Closed-Form Power and Sample Size Calculations for Bayes Factors
Supplementary Material

Samuel Pawel **Leonhard Held**

Epidemiology, Biostatistics and Prevention Institute (EBPI)
Center for Reproducible Science (CRS)
University of Zurich
E-mail: samuel.pawel@uzh.ch

1 Distribution of the Bayes factor

The Bayes factor (2) with point analysis prior ($\tau = 0$) can be rewritten as

$$\mathrm{BF}_{01} = \exp\left[\frac{n}{\sigma^2_{\hat{\theta}}}\left\{\hat{\theta}(\theta_0 - \mu) - \frac{\theta_0^2 - \mu^2}{2}\right\}\right]. \tag{1}$$

Suppose that compelling evidence for H_1 is achieved when $\mathrm{BF}_{01} \leq k$. In this case, $\mathrm{BF}_{01} \leq k$ can be rewritten as

$$\hat{\theta}(\theta_0 - \mu) \leq \frac{\sigma^2_{\hat{\theta}} \log k}{n} + \frac{\theta_0^2 - \mu^2}{2}.$$

Dividing by $(\theta_0 - \mu)$ changes the inequality if $\mu > \theta_0$. We then have that under a normal distribution $\hat{\theta} \mid n, \mu_d, \tau_d \sim \mathrm{N}(\mu_d, \tau_d^2 + \sigma^2_{\hat{\theta}}/n)$, the probability of compelling evidence is given by (3).

The Bayes factor (2) with normal analysis prior ($\tau > 0$) can be rewritten as

$$\mathrm{BF}_{01} = \sqrt{1 + \frac{n\tau^2}{\sigma^2_{\hat{\theta}}}}\exp\left(-\frac{1}{2}\left[\frac{\{\hat{\theta} - \theta_0 - \frac{\sigma^2_{\hat{\theta}}}{n\tau^2}(\theta_0 - \mu)\}^2}{\frac{\sigma^2_{\hat{\theta}}}{n}(1 + \frac{\sigma^2_{\hat{\theta}}}{n\tau^2})} - \frac{(\theta_0 - \mu)^2}{\tau^2}\right]\right). \tag{2}$$

Suppose that compelling evidence for H_1 is achieved when $\mathrm{BF}_{01} \leq k$, which can be rearranged to

$$\left\{\hat{\theta} - \theta_0 - \frac{\sigma^2_{\hat{\theta}}}{n\tau^2}(\theta_0 - \mu)\right\}^2 \geq \left\{\log\left(1 + \frac{n\tau^2}{\sigma^2_{\hat{\theta}}}\right) + \frac{(\theta_0 - \mu)^2}{\tau^2} - \log k^2\right\}\left(1 + \frac{\sigma^2_{\hat{\theta}}}{n\tau^2}\right)\frac{\sigma^2_{\hat{\theta}}}{n}.$$

Therefore, under a normal distribution $\hat{\theta} \mid n, \mu_d, \tau_d \sim \mathrm{N}(\mu_d, \tau_d^2 + \sigma^2_{\hat{\theta}}/n)$, the probability of compelling evidence is given by (6).

2 Limiting power of Bayes factor with normal analysis prior

We have that

$$\lim_{n \to \infty} M = \frac{\mu_d - \theta_0}{\tau_d}$$

and

$$\lim_{n\to\infty} X = \lim_{n\to\infty}\left[\left\{\log\left(1+\frac{n\tau^2}{\sigma^2_{\hat{\theta}}}\right)+\frac{(\theta_0-\mu)^2}{\tau^2}-\log k^2\right\}\frac{\sigma^2_{\hat{\theta}}}{n\tau_d^2+\sigma^2_{\hat{\theta}}}\right].$$

