Bayes factors in multiple $N = 1$ studies

Fayette Klaassen*, Claire Zedelius, Harm Veling, Henk Aarts and Herbert Hoijtink

* Department of Methodology and Statistics, Utrecht University, f.klaassen@uu.nl

Research goals

This research is concerned with individual centered analyses, in the form of multiple $N = 1$ studies. A central feature of this paper is that *multiple informative hypotheses* are formulated for each person. These hypotheses are first evaluated at the *individual level* and subsequently conclusions are formed at the *group level*. Specifically, this will be done in the context of a within-subject experiment. Three questions are of interest when considering a set of hypotheses and multiple $N = 1$ studies:

1. For each person, which hypothesis of a set is supported most?
2. For each hypothesis, what is the support that it holds for every person?
3. For a set hypotheses, are the persons homogeneous in which hypothesis is supported most?

These questions are assessed by means of individual Bayes factors and 2 additional measures.

Example

Zedelius, Veling, and Aarts (2011) conducted a within-subject experiment in which they investigated the effect of distraction and reward on memory.
– $P = 26$ persons participated in the experiment
– $J = 8$ experimental conditions (high or low reward (hr - lr); high or low interference (hi - li); supra- or subliminal cue (sup - sub); resulting in for example *hr-hi-sup*)
– $R = 7$ replications in each condition (0 = failure; 1 = success)
The number of successes $\mathbf{x}^i = [x_1^i, ..., x_J^i]$ of person i can be modeled using a binomial model with R trials and unknown success probabilities $\boldsymbol{\pi}_j^i = [\pi_1^i, ...\pi_J^i]$.

Informative hypotheses

Researchers can formulate informative hypotheses based on (competing) theories or expectations (Hoijtink, 2012). A general form of the informative hypotheses $m = 1, ..., M$ is:

$$H_m^i : \mathbf{R}_m \boldsymbol{\pi}^i > 0, \tag{1}$$

an informative hypothesis for person i using K constraints specified in the K rows of $\mathbf{R}_m$.

E.g. for $J = 4$, $\mathbf{R}_{m'} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$ results in $H_{m'}^i : \pi_1^i > \pi_2^i > \pi_3^i > \pi_4^i$.

Example

Zedelius et al. (2011) formulated three informative hypotheses:

$H_1^i : \pi_1^i > \pi_2^i > ... > \pi_7^i > \pi_8^i$

$H_2^i : \frac{\pi_1^i + \pi_2^i}{2} > \frac{\pi_3^i + \pi_4^i}{2} > \frac{\pi_5^i + \pi_6^i}{2} > \frac{\pi_7^i + \pi_8^i}{2}$

$H_3^i : \frac{\pi_1^i + \pi_5^i}{2} > \frac{\pi_2^i + \pi_6^i}{2} > \frac{\pi_3^i + \pi_7^i}{2} > \frac{\pi_4^i + \pi_8^i}{2}$

H_1^i states that for each person i the success probabilities are ordered from large to small over the conditions $1, ..., J$. For sake of notation, π_1^i is used rather than $\pi_{hr\text{-}hi\text{-}sup}^i$ and similarly for other conditions.
H_2^i and H_3^i state that for each person i, pairs of success probabilities are ordered from large to small.

Density, prior, posterior, Bayes factor

The density of the data (2) is a product over J binomial distributions. The prior (3) is a product over Beta distributions with $\alpha_0 = \beta_0 = 1$, that is, a uniform distribution. This prior is neutral with respect to similar hypotheses. The posterior (4) is a product over Beta distributions with $\alpha_1 = x_j^i + \alpha_0, \beta_1 = (R - x_j^i) + \beta_0$.

$$f(\mathbf{x}^i|\boldsymbol{\pi}^i) = \prod_{j=1}^{J} \binom{R}{x_j^i} (\pi_j^i)^{x_j^i} (1 - \pi_j^i)^{R - x_j^i} \quad \text{(density)} \tag{2}$$

$$h(\boldsymbol{\pi}^i|H_u^i) = \prod_{j=i}^{J} \frac{\Gamma(\alpha_0 + \beta_0)}{\Gamma(\alpha_0) \cdot \Gamma(\beta_0)} (\pi_j^i)^{\alpha_0 - 1} (1 - \pi_j^i)^{\beta_0 - 1} \quad \text{(unconstrained prior)} \tag{3}$$

$$g(\boldsymbol{\pi}^i|\mathbf{x}^i, H_u^i) = \prod_{j=1}^{J} \frac{\Gamma(\alpha_1 + \beta_1)}{\Gamma(\alpha_1) \cdot \Gamma(\beta_1)} (\pi_j^i)^{\alpha_1 - 1} (1 - \pi_j^i)^{\beta_1 - 1} \quad \text{(unconstrained posterior)} \tag{4}$$

