Pilot Analysis Demonstration

Aaron R. Caldwell

9/8/2020

Introduction

In this document I will detail our planned analyses using our pilot data from 15 manuscripts. Please note that additional analyses may be present in the final manuscript. All results reported in brackets `[XX, XX]` represent 95% confidence intervals.

```
library(readr)
library(tidyverse)
library(brms)
library(epitab)
df_pilot <- read_csv("df_pilot.csv") %>%
  mutate_if(is.character,as.factor)
```
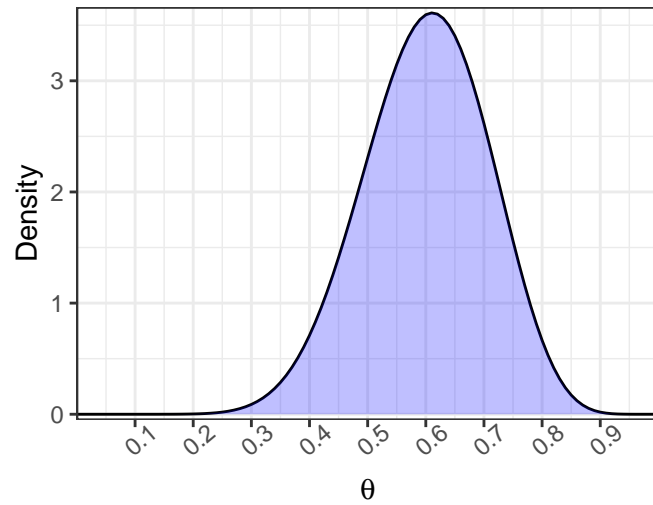
Hypothesis Tests

First, we have two hypotheses included in our manuscript.

1. For manuscripts with hypotheses, more than 80% will include some support for their hypothesis.

- Prior of $\beta(17,3)$
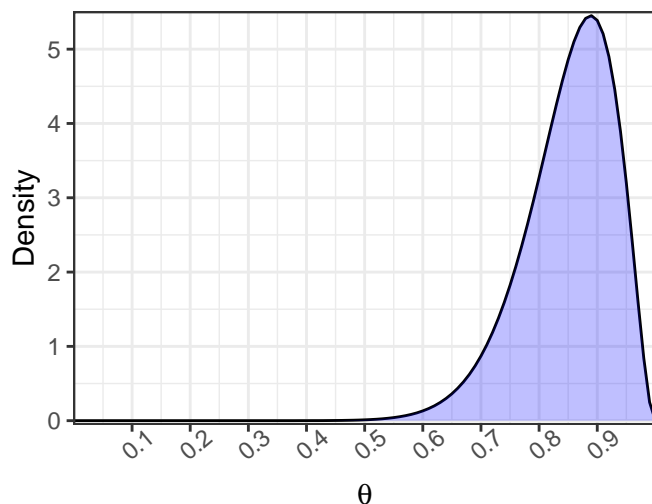
2. The proportion of manuscripts that explicitly state they are testing hypotheses is greater than 60%.

- Prior of $\beta(12, 8)$

Rate of Positive Results

From the pilot results, we would conclude that our hypothesis is less likely than the null hypothesis and therefore the positive result rate is more likely to be less than 80%

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

#transform data
hyp_pos = df_pilot %>%
  filter(hypo_tested == "Yes") %>%
  summarize(
    pos = sum(di_sup == "Y"),
    N = length(di_sup),
    rate = pos/N
  )
```

```r
#Build model
m_test <- brm(
  pos | trials(N) ~ 0 + Intercept,
  family = binomial(link = "identity"),
  prior = prior_1,
  data = hyp_pos,
  refresh = 0
)
```

```r
h_test <- hypothesis(m_test, "Intercept > 0.8")
knitr::kable(h_test$hypothesis, caption = "Hypothesis Test #1")
```

Table 1: Hypothesis Test #1

Hypothesis	Estimate	Est.Error	CI.Lower	CI.Upper	Evid.Ratio	Post.Prob	Star
(Intercept)-(0.8) > 0	-0.0347651	0.0750573	-0.1656635	0.0763424	0.5631106	0.36025	

