Power Analysis

Ryan Wilson, Sumedh Shah, Devashish Kulkarni

09/29/2022

1 Introduction

For our experiment, we plan to recruit participants from Amazon mTurk to solve 1 puzzle for a financial reward. After the puzzle is solved successfully, we will ask them to complete additional optional puzzles, with variation in the text of the request. For example, one set of instructions would mention a specific benefit for others to solving the puzzle (e.g. researchers are building a math course for disadvantaged children in grade 8-10), while another set would not specify this information. Our outcome would the number of puzzle solved by participants in each group, with a maximum 5 puzzles. We plan to collect information on age, gender and ethnicity as covariates. The number of puzzles solved would be an integer from 0-5. Age will be measured in one of 5 buckets (<20, 20-35, 36-50, 50-65, >65), gender will take one of 3 values (Male, Female, Other) and ethnicity will take one of 5 values (White, Black, Asian, Hispanic, Other). Previous research on users on Amazon mTurk is used to simulate the distribution of the covariates. A previous field experiment on charitable giving in a Salvation army store found that a verbal request for a donation increased the number of donors by 55% and the total donations by 69% as compared to a silent opportunity to donate, without a verbal request. Considering both these numbers and keeping in mind and the difference in setting, our conservative estimate for the outcomes is 1 puzzle solved on average by the control group (i.e. the minimum required for the reward) and 1.5 puzzles solved on average by the treatment group. Our three scenarios would test for different effect sizes and dispersion in data. We generate a dataset with 1000 observations. Various samples of different sample sizes are sampled from this population. The sampling process is repeated 1000 times. Regression is used to fit a model on the outcome with only the treatment assignment as the regressor for each sample. The p-values generated from this regression are used to calculate power by counting the fraction of p-values < 0.05.

2 Scenarios

2.1 Low effect size (0.5), low dispersion

- Control group: 10% fails to solve the the first puzzle, 80% solve only 1 puzzle, 10% solve 2 puzzles. (mean = 1)
- Treatment group: 10% fails to solve the first puzzle, 40% solve only 1 puzzle, 40% solve 2 puzzles, 10% solve 3 puzzles. (mean = 1.5)
- No correlation between any covariates on the outcome.

2.2 Low effect size (0.5), high dispersion

- Control group: 10% fails to solve the the first puzzle, 40% solve only 1 puzzle, 40% solve 2 puzzles, 10% solve 3 puzzles. (mean = 1.5)
- Treatment group: 10% fails to solve the first puzzle, 20% solve only 1 puzzle, 40% solve 2 puzzles, 20% solve 3 puzzles, 10% solve 4 puzzles. (mean = 2)
- No correlation between any covariates on the outcome.

2.3 High effect size (1.5), high dispersion

- Control group: 10% fails to solve the the first puzzle, 40% solve only 1 puzzle, 40% solve 2 puzzles, 10% solve 3 puzzles. (mean = 1.5)
- Treatment group: 0% fails to solve the first puzzle, 10% solve only 1 puzzle, 20% solve 2 puzzles, 40% solve 3 puzzles, 20% solve 4 puzzles, 10% solve 5 puzzles. (mean = 3)
- No correlation between any covariates on the outcome.

