**In statistics, a hypothesis is a claim or statement about a property of a population.**

**A hypothesis test (or test of significance) is a standard mathematical procedure for testing a claim about a property of a population.**

**If, under a given assumption, the probability of a particular observed event is exceptionally small, we conclude that the assumption is probably not correct.**

**Example:** Suppose we conduct an experiment with 100 couples who want to have baby girls, and they all follow the Gender Choice "easy-to-use in-home system" described in the pink package. For the purpose of testing the claim of an increased likelihood for girls, we will assume that Gender Choice has no effect.
Using common sense and no formal statistical methods, what should we conclude about the assumption of no effect from Gender Choice if 100 couples using Gender Choice have 100 babies consisting of
a) 52 girls?; b) 97 girls?

a) **We normally expect around 50 girls in 100 births. The result of 52 girls is close to 50, so we should not conclude that the Gender Choice product is effective.**
**If the 100 couples used no special method of gender selection, the result of 52 girls could easily occur by chance.**
**The assumption of no effect from Gender Choice appears to be correct. There isn't sufficient evidence to say that Gender Choice is effective.**

b) **The result of 97 girls in 100 births is extremely unlikely to occur by chance. We could explain the occurrence of 97 girls in one of two ways:**
**Either an extremely rare event has occurred by chance, or Gender Choice is effective.**
**The extremely low probability of getting 97 girls is strong evidence against the assumption that Gender Choice has no effect. It does appear to be effective.**

## Key Components of a Hypothesis Test

**First Goal:** **Given a claim, identify the null hypothesis and the alternative hypothesis, and express them both in symbolic form.**

- **The null hypothesis (denoted by $H_0$) is a statement that the value of a population parameter (such as proportion p , mean $\mu$ , or standard deviation $\sigma$) is equal to some claimed value.**

$H_0$ : p = (NUMBER)
$H_0$ : $\mu$ = (NUMBER)
$H_0$ : $\sigma$ = (NUMBER)

- **Any decision is always made about the null hypothesis:**

**Either reject $H_0$ or fail to reject $H_0$.**

- **The alternative hypothesis (denoted by $H_1$ or $H_a$ or $H_A$) is the statement that the parameter has a value that somehow differs from the null hypothesis.**
- **The symbolic form of the alternative hypothesis must use one of these symbols: $\neq$, <, >.**

$H_A$ : p $\neq$ **or** < **or** > ( Same NUMBER)
$H_A$ : $\mu$ $\neq$ **or** < **or** > (Same NUMBER)
$H_A$ : $\sigma$ $\neq$ **or** < **or** > (Same NUMBER)

Example: Identify the Null and Alternative Hypothesis. Use the given claims to express the corresponding null and alternative hypotheses in symbolic form. $h_0$ represents the given value, $h_a$ should be a compliment of it

**a) The proportion of drivers who admit to running red lights is greater than 0.5.**

$H_0$ : $p = 0.5$.

$H_a$ : $p > 0.5$,

**and we let p be the true proportion of drivers who admit to running red lights**

**b) The mean height of professional basketball players is at most 7 ft.**

b) $H_0$ : $\mu = 7$

$H_a$ : $\mu < 7$

and we let $\mu$ be the true mean height of professional basketball players

**c) The standard deviation of IQ scores of actors is equal to 15.**

$H_0$ : $\sigma = 15$

$H_a$ : $\sigma \neq 15$

Where $\sigma$ is the true standard deviation of IQ scores of actors

Second Goal:

**Given a claim and sample data, calculate the value of the test statistic.**

**The test statistic is a value used in making a decision about the null hypothesis, and is found by converting the sample statistic to a score with the assumption that the null hypothesis is true.**

# Test Statistic - Formulas

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

**Test statistic for proportions**

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}}$$

**Test statistic for mean**

Very similar to critical values from the last unit. We just account for less than and greater than with one or two tails.