```r
test_pos = posterior_interval(m_test,
                  prob = .95)
knitr::kable(test_pos, caption = "Hyp Test #1: 95% Posterior C.I.")
```

Table 2: Hyp Test #1: 95% Posterior C.I.

	2.5%	97.5%
b_Intercept	0.5990816	0.8935397
lp__	-5.5427593	-2.9766924

Rate of Hypotheses Tested

However, we would conclude that there is some weak evidence for our hypothesis that more than 60% of manuscripts explicitly test a hypothesis.

```r
#Set prior
prior_2 = set_prior("beta(12, 8)", class = "b", lb = 0, ub = 1)

#Generate test data
hyp_test = df_pilot %>%
  summarize(
    pos = sum(hypo_tested == "Yes"),
    N = length(hypo_tested),
    rate = pos/N
  )
```

```r
#Build model
m_test2 <- brm(
  pos | trials(N) ~ 0 + Intercept,
  family = binomial(link = "identity"),
  prior = prior_2,
  data = hyp_test,
  refresh = 0
)
```

```r
h_test2 <- hypothesis(m_test2, "Intercept > 0.6")
knitr::kable(h_test2$hypothesis, caption = "Hypothesis Test #2")
```

Table 3: Hypothesis Test #2

Hypothesis	Estimate	Est.Error	CI.Lower	CI.Upper	Evid.Ratio	Post.Prob	Star
(Intercept)-(0.6) > 0	0.0310892	0.0792066	-0.1021272	0.1563381	1.928258	0.6585	

```r
test_pos2 = posterior_interval(m_test2,
                 prob = .95)
knitr::kable(test_pos2, caption = "Hyp Test #2: 95% Posterior C.I.")
```

Table 4: Hyp Test #2: 95% Posterior C.I.

	2.5%	97.5%
b_Intercept	0.4703159	0.7747509
lp__	-4.1975388	-1.7706237

Table 5: Statistics Reported

| | All | Significance Testing | |
		No	Yes
Total	10	1 (100)	9 (100)
Effect Size Reported			
No	5 (50)	0 (0)	5 (55.6)
Yes	5 (50)	1 (100)	4 (44.4)
Significant p-value			
No	NA (NA)	NA (NA)	3 (33.3)
Yes	NA (NA)	NA (NA)	6 (66.7)
p-value Type			
Exact (e.g., p = .049)	NA (NA)	NA (NA)	4 (44.4)
Relative (e.g., p < .05)	NA (NA)	NA (NA)	5 (55.6)

Descriptive Statistics

For the rest of the paper, since we do not have any explicit hypotheses we will only use descriptive statistics to how the general rates and trends of other good, or bad, practices.

Statistics

There are 3 primary descriptive statistics we are interested in: whether an effect size was reported, whether or not significance testing was utilized, and how the authors interpreted the p-values.

```
df_pilot %>%
  filter(hypo_tested == "Yes") %>%
contingency_table(independents=list("Effect Size Reported" = "effect_size",
                                    "Significant p-value" = "pval_sig",
                                    "p-value Type" = "pval_type"),
                  outcomes=list("Significance Testing" = "sig_test"),
                  crosstab_funcs=list(freq())) %>%
  neat_table(format = "latex",
             caption = "Statistics Reported")
```

```
ct_es = table(df_pilot$effect_size)

binom_es = binom.test(ct_es[[2]],sum(ct_es))
es_pr = paste0(round(binom_es$estimate*100,2),"% [",
       round(binom_es$conf.int[1]*100,2),", ",
       round(binom_es$conf.int[2]*100,2)
       ,"] of manuscripts reported some estimate of effect size.")

ct_sig = table(df_pilot$pval_sig)

binom_sig = binom.test(ct_sig[[2]],sum(ct_sig))
sig_pr = paste0(round(binom_sig$estimate*100,2),"% [",
       round(binom_sig$conf.int[1]*100,2),", ",
       round(binom_sig$conf.int[2]*100,2)
       ,"] of manuscripts reported a \"significant\" p-value.")

