

# *Contents*

## *goodness of fit and* **42**

# **contingency tables**

1. Goodness of fit
2. Contingency tables

### *Learning* **outcomes**

*You will learn how to decide whether a set of data fits a particular distribution. You will also meet a situation in which hypothesis tests are applied to non-numeric data.*

### *Time* **allocation**

*You are expected to spend approximately ten hours of independent study on the material presented in this workbook. However, depending upon your ability to concentrate and on your previous experience with certain mathematical topics this time may vary considerably.*

# Goodness of Fit

# 42.1



## Introduction

If you are applying statistics to practical problems in industry, you may find that much of your work is concerned with making decisions concerning probability distributions. Sometimes it is advantageous to be able to describe the approximate probability distribution followed by a data set obtained experimentally. For example you may be asked to decide whether a data set is approximately normal. In order to make such decisions, you will find that you may use the chi-squared test provided that certain conditions are satisfied. On other occasions you may be given data concerning non-numeric variables in the form of a contingency table. This is one of those occasions when hypothesis tests can be applied to non-numeric variables.



## Prerequisites

Before starting this Section you should ...

- ① understand the principles of hypothesis testing (Workbook 41)



## Learning Outcomes

After completing this Section you should be able to ...

- ✓ understand what is meant by the term goodness-of-fit.
- ✓ understand and perform hypothesis tests based on the chi-squared distribution.

# 1. Goodness-of-Fit Tests

The aim of a goodness-of-fit test is to determine the underlying nature of the probability distribution describing the population from which a random sample has been drawn. For example, we may wish to determine whether the population from which a sample has been drawn is normal, binomial or Poisson.

While a variety of goodness-of-fit tests exist, the test described here depends on the  $\chi^2$ -distribution and is usually called the chi-squared test.

We assume that a random sample of size  $n$  has been drawn from a population with an unknown probability distribution and that we wish to determine the nature of that distribution.

Firstly, if the data are continuous we organize the data into  $k$  intervals (often equal but not necessarily so) so that we can write down the observed frequency, say  $O_i$ , of the  $i$ th interval for  $1 \leq i \leq k$ .

Secondly, we form a hypothesis about the nature of the unknown distribution. That is, we assume that it is normal, binomial, Poisson or some other appropriate probability distribution. Thirdly, we calculate, on the basis of the hypothesis outlined above, the expected frequency, say  $E_i$ , of the  $i$ th interval for  $1 \leq i \leq k$ . The values of  $E_i$  are calculated using the formula

$$E_i = nP_i$$

where  $P_i$  is the probability associated with the interval  $i$ .

Next we calculate the goodness-of-fit statistic

$$X^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

It can be shown that, if the assumption made about the nature of the population (normal, binomial, Poisson etc.) is true then  $X^2$  follows (approximately) a chi-squared distribution with  $k - p - 1$  degrees of freedom. Note that  $p$  represents the number of parameters needed to describe the probability distribution of the population. For example the normal distribution has two parameters  $\mu$  and  $\sigma$ , the binomial distribution has two parameters  $n$  and  $p$ , while the Poisson distribution has one parameter,  $\mu$ .

Fourthly, we reject the hypothesis concerning the nature of the underlying probability distribution if the calculated value of  $X^2$  exceeds the value of  $\chi_{\alpha, k-p-1}^2$  where  $\alpha$  is the area in the tail of the  $\chi^2$ -distribution, typically 5% or 1%.

Notes

- (a) The larger the sample, the more reliable the result since the assertion that  $X^2$  follows (approximately) a chi-squared distribution improves with increasing sample size.
- (b) The size of the expected frequencies should be monitored carefully. Various authors recommend that minimum expected frequencies of 3, 4 or 5 are acceptable. It is reasonably safe to accept expected frequencies provided that they are greater than 5, 10 is certainly acceptable.
- (c) Some authors recommend that the  $k$  intervals into which the data are organized are chosen so that the frequencies in each interval are roughly equal - remember that equal intervals are not necessary for the test to be performed.

We will now look at two examples of goodness-of-fit tests, the first example uses a (discrete) Poisson distribution and the second uses a (continuous) normal distribution. Each worked example is immediately followed by an exercise for you to do.

