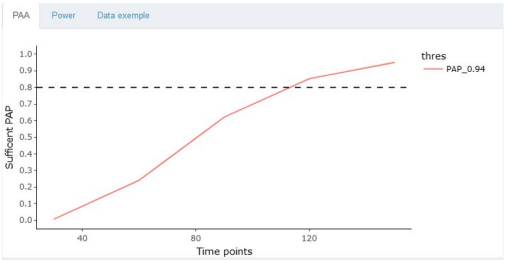


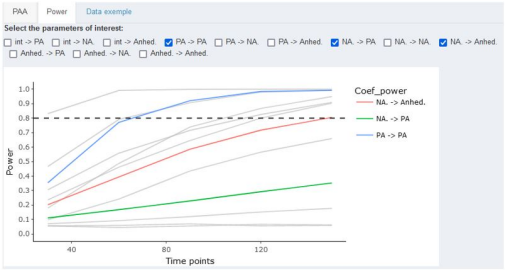
	PA	NA	Anhed.
PA	.5	-.4	.2
NA	-.3	.51	.4
Anhed.	.4	.6	.5

Transition

- A counter-intuitive effect: higher coefficients, higher recommendation

Exercise: PAA and power analysis





Factors influencing sample size recommendations based on PAA

- auto- and cross-regressive coefficient values
- complexity of the model (#variables)

Sample size planning for $N > 1$ IL designs

We now consider IL design where $N > 1$

- Repeated measurements are nested in persons: the simple regression assumption that errors across all observations are independent is violated
- Relationship between predictor and criterion can be different within or between individuals

Multilevel models

- Multilevel models extend the regression models by incorporating 'random effects' to account for between-person differences: within-person relations may differ across individuals
- Multilevel models are estimated using maximum likelihood or restricted maximum likelihood

Power-based sample size planning for multilevel models

Goal: Select the number of persons N and the number of repeated measurements T to test hypotheses that can be investigated with multilevel models with high power

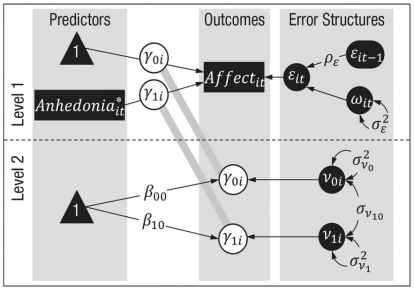
In this workshop, we will focus on:

- Power analysis for multilevel models that account for serial dependency
- The target will be the number of participants N given a predefined number of measurement occasions T
- Use analytic and simulation-based approaches
- Sensitivity analysis: varying the number of repeated measurement occasions T and the value of the model parameters

Concurrent within-person relationships

Goal: investigate relationships between time-varying variables at time t

Example: Does anhedonia at time t predict NA at time t?



where γ_{00} and γ_{10} denote the random intercept and slope and β_{00} and β_{10} denote the fixed intercept and slope

Multilevel model to investigate concurrent within-person relationships

Include random effects to account for between-person differences in the model coefficients (i.e., intercepts and slopes)

Level 1:

$$\text{NA}_{it} = \gamma_{0i} + \gamma_{1i}\text{Anhedonia}_{it} + \epsilon_{it}$$

Level 2:

$$\gamma_{0i} = \beta_{00} + \nu_{0i} \quad \text{random intercept}$$
$$\gamma_{1i} = \beta_{10} + \nu_{1i} \quad \text{random slope}$$

- β_{00} is the fixed intercept and β_{10} is the fixed slope

- ϵ_{it} is the within-person error: serially correlated following an AR(1) process with variance σ^2_ϵ and autocorrelation ρ_ϵ

- ν_{0i} and ν_{1i} are the random effects which are bivariate normal distributed: $\sigma^2_{\nu_0}$, $\sigma^2_{\nu_1}$, $\rho_{\nu_{01}}$

Example. Power analysis to investigate if anhedonia at time t predicts NA at time t

- Design a new IL study to investigate if anhedonia at time t predicts NA at time t in a sample of persons diagnosed with depression
- The new study will include 70 repeated measurement occasions
- Research hypothesis: the effect of anhedonia at time t on NA at time t is different from zero

$$H_0 : \beta_{10} = 0$$
$$H_1 : \beta_{10} \neq 0$$

How many participants are needed to test the hypothesis of interest with high statistical power?

