bain: Bayesian informative hypotheses evaluation

H. Hoijtink[1] C. van Lissa[1] J. Mulder[2] X. Gu[3]

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

[1]Department of Methodology and Statistics
Utrecht University C.J.vanLissa@uu.nl

[2]Department of Methodology and Statistics
Tilburg University J.Mulder3@uvt.nl

[3]Department of Educational Psychology
East China Normal University GuXin57@hotmail.com

Outline

The Bayes Factor

The Bayes factor (Jeffreys, 1961, Kass and Raftery, 1995) is the ratio of two marginal likelihoods. The simplest example:

$$BF_{12} = \frac{m_1}{m_2} = \frac{\int_\mu f(x \mid \mu, \sigma^2 = 1) h(\mu \mid H_1) d\mu}{\int_\mu f(x \mid \mu, \sigma^2 = 1) h(\mu \mid H_2) d\mu},$$

where,

$$f(x \mid \mu, \sigma^2 = 1) = \prod_{i=1}^{N} -\frac{1}{\sqrt{2\pi}} \exp \frac{1}{2} (x_i - \mu)^2,$$

with $H_1 : \mu = 0$ and $H_2 : \mu \neq 0$ leading to

$$h(\mu \mid H_1) = I_{\mu=0},$$

and

$$h(\mu \mid H_2) \sim \mathcal{N}(0, \tau^2)$$

The Bayes Factor

The choice of the prior distribution is crucial when using the Bayes factor for hypothesis evaluation!

The "normal" could have been a "t" or a "Cauchy". All these choices appear in the literature.

Great "unanswered" research question: which choice is good, when, and why.

The Bayes Factor

Note that m_1 is simply

$$m_1 = f(x \mid \mu = 0, \sigma^2 = 1)$$

Note that a simple estimator to compute m_2 is

$$m_2 \approx \frac{1}{Q} \sum_{q=1}^{Q} f(x \mid \mu_q, \sigma^2 = 1),$$

where μ_q for $q = 1, ..., Q$ is sampled from $h(\mu \mid H_2) \sim \mathcal{N}(0, \tau^2)$

The Bayes Factor

Note that, the simple estimator is very inadequate (Kass and Raftery, 1995). Much better options to compute the Bayes factor are presented in Chib (1995) and Chib and Jeliazkov (2001). In fact, except in this introduction, never use the simple estimator.

The Bayes Factor

Using R code in which the computation of BF_{12} is implemented the following will be illustrated:

1. That BF_{12} is largest if $\overline{x} = 0$ and decreasing for $\overline{x} \rightarrow \infty$
2. That the size of BF_{12} depends on the sample size
3. That BF_{12} increases if the prior standard deviation is increased (Lindley-Bartlett paradox), that is, always choose a meaningful prior variance and never a vague prior

The One Group Approximate Adjusted Fractional Bayes Factor

1. Based on the work of Gu, Mulder, and Hoijtink (2018), Mulder (2014), O'Hagan (1995).
2. First, the one group situation is illustrated. The multiple group situation (Hoijtink, Gu, and Mulder, 2018, based on De Santis and Spezzaferri, 2001) will be illustrated next.

The One Group Approximate Adjusted Fractional Bayes Factor

Consider the following multiple regression model with two predictors that are measured on the same scale:

$$y_i = \alpha_0 + \alpha_1 x_{1i} + \alpha_2 x_{2i} + \epsilon_i \text{ with } \epsilon_i \sim \mathcal{N}(0, \sigma^2)$$

with density of the data

$$f(\boldsymbol{y} \mid \boldsymbol{\alpha}, \sigma^2, \boldsymbol{x}) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp -\frac{1}{2} \frac{(y_i - \alpha_0 - \alpha_1 x_{1i} - \alpha_2 x_{2i})^2}{\sigma^2}$$

Values of $\boldsymbol{\alpha}, \sigma^2$ supported by the data lead to a large value for $f(.)$. Values of $\boldsymbol{\alpha}, \sigma^2$ not supported by the data lead to a small value for $f(.)$.