**Example** A manufacturer produces high-quality sheet aluminium for use in highly stressed aircraft wings. A random sample of 100 sheets is inspected and the number of faults per sheet recorded. The results are given in the table below.

Number of Faults per Sheet	Frequency of Occurrence
0	50
1	24
2	14
3	8
4	4

Suggest a possible probability distribution from which the sample may have been drawn and perform a chi-squared test to determine the validity of your suggestion.

### Solution

The data are already given in 5 classes with observed frequencies as shown. We will assume that the underlying distribution is Poisson and calculate the expected frequencies accordingly using the Poisson formula

$$P(X = r) = \frac{e^{-\mu} \mu^r}{r!}$$

We need the value of the mean. This is calculated as

$$\mu = \frac{50 \times 0 + 24 \times 1 + 14 \times 2 + 8 \times 3 + 4 \times 4}{100} = 0.92$$

Hence the Poisson probabilities and the corresponding expected frequencies are:

$$p_0 = P(X = 0) = \frac{e^{-\mu} \mu^0}{0!} = e^{-0.92} = 0.399, \quad E_0 = 39.9$$

$$p_1 = P(X = 1) = \frac{e^{-\mu} \mu^1}{1!} = e^{-0.92} \times 0.92 = 0.367, \quad E_1 = 36.7$$

$$p_2 = P(X = 2) = \frac{e^{-\mu} \mu^2}{2!} = \frac{e^{-0.92} \times 0.92^2}{2} = 0.169, \quad E_2 = 16.9$$

$$p_3 = P(X = 3) = \frac{e^{-\mu} \mu^3}{3!} = \frac{e^{-0.92} \times 0.92^3}{6} = 0.052, \quad E_3 = 5.2$$

$$p_4 = P(X \geq 4) = 1 - (0.399 + 0.367 + 0.169 + 0.052) = 0.015, \quad E_4 = 1.5$$

Note that in calculating  $p_4$  we have ensured that our probabilities sum to unity.

Since the last frequency is very small we will combine the last two and use 4 classes so that  $O_3 = 12$  and  $E_3 = 6.5$ .

The test statistic is

$$X^2 = \sum_{i=0}^3 \frac{(O_i - E_i)^2}{E_i} = \frac{(50 - 39.9)^2}{39.9} + \frac{(24 - 36.7)^2}{36.7} + \frac{(14 - 16.9)^2}{16.9} + \frac{(12 - 6.7)^2}{6.7} = 11.642$$

and the number of degrees of freedom is  $k - p - 1 = 4 - 1 - 1 = 2$  so that the critical value from tables is  $\chi_{0.05,2}^2 = 5.99$ . Clearly 11.462 and we must reject the null hypothesis that the underlying distribution is Poisson.



A manufacturer produces electronic components for use in computer controlled monitoring systems. A random sample of 100 components is inspected and the number of faults per component recorded. The results are given in the table below.

Number of Faults per Component	Frequency of Occurrence
0	45
1	35
2	16
3	4

Perform a chi-squared test to determine the validity of the assumption that the occurrence of faults in the components is Poisson.

**Your solution**

The data are given in 4 classes with observed frequencies as shown. The expected frequencies using the Poisson formula with a mean

$$\mu = \frac{45 \times 0 + 35 \times 1 + 16 \times 2 + 4 \times 3}{100} = 0.79$$

are

$$p_0 = P(X = 0) = \frac{e^{-\mu} \mu^0}{0!} = e^{-0.79} = 0.454, E_0 = 45.4$$

$$p_1 = P(X = 1) = \frac{e^{-\mu} \mu^1}{1!} = e^{-0.79} \times 0.79 = 0.359, E_1 = 35.9$$

$$p_2 = P(X = 2) = \frac{e^{-\mu} \mu^2}{2!} = \frac{e^{-0.79} \times 0.79^2}{2} = 0.142, E_2 = 14.2$$

$$p_3 = P(X \geq 3) = 1 - (0.454 + 0.359 + 0.142) = 0.045, E_3 = 4.5$$

The last frequency is small but since it is greater than 3 we will allow its use.

