Inference for Proportions

AP Statistics ▪ Unit 6

AP Statistics

What we will cover

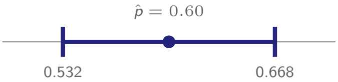
1. Confidence interval for p
2. Interpreting an interval
3. Setting up a test
4. p-values and conclusions
5. Type I and II errors
6. Comparing two proportions

1 Interval

Confidence interval for a proportion

A **confidence interval** 置信区间 is estimate $\pm$ margin:

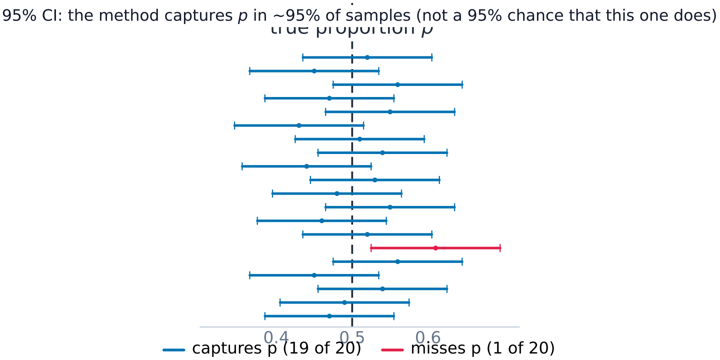
$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$



- **margin of error** 误差幅度 $= z^* \cdot$**standard error** 标准误
- for 95%: **critical value** 临界值 $z^* = 1.96$

> Check: random, $\leq 10\%$ of population, and large counts.

Over many samples



Over many samples, about 95% of 95% confidence intervals capture the true proportion

Worked example – choosing the sample size

How many people must be surveyed for a 95% interval with margin of error at most 0.03?

- $z^* \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}} \leq m$, solved for n
- use $\hat{p} = 0.5$ (the worst case, which maximises the product) and $z^* = 1.96$
- $n \geq \left(\dfrac{1.96}{0.03}\right)^2 (0.25) = 1067.1$
- round up: $n = 1068$

Two things are always marked: using $\hat{p} = 0.5$ when no estimate is given, and rounding up – rounding down would leave the margin too wide.

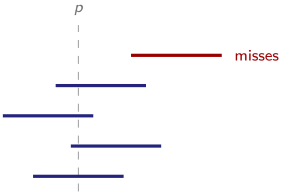
2 Interpret

Interpreting an interval

Two layers of meaning:

- *interval*: "95% confident the true p is in $(0.532, 0.668)$"
- *level*: "95% of such intervals capture p" – confidence is in the method

> A value outside the interval is not plausible. Bigger $n \Rightarrow$ narrower; higher confidence $\Rightarrow$ wider.



Most intervals catch p; some miss.

Judging a claim from an interval

the test	if the claimed value lies inside the interval, the data are consistent with it
if outside	the data give evidence against the claim
what it does not say	an interval containing the value does not *prove* it – it fails to rule it out
the correct interpretation	"we are 95% confident that the true proportion lies between . . . and . . ."
what 95% refers to	the method: 95% of intervals built this way capture the true value. It is not a probability about this one interval

> Never say "there is a 95% chance the parameter is in this interval". The parameter is fixed; it is the interval that varies from sample to sample.

3 Test setup

Setting up a test

A **significance test** 显著性检验 weighs evidence against a claim.

- **null** 原假设 $H_0 : p = p_0$ ("no effect")
- **alternative** 备择假设 $H_a : p >, <, \text{or} \neq p_0$

The **test statistic** 检验统计量 (one-sample z-test), using p_0:

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

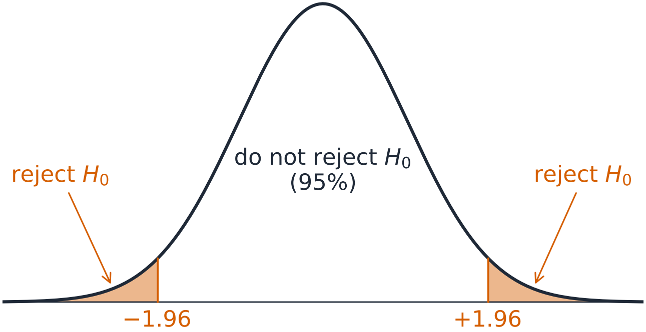
> Check conditions with the assumed p_0, not $\hat{p}$.

A coin

58 heads in 100: $\hat{p} = 0.58$, $H_0 : p = 0.5$, $H_a : p > 0.5$.

$$z = \frac{0.58 - 0.50}{0.05} = 1.6$$

Setting Up a Test for a Proportion



> A two-tailed 5% test rejects the null hypothesis in the shaded tails

Worked example – a one-proportion z-test

A company claims 90% satisfaction. A sample of 100 finds 84 satisfied. Test $H_0 : p = 0.90$ against $H_a : p < 0.90$.

