Mangoes: Ripe for Analysis

I. Introduction

Summer is nearly here, and that means mango season has arrived. Many fruits are ready for harvest during this time of year like strawberries, watermelons, bananas, and plums. However, the mango especially brings immense joy and excitement to people across the world for its vibrant yellow and orange colors and unique sweet taste, solidifying its position as one of the world's most beloved fruits. During the summer months, more stores, grocers, and farmers markets feature mangoes as signature items that allow more people to buy and enjoy this wonderful fruit.

However, one problem buyers consistently face with mangoes is the time it takes for them to ripen. Mangoes, like other fruits that undergo ripening, release an odorless gas called ethylene that promotes the ripening process, and the accumulation of this gas eventually leads to the fruit becoming ripe. To speed up this often lengthy process, many people utilize various methods to expedite the ripening process. Some use the paper bag method, which aims to trap the ethylene produced by the mangoes in the enclosed space of the paper bag and thus hasten the ripening of the mango, while others use the rice method, in which mangoes are placed in a bowl and covered by uncooked rice which absorbs the ethylene instead of letting it escape. Do these methods really affect the time it takes for mangoes to ripen?

II. Statistical Question

Is there a difference in the stage of ripening reached by mangoes when placed in unsealed empty containers, sealed empty containers, and unsealed containers filled with rice?

Hypotheses:
H₀: There is no difference in the true distribution of the stage of ripeness reached by mangoes when stored in unsealed empty containers, sealed empty containers, and unsealed containers filled with rice.

Hₐ: There is a difference in the true distribution of the stage of ripeness reached by mangoes when stored in unsealed empty containers, sealed empty containers, and unsealed containers filled with rice.

III. Data Collection

Due to there being a wide variety of different mangoes, we decided to test only one species in order to eliminate the confounding possibility of different mangoes ripening at different times. We concluded that we should test mangoes from New India Bazaar, a supermarket in Dublin, California. We decided to use a cluster sampling method to obtain our sample of mangoes, with crates of mangoes being our clusters. In New India Bazaar, they only sell mangoes of a particular species called Haden mangoes in crates of 9 each. The store had 42 total crates of mangoes available, so we decided to buy 4 crates of Haden mangoes. First, we assigned each of the 42 crates of mangoes available at the store to a number from 01 - 42. Then, using a random number generator, we generated numbers until we got 4 unique 2-digit numbers, ignoring 00 and 43-99, which were 16, 24, 32, and 37. We then bought the crates that corresponded to these numbers. This random sampling method allows us to make an inference about a population, meaning we can generalize the results of our experiment to all Haden mangoes sold at New India Bazaar in Dublin, California.

With the mangoes bought and ready for experimentation, we needed to assign the mangoes to the different treatments. In total, we would test 3 treatments for speeding up the
ripening process for mangoes. We decided to keep all mangoes in all treatment groups inside the same type of plastic container, to remove the confounding variable of mangoes being kept in different materials as we do not know the effect of paper on the mango ripening process and if it is different than that of plastic. Treatment 1 would serve as a control, with the mangoes being placed in an empty unsealed plastic container. Treatment 2 would simulate the paper bag method by creating a sealed environment to trap the ethylene, but we would use plastic containers instead of paper. Treatment 3 would be placing the mangoes inside an unsealed plastic container filled with equal amounts of uncooked rice, to simulate the rice method. The purpose of Treatment 1, our control treatment, is to establish a baseline that allows us to determine if the other treatments actually affect mango ripening compared to a normal environment.

To randomly assign each of our 36 mangoes to our three treatments, we once again used a random number generator. We labeled each mango with a number from 1 to 36, and we generated 12 random and unique numbers, without skips or repetitions, and assigned this first group of the corresponding 12 mangoes to Treatment 1. Then, we generated 12 more unique numbers that were different from those assigned to our control treatment, and we assigned the mangoes corresponding to these 12 numbers to Treatment 2. The remaining 12 mangoes were assigned to Treatment 3. This random assignment of mangoes to treatments is crucial for our experiment because it eliminates any chance differences in the individual mangoes in our sample, thus creating groups of mangoes that were largely similar to each other. Therefore, we are confident in claiming that any differences in our experimental results between the three treatment groups were a result of the treatments themselves, and not chance variation.

