Playing Smart: Statistics of the Carolina Pick 3 Lottery

ST 380, Fall 2009

Allison Camras
Michelle Cipullo
Megan Embrey
Cassie Mentha
Executive Summary

The purpose of the study was to analyze lottery data from the North Carolina Education Lottery website to determine if the differences in the frequencies of digits drawn were statistically significant. The study focused on two null hypotheses: the probability of drawing any of the digits is 0.1 and that each draw is statistically independent of each other. Several tests were performed on the data to establish frequency relationships on each of the digits drawn in each slot over many parameters. These tests/methods included histogram graphs, probability comparisons, probability density graphs and chi-squared tests. From these methods, p-values were obtained to determine statistical significance and statistical independence. The analysis of the data led to some key findings. High p-values (>0.05) were found from the cumulative lottery data for each ball slot which confirmed the first null hypothesis. The high p-values also confirmed the second null hypothesis when comparing the daytime/evening lottery data over the course of the lottery history. From the acceptance of the null hypotheses, each digit drawn is statistically independent of another digit drawn and the probabilities of each of the digits being drawn is the same. The frequency of each ball was analyzed to use in actual play of the lottery. The goal of the lottery play was to analyze all data in the various parameters (e.g. by year, by month, daytime/evening) to get the most frequently drawn digits to play for each slot. This was then compared to randomly selected digits that were chosen without analysis. Although the sample size for the randomly selected digits was small, in comparison with the analyzed digits, the analyzed digits appeared to have more common digits with the winning number. Even though different frequencies were found for each of the digits, the differences were not large enough to be considered statistically significant.
Table of Contents

Description of Data...........................................................................................................4
Methods and Results........................................................................................................5
Statistical Analysis............................................................................................................7
Major Findings..................................................................................................................8
Discussion.........................................................................................................................9
Appendix A
  2006 Evenings.............................................................................................................A1
  2007 Evenings.............................................................................................................A2
  2008 Monthly Data.....................................................................................................A3
  2009 Monthly Data.....................................................................................................A4
  Cumulative Evening Data..........................................................................................A5
  Cumulative Years by Month Data .............................................................................A6
  Weekdays/Weekend Data.........................................................................................A7
  Daily Data...................................................................................................................A8
  Total Lottery History.................................................................................................A9
Appendix B
  Lottery Number Analysis.........................................................................................B1
Appendix C
  Probability of Data.................................................................................................C1
Appendix D
  Frequencies of Total Data.......................................................................................D1
  Frequencies of Data by Year....................................................................................D2
  Frequencies of Data by Evening..............................................................................D3
Appendix E
  November Data and Analysis..................................................................................E1
Description of Data

The Carolina Pick 3 Lottery was started in North Carolina in 2006. The way you play is by first picking 3 digits between 0 and 9 (or any number between 000 to 999), or having a computer randomly pick your lottery digits (known as the “Quick Pick”). Next, you choose which version you would like to play.

The first way is **Exact**, which means your selected digits must be in exactly the right order; the odds of winning this way are 1 in 1,000. The next option is to do **Any Order**, and this way you must have the correct digits, but they can be in any order. If by chance your digits end up in exact order, you still only win the “any order” amount with this option. Within every option except the exact order, you have a subset of an option of 3-way or 6-way digits. The 3-way digits mean that you have a number with two of the same digits (ex. 332). Because of this there are only 3 ways to combine those digits, therefore, the probability of winning is less likely. In a 6-way option you have a number that has 3 different digits, and because of this there are 6 possible ways the number can be drawn. A 3-way draw in “any order” odds are 1 in 333.33, a 6-way draw in any order is 1 in 166.67.

The next option is **50/50 (Exact/Any Order)**, this is a combination of the first two; exact and any order. The 50/50 option allows you to pick 3 digits, and if they show up in exact order the payoff is greater. However, if they show up in any order you still win, but the payoff is smaller. The odds with a 3-way pick for exact is 1 in 1000, and for 6-way 1 in 333.33, for any order the odds for a 3 way are 1 in 333.33 and for a 6 way the odds are 166.67.