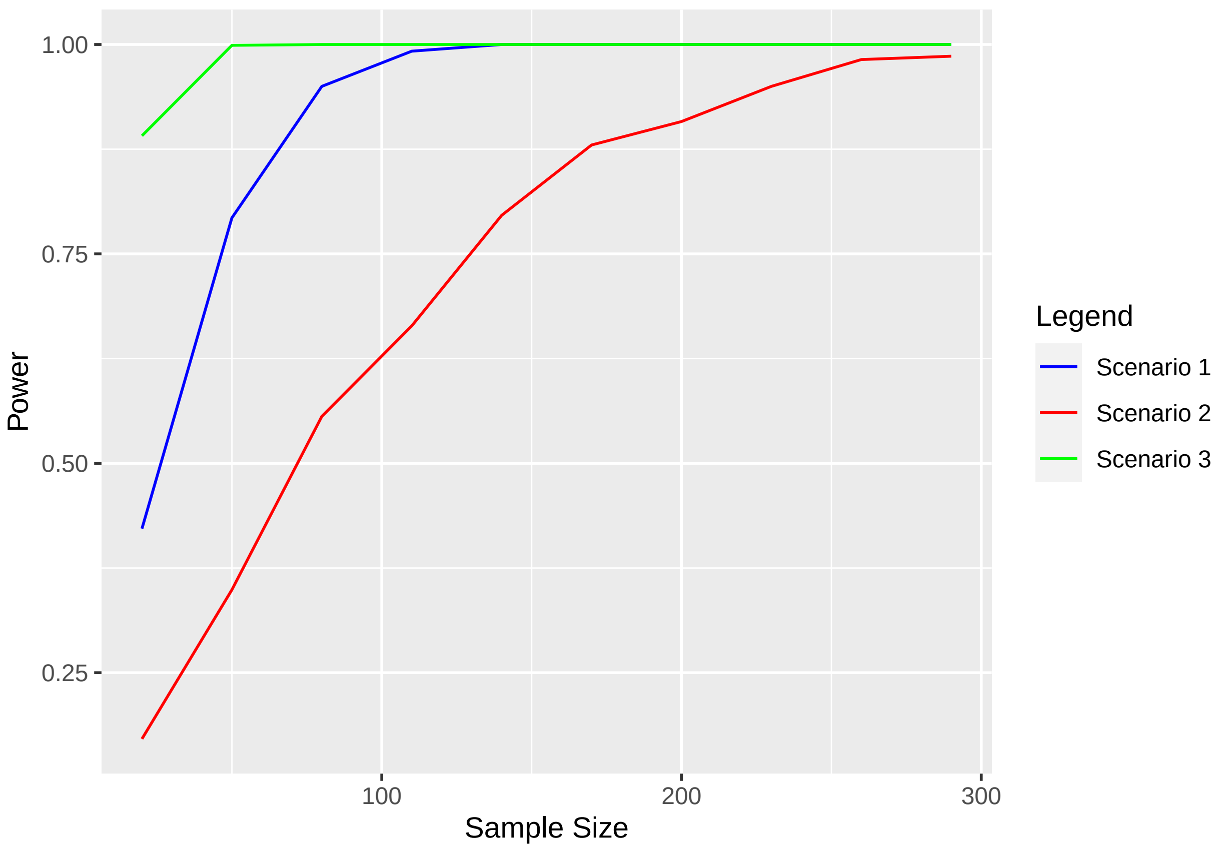
```r
rows <- 1000

d <- data.table(id = 1:rows)

# Create treatment assignment
d[ , ':='(treat = sample(c(0,1), .N, replace = TRUE))]

# Create gender covariate (55% female, 40% male, 5% other)
d[ ,  gender := sample(c('M', 'F', 'O'), size = .N, replace = TRUE,
                       prob = c(0.55, 0.40, 0.05))]

# Create Age covariate, values from 1-5:
# 1 -> <20 | 2 -> 20-35 | 3 -> 36-50 | 4 -> 50-65 | 5-> >65
# Probability estimates from the reference
d[ ,  age := sample(c(1, 2, 3, 4, 5), size = .N, replace = TRUE,
                    prob = c(.3, .37, .17, .11, .05))]

# Create Ethnicity covariate:
# W -> White, B -> Black, A -> Asian, H -> Hispanic, O -> Other
# Probability estimates from the reference
d[ ,  ethnicity := sample(c('W', 'B', 'A', 'H', 'O'), size = .N, replace = TRUE,
                          prob = c(.7, .07, .1, .1, .03))]

# Scenario 1
# Create control outcomes
d[treat == 0,  outcome_1 := sample(c(0, 1, 2), size = .N, replace = TRUE,
                                   prob = c(.1, 0.8, 0.1))]

# Create treatment outcome
d[treat == 1,  outcome_1 := sample(c(0, 1, 2, 3), size = .N, replace = TRUE,
                                   prob = c(.1, 0.4, 0.4, 0.1))]

# Scenario 2
# Create control outcomes
d[treat == 0,  outcome_2 := sample(c(0, 1, 2, 3), size = .N, replace = TRUE,
                                   prob = c(.1, 0.4, 0.4, 0.1))]

# Create treatment outcome
d[treat == 1,  outcome_2 := sample(c(0, 1, 2, 3, 4), size = .N, replace = TRUE,
                                   prob = c(.1, 0.2, 0.4, 0.2, 0.1))]

# Scenario 3
# Create control outcomes
d[treat == 0,  outcome_3 := sample(c(0, 1, 2, 3), size = .N, replace = TRUE,
                                   prob = c(.1, 0.4, 0.4, 0.1))]

# Create treatment outcome
```

```r
d[treat == 1,  outcome_3 := sample(c(0, 1, 2, 3, 4, 5), size = .N, replace = TRUE,
                                    prob = c(0, .1, 0.2, 0.3, 0.2, 0.1))]
```