The One Group Approximate Adjusted Fractional Bayes Factor

Using the maximum likelihood estimates $\hat{\alpha}_1$ and $\hat{\alpha}_2$ and the corresponding covariance matrix $\Sigma_{\alpha_1,\alpha_2}$ and a standard uninformative prior $h(\alpha_1, \alpha_2) = 1$ for α_1, α_2 a normal **approximation** of the posterior distribution of α_1, α_2 is

$$g_u(\alpha_1, \alpha_2 \mid \boldsymbol{y}, \boldsymbol{x}) \approx \mathcal{N}([\alpha_1, \alpha_2] \mid [\hat{\alpha}_1, \hat{\alpha}_2], \Sigma_{\alpha_1,\alpha_2}) \times 1 =$$

$$\mathcal{N}([\alpha_1, \alpha_2] \mid [\hat{\alpha}_1, \hat{\alpha}_2], \Sigma_{\alpha_1,\alpha_2})^{1-b} \times \mathcal{N}([\alpha_1, \alpha_2] \mid [\hat{\alpha}_1, \hat{\alpha}_2], \Sigma_{\alpha_1,\alpha_2})^{b} \times 1$$

Note that b denotes the "**fraction**" of the density of the data used to function as prior distribution (inspired by O'Hagan, 1995).

The One Group Approximate Adjusted Fractional Bayes Factor

The size of b is inspired by the concept of minimal training samples. Here $b = J/N$ where J denotes the number of independent constraints in the hypotheses under consideration.

Other choices are conceivable. Therefore, it is always recommended to do a sensitivity analysis (vary J) when hypotheses are formulated using equality constraints.

The **fractional** prior distribution then is:

$$h_u(\alpha_1, \alpha_2 \mid [\boldsymbol{y}, \boldsymbol{x}]^b) = \mathcal{N}([\alpha_1, \alpha_2] \mid [\hat{\alpha}_1, \hat{\alpha}_2], \Sigma_{\alpha_1,\alpha_2})^b \times 1 =$$

$$\mathcal{N}([\alpha_1, \alpha_2] \mid [\hat{\alpha}_1, \hat{\alpha}_2], \frac{\Sigma_{\alpha_1,\alpha_2}}{b})$$

The One Group Approximate Adjusted Fractional Bayes Factor

The prior distribution has to be **adjusted** such that its means are located on the boundaries of the hypotheses under consideration. Otherwise, as show earlier, the resulting Bayes factor will be inconsistent.

For the hypotheses used in this example this renders:

$$h_u(\alpha_1, \alpha_2 \mid [\boldsymbol{y}, \boldsymbol{x}]^b) = \mathcal{N}([\alpha_1, \alpha_2] \mid [0, 0], \frac{\Sigma_{\alpha_1,\alpha_2}}{b})$$

The One Group Approximate Adjusted Fractional Bayes Factor

Using this setup

$$BF_{iu} = \frac{m_i}{m_u} =$$

$$\frac{\int_{\alpha_1,\alpha_2} \mathcal{N}([\alpha_1,\alpha_2] \mid [\hat{\alpha}_1,\hat{\alpha}_2], \frac{\Sigma_{\alpha_1,\alpha_2}}{1-b}) h_i(\alpha_1,\alpha_2 \mid [\boldsymbol{y},\boldsymbol{x}]^b) d\alpha_1,\alpha_2}{\int_{\alpha_1,\alpha_2} \mathcal{N}([\alpha_1,\alpha_2] \mid [\hat{\alpha}_1,\hat{\alpha}_2], \frac{\Sigma_{\alpha_1,\alpha_2}}{1-b}) h_u(\alpha_1,\alpha_2 \mid [\boldsymbol{y},\boldsymbol{x}]^b) d\alpha_1,\alpha_2},$$

where

$$h_i(\alpha_1,\alpha_2 \mid [\boldsymbol{y},\boldsymbol{x}]^b) = \frac{h_u(\alpha_1,\alpha_2 \mid [\boldsymbol{y},\boldsymbol{x}]^b) I_{\alpha_1,\alpha_2 \in H_i}}{\int_{\alpha_1,\alpha_2} h_u(\alpha_1,\alpha_2 \mid [\boldsymbol{y},\boldsymbol{x}]^b) I_{\alpha_1,\alpha_2 \in H_i} d\alpha_1,\alpha_2}$$

