PSY.RES11 The Chi-Squared Test: Rationale, Application and Critical Values

AQA 7182 · Calculators not allowed · about 60 minutes

Total Marks

Name: ______________________________ Date: ____ / ____ / ______

Answer ALL questions. Show all your working.

1 Define what the Chi-Squared test of association is used for in psychological research, naming the type of data and a typical design where it applies.

Define the Chi-Squared test of association and state when it is appropriate.

(Total for Question 1 is 1 mark)

2 Write the Chi-Squared formula as used in contingency-table calculations, and explain in one sentence what each symbol represents in the context of observed and expected frequencies.

State the Chi-Squared calculation formula and explain its symbols.

(Total for Question 2 is 2 marks)

3 Explain how degrees of freedom (df) are calculated for a contingency table in a Chi-Squared test, giving the formula and an example in words for a 2 by 3 table.

Describe calculation of degrees of freedom for a contingency table and give a 2x3 example in words.

(Total for Question 3 is 2 marks)

4 When deciding between the Chi-Squared test and other inferential tests, name two features of the data or design that indicate Chi-Squared is the correct choice, and give one feature that would make a t-test more appropriate.

Identify two features pointing to Chi-Squared and one feature favouring a t-test.

(Total for Question 4 is 3 marks)

5 Worked calculation context: A researcher investigates whether offering a short stress-management leaflet affects whether students seek counselling. In two independent groups of 60 students each, the researcher records whether a student contacted counselling services within four weeks. The observed frequencies are:

Group A (leaflet): contacted = 24, not contacted = 36
Group B (no leaflet): contacted = 15, not contacted = 45

Use a 2x2 contingency table and perform a Chi-Squared test of association at $p \leq 0.05$. Show all working to gain full credit and state the conclusion.

Complete the full Chi-Squared calculation for the 2x2 table above and state the conclusion at $p \leq 0.05$.

(Total for Question 5 is 9 marks)

---

6 State the critical value from a Chi-Squared table for df = 1 at $p \leq 0.05$ (two-tailed convention used for association tests).

Give the critical value for df = 1 at $p \leq 0.05$.

(Total for Question 6 is 1 mark)

---

7 For the 2x2 contingency table in Question 5, calculate the degrees of freedom and state it.

Calculate the degrees of freedom for the 2x2 table used in Q5.

(Total for Question 7 is 1 mark)

---

8 Explain the difference between one-tailed and two-tailed use of a critical values table in Chi-Squared testing, and state which is conventionally used for tests of association.

Explain one-tailed vs two-tailed in Chi-Squared and state the convention for association tests.

(Total for Question 8 is 2 marks)

---

9 State one limitation of the Chi-Squared test related to expected cell frequencies in small samples, and name an alternative test that can be used instead when this limitation applies.

Give one limitation about expected frequencies and an alternative test.

(Total for Question 9 is 2 marks)

---

10 Explain briefly why the Chi-Squared test is not suitable for continuous interval-scale data and name the family of tests more appropriate in that case.

Explain why Chi-Squared is unsuitable for continuous data and name appropriate tests.

(Total for Question 10 is 2 marks)

11 Using the contingency table from Question 5, calculate the expected frequency for the cell 'not contacted, no-leaflet' and show the working.

Calculate the expected frequency for not contacted, no-leaflet from Q5.

(Total for Question 11 is 1 mark)

---

12 Evaluate the use of the Chi-Squared test of association in psychological research. In your answer discuss strengths, limitations, assumptions, and practical considerations such as sample size and interpretation of significant results. Refer to how these points would matter in studies using categorical measures such as diagnosis categories, treatment uptake, or survey categories.

Evaluate the use of the Chi-Squared test of association in psychological research.

