

Chi-Square Tests for Two-Way Tables

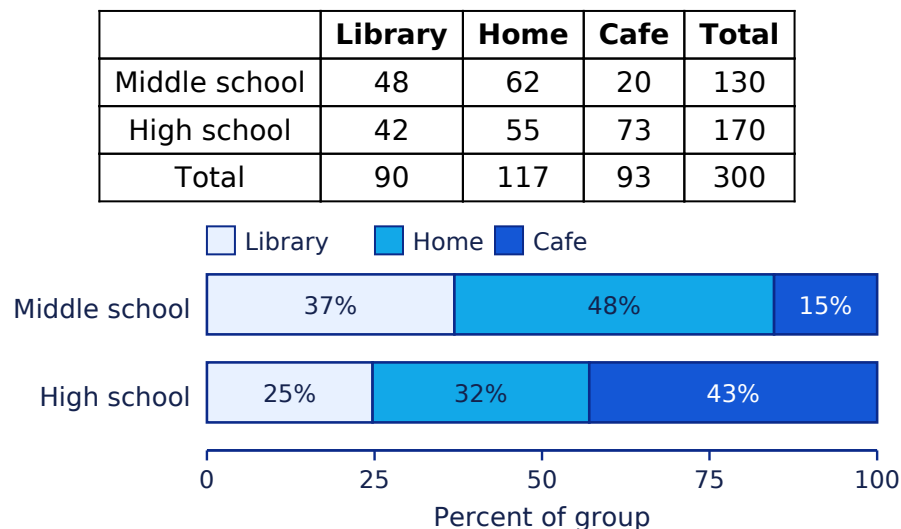
Grade 12 – AP Statistics math · Lesson · Standards: AP Statistics Unit 3 (2026-27)

Quick answer: A chi-square test compares observed counts in a two-way table with the counts expected if the null hypothesis is true. The statistic is the sum of $(O - E)^2/E$, with $df = (rows - 1)(columns - 1)$, and it tests whether distributions are the same (homogeneity) or two variables are associated (independence).

Standards: AP Statistics Unit 3 (2026-27) · **Level:** Grade 12 AP Statistics

What does a chi-square test compare?

A chi-square test uses counts in a two-way table. The **test for homogeneity** asks whether the distribution of one categorical variable is the same in several populations or treatment groups. The **test for independence** asks whether two categorical variables are associated in a single population. The calculations are identical; only the hypotheses and sampling design differ.



Conditional distribution of study spot by school level.

How do you compute expected counts and the statistic?

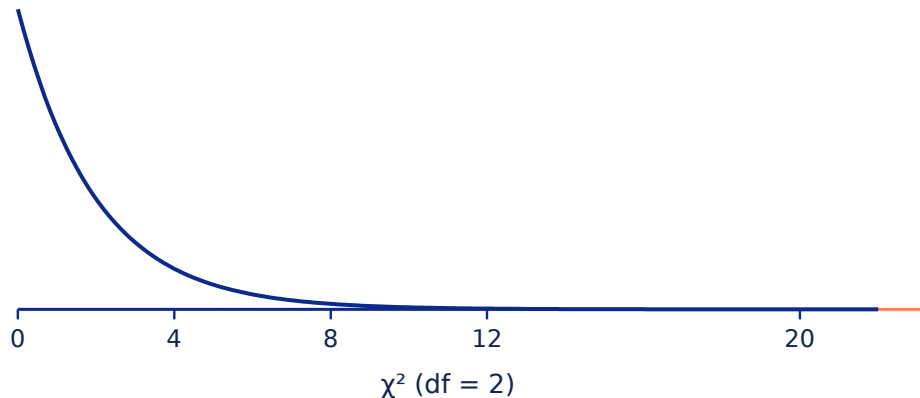
Expected count, statistic and degrees of freedom

$$\text{Expected} = \frac{((\text{row total})(\text{column total}))}{(\text{table total})}$$

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

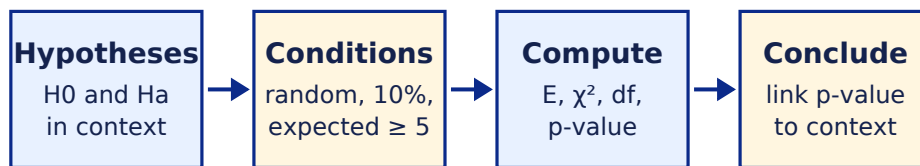
$$df = (rows - 1)(columns - 1)$$

For high school and cafe, the expected count is $170 \times 93/300 = 52.70$, compared with 73 observed. Summing the six cells gives $\chi^2 = 26.15$ with $df = (2 - 1)(3 - 1) = 2$. The largest contributions come from the cafe column (10.23 and 7.82).



Chi-square curve ($df = 2$); the p-value is the shaded right tail.

What are the conditions and the steps?



Chi-square test steps.

Method: chi-square test

1. State H_0 (no difference in distributions, or no association) and H_a (there is a difference or an association).
2. Check: random samples or random assignment, the 10% condition, and every expected count at least 5.
3. Compute the expected counts, χ^2 , degrees of freedom and the right-tail p-value.
4. If the $p\text{-value} \leq \alpha$, reject H_0 ; state the conclusion in context.

Worked example

H_0 : the distribution of preferred study spot is the same for middle and high school students; H_a : it is not. All expected counts exceed 5 (39.0 is the smallest). With $\chi^2 = 26.15$, $df = 2$, the p-value is $2.1e-06$, which is below 0.05. There is convincing evidence that the distribution of preferred study spots differs between middle and high school students.

The test shows *whether* distributions differ, not how. Compare the conditional percentages in the segmented bar graph or look at the cells with the largest contributions. A significant result from an observational study shows association, not causation; only random assignment in an experiment supports a causal conclusion.

Mathbot's tip

The chi-square statistic is always positive and the p-value is always the area to the right. A big statistic means the observed counts are far from what H_0 predicts.

Key points to remember

- Expected count = (row total)(column total)/table total.
- $\chi^2 = \sum (O - E)^2/E$; $df = (r - 1)(c - 1)$; p-value is the right tail.
- Conditions: random, 10%, and all expected counts at least 5.
- Homogeneity: several groups, one variable. Independence: one sample, two variables.
- Only a randomized experiment supports a cause-and-effect conclusion.

Chi-Square Tests for Two-Way Tables: frequently asked questions

What is the difference between a test for homogeneity and a test for independence?

Homogeneity compares one categorical variable across several separate populations or treatments, using random samples from each. Independence uses one random sample and two categorical variables. The expected counts, statistic and degrees of freedom are computed the same way.

How do I compute an expected count?

Multiply the row total by the column total and divide by the table total. This gives the count you would expect in that cell if the null hypothesis were true, with the same marginal totals as the observed table.

What are the conditions for a chi-square test?

The data come from random sampling or random assignment, the sample is less than 10% of the population, and every expected count is at least 5. Check expected counts, not observed counts.

How are the degrees of freedom found?

For a table with r rows and c columns, $df = (r - 1)(c - 1)$. A 2 by 2 table has $df = 1$, a 2 by 3 table has $df = 2$, and a 3 by 3 table has $df = 4$.

Does a small p-value prove that one variable causes the other?

No. A small p-value only shows convincing evidence of a difference or association. Causation can be concluded only if the data come from a randomized experiment. Observational studies may have confounding variables.