Thus, when also $\tau_d \downarrow 0$ and $\mu_d \neq \theta_0$, both M and X diverge but the M term diverges faster than the X term. When $\tau_d > 0$, the M term approaches a constant while the X term approaches zero. Consequently, in both cases it holds that

$$\lim_{n\to\infty} \Pr(\mathrm{BF}_{01} \leq k \mid n, \mu_d, \tau_d, \tau > 0) = \lim_{n\to\infty}\left\{\Phi(-\sqrt{X}-M)+\Phi(-\sqrt{X}+M)\right\} = 1.$$

3 Sample size for Bayes factor with local normal prior

Equating the power function (11) to $1-\beta$ and applying algebraic manipulations, we have that

$$\begin{aligned} z^2_{(1-\beta)/2} &= \left\{\log\left(1+\frac{n\tau^2}{\sigma^2_{\hat{\theta}}}\right)-\log k^2\right\}\frac{\sigma^2_{\hat{\theta}}}{n\tau^2} \\ &\approx \left\{\log\left(\frac{n\tau^2}{\sigma^2_{\hat{\theta}}}\right)-\log k^2\right\}\frac{\sigma^2_{\hat{\theta}}}{n\tau^2} \\ &= \log\left(\frac{n\tau^2}{\sigma^2_{\hat{\theta}}k^2}\right)\frac{\sigma^2_{\hat{\theta}}}{n\tau^2} \end{aligned}$$

Multiplying by $-k^2$ and rewriting the second factor on the right-hand-side as exponential leads to

$$-k^2\, z^2_{(1-\beta)/2} = -\log\left(\frac{n\tau^2}{\sigma^2_{\hat{\theta}}k^2}\right)\exp\left\{-\log\left(\frac{n\tau^2}{\sigma^2_{\hat{\theta}}k^2}\right)\right\}.$$

Hence, we can apply the Lambert W function to obtain

$$-\log\left(\frac{n\tau^2}{\sigma^2_{\hat{\theta}}k^2}\right) = \mathrm{W}\left(-k^2\, z^2_{(1-\beta)/2}\right)$$

from which we obtain the sample size

$$n = \frac{\sigma^2_{\hat{\theta}}}{\tau^2}\,k^2\,\exp\left\{-\mathrm{W}\left(-k^2\, z^2_{(1-\beta)/2}\right)\right\}.$$

For arguments $y \in (-1/e, 0)$, the Lambert W function has two branches. The sample size is obtained from the branch commonly denoted as $\mathrm{W}_{-1}(\cdot)$ which satisfies $\mathrm{W}(x) < -1$ for $y \in (-1/e, 0)$ (Corless et al., 1996). This is because this branch always leads to larger sample sizes than the other and guarantees that unit information sample sizes are always larger than one.

4 Simulation-based evaluation of the bfpwr package

Figure 1 shows a simulation-based evaluation of the power and sample size calculation methods for the Bayesian z-test as implemented in our `bfpwr` R package. The values for the design and analysis prior means were chosen to represent conventions for no (0), small (0.2), medium (0.5), and large (0.8) standardized mean differences (Cohen, 1992). The null and alternative hypotheses were defined as $H_0\colon \theta = 0$ against $H_1\colon \theta \neq 0$. The standard deviations were chosen to include point and normal priors. For each combination of analysis/design prior mean/standard deviation, the sample size to obtain a Bayes factor equal or below $k = 1/10$ with a target power of 80% was computed (shown at the top of each plot). This sample size along with the design prior was subsequently used to simulate 50'000 standardized mean difference parameter estimates based on which 50'000 Bayes factors were computed. The power was then estimated from the proportion of Bayes factors equal or below the level $k = 1/10$. Note that for certain design/analysis prior combinations, it is impossible to achieve the target power with a finite sample size. In this case an "x" is shown in the plot.

