Power Calculation & Sample Size Estimation

By: Ashok Kumar Sharma, Ph.D.

Senior Research Bioinformatician

Digestive and Liver Diseases

Cedars Sinai Medical Center

Mail: compbiosharma@gmail.com

@ashoks773 @ashoks773 @sharma-ak

Introduction to power calculation

Power Calculation

- Power calculation is essential for ensuring that a study is properly designed to detect meaningful effects.
- **Why it matters**: Too few samples can miss real effects (underpowered), and too many can waste resources (overpowered).
- **Example**: Testing a new drug – if underpowered, we might falsely conclude has no effect when

Null Hypothesis (H_0)

- **Definition**: The null hypothesis is a statement that there is no effect or no difference in a statistical test. It serves as a starting point for statistical analysis, which researchers aim to test against.
- **Example**: In a clinical trial comparing a new drug to a placebo:
  - **H_0**: The new drug has no effect on patient recovery compared to the placebo.

H_0 ???

A null hypothesis is a form of hypothesis that is deemed TRUE until proven wrong based on experimental data

Key Concepts

Error Types

- Type I Error (α) - False Positive Rate (FPR)
- Type II Error (β) - False Negative Rate (FNR)

H_0 = The new drug has no effect on patient recovery compared to the placebo

	Study rejects H_0	Study fails to rejects H_0
H_0 is True		**True Negative Correct Outcome**
H_0 is False	**True Positive Correct Outcome**	

Example: If a clinical trial concludes that the new drug is effective when it actually isn't:

Probability of Type I Error (α): If α = 0.05, there is a 5% chance of making this error.

Example: If a clinical trial concludes that the new drug is not effective when it actually is:

Probability of Type II Error (β): If β = 0.20, there is a 20% chance of making this error.

Power of Test (1 - β)

Definition: Probability of **correctly rejecting the null hypothesis when it is false**. Higher power means a greater ability to detect an effect if there is one.

When True State = Null Hypothesis (**H_0**) is False

Study rejects **H_0** → Power (**1 – β = 0.80**) (True Positive)

Study fails to rejects **H_0** → Type II Error **(β = 0.20)** (False Negative)

Example: In the context of the drug trial, if the power is 0.80, there is an 80% chance of detecting a true effect of the drug if it exists.

Effect Size; Sample Size; and Power

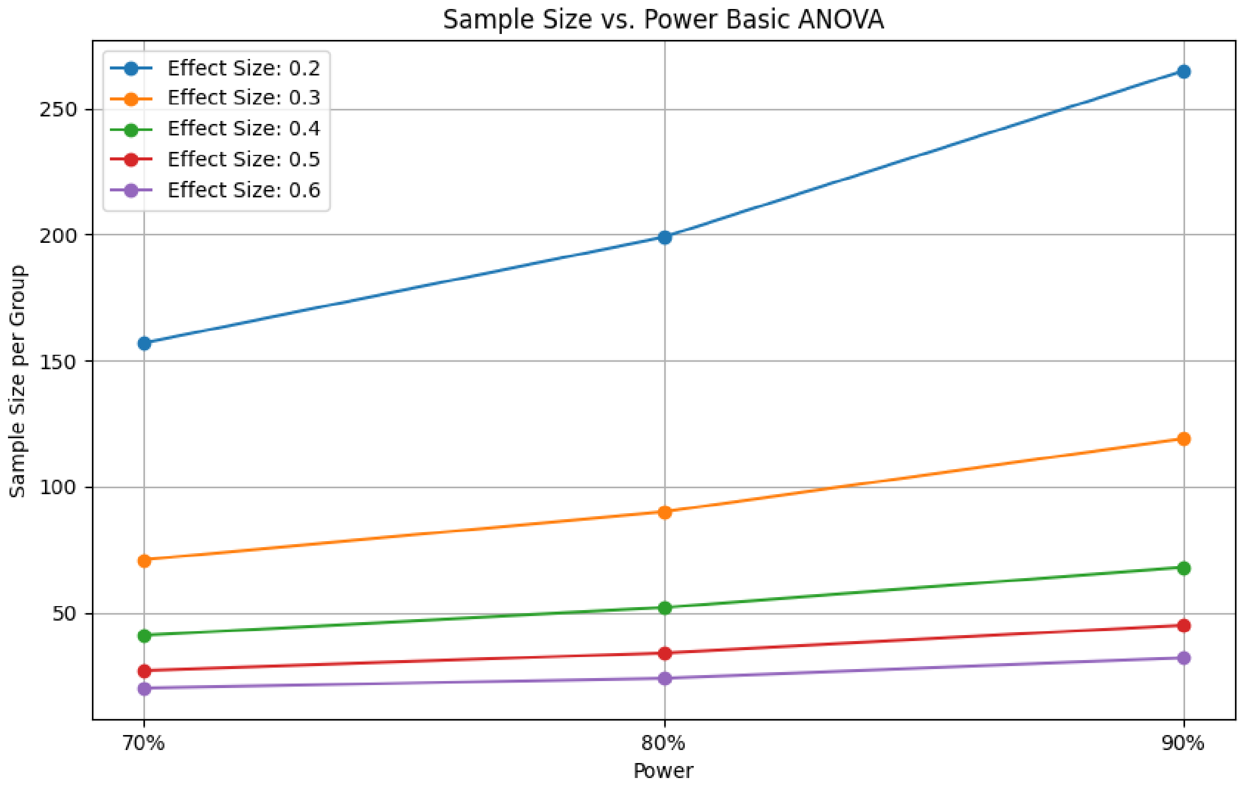
Effect Size (ES): The magnitude of the difference you're aiming to detect.