The One Group Approximate Adjusted Fractional Bayes Factor

Note that,

$$g_i(\alpha_1, \alpha_2 \mid \boldsymbol{y}, \boldsymbol{x}) = \frac{f(\boldsymbol{y} \mid \boldsymbol{\alpha}, \sigma^2, \boldsymbol{x}) h_i(\alpha_1, \alpha_2 \mid [\boldsymbol{y}, \boldsymbol{x}]^b)}{m_i}$$

and

$$g_u(\alpha_1, \alpha_2 \mid \boldsymbol{y}, \boldsymbol{x}) = \frac{f(\boldsymbol{y} \mid \boldsymbol{\alpha}, \sigma^2, \boldsymbol{x}) h_u(\alpha_1, \alpha_2 \mid [\boldsymbol{y}, \boldsymbol{x}]^b)}{m_u}$$

The One Group Approximate Adjusted Fractional Bayes Factor

Therefore

$$BF_{iu} = \frac{m_i}{m_u} =$$

$$\frac{f(\boldsymbol{y} \mid \boldsymbol{\alpha}, \sigma^2, \boldsymbol{x}) h_i(\alpha_1, \alpha_2 \mid [\boldsymbol{y}, \boldsymbol{x}]^b)}{g_i(\alpha_1, \alpha_2 \mid \boldsymbol{y}, \boldsymbol{x})} / \frac{f(\boldsymbol{y} \mid \boldsymbol{\alpha}, \sigma^2, \boldsymbol{x}) h_u(\alpha_1, \alpha_2 \mid [\boldsymbol{y}, \boldsymbol{x}]^b)}{g_u(\alpha_1, \alpha_2 \mid \boldsymbol{y}, \boldsymbol{x})} =$$

$$\frac{1/c_i \times h_u(\alpha_1, \alpha_2 \mid [\boldsymbol{y}, \boldsymbol{x}]^b)}{1/f_i \times g_u(\alpha_1, \alpha_2 \mid \boldsymbol{y}, \boldsymbol{x})} / \frac{h_u(\alpha_1, \alpha_2 \mid [\boldsymbol{y}, \boldsymbol{x}]^b)}{g_u(\alpha_1, \alpha_2 \mid \boldsymbol{y}, \boldsymbol{x})} = \frac{f_i}{c_i},$$

The One Group Approximate Adjusted Fractional Bayes Factor

where

$$f_i = \int_{\alpha_1,\alpha_2 \in H_i} \mathcal{N}([\alpha_1, \alpha_2] \mid [\hat{\alpha}_1, \hat{\alpha}_2], \Sigma_{\alpha_1,\alpha_2}) d\alpha_1, \alpha_2$$

and

$$c_i = \int_{\alpha_1,\alpha_2 \in H_i} \mathcal{N}([\alpha_1, \alpha_2] \mid [0, 0], \frac{\Sigma_{\alpha_1,\alpha_2}}{b}) d\alpha_1, \alpha_2$$

cf. a few slides back were $h_i() = h_u()/c_i$.

The One Group Approximate Adjusted Fractional Bayes Factor

Consider the hypotheses:

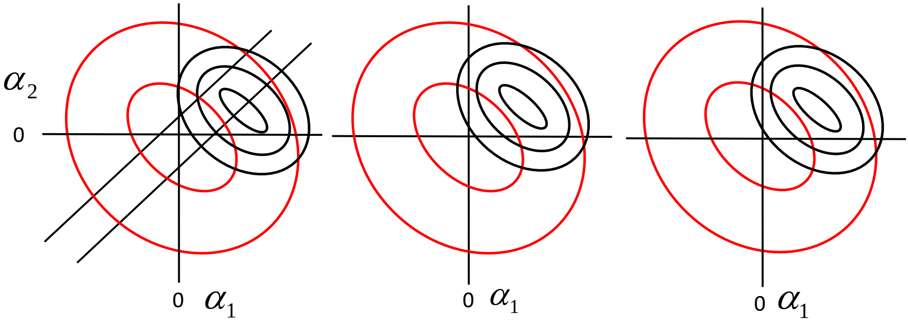
$$H_1 : \alpha_1 \approx \alpha_2, \text{ that is, } |\alpha_1 - \alpha_2| < .1$$

$$H_2 : \alpha_1 > 0, \alpha_2 > 0$$

$$H_3 : \alpha_1, \alpha_2$$

These hypotheses can be "translated" into prior distributions, posterior distributions can be superimposed and Bayes factors can "visually" be computed.

$$H_1 : \alpha_1 \approx \alpha_2 \qquad H_2 : \alpha_1 > 0, \alpha_2 > 0, \qquad H_u : \alpha_1, \alpha_2$$



$$BF_{1u} = \frac{f_1}{c_1} = \frac{.3}{.2} = 1.5 \qquad BF_{2u} = \frac{f_2}{c_2} = \frac{.95}{.25} = 3.8$$

$$BF_{21} = \frac{3.8}{1.5} = 2.5$$

The One Group Approximate Adjusted Fractional Bayes Factor

Note that, the computation of c_i and f_i is based on sampling from the posterior and prior distributions (the proportions of parameter vectors sampled in agreement with H_i.