Simulation-based power analysis for multilevel models

PowerAnalysisIL: a shiny app to perform power analysis for multilevel models applied to in IL studies

- The application implements a simulation-based approach to calculate statistical power

- Link to the shiny app: https://github.com/ginettelafit/PowerAnalysisIL

Script to run the shiny app

```
devtools::install_github("ginettelafit/PowerAnalysisIL", force = T)
library(PowerAnalysisIL)
shiny::runGist("6bac9d35c2521cc4fd91ce4b82490236")
```

Let's conduct the simulation-based power analysis!

Step 0: determine the values of the model parameter

To obtain the values of the model parameters we will use data from the Leuven clinical study

Estimation of the multilevel model using REML

REML estimation using R

```
fit.Model.1 = lme(NA.   1 + anhedonia.c, random =  1
+ anhedonia.c|PID,na.action=na.omit, data=data.MDD,
correlation=corAR1(), method="REML")

summary(fit.Model.1)
```

Estimation output

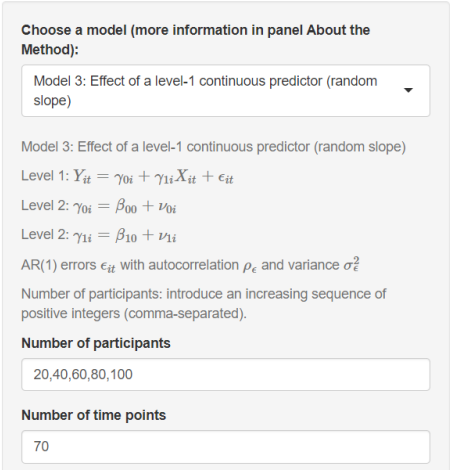
```
## Random effects:
##  Formula: ~1 + anhedonia.c | PID
##  Structure: General positive-definite, Log-Cholesky parametrization
##             StdDev     Corr
## (Intercept) 14.7788369 (Intr)
## anhedonia.c  0.1162717 0.003
## Residual    11.9150995
##
## Correlation Structure: AR(1)
##  Formula: ~1 | PID
##  Parameter estimate(s):
##       Phi
## 0.4293834
## Fixed effects:  NA. ~ 1 + anhedonia.c
##                 Value Std.Error   DF   t-value p-value
## (Intercept) 42.98279 2.4299656 2216 17.688641       0
## anhedonia.c  0.13900 0.0233386 2216  5.955753       0
```

Mean anhedonia: 51.66162

Std. deviation anhedonia: 23.6734

Step 1: in the PowerAnalysisIL app select the model and set the sample size

i. Indicate the model of interest
ii. Input the number of participants N (comma-separated): $N = 20, 40, 60, 80, 100$
iii. Input the number of repeated measurement occasions: $T = 70$



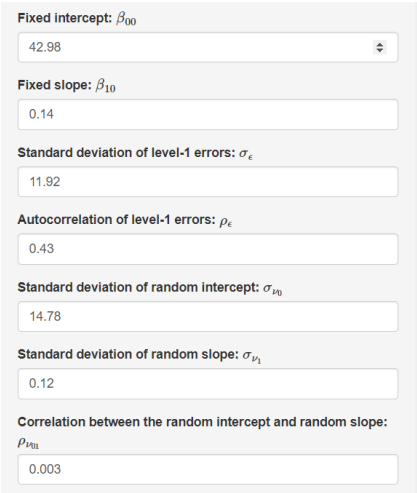
Step 2: in the PowerAnalysisIL app set the value of the model parameters

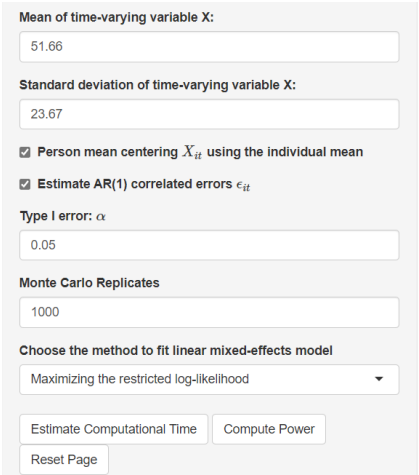
- We use the values obtained using the data from the Leuven clinical study
- Multilevel model:

$$\text{NA}_{it} = \beta_{00} + \beta_{10}\text{Anhedonia}_{it} + \nu_{0i} + \nu_{1i}\text{Anhedonia}_{it} + \epsilon_{it}$$