- **Definition**: Effect size is a quantitative measure of the magnitude of a phenomenon. It provides an indication of the strength of the relationship between two variables.

- **Example**: In comparing the effectiveness of two drugs:
  If the mean recovery time for Drug A is 8 days and for Drug B is 10 days, the effect size (Cohen's d) can be calculated as:

- Effect Size (Cohen's d)=$\frac{Mean(a)-Mean(b)}{Standard\ Deviation}$

Sample Size (n): How many samples you need, based on ES, and power.



Methods for power calculation

Non-Parametric Tests: When data don't meet the assumptions of parametric tests; such as normality
- Mann-Whitney U Test & Wilcoxon Signed-Rank Test
- Kruskal-Wallis Test: For two or More groups

Regression analysis:
- Linear regression: Sample size calculations for testing the effect of predictors on a continuous outcome
- Logistic regression: For binary outcomes, sample size can be estimated to detect effects of predictors on categorical outcomes

Mixed effect models: For longitudinal or clustered data
- Linear mixed models:
- Generalized linear mixed models:

Multivariate models: For Multivariate and High-Dimensional Data
- Multivariate ANOVA (MANOVA): For studies comparing multiple outcomes simultaneously across groups.
- Principal Component Analysis (PCA) and Factor Analysis: when dimensionality reduction is essential

Advanced power calculation for longitudinal studies

What is Intra Class Correlation (ICC)

- Measures how much of the total variation in the data is due to **<u>differences between subjects</u>** rather than within subjects
- A higher ICC indicates that more of the variation is due to differences between subjects (e.g., individual baseline differences), which is helpful in analyzing data with repeated measures over time

Adjusting for ICC in Power Calculations

Calculating ICC:

- **Random Effect Variance**: Variance attributable to the subject
- **Residual Variance**: Variance within subjects
- **Total Variance**: Sum of random effect and residual variances.
- **ICC Calculation**: ICC = Random Effect Variance / Total Variance

Adjusting sample size for ICC:

- Adjusted Sample Size = Sample Size / (1 + (Timepoints - 1) * ICC)

Removing Covariates Effects with LLM

Purpose: To isolate the effect of the primary variable of interest by adjusting for potential confounding variables (covariates).

Example formula: Outcome ~ Antibiotic_Use + Illness + Diet

Linear model to estimate the relationship between the outcome and covariates.