In conducting the experiment, we placed all of the mangoes from each treatment group in the same room with the same lighting conditions, and we conducted the experiment for each
treatment group simultaneously. We also set the thermostat to maintain the same temperature for the entire duration of the experiment. This creates similar environmental conditions for each treatment group, thus removing time, temperature, and lighting conditions as possible confounding variables. We ran the experiment for four days in total due to our research clarifying that the average minimum amount of time it takes for a mango to ripen is four days. After 96 hours of leaving the mangoes in the room to ripen, we checked the mangoes and determined their stage of ripeness. We are using a chi-square test for homogeneity because the random assignment of mangoes to three treatment groups establishes separate populations with a categorical variable of what stage of ripeness the mango is in.

(All 36 mangoes, organized in rows based on their treatment group. Note: The mangoes in rice are not fully submerged in the picture to show that the mangoes are in the container. When conducting the experiment, the mangoes were fully engulfed in an equal amount of rice.)

To determine the stage of ripeness that each mango was in, we referred to an article by a reliable organization called Mango.org, which was linked on the crate of mangoes we bought. That detailed the five stages of ripeness that a mango goes through. These stages are determined by three variables, namely the color of the internal flesh, the taste, and the firmness or rigidity of the mango. Stage 1 is where the mango is pale, crunchy and sour, while at Stage 5 the mango is bright orange, soft, and sweet.
Screenshots of the raw data of the stage of ripeness that the mangoes are in for each treatment group from top to bottom, control, sealed container, and rice container. We individually classified the look or inner flesh, texture, and taste, then took the most common of those numbers and determined the overall stage of ripeness, as shown in red text. We cut the mangoes in halves for the pictures, so the unlabelled halves are from the same mango as the labeled half to its right.)
IV. Data Display

Observed Counts:

<table>
<thead>
<tr>
<th></th>
<th>≤ Stage 3</th>
<th>&gt; Stage 3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>9</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>Sealed</td>
<td>4</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>Rice</td>
<td>3</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>20</td>
<td>36</td>
</tr>
</tbody>
</table>

Expected Counts:

<table>
<thead>
<tr>
<th></th>
<th>≤ Stage 3</th>
<th>&gt; Stage 3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>5.33</td>
<td>6.67</td>
<td>12</td>
</tr>
<tr>
<td>Sealed</td>
<td>5.33</td>
<td>6.67</td>
<td>12</td>
</tr>
<tr>
<td>Rice</td>
<td>5.33</td>
<td>6.67</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>20</td>
<td>36</td>
</tr>
</tbody>
</table>

Chi-Square Contributions

<table>
<thead>
<tr>
<th></th>
<th>≤ Stage 3</th>
<th>&gt; Stage 3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>2.53</td>
<td>2.02</td>
<td>4.55</td>
</tr>
<tr>
<td>Sealed</td>
<td>0.33</td>
<td>0.27</td>
<td>0.60</td>
</tr>
<tr>
<td>Rice</td>
<td>1.02</td>
<td>0.81</td>
<td>1.83</td>
</tr>
<tr>
<td>Total</td>
<td>3.88</td>
<td>3.10</td>
<td>6.98</td>
</tr>
</tbody>
</table>

**Frequency of Mangoes that Reached At Most and Greater than Stage 3 Ripeness**
(Uses frequency, not relative frequency, as the number of mangoes in each treatment is the same)

V. Data Analysis

The bar graph above indicates that Treatment 2 and Treatment 3 both had a larger number of mangoes above Stage 3 ripeness than Treatment 1, as the control only has 3 mangoes with greater than Stage 3 ripeness while the sealed empty container and the unsealed rice container have 8 and 9, respectively. Treatment 2 and Treatment 3 also have less mangoes with Stage 3 or less ripeness, only 4 and 3 respectively, while Treatment 1 has 9 mangoes with less than or equal to Stage 3 ripeness. Thus, the bar graph shows a clear difference in the stage of ripeness of the mangoes between the three treatments, and now we can perform an inference procedure to determine whether this difference is statistically significant.

Conditions for a Chi-Square Test for Homogeneity:

1. **Random**: Our mangoes were randomly assigned to each of our three treatment groups using a random number generator, so this condition is met in our experiment.

2. **Independence**: Since this is an experiment, we do not need to check the independence, or 10%, condition.

3. **Large Counts Condition**: Our expected counts of 5.33 and 6.67 are all greater than 5, as shown in the expected counts table above, so this condition is met in our experiment.

