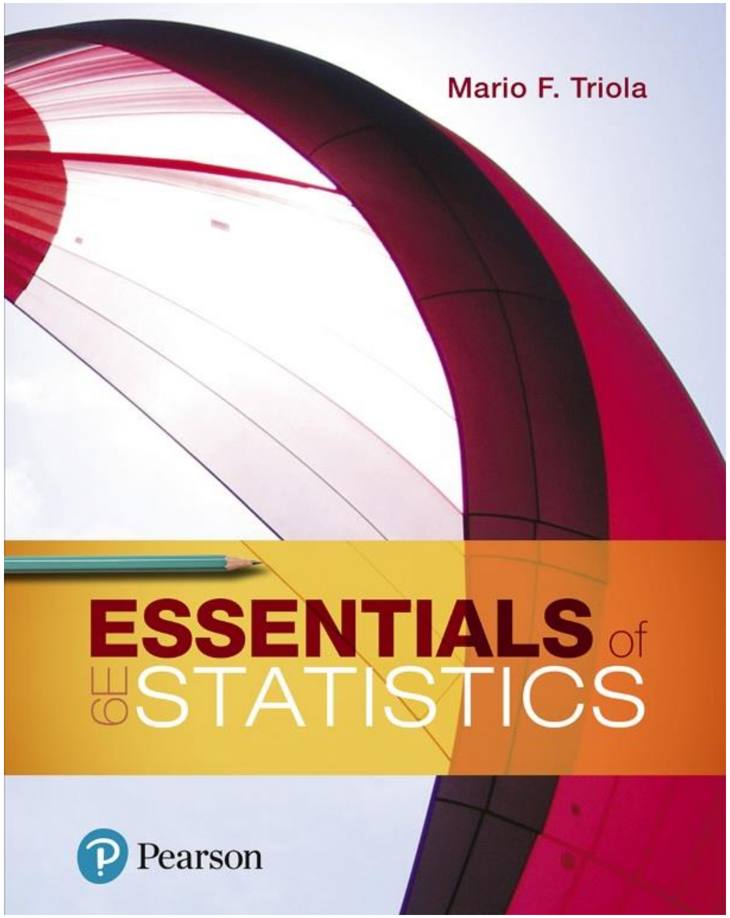
Elementary Statistics

Sixth Edition



Chapter 8

Hypothesis Testing

Hypothesis Testing

Example ZERO:

In a jury trial, the hypotheses are:

H_0: defendant is innocent;
H_1: defendant is guilty.

H_0 (innocent) is rejected if H_1 (guilty) is supported by evidence beyond "reasonable doubt."
Failure to reject H_0 (prove guilty) does not imply innocence, only that the evidence is insufficient to reject it.

Hypothesis and Hypothesis Test

- Hypothesis
  - In statistics, a **hypothesis** is a claim or statement about a property of a population.
- Hypothesis Test
  - A **hypothesis test** (or **test of significance**) is a procedure for testing a claim about a property of a population.
  - **Null hypothesis** H_0 is the symbolic expression that the parameter **equals** = the fixed value being considered.
  - **Alternative hypothesis** H_1 contains, $<$, $>$ or $\neq$

Hypothesis Testing

Example 1:

A chemist invents an additive to increase the life of an automobile battery. If the mean lifetime of the battery is 36 months, then his hypotheses are

H_0: $\mu = 36$ --status quo

H_1: $\mu > 36$ --researcher idea

Example: Fair coin?

Suppose that you are trying to decide whether a coin is fair or biased in favor of heads. The null hypothesis is H_0: the coin is fair (i.e., the probability of a head is 0.5), and the alternative hypothesis is Ha: the coin is biased in favor of a head (i.e. the probability of a head is greater than 0.5).

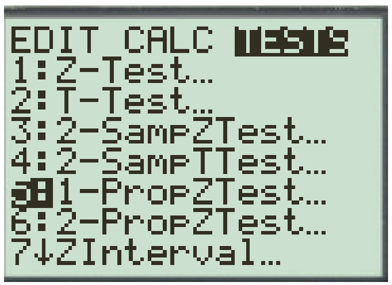
H_0: $p = 0.5$
H_1: $p > 0.5$

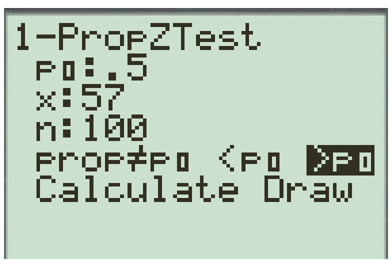
Take a sample [conduct an experiment in this case]: Toss the coin, say 100 times and count the number of heads. Number of heads = 57. Does this result is evidence enough to reject the null hypothesis?