(Total for Question 12 is 16 marks)

Mark scheme · PSY.RES11 The Chi-Squared Test: Rationale, Application and Critical Values

Question 1

- **B1** A test for association or difference between two categorical/nominal variables using frequency counts, typically applied to an independent-groups contingency table oe
- **Answer: A test for association or difference between two categorical (nominal) variables using frequency counts, usually applied to data in an independent-groups contingency table.**

Question 2

- **M1** formula given correctly as X^2 = sum((O - E)^2 / E) cao
- **A1** explains O = observed frequency in a cell and E = expected frequency in that cell under the null hypothesis oe
- **Answer: X^2 = sum((O - E)^2 / E). O is the observed frequency in a cell and E is the expected frequency for that cell under the null hypothesis.**

Question 3

- **M1** states df = (number of rows - 1) x (number of columns - 1) cao
- **A1** gives example in words: for a 2 by 3 table df = (2 - 1) x (3 - 1) = 1 x 2 = 2 oe
- **Answer: df = (rows - 1) x (columns - 1). For a 2 by 3 table df = (2 - 1) x (3 - 1) = 2.**

Question 4

- **M1** feature indicating Chi-Squared: both variables are nominal/categorical and recorded as counts/frequencies oe
- **M1** feature indicating Chi-Squared: independent groups design where participants are in separate categories (contingency table) oe
- **A1** feature favouring a t-test: the dependent variable is continuous (interval/ratio) and approximately normally distributed so group means are compared oe
- **Answer: Chi-Squared: two nominal categorical variables recorded as frequencies, and an independent-groups design with a contingency table. A t-test is appropriate if the DV is continuous and approximately normally distributed so that group means are compared.**

Question 5

- **M1** calculates the row totals, column totals and grand total correctly, e.g. leaflet total 60, no-leaflet total 60, contacted total 39, not contacted total 81, grand total 120 oe
- **M1** shows how to calculate an expected value using E = (row total x column total) / grand total and substitutes correct numbers for at least one cell oe
- **A1** gives the correct expected value for the contacted, leaflet cell as 19.5 and for the contacted, no-leaflet cell as 19.5 oe
- **A1** gives the correct expected value for the not-contacted, leaflet cell as 40.5 and for the not-contacted, no-leaflet cell as 40.5 oe
- **M1** substitutes O and E into the formula X^2 = sum((O - E)^2 / E) for each of the four cells oe
- **A1** calculates and states the numerical value for X^2 to two decimal places as 3.08 awrt 3.08 oe
- **M1** calculates degrees of freedom correctly for the 2x2 table as df = (2 - 1) x (2 - 1) = 1 oe
- **A1** states the correct critical value at $p \leq 0.05$ for df = 1 as 3.841 and compares it with the calculated X^2 oe
- **A1** draws the correct conclusion that $X^2 = 3.08 < 3.841$ so the result is not significant at $p \leq 0.05$ and the null hypothesis is retained oe
- **Answer: Totals: leaflet 60, no-leaflet 60, contacted 39, not contacted 81, grand total 120. Expected frequencies: contacted, leaflet = (60 x 39)/120 = 19.5; contacted, no-leaflet = 19.5; not contacted, leaflet = (60 x 81)/120 = 40.5; not contacted, no-leaflet = 40.5. Components: (24 - 19.5)^2/19.5 = 1.0385, (36 - 40.5)^2/40.5 = 0.5, (15 - 19.5)^2/19.5 = 1.0385, (45 - 40.5)^2/40.5 = 0.5. Sum X^2 = 3.077 = 3.08 (2 dp). df = 1. Critical value at $p \leq 0.05$ is 3.841. Since 3.08 < 3.841 the result is not significant at $p \leq 0.05$; retain the null hypothesis, no clear association between leaflet and contacting counselling in this sample.**

Question 6

- **B1** critical value = 3.841 cao

Question 7

- **B1** df = (2 - 1) x (2 - 1) = 1 cao
- **Answer: df = 1**

Question 8

- **M1** explains that one-tailed critical values test for an effect in a specified direction, while two-tailed test for any difference in either direction oe
- **A1** states that Chi-Squared tests of association use the upper-tail critical value (conventionally the two-tailed table is not relevant because Chi-Squared is always non-negative, so the standard critical values give the threshold for significance in the upper tail) oe
- **Answer: One-tailed looks for an effect in a specified direction while two-tailed looks for any effect; for Chi-Squared association tests we use the upper-tail critical value (effectively the standard table threshold for the upper tail) since X^2 is always non-negative.**