$$\begin{aligned}
\beta_{00} &= 42.98 \quad \text{fixed intercept} \\
\beta_{10} &= 0.14 \quad \text{fixed slope} \\
\sigma_{\epsilon} &= 11.92 \quad \text{std. deviation Level 1 errors} \\
\rho_{\epsilon} &= 0.43 \quad \text{std. deviation Level 1 errors} \\
\sigma_{\nu_0} &= 14.78 \quad \text{std. deviation random intercept} \\
\sigma_{\nu_1} &= 0.12 \quad \text{std. deviation random slope} \\
\rho_{\nu_{01}} &= 0.003 \quad \text{correlation between the random effects} \\
\mu_{\text{Anhedonia}} &= 51.66 \quad \text{mean anhedonia} \\
\sigma_{\text{Anhedonia}} &= 23.67 \quad \text{std. deviation anhedonia}
\end{aligned}$$

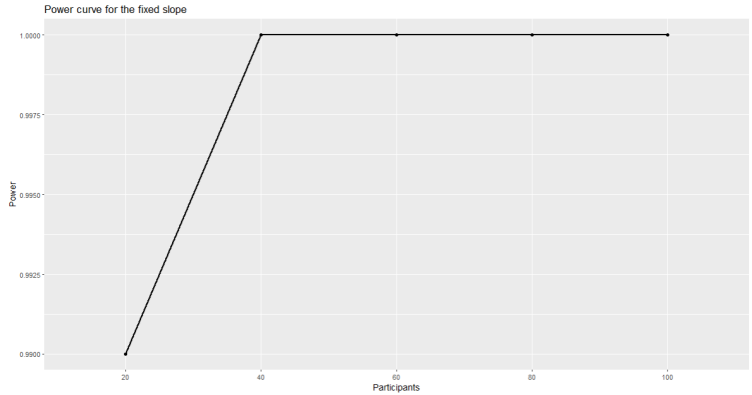
Step 2: in the PowerAnalysisIL app set the value of the model parameters





Step 3: inspect simulation results

- Statistical power is higher than 90% when the number of participants is equal to or higher than 20



Power analysis for multilevel models using the analytic approach

ApproxPowerIL: a shiny app to perform power analysis for multilevel models using analytical derivations

- The application uses asymptotic approximations for the standard errors to calculate statistical power

- Link to the shiny app:
https://gitlab.kuleuven.be/ppw-okpiv/researchers/u0119584/ApproxPowerIL

Script to run the shiny app

```
remotes::install_github("ginettelafit/ApproxPowerIL", force = T)
library(ApproxPowerIL)
shiny::runGist("302737dc046b89b7f09d15843389161c")
```

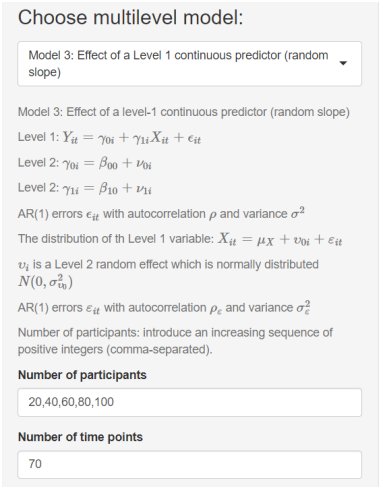
Let's conduct the power analysis using the analytic approach!

Step 1: in the ApproxPowerIL app select the model and set the sample size

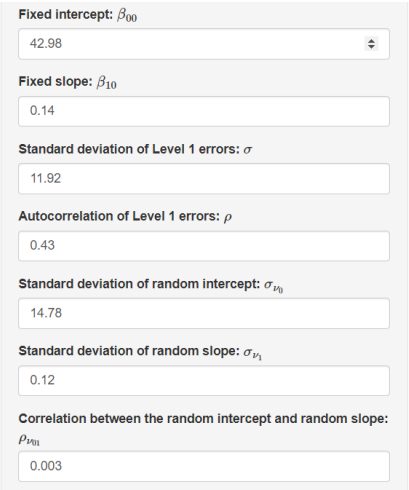
i. Indicate the model of interest

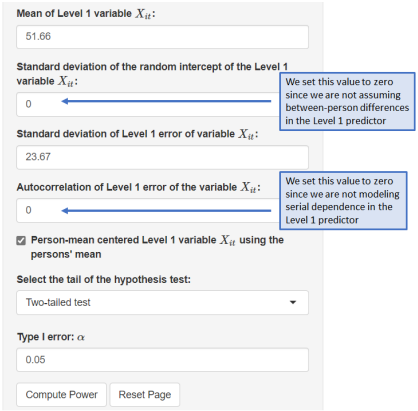
ii. Input the number of participants N (comma-separated): $N = 20, 40, 60, 80, 100$