This can be done easily and accurately with the algorithm presented in Gu, Hoijtink, Mulder, and Rosseel (unpublished).

The Multiple Group Approximate Adjusted Fractional Bayes Factor

Consider a "Students t-test" setup:

$$y_{ig} = \mu_g + \epsilon_i \text{ for } g = 1, 2, \epsilon_i \sim \mathcal{N}(0, \sigma^2),$$

with $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$ and sample sizes N_1 and N_2.

The hypotheses of interest are:

$H_1 : \mu_1 = \mu_2$

and

$H_u : \mu_1 \neq \mu_2$

The Multiple Group Approximate Adjusted Fractional Bayes Factor

Assuming that $\hat{\sigma}^2 = 1$, the normal approximation of the posterior distribution of μ_1 and μ_2 is:

$$g_u(\mu_1, \mu_2 \mid \boldsymbol{y}) \approx \mathcal{N}\left(\begin{bmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \end{bmatrix}, \begin{bmatrix} 1/N_1 & 0 \\ 0 & 1/N_2 \end{bmatrix}\right),$$

The Multiple Group Approximate Adjusted Fractional Bayes Factor

The "fractional" prior corresponding to the posterior is:

$$h_u(\mu_1, \mu_2 \mid [\boldsymbol{y}]^b) = \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/b \times 1/N_1 & 0 \\ 0 & 1/b \times 1/N_2 \end{bmatrix}\right),$$

which has

1. means on the boundary of the two hypotheses of interest
2. $b = J/N = 1/N$, where J denotes the number of constraints used to formulate H_1 and H_u

The Multiple Group Approximate Adjusted Fractional Bayes Factor

Using $\delta = \mu_1 - \mu_2$, the BF reduces to the Savage-Dickey density ratio

$$\mathrm{BF}_{1u} = \frac{f_1}{c_1} = \frac{g_u(\delta = 0 \mid \boldsymbol{y})}{h_u(\delta = 0 \mid [\boldsymbol{y}]^b)} = \frac{\mathcal{N}(0 \mid \hat{\delta}, N_1^{-1} + N_2^{-1})}{\mathcal{N}(0 \mid 0, (N_1 + N_2)(N_1^{-1} + N_2^{-1}))}.$$

The Multiple Group Approximate Adjusted Fractional Bayes Factor

Lets us first of all consider the situation in which N_1 and N_2 go to ∞ with the same rate, that is, let $N_g = a_g n$, for some positive constant a_g and let $n \rightarrow \infty$.

If $\hat{\delta} = 0$, then

$$f_1 = \mathcal{N}(0 \mid 0, N_1^{-1} + N_2^{-1}) \rightarrow \infty,$$

and

$$c_1 = \mathcal{N}(0 \mid 0, (N_1+N_2)(N_1^{-1}+N_2^{-1}) = \mathcal{N}(0 \mid 0, (a_1+a_2)(a_1^{-1}+a_2^{-1}),$$

which is a constant independent of n.

Consequently, $\text{BF}_{1u} \rightarrow \infty$ if $n \rightarrow \infty$ which is consistent.

The Multiple Group Approximate Adjusted Fractional Bayes Factor

If $\hat{\delta} \neq 0$,

$$f_1 = \mathcal{N}(0 \mid \hat{\delta}, N_1^{-1} + N_2^{-1}) \rightarrow 0,$$

if $n \rightarrow \infty$

and c_1 remains a constant

This implies that $\text{BF}_{1u} \rightarrow 0$ if $n \rightarrow \infty$ which is consistent.

