Bayesian Sample Size Justification

Aaron R. Caldwell

01 September, 2020

Introduction

In this study we plan on collecting data on 300 sport and exercise science research articles (100 from 3 journals). Based on the work of Büttner et al. (2020), we anticipate at least 150 (50%) of the articles will include a hypothesis that was tested. Further, the work of Fanelli (2010), Scheel, Schijen, and Lakens (2020), and Büttner et al. (2020) we believe that the percentage of articles that find support for their hypothesis is greater than 80%. Given this existing data, we believe we have an informative prior on the underlying distribution of postive results, and have opted for a Bayesian analysis of this primary endpoint. For this analysis, we will use the `brms` R package (Bürkner 2017).

Hypothesis

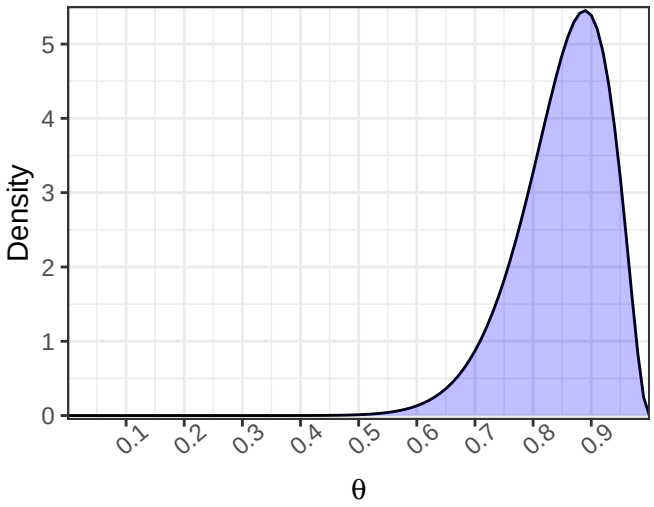
For this study, we hypothesize that the rate of positive results (i.e., studies that find at least partial support for their hypothesis) is greater than 80%. Therefore, the null hypothesis (H_0) is that the proportion of positive results is less than .8 and our alternative is greater than .8. There is no other effect being estimated in this study therefore the intercept of the model is what will be tested.

$H_0 : Intercept \leq .8$

$H_1 : Intercept > .8$

Prior Choice

The prior we selected for this analysis was informed by the previous studies assuming the true positive rate is approximately 85% (Fanelli 2010). However, we would like to avoid "spiking" the prior in favor of our hypothesis and therefore want a skeptical prior. Based on the work of Scheel, Schijen, and Lakens (2020) and Büttner et al. (2020) the estimated positive rates in original research investigations ranged from 82%-92%, and even some fields included in the survey by Fanelli (2010) observed rates at low as ~70%. Therefore, we selected a prior of $\beta(17, 3)$, and is visualized it below. This prior is centered around .85, but includes the possibility of higher (.9) and much lower (.7) proportions as compatible parameter estimates.



