Using the bfpwr package

Samuel Pawel
https://orcid.org/0000-0003-2779-320X

Package version 0.3
November 6, 2026

The **bfpwr** package provides functions to compute commonly used Bayes factors and perform corresponding power and sample size calculations. The theoretical background of the package is described in Pawel and Held (2025). Calculations in **bfpwr** are performed analytically or with numerical (non-simulation based) methods. This differs from most other packages that use simulation methods for the same purposes (e.g., the **BFDA** package from Schönbrodt and Stefan, 2019). This typically leads to faster computations without simulation error, but at the cost of being restricted to certain data distributions and analysis methods. This vignette illustrates how the package can be used in some typical situations. Table 1 summarizes the main functions of the package. It is recommended to look at the function documentation (`?functionname`) before using them for the first time as this vignette provides only a broad overview of the package.

Table 1: Main functions for Bayes factor analysis and design in the **bfpwr** package.

Bayes factor type	**Analysis function**	**Design function**
z-test Bayes factor (Section 1)	`bf01`	`powerbf01`
t-test Bayes factor (Section 2)	`tbf01`	`powertbf01`
Normal moment Bayes factor (Section 3)	`nmbf01`	`powernmbf01`

The **bfpwr** package can be installed from CRAN by running

```
install.packages("bfpwr")
```

followed by loading it with

```
library("bfpwr")
```

Development of **bfpwr** is being done on GitHub. Anyone with ideas for new features, bug reports, or other contributions to the package is invited to get in touch there (https://github.com/SamCH93/bfpwr).

1 z-test Bayes factor

The z-test Bayes factor is a fairly general analysis method that is applicable to data summarized by an estimate $\hat{\theta}$ of an unknown parameter θ, along with a standard error of the estimate. The estimate is assumed to be approximately normally distributed around the true parameter with a standard deviation equal to the standard error. The z-test Bayes factor then quantifies the evidence for the null hypothesis that the parameter takes a certain null value $H_0: \theta = \theta_0$ against the alternative hypothesis that it takes another value $H_1: \theta \neq \theta_0$, in light of the observed data and assuming a normal 'analysis prior' for the parameter θ under the alternative H_1. In all functions related to the z-test Bayes factor, the normal analysis prior is specified by the prior mean (argument `pm`) and the prior standard deviation (argument `psd`). Setting the prior standard deviation to zero corresponds to a point prior and reduces the Bayes factor to a likelihood ratio. As such, the **bfpwr** package can also be used for 'likelihoodist' (sometimes also known as 'evidential') analysis and design (see e.g., Royall, 1997, for an overview of the likelihoodist/evidential paradigm).

1.1 Analysis with the z-test Bayes factor

The z-test Bayes factor can be calculated using the function `bf01`, see `?bf01` for a detailed description of its arguments. The following code demonstrates application to data in the form of an estimated regression coefficient from a logistic regression model and its standard error

```
## glm to quantify association between Virginica species and sepal width
iris$virginica <- as.numeric(iris$Species == "virginica")
irisglm <- glm(virginica ~ Petal.Width, data = iris, family = "binomial")
estimate <- summary(irisglm)$coefficients[2,1] # logOR estimate
se <- summary(irisglm)$coefficients[2,2] # standard error

## Bayes factor parameters
null <- 0 # null value
pm <- 0 # analysis prior mean
psd <- 2 # analysis prior sd

## compute z-test Bayes factor
bf01(estimate = estimate, se = se, null = null, pm = pm, psd = psd)

## [1] 0.03945687
```

We can see that the data provide strong evidence ($\text{BF}_{01} = 1/25.3$) for the alternative hypothesis of a log odds ratio unequal to zero (i.e., an association between the species Virginica and sepal width), over the null hypothesis of a log odds ratio equal to zero (i.e., no association). Note that the Bayes factor calculated using the `bf01` function and all other functions in **bfpwr** are oriented *in favor of the null* hypothesis over the alternative, so $\text{BF}_{01} > 1$ indicates evidence for the null hypothesis, whereas $\text{BF}_{01} < 1$ indicates evidence for the alternative hypothesis. If a Bayes factor orientation in favor of the alternative over the null is desired instead, one can invert the Bayes factor by $\text{BF}_{10} = 1/\text{BF}_{01}$.

