3. Permutation test

A **permutation test** is a non-parametric statistical test used to determine if two or more groups differ from each other by comparing observed data to the distribution of data generated under the null hypothesis. The core idea of a permutation test is that under the null hypothesis, the labels of the groups (e.g., "before" and "after") can be permuted randomly because the groups are assumed to come from the same distribution.

**Key Concepts of the Permutation Test:**

1. **Null Hypothesis**: The null hypothesis (H0H\_0H0​) assumes that there is no difference between the groups being compared. For example, in your case, the null hypothesis is that the distribution of defects before and after the intervention is the same.
2. **Test Statistic**: This could be a difference in means, medians, or other statistics that summarize the difference between the groups. In the example earlier, we used the difference in means.
3. **Shuffling/Permuting**: Under the null hypothesis, the group labels (e.g., "before" and "after") are arbitrary. So, to test the null hypothesis, we can randomly shuffle (or permute) the data and assign them to groups many times. Each time, we recalculate the test statistic (e.g., the difference in means) for the permuted groups.
4. **Permutations**: By repeating the random shuffling many times (e.g., 10,000 times), we build up a distribution of the test statistic under the null hypothesis. This distribution shows what kinds of differences in the test statistic could happen just by chance.
5. **Comparing the Observed Test Statistic**: Once we have the permutation distribution, we compare the observed test statistic (from the original, unshuffled data) to this distribution. The proportion of permuted test statistics that are as extreme as, or more extreme than, the observed test statistic gives us the **p-value**.

**Steps of the Permutation Test:**

1. **Calculate the observed test statistic**: Compute the test statistic (e.g., the difference in means) from the original data.
2. **Combine the data**: Pool all the data together (ignoring group labels) to simulate the null hypothesis that the groups are from the same distribution.
3. **Shuffle the data**: Randomly shuffle the data and split it back into groups of the same size as the original groups. Compute the test statistic for this shuffled data.
4. **Repeat**: Repeat the shuffling process many times (typically 1,000 or more) to generate the distribution of the test statistic under the null hypothesis.
5. **Calculate the p-value**: Compare the observed test statistic to the permutation distribution. The p-value is the proportion of permuted test statistics that are as extreme as, or more extreme than, the observed statistic. This tells us how likely we are to observe such a difference by random chance.

**Example of a Permutation Test (Summary):**

Let’s take the **before** and **after** intervention defect data:

* Null Hypothesis (H0​) The defects before and after intervention come from the same distribution (i.e., the intervention had no effect).
* Test Statistic: The difference in means between the two groups.
* Process:
  1. **Observed statistic**: Calculate the actual difference in means between the "before" and "after" groups.
  2. **Permutation**: Shuffle the data randomly to mix up the "before" and "after" labels and calculate a new difference in means.
  3. **Repeat**: Do this many times (e.g., 10,000 times) to generate a distribution of the test statistic under the null hypothesis.
  4. **Compare**: Check how often the shuffled differences are as large or larger than the observed difference to estimate the p-value.

**Conclusion:**

* If the p-value is small (usually <0.05), you reject the null hypothesis, indicating that the intervention likely had an effect.
* If the p-value is large, you fail to reject the null hypothesis, meaning there’s not enough evidence to conclude that the intervention had an effect.

**Why Use a Permutation Test?**

* **No distribution assumptions**: Unlike parametric tests (like t-tests), permutation tests don’t assume that the data follow a normal distribution or any specific distribution.
* **Works with small sample sizes**: Permutation tests are particularly useful when you have small samples or when standard assumptions of parametric tests are questionable.