Josh Hauser

Identifying Patterns

Little

February 26, 2010

Streaks

Joe DiMaggio made history in the year of 1941. Joe had done something that no one in baseball had ever done before. In this year, Joe had 56 consecutive games with getting at least one hit. Many people say that this record will never be broken. Joe DiMaggio was one of the greats in baseball, but was he the greatest hitter there ever was? Many people argue that this was bound to happen over time and DiMaggio was just the lucky one who ended up getting the record. Can these crazy streaks be explained mathematically? Leonard Mlodinow summarized his thoughts about streaks in his article called “The Triumph of the Random From banking to baseball, winning streaks owe much to the laws of chance.” This article argues that over time there is bound to be a person who has a ridiculous season, and this ridiculous season will be averaged out by other seasons. The mathematics that the article provides says that 42 percent of the simulated histories had streaks of Joe DiMaggio’s length or longer. The longest of all of these streaks was said to have lasted 109 games. The simulated histories also had many other players that had the longest streak. Ty Cobb held it nearly 300 times in the simulated histories. Joe DiMaggio was definitely amazing baseball player, but was this streak just do to raw talent or was there some randomness involved in the whole process?

In Mlodinow’s article he explains streaks very generally. He compares it to flipping a coin. Leonard said that Joe DiMaggio had more than a 75 percent chance of getting a hit every game. Mlodinow shows an example on how this is relevant. He computes that if a coin were to be weighted to where one side had a 75 percent chance of being landed (which would represent a hit) and a side that had 25 percent chance of being landed on (which is anything other than a hit). With a coin weighted this way there would be streaks of 10 or more hits 10 percent of the time. Leonard makes a very good point at this part of his writing. He says that if someone were to have a streak of ten in a row in flipping coins then no one would call this talent. Yet everyone says it is tremendous talent when someone goes on a streak like this in baseball, basketball, or any other sport. This is just due to chance alone, so if other factors were at play then this streak does not seem completely unbreakable. Many extremely talented players compete to break this streak every year. Some of these players are even more talented than DiMaggio was hitting wise. The fact that all of these players, who have tremendous talent, compete to break this record shows that the streak that DiMaggio made could have had a lot to do with chance and randomness.

Another argument that Leonard brings up is from “The hot hand in basketball: On the misperception of random sequences.” This brings up the fact that everyone feels like they are “in the zone” at some point. Everyone has had that feeling that they are unstoppable in their sport. Whether it is not being able to miss a shot in basketball or being able to hit off everyone in baseball. People believe that this is an increase in their ability, but “being in the zone” usually ends pretty quickly. There are also cold streaks in all sports, whether it is a hitting slump in baseball or not being able to hit a shot in basketball. The article says that these hot and cold streaks are not even real; they are just the results you would get from flipping the appropriately weighted coin for that particular person. There is no doubt that Joe DiMaggio was “in the zone” at this time, but many other factors related to chance had to work out for this record to come into existence.

Leonard brings up another very interesting point. He says that one of the reasons that we may believe things like this aren’t due to randomness is our need for control. We want to believe that Joe DiMaggio did this from pure struggle, talent, and effort. Humans do not want to believe that we may not have control over all of the things in our environment.

The chances of this also could be measured by using a box model. Joe DiMaggio’s batting average over his career was .325. We’ll just round that to one third. His chance of getting a hit in a game would come from a box with one ticket having the value of one and two other tickets with the value of zero. The number of draws from the box would be the number of at bats that DiMaggio would get in a given game. If Joe were to have 4 at bats a game then there would be a very good chance that he would get a hit each game.

There are many different aspects of randomness that make this streak very hard to figure out. The main part is figuring out the likelihood of Joe getting a hit per game. It is tough to choose which batting average to use. Should we use the one for his whole career, his season, or even certain games? This aspect, along with walks and his number of at bats per game, make it very hard to find a specific formula for measuring the likelihood that Joe DiMaggio would get this streak.

The fact is that raw talent can only take you so far. There are numerous things that could have happened that could have thrown this record off path so easily. What if one day the pitcher that was throwing couldn’t throw a strike and walked DiMaggio four times? This would make the record nonexistent. What if the first time DiMaggio hit the ball he sprained his ankle running to first base? He would have to sit out the rest of the game, and maybe even the next, and would not achieve his record. There are even situations that do not even have to do with Joe’s performance. What if his team is struggling at the plate and getting many outs? DiMaggio may not get his opportunity to get his normal number of at bats and not have a chance to get a hit that particular game. There are so many things that had to come in line. Some of these things include pitches, weather, different situations in the game, and even personal things that were going on in his life that may have affected his mood. The truth is that DiMaggio could have been completely honed in on his skills and could be able to hit anything, but still not be able to produce a hit in the game. These are all confounders in this scenario. These are some of the reasons why the coin flipping experiment and the box model experiment don’t work. These things cannot be assigned a certain number to be plugged into the formula in order to have the perfect statistic. There are so many things that had to come together and align in order for this streak to happen. This record that Joe DiMaggio made had to do with chance and randomness as much as it had to do with his skill in the game.

I personally think that this streak was bound to happen over time and it was much more likely to happen in the past. With 30 teams in the Major League there is bound to be one person who has a standout streak over time. A great number of players have really good chances of getting hits each game. I do think that this streak was much easier to attain in the earlier years of baseball. I think that the game has gotten a lot harder and pitchers become a lot smarter. I believe if someone were to be hitting like Joe DiMaggio was back in the day then a pitcher would do anything in his power to not let him get a hit. I think the pitcher would not give him any good pitches to hit, even if it meant a walk. A person in this generation would have to be a ridiculous hitter and have everything going for him in order to attain this record. It may seem very hard these days, but with the talent that is in the Major League right now and with the help of a little randomness this could be broken.

There is no doubt in anyone’s mind that Joe DiMaggio was an amazing baseball player, but there were many different things that had to happen in order for him to get this streak. His talent really helped him have to opportunity to be able to pull it off. There can be an estimate on the randomness of this streak to occur, but there is no formula in order to figure out the exact probability. There would have to be a part of the formula to account for weather, the other team’s pitcher, number of at bats, and many other aspects of the game in order for there to be a legitimate formula for the probability. Joe DiMaggio does deserve credit for his wonderful feat, but this record was bound to be broken over time by some player due to chances alone. Joe was just one of the great players who was put in a great situation and came out a legend and will be remembered forever for what he did.