Chapter 10: Chi-Square Tests: Solutions

10.1 Goodness of Fit Test

In this section, we consider experiments with *multiple outcomes*. The probability of each outcome is fixed.

Definition: A **chi-square goodness-of-fit test** is used to test whether a frequency distribution obtained experimentally fits an "expected" frequency distribution that is based on the theoretical or previously known probability of each outcome.

An experiment is conducted in which a *simple random sample* is taken from a population, and each member of the population is grouped into exactly one of k categories.

Step 1: The **observed frequencies** are calculated for the sample.

Step 2: The **expected frequencies** are obtained from previous knowledge (or belief) or probability theory. In order to proceed to the next step, it is necessary that each expected frequency is at least 5.

Step 3: A hypothesis test is performed:

(i) The *null hypothesis* H_0: the population frequencies are equal to the expected frequencies.

(ii) The *alternative hypothesis*, H_a: the null hypothesis is false (what does this imply about the population frequencies?).

(iii) α is the *level of significance.*

(iv) The *degrees of freedom*: $k - 1$.

(v) A *test statistic* is calculated:

$$\chi^2 = \sum \frac{(\text{ observed } - \text{ expected })^2}{\text{expected}} = \sum \frac{(O-E)^2}{E}$$

(vi) From α and $k - 1$, a *critical value* is determined from the chi-square table.

(vii) Reject H_0 if χ^2 is larger than the *critical value* (right-tailed test).

Example 1: Researchers have conducted a survey of 1600 coffee drinkers asking how much coffee they drink in order to confirm previous studies. Previous studies have indicated that 72% of Americans drink coffee. The results of previous studies (left) and the survey (right) are below. At $\alpha = 0.05$, is there enough evidence to conclude that the distributions are the same?

Response	**% of Coffee Drinkers**
2 cups per week	15%
1 cup per week	13%
1 cup per day	27%
2+ cups per day	45%

Response	**Frequency**
2 cups per week	206
1 cup per week	193
1 cup per day	462
2+ cups per day	739

(i) The *null hypothesis* H_0:the population frequencies are equal to the expected frequencies (to be calculated below).

(ii) The *alternative hypothesis*, H_a: the null hypothesis is false.

(iii) $\alpha = 0.05$.

(iv) The *degrees of freedom*: $k - 1 = 4 - 1 = 3$.

(v) The *test statistic* can be calculated using a table:

Response	**% of Coffee Drinkers**	E	O	$O - E$	$(O-E)^2$	$\frac{(O-E)^2}{E}$
2 cups per week	15%	$0.15 \times 1600 = 240$	206	-34	1156	4.817
1 cup per week	13%	$0.13 \times 1600 = 208$	193	-15	225	1.082
1 cup per day	27%	$0.27 \times 1600 = 432$	462	30	900	2.083
2+ cups per day	45%	$0.45 \times 1600 = 720$	739	19	361	0.5014

$$\chi^2 = \sum \frac{(\text{ observed } - \text{ expected })^2}{\text{expected}} = \sum \frac{(O-E)^2}{E} = \underline{8.483}.$$

(vi) From $\alpha = 0.05$ and $k - 1 = 3$, the *critical value* is 7.815.

(vii) Is there enough evidence to reject H_0? Since $\chi^2 \approx 8.483 > 7.815$, there is enough statistical evidence to reject the null hypothesis and to believe that the old percentages no longer hold.

Example 2: A department store, A, has four competitors: B,C,D, and E. Store A hires a consultant to determine if the percentage of shoppers who prefer each of the five stores is the same. A survey of 1100 randomly selected shoppers is conducted, and the results about which one of the stores shoppers prefer are below. Is there enough evidence using a significance level $\alpha = 0.05$ to conclude that the proportions are really the same?

Store	A	B	C	D	E
Number of Shoppers	262	234	204	190	210

(i) The *null hypothesis* H_0:the population frequencies are equal to the expected frequencies (to be calculated below).

(ii) The *alternative hypothesis*, H_a: the null hypothesis is false.

(iii) $\alpha = 0.05$.

(iv) The *degrees of freedom*: $k - 1 = 5 - 1 = 4$.

(v) The *test statistic* can be calculated using a table:

Preference	**% of Shoppers**	E	O	$O - E$	$(O - E)^2$	$\frac{(O-E)^2}{E}$
A	20%	$0.2 \times 1100 = 220$	262	42	1764	8.018
B	20%	$0.2 \times 1100 = 220$	234	14	196	0.891
C	20%	$0.2 \times 1100 = 220$	204	-16	256	1.163
D	20%	$0.2 \times 1100 = 220$	190	-30	900	4.091
E	20%	$0.2 \times 1100 = 220$	210	-10	100	0.455

$$\chi^2 = \sum \frac{(\text{ observed } - \text{ expected })^2}{\text{expected}} = \sum \frac{(O - E)^2}{E} = \underline{14.618}.$$

(vi) From $\alpha = 0.05$ and $k - 1 = 4$, the *critical value* is 9.488.

