UNIT 2: Is There a Home Field Advantage in the NFL?
Comparing Two Proportions

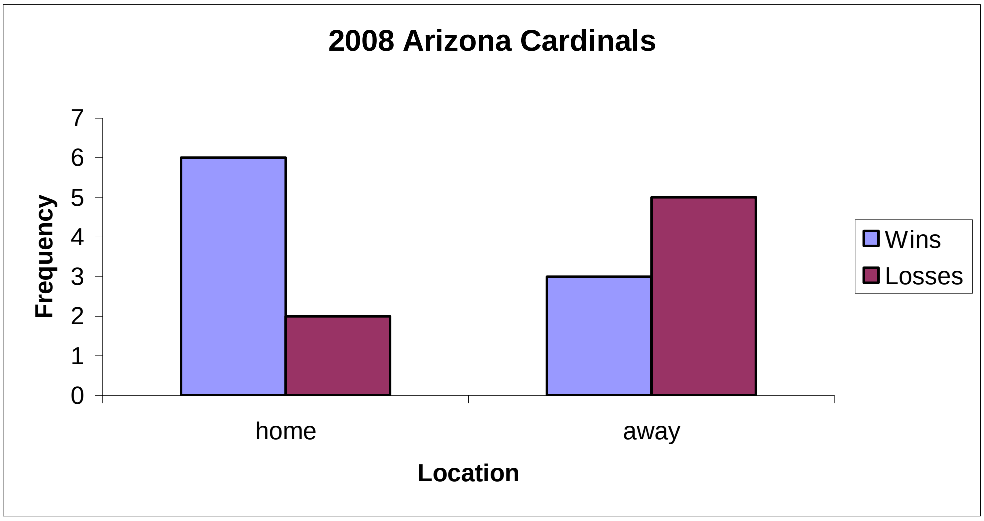
In the National Football League—and in virtually every other sports league—it is assumed that the home team has an advantage. After all, having 60,000 fanatics cheering for you can't hurt. This is why teams with the best records are awarded playoff games at home—to reward their regular season success and give them a more favorable chance to advance into the next round. In this Unit, we will investigate whether the home field advantage really does exist in the NFL.

In the American judicial system, when a person is on trial, the jury is instructed to consider the defendant innocent until proven guilty. Furthermore, to declare a person guilty requires that there is evidence beyond a reasonable doubt. These same principles apply when testing a statistical claim. That is, we start by assuming there is no home field advantage until proven otherwise. Furthermore, we cannot conclude that there is a home field advantage unless we have convincing evidence. To be convincing, it must be clear that the evidence was not due to *Random Chance*.

The 2008 Arizona Cardinals

One of the most surprising teams in the 2008 NFL season was the Arizona Cardinals. Long a laughing stock of the league, they cruised to a NFC West division title and then beat 3 favored teams in a row to reach the Super Bowl. In the Super Bowl they narrowly lost to the Pittsburgh Steelers in a thrilling game. However, as exciting as the post-season was for Cardinal fans, we are not going to include it in our analysis. Instead we will focus just on the regular season.

During the regular season, the Cardinals were 9-7 overall (9 wins and 7 losses) with a 6-2 home record (winning percentage = 75%) and a 3-5 road record (winning percentage = 37.5%). As always, we start the analysis with a graph to compare the Home *Performance* and Road *Performance* of the Cardinals:



The graph certainly suggests that the Cardinals were a better team at home. However, it is possible that the Cardinals had no extra *Ability* at home and the favorable *Performances* were just due to *Random Chance*.

Stating Hypotheses:
The claim that the Cardinals have the same *Ability* to win at home and on the road is called the null hypothesis. The claim that the Cardinals have a greater *Ability* to win at home is called the alternative hypothesis. The alternative hypothesis describes our suspicion (what we want to prove) while the null hypothesis describes a neutral situation.

Here are the same hypotheses written out formally, with the null hypothesis denoted H_0 and the alternative hypothesis denoted H_a.

H_0: The 2008 Cardinals have the same *Ability* to win at home and on the road.

H_a: The 2008 Cardinals have a greater *Ability* to win at home than on the road.

