# Teacher Delivery Guide Statistics: Chi-Squared Tests

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Specification** | **Ref.** | **Learning outcomes** | **Notes** | **Notation** | **Exclusions** |
| **Y422 STATISTICS MAJOR: CHI-SQUARED TESTS (a)**  **Y432 STATISTICS MINOR: CHI-SQUARED TESTS**  **Y412 STATISTICS a: CHI-SQUARED TESTS** | | | | | |
| Contingency tables | b16 | Be able to interpret bivariate categorical data in a contingency table. | Numerical data can be put into categories, but this loses information. |  |  |
| test for a contingency table | H1 | Be able to apply the  test (chi-squared) to a contingency table. | Only ‘H0: No association between the factors’ or H0: ‘variables are independent’ will be tested.  Calculating degrees of freedom is expected.  Knowing how to calculate observed values and contributions to the test statistic are expected, but repetitive calculations will not be required.  Learners should state whether there is sufficient evidence or not to reject H0 and then give a non-assertive conclusion in context e.g. ‘There is not sufficient evidence to believe that there is association between … and …’. |  | Yates’ continuity correction is not expected, though its appropriate use will not be penalised. |
| H2 | Be able to interpret the results of a  test using tables of critical values or the output from software. | Output from software may be given as a *p*-value.  Interpretation may involve considering the individual cells in the table of contributions to the test statistic. |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Specification** | **Ref.** | **Learning outcomes** | **Notes** | **Notation** | **Exclusions** |
| test for goodness of fit | H3 | Be able to carry out a  test for goodness of fit of a uniform, binomial or Poisson model. | Only ‘H0: the given model fits the data’ or ‘H0; the given model is suitable’ will be tested.  Calculating degrees of freedom is expected.  Knowing how to calculate observed values and contributions to the test statistic is expected, but repetitive calculations will not be required.  Learners should be aware that cells are often combined when there are small expected frequencies, but will not have to make such decisions in examination questions.  Learners should state whether there is sufficient evidence or not to reject H0 and then give a non-assertive conclusion in context e.g. ’It is reasonable to believe that the … model is suitable.’ |  |  |
| H4 | Be able to interpret the results of a  test using tables of critical values or the output from software. | Output from software may be given as a *p*-value. |  |  |

**Thinking Conceptually**

**General approaches**

The different stages of using the chi-squared test cannot really be separated, and so it is probably better to do a single example from start to finish before practising with students. Look first at contingency tables, and then fitting a given distribution to a set of data.

Introduce the formula for the chi-squared statistic first and then show how to calculate the degrees of freedom in each case; as when fitting a distribution where no parameters are estimated or as for a contingency table. Emphasise the meaning of degrees of freedom and show how the formula for a contingency table is related to the number of “free” cells in the contingency table once the totals are given.

Show students how to use appropriate technology to check their working for a chi-squared test for a contingency table.

Demonstrate the use of tables for critical values.

It may be useful to refer to the curve of the distribution to demonstrate the critical region and the calculated value against the significant value to aid students’ understanding and clarify the interpretation of the result of the test, which is after all the whole point.

**Common misconceptions**

Learners often get the formula for the chi-squared test statistic wrong, either forgetting to square the numerator or using observed values instead of expected ones for the denominator.

Using the wrong degrees of freedom can also be a problem, especially if learners remember the formula and try to apply that in the case of fitting a distribution such as the Poisson – it will lead them to try and use 0 which is meaningless. Students also need to remember that for the goodness of fit test, any parameters estimated from the data reduce the number of degrees of freedom.

Some learners mistakenly think that percentages rather than frequencies can be used in a chi-squared test.

**Conceptual links to other areas of the specification**

This topic links back to the various distributions that have been studied and also can be linked forward to work on Non-Parametric Tests, in the sense that it is possible to talk about the sorts of tests that can be carried out on collected data. The concept of degrees of freedom is also needed in the non-parametric testing section, and clarifying this as early as possible will avoid confusion.

**Thinking Contextually**

The chi-squared test is best introduced by an example. It might be possible to use one of the video links to examples suggested below to work through, stopping it at important stages and discussing the theory that is associated with it. Work through a basic example with a simple contingency table first, Using the Integrated Mathematical Problem Solving resource will allow you to focus on the concepts rather than the calculations.

As a separate lesson/section, recap the distributions done so far and look at fitting a distribution to a set of data – calculating probabilities and converting to expected values, and then carrying out a goodness of fit test. Look at data such as the number of students in the class absent for every lesson (Poisson?), the number of pieces of chewing gum stuck on each square metre of playground (Poisson?), the number of heads when 5 coins are tossed repeatedly (binomial?). Set up some of these practical situations and allow students to investigate them, drawing out at the end what the chi-squared test has told them about their assumptions.