The final option for play is the **Combo Play**. The combo play allows you to win if you have the same 3 numbers in any order, and will still give you the pay out of the exact order option. The same pay out is possible because these tickets are more expensive than the exact order tickets. The prices of these tickets vary by the amount of possible winnings, costing more for a higher pay out. With the Exact, Any Order, and 50/50(Exact/Any Order) the tickets cost $0.50 or $1.00 per play. Choosing the 3-way option will yield a higher pay out because the probability of winning is less likely, whereas the 6-way yields a lower pay out, but you have a better chance of winning.

A condensed overview of the options is seen in the figure below (courtesy of the NCEL):
Methods and Results

The lottery is a game of chance, and one would expect to find a uniform distribution if drawings were analyzed statistically. The drawing data was analyzed over the course of the NC Pick 3 Lottery history to find non-uniform distributions and to see if one digit was drawn more frequently than others in each ball slot. The 50/50 game playing option was used to show results of getting the digits in exact order, as well as if the digits appeared in any order. With the data, the goal was to reject the null hypothesis that the probability of drawing one digit was equal to another digit. To further explore this we analyzed each ball drawing separately over various parameters.

To carry out our experiment, we downloaded all data from the start of the pick 3 lottery and then broke it down into several groups to provide a better view of the data. The following are the groups analyzed:

- Cumulative all numbers drawn
- Cumulative yearly
- Cumulative monthly
- Cumulative daily/evening
- Yearly by Month (daytime/evening)
- In-depth November Analysis

In-depth daytime drawings were not analyzed, as less data was available. The daytime drawings started on March 31, 2008, whereas evening drawings have existed since the start of the Carolina Pick 3 on October 6, 2006. For the in-depth analysis of November we looked at the
cumulative November drawings, November evenings, just 2009 for November, and the weekday evening draws for each day in November. November was a good candidate for an in-depth analysis because we have data from 2006 to the present. Since it is currently the month of November, we have been able to put these numbers into play from Friday, November 20 to Saturday, November 28.

Once the data was downloaded and formatted into excel spreadsheets, it was divided into the groups mentioned above. Next, R code was created to display a histogram showing the distribution for each ball drawn (See appendix A). The graphs were then analyzed and one (or more) digits were chosen based upon frequency to create a number (or numbers) for lottery play for each category. This became the set of numbers that was found to be the most likely candidate to be drawn. If one were to play the lottery, this is the information that would be useful to consider when choosing a number to play.

To put our theory into practice, we started playing numbers based on this data. For each day played the following number sets were used:

- Most common cumulative number
- Most common number(s) for the current month
- Most common number(s) for the year played
- Most common number(s) for the time of the drawing by month and day of the week (daytime/evening)
- Most common number(s) of “all evenings” or “all days”
- One random draw ("Quick Pick")

In analyzing the 83 tickets purchased over the lottery play, the following results were obtained:

- 53 chosen numbers and 4 random numbers contained at least one correct digit.
- 13 chosen numbers and 2 random numbers contained at least two correct digits.
- 1 chosen number and 0 random numbers contained all three correct digits.
- None of the chosen or random numbers had all three digits in the correct order.

(A complete analysis of played numbers can be viewed in Appendix B)
Statistical Analysis

The data was analyzed using several statistical methods: histogram graphs, probability comparisons, probability density graphs, and chi-squared tests. The methods were tested on three categories of data: the entire lottery history, all data grouped by year (2006-2009), and all data grouped by time (day or evening). The first two analyses – histograms and probability comparisons – gave an idea of the spread and depth of the data, but did not allow a true statistical analysis. These tables and graphs can be viewed in Appendix A and Appendix C. The graphs for the histograms were used to determine the frequencies of the numbers drawn (the data used in the picking of the lottery numbers to play). The other two analyses provided a more analytical view of the data. Chi-squared tests were performed on the data, and all p-values were found to be well above the 0.05 threshold used to determine statistical significance. From this, it was determined that, while the data originally appeared to have non-uniform densities, the frequencies for the balls were not statistically different. Therefore, the null hypothesis could not be rejected; the probability of any ball being selected is equal to that of any other ball being selected. Although different frequencies and probabilities among the selected balls were found, there was no statistical significance to the difference.

