

AP Statistics · Chi-Square Tests for Homogeneity and Independence · Lesson

Common Core AP Statistics Unit 3 (2026-27) · mathovore.fr

This ap statistics chi square lesson explains chi-square tests for homogeneity and independence in one printable page, with definitions, a method and a worked example.

A chi-square test uses a two-way table of counts to decide whether several groups share one distribution of a categorical variable (homogeneity) or whether two variables are associated (independence). Compare observed and expected counts with $\chi^2 = \sum \frac{(O-E)^2}{E}$, use $df = (r-1)(c-1)$, and read the p-value.

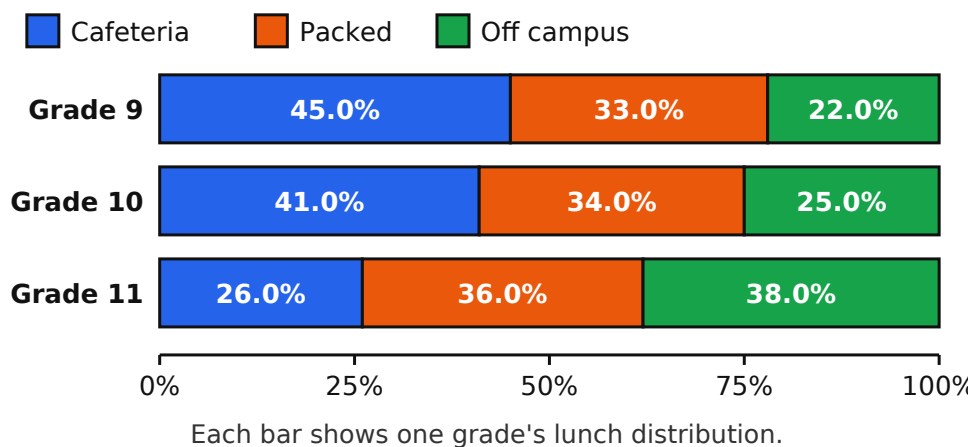
AP StatisticsCommon Core: AP Statistics Unit 3 (2026-27)

AP Statistics Chi Square Lesson: key facts

- Unit 3 of AP Statistics: chi-square tests for homogeneity and independence use a two-way table of counts.
- Expected count: $E = \frac{\text{row total} \times \text{column total}}{\text{grand total}}$.
- Test statistic: $\chi^2 = \sum \frac{(O-E)^2}{E}$ with $df = (r-1)(c-1)$.
- Conditions: random sampling or assignment, the 10% condition, and all expected counts at least 5.
- A conclusion names the variables in context and states whether the evidence is convincing.

What does a chi-square test do with a two-way table?

A two-way table sorts each person by two categorical variables, such as grade level and lunch choice. The chi-square test asks whether the differences between the groups are too large to blame on chance.



Two tests, one calculation

- **Homogeneity:** separate random samples from several populations (or random assignment to several treatments), one categorical response. H_0 : the response has the same distribution in every population.
- **Independence:** one random sample, two categorical variables recorded for each individual. H_0 : the two variables are independent in the population.

How do you find expected counts, the statistic and the degrees of freedom?

Formulas

$$E = \frac{\text{row total} \times \text{column total}}{\text{grand total}}$$

$$\chi^2 = \sum \frac{(O - E)^2}{E} \quad df = (r - 1)(c - 1)$$

Expected counts show what the table would look like if the null hypothesis were true. Large gaps between O and E make χ^2 large and the p-value small.

free	free	fixed
free	free	fixed
fixed	fixed	fixed

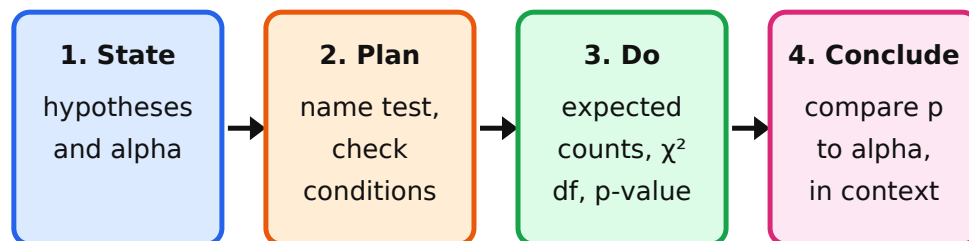
$$\text{free cells: } (3 - 1) \times (3 - 1) = 4$$

Row and column totals fix the last row and column, so a 3 by 3 table has 4 degrees of freedom.