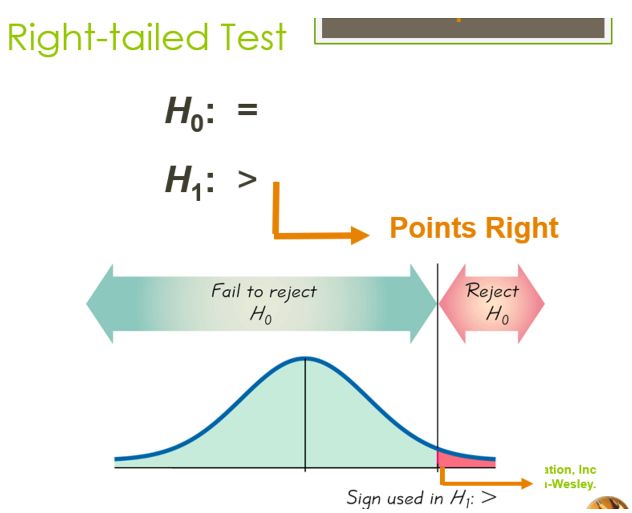
Example: A survey of n=880 randomley selected adult drivers showed that 56% (or $\hat{p} = .56$) of those responded admitted to running red lights. Find the value of the test statistic for the claim the the majority of all adult drivers admit to running red lights.

Solution: The preceding example showed that the given claim results in the following null and alternative hypotheses: $H_0$: $p = 0.5$ and $H_a$: $p > 0.5$. Because we work under the assumption that the null hypothesis is true with $p = 0.5$, we get the following test statistic:

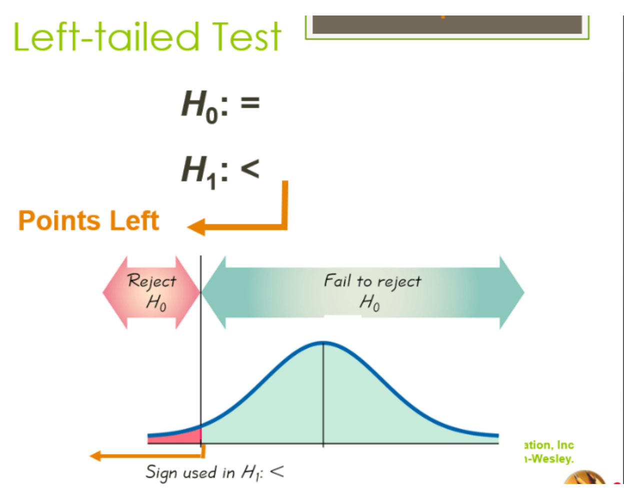
$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.56 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{880}}} = 3.56$$

Third Goal: **Given a value of the test statistic, identify the *P*-value**

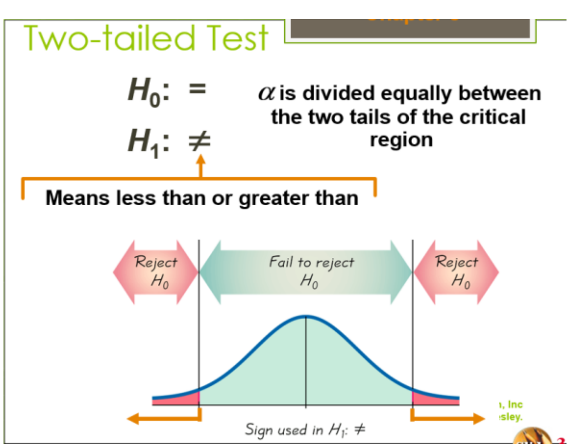
**The *P*-value (or *p*-value or probability value) is the probability of getting a value of the test statistic that is at least as extreme as the one representing the sample data, assuming that the null hypothesis is true. The null hypothesis is rejected if the *P*-value is very small, such as 0.05 or less.**




Use for Greater than




Use for less than




Use when not equal too

Critical Regions are in red!!!!!!

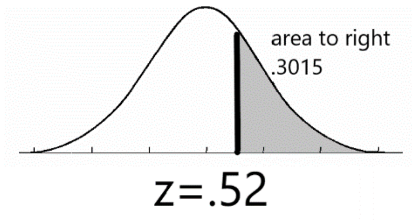
**We always test the null hypothesis. The initial conclusion will always be one of the following:**

**1. Reject the null hypothesis. Reject $H_0$ if the test statistic falls within the critical region.**

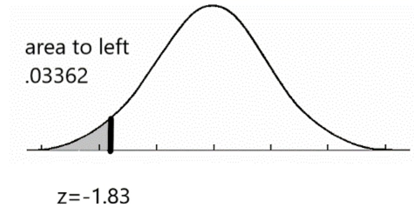
**2. Fail to reject the null hypothesis. Fail to reject $H_0$ if the test statistic does not fall within the critical region.**

Example: Given the following information find the p-value

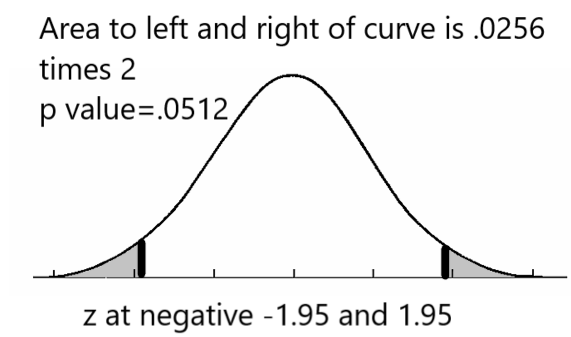
1. The test statistic is z=.52 and it is right tailed.