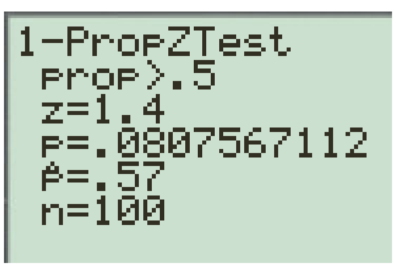
Example: Tossing a coin

Using TI 83 or 84 and Casio 9750

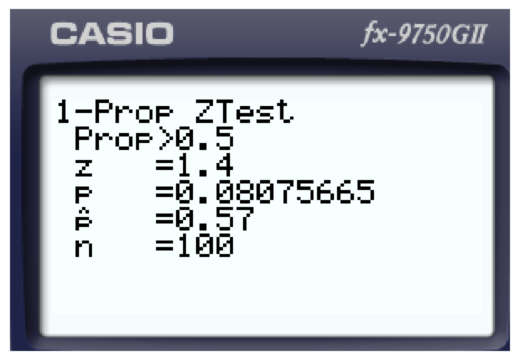
STAT-TESTS-1-PropZTest







F3 (TEST) - Z – 1P

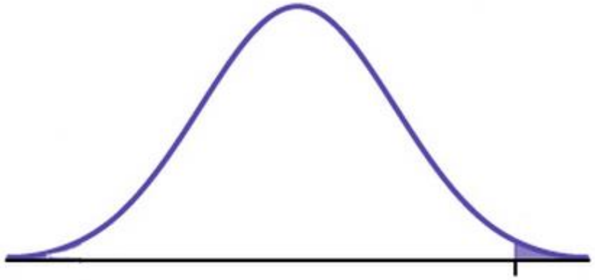
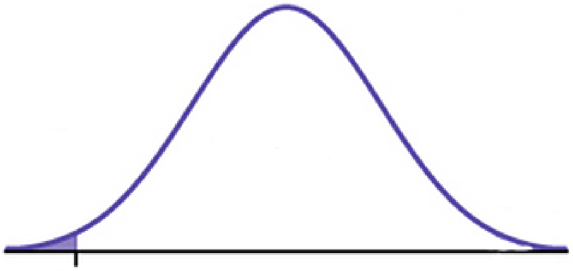
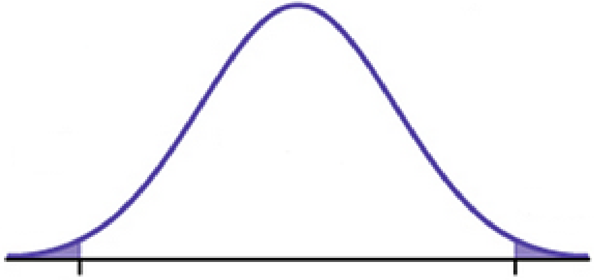


The probability of obtaining 57 heads or more in 100 tosses, assuming the coin is fair is 0.0808.

Pearson

Significance Level α

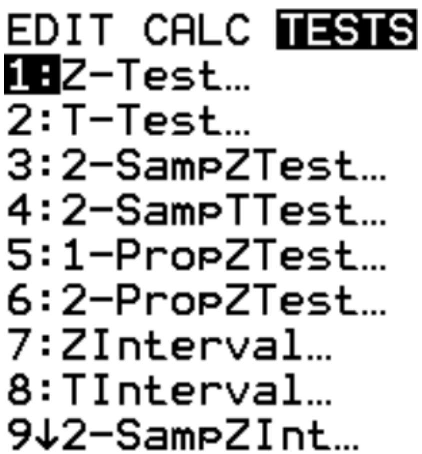
- Significance Level
  - The significance level α for a hypothesis test is the probability value used as the cutoff for determining when the sample evidence constitutes significant evidence against the null hypothesis.
  - The significance level α is the same α introduced earlier in this book, where we defined “critical value”. Common choices for α are 0.05, 0.01, and 0.10; 0.05 is most common.

