10.1: The Chi-Square Distribution

In previous lessons, we have learned about binomial experiments. A binomial experiment has n independent trials with only two outcomes per trial (success and failure). We will now consider **multinomial experiments** in which we allow two or more outcomes per trial in n independent trials. A binomial experiment is a special case of a multinomial experiment. We will explore multinomial experiments by returning to a hypothesis test for a single population proportion.

College Completion Rates

Research on college completion has shown that about 60% of students who begin college eventually graduate. A publication of higher education claims that the proportion for STEM (science, technology, engineering, math) majors is different. Researchers randomly select 102 STEM majors and determine that 51 eventually graduate.

Step 1.

We will let p represent the proportion of all STEM majors who begin college and ultimately graduate. The null hypothesis is $H_0 : p = 0.60$. The alternative hypothesis is $H_a : p \neq 0.60$. The publication authors have the burden of proof and must produce evidence to support their claim that the proportion of college graduates among STEM majors is different against the assumption that it is not.

Step 2.

Recall, the null hypotheses is

$$H_0 : p = 0.60$$

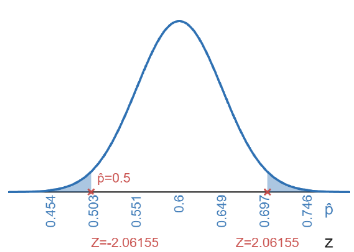
The number of expected successes in the sample is $np = 102(0.60) = 61.2 \geq 10$. The number of expected failures in the sample is $n(1-p) = n - np = 102 - 61.2 = 40.8 \geq 10$. Therefore, the sampling distribution of sample proportions is approximately normal.

The sample proportion is

$$\hat{p} = \frac{x}{n} = \frac{51}{102} = 0.5$$

Step 3.

The sampling distribution is shown below. The major tick marks have been labeled with values of $\hat{p}$ and the corresponding Z-scores.



Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

We compute the Z-score for the sample statistic,

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.5 - 0.60}{\sqrt{\frac{0.60)(1-0.60)}{102}}} \approx -2.06155$$

The sample statistic is 2.06155 standard errors below the assumed population proportion. We perform a two-tailed test because the alternative hypothesis, $H_a : p \neq 0.60$, contains a not equal to inequality. We can now find the P-value, which is the probability of seeing a sample proportion as extreme or more extreme than 0.5, by finding the probability from the standard normal distribution. The P-value is approximately 0.03925.

Step 4.

The level of significance is 5% which is 0.05 as a decimal. 0.03925<0.05 so we reject the null hypothesis and support the alternative hypothesis. The sample data support the claim that the proportion of all STEM majors who eventually graduate is different than 60%.

Chi-Square Statistic

The test statistic, Z, as in the example above, works well for binomial experiments consisting of only two outcomes per trial. When there are three or more outcomes per trial, things get more complicated and the Z statistics are insufficient. For multinomial experiments, we can use a different test statistic. This statistic is denoted χ^2 (χ is the Greek letter chi and is pronounced like in Cobra Kai, and not like tea). The formula for the test statistic is

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}$$

Each outcome corresponds to one term in the sum above. E_i stands for the i^{th} **expected frequency** and O_i stands for the i^{th} **observed frequency**.

Example

To understand the terms in this test statistic, let's apply it to a basic example.

Robin is throwing a party for her daughter's 16th birthday. She has party favors for all the guests to leave with.

Scenario 1. The party is a small gathering and five people attend. Robin only has four party favors.

Scenario 2. The party is huge and 51 people attend. Robin only has 50 party favors.

In which scenario above is Robin's error more noticeable?

We can compute the relative error in each scenario:

Scenario 1. $O = 5, E = 4$, so $\frac{(O-E)^2}{E} = \frac{(5-4)^2}{4} = \frac{1}{4} = 0.25$

Scenario 2. $O = 51, E = 50$, so $\frac{(O-E)^2}{E} = \frac{(51-50)^2}{50} = \frac{1}{50} = 0.02$

When there are fewer people expected to attend, the impact of the error is significantly larger than if the same error is made when many people are expected to attend. The sum of these relative errors make up the chi-square test statistic.

The Normal Distribution vs the Chi-Square Distribution

We return to our original problem: Research on college completion has shown that about 60% of students who begin college eventually graduate. A publication of higher education claims that the proportion for STEM (science, technology, engineering, math) majors is different. Researchers randomly select 102 STEM majors and determine that 51 eventually graduate.

1. What is the observed number of STEM majors who eventually graduate? (These are successes in the experiment).

   $O_1 =$ ________________

2. How many STEM majors are expected to graduate based on the null hypothesis?

   $E_1 =$ ________________

3. Compute the contribution of this outcome to the χ^2 test statistic.

   $\frac{(O_1-E_1)^2}{E_1} =$ ________________

4. What is the observed number of STEM majors who do not eventually graduate? (These are failures in this experiment).

   $O_2 =$ ________________

5. How many STEM majors are not expected to graduate based on the null hypothesis?

   $E_2 =$ ________________

6. Compute the contribution of this outcome to the χ^2 test statistic.

   $\frac{(O_2-E_2)^2}{E_2} =$ ________________

7. Think about the test statistic, where expected frequencies are based on p from the null hypothesis. If the null hypothesis is false, would we expect this statistic to be small or large?

8. The number of STEM majors who did eventually graduate was observed from a random sample, and varies from sample to sample. Once this value is known, is the number of STEM majors who did not eventually graduate random?

9. If **degrees of freedom** represent the number of observed frequencies that vary freely (randomly), how many degrees of freedom are present among the two observed frequencies (of those who did graduate and those that did not)?

10. Add the values from numbers 3 and 6 to find the test statistic (round to three decimal places).

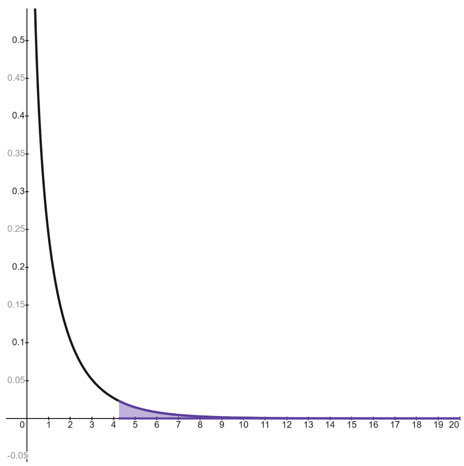
$$\chi^2 = \sum_i \frac{(o_i - E_i)^2}{E_i} = \_\_\_\_\_$$

11. Compute the square of the Z-statistic (round to three decimal places):

$$z^2 = (-2.06155)^2 = \_\_\_\_\_$$

12. How is the χ^2 statistic, with one degree of freedom, related to the Z statistic?

13. Go to this desmos graph, https://www.desmos.com/calculator/bjohldwaym, and enter the degrees of freedom and the test statistic from number 10 to calculate the P-value.



Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

In summary, because the Z test for a population proportion yields the same P-value as the corresponding χ^2 test (with one degree of freedom), the tests are equivalent. We will use this distribution to compare three or more proportions in the next section.