W241 Power Analysis

2023-02-21

```
library(data.table)

knitr::opts_chunk$set(dpi = 300)

set.seed(3)
```

```
power_test_t <- function(
  mean_control = 10,
  mean_treat = 11,
  sd_control = 3,
  sd_treat = 3.5,
  number_per_condition = 40,
  power_loops = 100,
  ri_loops = 100,
  verbose = TRUE) {

    p_values <- NA
    ri <- NA
    d <- data.table()

    d[ , condition := rep(c('control', 'treatment'), each = number_per_condition)]

    for(power_loop in 1:power_loops) {
      if(verbose == TRUE) {
        if(power_loop %% 10 == 0) {
          cat(sprintf('Loop Number: %.0f\n', power_loop))
        }
      }

      # d[condition == 'control',   Y := rnorm(.N, mean = mean_control, sd = sd_control)]
      # d[condition == 'treatment', Y := rnorm(.N, mean = mean_treat, sd = sd_treat)]
      #
      # ate <- d[ , mean(Y), keyby = condition][ , diff(V1)]
      #
      # for(ri_loop in 1:ri_loops) {
      #   ri[ri_loop] <- d[ , mean(Y), keyby = sample(condition)][ , diff(V1)]
      #   }
      #
      # p_values[power_loop] <- mean(abs(ri) > abs(ate))

      p_values[power_loop] <- t.test(
        x = rbinom(number_per_condition, size = 1, prob = mean_control),
        y = rbinom(number_per_condition, size = 1, prob = mean_treat)
      )$p.value
    }
```

```
    return(list(
      'p_values' = p_values,
      'power' = mean(p_values < 0.05)
      ))
}
```

A 2012 study evaluated factors that increased response rates to web surveys. 1926 participants were invited by e-mail in 2008 to particpiate in a web survey. As our experiment also involves evaluating survey partcipation through e-mail invitation, we believe this study while different in design will provide a reasonable estimate of the magnitude of treatment effect size. The survey had roughly 19.7% to 32.3% between different treatment factors and a roughly 5% to 7% average treatment effect when comparing between control and treatment.

Keusch 2012 https://tinyurl.com/4dz6sb5w

```
samples_per_condition <- c(50, 60, 70, 80, 90, 100, 300, 500, 1000, 1500, 2000, 3000, 5000)

size_power <- NA
effect_power_high <- NA
effect_power_low <- NA

se = 0.112
std = se * sqrt(1926)
std_con = std
std_tre = std * 1.1

total_percent = 0.25
control = 0.25
treat = 0.30

treat_2 = 0.32
treat_3 = 0.28

for(i in 1:length(samples_per_condition)) {
  size_power[i] <- power_test_t(
    mean_control = control, mean_treat = treat,
    power_loops = 1000, verbose = FALSE,
    number_per_condition = samples_per_condition[i]
    )$power
}

for(i in 1:length(samples_per_condition)) {
  effect_power_high[i] <- power_test_t(
    mean_control = control, mean_treat = treat_2,
    power_loops = 1000, verbose = FALSE,
    number_per_condition = samples_per_condition[i]
    )$power
}

for(i in 1:length(samples_per_condition)) {
  effect_power_low[i] <- power_test_t(
    mean_control = control, mean_treat = treat_3,
    power_loops = 1000, verbose = FALSE,
    number_per_condition = samples_per_condition[i]
    )$power
```