The test statistic is

$$X^2 = \sum_{i=0}^3 \frac{E_i}{(O_i - E_i)^2} = \frac{45.4}{(45 - 45.4)^2} + \frac{35.9}{(35 - 35.9)^2} + \frac{14.2}{(16 - 14.2)^2} + \frac{4.5}{(4 - 4.5)^2} = 0.310$$

and the number of degrees of freedom is  $k - p - 1 = 4 - 1 - 1 = 2$  so that the critical value from tables is  $\chi_{0.05,2}^2 = 5.99$ . Clearly  $0.310 < 5.99$  and we accept the null hypothesis that the underlying distribution is Poisson. Note that the decision to accept the value  $E_3 = 4.5$  is fairly marginal and that some personal judgement in such situations as to whether such values should be accepted or combined with another class is unavoidable.



Using the data of the previous exercise but combining the expected frequencies of the last two classes, perform a chi-squared test to determine the validity of the assumption that the occurrence of faults in the components is Poisson.

### Your solution

The data are given in 4 classes with observed frequencies as shown. The expected frequencies using the Poisson formula with a mean

$$\mu = \frac{45 \times 0 + 35 \times 1 + 16 \times 2 + 4 \times 3}{100} = 0.79$$

are

$$\mu = \frac{100}{45 \times 0 + 35 \times 1 + 16 \times 2 + 4 \times 3} = 0.79$$

$$p_0 = P(X = 0) = \frac{e^{-\mu} \mu^0}{0!} = e^{-0.79} = 0.454, E_0 = 45.4$$

$$p_1 = P(X = 1) = \frac{e^{-\mu} \mu^1}{1!} = e^{-0.79} \times 0.79 = 0.359, E_1 = 35.9$$

$$p_2 = P(X = 2) = \frac{e^{-\mu} \mu^2}{2!} = \frac{e^{-0.79} \times 0.79^2}{2} = 0.142, E_2 = 14.2$$

$$p_3 = P(X \geq 3) = 1 - (0.454 + 0.359 + 0.142) = 0.045, E_3 = 4.5$$

We will combine the expected frequencies of the last two classes and use 3 classes in total with expected frequencies of  $E_0 = 45.4, E_1 = 35.9, E_2 = 18.7$ .

The test statistic is

$$X^2 = \sum_{i=0}^3 \frac{E_i}{(O_i - E_i)^2} = \frac{45.4}{(45 - 45.4)^2} + \frac{35.9}{(35 - 35.9)^2} + \frac{18.7}{(20 - 18.7)^2} = 0.113$$

and the number of degrees of freedom is  $k - p - 1 = 4 - 1 - 1 = 2$  so that the critical value from tables is  $\chi_{0.05,2}^2 = 5.99$ . Clearly  $0.113 < 5.99$  and we accept the null hypothesis that the underlying distribution is Poisson. Note that the decision to combine the last two classes has not, in this case, affected the acceptance of the null hypothesis.

**Example** A quality control engineer is given the job of checking the voltage output characteristics of a circuit component in a CD player. After checking 100 randomly selected components and plotting a histogram of the results, the engineer concludes that the mean output of the 100 checked components is  $\bar{x} = 6.12$  volts, that the standard deviation is  $s = 0.1$  volts and that the voltage distribution is probably normal. Choose a suitable test to decide whether the assumption of normality is valid at the 5% level of significance.



### Solution

The engineer decides to use a chi-squared test to test the assumption of normality and follow the (common) practice of ensuring that the *expected* frequencies are equal. To do this, the data is put into eight equal classes and the class boundaries calculated as follows.

