Fact Sheet – Goodness-of-Fit (12.1)

A goodness-of-fit test is an inferential procedure used to determine whether a frequency distribution follows a specific (probability) distribution.

Example: A 6-sided die is fair, i.e. $p_1 = p_2 = \ldots = p_6 = \frac{1}{6}$.

Example: The distribution of plain M&M candies in a bag is 13% brown, 14% yellow, 13% red, 20% orange, 24% blue, and 16% green.

Conditions

To perform a Goodness-of-Fit test for *k* categories, the following two conditions must be met.

- Each expected frequency must be at least 1: $E_i \geq 1$ for all *i* = 1 to *k*.
- No more than 20% of all expected frequencies are less than 5.

You can check the conditions by computing the expected frequencies once you have set up the null hypothesis H_0. The expected frequency for each category can be computed by multiplying the sample size (*n*) by the claimed proportion (p_i) for that category: $E_i = n \cdot p_i$

Hypothesis Test

Step 1

The null hypothesis will be of the form $p_1 = \#,\ p_2 = \#,\ \ldots, p_k = \#$.

H_1 will always be "At least one proportion is different than claimed."

Step 3

The test statistic is $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$, where O_i is the observed sample frequency for category *i* and E_i is the computed expected frequency for category *i*.

Just write "Goodness-of-Fit", rather than writing the test statistic.

Step 4

To compute the test statistic and P-value using StatCrunch …

1. Enter the observed counts in the first column.
Enter the expected counts in the second column.
Name the columns observed and expected.
2. Select **Stat**, highlight **Goodness-of-fit**, then highlight **Chi-Square test**.
3. Select the column that contains the observed counts and select the column that contains the expected counts. Click Calculate.

Examples

1) Forty-five percent of Americans have type O blood, 40% have type A, 11% have type B, and 4% have type AB. A biology class tests the blood of 250 students. Of these students, 124 have type O, 108 have type A, 14 have type B and 4 have type AB. At the 0.05 level of significance, test the claim that the claimed blood type percentages are correct.

(Step 1)

H_0: H_1:

$E_O =$	$E_A =$	$E_B =$	$E_{AB} =$

(Step 2) $\alpha =$_____

(Step 3) Goodness-of-Fit

(Step 4) $\chi^2 =$____________ *P*-value = ______________

(Step 5) ____________________ H_0.

There _______________ sufficient evidence to support that the proportion for at least one blood type is different than claimed.

2) The maker of a non-dairy "ice cream" holds a taste test. Subjects taste the non-dairy "ice cream", as well as three popular brands of ice cream. Of 120 ice cream eaters, 21 selected the non-dairy "ice cream" as their favorite, while 42 selected ice cream A, 28 selected ice cream B and 29 selected ice cream C. At the 0.01 level of significance, test the claim that ice cream eaters equally prefer the four different products.

3) An algebra instructor tells his students on the first day of class that 40% will pass, 30% will fail, and 30% will withdraw that semester if historical patterns hold true. The class began with 38 students. Seventeen of the students passed the class, 6 failed, and the rest dropped. At the 0.05 level of significance, test the claim that the instructor made on the first day of class.

4) A random sample of 80 moviegoers was asked which of the following movie genres was their favorite: Action/Adventure, Comedy, Horror, Romance.
Here are the results:

Action/Adventure: 17
Comedy: 31
Horror: 20
Romance: 12

At the 0.01 level of significance, test the claim that the four genres are equally likely to be listed as a moviegoer's favorite.