(vii) Is there enough evidence to reject H_0? Since $\chi^2 \approx 14.618 > 9.488$, there is enough statistical evidence to reject the null hypothesis and to believe that customers do not prefer each of the five stores equally.

10.2 Independence

Recall that two events are *independent* if the occurrence of one of the events has no effect on the occurrence of the other event.

A **chi-square independence test** is used to test whether or not two variables are independent.

As in section 10.1, an experiment is conducted in which the frequencies for *two* variables are determined. To use the test, the same assumptions must be satisfied: the observed frequencies are obtained through a simple random sample, and each expected frequency is at least 5. The frequencies are written down in a table: the columns contain outcomes for one variable, and the rows contain outcomes for the other variable.

The procedure for the hypothesis test is essentially the same. The differences are that:

(i) H_0 is that the two variables are independent.

(ii) H_a is that the two variables are not independent (they are dependent).

(iii) The expected frequency $E_{r,c}$ for the entry in row r, column c is calculated using:

$$E_{r,c} = \frac{(\text{ Sum of row } r) \times (\text{ Sum of column } c)}{\text{Sample size}}$$

(iv) The degrees of freedom: (number of rows - 1)×(number of columns - 1).

Example 3: The results of a random sample of children with pain from musculoskeletal injuries treated with acetaminophen, ibuprofen, or codeine are shown in the table. At $\alpha = 0.10$, is there enough evidence to conclude that the treatment and result are independent?

	Acetaminophen (c. 1)	Ibuprofen (c. 2)	Codeine (c. 3)	Total
(r. 1) Significant Improvement	58 (66.7)	81 (66.7)	61 (66.7)	200
(r. 2) Slight Improvement	42 (33.3)	19 (33.3)	39 (33.3)	100
Total	100	100	100	300

First, calculate the column and row totals.

Then find the expected frequency for each item and write it in the parenthesis next to the observed frequency.

Now perform the hypothesis test.

(i) The *null hypothesis* H_0: the treatment and response are independent.

(ii) The *alternative hypothesis*, H_a: the treatment and response are dependent.

(iii) $\alpha = 0.10$.

(iv) The *degrees of freedom*:
(number of rows - 1)×(number of columns - 1) $= (2-1)\times(3-1) = 1\times 2 = 2$.

(v) The *test statistic* can be calculated using a table:

Row, Column	E	O	$O-E$	$(O-E)^2$	$\frac{(O-E)^2}{E}$
1,1	$\frac{200\cdot 100}{300} = 66.7$	58	-8.7	75.69	1.135
1,2	$\frac{200\cdot 100}{300} = 66.7$	81	14.3	204.49	3.067
1,3	$\frac{200\cdot 100}{300} = 66.7$	61	-5.7	32.49	0.487
2,1	$\frac{100\cdot 100}{300} = 33.3$	42	8.7	75.69	2.271
2,2	$\frac{100\cdot 100}{300} = 33.3$	19	-14.3	204.49	6.135
2,3	$\frac{100\cdot 100}{300} = 33.3$	39	5.7	32.49	0.975

$$\chi^2 = \sum \frac{(\text{ observed } - \text{ expected })^2}{\text{expected}} = \sum \frac{(O-E)^2}{E} = \underline{14.07}.$$

(vi) From $\alpha = 0.10$ and d.f $= 2$, the *critical value* is $\underline{4.605}$.

(vii) Is there enough evidence to reject H_0? Since $\chi^2 \approx 14.07 > 4.605$, there is enough statistical evidence to reject the null hypothesis and to believe that there is a relationship between the treatment and response.

Practice Problem 1: A doctor believes that the proportions of births in this country on each day of the week are equal. A simple random sample of 700 births from a recent year is selected, and the results are below. At a significance level of 0.01, is there enough evidence to support the doctor's claim?

Day	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Frequency	65	103	114	116	115	112	75

(i) The *null hypothesis* H_0:the population frequencies are equal to the expected frequencies (to be calculated below).

(ii) The *alternative hypothesis*, H_a: the null hypothesis is false.

(iii) $\alpha = 0.01$.

(iv) The *degrees of freedom*: $k - 1 = 7 - 1 = 6$.