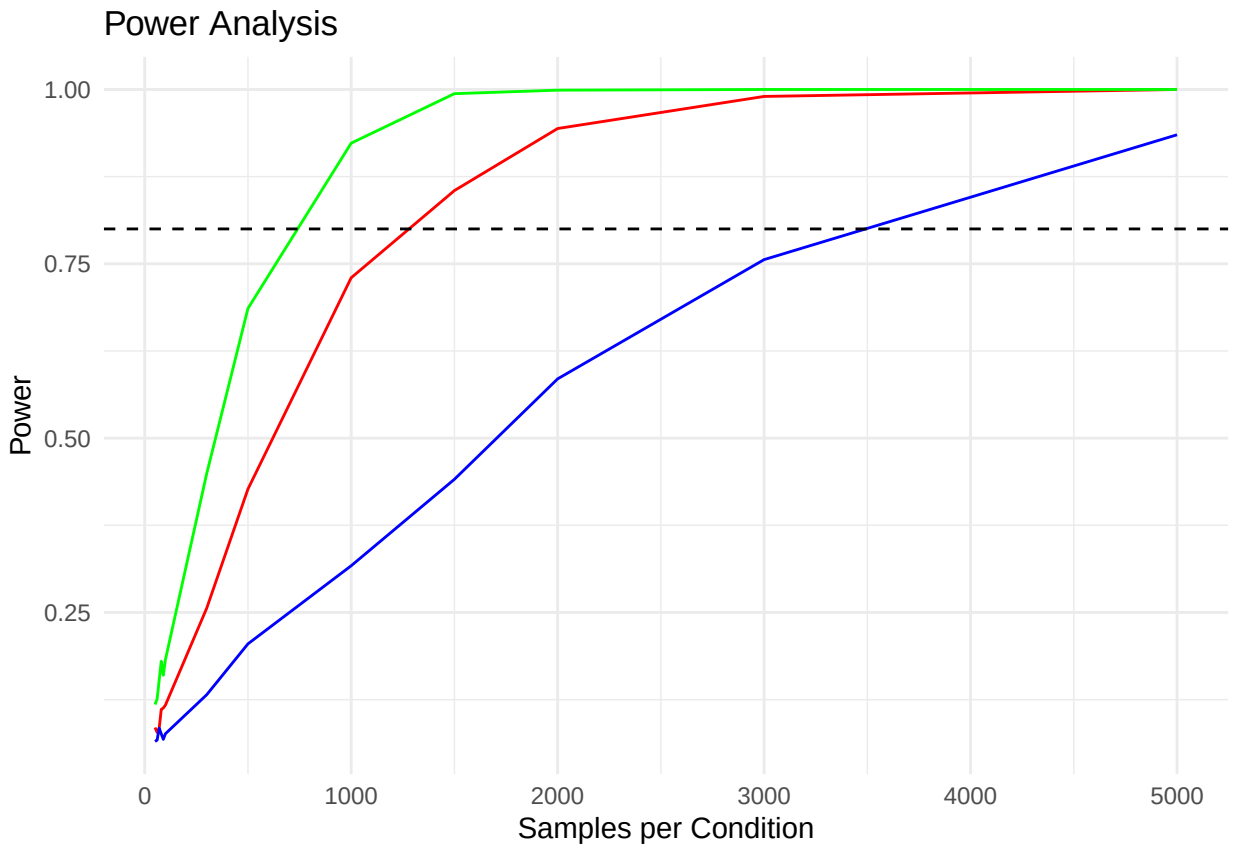
```
}
```

```
require(ggplot2)
```

```
## Loading required package: ggplot2
```

```
df <- data.frame(samples_per_condition, size_power, effect_power_high, effect_power_low)
```

```
ggplot(df, aes(samples_per_condition)) +
  geom_line(aes(y=size_power), colour="red") +
  geom_line(aes(y=effect_power_high), colour="green") +
  geom_line(aes(y=effect_power_low), colour="blue") +
  labs(x = "Samples per Condition", y = "Power", title = "Power Analysis") +
  geom_hline(yintercept = 0.8, linetype = "dashed", color = "black") +
  theme_minimal() +
  scale_colour_manual("",
                      values = c("Size Power" = "red",
                                 "Effect Power High" = "green",
                                 "Effect Power Low" = "blue"))
```



Based on the study results, we used a t-test as our statistical test as we will in our experiment. We used a binomial distribution to simulate the data of whether or not the participant responds with 1 being responded and 0 being not responded. The probability was based on the percentage response rate on the study we used as reference. We simulated 3 different scenarios with 3 different treatment effect sizes.

The red line represents our base scenario with the control group having 25% response rate and treatment with 30% response rate for sample size of 100 people split 50 for control and 50 for treatment.

The green line represents our scenario where we increase the average treatment effect from 5% to 7%. The control group has a 25% response rate and the treatment group has a 32% response rate.

The blue line represents our scenario where we lower the average treatment effect from 5% to 3%. The control group has a 25% response rate and the treatment group has a 28% response rate.

The black line represents the conventional minimum acceptable statistical power of 0.8

Based on our calculations, we will be sending out around than 7000 emails so we will have enough statistical power for our experiment based on our power analysis given that 3% is a much more conservative estimate than our lower estimate of 5% based on the study. At 3500 sample per condition size, at 3% we will meet the 0.8 threshold for statistical power.