The Multiple Group Approximate Adjusted Fractional Bayes Factor

A numerical illustration:

In the table for H_1, $\bar{\mathbf{x}}_1 = (0,0)$ and for H_u, $\bar{\mathbf{x}}_2 = (-.35, .35)$

		(M)BF				
					$\bar{\mathbf{x}}_1$	$\bar{\mathbf{x}}_2$
N_1	N_2	b	$\hat{\mathbf{\Sigma}}_{11}/b$	$\hat{\mathbf{\Sigma}}_{22}/b$	BF_{1u}	BF_{1u}
10	10	.05	2	2	4.47	*1.31*
25	25	.02	2	2	7.07	*.33*
50	50	.01	2	2	10.00	*.02*
100	100	.005	2	2	14.14	*.00*

The Multiple Group Approximate Adjusted Fractional Bayes Factor

Now if we fix N_1 and let $N_2 \to \infty$, then if $n \to \infty$

$$f_1 = \mathcal{N}(0 \mid \hat{\delta}, N_1^{-1} + N_2^{-1}) = \mathcal{N}(0 \mid \hat{\delta}, N_1^{-1}),$$

which is a constant, and

$$c_1 = \mathcal{N}(0 \mid 0, (N_1 + N_2)(N_1^{-1} + N_2^{-1}) = \mathcal{N}(0 \mid 0, \infty) \to 0,$$

Consequently, in the limit $BF_{1u} \to \infty$ also if H_u is true, which is inconsistent behavior.

The Multiple Group Approximate Adjusted Fractional Bayes Factor

A numerical illustration:

In the table for H_1, $\bar{\mathbf{x}}_1 = (0, 0)$ and for H_u, $\bar{\mathbf{x}}_2 = (-.35, .35)$

		BF				
					$\bar{\mathbf{x}}_1$	$\bar{\mathbf{x}}_2$
N_1	N_2	b	$\hat{\mathbf{\Sigma}}_{11}/b$	$\hat{\mathbf{\Sigma}}_{22}/b$	BF_{1u}	BF_{1u}
10	10	.05	2	2	4.47	*1.31*
10	25	.029	3.5	1.4	5.92	*1.03*
10	50	.017	6.0	1.2	7.74	*1.01*
10	100	.009	11	1.13	10.48	*1.13*
10	200	.005	21	1.05	14.94	*1.41*
10	1000	.001	101	1.01	31.78	*2.81*

The Multiple Group Approximate Adjusted Fractional Bayes Factor

This inconsistent behavior can be avoided if the fraction of information b used to specify the prior distribution is group specific:

$$h_u(\mu_1, \mu_2 \mid [\mathbf{y}]^{\mathbf{b}}) \approx \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/b_1 \times 1/N_1 & 0 \\ 0 & 1/b_2 \times 1/N_2 \end{bmatrix}\right) =$$

$$\mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\right),$$

where $b_1 = \frac{1}{2}\frac{1}{N_1}$ and $b_2 = \frac{1}{2}\frac{1}{N_2}$.

The Multiple Group Approximate Adjusted Fractional Bayes Factor

The MGAAFBF (what's in a name) based on this prior is:

$$MBF_{1u} = \frac{f_1}{c_1} = \frac{g_u(\delta = 0 \mid \boldsymbol{y})}{h_u(\delta = 0 \mid [\boldsymbol{y}]^{\boldsymbol{b}})} = \frac{\mathcal{N}(0 \mid \hat{\delta}, N_1^{-1} + N_2^{-1})}{\mathcal{N}(0 \mid 0, 4)} =$$

The Multiple Group Approximate Adjusted Fractional Bayes Factor

If $\hat{\delta} = 0$ and $n \rightarrow \infty$, then $f_1 \rightarrow \infty$ and c_1 is constant. This implies that $MBF_{1u} \rightarrow \infty$.

If $\hat{\delta} \neq 0$ and $n \rightarrow \infty$, then $f_1 \rightarrow 0$ and c_1 is constant. This implies that $MBF_{1u} \rightarrow 0$.

Stated otherwise, for $n \rightarrow \infty$ MBF_{1u} is consistent.

The Multiple Group Approximate Adjusted Fractional Bayes Factor

If $N_2 \rightarrow \infty$ while N_1 is fixed:

Then if $\hat{\delta} = 0$, in the limit

$$MBF_{1u} = \frac{f_1}{c_1} = \frac{\mathcal{N}(0 \mid 0, N_1^{-1} + N_2^{-1})}{\mathcal{N}(0 \mid 0, 4)} = \frac{\mathcal{N}(0 \mid 0, N_1^{-1})}{\mathcal{N}(0 \mid 0, 4)}$$

Although for $N_2 \rightarrow \infty$ MBF_{1u} does not approach ∞ (but for $N_1 \geq 1$ is has a value larger than 1), this is reasonable behavior and the inconsistent behavior of the BF is avoided.

