The Evaluation of Multiple N=1 Studies

H. Hoijtink[1], Fayette Klaassen[1]

[1]Department of Methodology and Statistics
Utrecht University H.Hoijtink@uu.nl

2018

An Example Data Set

A person is presented with 20 movie clips. The clips vary with respect to emotional value and arousal (there is a clip about a starving lion and one about a coconut shell)

- 10 clips are watched using virtualy reality glasses, 10 clips are watched on a movie screen. This is the variable "condition".
- The person scores each clip with respect to:
  - duration in seconds
  - emotional value on a 1-9 Likert scale
  - arousal on a 1-9 Likert scale

1. Van der Ham, I. J. M., Klaassen, F., Van Schie, K., and Cuperus, A. (2019). Elapsed time estimates in virtual reality and the physical world: The role of arousal and emotional valence. *Computers in Human Behavior, 94,* 77-81.

The data rendered by the person

clip number	duration (d)	condition (c)	emotional (e) value	arousal (a)
1	15	vr	2	3
2	8	vr	5	4
...	...	...	...	...
11	7	ms	5	2
12	12	ms	7	2
...	...	...	...	...

Note that, such a data set is available for each of 29 persons. These data can (and will be in the sequel) be treated as multiple N=1 studies.

A model for the data of each person

A multiple regression model for the data of each person:

$$d_t = \beta_0 + \beta_c c_t + \beta_e e_t + \beta_a a_t + e_t,$$

where $t = 1, ..., 20$, that is, the model is fitted to data as presented on the previous slide. The residuals e have a normal distribution with mean 0 and variance σ^2.

The results of executing N=29 regression analyses

person	$\hat{\beta}_c$	$\hat{\beta}_e$	$\hat{\beta}_a$	Σ_β	N-trials
1	1.2	2.0	-.4	3x3	20
2	.9	1.1	2.9	3x3	20
...	...	...	...	...	...
29	-.2	1.5	1.9	3x3	20

Note that, to compute Bayes factors for hypotheses using the data of person 1, only the estimates of β, their covariance matrix (denotes by 3x3) and the sample size (N-trials = 20) corresponding to person 1 are needed. The same holds for the other 28 persons.

The Hypotheses

$$H_1^i : \beta_c = 0 \; \& \; \beta_e > 0 \; \& \; \beta_a > 0$$

$$H_2^i : \beta_c > 0 \; \& \; \beta_e > 0 \; \& \; \beta_a > 0$$

$$H_{1c}^i : \text{ not } H_1$$

Note that, the superscript i indicates that the hypotheses are evaluated **independently** for each of the $i = 1, ..., 29$ persons in the experiment.

1. Klaassen, F., Zedelius, C., Veling, H., Aarts, H., and Hoijtink, H. (2018). All for one or some for all? Evaluating informative hypotheses for multiple N=1 studies. *Behavior Research Methods, 50*, 2276-2291.
2. Klaassen, F. (2019). Combining evidence over multiple individual analyses. In: UNDER CONSTRUCTION.

The results of evaluation N=29 sets of hypotheses

person	BF^i_{1c}	BF^i_{12}
1	.16	7.25
2	.96	1.74
3	2.77	1.77
4	0,17	6.97
5	3.25	2.83
6	1.52	2.64
7	8.10	1.57
9	5.48	0,96
10	.70	9.55
11	0.03	0.01
12	3.66	1.79
...	...	...
29	1.49	3.12

How to summarize the results?

The "overall persons" hypotheses are that the constraints in each hypothesis apply to each person:

$$H_1^{total} : H_1^1 \& ... \& H_1^{29}$$

$$H_2^{total} : H_2^1 \& ... \& H_2^{29}$$

$$H_{1c}^{total} : H_{1c}^1 \& ... \& H_{1c}^{29}$$

The geometric mean of the Bayes factors for $i = 1, \ldots, 29$

To quantify the support in the data for H_1^{total}, H_2^{total}; and H_{1c}^{total} we could simply multiply the Bayes factors for the 29 persons, that is,

$$\mathrm{BF}_{12}^{total} = \mathrm{BF}_{12}^{1} \times \ldots \times \mathrm{BF}_{12}^{29} \tag{1}$$

However, this will render numbers that get bigger and bigger when adding more persons to the experiment, that is, it is a number that depends on the number of persons in the experiment. Therefore we proposed to use the geometric mean, that is, the "average of a multiplication of numbers":

$$\mathrm{GBF}_{12}^{total} = (\mathrm{BF}_{12}^{total})^{1/29} \tag{2}$$

The resulting number is the expected value of BF_{12}^{29+1}, that is the expected value for a 30th person that is added to the experiment.

"Averaged" Results

$\text{GBF}^{total}_{1,1c} = .65$

$\text{GBF}^{total}_{12} = 1.13$

How to interpret these results?

Additional statistics: The evidence rate

The proportion of persons whose BF "prefers" the same hypothesis as the GBF.

Evidence rate for $\text{GBF}_{1,1c}^{total} = .65$ equals .45

Evidence rate for $\text{GBF}_{12}^{total} = 1.13$ equals .34

How to interpret these results?

Additional statistics: The stability rate

The proportion of persons whose BF "prefers" the same hypothesis as the GBF but with stronger evidence.

Stability rate for $\text{GBF}^{\textit{total}}_{1,1c} = .65$ equals .35

Stability rate for $\text{GBF}^{\textit{total}}_{12} = 1.13$ equals .69

How to interpret these results?

The evaluation of multiple N=1 studies

Discussion ...