From the standard normal distribution the  $Z$  values corresponding to class boundaries giving a probability of 0.125 (i.e.  $1/8$ ) may be read off from tables as 0, 0.32, 0.675, 1.15 and  $\infty$  for positive values and 0,  $-0.32$ ,  $-0.675$ ,  $-1.15$  and  $-\infty$  for negative values. Using

$$Z = \frac{x - \bar{x}}{s} \rightarrow x = \bar{x} + Z.s$$

the class boundaries are calculated to be: 6.005, 6.053, 6.088, 6.120, 6.152, 6.188, 6.235. This gives the eight classes, the observed frequencies found by the engineer (you are given this information here), and the expected frequencies as:

Classes	Observed Frequencies $O_i$	Expected Frequencies $E_i$
$x < 6.005$	8	12.5
$6.005 \leq x < 6.053$	11	12.5
$6.053 \leq x < 6.088$	16	12.5
$6.088 \leq x < 6.120$	19	12.5
$6.120 \leq x < 6.152$	18	12.5
$6.152 \leq x < 6.188$	13	12.5
$6.188 \leq x < 6.235$	9	12.5
$6.235 \leq x$	6	12.5

The hypotheses are:  $H_0$ : distribution is normal,  $H_1$ : distribution is not normal

The test statistic is

$$\begin{aligned} X^2 &= \sum_{i=1}^8 \frac{(O_i - E_i)^2}{E_i} \\ &= \frac{(8 - 12.5)^2}{12.5} + \frac{(11 - 12.5)^2}{12.5} + \frac{(16 - 12.5)^2}{12.5} + \frac{(19 - 12.5)^2}{12.5} + \frac{(18 - 12.5)^2}{12.5} \\ &\quad + \frac{(13 - 12.5)^2}{12.5} + \frac{(9 - 12.5)^2}{12.5} + \frac{(6 - 12.5)^2}{12.5} \\ &= 1.62 + 0.18 + 0.98 + 3.38 + 2.42 + 0.02 + 0.98 + 3.38 = 12.96 \end{aligned}$$

and the number of degrees of freedom is  $k - p - 1 = 8 - 2 - 1 = 5$  so that the critical value from tables is  $\chi_{0.05,5}^2 = 11.07$ .

Since  $11.07 < 12.96$  we have sufficient evidence to reject the null hypothesis and so the engineer should conclude that the distribution of voltages is not normal.



An electrical engineer working for a Health and Safety Executive measures the radiation emitted through the closed doors of 100 used microwave ovens. The measurements, in  $\text{mw}/\text{cm}^2$ , are given in the table below.

0.19	0.16	0.14	0.20	0.17	0.21	0.18	0.22	0.26	0.23
0.13	0.17	0.16	0.21	0.18	0.22	0.20	0.23	0.16	0.26
0.19	0.16	0.14	0.20	0.18	0.21	0.19	0.22	0.27	0.24
0.12	0.17	0.15	0.20	0.18	0.22	0.19	0.23	0.29	0.25
0.06	0.16	0.14	0.20	0.17	0.21	0.18	0.22	0.26	0.23
0.13	0.17	0.16	0.20	0.18	0.22	0.19	0.23	0.30	0.25
0.19	0.17	0.14	0.20	0.18	0.21	0.19	0.22	0.27	0.24
0.11	0.17	0.15	0.20	0.18	0.21	0.19	0.23	0.27	0.24
0.13	0.17	0.16	0.21	0.18	0.22	0.19	0.23	0.33	0.25
0.13	0.17	0.16	0.21	0.18	0.22	0.19	0.23	0.36	0.26

The mean radiation of the checked ovens is  $\bar{x} = 0.20 \text{ mw}/\text{cm}^2$ , and the standard deviation is  $s = 0.05 \text{ mw}/\text{cm}^2$ . Verify that the table below giving the eight classes corresponding to the observed and expected frequencies shown is correct.

Classes	Observed Frequencies $O_i$	Expected Frequencies $E_i$
$x < 0.143$	11	12.5
$0.143 \leq x < 0.166$	10	12.5
$0.166 \leq x < 0.184$	19	12.5
$0.184 \leq x < 0.200$	10	12.5
$0.200 \leq x < 0.216$	16	12.5
$0.216 \leq x < 0.234$	17	12.5
$0.234 \leq x < 0.258$	6	12.5
$0.258 \leq x$	11	12.5

Use a chi-squared test to decide whether the radiation readings obtained from the ovens are normally distributed at the 5% level of significance.