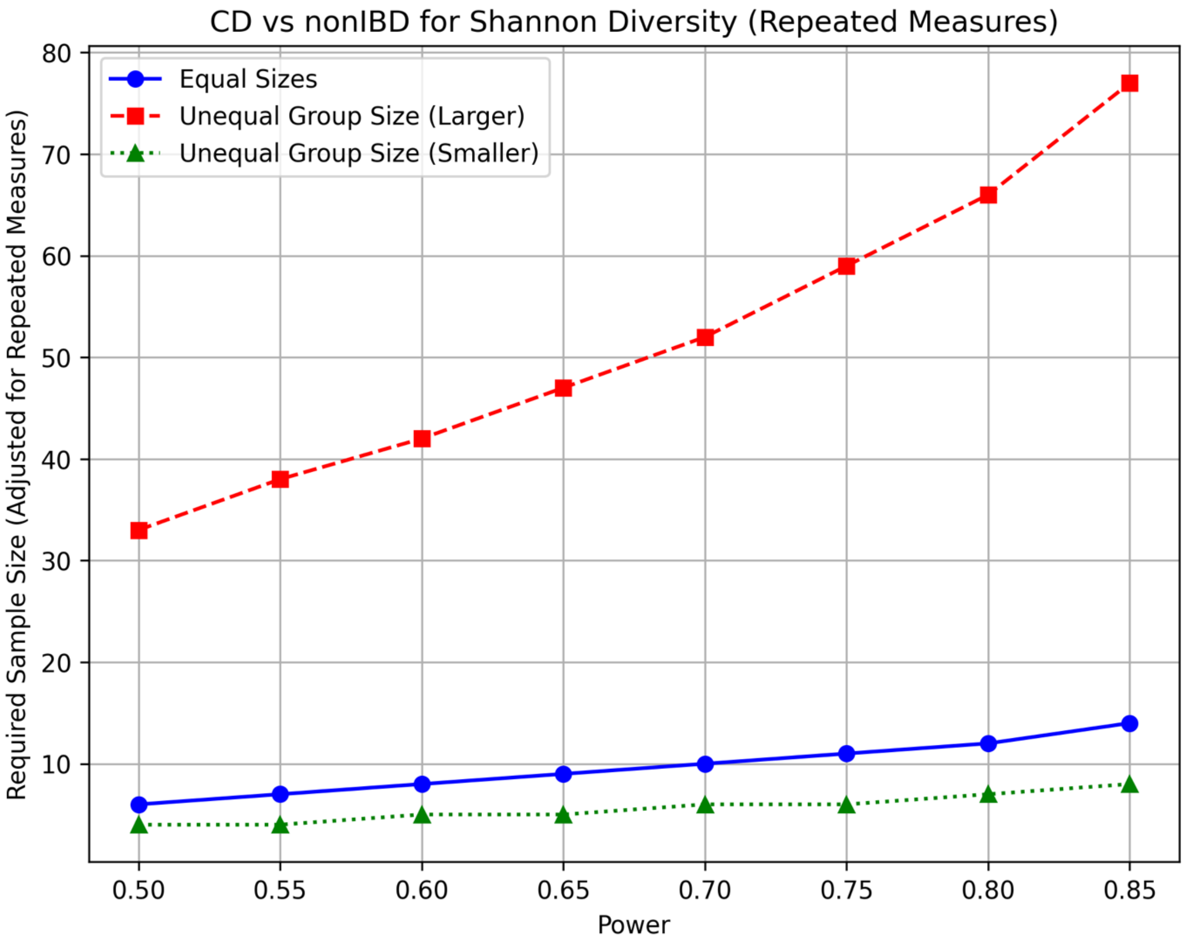
Interpretation: Coefficients reflect the adjusted relationship of each covariate with the outcome

Example study to check estimate the sample size based on microbiome diversity differences among CD vs non-IBD individuals – Highly imbalanced dataset

Dataset: HMP2- IBDMDB
https://github.com/ashoks773/PowerAnalysis/blob/main/data/CS_diversity.csv
Site: Cedars-Sinai
Total Patients: 21 with multiple time points
Total samples: 244
- CD: 224
- nonIBD: 22

Measure: Shannon diversity

Tasks:
1. Effect Size calculation
2. Interclass correlation coefficient (ICC) calculation
3. Sample size adjustment for ICC
4. **Estimated samples** for range of powers



