

AP Statistics - Significance Tests for a Proportion

- Lesson

Common Core AP STAT 6.4, AP STAT 6.5, AP STAT 6.6, AP STAT 6.7, HSS-IC.A.2, HSS-IC.B.5 · maths-pdf.fr/en/

A **significance test for a proportion** compares $H_0: p = p_0$ with an alternative using $z = (\hat{p} - p_0) / \sqrt{p_0(1-p_0)/n}$. The **p-value** is the chance, if H_0 is true, of a result this extreme. If it is below α , reject H_0 . A **Type I error** rejects a true H_0 ; a **Type II error** fails to reject a false H_0 .

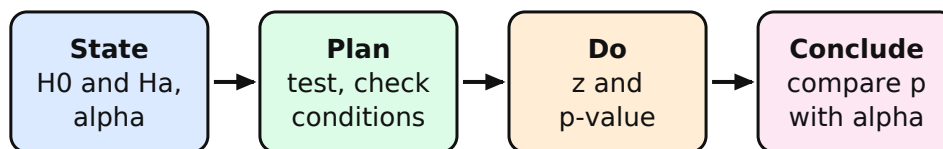
AP Statistics · Common Core State Standards: **AP STAT 6.4, AP STAT 6.5, AP STAT 6.6, AP STAT 6.7, HSS-IC.A.2, HSS-IC.B.5**

Key facts

- H_0 states no effect ($p = p_0$); H_a is $p > p_0$, $p < p_0$, or p not equal to p_0 .
- $z = (\hat{p} - p_0) / \sqrt{p_0(1 - p_0)/n}$; check random, 10%, and large counts with p_0 .
- p-value: probability of a result at least as extreme as observed, assuming H_0 is true.
- p-value $< \alpha$: reject H_0 (convincing evidence); otherwise fail to reject; never 'accept H_0 '.
- Type I error: reject true H_0 (probability α). Type II: fail to reject false H_0 ; power = $1 - P(\text{Type II})$.

What is a significance test?

A **significance test** uses sample data to decide between two competing claims about a population parameter. The **null hypothesis** H_0 is the claim of no effect or no change (such as $p = 0.45$). The **alternative hypothesis** H_a is what we are trying to find evidence for ($p > 0.45$, $p < 0.45$, or $p \neq 0.45$).



The four-step process for a test.

Example: a basketball player usually makes 45% of her three-point shots. After training, she makes 58 of 100. Is there evidence that her true percentage has increased? $H_0: p = 0.45$ and $H_a: p > 0.45$, where p is the true proportion of three-pointers she makes after training.

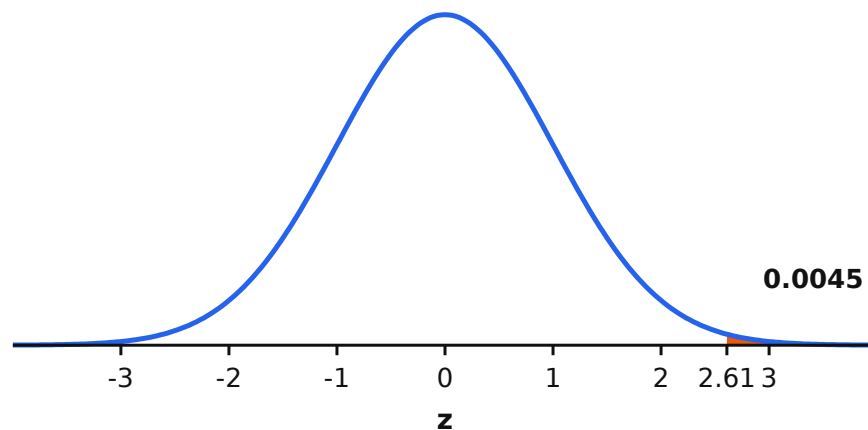
How do you compute the test statistic and p-value?

For a one-sample z test for a proportion:

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

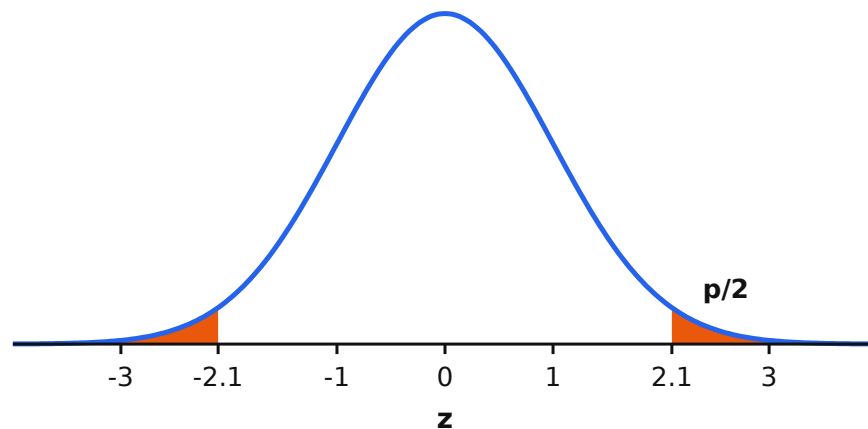
Conditions: random sample, 10% condition, and large counts using p_0 : np_0 and $n(1-p_0)$ both at least 10. Here $np_0 = 45$ and $n(1-p_0) = 55$.

$$\hat{p} = \frac{58}{100} = 0.58, SE = \sqrt{\frac{0.45(0.55)}{100}} = 0.0497, z = \frac{0.58 - 0.45}{0.0497} = 2.61.$$



Right-tailed p-value for $z = 2.61$: 0.0045.

The **p-value** is the probability, assuming H_0 is true, of getting a sample result at least as extreme as the one observed. Here $P(Z > 2.61) = 0.0045$. For a two-sided test, double the tail area.



A two-sided p-value adds both tails.

How do you reach a conclusion?

Compare the p-value with the significance level α (often 0.05). If $p\text{-value} < \alpha$, reject H_0 : there is convincing evidence for H_a . Otherwise fail to reject H_0 : there is not convincing evidence. Here $0.0045 > 0.05$, so we fail to reject H_0 . The data do not give convincing evidence that her true three-point percentage is above 0.45. We never say that we "accept" H_0 .

What are Type I and Type II errors?

	H_0 true	H_0 false
Reject H_0	Type I error	Correct (power)
Fail to reject H_0	Correct	Type II error

Type I and Type II errors.

A **Type I error** rejects a true H_0 ; its probability is α . A **Type II error** fails to reject a false H_0 . The **power** of a test is the probability of rejecting a false H_0 ; it increases with a larger sample size, a larger α , or a parameter farther from $\mathcal{P}_0$.

Questions about significance tests for a proportion

What does a p-value of 0.03 mean?

If the null hypothesis were true, there would be only a 3% chance of getting a sample result as extreme as, or more extreme than, the one we observed. It is not the probability that H_0 is true.

Why do we say 'fail to reject' instead of 'accept' H_0 ?

A test can only show that the data are inconsistent with H_0 . A large p-value means the sample is plausible under H_0 , but other values of p might also be plausible.

How do I choose between one-sided and two-sided alternatives?

Use the direction in the question before collecting data. If the claim is 'greater than' or 'less than,' use a one-sided test; if it asks whether the parameter is 'different' from a value, use a two-sided test.

What is the relationship between alpha and a Type I error?

The significance level α is the probability of making a Type I error when H_0 is true. Choosing $\alpha = 0.05$ means a 5% chance of rejecting a true null.

How can I increase the power of a test?

Use a larger sample size, a larger significance level, or study a situation where the true parameter is farther from the null value. Lower variability also raises power.