MATHEMATICS 360-255-LW
Quantitative Methods II
Martin Huard
Fall 2009

XIX – Chi-Square Goodness of Fit
SOLUTIONS

1. The following table lists the frequency distribution for a sample of 50 absences of college students from classes according to the day of occurrence.

Day of the week	Mon	Tue	Wed	Thu	Fri	*Total*
Number of absences	14	6	4	10	16	50
E	10	10	10	10	10	

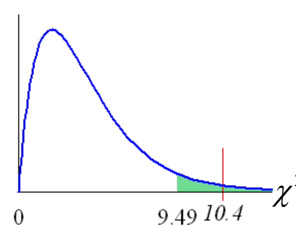
Test at the 5% significance level if the number of absences is equally distributed over all days of the week. Try with both approaches, the classical and the p-value.

Classical approach

Step 1 Assumptions: The classes are all inclusive and mutually exclusive

Step 2 H_0: The number of absences are equally distributed over all days of the week.
H_A: The number of absences are not equally distributed over all days of the week.

Step 3 a) Test statistic: χ^2 with $df = 4$
b) Right-tailed test with $\alpha = 0.05$
c) $\chi^2_{(df,\alpha)} = \chi^2_{(4,0.05)} = 9.49$



Step 4
$$\chi^2 = \sum \frac{(O-E)^2}{E}$$
$$= \frac{(14-10)^2}{10} + \frac{(6-10)^2}{10} + \ldots + \frac{(16-10)^2}{10}$$
$$= 10.40$$

Step 5 a) χ^2 is in the critical region
b) Reject H_o.
∴ There is not sufficient evidence at the 5% level of significance to conclude that the number of absences is not equally distributed over all days of the week.

***p*-value approach**

Step 1 Assumptions: The classes are all inclusive and mutually exclusive

Step 2 H_0: The number of absences are equally distributed over all days of the week.
H_A: The number of absences are not equally distributed over all days of the week.

Step 3 a) Test statistic: χ^2 with $df = 4$
b) Right-tailed test with α = 0.05

Step 4
a) $\chi^2 = \sum \frac{(O-E)^2}{E}$

$$= \frac{(14-10)^2}{10} + \frac{(6-10)^2}{10} + \ldots + \frac{(16-10)^2}{10}$$

$= 10.40$

b) $0.025 < p\text{-value} < 0.05$

Step 5 a) p-value < α = 0.05
b) Reject H_o.
∴ There is not sufficient evidence at the 5% level of significance to conclude that the number of absences is not equally distributed over all days of the week.

2. To check a die for fairness, it is rolled 300 times. If the die is fair, then all of the outcomes should be equally likely. Here are the results of the rolls.

Number on die	1	2	3	4	5	6
Number of rolls	57	43	44	53	46	57
E	50	50	50	50	50	50

At the 5% level of significance, test the claim that the die is fair. Try with both approaches, the classical and the p-value.

Classical approach

Step 1 Assumptions: The classes are all inclusive and mutually exclusive

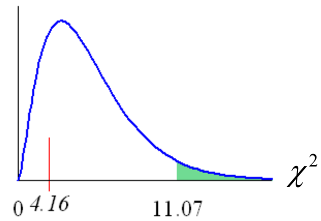
Step 2 H_0: Each outcome of the die is equally likely (the die is fair).

H_A: Not all outcomes of the die are equally likely.

Step 3 c) Test statistic: χ^2 with $df = 5$

d) Right-tailed test with $\alpha = 0.05$

d) $\chi^2_{(df,\alpha)} = \chi^2_{(5,0.05)} = 11.07$



Step 4

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

$$= \frac{(57-50)^2}{50} + \frac{(43-50)^2}{50} + \ldots + \frac{(57-50)^2}{50}$$

$$= 4.16$$

Step 5 a) χ^2 is not in the critical region

b) Fail to reject H_o.

∴ There is not sufficient evidence at the 5% level of significance to conclude that the die is not fair.

***p*-value approach**

Step 1 Assumptions: The classes are all inclusive and mutually exclusive

Step 2 H_0: Each outcome of the die is equally likely (the die is fair).

H_A: Not all outcomes of the die are equally likely.

Step 3 a) Test statistic: χ^2 with $df = 5$

b) Right-tailed test with $\alpha = 0.05$

Step 4

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

a) $$= \frac{(57-50)^2}{50} + \frac{(43-50)^2}{50} + \ldots + \frac{(57-50)^2}{50}$$

$$= 4.16$$

b) $0.500 < p\text{-value} < 0.750$

Step 5 a) $p\text{-value} > 0.500 > \alpha = 0.05$

b) Fail to Reject H_o.

∴ There is not sufficient evidence at the 5% level of significance to conclude that the die is not fair.

3. A study of high school professors asked them how prepared they were to integrate technology into their instructions. Thirteen percent were very well prepared, 26% were well prepared, 51% were somewhat prepared, and 10% were not prepared at all. A random sample of 60 small town high school teachers showed the following results.

Level of Preparedness	Number of Respondents	E
Very well	5	7.8
Well	11	15.6
Somewhat	33	30.6
Not at all	11	6

At the 10% level of significance, test the claim that teachers in this small town follow the distribution reported in the study. Try with both approaches, the classical and the p-value.

