Cost-Effective ESM Studies

Application to another research question and statistical model

March 14, 2025

Contents

1 Introduction

This supplementary material is associated with the article *Cost-Effective ESM Studies: Integrating Budget Constraints into Sample Size Decisions*.

In this supplementary material, we built on the step-by-step approach from the manuscript to explore sample size recommendations for another research question and derived model. Peter, Laure, and Jean are now interested in how the time-varying predictor anhedonia is contemporaneously associated with negative affect (NA). Their study scenario and design remain the same, leading to the same cost functions.

2 Method

As in the main manuscript, we will proceed systematically through each step of the approach. Elements identical to those in the paper will not be discussed here.

2.1 Step 1: Research Question and Study designs of Peter, Laure, and Jean

Peter, Laure, and Jean want to study the effect of anhedonia on negative affect in daily life. Previous research found that a higher level of anhedonia was momentarily associated with a higher level of negative affect. In other words, Peter, Laure, and Jean expect anhedonia to be positively associated with negative affect within-person.

For this alternative research question, Peter, Laura, and Jean will use the ESM designs presented in the manuscript to study this question.

2.2 Step 2: Select the Statistical Model

Laure and Jean's research question and hypothesis involve estimating and testing if anhedonia is positively associated with negative affect, within-person. To estimate this within-person effect, the researchers used a multilevel regression model with anhedonia as a predictor of momentary time-varying negative affect. The equation of the model reads as follows:

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

where $i = (1, ..., N)$ denotes the person and $t = (1, ..., T)$ the measurement occasion. In this model, β_{00} represents the fixed intercept, indicating the expected level of negative affect (NA) of an average person when anhedonia is zero. The fixed slope β_{10} captures the average within-person relationship between anhedonia and NA across all participants.

The model includes two random effects. ν_{0i} represents the person-specific deviation from the overall intercept, allowing each individual to have their own baseline level of NA. Similarly, ν_{1i} captures the person-specific deviation from the average anhedonia-NA relationship, acknowledging that the strength of this relationship may vary across individuals.

Finally, ϵ_{it} represents the within-person errors, which account for unexplained variation in NA that cannot be attributed to anhedonia. These errors are assumed to follow a Gaussian distribution centered on zero with variance σ^2_ϵ. The model assumes that the errors are identically and independently distributed and that the time intervals between measurement occasions are equal. It's worth noting that this multilevel model allows us to simultaneously examine both within-person processes (how changes in anhedonia relate to changes in NA within individuals) and between-person differences (how individuals differ in their baseline NA and in the strength of the anhedonia-NA relationship).

Given the research question at hand, the parameter of interest is β_{10} as this parameter quantifies the extent to which negative affect is associated with an increase or decrease in anhedonia. Hence, if β_{10} is positive and significantly larger than zero, this would support Peter, Laure, and Jean's hypothesis. Different tools and estimation methods can be used to fit this model (see Lafit et al. 2021). Jean and Laure have decided to rely on the Restricted Maximum Likelihood Estimation (RMLE) implemented in the 'lme' function from the nlme R package (Pinheiro, Bates, and R Core Team 2023) and, as commonly done, Peter, Laure, and Jean decided to person-mean center the predictor Anhedonia_{it} (Hamaker and Grasman 2015).

2.3 Step 3: Power Analysis

Peter, Jean, and Laure will again rely on a simulation-based power analysis to estimate power (Revol, Lafit, and Ceulemans 2024) and to obtain sample size recommendations to reach a power of 80%. Again, they will also implement missing data following an expected compliance distribution.

To estimate the hypothesized model, our researchers can rely on the same existing dataset as in the manuscript. Indeed, during the data collection, Heininga et al. (2019) also collected anhedonia rated on a slider scale ranging from 0 ("not at all") to 100 ("very much"). When estimating the model, the large effect size of β_{10} yields high statistical power even with small sample sizes, making cost considerations less critical in the sample size determination. For demonstrative purposes, we deliberately reduced this effect size to illustrate the critical role that cost can play with more modest effect sizes. The dataset is available on the EMOTE database (Kalokerinos et al. Manuscript in preparation).

Peter, Jean, and Laure built a power contour plot using a grid based on N and T values ranging from 50 to 200 with steps of 5 for both. The power contours are represented by continuous colored lines, respectively purple for 70% power, blue for 80% power, and yellow for 90% power. From Figure 1, we can deduce that increasing N has a greater impact on power than increasing T. If Peter, Laure, and Jean were to prioritize statistical power alone, they might both select, for instance, $T = 120$ and $N = 146$ achieving 80% power. However, other combinations are also viable, such as $T = 70$ and $N = 145$, $T = 100$ and $N = 146$, or $T = 150$ and $N = 145$.

2.4 Step 4: Cost Function

Given that Peter, Laure, and Jean operate under the same study design and cost constraints as described in the manuscript, the cost function and associated parameter values remain identical. The implementation of cost information within the power contour plots follows the same approach as previously described.