iii. Input the number of repeated measurement occasions: $T = 70$



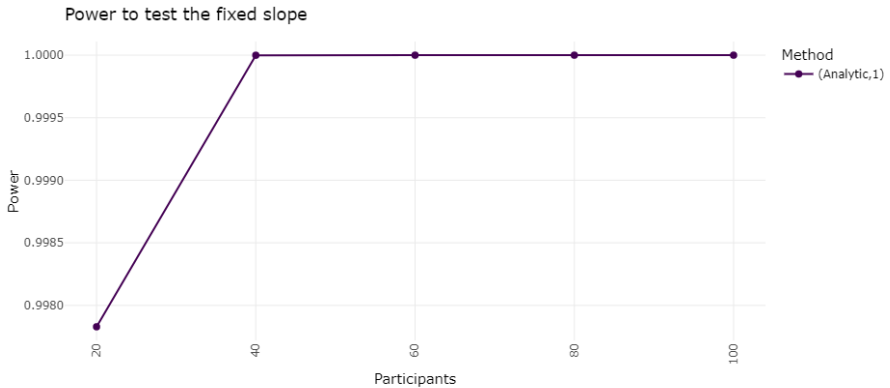
Step 2: in the ApproxPowerIL app set the value of the model parameters





Step 3: inspect results

- Statistical power is higher than 90% when the number of participants is equal to or higher than 20



Remark

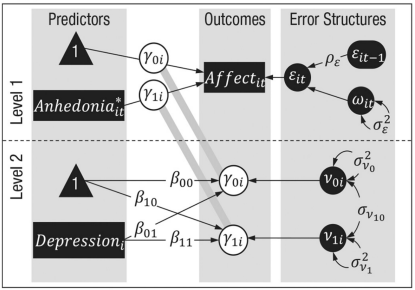
Differences between the two approaches to calculate statistical power

- The analytic approach uses asymptotic approximations for deriving the standard errors of the estimates of the fixed effect, whereas in the simulation-based approach, the model of interest is fitted to the simulated data
- The two approaches yield different sample size recommendations when either N or T are small
- The analytic approach can be used first to obtain the power curve over different sample size values. Next, the simulation-based approach can be used over a restricted range of N to save computational time

Between-person differences in within-person relationships

Extending the previous model by incorporating the interaction effect between a time-invariant (Level 2) and time-varying (Level 1) predictor

Example: Does depression moderates the effect of Anhedonia on NA?



where β_{11} is the fixed cross-level interaction effect between Anhedonia and Depression

Multilevel model to investigate cross-level interaction effects

The model includes a Level 2 continuous predictor: Depression

Level 1:

$$\text{NA}_{it} = \gamma_{0i} + \gamma_{1i}\text{Anhedonia}_{it} + \epsilon_{it}$$

Level 2:

$$\gamma_{0i} = \beta_{00} + \beta_{01}\text{Depression}_i + \nu_{0i} \quad \text{random intercept}$$
$$\gamma_{1i} = \beta_{10} + \beta_{11}\text{Depression}_i + \nu_{1i} \quad \text{random slope}$$

- β_{11} represents the cross-level interaction effect between depression and anhedonia

Exercise

Goal: select the number of persons to investigate if depression moderates the relationship between anhedonia and NA: $H_0 : \beta_{11} = 0$ vs. $H_1 : \beta_{11} \neq 0$

Using the Leuven clinical study, we estimated the parameters of the multilevel model including the cross-level interaction effect

Follow the steps:

1. Select sample size using the analytic approach ($N = 20, 40, 60, 100, 120$)
2. Compare the results with the ones obtained using the simulation-based approach

