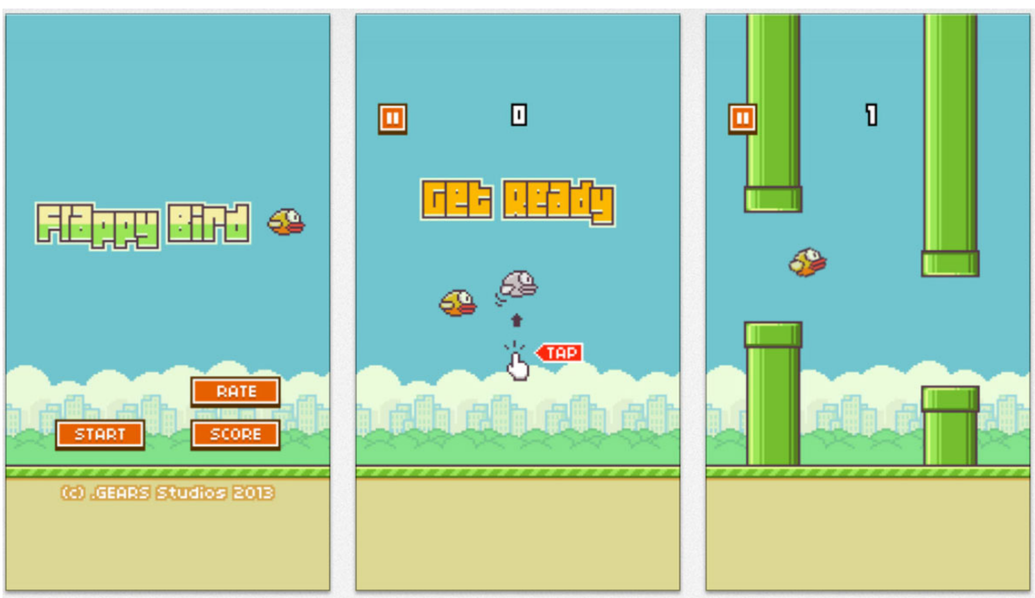
Flappy Bird

Albert Kuo
Anthony Pan

February 20, 2016

Abstract

Albert and Anthony played Flappy Bird and calculated some means, made some plots, and wrote this report.



1 Introduction

In early 2014, an iOS game, "Flappy Bird," became a viral hit[1], making its way to our phones. At that time, we were also taking statistics together with Professor Stein. Inspired by the class, we decided to do a statistical analysis of the game we were playing so much of anyway.

Thus, the boundaries of work and play were blurred, paralleling the development of our myopia from staring at our phone screens so intensely. But it was all worth it[2]. We developed very specialized skills in tapping our phone screens at just the right frequency and, as a side benefit, learned how to use R. The results of our work follow in the remainder of this paper.

Flappy Bird In case you don't already know what Flappy Bird is, here is a short description. The objective of the game is to get a bird through as many tubes as possible. You control the bird by tapping the phone screen. The bird flies up when you tap the screen and drops down when you don't do anything. To get the bird through a tube, you will need to tap the phone screen at the right time, which changes depending on how high or low the tubes are. If the bird hits a tube, the ceiling, or the ground, it dies. The number of tubes you pass through is your score. A video of Flappy Bird gameplay can be found at this link.

Version II You are reading the second version of this report. The first version was completed in March 2014. When Anthony looked at the original report a year later, however, he was so appalled by what he saw that he vowed to redo the report. Together, we worked to revamp the report, producing better plots, analysis, and just about everything else. You should expect a third version of this report, to be done when we are even older, wiser, and more able to answer the questions we raised in this version.

[1] So popular was Flappy Bird that the developer, believing that the game was too addictive, took down the game from the App Store. Luckily, we downloaded the game before it was taken down, allowing our enjoyment of the game and this project to continue unabated.

[2] Anthony has not actually expressed this sentiment out loud, but Albert is taking some liberty in speaking for both of them.

2 Data

2.1 Data collection

We recorded the score of every game of Flappy Bird played over the course of about two months, from January to March 2014. Since Albert had never played Flappy Bird before, he was able to record every single game of Flappy Bird he has ever played in his life up to that point. In total, he recorded 800 games of Flappy Bird. Anthony, on the other hand, had been playing Flappy Bird for a few weeks prior to when he began recording his scores. He recorded 300 games of Flappy Bird.