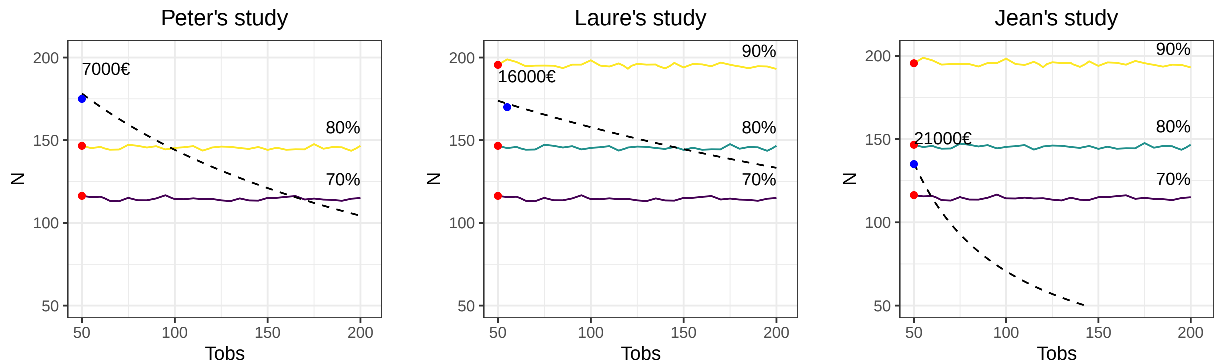

Figure 1: The red dot indicates the most cost-effective combinations of N and T that achieve 70%, 80% and 90% power. The blue dot indicates the highest achievable power within Peter, Laure, and Jean's budgets. The solid lines are the power contours, respectively representing 70%, 80% and 90% power; note that these power contours are the same in the three plots. The dashed cost contour line reflects the researcher's maximum budget.

As in the manuscript, the power and cost contour plot returns different information (see Figure 1). For a targeted power level (e.g., 80%), we can identify the combination of N and T that minimizes the total cost of the study (red dots in Figure 1), being $N = 147$ and $T = 50$ for our three researchers, with a total cost of 5773.18 euros for Peter, 13764.34 euros for Laure, and 22441.83 euros for Jean.

On the other hand, for a fixed maximum budget, we can identify the combination of N and T that maximizes power (blue dots in Figure 1), being $N = 175$ and $T = 50$ achieving 0.867 % power for Peter (6872.83 euros), $N = 170$ and $T = 55$ achieving 0.86 % power for Laure (15822.59 euros), and $N = 135$ and $T = 50$ achieving 0.767 % for Jean (20918.41 euros).

2.5 Step 5: Sample size decision

It is time for our researchers to decide on their sample size. Remember that if Peter, Laure and Jean were to prioritize statistical power alone, they would obtain the same power contours and might all select, for instance, $T = 70$ and $N = 145$ achieving 80% power

Having analyzed those outputs, Laure and Jean must finalize their sample size decisions. Remember that based on the power and cost contour plot, researchers choose any T and N combination between her 80% power contour (blue) and maximum budget contour (dashed line), as all these solutions yield power of at least 80% and remain within their maximum budget. Following the same logic as in the paper:

- **Peter** aims to identify the optimal sample size that achieves 80% power with minimal cost. The power and contour plot analysis indicates optimal parameters of $N = 147$ and $T = 50$, costing 5773.18 euros.
- **Laure** aimed to maximize her chances of detecting the effect of interest by selecting a sample size achieving 90% power. The power and contour plot analysis indicates optimal parameters of $N = 196$ and $T = 50$, costing 17845.12 euros.
- **Jean** can opt for the highest achievable power within his budget, choosing $N = 135$ and $T = 50$, to achieve 76.7 % power for 12764.97 euros.

3 Discussion

In the manuscript, our illustration showcased that even when focusing on the same research question, statistical model, and parameter of interest, dealing with different cost structures results in vastly distinct optimal sample sizes. However, This is not the case for this example. The power and cost contour plots in this supplementary material show that the optimal sample sizes to reach .8 for Peter, Laure, and Jean are the same, although the associated costs are different. This is mainly due to the strong influence of N compared to T on the power - above $T = 50$ only N has an impact on power.

However, cost is still an important piece of information to evaluate the feasibility of the study and to ensure that the research is conducted within the budget constraints.

4 References

Hamaker, Ellen L., and Raoul P. P. P. Grasman. 2015. "To Center or Not to Center? Investigating Inertia with a Multilevel Autoregressive Model." *Frontiers in Psychology* 5 (January): 1492. https://doi.org/10.3389/fpsyg.2014.01492.