Data Analysis Example

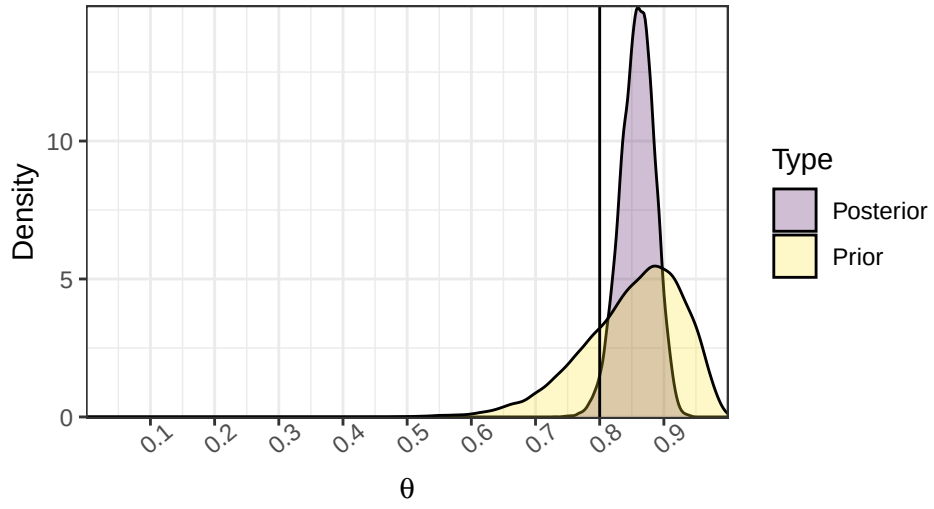
Below, we have incorporated this prior (`prior_1`) into a simulated dataset (`test_df`) and then analyzed this with the `brm` function (saved as `m_test`).

```
#Set prior
prior_1 = set_prior("beta(17, 3)", class = "b", lb = 0, ub = 1)

#Generate test data
test_df = data.frame(run = 1,
                     pos = rbinom(1, 150, .85),
                     N = rep(150, 1)) %>%
  mutate(rate = pos/N)

#Build model
m_test <- brm(
  pos | trials(N) ~ 0 + Intercept,
  family = binomial(link = "identity"),
  prior = prior_1,
  data = test_df,
  sample_prior = TRUE,
  iter = 1e4,
  cores = 4,
  refresh = 0
)
```

We can then visualize the performance of the prior and the posterior from this model.



In addition, the hypothesis can be tested with the `hypothesis` function and the posterior compatibility intervals (C.I.).

```
h_test <- hypothesis(m_test, "Intercept > 0.8")
knitr::kable(h_test$hypothesis, caption = "Hypothesis Test")
```

Table 1: Hypothesis Test

Hypothesis	Estimate	Est.Error	CI.Lower	CI.Upper	Evid.Ratio	Post.Prob	Star
(Intercept)-(0.8) > 0	0.0583896	0.0266125	0.0126561	0.0998934	48.62779	0.97985	*

```
test_pos = posterior_interval(m_test,
                      prob = .95)
knitr::kable(test_pos, caption = "95% Posterior C.I.")
```

Table 2: 95% Posterior C.I.

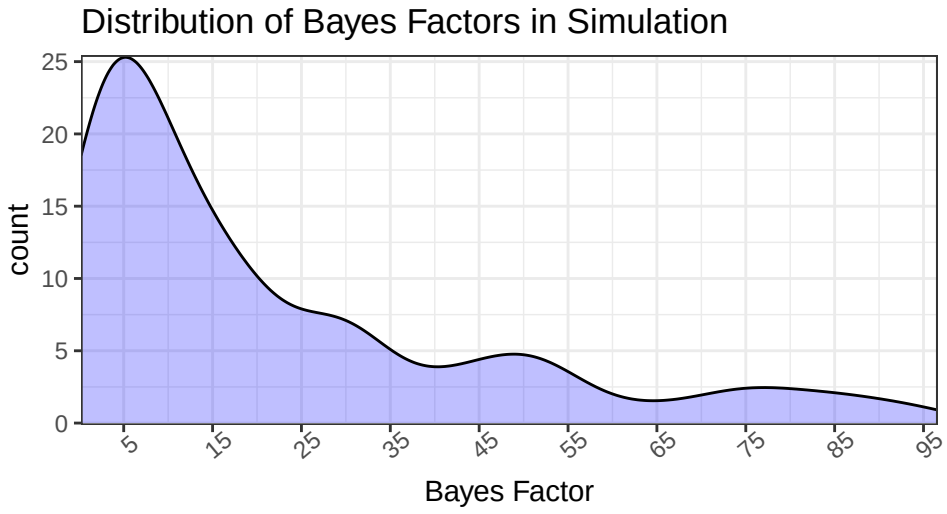
	2.5%	97.5%
b_Intercept	0.8027355	0.9068057
prior_b	0.6707599	0.9668113
lp___	-5.3816301	-2.8567919

From the simulated scenario we find that given the data the hypothesis that the true positive result rate is greater than 80% is 48.63 times more likely than the true value being less than 80%. Now, this is only over 1 simulated dataset, and, in order to estimate our "power, we will need to replicate this process over a thousand simulations.