Classical approach

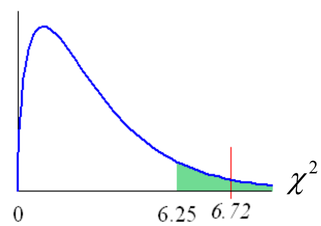
Step 1 Assumptions: The classes are all inclusive and mutually exclusive

Step 2 H_0: Teachers in the small town follow the distribution reported in the study.
H_A: Teachers in the small town do not follow the distribution reported in the study.

Step 3 a) Test statistic: χ^2 with $df = 3$

b) Right-tailed test with $\alpha = 0.10$

c) $\chi^2_{(df,\alpha)} = \chi^2_{(3,0.10)} = 6.25$



Step 4

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

$$= \frac{(5-7.8)^2}{7.8} + \frac{(11-15.6)^2}{15.6} + \frac{(33-30.6)^2}{30.6} + \frac{(11-6)^2}{6}$$

$$= 6.72$$

Step 5 a) χ^2 is in the critical region

b) Reject H_o.

∴ There is sufficient evidence at the 10% level of significance to conclude that teachers in the small town do not follow the distribution reported in the study.

***p*-value approach**

Step 1 Assumptions: The classes are all inclusive and mutually exclusive

Step 2 H_0: Teachers in the small town follow the distribution reported in the study.
H_A: Teachers in the small town do not follow the distribution reported in the study.

Step 3 a) Test statistic: χ^2 with $df = 3$
b) Right-tailed test with α = 0.10

Step 4
a) $\chi^2 = \sum \frac{(O-E)^2}{E}$

$$= \frac{(5-7.8)^2}{7.8} + \frac{(11-15.6)^2}{15.6} + \frac{(33-30.6)^2}{30.6} + \frac{(11-6)^2}{6}$$

$= 6.72$

b) $0.050 < p\text{-value} < 0.100$

Step 5 a) p-value < α = 0.10
b) Reject H_o.
∴ There is sufficient evidence at the 10% level of significance to conclude that teachers in the small town do not follow the distribution reported in the study.

4. A pair of dice are rolled 200 times. Here are the sums produced by the 200 rolls.

Sum	2	3	4	5	6
Number of Rolls	4	16	17	16	27
E	$\frac{200}{36} = 5.56$	$\frac{2\cdot200}{36} = 11.11$	$\frac{3\cdot200}{36} = 16.67$	$\frac{4\cdot200}{36} = 22.22$	$\frac{5\cdot200}{36} = 27.78$

7	8	9	10	11	12
43	25	18	15	16	3
$\frac{6\cdot200}{36} = 33.33$	$\frac{5\cdot200}{36} = 27.78$	$\frac{4\cdot200}{36} = 22.22$	$\frac{3\cdot200}{36} = 16.67$	$\frac{2\cdot200}{36} = 11.11$	$\frac{200}{36} = 5.56$

At the 1% significance level, can we conclude that the dies are fair? That is, the sums produced by this pair of dice follow the expected distribution? Try with both approaches, the classical and the p-value.

Classical approach

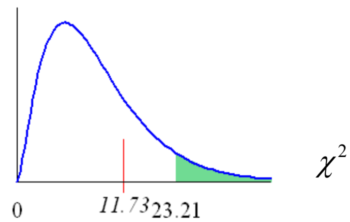
Step 1 Assumptions: The classes are all inclusive and mutually exclusive

Step 2 H_0: Sums produced by the pair of dice follow the expected distribution.
H_A: Sums produced by the pair of dice do not follow the expected distribution.

Step 3 a) Test statistic: χ^2 with $df = 10$

b) Right-tailed test with $\alpha = 0.01$

c) $\chi^2_{(df,\alpha)} = \chi^2_{(10,0.01)} = 23.21$



Step 4

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

$$= \frac{(4-5.56)^2}{5.56} + \frac{(16-11.11)^2}{11.11} + \ldots + \frac{(3-5.56)^2}{5.56}$$

$$= 11.73$$

Step 5 a) χ^2 is not in the critical region

b) Fail to reject H_o.

∴ There is not sufficient evidence at the 1% level of significance to conclude that the dies are not fair.

***p*-value approach**

Step 1 Assumptions: The classes are all inclusive and mutually exclusive

Step 2 H_0: Sums produced by the pair of dice follow the expected distribution.
H_A: Sums produced by the pair of dice do not follow the expected distribution.

Step 3 a) Test statistic: χ^2 with $df = 10$

b) Right-tailed test with $\alpha = 0.01$

Step 4

a) $\chi^2 = \sum \frac{(O-E)^2}{E}$

$$= \frac{(4-5.56)^2}{5.56} + \frac{(16-11.11)^2}{11.11} + \ldots + \frac{(3-5.56)^2}{5.56}$$

$$= 11.73$$

b) $0.250 < p\text{-value} < 0.500$

Step 5 a) $p\text{-value} > 0.250 > \alpha = 0.01$

b) Fail to reject H_o.

