10.2: Hypothesis Tests for a Population Proportion

The Logic of Hypothesis Testing

Consider this:
A researcher obtains a random sample of 1000 people and finds that 534 are in favor of the banning cell phone use while driving, so = 534/1000.

- Does this suggest that more than 50% of people favor the policy?
- Or is it possible that the true proportion of registered voters who favor the policy is some proportion less than 0.5 and we just happened to survey a majority in favor of the policy?
- In other words, would it be unusual to obtain a sample proportion of 0.534 or higher from a population whose proportion is 0.5?
- What is convincing, or statistically significant, evidence?

When observed results are unlikely under the assumption that the null hypothesis is true, we say the result is **statistically significant**.
When results are found to be statistically significant, we reject the null hypothesis.

Let's Check the Null Hypothesis

H_0="The true population proportion of registered voters who favor the policy is 50%"

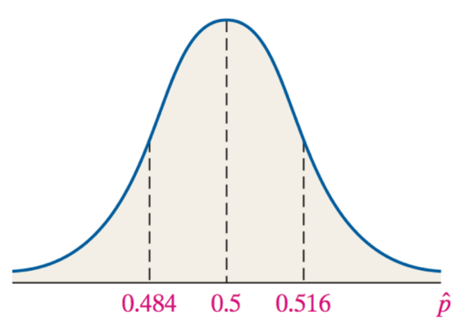
- *Check Sample size: is* $np(1-p) \geq 10$*?*

  Can we use the Normal Model?

- *What is the Mean of the Distribution of* $\hat{p}$*?* $\mu_{\hat{p}} = p$

- *Find the Standard Deviation*

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

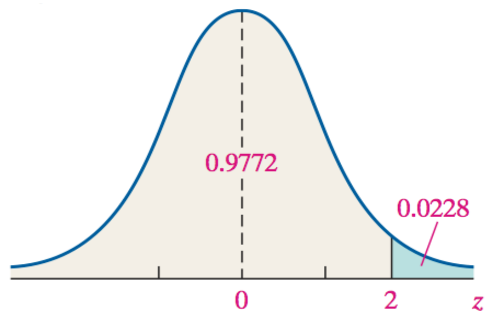


- The sample would be considered **statistically significant** (or sufficient) if the sample proportion is more than 2 standard deviations away from the population proportion of 0.5. Is it?

- Conclusion: Should we reject the null hypothesis?

Why does it make sense to reject the H_0?

If the null hypothesis were true (i.e. the population proportion is 0.5), 97.2% of all sample proportions would be less than .5+2(0.016)=.532, and only
2.28% of the sample populations would be more than .532



Hypothesis Testing using the Classical Approach

- **We Reject the Null Hypothesis** If the sample proportion is greater than 2 standard deviations from the proportion stated in the null hypothesis.
- **We *Do Not Reject* the Null Hypothesis** if the sample proportion is within 2 standard deviation from the proportion stated in the null hypothesis.

The Logic of the P-Value Approach

A second criterion we may use for testing hypotheses is to answer the question,

"How likely it is to obtain a sample proportion of 0.534 or higher from a population whose proportion is 0.5?"

If a sample proportion of 0.534 or higher is unlikely (or unusual), we have evidence against the statement in the null hypothesis.
Otherwise, we do not have sufficient evidence against the statement in the null hypothesis.

Test statistic for Proportions:

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

Try it:

Find the z-value for the sample proportion of $\hat{p} = 0.534$ if $p_0 = 0.5$

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

$P(\hat{p} > 0.534) = P(z \geq 2.15) =$

This number is called our "P-value".

If the P-value is significantly small (usually < 5%) we reject the Null Hypothesis.

Should we reject H_0="Population Proportion =.5"?

Hypothesis Testing Using the *P*-value Approach:

If the probabilities $P(p > \hat{p})$ or $P(p < \hat{p})$ for of getting a sample proportion p is small under the assumption that the statement in the null hypothesis is true, reject the null hypothesis.

Test Hypothesis about a Population Proportion

Remember:

1. The best point estimate of p, the proportion of the population with a certain characteristic, is given by
$$\hat{p} = \frac{x}{n}$$
where x is the number of individuals in the sample with the specified characteristic and n is the sample size.
2. The sampling distribution of $\hat{p}$ is approximately normal, with mean $\mu_{\hat{p}} = p$ and standard deviation
$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$
provided that the following requirements are satisfied:
    a. The sample is a simple random sample.
    b. $np(1-p) \geq 10$.
    c. The sampled values are independent of each other.

