Goodness of Fit Test

A goodness of fit test is a hypothesis test used to decide whether a population fits a hypothesized probability distribution.

There is a probability distribution that we think the population fits.

We use observed sample data to draw a conclusion about whether the entire population fits that probability distribution, or does not fit it.

Null Hypothesis: Ho: The population fits the hypothesized distribution
(explain in context of the problem, refer to the distribution)
Alternate Hypothesis: Ha: The population does not fit the hypothesized distribution
(explain in context of the problem, refer to the distribution)

We compare the observed sample data to the expected data that reflects what the sample should look like if it was a perfect fit to the expected distribution.

We find a test statistic that is a measure of whether the observed sample data fits (is close to) the expected data or whether the observed sample data does not fit (is far from) the expected data.

If the test statistic is small, the observed sample data fits (is close to) the expected data.
If the test statistic is big, the observed sample data does not fit (is far from) the expected data.

We deicide whether the test statistic is big or small by seeing how far out it is in the right tail of the chi-square probability distribution.

To see if it is far out in the right tail we find the p-value, the probability (area) to the right of the test statistic. Probability distribution is chi-square with degrees of freedom = number of categories minus 1

Sample fits Expected	⇔	Test Statistic is Small	⇔	Not far out in right tail	⇔	p-value is large	⇔	Do NOT Reject Ho

Sample does NOT fit Expected	⇔	Test Statistic is Big	⇔	Far out in right tail	⇔	p-value is small	⇔	Reject Ho

Use the usual decision rule by comparing p-value to significance level.

Write a conclusion in words in the context of the problem explaining whether the population fits or does not fit the hypothesized distribution, explaining what the population represents and what the distribution is.

Goodness of Fit Test

At Cathy's Campus Coffee Cart, Cathy has a theory about what her customers order. She thinks that 30% of customers order large caffeinated coffee, 20% order small caffeinated coffee, 10% order decaf coffee, 10% order tea, 20% order iced drinks, and 10% order juice.
Cathy wants to know whether this is really the true distribution of orders for all her customers.
Cathy decides to test her theory by selecting a sample of 200 customers and finds that the orders are:
70 large caffeinated coffee; 33 small caffeinated coffee; 35 decaf coffee; 10 tea; 37 iced drinks; 15 juice

A. Finding the Expected Frequencies: For 200 customers, how many people would she have expected to order each type of drink, based on her theory at the beginning of the problem?
Multiply 200 by the percentages given in the top paragraph

B. Organizing the data: Make a table, showing three columns: Type of beverage, observed frequency and expected theoretical frequency. (Remember that frequencies are counts, not percents, fractions, or decimal proportions.).

C. Calculating the test statistic by hand
In a fourth column and find Observed – Expected $(Obs - E)$
In a fifth column, square the $(Obs - E)$ value in the third column and divide by the value in the Expected column
Add up the fifth column $(Obs - E)^2/E$ to get $\sum \frac{(Obs. - E)^2}{E}$

D. Calculating the test statistic by calculator: You can do step C in your calculator :
In STAT EDIT:
Enter the Observed data frequencies in list L1
Enter the Expected Theoretical frequencies in list L2
Arrow up to the very top (title line) in list L3 and input the instruction **(L1 –L2)^2/L2**
In STAT CALC:
Do one variable statistics for list L3. The item Σx is the total of all items in list L3

E. Understanding the test statistic: The sum of the third column is your "test statistic".

$$\sum \frac{(Obs. - E)^2}{E}$$

Since the numerator contains differences of "observed minus expected data", it is a measure of how close or how far the observed data is to the expected.

a. If the observed data and expected data are close together, so that the observed data is a good fit to the expected, will the test statistic be small or large? Explain why.

b. If the observed data and expected data are far apart, so that the observed data is not a good fit to the expected, will the test statistic be small or large? Explain why.

F. Deciding if the test statistic is small or large.

If the test statistic is large, when we look at the appropriate probability distribution it will be far out in the right tail. Then the area to the right will be small. The area to the right of the test statistic is the p-value. A large test statistic goes along with a small p-value.

If the test statistic is small, when we look at the appropriate probability distribution it will not be far out in the right tail. The area to the right will not be small. The area to the right of the test statistic is the p-value. A small test statistic goes along with a large p-value.

Writing the Hypothesis Test for a Goodness Of Fit Test

1. Write the hypotheses
   Null Hypothesis: The data fit the expected (*hypothesized*) distribution
   (Describe the expected hypothesized distribution either in sentence form, by using its appropriate name, by showing a formula or referring to a nearby visible table or list)
   Alternate Hypothesis: The data do not fit the expected (*hypothesized*) distribution
2. Decide upon α
3. Collect data and get the observed and expected data ready to use (in questions A and B above)
4. Find the test statistic and p-value:
   Test statistic χ^2 = ________ (from question C above).
   Find the p-value and draw the graph and write the interpretation of the p-value:
   To determine what is the appropriate cutoff for a large or small value for the test statistic, we use the chi-square distribution to calculate a p-value.
   The degrees of freedom is n – 1 where n is the number of cells (categories) in the distribution.
   Find the p-value: Prob($\chi^2 \geq$ value of test statistic) = χ^2cdf (test statistic, 10^99, df)
5. Compare p-value to the significance level and make a decision.
6. Write a conclusion in the context of the problem.

Hypothesis test for a goodness of fit test for Cathy's Campus Coffee Cart

Ho: The true distribution of the purchases of all Cathy's customers fits the expected distribution of:
30% large caffeinated coffee; 20% small caffeinated coffee; 10% decaf coffee; 10% tea; 20% iced drinks; 10% juice
H_A: The true distribution of the purchases of all Cathy's customers does not fit the expected distribution

α = 0.5

	L1 (Observed data)	**L2 (Expected Frequencies)**	**L3 =(L1 –L2)^2/L2**
large caffeinated coffee	**70**	**60** *(30% of 200)*	**1.6667**
small caffeinated coffee	**33**	**40** *(20% of 200)*	**1.225**
decaf coffee	**35**	**20** *(10% of 200)*	**11.25**
tea	**10**	**20** *(10% of 200)*	**5**
iced drinks	**37**	**40** *(20% of 200)*	**.225**
juice	**15**	**20** *(10% of 200)*	**1.25**

STAT CALC 1-Var Stats L3: *we need the SUM* **Σx = 20.6167** ← *this is the Test Statistic*
degrees of freedom is df = number of cells – 1 = 6 cells – 1 = 5
Find the p-value: χ^2cdf (20.6167, 10^99, 5) = .00096 = pvalue

Graph drawn in class; see your class notes or similar examples in textbook in Ch. 11 for how to draw the graph
Draw your own here, shading to the right of the test statistic – this is always a right tailed test!

Decision: p-value = .00096 is smaller than α = .05 REJECT THE NULL HYPOTHESIS

Conclusion: The true distribution of the purchases of all Cathy's customers does NOT fit the expected distribution of 30% large caffeinated coffee; 20% small caffeinated coffee; 10% decaf coffee; 10% tea; 20% iced drinks; 10% juice
.

TRY IT: Suppose instead that Cathy's sample of 200 customers orders looked like this:
63 large caffeinated coffee, 37 small caffeinated coffee, 21 decaf coffee, 20 tea, 38 iced drinks, 21 juice.

Redo the test and draw a conclusion. (If you do it correctly you should find χ^2 = .575 and p-value = 0 **.**989)