(v) The *test statistic* can be calculated using a table:

Day	E	O	$O-E$	$(O-E)^2$	$\frac{(O-E)^2}{E}$
Sunday	$700/7 = 100$	65	-35	1225	12.25
Monday	$700/7 = 100$	103	3	9	0.09
Tuesday	$700/7 = 100$	114	14	196	1.96
Wednesday	$700/7 = 100$	116	16	256	2.56
Thursday	$700/7 = 100$	115	15	225	2.25
Friday	$700/7 = 100$	112	12	144	1.44
Saturday	$700/7 = 100$	75	-25	625	6.25

$$\chi^2 = \sum \frac{(\text{ observed } - \text{ expected })^2}{\text{expected}} = \sum \frac{(O-E)^2}{E} = 26.8.$$

(vi) From $\alpha = 0.01$ and $k - 1 = 6$, the *critical value* is 16.812.

(vii) Is there enough evidence to reject H_0? Since $\chi^2 \approx 26.8 > 16.812$, there is enough statistical evidence to reject the null hypothesis and to believe that the proportion of births is not the same for each day of the week.

Practice Problem 2: The side effects of a new drug are being tested against a placebo. A simple random sample of 565 patients yields the results below. At a significance level of $\alpha = 0.05$, is there enough evidence to conclude that the treatment is independent of the side effect of nausea?

Result	Drug (c.1)	Placebo (c.2)	Total
Nausea (r.1)	36	13	49
No nausea (r.2)	254	262	516
Total	290	275	565

(i) The *null hypothesis* H_0: the treatment and response are independent.

(ii) The *alternative hypothesis*, H_a: the treatment and response are dependent.

(iii) $\alpha = 0.01$.

(iv) The *degrees of freedom*:
(number of rows - 1)$\times$(number of columns - 1) $= (2-1) \times (2-1) = 1 \times 1 = 1$.

(v) The *test statistic* can be calculated using a table:

Row, Column	E	O	$O-E$	$(O-E)^2$	$\frac{(O-E)^2}{E}$
1,1	$\frac{49 \cdot 290}{565} = 25.15$	36	10.85	117.72	4.681
1,2	$\frac{49 \cdot 275}{565} = 23.85$	13	-10.85	117.72	4.936
2,1	$\frac{516 \cdot 290}{565} = 264.85$	254	-10.85	117.72	0.444
2,2	$\frac{516 \cdot 275}{565} = 251.15$	262	10.85	117.72	0.469

$$\chi^2 = \sum \frac{(\text{ observed } - \text{ expected })^2}{\text{expected}} = \sum \frac{(O-E)^2}{E} = \underline{10.53}.$$

(vi) From $\alpha = 0.10$ and d.f $= 1$, the *critical value* is $\underline{2.706}$.

(vii) Is there enough evidence to reject H_0? Since $\chi^2 \approx 10.53 > 2.706$, there is enough statistical evidence to reject the null hypothesis and to believe that there is a relationship between the treatment and response.

Practice Problem 3: Suppose that we have a 6-sided die. We assume that the die is unbiased (upon rolling the die, each outcome is equally likely). An experiment is conducted in which the die is rolled 240 times. The outcomes are in the table below. At a significance level of $\alpha = 0.05$, is there enough evidence to support the hypothesis that the die is unbiased?

Outcome	1	2	3	4	5	6
Frequency	34	44	30	46	51	35

(i) The *null hypothesis* H_0: each face is equally likely to be the outcome of a single roll.

(ii) The *alternative hypothesis*, H_a: the null hypothesis is false.

(iii) $\alpha = 0.05$.

(iv) The *degrees of freedom*: $k - 1 = 6 - 1 = 5$.

(v) The *test statistic* can be calculated using a table:

Face	E	O	$O-E$	$(O-E)^2$	$\frac{(O-E)^2}{E}$
1	$240/6 = 40$	34	-6	36	0.9
2	$240/6 = 40$	44	4	16	0.4
3	$240/6 = 40$	30	-10	100	2.5
4	$240/6 = 40$	46	6	36	0.9
5	$240/6 = 40$	51	11	121	3.025
6	$240/6 = 40$	35	-5	25	0.625

$$\chi^2 = \sum \frac{(\text{ observed } - \text{ expected })^2}{\text{expected}} = \sum \frac{(O-E)^2}{E} = \underline{8.35}.$$

(vi) From $\alpha = 0.01$ and $k - 1 = 6$, the *critical value* is $\underline{15.086}$.

(vii) Is there enough evidence to reject H_0? Since $\chi^2 \approx 8.35 < 15.086$, we fail to reject the null hypothesis, that the die is fair.