

Drug Safety Test

The dataset comes from a randomized controlled drug trial conducted by a medical group and shared by Vanderbilt University Department of Biostatistics. The study evaluates drug safety by comparing a Drug and Placebo group while tracking adverse effects, vital signs, and lab measures. You will conduct hypothesis testing using t-tests to determine statistical differences between groups.

Columns Explanation:

- **age**: Age of the participant
- **sex**: Gender of the participant (**male** or **female**).
- **trx**: Treatment group:
 - "Drug" → Received the actual drug.
 - "Placebo" → Received a placebo (control group).
- **week**: Week number in the study.
- **wbc**: White blood cell count (WBC) measurement.
- **rbc**: Red blood cell count (RBC) measurement.
- **adverse_effects**: Presence of adverse effects (**Yes** or **No**).
- **num_effects**: Number of adverse effects experienced by the participant.

1. Load the `drug_safety.csv` into a Pandas DataFrame.
2. Remove samples that contain nan whenever it is needed.
3. Display basic statistics (e.g., mean, standard deviation) for numeric columns.
4. Group the dataset by **trx** (Drug vs. Placebo) and summarize key statistics for **wbc**, **rbc**, and **num_effects**.
5. Change **adverse_effects** column so that you can define **mean** for it.
6. For each metric below, determine if they differ significantly between the Drug and Placebo groups?
 - a. mean white blood cell count

b. mean red blood cell count

c. mean num effects

d. mean adverse effect

- Formulate null and alternative hypotheses:
 - H_0 : There is no significant difference
 - H_1 : There is a significant difference
- perform an independent t-test.
- Interpret the p-value and state whether to reject or fail to reject the null hypothesis.
- If we set the p-value significant level to 0.05, which tests will fail? What about 0.1? What does this significant level mean?
- use `scipy` library for it and report:
 - What is the `alternative` argument and what did you choose for each metric, why?
 - What is the `equal_var` argument and what does it do?