The Multiple Group Approximate Adjusted Fractional Bayes Factor

If $N_2 \to \infty$ while N_1 is fixed:

Then if $\hat{\delta} \neq 0$ in the limit

$$MBF_{1u} = \frac{f_1}{c_1} = \frac{\mathcal{N}(0 \mid \hat{\delta}, N_1^{-1} + N_2^{-1})}{\mathcal{N}(0 \mid 0, 4)} = \frac{\mathcal{N}(0 \mid \hat{\delta}, N_1^{-1})}{\mathcal{N}(0 \mid 0, 4)}$$

If, for example, $N_1 = 25$ and $\hat{\delta} = .1$, $\text{MBF}_{1u} = 8.8$, that is, H_1 is supported. This is reasonable, because both the sample size of Group 1 and the effect size are small and therefore the effect is not convincingly different from zero.

If both are larger, for example, $N_1 = 49$ and $\hat{\delta} = .5$, $\text{MBF}_{1u} = .03$, that is, H_u is supported.

The Multiple Group Approximate Adjusted Fractional Bayes Factor

A numerical illustration:

In the table for H_1, $\bar{\mathbf{x}}_1 = (0,0)$ and for H_u, $\bar{\mathbf{x}}_2 = (-.35, .35)$

		MBF					
						$\bar{\mathbf{x}}_1$	$\bar{\mathbf{x}}_2$
N_1	N_2	b_1	b_2	$\hat{\boldsymbol{\Sigma}}_{11}/b_1$	$\hat{\boldsymbol{\Sigma}}_{22}/b_2$	MBF_{1u}	MBF_{1u}
10	10	.05	.05	2	2	4.47	1.31
10	25	.05	.02	2	2	5.34	.93
10	50	.05	.01	2	2	5.77	.75
10	100	.05	.005	2	2	6.03	.65
10	200	.05	.0025	2	2	6.17	*.60*
10	1000	.05	.0005	2	2	6.29	.56

The Multiple Group Approximate Adjusted Fractional Bayes Factor

The normal approximation used in the (M)BF makes many things easy:

1. Proofs of consistency
2. Creating a Bayes factor that is robust against outliers
3. As will be elaborated now, dealing with missing data

Computing the Bayes Factor from Data with Missing Values

This presentation will be illustrated using a normal linear regression model:

$$y_i = \alpha_0 + \alpha_1 x_{i1} + \alpha_2 x_{i2} + e_i,$$

The goal is to compute the Bayes factor BF_{ku} where:

$$H_k : \alpha_1 > 0, \alpha_2 > 0, \alpha_1 > \alpha_2$$

and,

$$H_u : \alpha_1, \alpha_2$$

in the presence of missing values for the y_i, x_{i1}, x_{i2}

Computing the Bayes Factor from Data with Missing Values

The Bayes factor implemented in `Bain` can be written as:

$$BF_{ku} = \frac{f_k}{c_k},$$

were f_k denotes the fit of H_k, that is, using $\alpha = [\alpha_1, \alpha_2]$

$$f_k = \int_{\alpha \in H_k} g_u(\alpha \mid \; Y_o, X_o) d\alpha$$

The complexity of H_k is denoted by c_k, that is,

$$c_k = \int_{\alpha \in H_k} h_u(\alpha) d\alpha$$

Computing the Bayes Factor from Data with Missing Values

`Bain` uses a normal approximation of $g_u(.)$:

$$g_u(\alpha \mid Y_o, X_o) \approx \mathcal{N}(\alpha \mid \hat{\alpha}, \Sigma_\alpha),$$

from which the prior distribution (fractional Bayes factor approach) is derived as:

$$h_u(\alpha) = \mathcal{N}(\alpha \mid [0,0], \Sigma_\alpha / b)$$

where, for the example at hand, the fraction of information from the data used to construct the prior equals $b = 2/N$.

Computing the Bayes Factor from Data with Missing Values

This unique setup enables the computation of Bayes factors in the presence of missing data. All that is needed is:

1. $\hat{\alpha}$ and Σ_α computed from data with missing values
2. An estimate N_{eff} of the effective number of persons in the data, that is, the N that accounts for the missing values.

