Optional Stopping with Bayes factor

Seo Wook Choi

[1]Department of Psychology

Study Motivation

- An optional stopping of collecting data until a researcher gets satisfying result can be interpreted in two ways, an effort or a cheating.
- Is there any possible standard procedure for researchers to use optional stopping?
- Thanks to Dr. DeBoeck for a candidacy exam question.

Sequential probability ratio test

- Wald's sequential probability ratio test.

$$H_0 : p = p_0 \tag{1}$$

$$H_1 : p = p_1 \tag{2}$$

$$S_i = S_{i-1} + log(LR_i) \tag{3}$$

$$if \;\; a < S_i < b, \;\; Keep \;\; Running \tag{4}$$

$$if \;\; S_i \geq b, \;\; Accept \;\; H_1 \tag{5}$$

$$if \;\; S_i \leq a, \;\; Accept \;\; H_0 \tag{6}$$

Optional stopping criterion

- Different behaviors of Bayes factor and p-value (Rouder et al. (2009))
- Which seems to the be more conservative criterion? BF=3 or p-value=0.05?
- Possible to accept null using BF.

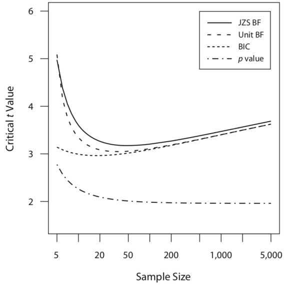
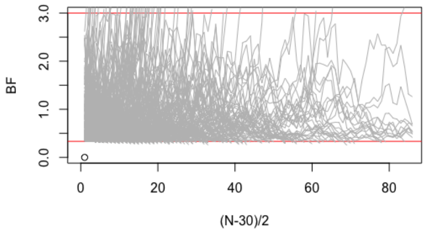


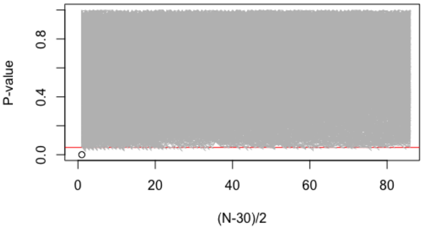
Figure 4. Critical *t* values needed for posterior odds of 10:1 favoring the alternative for the JZS Bayes factor (solid line), the unit-information Bayes factor (longer dashed line), and the BIC (shorter dashed line), as well as critical *t* values needed for $p < .05$ (dashed-and-dotted line).

T-test

- H_0:$\mu = 0$, H_1:$\mu \neq 0$. Data is sampled from $N(\mu, 1)$. For H_1, $\mu = X$.
- Starting with 30 samples, add 1 more subjects to each condition if the current samples do not meet the criterion.
- Using criteria BF=3 and 1/3, p-value < 0.05
- Bayes factor is calculated using R package Bayesfactor using r=0.707 for Cauchy prior parameter. Meaning: a probability of standardized effect size is greater than r is 1/2.
- This procedure is very similar to Schonbrodt et al. (2017).

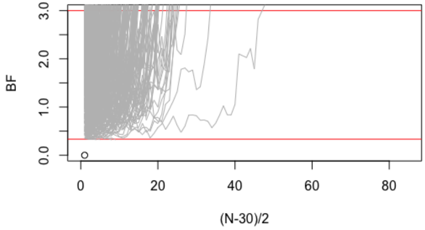
Samples from H_0

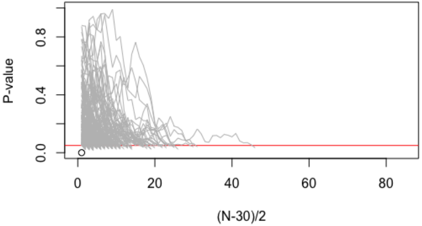




- Null: 892, Alternative: 92
- Alternative: 269

Samples from H_1



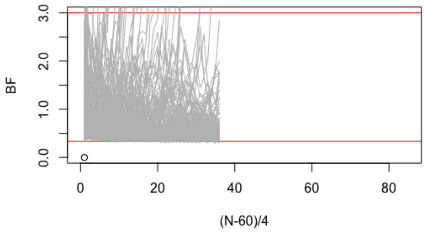


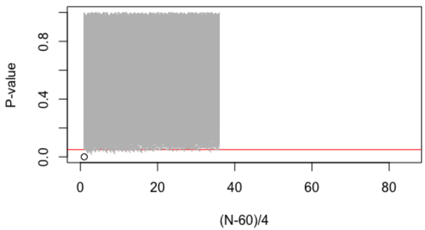
- Null: 8, Alternative: 992
- Alternative: 1000

