0.0.1 Question 1c

Before we write any code, let's review the idea of hypothesis testing with the permutation test. We first simulate the experiment many times (say, 10,000 times) through random permutation (i.e., without replacement). Assuming that the null hypothesis holds, this process will produce an empirical distribution of a predetermined test statistic. Then, we use this empirical distribution to compute an empirical p-value, which is then compared against a particular cutoff threshold in order to accept or reject our null hypothesis.

In the below cell, answer the following questions: * What does an empirical p-value from a permutation test mean in this particular context of birthweights and maternal smoking habits? * Suppose the resulting empirical p-value $p \leq 0.01$, where 0.01 is our p-value cutoff threshold. Do we accept or reject the null hypothesis? Why?

- The empirical p-value in this scenario is the percentage of permutations whose absolute value of the difference between the average weight of smoker parents' and nonsmoker parents' babies (the empirical distribution of the test statistic) is greater than or equal to the observed absolute value of the difference (the observed test statistic). Since our null hypothesis is that the two groups of parents should have the same weights of children, then the absolute value of the difference of their average weights should be zero, and we should expect a distribution centered around zero for our empirical distribution of the test statistic under the null hypothesis. Because our alternative hypothesis is that the average weights of smokers' children will just be *different* than the average weights of nonsmokers' children, we can just count the number of absolute value of differences of weights in the empirical distribution that are greater than or equal to the observed absolute value difference in weight, and divide that by the total number of permutations/repetitions performed.

- If the resulting empirical p-value is less than our threshold, than we can reject the null hypothesis. The empirical p-value tells us that over 99% of the time, the permutation test yielded a simulated test statistic of absolute difference of average weight between smoker and nonsmoker mothers that was greater than the observed test statistic; in other words, there was a difference in mean weights between the two groups. There is still a chance that the null-hypothesis holds true, and that the simulations for the permutation test simply resulted in a changed test statistic while the reality is that the average weights of the two groups are indeed the same; however, this chance is extremely small, and thus we will reject the null-hypothesis.
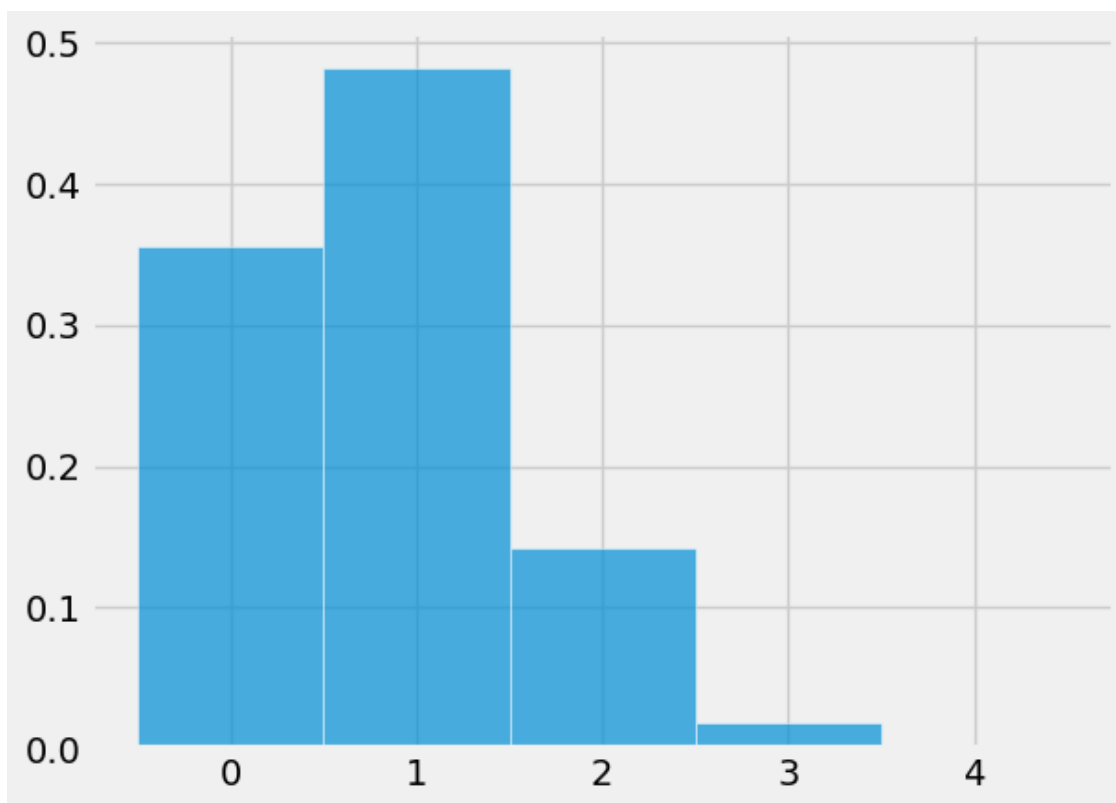
0.0.2 Question 1e

The array `differences` is an empirical distribution of the test statistic simulated under the null hypothesis. This is a prediction about the test statistic, based on the null hypothesis.

Use the `plot_distribution` function you defined in an earlier part to plot a histogram of this empirical distribution. Because you are using this function, your histogram should have unit bins, with bars centered at integers. No title or labels are required for this question.

Hint: This part should be very straightforward.

```
In [20]: plot_distribution(differences)
```

```
Out[20]: (array([0.3558, 0.4825, 0.142 , 0.0179, 0.0018]),
          array([-0.5,  0.5,  1.5,  2.5,  3.5,  4.5]),
          <BarContainer object of 5 artists>)
```

0.0.3 Question 1g

Based on your computed empirical p-value, do we reject or fail to reject the null hypothesis? Use the p-value cutoff proposed in Question 1c of 0.01, or 1%.

We can reject the null-hypothesis, because our computed empirical p-value of .0026 is less than the cutoff of .01. This means that our observed statistic is statistically relevant, and thus we reject the null-hypothesis that, on average, smoking and non-smoking mothers give birth to children with the same weight