2.2 Exploratory Analysis

Below are scatter plots of both players' scores, with the game number along the x-axis and the score along the y-axis. For Albert's scores, new high scores are marked as a red diamond. For Anthony's scores, his overall high score is marked as a red diamond.

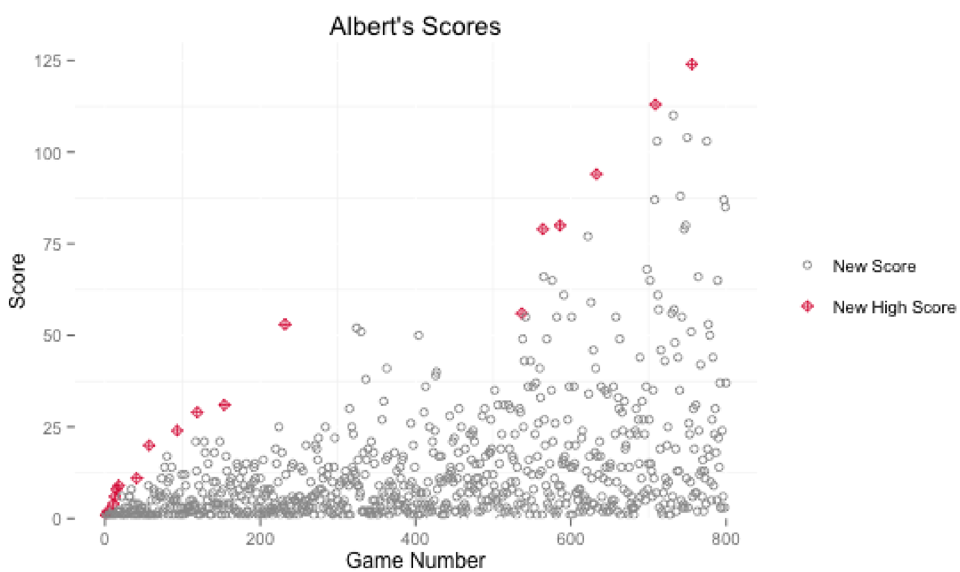


Figure 1: Scatter plot of Albert's Scores

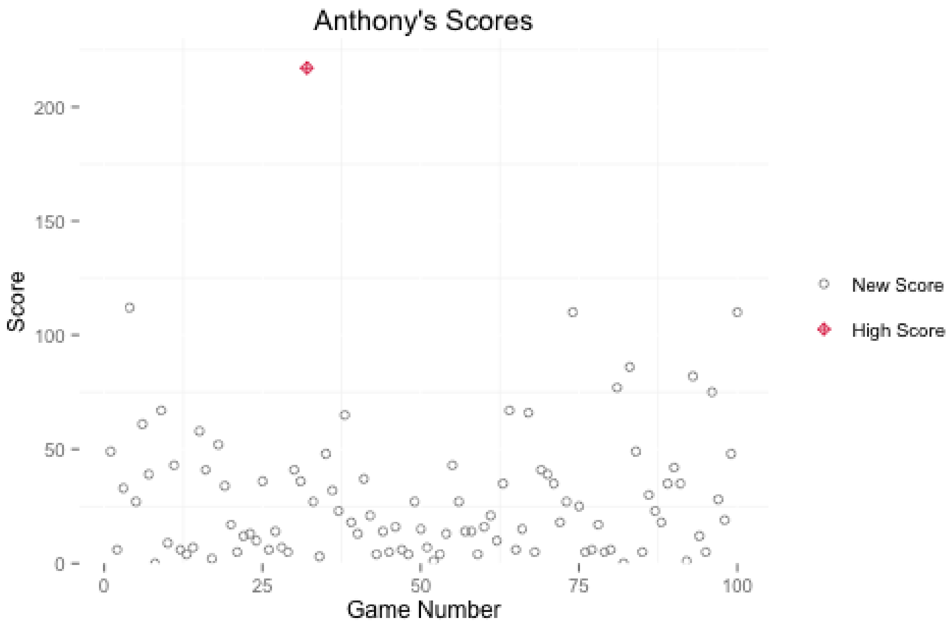


Figure 2: Scatter plot of Anthony's Scores

From the scatter plots, a few things already appear strange. For instance, in Albert's scores, there is a long stretch of about 300 games when he wasn't able to break his high score. Three times, he came close, but he failed each time. In Anthony's scores, his high score of 217 is far above the rest of his scores, almost two times his next highest score of 112. We will investigate these quirks and other observations in the remainder of this report.

