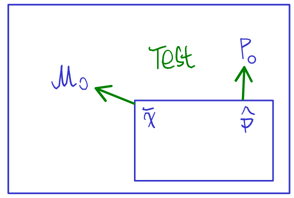
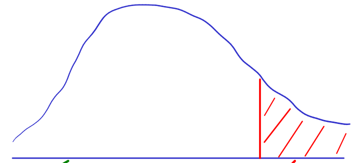
Hypothesis Testing - 1 Mean 1 Proportion

David Armstrong

UCI





Hypothesis Tests

- ASSESS the **evidence** provided by the data and to decide between two competing claims (hypotheses) about the population parameter

- The **critical value** gives us the first value for which we would **reject** H_o.

  Null Hypothesis

  If the **test statistic** is further in the tails compared to the **critical value**, we **reject** H_o.

  If the **test statistic** is not further in the tails compared to the **critical value**, we **fail to reject** H_o.

- The **significance level** gives the area of the **rejection region**

  Common α levels of rejection are $\alpha = 0.10, 0.05, 0.01$.

  We use the significance level to find the critical value.

  If the p-value is low, we reject H_o

  Set before we start our experiment.

  This is the threshold that we can compare the p-value to.

  If the p-value $\leq \alpha$ we reject the null hypothesis.

- The **p-value** for the observed data measures the **weight of evidence** against the null hypothesis.

  The **p-value** is the probability of seeing a value as extreme or more extreme than the one in our sample, **assuming the null hypothesis is true**.

  The smaller the **p-value**, the **more evidence** we have against H_o.

  The larger the **p-value**, the **less evidence** we have against H_o.

- In order to test the hypothesis for a population parameter, we use two things

  We use the sample statistic as the best estimate of the population parameter.

  We also need a measure of spread.

Stating Hypotheses

- **Null Hypothesis**

  This states what is generally believed to be the true population parameter.

  $H_o : parameter = hypothesized\ value$

- **Alternative Hypothesis**

  The "research hypothesis". What we are trying to prove.

  Less than test: $H_o : parameter < hypothesized\ value$

  Greater than test: $H_o : parameter > hypothesized\ value$ One tailed

  Different than test: $H_o : parameter \neq hypothesized\ value$

  2 sided

Draw a Picture

Use the sampling distribution to draw a picture of your hypothesis.

Find the critical value

qnorm

Use the significance level to find the critical value (t_{crit} or z_{crit}).

qt

Calculate the test statistic

Use the sampling distribution to calculate the test statistic (t_{TS} or z_{TS}).

Make a statistical decision and state your conclusion in context of the problem

- Since the test statistic IS/IS NOT more extreme than the critical value, we REJECT/FAIL TO REJECT the null hypothesis at the α = **Value** level. There IS/IS NOT statistically significant evidence that **Context**

- Since the p-value = **Value** $\leq$ vs $>$ **Value** = α we REJECT/FAIL TO REJECT the null hypothesis. There IS/IS NOT statistically significant evidence that **Context**

- The most important part is to use complete sentences to convey the context of the problem as well as the conclusion.

Hypothesis Alternative

$H_A \neq$

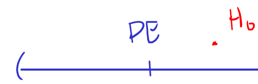
Confidence Intervals and Two Sided Hypothesis Tests

- Another way to interpret a Confidence interval

  The set of all values in which a 2 sided hypothesis test would fail to reject H_o

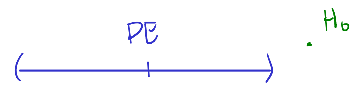
- Create a Confidence Interval

  If the null value is INSIDE the confidence interval

  Fail to Reject the null hypothesis

  If the null value is OUTSIDE the confidence interval

  Reject the null hypothesis

- Conduct a Hypothesis Test

  If we Fail to Reject the null hypothesis

  The null value is INSIDE the confidence interval

  If we Reject the null hypothesis

  The null value is OUTSIDE the confidence interval

- Since the claimed value IS NOT/IS captured within the interval we REJECT/FAIL TO REJECT the claim at the $\alpha =$ **Value** level.