ct_ptype = table(df_pilot$pval_type)

binom_ptype = binom.test(ct_ptype[[2]],sum(ct_ptype))
ptype_pr = paste0(round(binom_ptype$estimate*100,2),"% [",
       round(binom_ptype$conf.int[1]*100,2),", ",
       round(binom_ptype$conf.int[2]*100,2)
       ,"] of manuscripts reported exact p-values (p = .045) versus relative p-values (p < .05).")
```

60% [32.29, 83.66] of manuscripts reported some estimate of effect size.

66.67% [29.93, 92.51] of manuscripts reported a "significant" p-value.

55.56% [21.2, 86.3] of manuscripts reported exact p-values (p = .045) versus relative p-values (p < .05).

Table 6: Preregistration Reported

	All	Preregistration No	Preregistration Yes
Total	15	14 (100)	1 (100)
Clinical Trial			
No	8 (53.3)	8 (57.1)	0 (0)
Yes	7 (46.7)	6 (42.9)	1 (100)
RCT			
No	11 (73.3)	10 (71.4)	1 (100)
Yes	4 (26.7)	4 (28.6)	0 (0)
Animal			
No	14 (93.3)	13 (92.9)	1 (100)
Yes	1 (6.7)	1 (7.1)	0 (0)

Preregistration

Recently, there has been an emphasis on trial preregistration. This should be a requirement for clinical trials and any randomized control trial (RCT), and is highly encouraged for animal studies. Overall, from the pilot data, we would conclude that most manuscripts are out of compliance for clinical trials and no animal studies included preregistration information

```
contingency_table(independents=list("Clinical Trial"="clin_trial",
                                     "RCT"="RCT",
                                     "Animal" = "animal"),
                  outcomes=list("Preregistration"="prereg"),
                  crosstab_funcs=list(freq()),
                  data=df_pilot) %>%
  neat_table(format = "latex",
             caption = "Preregistration Reported")
```

```
ct_prereg = table(df_pilot$prereg)

binom_prereg = binom.test(ct_prereg[[2]],sum(ct_prereg))
prereg_pr = paste0(round(binom_prereg$estimate*100,2),"% [",
       round(binom_prereg$conf.int[1]*100,2),", ",
       round(binom_prereg$conf.int[2]*100,2)
       ,"] of manuscripts reported preregistration or clinical trial registration information.")
```

6.67% [0.17, 31.95] of manuscripts reported preregistration or clinical trial registration information.

Sample Size Information

All current guidelines for these three types of studies also require sample size reporting and encourage sample size justification. In our pilot sample all of the studies at least included some sample size information. However, only 2 studies included some form of sample size justification. Therefore, Kinesiology studies are 6.5 times more likely to *not* report a sample size justification.

```
contingency_table(independents=list("Clinical Trial"="clin_trial",
                                     "RCT"="RCT",
                                     "Animal" = "animal"),
```

Table 7: Sample Size Justification

| | All | Sample Size Justification | |
		No	Yes
Total	15	13 (100)	2 (100)
Clinical Trial			
No	8 (53.3)	7 (53.8)	1 (50)
Yes	7 (46.7)	6 (46.2)	1 (50)
RCT			
No	11 (73.3)	10 (76.9)	1 (50)
Yes	4 (26.7)	3 (23.1)	1 (50)
Animal			
No	14 (93.3)	12 (92.3)	2 (100)
Yes	1 (6.7)	1 (7.7)	0 (0)

```
                outcomes=list("Sample Size Justification"="N_just"),
                crosstab_funcs=list(freq()),
                data=df_pilot) %>%
  neat_table(format = "latex",
             caption = "Sample Size Justification")
```