3 Analysis

3.1 Bernoulli Process

As we were learning about the Bernoulli process when we were collecting our data, we theorized that the game follows a Bernoulli process quite well. We can denote the probability that the bird flies through each tube as $1 - p$. Each tube is a trial and can be considered as independent and identically distributed (i.i.d). That is, the difficulty of passing through each tube is independent of passing through another. This may not be completely true, since it is likely that one tube has an impact on how difficult the next tube is to fly through. However, since the probability that the bird flies through the tube still depends almost entirely on the player's skill, this is not a terribly inaccurate assumption to make.

The score is the number of tubes the bird successfully flies through. In other words, it is the number of trials until a failure occurs, so it can be modeled by a geometric distribution. We have plotted a histogram of our scores along with a theoretical geometric distribution density, using the average as maximum likelihood estimate parameter ($p = 1/\text{mean score}$). For Albert's scores, the theoretical geometric density underpredicts for low scores (< 10) and overpredicts for higher scores. This could be a reflection of the fact that one of the assumptions for our model is violated, which is that the probability p for Albert is changing significantly with each game. Since Albert is getting better at the game each time, p, the probability that the bird does not pass through a tube, is decreasing with each game.

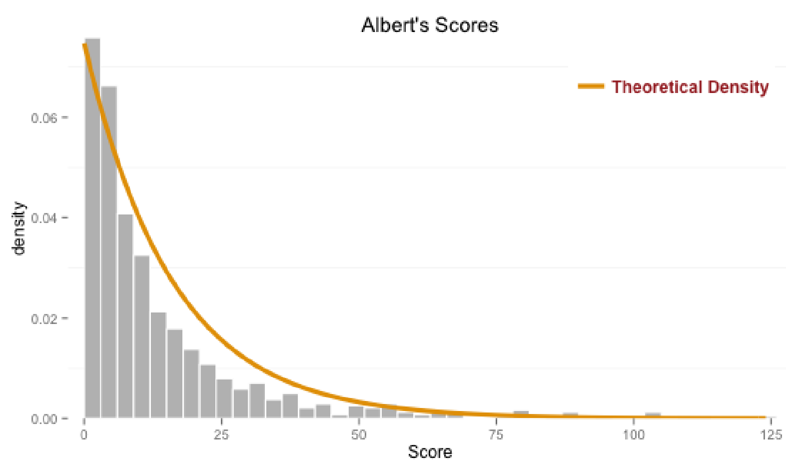


Figure 3: Histogram with Estimated Density of Albert's Scores

However, when we apply the same hypothesis of a geometric distribution to Anthony's scores, the same problem appears. Namely, the theoretical geometric density underpredicts for low scores and overpredicts for higher scores. The pattern is less clear in Anthony's scores, which is noisier, likely due to the smaller sample size.

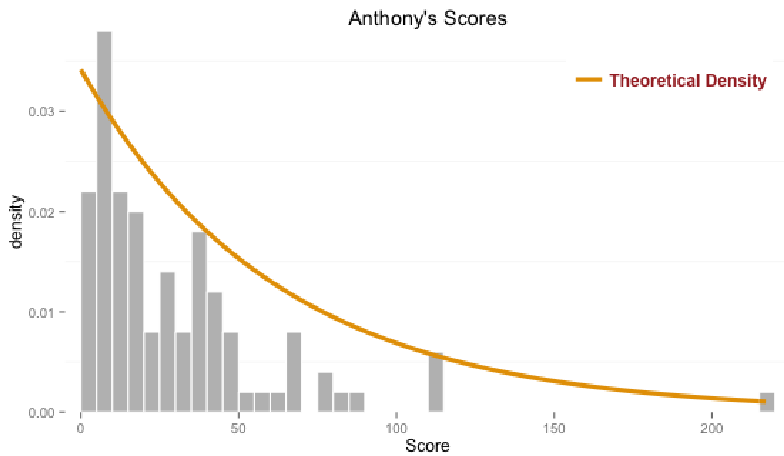


Figure 4: Histogram with Estimated Density of Anthony's Scores

As we discussed earlier, Anthony's high score is an outlier. The removal of this high score does result in a better fit to the data using the theoretical geometric density, as we can see in Figure 5. But there is no valid reason for why we should remove it and so we are still left with this puzzling question: why do Anthony's scores not follow a geometric distribution?