Steps for Testing Hypothesis Regarding Population Proprtion, *p*

Step 1: *Determine the null and alternative hypotheses.*
The hypotheses can be structured in one of three ways:

Two-Tailed	Left-Tailed	Right-Tailed
H_0: $p = p_0$	H_0: $p = p_0$	H_0: $p = p_0$
H_1: $p \neq p_0$	H_1: $p < p_0$	H_1: $p > p_0$

Note: p_0 is the assumed value of the population proportion.

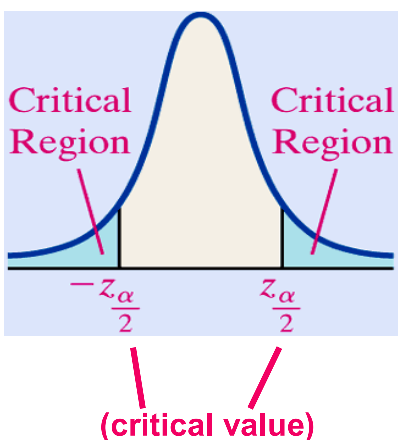
Step 2: *Select a level of significance, α, based on the seriousness of making a Type I error.*
Step 3: *Compute the* ***test statistic***

$$z_0 = \frac{\hat{p} - p_0}{\sigma_{\hat{p}}} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

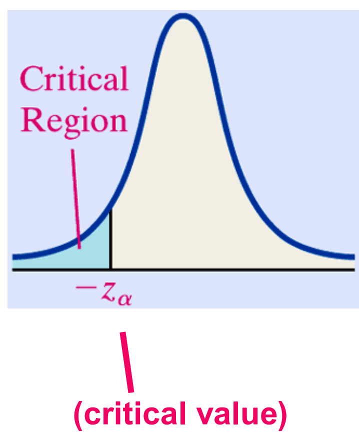
Note: *We use p_0 in computing the standard error rather than $\hat{p}$. This is because, when we test a hypothesis, the null hypothesis is always assumed true.*

Classical Approach: Now Determine the Critical value using z-table

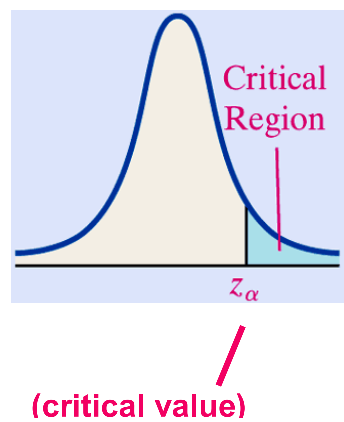
Two-Tailed



Left-Tailed



Right Tailed



Classical Approach Step 4: *Compare the critical value with the test statistic:*

Two-Tailed	Left-Tailed	Right-Tailed
If $z_0 < -z_{\frac{\alpha}{2}}$ or $z_0 > z_{\frac{\alpha}{2}}$, reject the null hypothesis.	If $z_0 < -z_\alpha$, reject the null hypothesis.	If $z_0 > z_\alpha$, reject the null hypothesis.

P-Value Approach:

Technology Step 3: [STAT]-[Test]-[1PropZtest]
Step 4: If the *P*-value < *α*, reject the null hypothesis.

Step 5: State the conclusion.

Example

In 1997, 46% of Americans said they did not trust the media "when it comes to reporting the news fully, accurately and fairly". In a 2007 poll of 1010 adults nationwide, 525 stated they did not trust the media.
At the $\alpha = 0.05$ level of significance, is there evidence to support the claim that the percentage of Americans that do not trust the media to report fully and accurately has increased since 1997?
Source: Gallup Poll

Step 0: Verify the requirements for hypothesis test
- Is it a simple random sample?
- Is $np_0(1-p_0) > 10$?
- Is the data independent? Since it's less than 5% of the population, assume it is.

Step 1: State Hypotheses

H_0:

H_1:

Step 2: Level of significance is $\alpha =$_______

Step 3: Find Sample Proportion:

$\hat{p} =$

Find test statistic:

$$z_0 = \frac{\hat{p}-p_0}{\sigma_{\hat{p}}} = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}n$$

TI-84: [STAT]-[Test]-[1PropZtest]

Classical Approach

Step 4: Find Critical Value

$z_{0.05} =$ *(This is the cutoff value for z_0)*

Step 5: Do you reject H_0?

Is $z_0 > z_{0.05}$?

P-value Approach

Step 4: Find Area to the right of $z_0 = 3.83$

That is, Find the P-value: $P(Z > 3.83) =$ /

(This is the cutoff value for z_0)

Step 5: Is $P(Z > z_0) < \alpha$?

Do you reject the null hypothesis?

Step 6: Conclusion

There (is / is not) *sufficient evidence at $\alpha = 0.05$ level of significance to conclude that the percentage of Americans that do not trust the media to report fully and accurately has increased since 1997.*