Parameter Description	Test Statistic	H_o:
1 population proportion	$\boxed{Z = \dfrac{\hat{p} - p_o}{\sqrt{\dfrac{p_o q_o}{n}}}}$	$\boxed{p = p_o}$
1 population mean (σ unknown)	$\boxed{t = \dfrac{\bar{X} - \mu_o}{\dfrac{s}{\sqrt{n}}} \text{ where } df = n - 1}$	$\boxed{\mu = \mu_o}$

Assumptions for Hypothesis Tests

1 population proportion

- $n\hat{p} \geq 10$ and $n\hat{q} \geq 10$
- Sample size is less than 10% of the population size; if we are sampling is w/out replacement
- The sample can be regarded as a simple random sample from the population of interest.
- The data values are assumed to be independent of each other.

1 population mean

- We need to have a large enough sample size ($n \geq 30$). For $n < 30$ with extreme skewness or outliers, you cannot use this method.
- Sample size is less than 10% of the population size; if we are sampling is w/out replacement
- The sample can be regarded as a simple random sample from the population of interest.
- The data values are assumed to be independent of each other.

$H_0 : P = P_0$

$H_A : P < P_0$

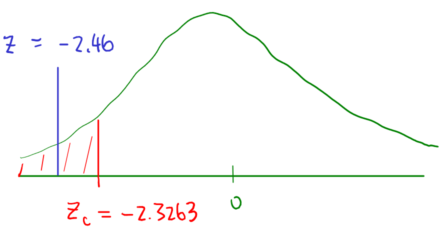
Ex: What is your decision?

Left Tailed Test:
Test Statistic: $z = -2.46$
$\alpha = 0.01$

$z_c = \text{qnorm}(0.01, 0, 1) = -2.3263$

Are you in the rejection region ?

Yes, since the test statistic is more extreme than the critical value we reject the null hypothesis at the $\alpha = 0.01$ level



Pvalue $= P(Z < -2.46)$

$= \text{pnorm}(-2.46, 0, 1)$

$= 0.0069$

Since Pvalue $= 0.069 \leq 0.01 = \alpha$

we reject the null hypothesis

$H_0 : P = P_0$

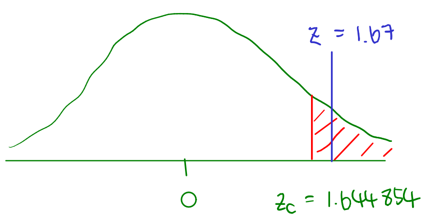
$H_A : P > P_0$

Right Tailed Test

Test Statistic: $z = 1.67$
$\alpha = 0.05$

$z_c = \text{qnorm}(1 - 0.05 = 0.95, 0, 1) = 1.644854$

Since the test statistic is more extreme than the critical value, we reject the null hypothesis at the $\alpha = 0.05$ level



Pvalue $= P(Z > 1.67) = 1 - P(Z < 1.67)$

$= 1 - \text{pnorm}(1.67, 0, 1) = 0.0475$

Since Pvalue $= 0.0475 \leq 0.05 = \alpha$

we reject the null hypothesis.

$H_0 = P = P_0$

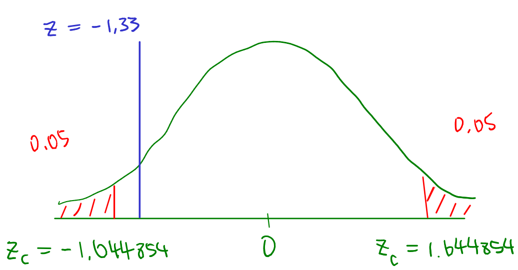
$H_A = P \neq P_0$

Two-Tailed Test

Test Statistic: $z = -1.33$
$\alpha = 0.10$ $\quad \frac{\alpha}{2} = 0.05$

$z_c = \text{qnorm}(\frac{1 - 0.10}{2} = 0.05, 0, 1) = -1.644854$

Since the test statistic is not more extreme than the critical value, we fail to reject the null hypothesis at the $\alpha = 0.10$ level.



multiple by 2 because it is two-tailed.