1.2 Design with the z-test Bayes factor

Another reason we would want to use **bfpwr** is that we have no observed data yet, and we want to either (i) compute a sample size that will produce a compelling Bayes factor with a given probability, or (ii) compute the probability of obtaining a compelling Bayes factor for a given sample size (the 'power').

The function `powerbf01` can be used to compute power and sample size, assuming that the data are normally distributed and that the parameter of interest is either a mean or a (standardized) mean difference. It is inspired by the `power.t.test` function from the `stats` package, with which many users will be familiar. One can either compute the power for a given sample size (by specifying the `n` argument) or compute the sample size for a given power (by specifying the `power` argument). The code below demonstrates usage of `powerbf01` for a two-sample test of a standardized mean difference parameter

```
## Bayes factor and sample size parameters
null <- 0 # null value
pm <- 0.3 # analysis prior mean
psd <- 1 # analysis prior sd
k <- 1/10 # desired Bayes factor threshold (BF01 < k)
type <- "two.sample" # two-sample z-test
sd <- 1 # sd of one observation, set to 1 for standardized mean difference scale

## compute sample size to achieve desired power
pow <- 0.9
(res1 <- powerbf01(k = k, power = pow, sd = sd, null = null, pm = pm, psd = psd,
                   type = type))

##
##      Two-sample z-test Bayes factor power calculation
##
##                          n = 1292.6
##                      power = 0.9
##                         sd = 1
##                       null = 0
##       analysis prior mean = 0.3
##         analysis prior sd = 1
##         design prior mean = 0.3
##           design prior sd = 1
##            BF threshold k = 1/10
##
## NOTE: BF oriented in favor of H0 (BF01 < 1 indicates evidence for H1 over H0)
##       n is number of *observations per group*
##       sd is standard deviation of one observation (assumed equal in both groups)

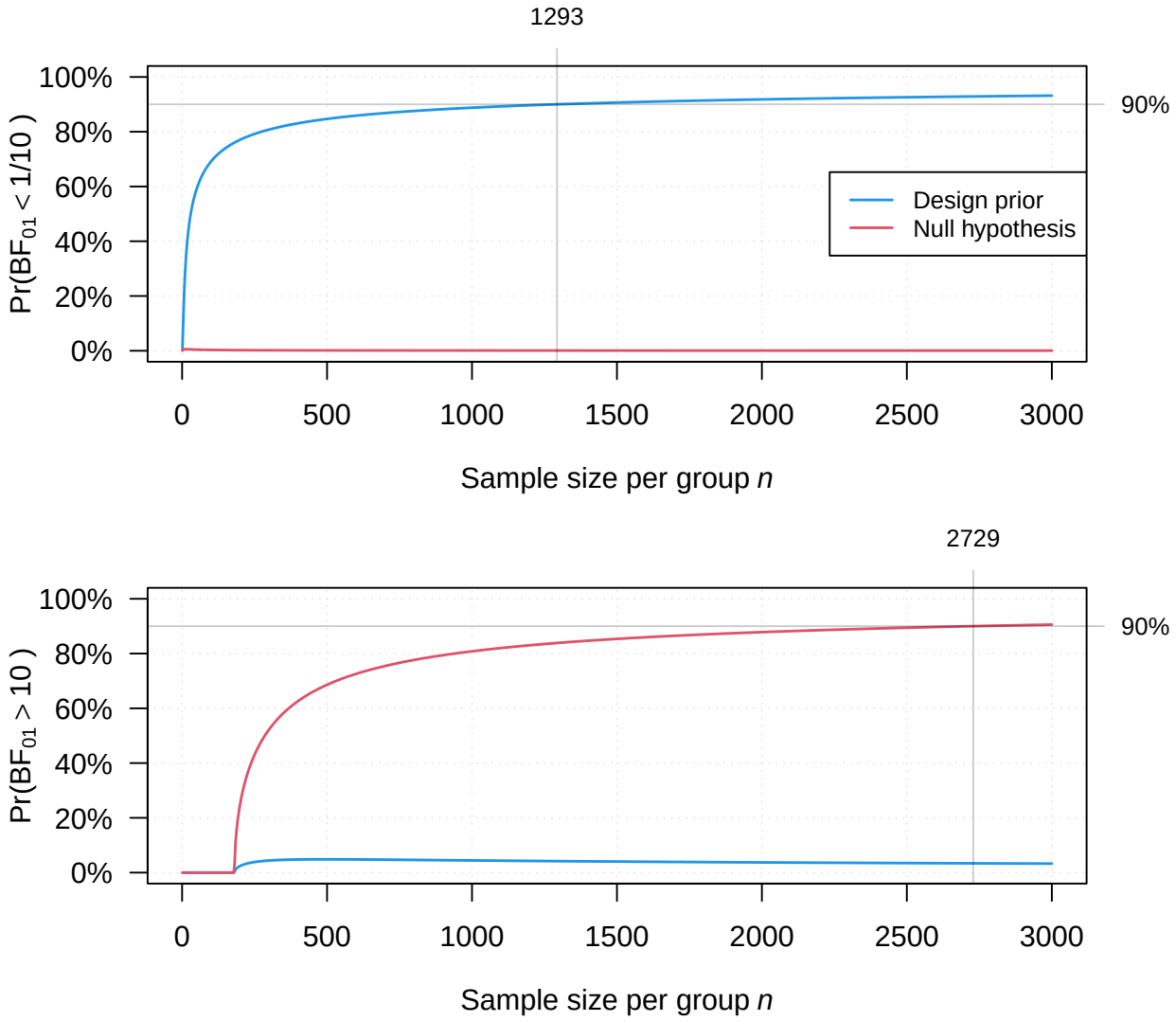
## compute power for a given sample size
n <- 500
(res2 <- powerbf01(k = k, n = n, sd = sd, null = null, pm = pm, psd = psd,
                   type = type))

##
```

```
##      Two-sample z-test Bayes factor power calculation
##
##                          n = 500
##                      power = 0.8467494
##                         sd = 1
##                       null = 0
##       analysis prior mean = 0.3
##         analysis prior sd = 1
##         design prior mean = 0.3
##           design prior sd = 1
##            BF threshold k = 1/10
##
## NOTE: BF oriented in favor of H0 (BF01 < 1 indicates evidence for H1 over H0)
##       n is number of *observations per group*
##       sd is standard deviation of one observation (assumed equal in both groups)
```

