Ap Statistics Chapter 14 Test

MASTERING THE AP STATISTICS CHAPTER 14 TEST: A COMPLETE GUIDE TO CHI-SQUARE INFERENCE

For many high school students, the AP Statistics curriculum represents a pivotal step into data-driven reasoning. By the time you reach Chapter 14, you have already navigated the worlds of probability, sampling distributions, and confidence intervals for means and proportions. Now, you are about to tackle one of the most versatile and frequently tested topics on the AP exam: inference for categorical data using chi-square tests. The **AP Statistics Chapter 14 test** typically covers three distinct procedures: the chi-square goodness-of-fit test, the chi-square test for homogeneity, and the chi-square test for independence. Understanding these methods is not just about memorizing formulas—it's about learning how to answer questions about how well observed data fit a theoretical model or whether two categorical variables are related. This article will walk you through everything you need to know to ace your Chapter 14 test, from conditions to calculations to common pitfalls.

WHY CHAPTER 14 MATTERS ON THE AP EXAM

The AP Statistics exam consistently includes free-response questions and multiple-choice items that test chi-square procedures. According to the College Board’s course description, inference for categorical data accounts for roughly 12–15% of the exam. That means mastering the **AP Statistics Chapter 14 test** not only boosts your grade in class but also gives you a solid foundation for the May exam. Many students find chi-square tests easier to conceptualize than t-tests because they rely on comparing observed counts to expected counts, a very intuitive idea. However, the devil is in the details: knowing when to use each test, checking conditions, and correctly interpreting the test statistic and p-value are essential skills.

What You Will Learn in This Chapter

Chapter 14 usually begins by introducing the chi-square distribution and its properties. Unlike the normal or t-distributions, the chi-square distribution is skewed to the right and takes only positive values. Its shape depends on the degrees of freedom. You will then apply this distribution to three testing scenarios:

- **Chi-Square Goodness-of-Fit Test**: Is there a significant difference between the observed distribution of a single categorical variable and an expected distribution?
- **Chi-Square Test for Homogeneity**: Do two or more populations have the same distribution of a categorical variable?
- **Chi-Square Test for Independence**: Are two categorical variables independent within a single population?

Each test uses a similar chi-square statistic, but the hypotheses, expected counts, and degrees of freedom differ. The **AP Statistics Chapter 14 test** will likely ask you to identify the correct procedure for a given context.

SETTING THE STAGE: THE CHI-SQUARE DISTRIBUTION

Before diving into each test, it is helpful to review the chi-square distribution. The test statistic is calculated as:

$\chi^2 = \Sigma$ (Observed – Expected)^2 / Expected

where the sum is taken over all cells in the table. When the null hypothesis is true and conditions are met, this statistic follows a chi-square distribution with degrees of freedom equal to (number of categories – 1) for a goodness-of-fit test, or (rows – 1)(columns – 1) for tests of homogeneity and independence. The shape of the distribution becomes more symmetric as degrees of freedom increase. On the **AP Statistics Chapter 14 test**, you may be asked to estimate p-values using a chi-square table or to interpret them from technology output.

Key Conditions for Chi-Square Tests

All chi-square tests require three main conditions. First, the data must come from a random sample or a randomized experiment. Second, the sample size must be large enough: the expected count for each cell should be at least 5 (some textbooks say at least 1, but the AP exam typically expects at least 5 for each expected count). Third, individual observations must be independent. For the goodness-of-fit test, you also assume the sample is of a single categorical variable from one population. For the test of homogeneity, you assume independent random samples from each population. For the test of independence, you assume one random sample from a single population, with two categorical variables measured on each individual. Remember these distinctions—they are classic multipart questions on the **AP Statistics Chapter 14 test**.

BREAKING DOWN THE CHI-SQUARE GOODNESS-OF-FIT TEST

The goodness-of-fit test answers a simple question: does the distribution of a categorical variable in a sample match a hypothesized distribution? For example, a candy company claims that its bag contains 30% red, 20% orange, 20% yellow, 10% green, 10% blue, and 10% brown. You buy a bag and count the colors. Is the observed distribution consistent with the claimed proportions? This is a classic goodness-of-fit problem that might appear on your **AP Statistics Chapter 14 test**.

Hypotheses:

H_0: The distribution of [variable] is the same as the claimed distribution.

H_a: The distribution of [variable] is different from the claimed distribution.

Procedure:

1. State the null and alternative hypotheses.
2. Check conditions: random sample, independence, large expected counts (all ≥5).
3. Calculate expected counts: multiply sample size by each claimed proportion.
4. Compute the chi-square statistic.
5. Find the p-value using the chi-square distribution with degrees of freedom = number of categories – 1.
6. Compare p-value to significance level (usually α=0.05) and draw a conclusion in context.

Be careful: the alternative hypothesis is always "not equal to" or "different from" the claimed distribution. There is no one-sided chi-square test—it is inherently two-sided because the test statistic is always positive. Common mistakes on the **AP Statistics Chapter 14 test** include forgetting to check expected counts or misstating the degrees of freedom. Also, when writing the conclusion, say something like: "Because the p-value (0.03) is less than $\alpha=0.05$, we reject H_0. There is convincing evidence that the actual distribution of candy colors differs from the company's claim."