3 Statistical Power

```r
## you can either write a for loop, use an apply method, or use replicate
## (which is an easy-of-use wrapper to an apply method)
percentages_to_sample <- seq(2,30,by=3)

#Sample size in each group (divides by 2)
sample_sizes = ceiling(percentages_to_sample/100*nrow(d)/2)
power_sample_size_1 <- NA
power_sample_size_2 <- NA
power_sample_size_3 <- NA

n_loops = 1000
count = 0

for(sample_size in sample_sizes){

  count = count + 1

  # fill this in with the p-values from your power analysis
  regression_p_values_1 <- NA
  regression_p_values_2 <- NA
  regression_p_values_3 <- NA

  ## you can either write a for loop, use an apply method, or use replicate
  ## (which is an easy-of-use wrapper to an apply method)

    for(sim in 1:n_loops) {

      # Sample from treatment and control groups
      treatment_group = d[treat == 1 , ][sample(.N, sample_size, replace = FALSE)]
      control_group =  d[treat == 0 , ][sample(.N, sample_size, replace = FALSE)]

      combined_group = data.table(rbind(treatment_group, control_group))

      # Fit regression model and extract p-value
      regression_p_values_1[sim] <- summary(lm(outcome_1 ~ treat,
                                    data = combined_group))$coefficients[2,4]

      regression_p_values_2[sim] <- summary(lm(outcome_2 ~ treat,
                                    data = combined_group))$coefficients[2,4]

      regression_p_values_3[sim] <- summary(lm(outcome_3 ~ treat,
                                    data = combined_group))$coefficients[2,4]

    }

  power_sample_size_1[count] = mean(regression_p_values_1<0.05)
  power_sample_size_2[count] = mean(regression_p_values_2<0.05)
```

```
  power_sample_size_3[count] = mean(regression_p_values_3<0.05)

}

df <- data.frame(sample_sizes*2, power_sample_size_1, power_sample_size_2)
colors <- c("Scenario 1" = "blue", "Scenario 2" = "red", "Scenario 3" = "green")
ggplot(df, aes(x = sample_sizes*2)) +
  geom_line(aes(y = power_sample_size_1, color = "Scenario 1")) +
  geom_line(aes(y = power_sample_size_2, color = "Scenario 2")) +
  geom_line(aes(y = power_sample_size_3, color = "Scenario 3")) +
  labs(x = "Sample Size",
       y = "Power",
       color = "Legend") +
  scale_color_manual(values = colors)
```



References

1 - https://www.cloudresearch.com/resources/blog/who-uses-amazon-mturk-2020-demographics/#:~:text=Age%20%26%20

2 - https://www.nber.org/system/files/working_papers/w17648/w17648.pdf