Defining a Bayes factor below the threshold $k = 1/10$ as target for compelling evidence, we see that $n = 1293$ observations per group (obtained from rounding the resulting sample size $n = 1292.6$ to the next larger integer) are required to obtain a power of 90%, or that a power of 84.7% is obtained for a sample size of $n = 500$. Both function calls assume that the underlying parameter is sampled from the normal analysis prior under the alternative hypothesis, as specified with the arguments `pm` and `psd`. By additionally specifying the arguments `dpm` and `dpsd`, it is also possible to specify a normal 'design prior' that differs from the analysis prior. For example, we may want to specify a design prior that encodes more optimistic assumptions about the parameter than the analysis prior, or we may want to set the design prior to the null hypothesis (`dpm = null` and `dpsd = 0`) to compute the probability of misleading evidence in favor of the alternative when the null is actually true. The latter is done automatically when plotting the resulting `power.bftest` object with the corresponding `plot` method, as shown below

```
## plot power curves (also tweaking the limits and resolution of the x-axis)
plot(res1, nlim = c(1, 3000), ngrid = 1000)
```



The first plot highlights the sample size required to achieve the specified power, along with power values for other sample sizes. The second plot shows a power curve for obtaining a Bayes factor in favor of the null hypothesis (using the reciprocal of the specified threshold `k`), it can be turned off with the argument `nullplot = FALSE`. Finally, the argument `plot = FALSE` can be specified to return only the data underlying the plot, for example, if one wants to use an alternative plotting package.