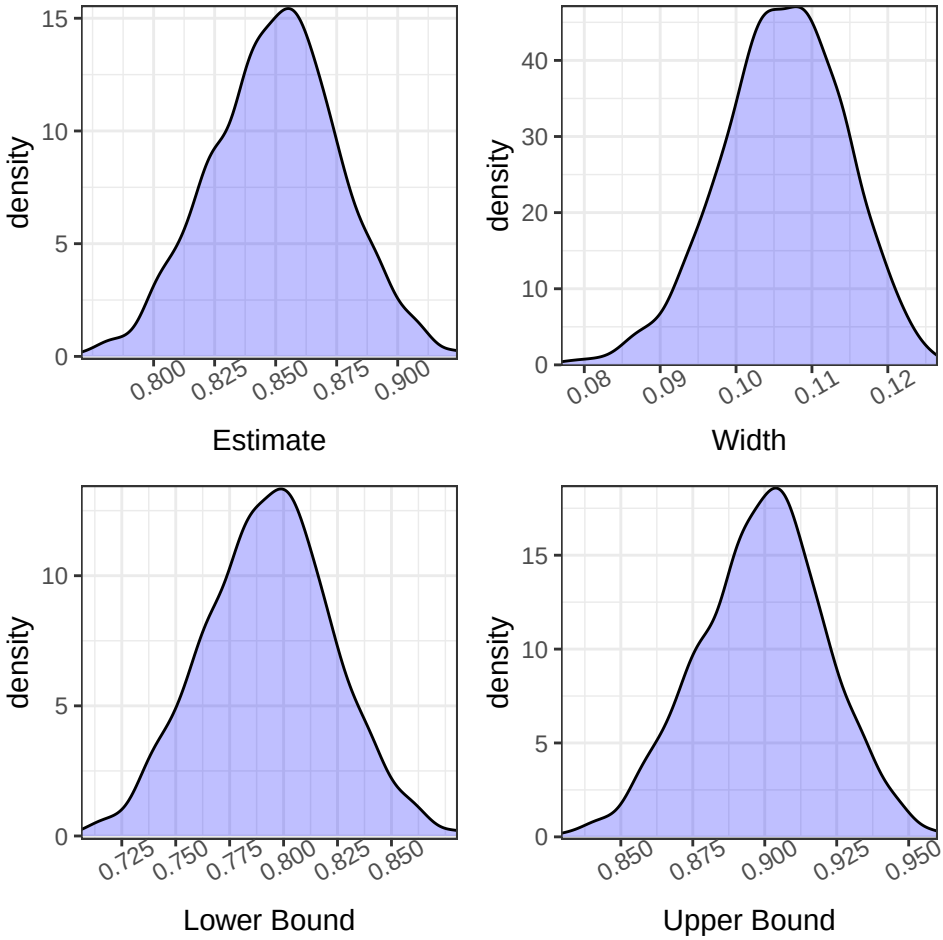
Simulations

Now that we have established the process by which the data are analyze I will summarize the results of a simulation (1000 iterations) of the performance of this model. Please note that code to reproduce these analyses can be found at the end of the document.

First, this analysis, under the previously stated assumptions, would be able to yield a Bayes Factor in favor of our hypothesis (BF > 3) 86.7% of the time. Below is a plot of the simulated Bayes Factors (excluding Bayes Factors > 100). As a note, in the final manuscript we also plan to report the posterior probabilites of our selected hypotheses.



Second, we have included a plot of the distribution of posterior credible intervals below. Approximately 41.6% of all CI lower bounds were greater than 80%.