The BF that expresses support for H_m^i against an unconstrained hypothesis H_u^i can be written as a ratio of complexity (c_m^i) and fit (f_m^i):

$$BF_{mu}^i = \frac{f_m^i}{c_m^i} = \frac{\int_{\boldsymbol{\pi}^i \in H_m^i} g(\boldsymbol{\pi}^i|\mathbf{x}^i, H_u^i) \delta \boldsymbol{\pi}^i}{\int_{\boldsymbol{\pi}^i \in H_m^i} h(\boldsymbol{\pi}^i, H_u^i) \delta \boldsymbol{\pi}^i}, \tag{5}$$

where c_m^i is that part (3) that is in agreement with f_m^i is that part of (4) that is in agreement with (1) (Klugkist, Laudy, & Hoijtink, 2010).

Example

Research goal 1. In order to select for each person the best hypothesis from a set, each hypothesis can be evaluated against the unconstrained hypothesis. The results are shown in Table 1.
For person 1, it is not clear whether H_1^1 or H_3^1 is preferred more, and both are not clearly preferred more than H_u^1.
For person 3, all three experimental hypotheses are clearly not supported by the data.
For person 25, H_1^{25} is clearly more supported than the other two hypotheses.

Table 1. *Individual BFs*

i	BF_{1u}^i	BF_{2u}^i	BF_{3u}^i
1	1.30	0.87	1.89
2	2.23	1.42	4.12
3	0.04	0.10	0.56
⋮	⋮	⋮	⋮
24	25.52	4.36	7.01
25	24.43	5.25	3.66
26	0.09	0.12	0.33

Research goal 2. For each hypothesis it can be seen that the individual Bayes factors vary in strength and direction of support. Thus it seems unlikely that any hypothesis holds for all individuals.
Research goal 3. As described above, the individuals differ in which hypothesis (or multiple hypotheses) are supported most.

Multiple $N = 1$

All P individual Bayes factors could be multiplied, which expresses the support for H_m^i relative to H_u^i for every person i. This term is a function of P which makes it difficult to interpret. Stephan and Penny (2007) have suggested using the geometric mean of the product of individual Bayes factors for better interpretability:

$$\text{gP-BF}_{mm'} = \left(\prod_{i=1}^{P} \text{BF}_{mu}^i \right)^{\frac{1}{P}}, \tag{6}$$

which describes the 'average' support in favor of H_m^i found over P persons. It can be interpreted as the Bayes factor that is expected for a next individual.
The gP-BF can be influenced by outliers, which is why we propose an additional measure, the Evidence Rate (ER). The ER is a measure o fthe consistency in the preferred hypothesis over multiple individual Bayes factors:

$$ER = \begin{cases} \frac{1}{P} \sum_{i=1}^{P} I_{BF^i > 1} & \text{if gP-BF} < 1 \\ \frac{1}{P} \sum_{i=1}^{P} I_{BF^i < 1} & \text{if gP-BF} > 1 \end{cases}, \tag{7}$$