To conduct such sample size and power calculations for other parameters than a (standardized) mean (difference), additional assumptions regarding the standard error and sample size are needed. The functions `pbf01` and `nbf01` can perform power and sample size calculations in more general cases than `powerbf01`. Both assume that the standard error is of the form $\sigma_{\hat{\theta}}/\sqrt{n}$, where $\sigma_{\hat{\theta}}$ is the standard deviation of one effective observation and n is the 'effective sample size'. In both functions, users have to specify the unit standard deviation $\sigma_{\hat{\theta}}$ that determines the scale of the calculation. Table 2 lists how some typically used parameters can be cast into this framework. The documentation of both functions (`?pbf01` and `?nbf01`) provides more information and examples.

Table 2: Different types of parameter estimates $\hat{\theta}$ with approximate standard error $\sigma_{\hat{\theta}}/\sqrt{n}$ and corresponding interpretation of sample size n and unit standard deviation $\sigma_{\hat{\theta}}$ (adapted from Chapter 2.4 in Spiegelhalter et al., 2004 and Chapter 1 in Grieve, 2022). The standard deviation of one continuous outcome observation is denoted by σ. Parameter estimates based on two groups assume an equal number of observations per group.

Outcome	**Parameter estimate $\hat{\theta}$**	**Interpretation of n**	**Unit standard deviation $\sigma_{\hat{\theta}}$**
Continuous	Mean	Sample size	σ
Continuous	Mean difference	Sample size per group	$\sigma\sqrt{2}$
Continuous	Standardized mean difference	Sample size per group	$\sqrt{2}$
Continuous	z-transformed correlation	Sample size minus 3	1
Binary	Log odds ratio	Total number of events	2
Binary	Arcsine square root difference	Sample size per group	$1/\sqrt{2}$
Survival	Log hazard ratio	Total number of events	2
Count	Log rate ratio	Total count	2

2 *t*-test Bayes factor

The t-test Bayes factor is an analysis method specifically tailored to testing a (standardized) mean (difference) parameter θ based on normally distributed data with unknown variance (assumed equal across both groups if more than one). This Bayes factor quantifies the evidence that the data (in the form of a t-statistic and sample sizes) provide for the null hypothesis that the parameter equals zero against the alternative hypothesis that it is not equal zero. Following Gronau et al. (2020), the t-test Bayes factor implemented in **bfpwr** assumes that a scale-location t prior distribution (potentially truncated to only positive or only negative parameters) is assigned to the standardized mean (difference) under the alternative hypothesis. When centering the prior on zero and setting its degrees of freedom to one, the Bayes factor reduces to the 'Jeffreys-Zellner-Siow' (JZS) Bayes factor (Jeffreys, 1961; Zellner and Siow, 1980), which is often used as a 'default' Bayes factor in the social sciences (Rouder et al., 2009). Setting other values for these parameters allows data analysts to incorporate directionality or prior knowledge about the parameter, potentially making the test more informative. Figure 1 shows examples of different prior distributions along with the arguments to specify them.

2.1 Analysis with the *t*-test Bayes factor

The t-test Bayes factor can be calculated with the function `tbf01`, see the documentation `?tbf01` for details on its arguments. The following code chunk illustrates usage of the function with two examples from the literature

```
## paired t-test JZS Bayes factor analyses from Rouder et al. (2009, p.232)
tbf01(t = 2.03, n = 80, plocation = 0, pscale = 1, pdf = 1, type = "paired")

## [1] 1.557447

## informed prior analysis from Gronau et al. (2020, Figure 1)
tbf01(t = -0.90, n1 = 53, n2 = 57, plocation = 0.350, pscale = 0.102, pdf = 3,
      alternative = "greater", type = "two.sample")
```