∴ There is not sufficient evidence at the 1% level of significance to conclude that the dies are not fair.

5. A philosophy 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 5% level of significance, test the claim that the instructor made on the first day of class. Try with both approaches, the classical and the p-value.

	Pass	*Fail*	*Withdraw*
E	$0.4 \cdot 38 = 15.2$	$0.3 \cdot 38 = 11.4$	$0.3 \cdot 38 = 11.4$
O	17	6	15

Classical approach

Step 1 Assumptions: The classes are all inclusive and mutually exclusive

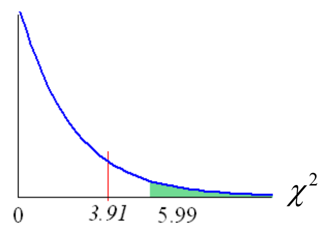
Step 2 H_0: The distribution of grades follows the given distribution.

H_A: The distribution of grades does not follow the given distribution.

Step 3 a) Test statistic: χ^2 with $df = 2$

b) Right-tailed test with $\alpha = 0.05$

c) $\chi^2_{(df,\alpha)} = \chi^2_{(2,0.05)} = 5.99$



Step 4

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

$$= \frac{(17-15.2)^2}{15.2} + \frac{(6-11.4)^2}{11.4} + \frac{(15-11.4)^2}{11.4} = 3.91$$

Step 5 a) χ^2 is not in the critical region

b) Fail to reject H_o.

∴ There is not sufficient evidence at the 5% level of significance to conclude that the distribution of grades follows the given distribution, that is, that the instructors' claim is true.

***p*-value approach**

Step 1 Assumptions: The classes are all inclusive and mutually exclusive

Step 2 H_0: The distribution of grades follows the given distribution.

H_A: The distribution of grades does not follow the given distribution.

Step 3 a) Test statistic: χ^2 with $df = 2$

b) Right-tailed test with $\alpha = 0.05$

Step 4

a) $\chi^2 = \sum \frac{(O-E)^2}{E}$

$$= \frac{(17-15.2)^2}{15.2} + \frac{(6-11.4)^2}{11.4} + \frac{(15-11.4)^2}{11.4} = 3.91$$

b) $0.100 < p\text{-value} < 0.250$

Step 5 a) $p\text{-value} > 0.100 > \alpha = 0.05$

b) Fail to reject H_o.

∴ There is not sufficient evidence at the 5% level of significance to conclude that the distribution of grades follows the given distribution, that is, that the instructors' claim is true.

6. A 1970 study showed that of Canadian married-couple families, 42.9% had no children, 18.3% had one child, 18.0% had two children, and 20.8% had three or more children. A recent survey of 500 Canadian married-couple, families revealed that 267 had no children, 87 had one child, 96 had two children, and 50 had three or more children. At the 1% level of significance, test the claim that the 1970 proportions are no longer valid. Try with both approaches, the classical and the p-value.

	No Child	*1 child*	*2 child*	*3 or more*
E	$0.429 \cdot 500 = 214.5$	$0.183 \cdot 500 = 91.5$	$0.18 \cdot 500 = 90$	$0.208 \cdot 500 = 104$
O	267	87	96	50

Classical approach

Step 1 Assumptions: The classes are all inclusive and mutually exclusive

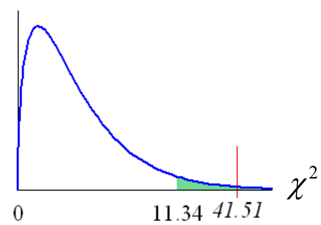
Step 2 H_0: The proportions are the same now as in 1970.

H_A: The proportions are not the same now as in 1970.

Step 3 a) Test statistic: χ^2 with $df = 3$

b) Right-tailed test with $\alpha = 0.01$

c) $\chi^2_{(df,\alpha)} = \chi^2_{(3,0.01)} = 11.34$



Step 4

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

$$= \frac{(267-214.5)^2}{214.5} + \frac{(87-91.5)^2}{91.5} + \frac{(96-90)^2}{90} + \frac{(50-104)^2}{104} = 41.51$$

Step 5 a) χ^2 is in the critical region

b) Reject H_o.

∴ There is sufficient evidence at the 1% level of significance to conclude that the 1970 proportions are not valid anymore.

***p*-value approach**

Step 1 Assumptions: The classes are all inclusive and mutually exclusive

Step 2 H_0: The proportions are the same now as in 1970.

H_A: The proportions are not the same now as in 1970.

Step 3 a) Test statistic: χ^2 with $df = 3$

b) Right-tailed test with $\alpha = 0.01$

Step 4

a) $\chi^2 = \sum \frac{(O-E)^2}{E}$

$$= \frac{(267-214.5)^2}{214.5} + \frac{(87-91.5)^2}{91.5} + \frac{(96-90)^2}{90} + \frac{(50-104)^2}{104} = 41.51$$

b) p-value < 0.005

Step 5 a) p-value $< 0.005 < \alpha = 0.01$

b) Reject H_o.

∴ There is sufficient evidence at the 1% level of significance to conclude that the 1970 proportions are not valid anymore.