```
ct_njust = table(df_pilot$N_just)

binom_njust = binom.test(ct_njust[[2]],sum(ct_njust))
njust_pr = paste0(round(binom_njust$estimate*100,2),"% [",
       round(binom_njust$conf.int[1]*100,2),", ",
       round(binom_njust$conf.int[2]*100,2)
       ,"] of manuscripts reported some form of sample size justification.")
```

13.33% [1.66, 40.46] of manuscripts reported some form of sample size justification.

Table 8: Sample Size by Category

sci_cat	median	IQR
Applied exercise physiology (human)	19	52623.75
Biomechanics	13	0.00
Clinical research	232	212.00
Environmental physiology (heat, cold, & altitude)	40	0.00
Epidemiology	986	159.00
Motor learning/control/behavior	335	0.00
Sport performance	32	15.50
Sport/exercise psychology	197	0.00

Further, we can illustrate the range of sample sizes with a histogram of the total reported sample size. And compare the sample sizes by sub-discipline. Due to the skewed distribution, the differences in the sample sizes for sub-disisplines will be compared with a heteroscedastic one-way ANOVA for trimmed means (WRS2 `t1way` function). Overall, from the given data we would conclude that the sample sizes have a large range, and unsurprisingly epidemiology has the largest sample size.

```
df_pilot %>%
  filter(N < 50000) %>%
ggplot(aes(x = N)) +
  geom_histogram()
```
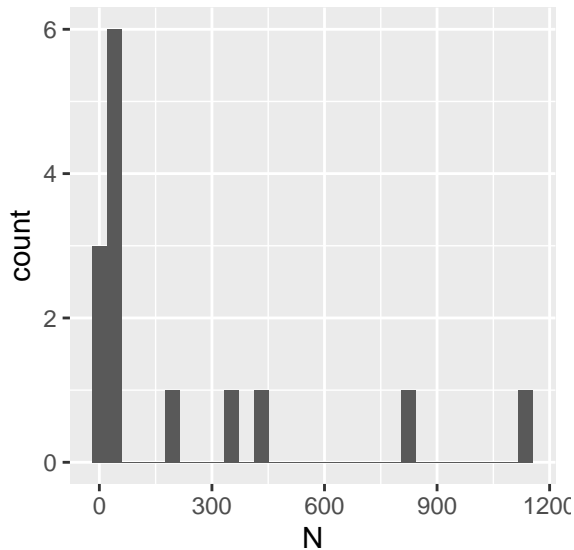


```
df_pilot %>%
  group_by(sci_cat) %>%
  summarize(median = median(N),
            IQR = IQR(N),
    .groups = 'drop'
  ) %>%
  knitr::kable(caption = "Sample Size by Category")
```

Table 9: Data Sharing Reported

Var1	Freq
No	11
Yes	4

Table 10: Replication Studies

Var1	Freq
No	15

Other Open Sciences Practices

We can estimate the proportion of studies that report some type of open data or data sharing in their manuscript.

```
knitr::kable(table(df_pilot$open_data),
             caption = "Data Sharing Reported")
```

```
ct_odat = table(df_pilot$open_data)

binom_odat = binom.test(ct_odat[[2]],sum(ct_odat))
odat_pr = paste0(round(binom_odat$estimate*100,2),"% [",
      round(binom_odat$conf.int[1]*100,2),", ",
      round(binom_odat$conf.int[2]*100,2)
      ,"] of manuscripts reported some form of data sharing or open data.")
```

26.67% [7.79, 55.1] of manuscripts reported some form of data sharing or open data.

There is some concern that the replicability of kinesiology is not know so we can also report the proportion of studies that explicitly stating they are replicating previous work.

```
knitr::kable(table(df_pilot$replic),
             caption = "Replication Studies")
```

```
ct_replic = table(df_pilot$replic)

binom_replic = binom.test(0,15)
replic_pr = paste0(round(binom_replic$estimate*100,2),"% [",
      round(binom_replic$conf.int[1]*100,2),", ",
      round(binom_replic$conf.int[2]*100,2)
      ,"] of manuscripts explicitly stated they were replicating a previous study.")
```

0% [0, 21.8] of manuscripts explicitly stated they were replicating a previous study.