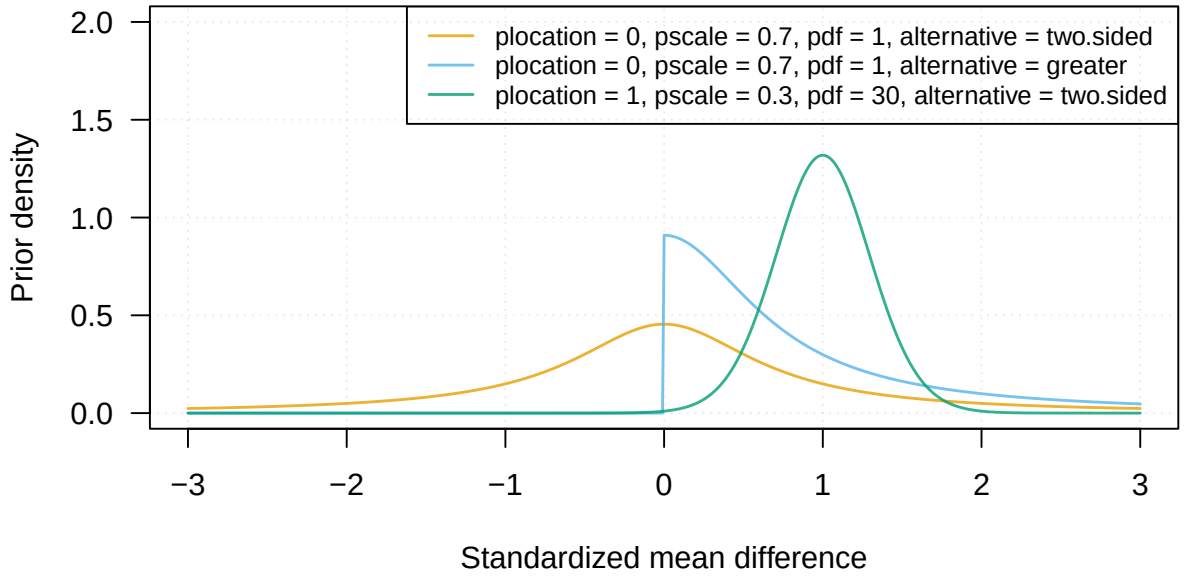


Figure 1: Illustration of different (truncated) scale-location t prior distributions for the t-test Bayes factor along with the arguments to specify them. The yellow and blue priors represent 'default' priors that are undirectional and directional, respectively, the green prior represents an informed prior.

```
## [1] 11.55065
```

The first example reproduces a Bayes factor calculation reported in Rouder et al. (2009, p. 232): A t-statistic of $t = 2.03$ obtained from $n = 80$ paired observations part of a psychological experiment, together with a standard Cauchy distribution assigned to the standardized mean difference parameter (a t distribution centered around zero with one degree of freedom and a scale of one), yields a Bayes factor of $\text{BF}_{01} = 1.56$, which provides anecdotal evidence in favor of the null hypothesis of no effect over the alternative hypothesis of an effect. The second example, from Gronau et al. (2020, p. 140-141), analyzes a t-statistic of $t = -0.90$ from a psychological experiment based on a mean comparison of two groups with size $n_1 = 53$ and $n_2 = 57$, respectively. Here, a more informed prior was elicited from an expert, and is also truncated to positive effects (with the argument `alternative = "greater"`). This yields a Bayes factor of $\text{BF}_{01} = 11.55$, which provides strong evidence for the null hypothesis of no effect over the alternative of an effect.

2.2 Design with the t-test Bayes factor

The `powertbf01` function can be used for both power and sample size calculations assuming that the future data are analyzed with the t-test Bayes factor, but still assuming known variances and a normal design prior for the design. As such, the function has a similar specification as `powerbf01`, differing only in the arguments of the analysis prior. The following code illustrates its usage

