Power Analysis

Datasci 241

Jared Feldman, Jerry Gonzalez, Reed Evans, Randy Louie

10/10/2023

Contents

Justification for simulation variables

We performed a pilot test of 10 games with our chess bot that plays at a consistent 1500 (average) level against real opponents on lichess.com. We found that the opponents played at an average of 72% accuracy and had a 12.3 standard deviation in their accuracy rating. We used these values in our power analysis.

We are using a conservative estimate of 5-15% effect size for our power analysis. We found in the paper *Trash-talking: Competitive incivility motivates rivalry, performance, and unethical behavior* by Jeremy A. Yip, Maurice E. Schweitzer, and Samir Nurohamed. https://www.sciencedirect.com/science/article/pii/S0749597816301157

They found that trash talking had an an indirect effect on competitive performance through creation of a perceived rivalry between the players. On page 131, they show that there was an effect of b=.32, with a 95% confidence interval of 0.02, 0.87. Trash talking resulted in the opponent performing better.

Create simulation function using linear regression

```
simulate_regression <- function(recruits, mean_bad_moves, effect_size, sd_bad_moves){
    d <- data.table(
      n = 1:recruits)

    half_recruits <- floor(recruits/2)

    ## no treatment data
    d_1 <- data.table(
      id    = 1:half_recruits,
      treat = 0)

    ## assign num bad moves to treatment group
    d_1[ , Bad_moves := rnorm(.N, mean=mean_bad_moves, sd=sd_bad_moves)]

    ## treatment data
    d_2 <- data.table(
      id    = (half_recruits+1):recruits,
      treat = 1)

    ## assign num bad moves to treatment group
    d_2[ , Bad_moves := rnorm(.N, mean=mean_bad_moves*(1+effect_size), sd=sd_bad_moves)]

    ## Stack data frames
    d <- rbind(d_1, d_2)

    model_1    <- lm(Bad_moves ~ treat, data = d)
    anova_m1 <- anova(model_1)

    return(anova_m1$`Pr(>F)`)
    }
```

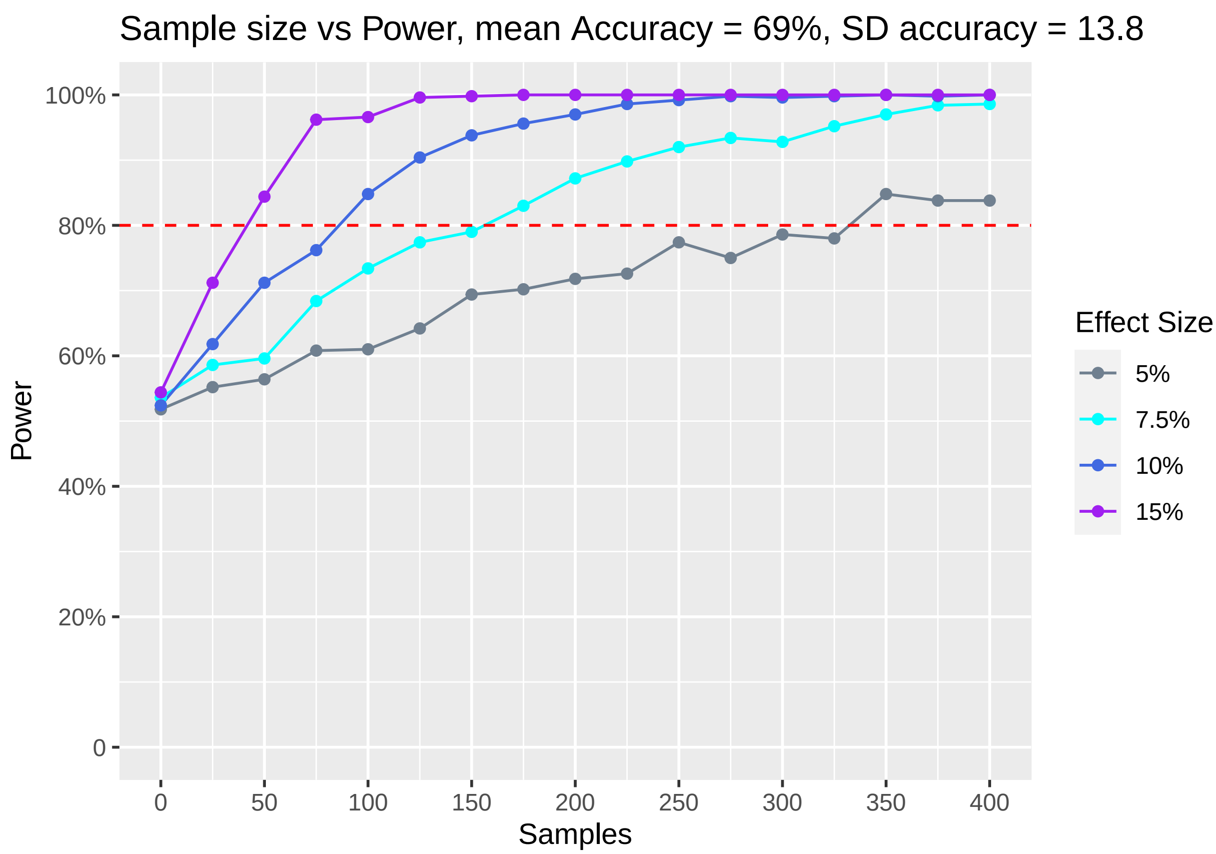
Simulate

```
# Sample sizes to simulate
steps <- seq(0, 400, by=25)

# Effect sizes to simulate
effects <- c(.05, .075, .10, .15)
```

```
# Create master data frame to aggregate the data into
g_total <- data.frame()

for (eff in effects){ #loop through effect sizes
  power_list <- data.frame()

  # Print for status
  # print(paste(toString(eff*100),"%", sep = ""))

  for (step in steps){ #loop through sample sizes

    # simulate_regression <- function(recruits, mean_bad_moves, effect_size, sd_bad_moves)
    p_vals <- replicate(250, simulate_regression(step, 69, eff, 13.7))

    power <- length(p_vals[p_vals < 0.05])/length(p_vals)
    power_list <- rbind(power_list, power)
  }
  g <- bind_cols(power_list, steps)
  colnames(g) <- c("y", "x")

  # Create column with effect size for this loop
  x<-rep(c(paste(toString(eff*100),"%", sep = "")),times=length(steps))
  g["Effect_size"] = x

  # append data to master data frame
  g_total <- rbind(g_total, g)
}
```

Plot

```
g_total$Effect_size <- factor(g_total$Effect_size, levels = c("5%", "7.5%", "10%", "15%"))

plot <- ggplot(data = g_total, aes(x=x, y=y, color = Effect_size)) +
  geom_line() +
  geom_point() +
  scale_y_continuous(name="Power", limits=c(0, 1), labels=c("0","20%","40%", "60%","80%","100%"), breaks
  scale_x_continuous(name="Samples", limits=c(0, 400), labels=c("0","50","100", "150","200","250","300"
  ylab("Power") +
  geom_hline(yintercept = 0.8, color = "red", linetype="dashed") +
  ggtitle("Sample size vs Power, mean Accuracy = 69%, SD accuracy = 13.8") +
  scale_color_manual(name="Effect Size", values =c("slategrey", "cyan", "royalblue", "purple"))
plot
```


Sample size vs Power, mean Accuracy = 69%, SD accuracy = 13.8
Power
100%
80%
60%
40%
20%
0
0
50
100
150
200
250
300
350
400
Samples
Effect Size
5%
7.5%
10%
15%