The first chi-squared tests involved the null hypothesis that each ball would have a frequency that provided a probability of 0.1. It was found that, in almost every case, probabilities were not equal to 0.1, but p-values determined by the tests on each ball suggest that these differences were not significant (see Appendix D for more graphs). This was true in all cases – the entire lottery history, each year, and each time.

An example showing comparison graphs between daytime and evening draws on ball three is located below:
The daytime and evening graphs above take into account the entire lottery history for data on ball three. Density lines are plotted on the graphs to represent the probability of each ball drawn for daytime and evening draws. As shown, the probabilities are close to 0.1 for each ball drawn. The p-value for the daytime draws was calculated to be 0.5602 and the p-value for the evening draw was 0.8792. The p-values, being above 0.05, show that the numbers drawn are statistically independent of each other. This also shows that there is no statistical significance concerning the frequency of the balls being drawn and that each ball drawn is independent of another ball drawn.

Concerning the density graphs, a correlation was shown between probability density and p-values. Graphs that demonstrated a more uniform density curve were found to have p-values closer to 1. This is expected because of the more uniform distribution of the numbers drawn for that particular data set. Those graphs that showed greater variation in frequencies had lower corresponding p-values, although none of these values were below the 0.05 threshold. This again shows statistical independence of the balls drawn for play (view graphs above).

**Major Findings**

The analysis led to some key findings. First, it was found that the null hypothesis that each draw is statistically independent of one another must be accepted. This means that a draw in one slot will not have an affect on any other draws. This is to be expected since each slot is drawn from separate containers, each having digits 0 through 9. Next it was found that each digit had a slightly different probability of being drawn in each parameter set, however, this difference is not seen as statistically significant. Statistical significance was ruled out after performing the chi-squared test on the data which gave p-values greater than 0.05. Since the p-value was greater than 0.05 the second null hypothesis was also accepted, that each digit has the same probability of being drawn. Again, since the lottery is a random draw, this would be expected.

Along with statistically backed findings, there were also some interesting finds that had no statistical significance, but stood out. The in-depth analysis of the November weekdays showed that some numbers were never drawn, for example on Fridays a 5, 6 and 7 have never been drawn for evening ball threes. The number drawn the most in any digit of the number over all of the drawings is 1, but it was never in a winning number during the duration of our play.
Many of the graphs showed what looked like a non-uniform distribution, even though the p-value specified otherwise (See Appendix E for a more in-depth analysis). With more data came more uniform distribution (as with daytime vs. evening), yet with the comparison of the cumulative data for all three ball slots, ball three showed the most uniform distribution and highest p-value even though the amount of data was equal to the other ball slots. The p-value for ball one was 0.18, while p-value for ball two was 0.21 and the p-value for ball three was 0.98, with higher values correlating to a more uniform distribution.

The non-cumulative data, especially in months like November, provided a more stratified graph allowing a more decisive choice of a number set. Some graphs had what appeared to be a normal distribution, while some were either left or right skewed. The number selections that have proven most favorable over the course of our trial have been November cumulative data. Because of these results, it can be concluded that the probability of picking a winning number is more favorable if the number is chosen from the analysis rather than at random. There is still luck involved, but an educated “guess” can be made when picking commonly drawn numbers. Again, this is not statistically proven, but in the experiment it did show that the chosen numbers gave better results than the randomly selected numbers.

