

Inferential Statistics

Lecture 13: Independence, homogeneity and goodness of fit

Jérémy Hervelin

Pforzheim University
j.hervelin@lb.hs-pforzheim.de

7 December 2026 13:45–15:15
Winter semester 2026/27

Adapted from the teaching materials of Prof. Dr. Thilo Klein.

Our 90-minute route

- First 15 minutes: retrieve prerequisites and introduce the question.
- Next 35 minutes: worked examples, with short checks after each concept.
- Next 25 minutes: paired activity and interpretation.
- Final 15 minutes: synthesis, questions and an applied discussion.

Categorical inference: count and compare

- Use contingency tables to test independence or homogeneity.
- Calculate expected counts and the Pearson chi-square statistic.
- Check the approximation before interpreting the p-value.
- Test goodness of fit to a prespecified categorical distribution.

Independence and homogeneity: different designs

Independence: one sampled population, two categorical variables; ask whether they are associated.

Homogeneity: independently sampled groups; ask whether the category distribution is the same across groups.

The Pearson calculation can be identical, but the sampling design and interpretation differ. Observations must be independent in the relevant sampling units.

Original education example

	Advanced maths	No advanced maths	Total
Vocational	37	55	92
Other	115	144	259
Total	152	199	351

Under independence, the advanced-maths share should be the same in both education groups. Compare observed counts with counts predicted by the null.

Expected counts under independence

For row i , column j and total N ,

$$E_{ij} = \frac{(\text{row total } i)(\text{column total } j)}{N}.$$

For the upper-left cell: $92 \times 152/351 = 39.84$.

Expected counts are approximately

$$\begin{pmatrix} 39.84 & 52.16 \\ 112.16 & 146.84 \end{pmatrix}.$$

They sum to the observed row and column totals. Use unrounded counts in the test statistic.

Classroom check – Q1

Exam-style

1–2 minutes

A table has row total 92, column total 152, grand total 351. Find the expected count under independence.

Think individually, then compare your reasoning with a neighbour.

Lecture 13, classroom quiz Q1. Same question number as the printed quiz.

Pearson chi-square statistic

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}.$$

Large values show disagreement with the null. With adequate expected counts, the null reference distribution is approximately $\chi^2_{(r-1)(c-1)}$.

Squaring prevents positive and negative deviations from cancelling; division by E_{ij} scales each deviation relative to its expected size.

Classroom check – Q2

Exam-style

1–2 minutes

A 3-by-2 independence table has how many degrees of freedom?

Think individually, then compare your reasoning with a neighbour.

Lecture 13, classroom quiz Q2. Same question number as the printed quiz.

Original table: compute and interpret

Without continuity correction, $X^2 = 0.48408$, $df = 1$, and $p = 0.4866$.

We fail to reject independence at 5%. This does not prove independence, and it gives no causal conclusion about education and subject choice.

For a two-by-two table, R's default applies a continuity correction. Use `correct=FALSE` to reproduce the uncorrected Pearson calculation taught here.

Classroom check – Q5

Exam-style

1–2 minutes

Failing to reject an independence null proves no association.

- A. True;
- B. False

Think individually, then compare your reasoning with a neighbour.

Lecture 13, classroom quiz Q5. Same question number as the printed quiz.

Check the approximation

- Use counts of mutually exclusive categories, not percentages supplied without their denominators.
- Inspect expected counts, not only observed counts.
- As a conservative classroom check, require every expected count to be at least 5.
- For sparse tables consider an appropriate exact or Monte Carlo method; do not merge categories solely to obtain significance.

Classroom check – Q3

Concept check

1–2 minutes

The Pearson approximation should be assessed using. . .

- A. Expected counts;
- B. Only observed percentages;
- C. The p-value alone

Think individually, then compare your reasoning with a neighbour.

Lecture 13, classroom quiz Q3. Same question number as the printed quiz.

Original homogeneity example

The original two-by-six gender/party table has 255 observations:

$$\begin{pmatrix} 13 & 10 & 3 & 11 & 5 & 23 \\ 55 & 30 & 20 & 26 & 24 & 35 \end{pmatrix}.$$

All expected counts exceed 5. Pearson $X^2 = 10.852$, $df = 5$, and $p \approx 0.0544$.

Interpret the result at the prespecified 5% level, and explain why 0.0544 does not establish equal preference distributions.

Sparse cells in the original R practical

The separate two-by-eight election table used in the R source has 4 of 16 expected counts below 5, with minimum 1.78.

The usual asymptotic p-value needs caution. In the revised practical, inspect the expected table and also calculate a conditional Monte Carlo p-value with a fixed seed.

Simulated p-values carry Monte Carlo error; record the number of replicates and the seed.

Goodness of fit: one categorical variable

For specified category probabilities $p_1, \dots, p_k$, expected counts are $E_j = np_j$ and

$$\chi^2 = \sum_{j=1}^k \frac{(O_j - np_j)^2}{np_j}.$$

Under the null and suitable expected counts, compare with χ_{k-1}^2 when no parameters are estimated from these data.

The categories must cover every possible outcome, with no overlap.

Original die-roll example

Thirty rolls give counts (6, 2, 4, 6, 8, 4) for faces 1 through 6.

Under a fair-die model, each expected count is 5:

$$\chi^2 = \frac{(6 - 5)^2 + (2 - 5)^2 + \cdots + (4 - 5)^2}{5} = 4.4.$$

With 5 degrees of freedom, $p = 0.4934$. There is no strong evidence against the fair-die probabilities in this sample.

Classroom check – Q4

Exam-style

1–2 minutes

Fair-die counts are 6,2,4,6,8,4 out of 30. Find the Pearson statistic.

Think individually, then compare your reasoning with a neighbour.

Lecture 13, classroom quiz Q4. Same question number as the printed quiz.

Fitted distributions need extra care

When m parameters are estimated, a $k - 1 - m$ reference degree of freedom can apply under appropriate regularity and fitting conditions.

This is not an automatic correction for arbitrary estimated models or data-selected bins. Include tail categories and check expected counts.

Sparse expected counts or data-dependent bins can invalidate the usual chi-square reference distribution. Choose bins before inspecting the data and check the expected counts.

A test does not prove a distribution

- Failing to reject a normal model does not establish normality.
- A goodness-of-fit result depends on the chosen categories and sample size.
- Plot the data, check tail behaviour and consider which model deviations matter for the planned analysis.
- Do not use a goodness-of-fit p-value as permission to ignore the sampling design.

Applied workshop: education and employment

A synthetic survey records education (school, vocational, university) and employment (yes/no).

State the independence null; form a three-by-two count table; calculate expected counts and degrees of freedom; choose an appropriate method.

Then report group employment shares. A global chi-square rejection indicates association but does not identify a causal education effect or explain which comparisons drive the association.

Independent practice

- Practise the five online questions for this lecture; 80% provides access to the next practice module.
- Course materials remain accessible regardless of your practice progress.
- Explain your reasoning before comparing answers.
- Review the method, assumptions and interpretation as well as the calculation.

Sources and teaching notes

Adapted from Prof. Dr. Thilo Klein's supplied Inferential Statistics materials.

Revised teaching edition: Jérémy Hervelin, Pforzheim University, winter semester 2026/27.

New youth-employment, programme and interview-score examples are synthetic illustrations.