Power_Analysis_micro_goal

Fred

Power Analysis and Assumptions:

This is a 3*2*2*2*2 between group factorial experiment, some references I have so far been reading:

References:

I highly recommend for people who haven't done a lot of power analysis to go through some of the references I have below, they were really helpful for me to understand the whole picture and come up with the idea of doing it manually instead of a black-box on G*Power

1. How to unify a general sample size instead of per-cell sample size:
    1. http://jakewestfall.org/blog/index.php/2015/05/26/think-about-total-n-not-n-per-cell/
2. Sample size and p-hacking issues, what might arise and how to solve them:
    1. http://datacolada.org/17
3. How to keep account of interaction effect, in a factorial design:
    1. https://approachingblog.wordpress.com/2018/01/24/powering-your-interaction-2/
4. Test case on a 2*2 factorial experiment power analysis run by R: A lot of coding was adapted from here and changed based on our choice
    1. https://www.markhw.com/blog/power-twoway
5. understanding factorial design and statistical analysis:
    1. https://www.youtube.com/watch?v=uaWEQj18zqI
6. How to choose an ambiguous cohen-d value:
    1. https://www.spss-tutorials.com/cohens-d/

    2. https://www.youtube.com/watch?v=GDe4M0xEghs
7. There's many more stackoverflow pages for me to look into to solve bugs but I don't keep account of those anymore
8. Simulating a factorial design experiment:
    1. https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-019-0883-9
    2. https://debruine.github.io/faux/
    3. https://stats.stackexchange.com/questions/57642/simulating-responses-from-a-factorial-experiment-for-power-analysis
    4. https://rdrr.io/cran/faux/man/sim_design.html

Coding Step by Step:

Step 1: Simulation Function:

Let's figure out in this step, how to simulate such data using three settings:

1. low effect size, with a Cohen's D = 0.2
2. medium effect size, with a Cohen's D = 0.5
3. Large effect size, with a Cohen's D = 0.8

```
## ReadME

# Let's first write the function of getting mode:
Modes <- function(x) {
  ux <- unique(x)
  tab <- tabulate(match(x, ux))
  ux[tab == max(tab)]
}

# note that if we are including interaction, by default it is FALSE to not include interac
# effect between factors, set to TRUE manually if we are testing for
sim_data <- function(n, eff_size, alpha, interaction = FALSE) {
  s <- eff_size/2 # how much each data is fluctuating around 0
  f1 <- factor(sample(c("l1", "l2", "l3"), n, TRUE))
  f2 <- factor(sample(c("l1", "l2"), n, TRUE))
  f3 <- factor(sample(c("l1", "l2"), n, TRUE))
  f4 <- factor(sample(c("l1", "l2"), n, TRUE))
  f5 <- factor(sample(c("l1", "l2"), n, TRUE))
```

```
  # prior generations at the level of eff_size, or in other words, standardized cohen's D
  # distance,
  mu_1 <- ifelse(f1 == 'l1', -1*s, ifelse(f1 == 'l2', 0, s))
  mu_2 <- ifelse(f2 == 'l1', -1*s, s)
  mu_3 <- ifelse(f3 == 'l1', -1*s, s)
  mu_4 <- ifelse(f4 == 'l1', -1*s, s)
  mu_5 <- ifelse(f5 == 'l1', -1*s, s)

  mu <- cbind(mu_1, mu_2, mu_3, mu_4, mu_5)

  # Now let us simulate the dat:
  dv <- c()
  for (i in 1:n){
    dv[i] <- rnorm(1, mean(Modes(mu[i,])), 1)
    # assume a standardeviation of 1 to follow the setup of  cohen's D
    # People can change the mean of mode function to mode by writing their own, or median
  }
  # so far for model checking, I only check the case when some factor levels are significa
  if (interaction == FALSE){
      return(min(summary(lm(dv ~ f1 + f2 + f3 + f4 + f5))$coef[-1,4]) < alpha)
    # coef[2,4] and coef[3,4] looks at the p-value of factor 1,
    # as it is explored the least amount, we want to see if any level is significant
  }
  else{
    return(min(summary(lm(dv~ f1 * f2 * f3 * f4 * f5))$coef[-1,4])< alpha)
    # row 38:48 are all the five level interactions item,
    # it is safe to show those are significant values
    # (Fred made the assumption here, double check, what is your model design?)
  }
}
```

P.S. I have tested multiple cases (around ~15), and I think the model converges under all the test cases I run, but I am not CS majored and I hate writing test cases, please let me know if there are things that it doesn't work out :D

Step 2: Get-Power Function:

Now we write the function to calculate the power for such simulated model:

```
# it takes the mean value for the number of times the model is signifcant in the safest wa
get_power <- function(n, eff_size, reps, alpha, interaction = FALSE) {
  mean(
    sapply(1:reps, function(placeholder) {
      sim_data(n, eff_size, alpha, interaction)
    })
  )
}
```

Step 3: Step-wise Sample Size Power analysis function:

Now we want to extend this to a range of sample size settings, for example we start from sample size 100, end at 400, and each time update by 25, something like this, so that it is easier for people to run the code…instead of hitting their head and solving everything :D

```
power_analysis <- function(eff_size, reps, start, end, by, alpha, interaction = FALSE) {
  set.seed(447447539) # oh I have good ways finding seed, see if you can crack the mistery
  out <- lapply(
    seq(start, end, by),
    get_power,
    eff_size, reps, alpha, interaction
  )
  out <- as.data.frame(do.call(rbind, out))
  names(out) <- "Interaction Term Power"
  out$`Sample Size` <- seq(start, end, by)
  return(
    out[, c(2, 1)]
  )
}
```

Result and Conclusion for no interaction effect

I will formulate this section into 3 sections, as described before, I used a rep number of 200, and I think my computer is already dying running the simulation…for whoever have a better computer, try increase the rep number larger, although 200 rep is quite good enough

Results: Small Effect Size: 0.2

```
power_analysis(eff_size = 0.2,
               reps = 200,
               start = 1300,
               end = 1500,
               by = 25,
               alpha = 0.05,
               interaction = FALSE)
```

```
  Sample Size Interaction Term Power
1        1300                  0.760
2        1325                  0.780
3        1350                  0.805
4        1375                  0.835
5        1400                  0.805
6        1425                  0.825
7        1450                  0.830
8        1475                  0.795
9        1500                  0.815
```

So for a small effect size, we will need around 1350 participants

Results: Meium Effect Size: 0.5

```
power_analysis(eff_size = 0.5,
               reps = 200,
               start = 200,
               end = 400,
               by = 25,
               alpha = 0.05,
               interaction = FALSE)
```

```
  Sample Size Interaction Term Power
1         200                  0.725
2         225                  0.790
3         250                  0.830
4         275                  0.875
5         300                  0.880
```

```
6          325                    0.865
7          350                    0.910
8          375                    0.930
9          400                    0.910
```

so for a medium effect size, we will need around 225-250 participants

Results: Large Effect Size: 0.8

```
power_analysis(eff_size = 0.8,
               reps = 200,
               start = 50,
               end = 200,
               by = 25,
               alpha = 0.05,
               interaction = FALSE)
```

```
  Sample Size Interaction Term Power
1          50                  0.535
2          75                  0.730
3         100                  0.805
4         125                  0.900
5         150                  0.940
6         175                  0.945
7         200                  0.945
```

So for a large effect size, we will need around 100 participants

Now:

Result and Conclusion for with interaction effect

Results: Small Effect Size: 0.2

```
power_analysis(eff_size = 0.2,
               reps = 200,
               start = 2000,
               end = 2200,
```

```
                  by = 25,
                  alpha = 0.05,
                  interaction = TRUE)
```

```
  Sample Size Interaction Term Power
1        2000                  0.745
2        2025                  0.720
3        2050                  0.745
4        2075                  0.800
5        2100                  0.775
6        2125                  0.775
7        2150                  0.695
8        2175                  0.760
9        2200                  0.720
```

so we will need about 2075 data here under a small effect size

Results: Medium Effect Size: 0.5

```
power_analysis(eff_size = 0.5,
               reps = 200,
               start = 400,
               end = 600,
               by = 25,
               alpha = 0.05,
               interaction = TRUE)
```

```
  Sample Size Interaction Term Power
1         400                  0.730
2         425                  0.760
3         450                  0.770
4         475                  0.795
5         500                  0.795
6         525                  0.755
7         550                  0.790
8         575                  0.825
9         600                  0.845
```

So we will need about 550-575 data point under a medium effect size of 0.5

Results: Medium Effect Size: 0.8

```
power_analysis(eff_size = 0.8,
               reps = 200,
               start = 200,
               end = 400,
               by = 25,
               alpha = 0.05,
               interaction = TRUE)
```

```
  Sample Size Interaction Term Power
1         200                  0.760
2         225                  0.740
3         250                  0.755
4         275                  0.880
5         300                  0.875
6         325                  0.890
7         350                  0.840
8         375                  0.860
9         400                  0.895
```

so we will need about 250- 275 samples for a large effect size of 0.8

Summary

Now to summary all of these, let's create a chart

Below are results for the model lm ~ factor1 + factor2 + factor3 + factor4 + factor5:

Table 1: Optimal Sample Size under Different Effect Size

Effect Size (or standardized Cohen's Distance)	Optimal Sample Size	Power expected
Small Effect size = 0.2	1350	80.5%
Medium Effect size = 0.5	225-250	79.0% ~ 83%
Large Effect Size = 0.8	100	80.5%

Below are results for the model lm ~ factor1 * factor2 * factor3 * factor4 * factor5:

Table 2: Optimal Sample Size under Different Effect Size

Effect Size (or standardized Cohen's Distance)	Optimal Sample Size	Power expected
Small Effect size = 0.2	2075	80.0%
Medium Effect size = 0.5	550-575	79.0% ~ 82.5%
Large Effect Size = 0.8	250-275	75.5% ~ 88.0%

Alright 5 hours :D, I will ask Ananya to buy me lunch/dinner :D