Note: Hypotheses are always stated in terms of *Ability* (what we are trying to estimate) not *Performance* (the data we have observed). We already know that the Cardinals *Performed* better at home. What we want to know is if they really have a greater *Ability* to win at home.

Testing Hypotheses:
When we test hypotheses, we always start by assuming the null hypothesis is true until we have convincing evidence otherwise. In other words, we start by assuming there is no difference in *Ability* until we can show the difference in *Performance* was not due to *Random Chance*.

The difference in their winning percentage at home and on the road was 75% - 37.5% = 37.5%. We want to know how likely it is for a team that goes 9-7 overall to have a difference of at least 37.5% in their home and road winning percentages simply due to *Random Chance* and not any difference in *Ability*. This probability is called a *p*-value.

To investigate what role *Random Chance* could have in a football team's performance at home and on the road we will conduct a simulation. First, begin with 16 equally sized index cards. On nine of these cards write a "W" on one side. On the other seven cards, write a "L" on one side. Then, turn them all upside down and shuffle them up. Now, deal 8 cards into one pile to represent the team's performance at home and put the remaining cards in another pile to represent the team's performance on the road.

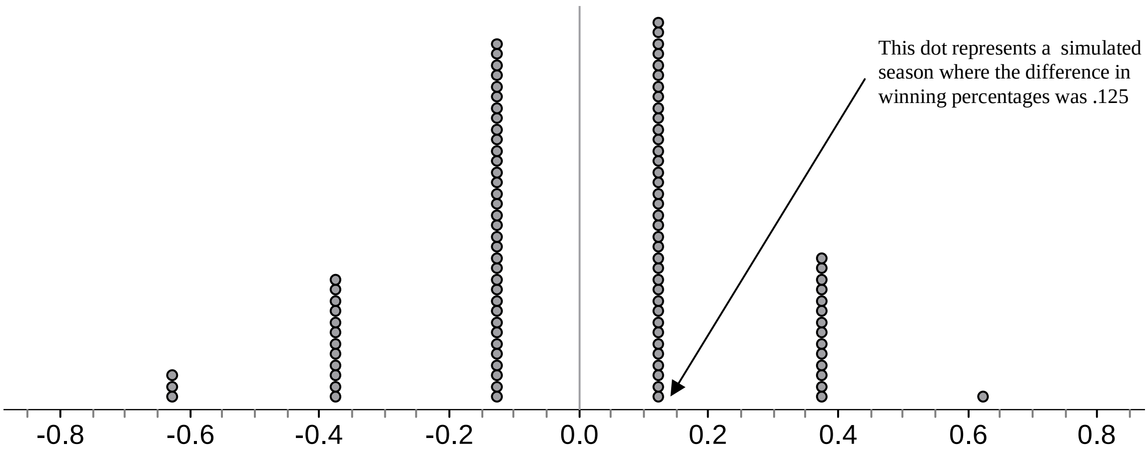
This method of simulation works well because it will always result in a 9-7 overall record for the Cardinals which matches their overall performance in the regular season. It also assumes that there is no difference in *Ability* by randomly allocating the wins to the home games and road games with equal likelihood.

Here are the results of one trial:

Home:	L	W	W	W	W	L	W	L	record: 5-3 (.625)
Road:	L	W	W	L	W	L	W	L	record: 4-4 (.500)
									difference: .125

In this trial, the Cardinals performed better at home, but only by .125 (instead of the .375 in the actual 2008 season). The result of this trial is marked with the arrow in the dotplot below.

Here are the results of 100 simulated “seasons”. The variable displayed is the difference between the home winning percentage and the road winning percentage.

