

Statistics: 1.4 Chi-squared goodness of fit test

Rosie Shier. 2004.

1 Introduction

A chi-squared test can be used to test the hypothesis that observed data follow a particular distribution. The test procedure consists of arranging the n observations in the sample into a frequency table with k classes. The chi-squared statistic is:

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

The number of degrees of freedom is $k - p - 1$ where p is the number of parameters estimated from the (sample) data used to generate the hypothesised distribution.

2 Example: The Poisson Distribution

Let X be the number of defects in printed circuit boards. A random sample of $n = 60$ printed circuit boards is taken and the number of defects recorded. The results were as follows:

Number of Defects	Observed Frequency
0	32
1	15
2	9
3	4

Source: Applied Statistics and Probability for Engineers - Montgomery and Runger

Does the assumption of a Poisson distribution seem appropriate as a model for these data?

The null hypothesis is $H_0 : X \sim \text{Poisson}$

The alternative hypothesis is $H_1 : X$ does not follow a Poisson distribution.

The mean of the (assumed) Poisson distribution is unknown so must be estimated from the data by the sample mean:

$$\hat{\mu} = ((32 \times 0) + (15 \times 1) + (9 \times 2) + (4 \times 3)) / 60 = 0.75$$

Using the Poisson distribution with $\mu = 0.75$ we can compute p_i , the hypothesised probabilities associated with each class. From these we can calculate the expected frequencies (under the null hypothesis):

$$p_0 = P(X = 0) = \frac{e^{-0.75}(0.75)^0}{0!} = 0.472 \Rightarrow E_0 = 0.472 \times 60 = 28.32$$

$$p_1 = P(X = 1) = \frac{e^{-0.75}(0.75)^1}{1!} = 0.354 \Rightarrow E_1 = 0.354 \times 60 = 21.24$$

$$p_2 = P(X = 2) = \frac{e^{-0.75}(0.75)^2}{2!} = 0.133 \Rightarrow E_2 = 0.133 \times 60 = 7.98$$

$$p_3 = P(X \geq 3) = 1 - (p_0 + p_1 + p_2) = 0.041 \Rightarrow E_3 = 0.041 \times 60 = 2.46$$

Note: The chi-squared goodness of fit test is not valid if the expected frequencies are too small. There is no general agreement on the minimum expected frequency allowed, but values of 3, 4, or 5 are often used. If an expected frequency is too small, two or more classes can be combined. In the above example the expected frequency in the last class is less than 3, so we should combine the last two classes to get:

Number of Defects	Observed Frequency	Expected Frequency
0	32	28.32
1	15	21.24
2 or more	9	10.44

The chi-squared statistic can now be calculated:

$$\chi^2 = \sum \frac{(O - E)^2}{E} = \frac{(32 - 28.32)^2}{28.32} + \frac{(15 - 21.24)^2}{21.24} + \frac{(13 - 10.44)^2}{10.44} = 2.94$$

The number of degrees of freedom is $k - p - 1$. Here we have $k = 3$ classes and we have $p = 1$ because we had to estimate one parameter (the mean, μ) from the data. So, our chi-squared statistic has $3 - 1 - 1 = 1$ df. If we look up 2.94 in tables of the chi-squared distribution with df=1, we obtain a p-value of $0.05 < p < 0.1$. We conclude that there is no real evidence to suggest the the data DO NOT follow a Poisson distribution, although the result is borderline.

3 Goodness of fit test for other distributions

The chi-squared goodness of fit test can be used for any distribution. For a discrete distribution the procedure is as described above. For a continuous distribution it is necessary to group the data into classes. Common practice used to carry out the goodness of fit test is to choose class boundaries so that the expected frequencies are equal for each class. For example, suppose we decide to use $k = 8$ classes. For the standard normal distribution the intervals that divide the distribution into 8 equally likely segments are $[0, 0.32)$, $[0.32, 0.675)$, $[0.675, 1.15)$, $[1.15, \infty)$ and their four “mirror image” intervals below zero. For each interval $p_i = 1/8$ so the expected frequencies would be $E_i = n/8$.