

## 10.2: Goodness-of-Fit

In the last section, we were introduced to the chi-square distribution and its relationship to the normal distribution. We will use the process for  $\chi^2$  tests in multinomial experiments (with more than two outcomes per observation). These tests are often referred to as **Goodness-of-Fit tests**.

### Step 1: Determine the Hypotheses

The goodness of fit test makes claims about the proportions or probabilities for each outcome of a multinomial experiment. If there are  $k$  outcomes per trial, then the null hypothesis would be

$$H_0 : p_1 = \text{value}_1, p_2 = \text{value}_2, \dots, p_k = \text{value}_k$$

The sum of the hypothetical proportions in the null hypothesis must add to 1.

For the null hypothesis to not be true, one or more of the proportions would be incorrect. Thus, the alternative is stated as a sentence, and takes the form

$$H_a : \text{At least one of these proportions is incorrect.}$$

### Step 2: Collect the Data

Using a sample of  $n$  independent trials, each having  $k$  outcomes, the multinomial experiment is conducted by collecting categorical/qualitative data from a random sample. For each outcome, the observed frequency,  $O_i$ , is the number of observed successes. The sum of these observed frequencies is always the sample size,

$$n = \sum_i O_i$$

so the last observed frequency does not vary freely. Therefore, there are  $k - 1$  degrees of freedom among the observed frequencies.

To compute the expected frequencies, we take the sample size and multiply by each proportion assumed in the null hypothesis,  $E_i = np_i$ .

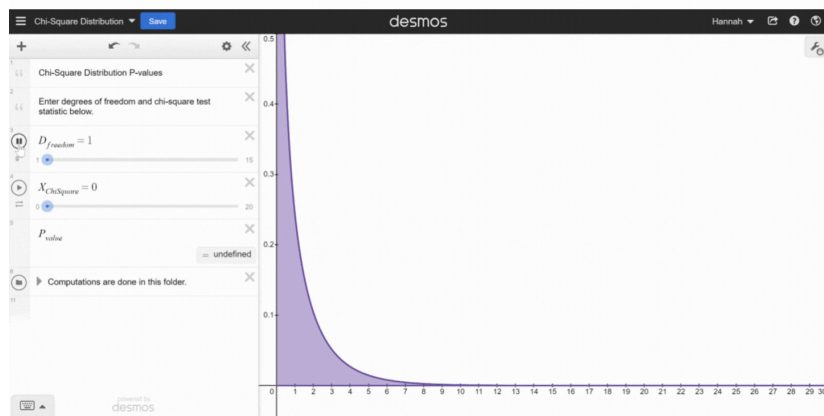
For approximate normality in a binomial experiment, we require at least 10 expected successes and failures in the sample. In a multinomial experiment, the criteria is softened because it can require unreasonably large sample sizes. We say the test statistic is approximately distributed according to the chi-square distribution if each expected frequency,  $E_i$ , is at least 5.

### Step 3: Assess the Evidence

For goodness of fit tests, the test statistic is

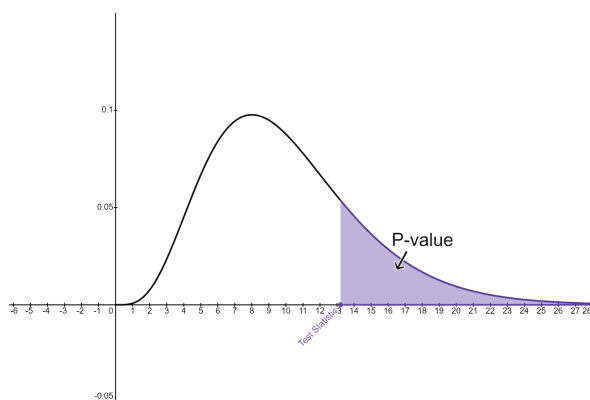
$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}$$

The shape of the distribution depends on the degrees of freedom. In every case, the distribution starts at 0 and is skewed to the right. The mode (peak) occurs at two less than the number of degrees of freedom.