Pvalue $= 2\,P(Z < -1.33) = 2\,\text{pnorm}(-1.33, 0, 1) = 0.1835$

Since Pvalue $= 0.1835 > 0.10 = \alpha$, we fail to reject the null hypothesis.

If $z = +$, then Pvalue $= 2\,(1 - P(Z < x))$

FTR $\longrightarrow$ ($\bullet H_0$)

If $\alpha = 0.05$ then it is a 95% CI

$\alpha = 0.10 = 90\%$ confidence interval

$p_0 = 0.409 \quad x = 32 \quad n = 100 \quad \hat{p} = \frac{32}{100} = 0.32 \quad \alpha = 0.05$

Ex: In a previous study, the National Highway Traffic Safety Administration reported that the proportion of traffic deaths attributable to alcohol was 40.9%. Out of 100 randomly selected traffic deaths this year, 32 were attributable to alcohol. Is there evidence that the true proportion of traffic deaths attributable to alcohol has decreased at the 0.05 significance level?

Step 1: Set up the null and alternative hypothesis

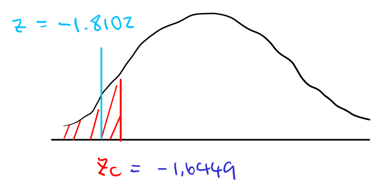
$H_0 : p = p_0 = 0.409$

$H_A : p < p_0 = 0.409$

The proportion of traffic deaths is less than 0.409

Step 2: Draw a picture

Assuming H_0 is true



Step 3: Calculate the test statistic

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{0.32 - 0.409}{\sqrt{\frac{(0.32)(1-0.32)}{100}}} = -1.8102$$

Step 4: Calculate the critical value

$z_c = \text{qnorm}(0.05, 0, 1) = -1.6449$

Step 5: Make and justify a statistical decision at 0.05 level, and state your conclusions in context of the problem.

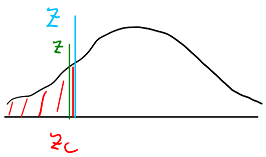
Since the test statistic is more extreme than the critical value, we reject the null hypothesis at the $\alpha = 0.05$ level. There is statistically significant evidence that the proportion of traffic deaths attributable to alcohol is less than 0.409.

Ex: Suppose you calculate a test statistic and find that the p-value is 0.0511 and your level is 0.05. What is your decision? Do you feel comfortable reporting that? What if the p-value is 0.04999?

P-value $\leq \alpha$ we reject

$0.0511 > 0.05$ Fail to reject

sampling variability   statistically significant evidence



Ex: According to the Society of Human Resource Management, 19% of human resource professionals are concerned with losing their jobs. Suppose in a recent sample of 1000 employees, 200 said their biggest concern was losing their job.

a. Calculate a 95% confidence interval for the true population proportion p of employees who are concerned about losing their job.

$n = 1000$ $\qquad PE \pm ME = \hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$

$x = 200$ $\qquad = 0.20 \pm \text{qnorm}(0.95 + \frac{1-0.95}{2} = 0.975, 0, 1) \sqrt{\frac{(0.2)(0.8)}{\sqrt{1000}}}$

$\hat{p} = \frac{200}{1000} = 0.20$

$p_0 = 0.19$ $\qquad = (0.1752, 0.2248)$

b. Without conducting a hypothesis test, do you think there is evidence that the proportion of current employees who are concerned about losing their job is different? Explain in a full sentence.

$p_0 = 0.19$

0.1752 0.2248

$H_A = p \neq p_0 = 0.19$

Since $p_0 = 0.19$ is captured within the interval, we would fail to reject the claim at the $\alpha = 0.05$ level.