2. Test statistic is left tailed and z=-1.83




3. The test statistic is 1.95 and is two tailed.




Identify the null hypothesis the alternative hypothesis, test statistic, p-value and conclusion about the null hypothesis and final conclusion that addresses the original claim

1. According to recent poll 53% of Americans would vote for the incumbent president. If a random sample of 11 people results in 45% who would vote for the incumbent, test the claim the actual percentage is 53%. Us a .10 significance level

$h_0\ p = .53$

p-value: Using a z of 1.6 $\propto = .0548$

$h_a\ p \neq 53$

Test Statistic $z = \frac{.45-.53}{\sqrt{\frac{(.53)(.47)}{100}}} = 1.60$

Critical value at 90% is 1.645, fail to reject null hypothesis. It falls within the critical value

$\propto > \hat{p}$ do not reject null hypothesis

$\propto < \hat{p}$ reje

2. Previously, an organization reported that teenagers spent 4.5 hours per week, on average, on the phone. The organization thinks that, currently, the mean is higher. Fifteen randomly chosen teenagers were asked how many hours per week they spend on the phone. The sample mean was 4.75 hours with a sample standard deviation of 2.0. Conduct a hypothesis test with 95% confidence level

$h_0\ \mu = 4.5$ $\quad h_a\ \ \mu > 4.5$ $\quad z = \frac{4.75-4.5}{\frac{2}{\sqrt{15}}} = 0.484$ $\quad$ p-value 0.314 $\quad$ Falls within the critical value

We will not reject this null hypothesis

Prob/Stat:

I. Identifying $H_o$ and $H_A$. State the null and Alternative Hypotheses in symbolic notation and identify the claim.

1. A university claims that the proportion of its students who graduate in four years is more than 82%.
   $H_o$ :

   $H_A$ :

2. A water faucet manufacturer claims that the mean flow rate of a certain type of faucet is less than 2.5 gallons per minutes.
   $H_o$ :

   $H_A$ :
3. A television manufacturer claims that the standard deviation of the life of a certain type of television is 3 years.
   $H_o$ :

   $H_A$ :

II. **Sketch** the normal distribution**, label the areas** for each significance level, **shade the "tails"** and find and **label the critical values**.

4. (Two-tailed test) $H_A$: $p \neq 0.5$, $\alpha = 0.01$

5. (Left tailed test) $H_A$: $\mu < 12$ feet, $\alpha = 0.05$

6. (Right-tailed test) $H_A$: $p > 0.5$, $\alpha = 0.01$

Hypothesis Testing Worksheet

1. For the following pairs, indicate which aren't legitimate hypotheses and explain why.
   a) $H_o: p \neq 0.4$ : $H_A: \hat{p} > 0.4$

   b) $H_o: \hat{p} = 0.16$ : $H_A: \hat{p} > 0.16$

   c) $H_o: \mu = 22$ : $H_A: \mu > 24$

2. For each situation, state the null and alternative hypothesis:
   a) In a random sample of 100 adult Americans, only 430 could name at least one justice who is currently serving on the U.S. Supreme Court. A claim is that half of adult Americans can name at least one justice who is currently serving on the U.S. Supreme Court.

   b) In a national survey of 2013 adults, 1283 indicated that they believe rudeness is a more serious problem than in years past. Does this indicate that more than three-quarters of American adults believe rudeness is a worsening problem?

   c) The Associated Press found that 730 of 1000 randomly selected adults preferred to watch movies at home rather than at a movie theater. Is there convincing evidence that the majority of adult Americans prefer watching movies at home?

   d) In a survey of 526 U.S. businesses, 400 indicated that they monitor employee's web site visits. Is there sufficient evidence that more than 70% of U.S. businesses monitor employees' web site visits?

   e) In a random sample of 1000 adult Americans, 700 indicated that they oppose the reinstatement of a military draft. Is there convincing evidence that the proportion of American adults who oppose the reinstatement of the draft is different from two-thirds?

For each of the following p-values, state whether you would "reject" or " fail to reject" $H_0$ for the given $\alpha$ level.