Discussion

The numbers chosen for lottery play were a reflection of the data provided by the NCEL for the Carolina Pick 3. Since the lottery is fairly new to North Carolina, there was significantly less data compared to other states. In other words, if these statistics had been performed on data from the Virginia Lottery (that has been played since 1989), different results may have been achieved. In addition, this study was not concerned with the frequencies of the three-digit numbers appearing in the lottery draws, but rather focused on individual digits; each digit was analyzed as a separate entity within the number (for example, the number 173 vs. the digits 1-7-3). Had we analyzed the frequency of numbers instead of digits, the results may not have provided us with discernable playing options (in other words, no one number may have had a significantly higher frequency of draw than another). Another limitation was the lack of results from random draws - only one number per day was a Quick Pick random draw. To thoroughly examine the efficacy of a chosen draw vs. a random draw, equal amounts of random and chosen numbers should have been played in order to make the comparisons fair. One final limitation of the study was time; a limited amount of time existed to examine the statistical results of the
lottery analysis and translate those results into playable numbers. A longer, more in-depth study could account for more categories of draws, as well as other lottery games.

Several questions were raised during the course of the study that could be examined through further research. Could this strategy of analysis be applied to other lottery games, such as the Carolina Pick 4 or Powerball, or would the addition of additional digits have no effect on individual ball draws? Would it be more beneficial, in the situation of the Carolina Pick 3, to focus on complete numbers rather than digits? The most important question, however, concerns the uniform distribution of draws. When examining specific digit draws, will a larger sample size provide the analyst with a more uniform distribution, or a more specific outlier? If this was investigated using lottery data from another state (such as Virginia), would different results be presented?

In conclusion, the main focus of the study was not to predict lottery draws, but analyze the distributions of draws to determine if outliers existed. While the differences in the frequencies of digits were not statistically significant, there were digits that appeared more frequently than others. Though this is possible, the probabilities of the digits being drawn are still equal and there is no statistical significance. The digits drawn are statistically independent of each other which support the two null hypotheses. Although one cannot accurately predict numbers, the experiment provided an outlet to other questions and theories that could be analyzed in subsequent experiments.
Appendix A-1
2006 Evenings

October, Ball One

October, Ball Two

October, Ball Three

November, Ball One

November, Ball Two

November, Ball Three

December, Ball One

December, Ball Two

December, Ball Three
oct=read.table("clipboard",header=T)
> octone=oct$One
> octtwo=oct$Two
> octthree=oct$Three
> hist(octone,breaks=seq(-1,9,by=1),main="October, Ball One",xlab="Number Drawn")
> hist(octtwo,breaks=seq(-1,9,by=1),main="October, Ball Two",xlab="Number Drawn")
> hist(octthree,breaks=seq(-1,9,by=1),main="October, Ball Three",xlab="Number Drawn")

nov=read.table("clipboard",header=T)
> novone=nov$One
> novtwo=nov$Two
> novthree=nov$Three
> par(mfrow=c(1,3))
> hist(novone,breaks=seq(-1,9,by=1),main="November, Ball One",xlab="Number Drawn")
> hist(novtwo,breaks=seq(-1,9,by=1),main="November, Ball Two",xlab="Number Drawn")
> hist(novthree,breaks=seq(-1,9,by=1),main="November, Ball Three",xlab="Number Drawn")

Dec=read.table("clipboard",header=T)
> decone=dec$One
> dectwo=dec$Two
> decthree=dec$Three
> hist(decone,breaks=seq(-1,9,by=1),main="December, Ball One",xlab="Number Drawn")
> hist(dectwo,breaks=seq(-1,9,by=1),main="December, Ball Two",xlab="Number Drawn")
> hist(decthree,breaks=seq(-1,9,by=1),main="December, Ball Three",xlab="Number Drawn")
Jan2007=read.table("clipboard",header=T)
Jan2007BallOne=Jan2007$One
Jan2007BallTwo=Jan2007$Two
Jan2007BallThree=Jan2007$Three
par(mfrow=c(1,3))
hist(Jan2007BallOne, breaks=seq(-1.9,by=1), main="January 2007, Ball One", xlab="Number Drawn")
hist(Jan2007BallTwo, breaks=seq(-1.9,by=1), main="January 2007, Ball Two", xlab="Number Drawn")
hist(Jan2007BallThree, breaks=seq(-1.9,by=1), main="January 2007, Ball Three", xlab="Number Drawn")