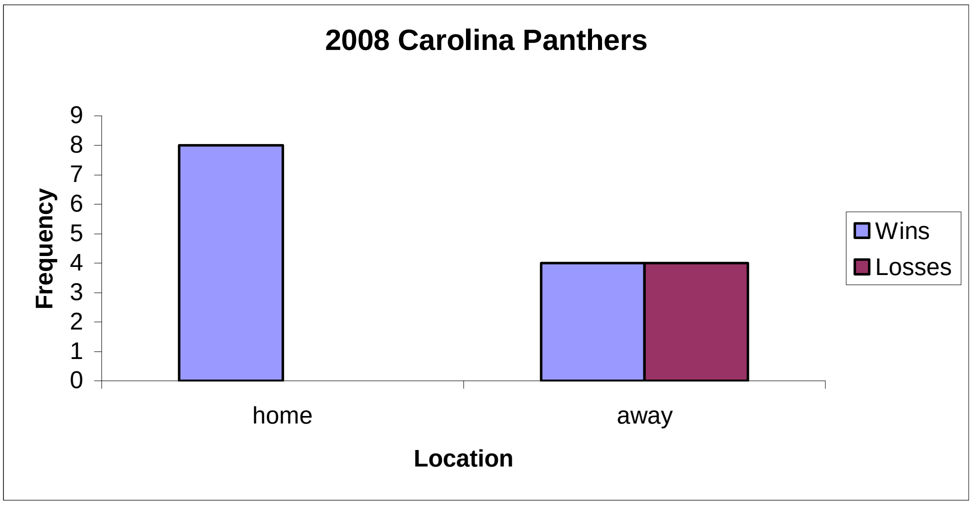


Simulated Differences in Winning Percentage
2008 Arizona Cardinals

In the 100 simulated seasons, the Cardinals had a difference equal to or greater than .375 in 15 of the 100 seasons (p-value = 0.15). Since this simulation was conducted assuming there is no difference in *Ability* and a difference in *Performance* of .375 was relatively common, the 2008 Cardinals do not give us convincing evidence that there really is a home field advantage. After all, their superior performance at home could simply be the results of *Random Chance*.

Using Technology for Simulation: 2008 Carolina Panthers

Now, let’s take a look at another team, the 2008 Carolina Panthers. Overall, the Panthers were 12-4 but were an astounding 8-0 at home (and just 4-4 on the road).



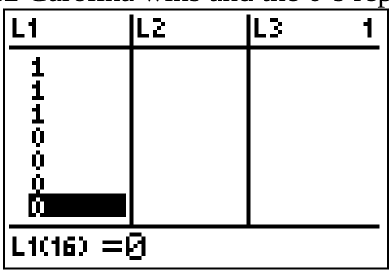
This means the difference in their home and road winning percentages was 1.000 - .500 = .500. Again, it is obvious that the Panthers *Performed* better at home than on the road. However, we need to investigate if they really have a greater *Ability* to win at home or if their *Performance* could have been a result of *Random Chance*.

Here are the hypotheses we will test:

H_0 : The 2008 Panthers have the same *Ability* to win at home and on the road.
H_a : The 2008 Panthers have a greater *Ability* to win at home than on the road.

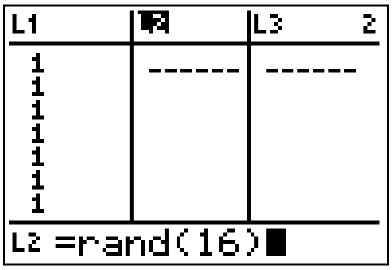
To simulate the season for the 2008 Carolina Panthers using a TI-84:

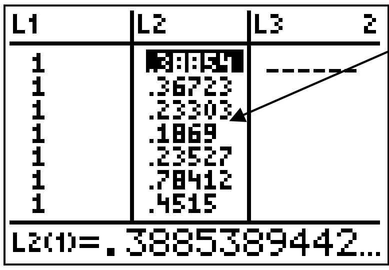
1. Open the list editor (Stat: Edit) and enter twelve 1's and four 0's into L1. In this case, the 1's represent the 12 Carolina wins and the 0's represent the 4 Carolina losses.



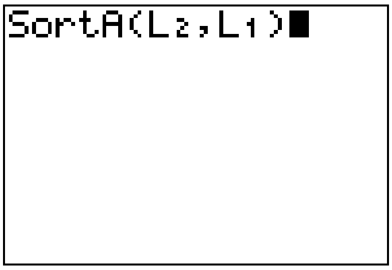
2. Next, we need to randomly shuffle these 1's and 0's into two groups of eight (to represent the 8 home games and the 8 road games). To do this, go to the heading of L2 and enter the following formula: Rand(16) and press enter (the Rand command is in the Math:Prb menu). This puts 16 random numbers into L2.