Question 9

- **M1** limitation: Chi-Squared may be inaccurate when expected frequencies in one or more cells are too small (commonly when any expected < 5 for a 2x2 table) oe
- **A1** alternative: Fisher's exact test can be used for small samples/low expected frequencies oe
- **Answer: Chi-Squared can be inaccurate when expected frequencies are small (often any expected < 5), so Fisher's exact test is an appropriate alternative for small samples or 2x2 tables.**

Question 10

- **M1** explains that Chi-Squared uses counts of categories and would require arbitrary binning of continuous data which loses information oe
- **A1** names appropriate tests such as t-tests or ANOVA for comparing group means of continuous data (parametric tests) oe
- **Answer: Chi-Squared uses categorical counts and would force arbitrary grouping of continuous data, losing information; for continuous interval/ratio data use t-tests or ANOVA to compare group means.**

Question 11

- **B1** correct expected frequency 40.5 with working E = (row total 60 x column total 81) / 120 = 4860/120 = 40.5 cao
- **Answer: Expected frequency = (60 x 81) / 120 = 4860 / 120 = 40.5.**

Question 12

- **Level 4** (13-16): Knowledge of the Chi-Squared test and its use is accurate and well developed. Evaluation is thorough and balanced, considering methodological, statistical and practical issues, and is applied to psychological examples. There is effective, integrated use of strengths and limitations, and conclusions are clear and justified. Specialist terminology is used correctly.
- **Level 3** (9-12): Knowledge is sound with some development. Evaluation addresses several relevant points, but may be less comprehensive or less well integrated. Application to psychological examples is present. Explanation is mostly clear and terminology is generally used correctly.
- **Level 2** (5-8): Knowledge of the test is partial and evaluation is limited or superficial. Some relevant points are made but development and application to psychology is limited. The answer may lack clarity or organisation in places.
- **Level 1** (1-4): Very limited knowledge and evaluation. Points may be simple or list-like and application to psychology is minimal or missing. Answer lacks structure and specialist terms are used incorrectly or not at all.
- **Level 0** (0): No creditable content.
- **Indicative content:**
  - AO1: Description of what Chi-Squared tests: a non-parametric test for association or difference between categorical variables using observed versus expected frequencies in contingency tables.
  - AO1: Formula and degrees of freedom: X^2 = sum((O - E)^2 / E); df = (rows - 1) x (columns - 1).
  - AO3: Strengths: simple to compute and interpret, widely applicable to categorical data such as diagnosis categories, treatment uptake, survey response categories; does not assume normal distribution of the underlying variables since data are categorical.
  - AO3: Strengths: provides a clear decision rule using critical values and p thresholds, useful for initial exploratory analysis of associations between categorical variables in applied settings.
  - AO3: Limitations: assumes expected frequencies are sufficiently large; when expected cell counts are small (commonly any expected < 5) the approximation to the chi-squared distribution is poor and results may be invalid.
  - AO3: Limitations: for small samples or 2x2 tables with low expected counts, exact tests such as Fisher's exact test or use of continuity corrections should be considered.
  - AO3: Limitations: Chi-Squared indicates whether an association exists but not its strength or direction; measures of effect size (phi, Cramer's V) are needed to assess practical significance, especially in large samples where tiny effects can be statistically significant.
  - AO3: Practical considerations: categorising continuous data to use Chi-Squared discards information and can be misleading; ensure categories are valid and meaningful in psychological contexts (e.g. diagnostic thresholds justified).
  - AO3: Consider the role of sample size: large samples increase power but can make trivial associations significant; small samples may fail assumptions and lead to Type II errors if effect exists.
  - AO3: Interpretation issues: significance does not imply causation, especially when data come from observational designs; researchers must consider confounding variables, sampling bias and independence of observations.
  - AO3: Ethical and reporting considerations: report expected frequencies, df, X^2 value and p value, and effect size; be transparent about any continuity corrections or use of exact tests.
  - AO3: Application to psychological examples: in a study of treatment uptake by diagnostic category, Chi-Squared can identify whether uptake differs by diagnosis, but follow-up with effect size and consideration of clinical importance is essential; discussion of generalisability if sample is clinic-based versus population-based.
  - AO3: Balanced judgment: Chi-Squared is a valuable, accessible tool for categorical data but must be used with attention to assumptions, supplemented by effect size and cautious interpretation about causality and practical importance.