$p_0 = 0.85$ $n = 125$ $x = 97$ $\hat{p} = \frac{97}{125} = 0.776$ $\alpha = 0.10$

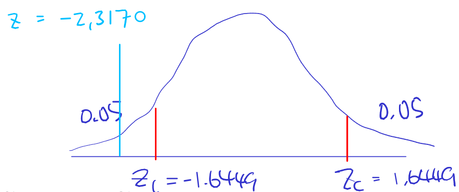
Ex: A local school board claims that 85% of its students score above 1200 on the SAT. A rival school board suspects the actual percentage is different. Of a random sample of 125 students who took the SAT, 97 of them scored above 1200 on the SAT. Test the claim at the 0.10 level.

Step 1: Set up the null and alternative hypothesis

$H_0: p_0 = 0.85$

$H_A: p_0 \neq 0.85$ The proportion of students who score above 1200 on the SAT is not 0.85.

Step 2: Draw a picture



p-value = 2 * pnorm(-2.3170)

2 tailed

Step 3: Calculate the test statistic Assume the Null hypothesis is true $p = p_0$

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{0.776 - 0.85}{\sqrt{\frac{(0.85)(0.15)}{125}}} = -2.3170$$

Step 4: Calculate the critical value

$$z_c = \text{qnorm}\left(1 - \frac{\alpha}{2}\right) = \text{qnorm}\left(1 - \frac{0.10}{2}\right) = \text{qnorm}(0.95)$$
$$= \pm 1.644854$$

Step 5: Make and justify a statistical decision at 0.10 level, and state your conclusions in context of the problem.

Since the test statistic is more extreme than the critical value, we reject the null hypothesis at the α at 0.10 level. There is statistically significant evidence that the proportion of students who score above 1200 on the SAT is not 0.85.

Would a 90% confidence interval contain the null proportion? Explain in a full sentence (verify by calculating the confidence interval)

( ) $\cdot\, H_0$

Since we rejected the claim at the $\alpha = 0.10$ level, a 90 % confidence interval would not capture the null value

$$PE \pm ME = \hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.776 \pm \text{qnorm}(0.95)\sqrt{\frac{(0.776)(0.224)}{125}}$$

$$= (0.7147, 0.8373)$$

Reject H_0 $p_0 = 0.85$ Outside

$\theta = ?$

$\mu_0 = 21$ $\bar{x} = 21.9$ $s = 3.5$ $n = 37$ $\alpha = 0.01$

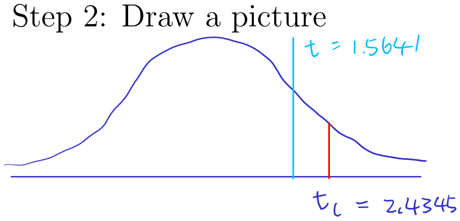
Ex: A feared predator in the ocean is the great white shark. It is known that the great white grows to an average length of 21 feet; however, one marine biologist feels that great whites of the coast off Bermuda grow much longer due to the type and amount of food they consume. Over a period of one year, the marine biologist used a capture/release procedure to measure the sharks. He captured 37 sharks; their average length was 21.9 feet and their standard deviation was 3.5 feet. Test the biologist's claim at the $\alpha = 0.01$ level.

Step 1: Set up the null and alternative hypothesis

$H_0: \mu = 21$

$H_A: \mu > 21$ The average length of Bermuda great whites is greater than 21 feet.

Step 2: Draw a picture



p-value = 1 - pt(1.5641, 36)

Step 3: Calculate the test statistic Assume H_0 is true

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{21.9 - 21}{\frac{3.5}{\sqrt{37}}} = 1.5641$$

Step 4: Calculate the critical value

t_c = qt(1 − α, df = n − 1) = qt(0.99, 36)

= 2.4345

Step 5: Make and justify a statistical decision at 0.01 level, and state your conclusions in context of the problem.

Since the test statistic is not more extreme than the critical value, we fail to reject the null hypothesis at the $\alpha = 0.01$ level. There is not statistically significant evidence that the average length of Bermuda great whites is greater than 21 feet.