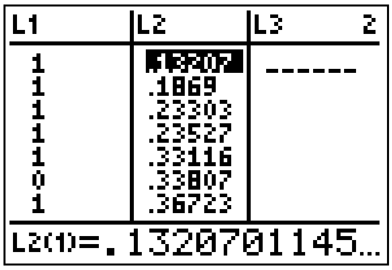
Your L2 will look different since the numbers are random!





3. Then, from the home screen, enter the command SortA(L2,L1) and press enter (the SortA command is in the Stat:Edit menu). This will put the random numbers in order and shuffle the 1's and 0's at the same time.



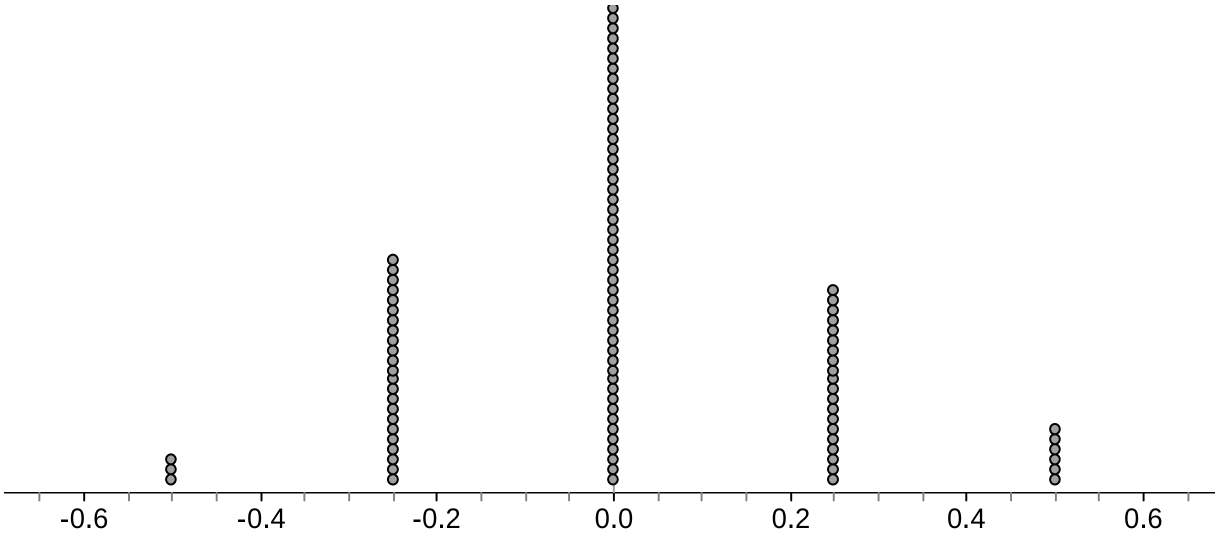


4. Finally, count how many 1's there are in the first 8 positions (the home games) and how many 1's there are in the last 8 positions (the road games) in L1 and compute the difference in the winning percentage (home – away). It is possible that this difference will be negative which means that the team won a higher percentage of games on the road.

Note: In the technology appendix you will find the code for a TI-84 program which automates this simulation process. There are also online applets that can automate this process for you.

Note: Many computer software packages, including Fathom, are programmed with a "shuffle" option which imitates the shuffling of the index cards used in the Cardinals example. Detailed instructions on how to use Fathom is given in the technology appendix.

Here are the results of 100 simulated seasons for the 2008 Carolina Panthers:



Simulated Differences in Winning Percentage
2008 Carolina Panthers

As you can see, in 6 of the 100 simulated seasons the Panthers had a difference of .500 in winning percentage (*p*-value = .06). So, when we assume the Panthers have the same *Ability* to win at home and on the road, it is possible that a 12-4 team can *Perform* 8-0 at home just by *Random Chance*. However, just because something is possible doesn't make it likely. After all, 6% is not a very big probability. The best conclusion in this situation is that we have moderately convincing evidence that the Panthers play better at home.