Estimation output

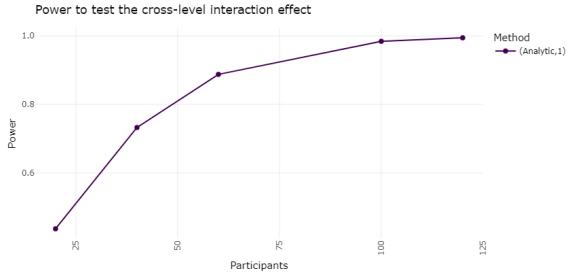
```
## Random effects:
##  Formula: ~1 + anhedonia.c | PID
##  Structure: General positive-definite, Log-Cholesky parametrization
##             StdDev     Corr
## (Intercept) 12.8555036 (Intr)
## anhedonia.c  0.1056154 0.249
## Residual    11.9234081
##
## Correlation Structure: AR(1)
##  Formula: ~1 | PID
##  Parameter estimate(s):
##       Phi
## 0.4302492
## Fixed effects:  NA. ~ 1 + anhedonia.c + anhedonia.c * QIDS.c
##                        Value Std.Error   DF   t-value p-value
## (Intercept)        42.97796 2.1228637 2215 20.245274  0.0000
## anhedonia.c         0.13747 0.0218391 2215  6.294553  0.0000
## QIDS.c              1.52600 0.4308459   36  3.541870  0.0011
## anhedonia.c:QIDS.c -0.01019 0.0046382 2215 -2.197910  0.0281
```

Mean anhedonia: 51.66162
Std. deviation anhedonia: 23.6734

Mean depression: 15.71
Std. deviation depression: 5.00

Exercise: Solution Using the Analytic Approach

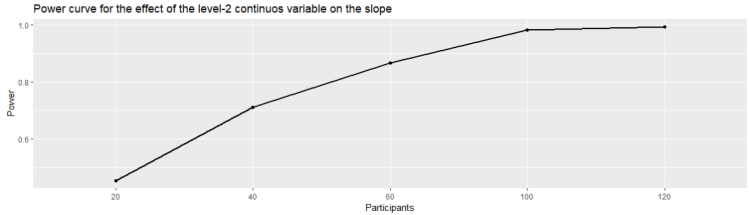
- Power curve to test the moderation effect of depression on the relationship between anhedonia and NA



- Statistical power is higher than 90% when the number of participants is equal to or higher than 100

Exercise: Solution Using the Simulation-based Approach

- Power curve to test the moderation effect of depression on the relationship between anhedonia and NA

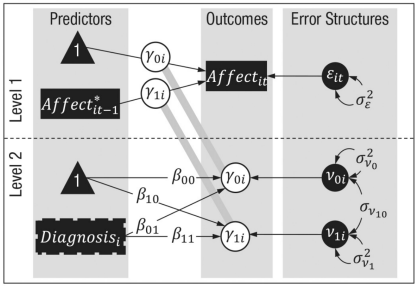


- Statistical power is higher than 90% when the number of participants is equal to or higher than 100

Lagged within-person relationships

Multilevel AR(1) Model

Example: estimate differences in the autoregressive effect of NA between persons diagnosed with depression and controls



where β_{11} denotes the difference in the fixed autoregressive effects between the two groups

Multilevel AR(1) model

Level 1:

$$\text{NA}_{it} = \gamma_{0i} + \gamma_{1i}\text{NA}_{it-1} + \epsilon_{it}$$

Level 2:

$$\gamma_{0i} = \beta_{00} + \beta_{10}\text{Diagnosis} + \nu_{0i} \quad \text{random intercept}$$
$$\gamma_{1i} = \beta_{10} + \beta_{11}\text{Diagnosis} + \nu_{1i} \quad \text{random autoregressive effect}$$

- β_{10} is the fixed autoregressive effect - β_{11} is the difference in the fixed autoregressive effect between the two groups

- ϵ_{it} is the within-person error: independent and identically distributed $N(0, \sigma^2_\epsilon)$

Power analysis for multilevel AR(1) models

- In the context of multilevel AR(1) models, power analysis can be conducted using the simulation-based approach
- There are no analytical formulas that can be used to calculate statistical power using the analytic approach

Power analysis for Multilevel AR(1) models

Power calculations for these models can be conducted using PowerAnalysisIL app

Exercise. Sensitivity analyses

Goal: select the number of persons in each group to investigate group differences in the fixed AR effect: $H_0 : \beta_{11} = 0$ **vs.** $H_1 : \beta_{11} \neq 0$

Follow the steps:

1. Conduct sensitivity analysis to investigate differences in statistical power when varying the number of measurement occasions T due to different levels of compliance (i.e., 60% and 80%)
2. Conduct sensitivity analysis to investigate differences in statistical power when varying the value of β_{11}: we assume $\beta_1 1$ is 10% lower/higher than the one obtained using the Leuven clinical data set