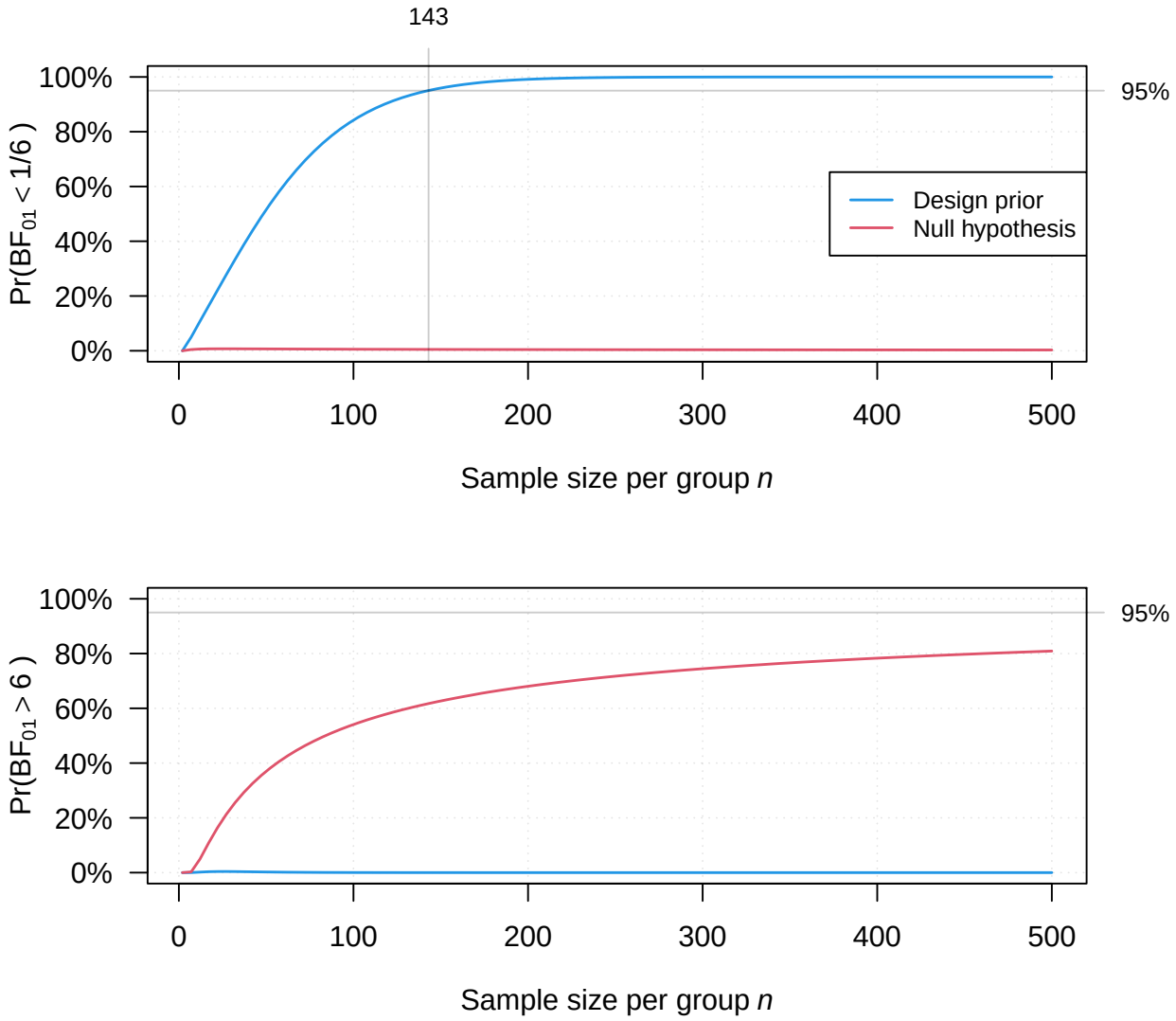
```
## determine sample size for a given power
(tres <- powertbf01(k = 1/6, # Bayes factor threshold
                    power = 0.95, # target power
```

```
                    type = "two.sample", # two-sample test
                    ## normal design prior
                    dpm = 0.5, # design prior mean at d = 0.5
                    dpsd = 0, # point design prior (standard deviation is zero)
                    ## directional JSZ analysis prior
                    plocation = 0,
                    pscale = 1/sqrt(2),
                    pdf = 1,
                    alternative = "greater"))

##
##      Two-sample t-test Bayes factor power calculation
##
##                              n = 142.4462
##                          power = 0.95
##                             sd = 1
##                           null = 0
##                    alternative = greater
##       analysis prior location = 0
##          analysis prior scale = 0.7071068
##             analysis prior df = 1
##             design prior mean = 0.5
##               design prior sd = 0
##                BF threshold k = 1/6
##
## NOTE: BF oriented in favor of H0 (BF01 < 1 indicates evidence for H1 over H0)
##       n is number of *observations per group*
##       sd is standard deviation of one observation (assumed equal in both groups)
```