Remember, the smaller the *p*-value, the more convinced we are that the results are not due to *Random Chance*. For example, if the *p*-value for Carolina had been .01 instead of .06, this would mean their *Performance* at home compared to on the road is even less likely to have occurred by *Random Chance* (and thus we have more evidence of a home field advantage). If the *p*-value is higher, such as *p*-value = .15 for the Cardinals, it is possible that the difference between home and road *Performance* could have been due to *Random Chance* alone (and thus we have less evidence of a home field advantage).

Another Example: The Schilling Theory

On February 19, 2009 Tom Verducci wrote an article on SportsIllustrated.com titled "In the Age of Baseball Parity, Pitching Health Rules." Here is a brief excerpt from the article:

> *But when you look at the depth in their respective divisions and the good fortune they enjoyed last season in pitching health, you begin to understand why repeating is so difficult. The 2008 Phillies and Rays were classic examples of The Schilling Theory: The team that gets the most starts out of its projected rotation has the best shot at winning.*
>
> *Want to see The Schilling Theory at work? Here are the only teams in 2008 to get 30 starts from four starters:*

1. Phillies (Won World Series)
2. Rays (Won AL pennant)
3. Angels (Won AL West)
4. White Sox (Won AL Central)

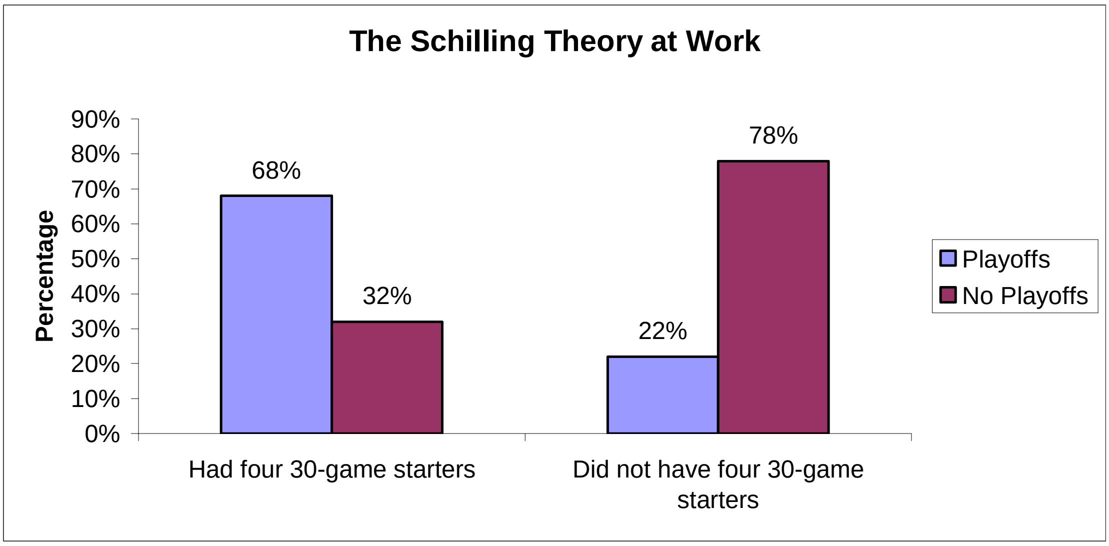
Since 2000, 28 teams, or about three per year, were fortunate enough to have four starters make 30 starts. Nineteen of those teams made the playoffs, or 68 percent of them.

Considering that only 8 out of 30 teams make the playoffs each year, it seems that Curt Schilling was on to something when he predicted that the teams with the healthiest starting pitching are more likely to make the playoffs.

Here is a two-way table summarizing the data presented in the original article:

Since 2000	Teams with four 30-game starters	Teams without four 30-game starters	Total
Teams that made the playoffs	19	53	72
Teams that did not make the playoffs	9	189	198
Total	28	242	270

Here is a comparative bar chart showing the same results. The graph shows the relative frequencies (percentages) instead of the raw numbers to make the categories easier to compare. Using relative frequencies is usually a good idea, especially when the number of observations are very different in the categories.