Interaction effect of two-way ANOVA

- H_0:$\beta_{12} = 0$, H_1:$\beta_{12} \neq 0$. Data is sampled from $N(\beta_{12}, 1)$. For H_1, $\beta_{12} = X_1 X_2$.
- Starting with 60 samples, add 1 more subjects to each condition (total of 4) if the current samples do not meet the criterion.
- P-values for a coefficient of the $X_1 X_2$ term is used.
- Bayes factors are calculated by $\mathrm{BF}(Y \sim X_1 + X_2 + X_1 X_2)/\mathrm{BF}(Y \sim X_1 + X_2)$.

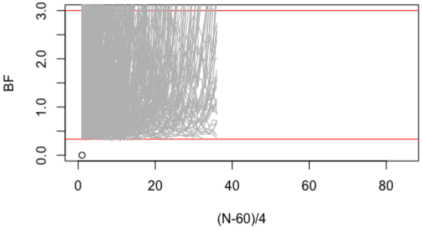
Samples from H_0

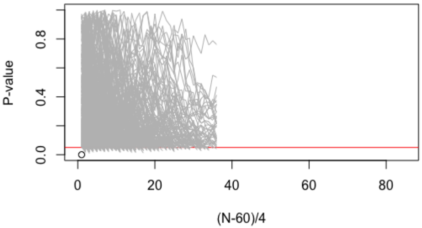




- Null: 821, Alternative: 60
- Alternative: 166

Samples from H_1



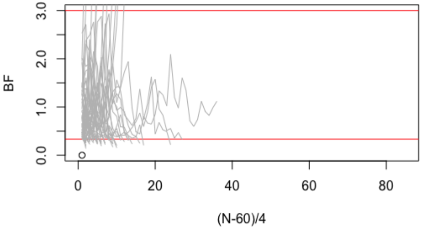


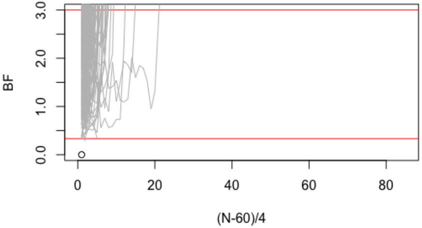
- Null: 177, Alternative: 798
- Alternative: 965

Bayesian Analysis

- Doing the optional stopping procedure with Bayesian analysis.
- For simple calculation of BF, approximate Bayes factor is used ($BF_{10} \approx exp(\frac{BIC_0 - BIC_1}{2})$).
- $Y \sim N(\mu, \sigma)$, $\mu = \beta X_1 X_2$, $\beta \sim N(0, 100)$, $\sigma \sim CU(0, 100)$.
- $\beta = 0$ and $\beta = 1$ for data generation.

Samples from H_0 and H_1

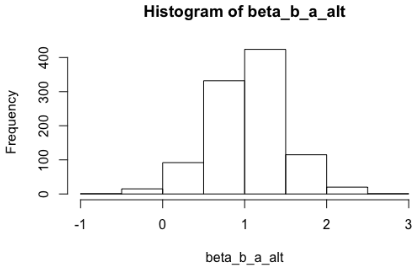




- Null: 979, Alternative: 20
- Null: 28 Alternative: 972

Estimation bias

- Optional stopping procedures tend to cause slight overestimation of the coefficients.
- With different seeds, mean of estimated coefficient tends to be overestimated. The bias is not huge, but there is a consistent pattern of bias.



Conclusion

- BF seems to be a good criterion for an optional stopping procedure.
- Approximate BF seems to be a possible way of setting optional stopping criterion for Bayesian analysis.
- Which level of BF to use as a criterion is still in question.
- Bias in the estimate can be a potential problem.