A goodness of fit test is always right tailed. This is because the test statistic involves squaring an error which will always result in a positive number. If the null hypothesis is false, we expect the test statistic to be large.

We use [this desmos graph](https://www.desmos.com/calculator/viuelise2r) (<https://www.desmos.com/calculator/viuelise2r>) to compute the P-value. The P-value is the area of the right tail (starting from the test statistic) under the chi-square distribution with  $k - 1$  degrees of freedom.



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

#### Step 4: Make a Decision and State a Conclusion

Compare the P-value to the level of significance ( $\alpha$ ). If the P-value is less than or equal to  $\alpha$ , we reject the null hypothesis in support of the alternative hypothesis. If the P-value is greater than  $\alpha$ , we fail to reject the null hypothesis, and we cannot support the alternative hypothesis.

## You try!

Below is the distribution of households total income in 2020 according to the United States Census Bureau:

Income per household	Proportion
Under \$35,000	26.2%
Between \$35,000 and \$100,000	40.3%
Over \$100,000	33.5%

You want to know if these proportions are different for black households.

### Step 1

1. How many possible outcomes ( $k$ ) for each black household are there?  $k =$  \_\_\_\_\_

2. State the null hypothesis:

$$H_0 : p_1 = \text{_____, } p_2 = \text{_____, } p_3 = \text{_____}$$

### Step 2

You randomly survey black households for their household income. The observed frequencies are below.

4. Calculate the sample size and expected frequencies for each category in the table below.

Income per black household	Observed frequency	Expected frequency
Under \$35,000	$O_1 = 40$	$E_1 = \text{_____}$
Between \$35,000 and \$100,000	$O_2 = 41$	$E_2 = \text{_____}$
Over \$100,000	$O_3 = 19$	$E_3 = \text{_____}$
Total	$n = \sum O_i = \text{_____}$	

5. Do the expected frequencies satisfy the conditions of an approximate chi-square distribution? Explain.

6. What are the degrees of freedom for this test?  $df = k - 1 = \text{_____}$

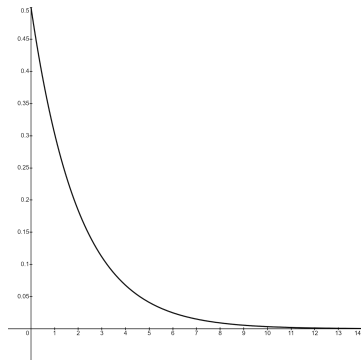
### Step 3

7. Fill in the blanks with the appropriate values and compute the  $\chi^2$  test statistic rounded to four decimal places:

$$\chi^2 = \sum_{i=1}^{k=3} \frac{(O_i - E_i)^2}{E_i} = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2} + \frac{(O_3 - E_3)^2}{E_3} = \frac{(\underline{\quad} - \underline{\quad})^2}{\underline{\quad}} + \frac{(\underline{\quad} - \underline{\quad})^2}{\underline{\quad}} + \frac{(\underline{\quad} - \underline{\quad})^2}{\underline{\quad}}$$

$\approx \underline{\hspace{2cm}}$

8. Use this desmos graph, <https://www.desmos.com/calculator/bjohldwaym>, to compute the P-value rounded to four decimal places. Plot the  $\chi^2$  statistic and shade the area to the right.



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

P-value =                     

### Step 4

9. Using a 1% level of significance, make a decision about the null and alternative hypotheses.

10. State the conclusion in context.

This page titled [10.2: Goodness-of-Fit](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Hannah Seidler-Wright](#).

- [Current page](#) by Hannah Seidler-Wright is licensed [CC BY-NC-SA 4.0](#).
- [1.2: The Statistical Analysis Process](#) by Hannah Seidler-Wright is licensed [CC BY-NC-SA 4.0](#).