From both the table and the graph, it appears that Schilling was correct: a team with healthy starting pitching seems more likely to make the playoffs. Specifically, the difference in the percentages of teams who makes the playoffs is 68% - 22% = 46%.

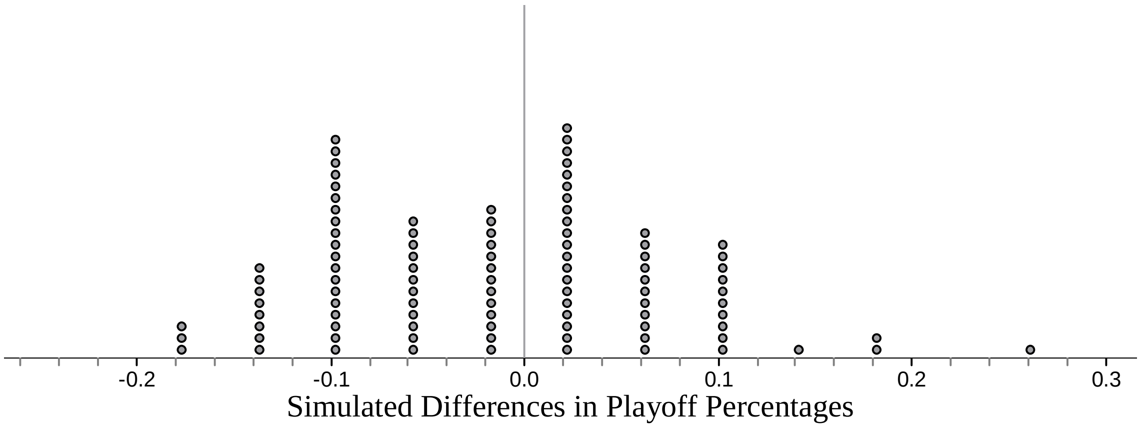
What we want to investigate is whether or not this could have occurred simply by *Random Chance*. In other words, is it possible that teams having four 30-game starters really have the same *Ability* to make the playoffs as teams with fewer than four 30-game starters and the difference in *Performance* during the last 9 years was due to random chance? To answer this question we will be testing the following hypotheses:

H_0 : Teams with four 30-game starters have the same *Ability* to make the playoffs as teams with fewer than four 30-game starters.

H_a : Teams with four 30-game starters have a greater *Ability* to make the playoffs compared to teams with fewer than four 30-game starters.

It's time for another simulation! In this case we are assuming that making the playoffs is just as likely for teams with healthy pitching as it is for teams without healthy pitching. So, we will be taking the 72 playoff appearances and randomly assigning them to the 28 healthy teams and the 242 unhealthy teams and then computing the difference in the percentage of time each type of team makes the playoffs.

For example, in one simulation, the healthy teams made the playoffs 32% of the time and the unhealthy teams made the playoffs 26% of the time, for a difference of 6% (or .06). The results of 100 trials of this simulation are shown below:



As you can see, none of the simulated differences are even close to the 46% difference we saw in the real data. That is, according to this simulation, the *p*-value is 0. Since getting a difference as high as 46% is *very* unlikely to happen by *Random Chance* alone, the data gives us strong evidence that teams with four 30-game starters have a greater *Ability* to make the playoffs than teams with fewer than four 30-game starters.

Caution about Causation:

One of the most repeated phrases in statistics is: "Association does not imply causation." That is, even if there is an association between two events, it doesn't mean that one caused the other. In this case, there is a strong association between having four healthy starting pitchers and making the playoffs. However, having four terrible pitchers pitch over and over again for the whole season won't get you anywhere but last place. So, having four pitchers start at least 30 games is no guarantee. A hidden assumption in the Schilling Theory is that the starting pitchers on a team must be decent to begin with. Only then will the health of the starters increase the probability of making the playoffs. Still, there is a lesson to be learned for general managers: when considering which starting pitchers to sign, durability is an important factor to consider.

Connections: Looking Forward…Looking Back

In Unit 1, we were testing if a previously established *Ability* level has changed. It was assumed that the previously established *Ability* level was a known constant (and not a variable). For example, we assumed that LeBron's true ability to make three-pointers was 31.5% and then investigated if his *Ability* to make three-pointers went down in the playoffs.