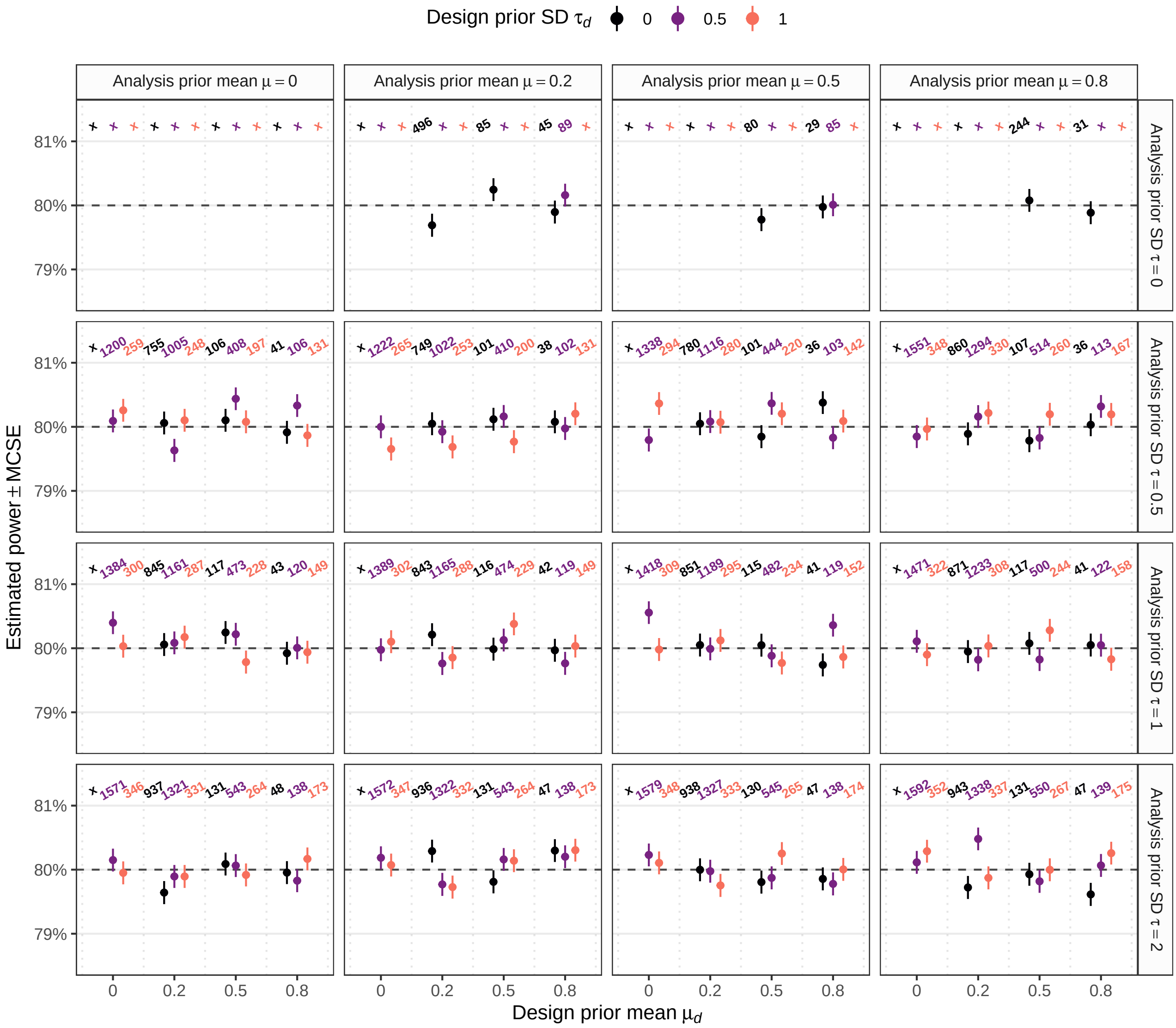

Figure 1: Simulation-based evaluation of power and sample size calculations related to the Bayesian z-test as implemented in the `bfpwr` package. The top of each plot shows the sample size to obtain a Bayes factor equal or below $k = 1/10$ with a target power of 80% for the corresponding combination of analysis and design prior (an "x" is shown if the target power is impossible to achieve for a given condition). 50'000 Bayes factors were then simulated based on this sample size, from which then the power was empirically estimated.