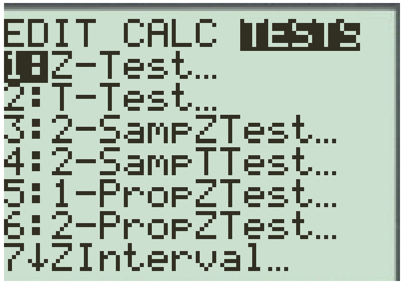




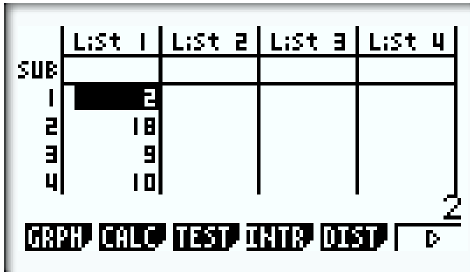
Identify the test statistic that is relevant to the test

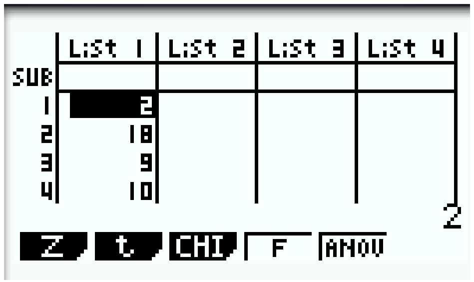
- **TI calc:** STAT - TESTS - 1PropZtest (proportions)
  - Z Test Z Test (Means, sigma known)
  - T Test T Test (Means, sigma unknown)

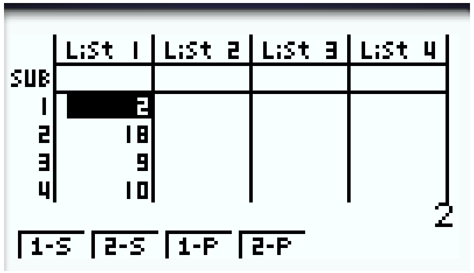
- **Casio calc**: F3 (Test) - F1 (Z) - F3 (1p) 1PropZtest (proportions)
  - F1 (Z) - F1 (1s) Z Test (Means, sigma known)
  - F2 (t) - F1 (1s) T Test (Means, sigma unknown)

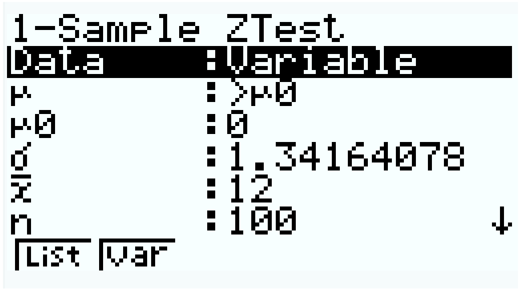


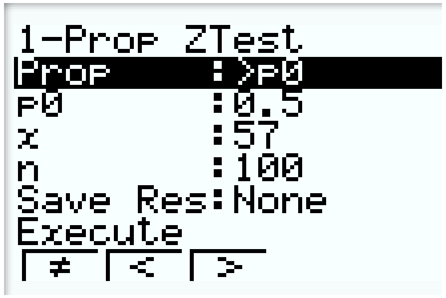


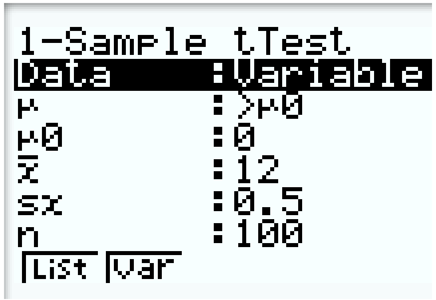












Pearson

Identify the test statistic that is relevant to the test and determine its sampling distribution

Parameter	Sampling Distribution	Requirements	Test Statistic
Proportion p	Normal (z)	$np \geq 5$ and $nq \geq 5$	$z = \dfrac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}}$
Mean μ	t	σ not known and normally distributed population or σ not known and $n > 30$	$t = \dfrac{\bar{x} - \mu}{\dfrac{s}{\sqrt{n}}}$
Mean μ	Normal (z)	σ known and normally distributed population or σ known and $n > 30$	$z = \dfrac{\bar{x} - \mu}{\dfrac{\sigma}{\sqrt{n}}}$

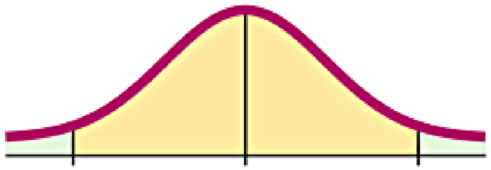
Find the Value of the Test Statistic, Then Find Either the *P*-Value or the Critical Values(s)

Find the value of the **test statistic** and the ***P*-value** or **critical value(s)**.