However, in Unit 2, we do not start with a known *Ability* level—instead we are comparing the estimated *Ability* levels for two different contexts in the same time period. For example, we compared the estimated *Ability* to win at home vs. estimated *Ability* to win on the road during a particular season.

In Unit 3 we will use the same methods learned in Unit 2 to investigate a long debated question: are repeated trials in sports independent? Later units will employ the same shuffling techniques but focus on differences in numerical data rather than categorical data.

Stats 101: The Traditional Approach

In a traditional statistics course, the tests we were performing in Unit 1 are usually called "one sample tests for a proportion" or "one proportion z-tests" and the tests we were performing in Unit 2 are usually called "two sample tests for a difference in proportions" or "two proportion z-tests." The methods typically used to perform these tests are not based on simulation, but rather by approximating the distribution of the difference in proportions with a normal curve. To be accurate, these tests require that there are a large number of observations in each category. This is not a requirement for the simulation approach, however. We will learn more about normal curves in Unit 7.

For Practice

1. State the null and alternative hypotheses for the LeBron James example in Unit 1.

2. State the null and alternative hypotheses for the Ervin Santana example in Unit 1.

3. On August 8, 2009, broadcasters on the Fox TV game of the week noted that Yankees outfielder Nick Swisher has his 3 home runs at home and 15 on the road. Does this indicate that Swisher has a greater *Ability* to hit home runs on the road? State the null and alternative hypotheses for this question.

4. Describe the errors in the following sets of hypotheses. Explain how they can be corrected.

a) H_0 : The University of Florida football team *Performs* equally well at home and on the road

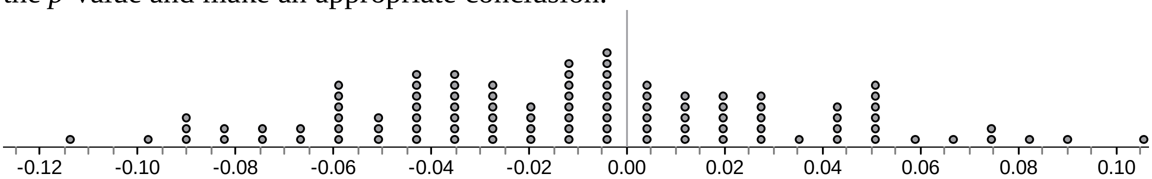
H_a : The University of Florida football team *Performs* better at home than on the road.

b) H_0 : The Boston Red Sox have a greater *Ability* to score runs against left handed pitchers than against right handed pitchers.

H_a : The Boston Red Sox have the same *Ability* to score runs against left handed pitchers and against right handed pitchers.

5. In the 2008 NFL season, the home teams won 57.3% (146/255) of their games while the road team won only 42.7% (109/255) for a difference of 14.6% (there was one tie, which is why the total for each is 255 instead of 256). Does this give evidence of a home field advantage in the NFL?

a) State the null and alternative hypotheses we are interested in testing.
b) Describe how you could use the methods of this Unit to simulate this situation.
c) A 100-trial simulation was conducted assuming that there is no home field advantage. The difference between the winning percentage for the home teams and the road teams was calculated for each trial and the results are below. Based on these results, estimate the *p*-value and make an appropriate conclusion.



Simulated Difference in Winning Percentage (Home – Road)

d) Why do you think that a 14.6% difference is very convincing evidence of a home field advantage when all NFL teams are combined but the 50% difference for the Carolina Panthers wasn't as convincing?

6. In sports such as tennis, players often think it makes a difference which side of the court they are on. This could be due to differences in visibility, wind, or other factors. Suppose that in a particular match, Kathy wins 22 of the 37 points when facing away from the sun but only 11 of the 35 points when facing the sun. Does this data give convincing evidence that Kathy plays better when not facing the sun?

a) How much better did Kathy perform when facing away from the sun? Calculate the difference in the percent of points won.
b) State the hypotheses we are interested in testing.
c) Describe how you could use the methods of this Unit to simulate this situation.
d) Suppose that the results of a simulation gave a *p*-value of .01. Interpret this value.
e) What conclusion would you make based on the *p*-value in part (d)?

