Power analysis for mixed models

Martial Foegel

2024-01-19

Why bother with a power analysis ?

Why bother with a power analysis ?

- Know, before conducting an experiment, how many items, participants or other experimental settings you need in order to achieve a given power,
- Higher power allows for better replicability of the given experiment (if you have correctly rejected your null hypothesis),
- Higher power means reduced Type S and Type M error,
- Forces you to think about your statistical model through and through before doing the experiment.

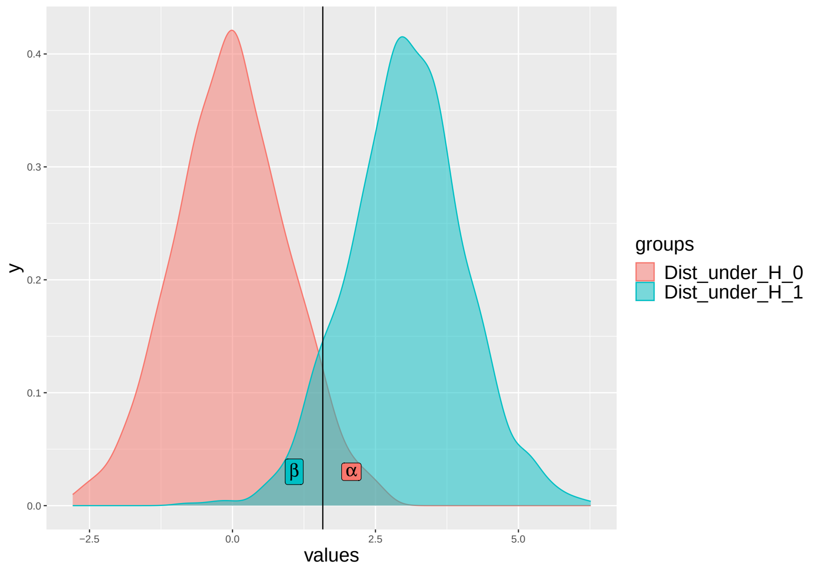
What is a power analysis ?

Power in relation to a statistical analysis

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="2">Statistical analysis result (sample)</td></tr>
<tr><td>Reject H_0</td><td>Don't reject H_0</td></tr>
<tr><td rowspan="2">Reality (population)</td><td>H_0 is true</td><td>FP
Type I error
α</td><td>TN
Correct decision
$1-\alpha$</td></tr>
<tr><td>H_0 is false</td><td>TP
Correct decision
$1-\beta$</td><td>FN
Type II error
β</td></tr>
</table>

Illustrating power

$H_0 : \mu = 0, \ H_a : \mu > 0$



Steps of a power analysis

1. Specify a null hypothesis, alternative hypothesis, alpha level and desired power.
2. Start with a small sample size and simulate your data with at least 1000 replications.
3. Calculate the power by taking the proportion of p-values that are smaller than the alpha level.
4.
    - If the power calculated is above the desired power level, stop the simulation.
    - If it is below, then increase the sample size and start again from step 2.

Power analysis for linear models

Power analysis for linear models

Here for a simple linear model with two groups where we are comparing their average.

$H_0 : \mu_1 = \mu_2$, $H_a : \mu_1 \neq \mu_2$

We are using the following parameters :

```
diff_means <- 1 #the estimated difference in means between the two groups

n_sims <- 1000 #number of simulations
alpha <- 0.01
desired_power <- 0.9
n <- 10  #sample size
n_step_size <- 10 #by how much we increase the sample size of each group


power_at_n <- c()
p_vals <- vector(length = n_sims)
```

```r
i <- 1

repeat {
  for (sim in 1:n_sims) {
    dt <- data.frame(groups = c(rep("group_1", n), rep("group_2", n)),
                     values = c(rnorm(n), rnorm(n, mean = diff_means)))
    p_vals[sim] <- anova(lm(values ~ groups, data = dt))$`Pr(>F)`[1]
  }
  # check power (i.e. proportion of p-values that are smaller than alpha-level)
  power_at_n[i] <- mean(p_vals < alpha)
  names(power_at_n[i]) <- n
  cat("Current power for sample size ", n, ": ", power_at_n[i], "\n")

  if(power_at_n[i] > desired_power){break}

  n <- n + n_step_size
  i <- i + 1
}
```

```
## Current power for sample size  10 :  0.291
## Current power for sample size  20 :  0.673
## Current power for sample size  30 :  0.876
## Current power for sample size  40 :  0.971
```

Power analysis for mixed models

Concluding remarks

Concluding remarks

- Have a good idea about what you are testing, and how to model it statistically.
- Start your power analysis with few parameters and big steps.
- If you need to check multiple p-values, plan the power analysis accordingly.
- If you are using random slopes and intercepts, you will need to work with a covariance matrix.
- A power analysis can take time : you can run it as a background job in RStudio, and parrallelize it to gain time. You can also make it run on the lab or the CNRS servers if you need to.
- Take a few more people then what your power analysis tells you to, just in case you end up excluding some.

References

Barr, Dale J. n.d. "Workshop: Simulating Power for Mixed-Effects Models | Simulating Power for Mixed-Effects Models." Accessed December 22, 2023. https://dalejbarr.github.io/aa-powersim.

Hoenig, John M, and Dennis M Heisey. 2001. "The Abuse of Power: The Pervasive Fallacy of Power Calculations for Data Analysis." *The American Statistician* 55 (1): 19–24. https://doi.org/10.1198/000313001300339897.

Kumle, Levi, Melissa L.-H. Võ, and Dejan Draschkow. 2021. "Estimating Power in (Generalized) Linear Mixed Models: An Open Introduction and Tutorial in R." *Behavior Research Methods* 53 (6): 2528–43. https://doi.org/10.3758/s13428-021-01546-0.

Miller, Jeff. 2009. "What Is the Probability of Replicating a Statistically Significant Effect?" *Psychonomic Bulletin & Review* 16 (4): 617–40. https://doi.org/10.3758/PBR.16.4.617.

Quandt, Julian. n.d. "Power Analysis by Data Simulation in R." Accessed December 19, 2023. https://julianquandt.com/post/power-analysis-by-data-simulation-in-r-part-i/#finally-an-actual-power-simulation.

Sonderegger, Morgan. n.d. *Regression Modeling for Linguistic Data*. Accessed January 5, 2024. https://mitpress.mit.edu/9780262045483/regression-modeling-for-linguistic-data/.