CHI-SQUARE TEST FOR HOMOGENEITY

The test for homogeneity compares the distribution of a categorical variable across two or more populations or treatment groups. For example, do students in different school districts have the same proportions of grade levels (freshman, sophomore, junior, senior)?

Hypotheses:
H_0: The distribution of the categorical variable is the same for each population.
H_a: The distribution of the categorical variable is not the same for at least one population.

Expected counts are calculated under the assumption that the null hypothesis is true: (row total × column total) / grand total. This is the same formula used for the test of independence, but the sampling design differs. For homogeneity, you have separate random samples from each population; for independence, you have one random sample. The **AP Statistics Chapter 14 test** often asks you to recognize which scenario applies. The degrees of freedom are (number of rows – 1)(number of columns – 1), where rows correspond to the categories of the variable and columns correspond to the different populations (or vice versa). Many students confuse rows and columns, but the formula works symmetrically.

A practical tip: when reading a problem, look for language like "random samples of ... from three different regions" to identify a homogeneity test. If the problem says "one random sample of ... and they were asked about two different characteristics," that is a test of independence.

CHI-SQUARE TEST FOR INDEPENDENCE

The test for independence examines whether two categorical variables are related in a single population. For instance, is handedness (left/right) independent of gender (male/female)? You take one random sample of individuals, record both variables, and create a two-way table.

Hypotheses:

H_0: Variable A and Variable B are independent.

H_a: Variable A and Variable B are not independent (they are associated).

The mechanics are identical to the homogeneity test—same expected counts formula, same chi-square statistic, same degrees of freedom. The only difference is the hypothesis wording and the sampling design. On the **AP Statistics Chapter 14 test**, you may be asked to explain the difference, so be prepared to articulate that homogeneity tests compare distributions across groups, while independence tests assess association within one group.

One classic mistake: students sometimes think that a significant chi-square test tells you the strength or direction of the relationship. The chi-square test only tells you if an association exists, not how strong it is. For strength, you would examine the difference between observed and expected counts or calculate a measure like Cramér's V (which is not on the AP syllabus but can be mentioned). Also, the test does not indicate which cells contribute most to the significance—you can look at the individual components (observed-expected)^2/expected to see which are large.

HOW TO PREPARE FOR YOUR AP STATISTICS CHAPTER 14 TEST

Now that you know the theory, let's focus on test-day strategies. The **AP Statistics Chapter 14 test** often includes multiple-choice questions that ask you to identify the correct test, interpret computer output, or check conditions. Free-response questions require a full four-step process: state, plan, do, conclude. Practice writing each step clearly. Use the acronym PHA (Parameters, Hypotheses, Alpha) for the state step, although many teachers use a different structure. Emphasize context—never just say "reject H_0" without relating it to the real-world situation.

Common Pitfalls to Avoid

- **Using observed counts instead of expected counts** in the denominator: Remember that the denominator uses expected counts, not observed counts.
- **Forgetting to check the large expected counts condition**: If any expected count is less than 5, the test may not be valid. On the AP exam, they usually design problems where this condition is met, but you still must state it.
- **Misidentifying degrees of freedom**: Goodness-of-fit df = k-1. Homogeneity/Independence df = (r-1)(c-1). A two-way table with 3 rows and 4 columns gives df = 6.
- **Confusing chi-square test for homogeneity with chi-square test for independence**: Look at the sampling method—separate samples vs one sample. If in doubt, check the wording.
- **Using chi-square test when the variable is quantitative**: Chi-square is for categorical variables only. If you have numerical data, you need a t-test or ANOVA.

Sample Problem

for Practice

Consider the following scenario typical of the **AP Statistics Chapter 14 test**: A researcher wonders if political affiliation (Democrat, Republican, Independent) is related to opinion on a new environmental policy (Support, Oppose, Undecided). She surveys a random sample of 200 voters and obtains the following two-way table (fictitious):

(Table omitted for brevity, but imagine observed counts and expected counts.)

Question: Is there convincing evidence of an association between political affiliation and opinion? Use a significance level of 0.05.

Solution: This is a chi-square test for independence because one sample of 200 voters was taken and two categorical variables were recorded. Hypotheses: H_0: Political affiliation and opinion are independent. H_a: They are not independent. Conditions: random sample (given), all expected counts ≥5 (check), independence (sample <10% of population, assume). Calculate χ^2, find df = (3-1)(3-1)=4, find p-value. If p-value < 0.05, reject H_0 and conclude that there is an association. Always interpret p-value in context.

Using

Technology: TI-84 and Output

On the **AP Statistics Chapter 14 test**, you will likely be allowed a graphing calculator. The TI-84 can run chi-square tests quickly. For a two-way table, enter data into matrix A, then run the χ^2-Test. The calculator will output the test statistic, degrees of freedom, and p-value. For goodness-of-fit, you need to enter observed and expected lists. Some free-response questions will ask you to interpret calculator output—for example, "A χ^2 value of 8.12 with df=3 gives a p-value of 0.045