Conclusion

Overall, the data from this study will be adequate to test our hypothesis since the 86.7% of the simulations demonstarted at least some evidence for our hypothesis. Also, we only assumed 150 manuscripts would be analyzed and the underlying distribution is exactly 85%. In reality, we anticipate that we should actually have more than 180 manuscripts (60% of the sample) with hypotheses to test which will only increase the "power" of this Bayesian analysis.

Appendix: Code to Reproduce the Simulations

```
# Data generating function ---------------------------
gen_data = function(run,n,prop){
  df = data.frame(run = run,
                  pos = rbinom(1, n, prop),
                  N = n) %>%
    mutate(rate = pos/N)
  return(df)
}

# Build the initial model ----------------------------
initial_form = function(n = 150,
                        prop = .85,
                        sim_prior = set_prior("beta(17, 3)", class = "b", lb = 0, ub = 1)){
  init_df = data.frame(run = 1,
                       pos = rbinom(1, n, prop),
                       N = n) %>%
    mutate(rate = pos/N)
  fit <- brm(
    pos | trials(N) ~ 0 + Intercept,
    family = binomial(link = "identity"),
    prior = sim_prior,
    data = init_df)
  return(fit)
}

# Set the parameters for the simulation ----------------------------
set.seed(08202020)
nsims = 1000
ci = .95
hyp_test = "Intercept > 0.8"
fit = initial_form(n = 150,
                   prop = .85,
                   sim_prior = set_prior("beta(17, 3)", class = "b", lb = 0, ub = 1))
bin_sims = data.frame(run = NA,
                      d = NA,
                      fit = NA)
bin_sims = bin_sims[FALSE,]

## Split simulations ----------------------------
# Must run in parts due to C error (possibly memory issues)
for (i in 1:10) {
bin_run = tibble(run = 1:(nsims/10)) %>%
  mutate(d = map(run, gen_data, n = 150, prop = .85)) %>%
  mutate(fit  = map(d, ~update(fit, newdata = .x, refresh = 0)))
bin_sims = rbind(bin_sims,bin_run)
}

## Calclulate estimates ----------------------------
bin_est = bin_sims %>%
  mutate(test = map(fit,tidy,prob=ci)) %>%
  unnest(test) %>%
```

```
  filter(term == "b_Intercept") %>%
  select(-d,-fit) %>%
  mutate(width = upper-lower)

## Calclulate hypothesis tests --------------------------
bin_hyp = bin_sims %>%
  mutate(hyp = map(fit,hypothesis,hyp_test)) %>%
  select(run,hyp)

hyp_df = data.frame(1,2,3,4,5,6,7,8)
colnames(hyp_df) = colnames(bin_hyp$hyp[[1]]$hypothesis)
hyp_df = hyp_df[FALSE,]

for (i in 1:nrow(bin_hyp)){
  hyp_df = rbind(hyp_df, as.data.frame(bin_hyp$hyp[[i]]$hypothesis))

}

#save.image(file = "sin_v2.RData")
```

References

Bürkner, Paul-Christian. 2017. "brms: An R Package for Bayesian Multilevel Models Using Stan." *Journal of Statistical Software* 80 (1): 1–28. https://doi.org/10.18637/jss.v080.i01.

Büttner, Fionn, Elaine Toomey, Shane McClean, Mark Roe, and Eamonn Delahunt. 2020. "Are Questionable Research Practices Facilitating New Discoveries in Sport and Exercise Medicine? The Proportion of Supported Hypotheses Is Implausibly High." *British Journal of Sports Medicine*, July, bjsports–2019–101863. https://doi.org/10.1136/bjsports-2019-101863.

Fanelli, Daniele. 2010. "'Positive' Results Increase down the Hierarchy of the Sciences." Edited by Enrico Scalas. *PLoS ONE* 5 (4): e10068. https://doi.org/10.1371/journal.pone.0010068.

Scheel, Anne M., Mitchell Schijen, and Daniel Lakens. 2020. "An Excess of Positive Results: Comparing the Standard Psychology Literature with Registered Reports," February. https://doi.org/10.31234/osf.io/p6e9c.