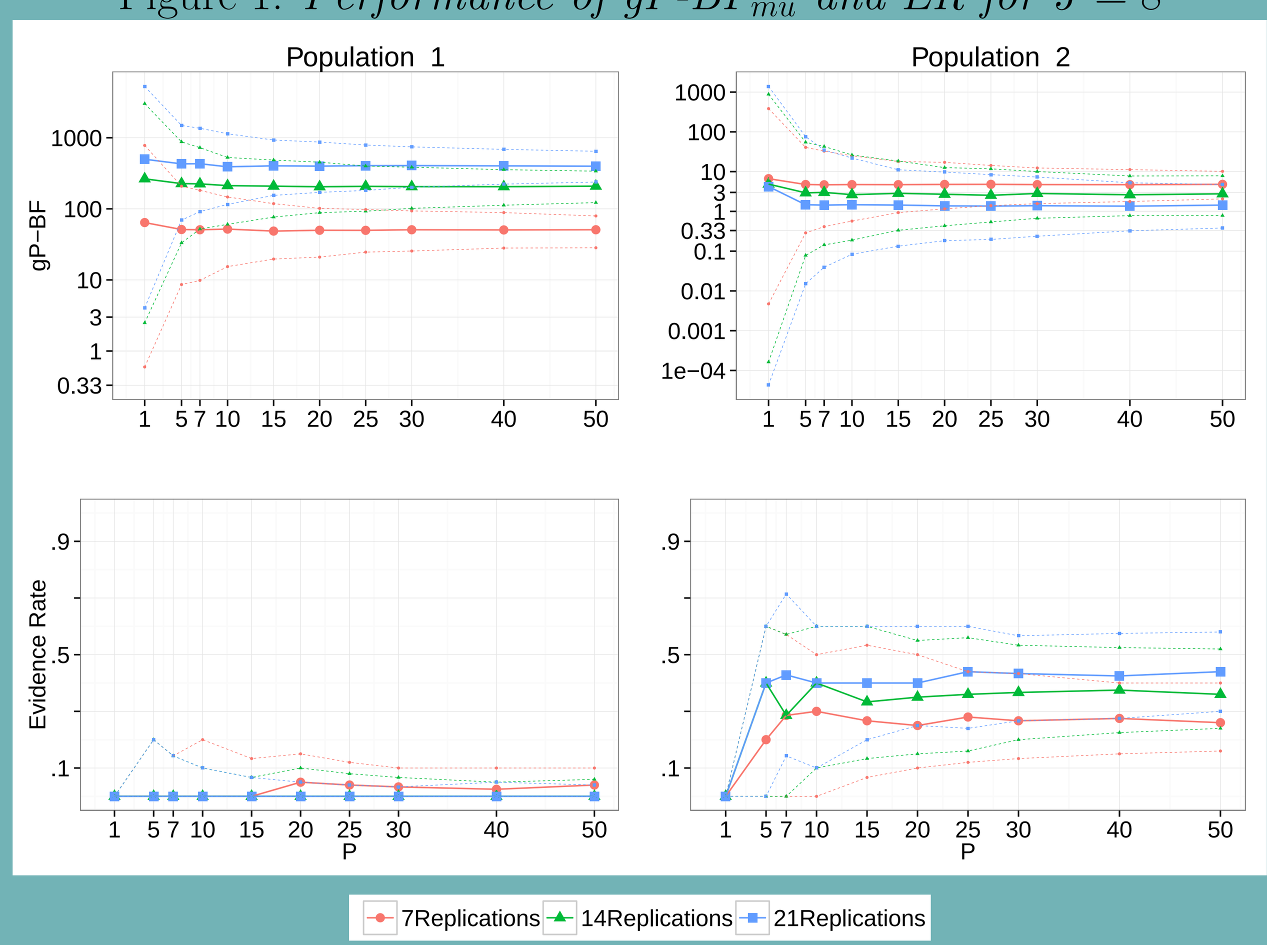
where $I_{BF^i > 1} = 1$ if $BF^i > 1$ and 0 otherwise.

Performance

Population 1 A population where all individuals agree with H_m^i.
All $\boldsymbol{\pi}^i$ are sampled from $h(\boldsymbol{\pi}^i|H_m^i) = h(\boldsymbol{\pi}^i|H_u^i) I_{\boldsymbol{\pi}^i \in H_m^i}$, where $I_{\boldsymbol{\pi}^i \in H_m^i} = 1$ if $\boldsymbol{\pi}^i \in H_m^i$ and 0 otherwise.

Population 2 A population where **not** all individuals agree with H_m^i.
All $\boldsymbol{\pi}^i$ are sampled from $\mathcal{N}(\log(\boldsymbol{\pi}), \mathbf{I}_J)$, where $\boldsymbol{\pi}$ is the average of all $\boldsymbol{\pi}^i$ from population 1, $\mathbf{I}_J$ is an identity matrix.

Figure 1. *Performance of gP-BF$_{mu}$ and ER for $J = 8$*



Conclusion

- Using individual Bayes factors, hypotheses can be evaluated for each person
- The gP-BF provides insight in the average evidence over multiple persons
- The Evidence Rate provides insight in whether subgroups exist
- The simulation shows that if subgroups do not exist, this is detected. If indeed a homogeneous population exists, this is reflected in both the gP-BF and the ER.

References

Hoijtink, H. (2012). *Informative Hypotheses. Theory and Practice for Behavioral and Social Scientists.* Boca Raton: Chapman & Hall/CRC.
Klugkist, I., Laudy, O., & Hoijtink, H. (2010). Bayesian evaluation of inequality and equality constrained hypotheses for contingency tables. *Psychological Methods*, *15*(3), 281-299.
Stephan, K. E., & Penny, W. D. (2007). Dynamic Causal Models and Bayesian Selection. In K. Friston, J. Ashbumer, S. Kievel, T. Nichols, & W. Penny (Eds.), *Statistical Parametric Mapping: The Analysis of Functional Brain Images* (p. 577-585). Academic Press.
Zedelius, C. M., Veling, H., & Aarts, H. (2011). Boosting or choking – How conscious and unconscious reward processing modulate the active maintenane of goal-relevant information. *Consciousness and Cognition*, *20*, 355-362.