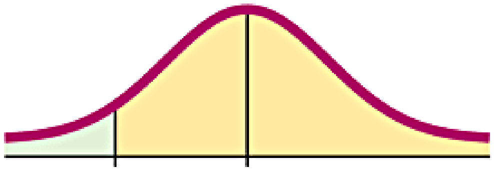
Critical Region: The **critical region** (or **rejection region**) is the area corresponding to all values of the test statistic that cause us to reject the null hypothesis.

Two-Tailed, Left-Tailed, Right-Tailed

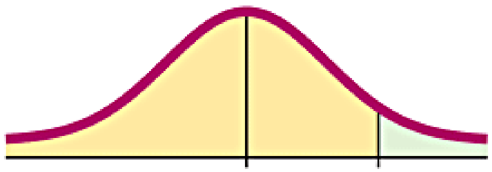
- **Two-tailed test:** The critical region is in the two extreme regions (tails) under the curve.
- **Left-tailed test:** The critical region is in the extreme left region (tail) under the curve.
- **Right-tailed test:** The critical region is in the extreme right region (tail) under the curve.



Sign used in H_1: $\neq$
Two-tailed test



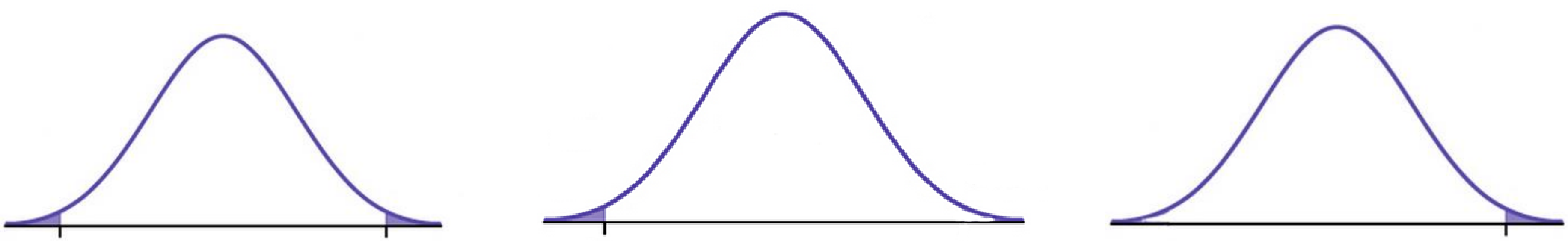
Sign used in H_1: $<$
Left-tailed test



Sign used in H_1: $>$
Right-tailed test

Critical Value Method

- Critical Values
  - In a hypothesis test, the **critical value(s)** separates the critical region (where we reject the null hypothesis) from the values of the test statistic that do not lead to rejection of the null hypothesis.



Make a decision to Either Reject H_0 or Fail to Reject H_0

Make a decision to either reject H_0 or fail to reject H_0.

Decision Criteria for the *P*-Value Method:

- If *P*-value $\leq \alpha$, reject
- If *P*-value $> \alpha$, fail to reject H_0.

Restate the Decision Using Simple and Nontechnical Terms

Wording the Final Conclusion

Condition	Conclusion
Original claim does not include equality, and you reject H_0.	"There is sufficient evidence to **support** the claim that … (original claim)"
Original claim does not include equality, and you fail to reject H_0.	"There is not sufficient evidence to **support** the claim that … (original claim)"
Original claim includes equality, and you reject H_0.	"There is sufficient evidence to warrant **rejection** of the claim that … (original claim)."
Original claim includes equality, and you fail to reject H_0.	"There is not sufficient evidence to warrant **rejection** of the claim that … (original claim)."

Type I and Type II Errors

- **Type I error:** The mistake of rejecting the null hypothesis when it is actually true. The symbol α (alpha) is used to represent the probability of a type I error.

$$\boldsymbol{\alpha} = P(\textbf{type I error}) = P(\textbf{rejecting } H_0 \textbf{ when } H_0 \textbf{ is true})$$

- **Type II error:** The mistake of failing to reject the null hypothesis when it is actually false. The symbol β(beta) is used to represent the probability of a type II error.

$$\boldsymbol{\beta} = P(\textbf{type II error}) = P(\textbf{failing to reject } H_0 \textbf{ when } H_0 \textbf{ is false})$$

Inferences About Two Proportions: Notation for Two Proportions

For population 1 we let

p_1 = **population** proportion

$\hat{p}_1 = \dfrac{x_1}{n_1}$ (**sample** proportion)

n_1 = size of the first sample

$\hat{q}_1 = 1 - \hat{p}_1$ (complement of $\hat{p}_1$)

x_1 = number of successes in the first sample

The corresponding notations p_2, n_2, x_2, $\hat{p}_2$, and $\hat{q}_2$ apply to population 2.

