

Goodness of Fit Homework

1 A grocery store sells four different sizes of a popular brand of corn flakes. For the past few years the proportion of boxes they sell of each size has been quite stable: 10% Small, 15% Medium, 60% Large, and 15% Jumbo. They decide to change the pricing of the four sizes and want to see if this changes the proportion of boxes they sell of each size. To test this, a few weeks after changing the prices they take a simple random sample of 120 transactions involving corn flakes and count how many boxes of each size were sold. Here are the results.

Observed number of boxes sold for each box size

| Small | Medium | Large | Jumbo |
|-------|--------|-------|-------|
| 8 | 24 | 61 | 27 |

(a) We wish to carry out a test of significance to see if the distribution of sizes of cereal boxes sold has changed. State the null and alternative hypotheses for this test.

(b) Find the expected counts for each size box under the assumption that the null hypothesis is true.

Expected number of boxes sold for each box size

| Small | Medium | Large | Jumbo |
|-------|--------|-------|-------|
| | | | |

(c) Discuss whether the conditions for this test have been met.

(d) Find the value of the test statistic and the P -value of the test, and make the appropriate conclusion. Use $\alpha = 0.05$.

(e) Based on your answer to (d), which error is it possible that you have made, Type I or Type II? Describe that error in the context of the problem.

(f) Use the components of the chi-square statistic to perform a follow-up analysis on the impact of the new prices on the sales of different sizes of cereal boxes.

2. Kellogg's Froot Loops cereal comes in six fruit flavors: orange, lemon, cherry, raspberry, blueberry, and lime. You pour out your morning bowl of cereal and methodically counted the number of cereal pieces of each flavor. Here is the data:

| Flavor: | Orange | Lemon | Cherry | Raspberry | Blueberry | Lime |
|---------|--------|-------|--------|-----------|-----------|------|
| Count: | 28 | 21 | 16 | 25 | 14 | 16 |

Test the null hypothesis that the population of Froot Loops produced by Kellogg's contains an equal proportion of each flavor.

3. Blast from the recent past:

Do students who read more books for pleasure tend to earn higher grades in English? A simple random sample from 79 students at a large high school was taken. Students classified as light readers if they read fewer than 3 books for pleasure per year. Otherwise, they were classified as heavy readers. Each student's average English grade for the previous two marking periods was converted to a GPA scale where 4.0 = A, 3.0 = B, and so on. The summary statistics for the two groups are below:

| Type of Reader | N | Mean | StDev | SE Mean |
|----------------|----|-------|-------|---------|
| Heavy | 47 | 3.640 | 0.324 | 0.047 |
| Light | 32 | 3.356 | 0.380 | 0.067 |

(a) Explain why it is acceptable to use a two-sample t -procedure in this setting (check conditions).

(b) The 95% confidence interval for the difference in the mean English grade for light and heavy readers is (0.1164, 0.4516). Interpret that interval in context to someone not taking statistics.

(c) Does the interval in part (b) provide convincing evidence that reading more causes an increase in students' English grades? Justify your answer.