Calculations:

Significance Level: \( \alpha = 0.05 \)

\[
\text{df} = (\text{rows} - 1)(\text{columns} - 1) = (3-1)(2-1) = (2)(1) = 2
\]

\[
\chi^2 = \frac{(\text{observed} - \text{expected})^2}{\text{expected}} = \frac{(9-5.33)^2}{5.33} + \frac{(3-6.67)^2}{6.67} + \frac{(4-5.33)^2}{5.33} + \frac{(8-6.67)^2}{6.67} = 6.98
\]

P-value = 0.0306
VI. Conclusion

As the P-value (0.0306) is less than the significance level (0.05), or \( p < \alpha \), we can reject the null hypothesis based on our data. We have convincing evidence that the time it takes for mangoes to ripen is different depending on if they are stored in an empty and unsealed container, an empty and sealed container, and an unsealed container filled with rice. Our P-value indicates that there is a 3.06\% chance of getting a difference as great or greater than we did in our results if the null hypothesis (the population distributions are the same) is correct.

VII. Reflection

Our project sought to determine whether there is a difference in the stages of ripeness reached by mangoes after being in a normal and unsealed environment, submerged in rice, or a sealed environment for 4 days. As a result of our research, we were able to find a reliable and comprehensive way to classify ripeness in a mango, through using stages, rather than relying on just exterior color or texture which would be inaccurate on their own. We also eliminated confounding variables through random assignment as well as controlling environmental factors, such as using the same plastic container, testing the mangoes in the same room at the same time, and using the same lighting conditions. However, there were areas where our experiment was limited and could be improved.

One aspect of our experiment that could be improved is that we utilized a relatively small sample size of 36 mangos total, with 12 in each treatment group. As mangoes are quite expensive compared to other fruits, we were limited in how large of a sample we could use. However, attaining a larger sample size of more mangoes would allow us to get more accurate results by reducing the effects of chance variability between the mangoes in the sample. Another limitation
of our experiment is that we randomly sampled mangoes from only one store, New India Bazaar, and from one species, the Haden mango, thus only enabling us to generalize our results to Haden mangoes from New India Bazaar in Dublin, California. We also only used a chi-square test for homogeneity so we could not determine which of the treatments is the best for ripening mangoes quicker, only that there is a difference in the stage of ripeness reached by the mango with each treatment. Instead, we could have performed another inference procedure like a two-sample z-test for proportions, a two-sample t-test for means, or a matched pairs t-test. This would be informative to do in the future, as we can now determine which of the treatments we tested, between the rice or sealed environment method, is actually better for quicker ripening.

Based on the conclusion of our experiment, we could have made a Type 1 error since we were able to reject the null hypothesis. In the context of our experiment, making a Type 1 error means that we found convincing evidence that there is a difference in the stage of ripeness reached between the Haden mangoes that are in an unsealed and empty container, a sealed and empty container, and an unsealed container filled with rice, when in reality there is no difference in the stage of ripeness reached by the Haden mangoes between these different treatments. The probability of making a Type 1 error in our experiment is equal to our significance level of 0.05. Looking at our Chi-Square Contributions table, we see that the two cells that contribute the most to our chi-square statistics are the control group with mangoes that achieved at most Stage 3 of ripeness in the 4 day experiment (2.53) and the control group that achieved greater than Stage 3 ripeness over the 4 days (2.02), meaning these contributed the most to our chi-square statistic, so it was the group that most deviated from what was expected.

For future study, we would like to have more accurate data and broaden the results of our experiment further. To do this and not be as affected by variability between mangoes in a sample,
we should collect a larger sample and thus have more mangoes in each treatment. This would require more funds, so we would have to plan to save money and allocate costs towards this experiment for future study. To be able to generalize our results to more than one mango species, it would be interesting to conduct a randomized block experiment where we take a random sample of all mangoes in a store, block the mangoes based on their species, and then conduct the experiment in each block. This would allow us to see whether the treatments are still effective or whether their effectiveness varies depending on the species of mango. We can also take a cluster sample of stores selling mangoes from all of the United States, thus enabling us to generalize the results of a future experiment to mangoes from the entire country.

However, the results of our experiment do have practical benefits and implications that anyone can use in their daily lives. A wide variety of other fruits, like avocados, bananas, plums, and peaches, undergo ripening very similar to mangoes, thus the treatments we used in our experiment may be extended to these fruits as well. With future experiments, we can determine which method is best for ripening and apply these results to many more fruits, including different species of mango, allowing us to help people make the most of their summer holidays.

VIII. Works Cited