In this example, taken from Schönbrodt and Wagenmakers (2018), it is assumed that the future data will be analyzed with a two-sample t-test Bayes factor using a directional JSZ analysis prior with a scale of $1/\sqrt{2}$, and that a Bayes factor below `k = 1/6` is interpreted as compelling evidence. In addition, a standardized mean difference of 0.5 was assumed for the underlying effect size, which corresponds to setting the design prior mean to `dpm = 0.5` and its standard deviation to zero (`dpsd = 0`). If parameter uncertainty is to be accounted for, one could alternatively set a positive value for the design prior standard deviation. Taken together, the calculation results in a sample size of $n = 142.45$, which means that at least 143 observations per group are required to obtain compelling evidence with a power of 95%. As with the z-test Bayes factor, the power curves corresponding to these design calculations can be plotted with

```
plot(tres)
```



3 Normal moment prior Bayes factor

The normal moment prior Bayes factor is another analysis method that is applicable to data in the form of a parameter estimate with standard error, similar to the z-test Bayes factor. The difference between the two approaches is that the former uses a so-called 'normal moment' prior distribution for the parameter under the alternative hypothesis while the z-test Bayes factor uses just a normal prior distribution. Normal moment priors are a type of 'non-local' prior distribution which allow evidence for a point null hypothesis to accumulate more quickly if it is indeed true (Johnson and Rossell, 2010). This is because normal moment priors have zero density at the null value, see Figure 2 for an illustration. In all functions related to the normal moment prior Bayes factor, the prior is specified by the point null hypothesis (argument `null`) and the prior spread (argument `psd`). The latter controls how much the prior is spread out, and determines the modes of the distribution which are located at $\pm$`psd`$\sqrt{2}$. As such, the prior could be specified by setting the spread parameter so that the modes equal two parameter values deemed plausible or relevant under the alternative.

Analysis and design with the normal moment prior Bayes factor is very similar to analysis with the