Exercise. Sensitivity analyses

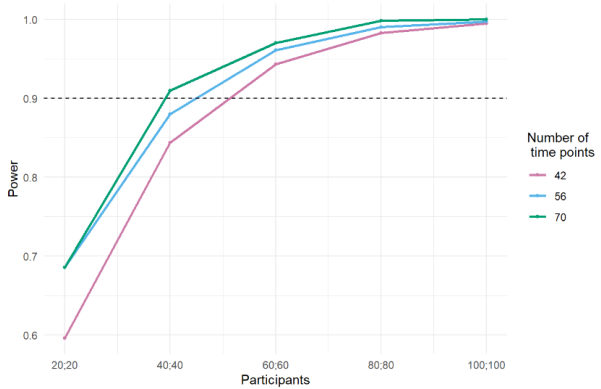
- Using the Leuven clinical study, we estimated the parameters of the multilevel AR(1) model

Estimation output

```
## Random effects:
##  Formula: ~1 + NA.lag | PID
##  Structure: General positive-definite, Log-Cholesky parametrization
##             StdDev    Corr
## (Intercept) 5.7874498 (Intr)
## NA.lag      0.1402727 -0.199
## Residual    8.7540300
##
## Fixed effects:  NA. ~ 1 + MDD + NA.lag + MDD * NA.lag
##                  Value Std.Error   DF   t-value p-value
## (Intercept)  6.824841 0.9800413 3911  6.963830  0.0000
## MDD         16.326600 1.5896204   76 10.270754  0.0000
## NA.lag       0.313887 0.0366665 3911  8.560574  0.0000
## MDD:NA.lag   0.116184 0.0472239 3911  2.460275  0.0139
```

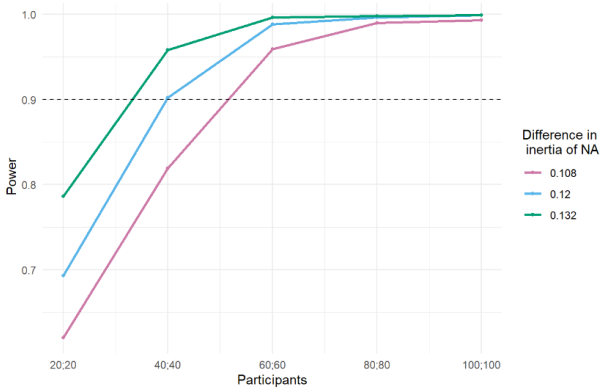
Exercise: Solution

Sensitivity analysis when varying the number of repeated measurements occasions T: statistical power decreases when T decreases



Exercise: Solution

Sensitivity analysis when varying the value of β_{11}: statistical power increases when the absolute value of $\beta_1 1$ increases



Remark I

The estimation framework proposed in this workshop assumes:

- repeated measurements are equidistant
- ignore night blocks
- cannot handle missing observations (i.e., missing values are listwise deleted)

How can we take these considerations into account when conducting a power analysis?

Remark I

Considerations related to the selection of T for power analysis

When conducting a power analysis select T considering the following:

- % of missing values (i.e., compliance)
- the number of observations that will be missing due to lagging the predictor within days

ID	Day	Beep	PA	PA.lag
1	1	1	NA	NA
1	1	2	27.33	NA
1	1	3	49.67	27.33
1	1	4	43.00	49.67
1	2	1	18.00	NA
1	2	2	33.33	18.00
1	2	3	41.00	33.33
1	2	4	52.00	41.00

Three missing observations: one missing observation for PA and two observations are missing after lagging PA

Remark II

How to tackle the uncertainty about the parameter values when computing statistical power?

- To calculate statistical power for the models presented in this workshop it is necessary to get information about the value of the intercept, autoregressive effect, and standard deviation of the within-person errors
- This is usually done by using data from previous studies, BUT findings from previous studies may be biased

Solution: conduct a sensitivity analysis to assess the influence of the values of the model parameters on power

We showcase how to conduct such a sensitivity analysis in this project: https://psyarxiv.com/7msh6/

Remark III

Considerations related to the validity of sample size recommendations

Both power analysis and PAA are conducted prior to data collection. Thus, the generalizability of sample size recommendation depends on the validity of the model when fitted to the data of the new study

Advanced methods for sample size planning

We will now present advanced methods for sample size planning [slides are available at http://samplesize.help/]