What conditions must hold?

- **Random:** random samples (homogeneity), one random sample (independence), or random assignment (experiment).
- **10% condition:** when sampling without replacement, each sample is less than 10% of its population.
- **Large counts:** every expected count is at least 5.

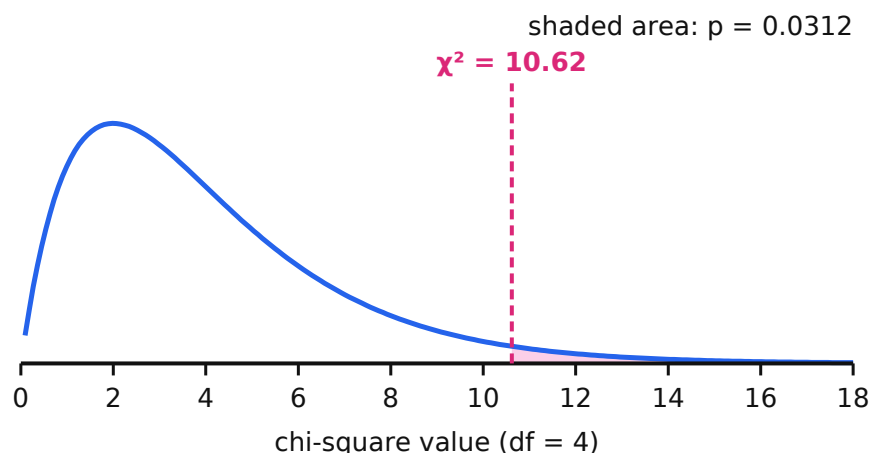
How do you carry out the test?



The four steps of every chi-square test.

Example. Random samples of 100 students from each of three grades (each with more than 1,000 students) chose a lunch type, as in the bars above. Is the distribution of lunch choice the same for all three grades?

1. **State.** H_0 : the distribution of lunch choice is the same in all three grades. H_a : it is not the same for at least one grade. Use $\alpha = 0.05$.
2. **Plan.** Chi-square test for homogeneity. The samples are random, each is under 10% of its grade, and the smallest expected count is 28.33.
3. **Do.** Totals are 112, 103 and 85, so each row expects 37.33, 34.33 and 28.33. Then $\chi^2 \approx 10.617$, $df = 4$, $p = 0.0312$.
4. **Conclude.** The p-value is less than 0.05, so reject the null hypothesis. There is convincing evidence that lunch choice is not distributed the same way in all three grades.



The p-value is the area to the right of the statistic.

Chi-Square Tests for Homogeneity and Independence: frequently asked questions

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

Homogeneity compares the distribution of one categorical variable across several separate samples or treatment groups. Independence uses a single sample and asks whether two categorical variables are associated. The expected counts, statistic and degrees of freedom are calculated the same way.

How do you find the expected count for a cell?

Multiply the row total by the column total and divide by the grand total. For a row total of 70, a column total of 60 and a grand total of 150, the expected count is $\frac{70 \times 60}{150} = 28$.

How many degrees of freedom does a two-way table have?

Use $df = (r-1)(c-1)$, where r is the number of rows and c is the number of columns. A table with 3 rows and 4 columns has $(3-1)(4-1) = 6$ degrees of freedom.

What does a small p-value mean in a chi-square test?

It means a statistic at least as large as the one observed would be unlikely if the null hypothesis were true. With a small p-value, reject the null hypothesis and conclude that the distributions differ or the variables are associated.

What if an expected count is less than 5?

The large counts condition fails, so the chi-square calculation may not give an accurate p-value. Collect more data, or combine categories only when the combined category still makes sense in context.

How to use this ap statistics chi square lesson

Use this ap statistics chi square lesson before the worksheet, then check what you learned with the quiz.