1) P-value = 0.234 when $\alpha$ = 0.05

2) P-value = 0.024 when $\alpha$ **= 0.05**

3) P-value = 0.024 when $\alpha$ = 0.01

4) P-value = 0.024 when $\alpha$ = 0.10

For each of the following, a) draw and shade the curve

b) calculate the p-value of the given test statistic

5) Right-tailed test $z = 2.05$

6) Left-tail test $z = -1.95$

7) Two-tail test $z = 1.75$

8) A newspaper article headline reads "One in Five Believe Path to Riches Is the Lottery". A survey was done to test the claim that 20% of adult Americans believe that the lottery is the path to riches. The hypothesis test of $H_o$: $p = 0.2$ versus $p > 0.2$ resulted in $z = 0.79$ for a right tailed test. Calculate the p-value and write the correct conclusion in context.

9) A researcher claims that less than 30% of adults are allergic to weeds. IN a random sample of 86 adults, 17 said they have such an allergy. At a significance level of $\alpha = 0.05$, is there enough evidence to support the researcher's claim.

10) Mendel"s Genetic Experiments. When Gregor Mendel conducted his famous hybridization experiment with peas, one such experiment in offspring consisting of 428 peas with green pods and 152 peas with yellow pods. According to his theory, ¼ of the offspring peas should have yellow pods. Use a 0.01 significance level to test the claim that the proportion of peas with yellow pods is equal to ¼.

# Probability Distribution Review

**1.** A survey asks a sample of students how many days they work in a typical week.

| x (number of days worked) | P(x) |
|---|---|
| 0 | .17 |
| 1 | .02 |
| 2 | .04 |
| 3 | .12 |
| 4 | .32 |
| 5 | .21 |
| 6 | .08 |
| 7 | .04 |

a. Find the expected number of days worked.

b. Find the standard deviation for the number of days worked.

c. Find the variance for the number of days worked.

d. Find the probability of selecting a student that works 2 days a week.

e. Find the probability of selecting a student that works less than 5 days a week.

f. Find the probability of selecting a student that works between 4 and 6 days, inclusive, per week.

**2.** A manufacturing report states that 6% of all products made by an old machine are defective. For the problems below, assume a sample size of 52.

S=
F=
p=
q=
n=

a. Find the probability that fewer than 8 products are defective.

b. Find the probability that exactly 4 products are defective.

c. Find the probability that more than 5 products are defective.

d. Find the probability that at most 6 products are defective.

e. Find the probability that at least 3 products are defective.

f. Find the expected number, variance, and standard deviation of defective items out of the 52 products in the sample.

# Probility Review

Listed below is a chart containing the color and number of marbles that are in a bag.

| Color | Number of Marbles |
|---|---|
| Red | 10 |
| Orange | 18 |
| Yellow | 6 |
| Green | 4 |
| Blue | 10 |
| Purple | 12 |


**1.** Find the probability of drawing a purple marble.

**2.** Find the probability of drawing two red marbles with replacement.

**3.** Find the probability of drawing a green marble then a yellow marble without replacement.

**4.** Find the odds in favor of drawing an orange marble.

**5.** Find the odds against drawing a yellow marble.

**6.** Find the probability of drawing a two blue marbles followed by a green marble without replacement.

**7.** Find the probability of drawing a yellow or blue marble.

**8.** Find the odds in favor of drawing a blue marble.

**9.** Find the probability of drawing a red marble then a green marble with replacement.

**10.** Find the odds against drawing a green or orange marble.

*Round to three decimal places.*

**11.** The table below summarizes results from a study of people who refused to answer survey questions.

| | | Age | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 18-21 | 22-29 | 30-39 | 40-49 | 50-59 | 60 and over | *Total* |
| Response to survey | Responded | 74 | 256 | 246 | 137 | 139 | 203 | *1055* |
| | Refused | 15 | 24 | 37 | 30 | 39 | 61 | *206* |
| | *Total* | *89* | *280* | *283* | *167* | *178* | *264* | *1261* |

Find the probability that the selected person…

a. refused to respond to the survey.

b. is 18-21, given they responded to the survey.

c. is below 22 or is older than 59.

d. is between 22-29 or refused to respond to the survey.

e. is age 30-59.

f. is age 30-39 and responded to the survey.

g. refused to respond to the survey, given they are 60 and over.

h. is age 60 and over, given they refused to respond to the survey.

i. responded to the survey and is age 50-59.

j. is age 22-39 or refused to respond to the survey.