**Your solution**

in

From the standard normal distribution the  $Z$  values corresponding to class boundaries giving a probability of 0.125 may be read off from tables as 0, 0.32, 0.675, 1.15 and *infity* for positive values and 0, -0.32, -0.675, -1.15 and  $-\infty$  for negative values. Using

$$Z = \frac{x - \bar{x}}{s} \rightarrow x = \bar{x} + Z \cdot s$$

the class boundaries are calculated to be:

$$0.143, 0.166, 0.184, 0.200, 0.216, 0.234, 0.258$$

This gives the eight classes, the observed frequencies found by the engineer and the expected frequencies as given in the table above.

The hypotheses are:  $H_0$ : distribution is normal,  $H_1$ : distribution is not normal.

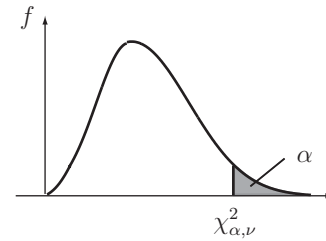
The test statistic is

$$\begin{aligned} X^2 &= \sum_{i=1}^8 \frac{E_i}{(O_i - E_i)^2} \\ &= \frac{12.5}{(11 - 12.5)^2} + \frac{12.5}{(10 + 12.5)^2} + \frac{12.5}{(19 - 12.5)^2} + \frac{12.5}{(10 - 12.5)^2} + \frac{12.5}{(16 - 12.5)^2} \\ &\quad + \frac{12.5}{(17 - 12.5)^2} + \frac{12.5}{(6 - 12.5)^2} + \frac{12.5}{(11 - 12.5)^2} + \frac{12.5}{(12.5)^2} \\ &= 0.18 + 0.5 + 3.38 + 0.5 + 0.98 + 1.62 + 3.38 + 0.18 = 10.72 \end{aligned}$$

and the number of degrees of freedom is  $k - p - 1 = 8 - 2 - 1 = 5$  so that the critical value from tables is  $\chi_{0.05,5}^2 = 11.07$ .

Since  $10.72 < 11.07$  we do not have sufficient evidence to reject the null hypothesis and so the engineer should conclude that the distribution of microwave radiation readings taken from the ovens is normal.

Percentage Points  $\chi_{\alpha,\nu}^2$  of the  $\chi^2$  distribution



$\alpha$	0.995	0.990	0.975	0.950	0.900	0.500	0.100	0.050	0.025	0.010	0.005
$\nu$											
1	0.00	0.00	0.00	0.00	0.02	0.45	2.71	3.84	5.02	6.63	7.88
2	0.01	0.02	0.05	0.01	0.21	1.39	4.61	5.99	7.38	9.21	10.60
3	0.07	0.11	0.22	0.35	0.58	2.37	6.25	7.81	9.35	11.34	12.83
4	0.21	0.30	0.48	0.71	1.06	3.36	7.78	9.49	11.14	13.28	14.86
5	0.41	0.55	0.83	1.15	1.61	4.35	9.24	11.07	12.83	15.09	16.75
6	0.68	0.87	1.24	1.64	2.20	5.35	10.65	12.59	14.45	16.81	18.55
7	0.99	1.24	1.69	2.17	2.83	6.35	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	7.34	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	8.34	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	9.34	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	10.34	17.28	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	6.30	11.34	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	12.34	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	13.34	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.27	7.26	8.55	14.34	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	15.34	23.54	26.30	28.85	31.00	34.27
17	5.70	6.41	7.56	8.67	10.09	16.34	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.87	17.34	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	18.34	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	19.34	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	20.34	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	21.34	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	22.34	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	23.34	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	24.34	34.28	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	25.34	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	26.34	36.74	40.11	43.19	46.96	49.65
28	12.46	13.57	15.31	16.93	18.94	27.34	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	28.34	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	29.34	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	39.34	51.81	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	49.33	63.17	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	46.46	59.33	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	55.33	69.33	85.53	90.53	95.02	100.42	104.22
80	51.17	53.54	57.15	60.39	64.28	79.33	96.58	101.88	106.63	112.33	116.32
90	59.20	61.75	65.65	69.13	73.29	89.33	107.57	113.14	118.14	124.12	128.30
100	67.33	70.06	74.22	77.93	82.36	99.33	118.50	124.34	129.56	135.81	140.17