Final Review – Part 5: Chi-Square Hypothesis Tests

Basics

- A hypothesis test problem will ask you to test a claim.
- Look for a significance level, not a confidence level.
- Use the 5 step process.
  1. State H0 & H1. *2. State the level of significance α.*
  3. State which test you are performing.
  4. State the calculated value of the test statistic and P-value from StatCrunch.
  5. Make a decision about H0. (Reject H0 or Fail to Reject H0)
  Make a conclusion about H1. (There {is/is not} sufficient evidence to conclude that "H1 is true".)
- These tests will use 1 sample of categorical data. Either you will have one sample broken into 3 or more categories, or 1 sample divided simultaneously by two categorical variables.

Goodness of Fit Hypothesis Test

- The wording should indicate that we are testing a claim about a population being divided into 3 or more categories.
- The claim will either list percentages for each category, or that all of the categories are equal.
- The sample information provided will be observed frequencies or counts for each category.
- The data gathered would be categorical, not numerical.
- StatCrunch steps:
  For each category enter the expected frequencies in one column.
  Compute the expected frequency for each category by multiplying the sample size (*n*) by the claimed proportion for each category.
  Stat > Goodness of Fit > Chi-Square Test
  Select the columns containing the observed and expected frequencies, and press compute! for test statistic & P-value.

Independence Hypothesis Test

- The wording should indicate that we are testing whether two categorical variables are related.
- The claim often contains the word "independent".
- The sample information provided will be in a contingency table.
- The data gathered would be categorical, not numerical.
- StatCrunch steps:
  Type table in the spreadsheet.
  Stat > Tables > Contingency > Summary
  Select the columns containing the numerical counts, and the column containing the row categories.
  Press compute! to find the test statistic and P-value.

Examples

1) A random sample of 200 COS students were asked their opinion on sushi. Here are the results, broken down by gender.

Gender	Love It	Hate It	Never Tried It
Male	30	20	30
Female	65	35	20

At the 0.05 level of significance, test the claim that a person's opinion on sushi is independent of gender.

"Test the claim that …" & "0.05 level of significance" → Hypothesis Test Problem

Contingency table is provided, "independent" in claim, the data are categorical → Hypothesis Test for Independence

***ANSWER**:*

1. H_0: Sushi opinion is independent of gender.

H_1: Sushi opinion is dependent on gender.

2. $\alpha = 0.05$

3. Independence Test

4. Chi Square = 11.44, P-value = 0.0033

5. Reject H_0.

There is sufficient evidence to conclude that sushi opinion is dependent on gender.

2) A baseball expert claims that 50% of players bat right-handed, 30% bat left-handed, and 20% are switch hitters. A sample of 250 college players produced the following results. At the 0.05 level of significance, test the expert's claim.

Right-Handed	Left-Handed	Switch Hitter
140	85	25

"Test the claim that …" & "0.05 level of significance" → Hypothesis Test Problem

The data are categorical, the population & sample are divided into 3 categories, the claim mentions 3 different %'s → Hypothesis Test for Goodness of Fit

***ANSWER**:*

1. H_0: p1 = 0.50, p2 = 0.30, p3 = 0.20

H_1: At least one proportion is different than claimed.

2. $\alpha = 0.05$

3. Goodness of Fit Test

4. Expected Frequencies: 250(0.50) = 125, 250(0.30) = 75, 250(0.25) = 50

Chi Square = 15.63, P-value = 0.0004

5. Reject H_0.

There is sufficient evidence to conclude that at least one proportion is different than claimed.

3) A random sample of 400 COS students were asked which season they were born in: spring, summer, fall, or winter. The results are shown below. At the 0.05 level of significance, test the claim that students are equally likely to be born in any of the 4 seasons.

Spring	Summer	Fall	Winter
115	96	106	83

"Test the claim that …" & "0.05 level of significance" → Hypothesis Test Problem

The data are categorical, the population & sample are divided into 4 categories, the claim refers to 4 equal categories → Hypothesis Test for Goodness of Fit

***ANSWER**:*

1. H_0: p1 = p2 = p3 = p4

H_1: At least one proportion is different than claimed.

2. $\alpha = 0.05$

3. Goodness of Fit Test

4. Expected Frequencies: 400(0.25) = 100 or 400/4 = 100 for each category

Chi Square = 5.66, P-value = 0.1294

5. Fail to Reject H_0.

There is NOT sufficient evidence to conclude that at least one proportion is different than claimed.