Computing the Bayes Factor from Data with Missing Values

Both 1. and 2. can be obtained using multiple imputation as, for example, implemented in the R package `mice`. This entails, among other things:

1. that auxiliary variables (not part of the regression model) can be used (to achieve MAR)
2. that $\hat{\alpha}$ and Σ_α are provided
3. that the fraction of missing information λ is provided, which renders $N_{eff} = (1 - \lambda)N$

Computing the Bayes Factor from Data with Missing Values

A simple illustration of the effectiveness of this approach.

Computing the Bayes Factor from Data with Missing Values

$$x_i \sim \mathcal{N}(\gamma, \omega) \text{ for } i = 1, ..., N,$$

$H_1 : \gamma = 0$ and $H_2 : \gamma > 0$ and $H_u : \gamma$.

Data will be generated such that the average of x_o is $\in \{-.2, 0, .2, .5\}$, and the sample standard deviation $s = 1$.

$N_o = 30$ and $N_m = 20$ missing values are added (MCAR).

This setup allows for inference with (that is, based on x_o, x_m) and without (that is, based on x_o) imputation. Both should render the same Bayes factors.

Computing the Bayes Factor from Data with Missing Values

Bayes Factors Computed Using only the Observed Values and Multiple Imputation with the Package `Bain`

	$\overline{x}$	$\overline{\gamma}$	Σ_γ	λ	BF_{1u}	BF_{2u}
observed	-.2	-0.20	.03		3.01	.27
imputed	-.2	-0.20	.03	.45	3.03	.29
observed	0	.00	.03		5.47	1
imputed	0	.00	.03	.45	5.26	1
observed	.2	.20	.03		3.01	1.72
imputed	.2	.20	.03	.45	2.87	1.73
observed	.5	.50	.03		.13	1.99
imputed	.5	.50	.03	.45	.13	1.99

Computing the Bayes Factor from Data with Missing Values

Conclusion

1. The approach presented can be used in the context of many statistical models and for informative hypotheses constructed using linear equality and inequality constraints among the parameters of interest
2. The approach presented can also be used for those modules of the `BayesFactor` R package that have prior distributions are independent of the data (t-test, ANOVA): compute Bayes factors for each imputed data set and compute their average.
3. For the technical details, elaborations, and examples, you are referred to Hoijtink, Mulder, Gu, and Rosseel (2019).

References

1. Chib, S. (1995). Marginal likelihood from the Gibbs output. *Journal of the American Statistical Association, 90*, 1313-1321.
2. Chib, S. and Jeliazkov, I. (2001). Marginal likelihood from the Metropolis-Hastings output. *Journal of the American Statistical Association, 96*, 270-281.
3. De Santis, F. and Spezzaferri, F. (2001). Consistent fractional Bayes factor for nested normal linear models. *Journal of Statistical Planning and Inference, 97,*, 305-321.
4. Gu, X., Mulder, J., and Hoijtink, H. (2018). Approximate adjusted fractional Bayes factors: A general method for testing informative hypotheses. *British Journal of Mathematical and Statistical Psychology, 71,* 229-261. DOI: 10.1111/bmsp.12110
5. Gu, X., Hoijtink, H., Mulder, J., and Rosseel, Y. (unpublished). Bain: A program for Bayesian testing of order constrained hypotheses in structural displaymath models.
6. Hoijtink, H., Mulder, J., van Lissa, C., and Gu, X. (2018). A tutorial on testing hypotheses using the Bayes factor. *Psychological Methods.* DOI: 10.1037/met0000201

References

7. Hoijtink, H., Gu, X., and Mulder, J. (2018). Bayesian evaluation of informative hypotheses for multiple populations. *British Journal of Mathematical and Statistical Psychology.* DOI: 10.1111/bmsp.12145
8. Hoijtink, H., Gu, X., Mulder, J., and Rosseel, Y. (2019). Computing Bayes Factors from Data with Missing Values. *Psychological Methods.* DOI: 10.1037/met0000187
9. Mulder, J. (2014). Prior adjusted default Bayes factors for testing (in)equality constrained hypotheses. *Computational Statistics and Data Analysis, 71*, 448-463.
10. O'Hagan, A. (1995). Fractional Bayes factors for model comparison (with discussion). *Journal of the Royal Statistical Society. Series B, 57*, 99-138.