Heininga, Vera E., Egon Dejonckheere, Marlies Houben, Jasmien Obbels, Pascal Sienaert, Bart Leroy, Joris van Roy, and Peter Kuppens. 2019. "The Dynamical Signature of Anhedonia in Major Depressive Disorder: Positive Emotion Dynamics, Reactivity, and Recovery." *BMC Psychiatry* 19 (1): 59. https://doi.org/10.1186/s12888-018-1983-5.

Kalokerinos, Elise K., Daniel Russo-Batterham, Peter Koval, Ella K. Moeck, Komal K. Grewal, Katie H. Greenaway, Kabir Manandhar Shrestha, et al. Manuscript in preparation. "The EMOTE Database: An Open, Searchable Database of Experience Sampling Data Mapping Everyday Life."

Lafit, Ginette, Janne K. Adolf, Egon Dejonckheere, Inez Myin-Germeys, Wolfgang Viechtbauer, and Eva Ceulemans. 2021. "Selection of the Number of Participants in Intensive Longitudinal Studies: A User-Friendly Shiny App and Tutorial for Performing Power Analysis in Multilevel Regression Models That Account for Temporal Dependencies." *Advances in Methods and Practices in Psychological Science* 4 (1): 251524592097873. https://doi.org/10.1177/2515245920978738.

Pinheiro, Jose, Douglas Bates, and R Core Team. 2023. "Nlme: Linear and Nonlinear Mixed Effects Models."

Revol, Jordan, Ginette Lafit, and Eva Ceulemans. 2024. "A New Sample-Size Planning Approach for Person-Specific VAR(1) Studies: Predictive Accuracy Analysis." *Behavior Research Methods* 56 (7): 7152–67. https://doi.org/10.3758/s13428-024-02413-4.

5 Session info

This section provides information about the session and environment in which the code was executed, including the version of R and the packages used.

```
## R version 4.4.1 (2024-06-14 ucrt)
## Platform: x86_64-w64-mingw32/x64
## Running under: Windows 11 x64 (build 26100)
##
## Matrix products: default
```

```
##
##
## locale:
## [1] LC_COLLATE=Dutch_Belgium.utf8      LC_CTYPE=Dutch_Belgium.utf8
## [3] LC_MONETARY=Dutch_Belgium.utf8     LC_NUMERIC=C
## [5] LC_TIME=English_United States.1252
##
## time zone: Europe/Brussels
## tzcode source: internal
##
## attached base packages:
## [1] grid      stats     graphics  grDevices utils     datasets  methods
## [8] base
##
## other attached packages:
##  [1] magick_2.8.5      gtable_0.3.6      tibble_3.2.1      officer_0.6.7
##  [5] flextable_0.9.7   kableExtra_1.4.0  viridis_0.6.5     viridisLite_0.4.2
##  [9] cowplot_1.1.3     ggExtra_0.10.1    latex2exp_0.9.6   tictoc_1.2.1
## [13] ggpubr_0.6.0      ggplot2_3.5.1     tidyr_1.3.1       dplyr_1.1.4
##
## loaded via a namespace (and not attached):
##  [1] xfun_0.48               rstatix_0.7.2           vctrs_0.6.5
##  [4] tools_4.4.1             generics_0.1.3          fansi_1.0.6
##  [7] highr_0.11              pkgconfig_2.0.3         data.table_1.16.2
## [10] uuid_1.2-1              lifecycle_1.0.4         farver_2.1.2
## [13] compiler_4.4.1          stringr_1.5.1           textshaping_0.4.0
## [16] munsell_0.5.1           tinytex_0.53            carData_3.0-5
## [19] fontLiberation_0.1.0    fontquiver_0.2.1        httpuv_1.6.15
## [22] htmltools_0.5.8.1       yaml_2.3.10             Formula_1.2-5
## [25] pillar_1.9.0            later_1.3.2             car_3.1-3
## [28] openssl_2.2.2           abind_1.4-8             fontBitstreamVera_0.1.1
## [31] mime_0.12               zip_2.3.1               tidyselect_1.2.1
## [34] digest_0.6.37           stringi_1.8.4           purrr_1.0.2
## [37] bookdown_0.41           labeling_0.4.3          rprojroot_2.0.4
## [40] fastmap_1.2.0           here_1.0.1              colorspace_2.1-1
## [43] cli_3.6.3               magrittr_2.0.3          utf8_1.2.4
## [46] broom_1.0.7             withr_3.0.2             gdtools_0.4.0
## [49] scales_1.3.0            promises_1.3.0          backports_1.5.0
## [52] rmarkdown_2.28          gridExtra_2.3           ggsignif_0.6.4
## [55] ragg_1.3.3              askpass_1.2.1           shiny_1.9.1
## [58] evaluate_1.0.1          knitr_1.48              miniUI_0.1.1.1
## [61] rlang_1.1.4             Rcpp_1.0.13             xtable_1.8-4
## [64] glue_1.8.0              xml2_1.3.6              svglite_2.1.3
## [67] rstudioapi_0.16.0       R6_2.5.1                systemfonts_1.1.0
```