Inferences About Two Proportions: Pooled Sample Proportion

The **pooled sample proportion** is denoted by $\bar{p}$ and it combines the two sample proportions into one proportion, as shown here:

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

$$\bar{q} = 1 - \bar{p}$$

Inferences About Two Proportions: Requirements

1. The sample proportions are from two simple random samples.
2. The two samples are **independent.** (Samples are **independent** if the sample values selected from one population are not related to or somehow naturally paired or matched with the sample values from the other population.)
3. For each of the two samples, there are at least 5 successes and at least 5 failures. (That is, $n\hat{p} \geq 5$ and $n\hat{q} \geq 5$ for each of the two samples).

Inferences About Two Proportions: Test Statistic for Two Proportions (with H_0: $p_1 = p_2$)

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{\bar{p}\bar{q}}{n_1} + \dfrac{\bar{p}\bar{q}}{n_2}}}$$

where $p_1 - p_2 = 0$ (assumed in the null hypothesis)

Where $\bar{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$ (**pooled** sample proportion) and $\bar{q} = 1 - \bar{p}$

Pearson

Confidence Interval Estimate of $p_1 - p_2 = 0$

The confidence interval estimate of the difference $p_1 - p_2$ is

$$(\hat{p}_1 - \hat{p}_2) - E < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + E$$

where the margin of error E **is given by**

$$E = z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}.$$

Hypothesis Tests

For tests of hypotheses made about two population proportions, we consider only tests having a null hypothesis of $p_1 = p_2$ (so the null hypothesis is $H_0 : p_1 = p_2$).

With the assumption that $p_1 = p_2$, the estimates of $\hat{p}_1$ and $\hat{p}_2$ are combined to provide the best estimate of the common value of $\hat{p}_1$ and $\hat{p}_2$, and that combined value is the pooled sample proportion $\bar{p}$ given in the preceding slides.

Inferences About Two Means: Independent Samples: Requirements

Requirements

1. The values of σ_1 and σ_2 are unknown and we do not assume that they are equal.
2. The two samples are **independent.**
3. Both samples are **simple random samples**.
4. Either or both of these conditions are satisfied: The two sample sizes are both **large** (with $n_1 > 30$ and $n_2 > 30$) or both samples come from populations having normal distributions.

Inferences About Two Means: Independent Samples: Hypothesis Test Statistic for Two Means: Independent Samples (with H_0: $\mu_1 = \mu_2$) (1 of 3)

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

(where $\mu_1 - \mu_2$ is often assumed to be 0)

Inferences About Two Means: Independent Samples: Hypothesis Test Statistic for Two Means: Independent Samples (with H_0: $\mu_1 = \mu_2$) (2 of 3)

Degrees of freedom

1. Use this simple and conservative estimate:
   df = smaller of $n_1 - 1$ and $n_2 - 1$.
2. Technologies typically use the more accurate but more difficult estimate given by the following formula:

$$\text{df} = \frac{(A+B)^2}{\dfrac{A^2}{n_1 - 1} + \dfrac{B^2}{n_2 - 1}} \quad \text{where } A = \frac{s_1^2}{n_1} \text{ and } B = \frac{s_2^2}{n_2}$$

Inferences About Two Means: Independent Samples: Hypothesis Test Statistic for Two Means: Independent Samples (with H_0: $\mu_1 = \mu_2$) (3 of 3)

***P*-Values:** *P*-values are automatically provided by technology. If technology is not available, refer to the *t* distribution in Table A-3. **Critical Values:** Refer to the *t* distribution in Table A-3.

Inferences About Two Means: Independent Samples: Confidence Interval Estimate of $\mu_1 - \mu_2$: Independent Samples

The confidence interval estimate of the difference $\mu_1 - \mu_2$ is

$$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$$

where the margin of error E **is given by**

$$E = t_{\frac{\alpha}{2}} \sqrt{\frac{s_1^{\,2}}{n_1} + \frac{s_2^{\,2}}{n_2}}.$$