

Milestone 2: Two Analysis Methods

CJ Hillbrand

Two-Tailed T-Test

The two tailed t-test is effective at comparing two distributions. This method will be used where the question asked attempts to compare two populations amongst each other. The questions that will be answered by this statistical method are:

- Q1: How does a vegetarian diet compare to a meat inclusive diet?
- Q2: Is there a price difference between the meals at Kroger, Safeway, and Trader Joes?

Question One

The procedure used to answer question one is as follows:

- Construct normalized recipes such that any given recipe can be classified as vegetarian or non-vegetarian then further normalize by nutritional value.
- Sample as many recipes as possible from each population constructing a distribution each having their own mean and variance.
- Use the mean and variance to find the required test statistic.

Question Two

The procedure used to answer question two is as follows:

- Construct normalized recipes and link ingredient information with products from each of the three stores under test.
- For each store sample as many recipes as possible from each population constructing a probability density function.
- For each pair of stores, develop the test statistic to characterize the price difference.

Null Hypothesis Testing

The null hypothesis test is effective at establishing statistical significance in a set of observations. This method is used where the question asked attempts to compare two similar or identical populations amongst each other. The questions that will be answered by this statistical method are:

- Q1: What is the price difference between buying organic and non-organic.
- Q2: How much money can consumers save by buying in bulk?

Question One

The procedure used to answer question one is as follows:

- Find the percent difference between organic products and their non-organic counterparts.

- Sample the percent difference and construct a distribution.
- Compare against the null hypothesis - the price change is 0. If the hypothesis is accepted, share the distribution of the change in price constructed in previous step.

Question Two

The procedure used to answer question two is as follows:

- Find the percent difference between buying a product in bulk and buying in smaller quantities.
- Sample the percent difference and construct a distribution.
- Compare against the null hypothesis - the price change is 0. If the hypothesis is accepted, share the distribution of the change in price constructed in previous step.

Regression Testing

Regression testing is effective in detecting a trend between two variables. This is used when the question being asked is examining the change given some variable. The questions that will be answered by this statistical method are:

- Q1: Is there a price difference between meal prepping 1 to n meals in a week?

Question One

The procedure used to answer question one is as follows:

- Construct normalized recipes.
- Construct a week's worth of meals by selecting a number of recipes such that their sum is equal to the recommend number of calories an adult should eat throughout the week.
- For a given meal plan take one meal out and replace it with a meal already existing in the plan. Do this until all meals in the meal plan are the same. Plot each point.
- Perform regression analysis on the data points, and display trend line such that the x-axis is the meals replaced, and the y axis is the cost.