7. In 2007, Angels pitcher Ervin Santana was 6-4 at home (60% winning percentage) but only 1-10 on the road (9.1% winning percentage) for a difference of 50.9%. Does this give evidence that Santana had a better *Ability* to win games at home compared to on the road?

a) State the null and alternative hypotheses we are interested in testing.
b) Describe how you could use the methods of this Unit to simulate this situation.
c) Conduct the simulation using the method you described in part (b). Do at least 10 trials and record the outcomes on a dotplot.
d) Based on the results of the simulation, what can you conclude about Santana's *Ability?*

8. Just like in tennis, beach volleyball players often prefer one side to another. Suppose that when Ashley was serving into the wind, she made 25 of 31 serves but while serving with the wind she made 27 of 30 serves. Does this data give convincing evidence that Ashley plays serves better with the wind?

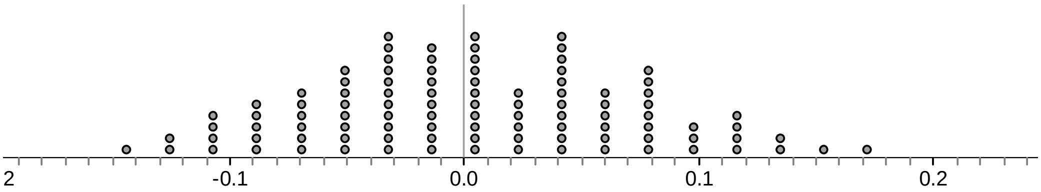
a) How much better did Ashley perform when serving with the wind? Calculate the difference in the percent of serves made.
b) State the hypotheses we are interested in testing.
c) Describe how you could use the methods of this Unit to simulate this situation.
d) Suppose that the results of a simulation gave a *p*-value of .16. Interpret this value.
e) What conclusion would you make based on the *p*-value in part (d)?

9. Suppose that Joe is an avid bowler and is testing two different ball weights: 14 pounds and 16 pounds. He rolls each of the balls 10 times and gets 4 strikes with the 16 pound ball but only 2 strikes with the 14 pound ball. Does this data give convincing evidence that Joe has a greater *Ability* to roll strikes with the 16 pound ball?

a) How much better did Joe perform with the 16 pound ball? Calculate the difference in the percent of strikes rolled.
b) State the hypotheses we are interested in testing.
c) Describe how you could use the methods of this Unit to simulate this situation.
d) Conduct the simulation using the method you described in part (c). Do at least 10 trials and record the outcomes on a dotplot.
e) Based on the results of the simulation, what can you conclude about Joe's *Ability?*

10. In the February 1, 2009, edition of the *Arizona Daily Star*, Ryan Finley wonders what effect a new logo has on the Arizona Cardinals. With the new, meaner looking Cardinal, the team is 30-37 but with the more gentle Cardinal, the team was 96-178.

a) How much better have the Cardinals *Performed* with the new logo? Calculate the difference in winning percentage with the new logo and with the old logo.
b) State the hypotheses we are interested in testing.
c) A 100-trial simulation was conducted assuming that the *Ability* of the Cardinals is the same with either logo. The difference between the winning percentage with the new logo and with the old logo was calculated for each trial and the results are below. Based on these results, estimate the *p*-value and make an appropriate conclusion.



d) If a student concluded that the *Ability* is higher with the new logo, would it be reasonable to conclude that the new logo is the cause of the increase in the *Ability*? Explain.

For Investigation:

1. Using a website such as www.baseball-reference.com, use the methods of this Unit to investigate if a certain player or team performs better (or worse) in particular contexts (home vs. away, day vs. night, dark uniform vs. white uniform, early in game vs. late in game, etc.). Simply enter a players name in the search box. Clicking on "splits" will break down the players/teams statistics into many subcategories.

Other similar websites include:

- www.pro-football-reference.com
- www.basketball-reference.com
- www.hockey-reference.com
- www.sports-reference.com/olympics/

2. Collect data about your own athletic performance (or a team's performance at your school) and investigate is you have a higher ability in certain circumstances (at home, after eating Chinese food, during the day, etc.)