# Resources

| **Title** | **Organisation** | **Description** | **Ref** |
| --- | --- | --- | --- |
| [Statistical Tests in Medicine](http://mei.org.uk/IMPS) | MEI | An introduction to chi-squared test for contingency table – this is a good initial example, it is in the Mathematics of Biology resources | b16 and H1 |
| [Chi-Square Test for Independence](http://stattrek.com/chi-square-test/independence.aspx) | Stat Trek | Notes explains how to conduct a chi-square test for independence. | b16 and H1 |
| [Chi-squared test of independence](https://www.youtube.com/watch?v=N6rCqIldUD8) | Stats with Mr. r | 7-minute video demonstration on using chi-squared to test independence. | b16 and H1 |
| [Chi-squared for 2-way tables](http://www.statstutor.ac.uk/resources/uploaded/chis-quared-tests-of-association.pdf) | Mathematics Learning Support Centre | A handout with a good worked example on the first two pages. The later stages are about use of the SPSS software and can be omitted. | b16 and H1 |
| [Chi-squared and Percentages](http://www.s253053503.websitehome.co.uk/msv/msv-10.html) | MSV | Investigation of the problems encountered if percentages are used rather than frequencies, promoting discussion and a further depth of understanding. | b16 and H1 |
| [Introduction to chi-squared tests](https://www.youtube.com/watch?v=-ltTwU9gu-U) | Stats with Mr. r | 11-minute video introducing the ideas of the chi-squared tests. | b16, H1 and H2 |
| [Chi-square Distribution Worksheet](http://worksheets.tutorvista.com/chi-square-distribution-worksheet.html) | Tutor Vista | An online multiple choice worksheet – would be good to use with a class via a digital projector. | b16, H1 and H2 |
| [Chi-square Goodness of Tit Test](http://www.biologyjunction.com/ChiSquare.doc) | Biology Junction | A explanation of the goodness of fit test, and some activities to do, leading on to the use of goodness of fit to analyse results in genetics. | b16, H1 and H2 |
| [Chi-square practice problems](https://www.biologycorner.com/worksheets/chi_square_practice.html) | Biology Corner | Four problems on a worksheet that can be tackled using the chi-squared test( although no answers). | b16, H1 and H2 |
| [Why do females sing?](http://kbsgk12project.kbs.msu.edu/wp-content/uploads/2013/12/female.song_.chi_.square.TEACHERGUIDEdocx.pdf) | Michigan State University | An example looking at the use of the chi-squared to investigate a problem in biology – looking at why female birds sing and whether it is related to their breeding cycle. Calculated values are in red and can be removed to turn this into a worksheet. Quite complex and perhaps appropriate for more able students or group discussion. | b16, H1 and H2 |
| [How to perform a chi-squared tests](https://www.youtube.com/watch?v=V4SRgabFbz0) | Eugene O’Loughlin | 7 minute video – useful for revision or recap. | b16, H1 and H2 |
| [Chi-squared test - explained](https://www.youtube.com/watch?v=1Ldl5Zfcm1Y) | Math Meeting | 13 minute video. Good summary, and useful for revision. | b16, H1 and H2 |
| [Goodness of Fit Tests: Writing Conclusions](https://www.geogebra.org/m/vD8jtZr2) | Geogebra | 24 calculations for learners to discuss the conclusion that can be drawn from the hypothesis test. | H3 and H4 |
| [Chi-Square Goodness of Fit Test](http://stattrek.com/chi-square-test/goodness-of-fit.aspx?Tutorial=AP) | Stat Trek | Notes on how to conduct a chi-square goodness of fit test**.** | H3 and H4 |

**OCR Resources**: *the small print*OCR’s resources are provided to support the delivery of OCR qualifications, but in no way constitute an endorsed teaching method that is required by the Board, and the decision to use them lies with the individual teacher. Whilst every effort is made to ensure the accuracy of the content, OCR cannot be held responsible for any errors or omissions within these resources.

© OCR 2018 - This resource may be freely copied and distributed, as long as the OCR logo and this message remain intact and OCR is acknowledged as the originator of this work.

OCR acknowledges the use of the following content: n/a

Please get in touch if you want to discuss the accessibility of resources we offer to support delivery of our qualifications: [resources.feedback@ocr.org.uk](mailto:resources.feedback@ocr.org.uk)

We’d like to know your view on the resources we produce. By clicking on ‘[Like](mailto:resources.feedback@ocr.org.uk?subject=I%20liked%20the%20AS%20and%20A%20Level%20Further%20Mathematics%20B%20(MEI)%20Teacher%20Delivery%20Guide%20Statistics:%20Chi%20Squared%20Tests)’ or ‘[Dislike](mailto:resources.feedback@ocr.org.uk?subject=I%20disliked%20the%20AS%20and%20A%20Level%20Further%20Mathematics%20B%20(MEI)%20Teacher%20Delivery%20Guide%20Statistics:%20Chi%20Squared%20Tests)’ you can help us to ensure that our resources work for you. When the email template pops up please add additional comments if you wish and then just click ‘Send’. Thank you.

Whether you already offer OCR qualifications, are new to OCR, or are considering switching from your current provider/awarding organisation, you can request more information by completing the Expression of Interest form which can be found here: [www.ocr.org.uk/expression-of-interest](http://www.ocr.org.uk/expression-of-interest)

Looking for a resource? There is now a quick and easy search tool to help find free resources for your qualification:   
[www.ocr.org.uk/i-want-to/find-resources/](http://www.ocr.org.uk/i-want-to/find-resources/)