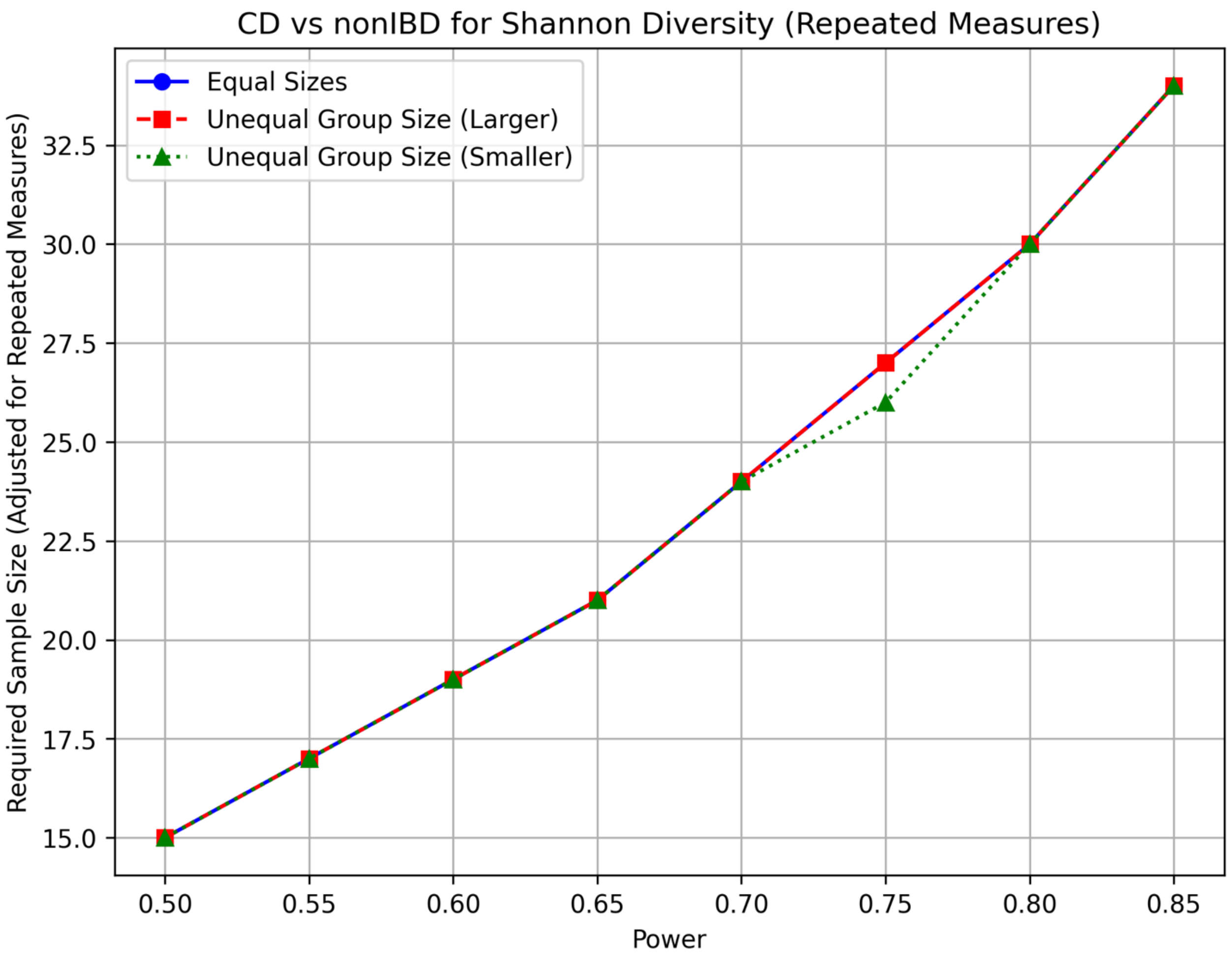
https://github.com/ashoks773/PowerAnalysis/blob/main/IBDMDB_dataset/IBDMDB_PowerAnalysis.ipynb

Dummy data to check estimate the sample size based on microbiome diversity differences among two CD groups

Dataset: Dummy (https://github.com/ashoks773/PowerAnalysis/blob/main/data/**df_cd_pilot_highICC.csv**)
Total Patients: 50
Timepoints : 11
Total samples: 550
- CD (CD_flare): 278
- nonIBD (CD_asymptomatic): 272

Measure: Shannon diversity

Tasks:
1. Effect Size calculation
2. Interclass correlation coefficient (ICC) calculation
3. Sample size adjustment for ICC
4. Get residuals to adjust for covariates
5. **Estimated samples** for range of powers



https://github.com/ashoks773/PowerAnalysis/blob/main/PowerAnalysis.ipynb