- $\hat{p} = 0.84$; use p_0 in the standard error, not $\hat{p}$
- $\text{SE} = \sqrt{\dfrac{0.90(0.10)}{100}} = 0.03$
- $z = \dfrac{0.84 - 0.90}{0.03} = -2.0$, so $p \approx 0.023$
- $0.023 < 0.05$: reject H_0

A *test* uses the claimed p_0 in the standard error; an *interval* uses $\hat{p}$. Mixing them is the commonest error here.

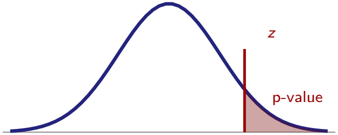
4 p-value

p-values and conclusions

The **p-value** p 值: *if H_0 were true*, the chance of a result at least this **extreme** 极端 – a tail area from z.

- p $\leq \alpha$: **reject** 拒绝 H_0
- p $> \alpha$: **fail to reject** 不拒绝 H_0

> Failing to reject is not proof of H_0 – just too little evidence.



5 Errors

Type I and Type II errors

A test on chance data can err two ways:

- **Type I** 第一类错误: reject a *true* H_0 (false alarm) – prob. α
- **Type II** 第二类错误: miss a *false* H_0 – prob. β

> **Power** 功效 $= 1 - \beta$ rises with larger n, a bigger true effect, or larger α.

	H_0 true	H_0 false
reject	Type I (α)	correct (power)
fail	correct	Type II (β)

The two errors, and writing the conclusion

the significance level 显著性水平	α – fixed before the data are seen; usually 0.05
the alternative hypothesis 备择假设	H_a – one-sided or two-sided, chosen from the question, never from the data
a Type I error 第一类错误	rejecting a true H_0 – a false alarm. Its probability is α
a Type II error	failing to reject a false H_0 – a missed detection. Its probability is β
power	$1-\beta$: the chance of detecting a real effect. Raised by a larger sample or a larger α
two proportions	use the **combined (pooled)** 合并 estimate of p in the standard error for the *test*, but not for the interval

Always write the conclusion in context, describing each error and its consequence in the problem's own terms – that sentence is where the marks are.

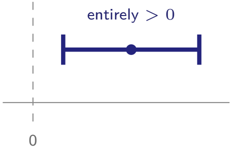
6 Two proportions

Comparing two proportions

A **two-sample z-interval** 双样本 z 区间:

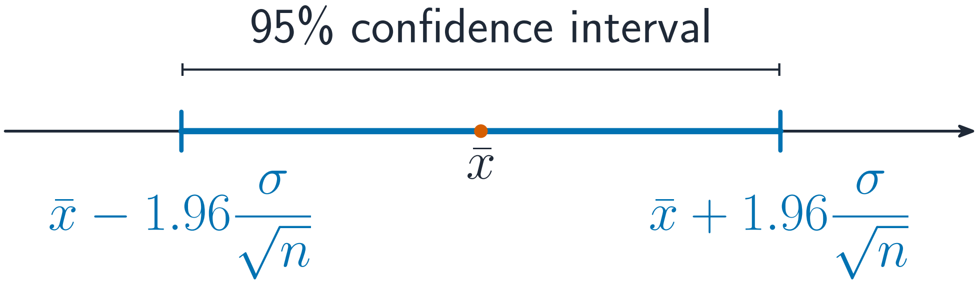
$$(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

> If the interval excludes 0, the groups truly differ. The test pools: $\hat{p}_c = \frac{x_1 + x_2}{n_1 + n_2}$.



Excludes 0 $\Rightarrow p_1 > p_2$.

Confidence Interval for a Proportion, continued



> A 95% confidence interval reaches 1.96 standard errors each side of the estimate

Why Be Normal?

Because a sample proportion $\hat{p}$ is approximately **normally** distributed (when the conditions hold).

> This is what makes **inference 推断** – drawing conclusions about a population from a sample – possible.

Exam tips

- State the conditions (random, 10%, large counts np, $nq \geq 10$) before any proportion inference.
- A **confidence interval** $=$ estimate $\pm$ margin of error; "95% confident" refers to the **method's** long-run capture rate.
- For a **test**, write H_0 and H_a, compute the test statistic, find the p-value, and compare to α.
- A small p-value is evidence **against** H_0; failing to reject does **not** prove H_0.
- Larger samples shrink the margin of error; a higher confidence level widens it.

7 Recap

Unit 6 – key takeaways

1 Interval

$\hat{p} \pm z^* \cdot \text{SE}$; 95% of intervals catch p

2 Test

$H_0 : p = p_0$; z uses p_0; small p rejects

3 Errors

Type I (α) false alarm; Type II (β); power $1 - \beta$

4 Two groups

interval excluding 0 $\Rightarrow$ real difference; pool for the test

Check yourself

1. What is z^* for a 95% confidence interval?
2. A small p-value is evidence for or against H_0?
3. Rejecting a true H_0 is which type of error?
4. What value must a difference interval exclude to show a real gap?

Answers **1.** 1.96 **2.** against H_0 **3.** Type I **4.** zero