We can see that in all conditions, the simulation-based estimate of the power closely spreads around the target power of 80%. The maximally observed discrepancy is 0.56% while the median discrepancy is 0.14%. This suggests that the power and sample size calculation methods work as intended.

5 Power with normal moment prior

Setting the Bayes factor (14) to less or equal than k and applying algebraic manipulations, we can bring the inequality into the form

$$\exp\left[1 + \frac{n(\hat{\theta} - \theta_0)^2}{\sigma^2_{\hat{\theta}}\{1 + \sigma^2_{\hat{\theta}}/(n\tau^2)\}}\right]\left[1 + \frac{n(\hat{\theta} - \theta_0)^2}{\sigma^2_{\hat{\theta}}\{1 + \sigma^2_{\hat{\theta}}/(n\tau^2)\}}\right] \geq \frac{\{1 + (n\tau^2)/\sigma^2_{\hat{\theta}}\}\sqrt{e}}{2k}.$$

Applying the Lambert W function on both sides, leads to

$$1 + \frac{n(\hat{\theta} - \theta_0)^2}{\sigma^2_{\hat{\theta}}\{1 + \sigma^2_{\hat{\theta}}/(n\tau^2)\}} \geq \mathrm{W}_0\left[\frac{\{1 + (n\tau^2)/\sigma^2_{\hat{\theta}}\}\sqrt{e}}{2k}\right]. \tag{3}$$

Since the argument of the Lambert W function is real and always non-negative, only the principal branch W_0 can satisfy the inequality. Assuming a $\hat{\theta} \mid n, \mu_d, \tau_d \sim \mathrm{N}(\mu_d, \tau_d^2 + \sigma^2_{\hat{\theta}}/n)$ distribution induced by a normal design prior, we can rearrange the inequality (3) and obtain the power function (15).

References

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1):155–159. doi:10.1037/0033-2909.112.1.155.

Corless, R. M., Gonnet, G. H., Hare, D. E. G., Jeffrey, D. J., and Knuth, D. E. (1996). On the Lambert W function. *Advances in Computational Mathematics*, 5(1):329–359. doi:10.1007/bf02124750.

Computational details

```
cat(paste(Sys.time(), Sys.timezone(), "\n"))

## 2024-11-13 13:44:08.41187 Europe/Zurich

sessionInfo()

## R version 4.4.1 (2024-06-14)
## Platform: x86_64-pc-linux-gnu
## Running under: Ubuntu 24.04.1 LTS
##
## Matrix products: default
## BLAS:   /usr/lib/x86_64-linux-gnu/blas/libblas.so.3.12.0
## LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.12.0
##
## locale:
##  [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
##  [3] LC_TIME=de_CH.UTF-8        LC_COLLATE=en_US.UTF-8
##  [5] LC_MONETARY=de_CH.UTF-8    LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=de_CH.UTF-8       LC_NAME=C
##  [9] LC_ADDRESS=C               LC_TELEPHONE=C
## [11] LC_MEASUREMENT=de_CH.UTF-8 LC_IDENTIFICATION=C
##
## time zone: Europe/Zurich
## tzcode source: system (glibc)
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] ggplot2_3.5.1 dplyr_1.1.4   bfpwr_0.1.3   knitr_1.48
##
## loaded via a namespace (and not attached):
##  [1] vctrs_0.6.5       cli_3.6.3         rlang_1.1.4       xfun_0.49
##  [5] highr_0.11        generics_0.1.3    labeling_0.4.3    glue_1.7.0
##  [9] colorspace_2.1-1  scales_1.3.0      fansi_1.0.6       grid_4.4.1
## [13] evaluate_0.24.0   munsell_0.5.1     tibble_3.2.1      lifecycle_1.0.4
## [17] compiler_4.4.1    pkgconfig_2.0.3   farver_2.1.2      digest_0.6.35
## [21] viridisLite_0.4.2 R6_2.5.1          tidyselect_1.2.1  utf8_1.2.4
## [25] pillar_1.9.0      magrittr_2.0.3    tools_4.4.1       withr_3.0.1
## [29] gtable_0.3.5
```