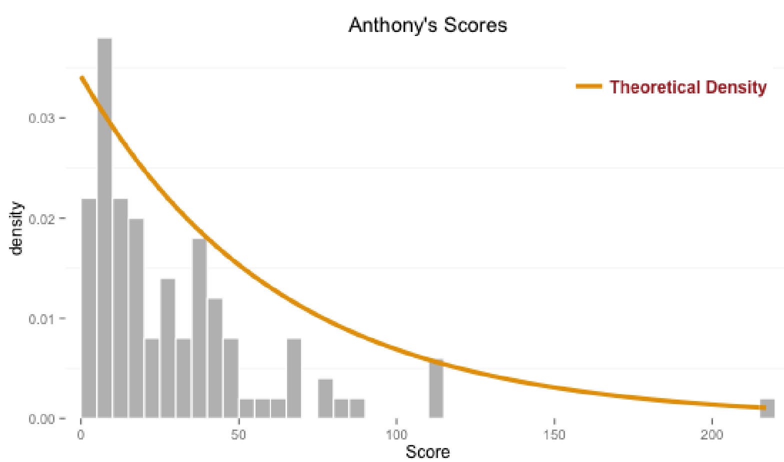


Figure 5: Histogram with Estimated Density of Anthony's Scores (Without High Score)

3.2 New High Scores

There is a cluster of lighter-colored points at the bottom left of the scatter plot. This tells us that the increases in high score are lower and more frequent when Albert is learning how to play the game, as one might expect. There is one extreme outlier: Albert played over 300 games without a new high score and when he did beat his high score, he improved by less than 5 points. As mentioned in the beginning of this report, this can also be seen in Figure 1, since there were no new high scores between Game 200-something and Game 500-something, although it is less obvious on that plot. The question of why this is will be answered in the next iteration of this report, where we will examine the learning curve of this game.

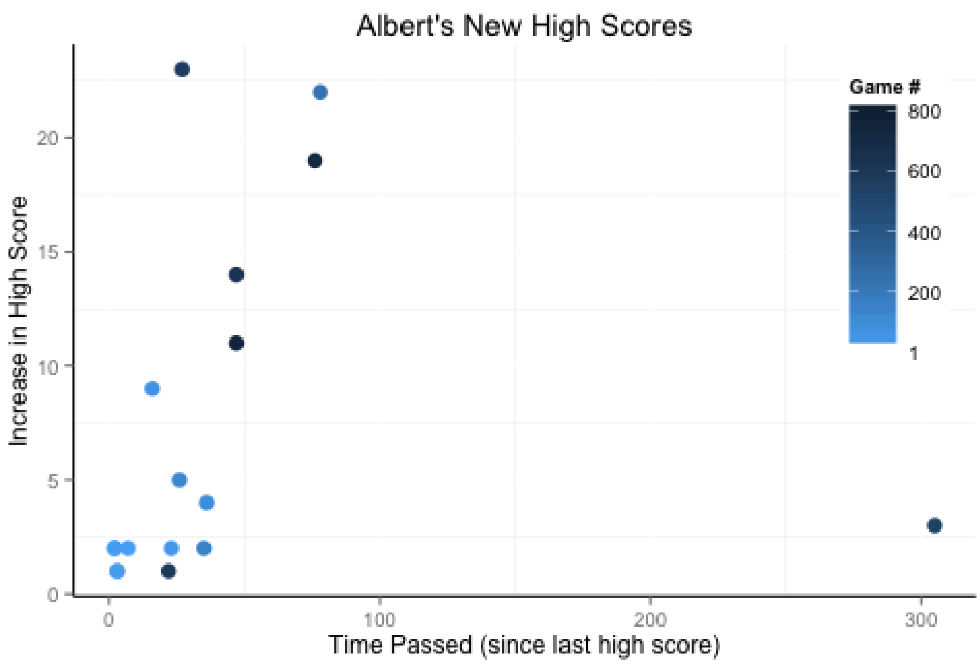


Figure 6: Scatter plot of Albert's New High Scores