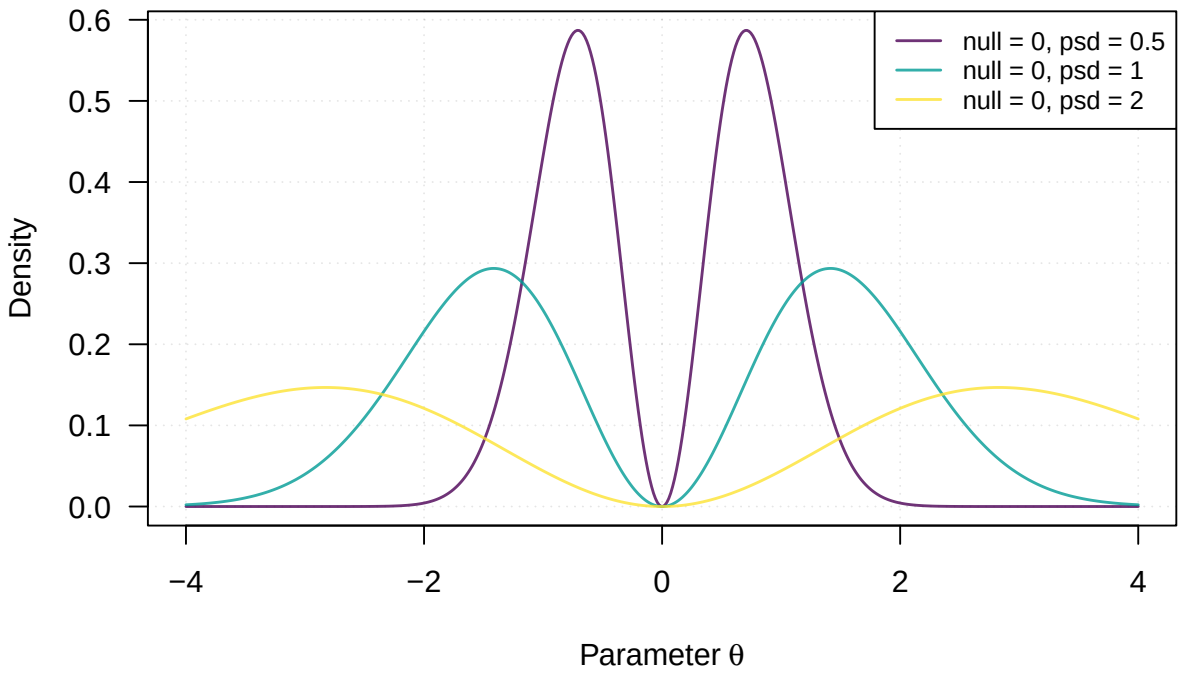


Figure 2: Illustration of different normal moment prior distributions along with the arguments to specify them.

z-test Bayes factor. There is again one analysis function (`nmbf01`), a design function for (standardized) mean (difference) parameters (`powernmbf01`), and two more general design functions that can also accommodate other parameter types (`pnmbf01` and `nnmbf01`). The documentation of these functions provides more information and examples.

References

Grieve, A. P. (2022). *Hybrid frequentist/Bayesian power and Bayesian power in planning clinical trials*. Chapman & Hall/CRC Biostatistics Series. Taylor & Francis, London. https://doi.org/10.1201/9781003218531.

Gronau, Q. F., Ly, A., and Wagenmakers, E.-J. (2020). Informed Bayesian t-tests. *The American Statistician*, 74(2):137–143. https://doi.org/10.1080/00031305.2018.1562983.

Jeffreys, H. (1961). *Theory of Probability*. Clarendon Press, Oxford, third edition.

Johnson, V. E. and Rossell, D. (2010). On the use of non-local prior densities in Bayesian hypothesis tests. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(2):143–170. https://doi.org/10.1111/j.1467-9868.2009.00730.x.

Pawel, S. and Held, L. (2025). Closed-form power and sample size calculations for Bayes factors. *The American Statistician*, 79(3):330–344. https://doi.org/10.1080/00031305.2025.2467919.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., and Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16(2):225–237. https://doi.org/10.3758/pbr.16.2.225.

Royall, R. (1997). *Statistical Evidence: A likelihood paradigm*. Chapman & Hall, London. https://doi.org/10.1201/9780203738665.

Schönbrodt, F. D. and Wagenmakers, E.-J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, 25(1):128–142. https://doi.org/10.3758/s13423-017-1230-y.

Schönbrodt, F. D. and Stefan, A. M. (2019). *BFDA: An R package for Bayes factor design analysis (version 0.5.0)*. https://github.com/nicebread/BFDA.

Spiegelhalter, D. J., Abrams, R., and Myles, J. P. (2004). *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. Wiley, Chichester. https://doi.org/10.1002/0470092602.

Zellner, A. and Siow, A. (1980). Posterior odds ratios for selected regression hypotheses. *Trabajos de Estadistica Y de Investigacion Operativa*, 31(1):585–603. https://doi.org/10.1007/bf02888369.

Computational details

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

## 2026-11-06 12:56:45.087822 Europe/Zurich

sessionInfo()

## R version 4.6.2 (2026-10-29)
## Platform: x86_64-pc-linux-gnu
## Running under: Ubuntu 24.04.5 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  LAPACK version 3.12.0
##
## locale:
##  [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
##  [3] LC_TIME=de_CH.UTF-8        LC_COLLATE=C
##  [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] bfpwr_0.3  knitr_1.52
##
## loaded via a namespace (and not attached):
## [1] compiler_4.6.2 cli_3.6.6      tools_4.6.2    otel_0.2.0     highr_0.12
## [6] xfun_0.57      rlang_1.2.0    evaluate_1.0.5
```