Feb2007=read.table("clipboard",header=T)
Feb2007BallOne=Feb2007$One
Feb2007BallTwo=Feb2007$Two
Feb2007BallThree=Feb2007$Three
par(mfrow=c(1,3))
hist(Feb2007BallOne, breaks=seq(-1.9,by=1), main="February 2007, Ball One", xlab="Number Drawn")
hist(Feb2007BallTwo, breaks=seq(-1.9,by=1), main="February 2007, Ball Two", xlab="Number Drawn")
hist(Feb2007BallThree, breaks=seq(-1.9,by=1), main="February 2007, Ball Three", xlab="Number Drawn")

Mar2007=read.table("clipboard",header=T)
Mar2007BallOne=Mar2007$One
Mar2007BallTwo=Mar2007$Two
Mar2007BallThree=Mar2007$Three
par(mfrow=c(1,3))
hist(Mar2007BallOne, breaks=seq(-1.9,by=1), main="March 2007, Ball One", xlab="Number Drawn")
hist(Mar2007BallTwo, breaks=seq(-1.9,by=1), main="March 2007, Ball Two", xlab="Number Drawn")
hist(Mar2007BallThree, breaks=seq(-1.9,by=1), main="March 2007, Ball Three", xlab="Number Drawn")

Apr2007=read.table("clipboard",header=T)
Apr2007BallOne=Apr2007$One
Apr2007BallTwo=Apr2007$Two
Apr2007BallThree=Apr2007$Three
par(mfrow=c(1,3))
hist(Apr2007BallOne, breaks=seq(-1.9,by=1), main="April 2007, Ball One", xlab="Number Drawn")
hist(Apr2007BallTwo, breaks=seq(-1.9,by=1), main="April 2007, Ball Two", xlab="Number Drawn")
hist(Apr2007BallThree, breaks=seq(-1.9,by=1), main="April 2007, Ball Three", xlab="Number Drawn")

May2007=read.table("clipboard",header=T)
May2007BallOne=May2007$One
May2007BallTwo=May2007$Two
May2007BallThree=May2007$Three
par(mfrow=c(1,3))
hist(May2007BallOne, breaks=seq(-1.9,by=1), main="May 2007, Ball One", xlab="Number Drawn")
hist(May2007BallTwo, breaks=seq(-1.9,by=1), main="May 2007, Ball Two", xlab="Number Drawn")
hist(May2007BallThree, breaks=seq(-1.9,by=1), main="May 2007, Ball Three", xlab="Number Drawn")

Jun2007=read.table("clipboard",header=T)
Jun2007BallOne=Jun2007$One
Jun2007BallTwo=Jun2007$Two
Jun2007BallThree=Jun2007$Three
par(mfrow=c(1,3))
hist(Jun2007BallOne, breaks=seq(-1.9,by=1), main="June 2007, Ball One", xlab="Number Drawn")
hist(Jun2007BallTwo, breaks=seq(-1.9,by=1), main="June 2007, Ball Two", xlab="Number Drawn")
hist(Jun2007BallThree, breaks=seq(-1.9,by=1), main="June 2007, Ball Three", xlab="Number Drawn")

Jul2007=read.table("clipboard",header=T)
Jul2007BallOne=Jul2007$One
Jul2007BallTwo=Jul2007$Two
Jul2007BallThree=Jul2007$Three
par(mfrow=c(1,3))
hist(Jul2007BallOne, breaks=seq(-1,9,by=1), main="July 2007, Ball One", xlab="Number Drawn")
hist(Jul2007BallTwo, breaks=seq(-1,9,by=1), main="July 2007, Ball Two", xlab="Number Drawn")
hist(Jul2007BallThree, breaks=seq(-1,9,by=1), main="July 2007, Ball Three", xlab="Number Drawn")

Aug2007=read.table("clipboard",header=T)
Aug2007BallOne=Aug2007$One
Aug2007BallTwo=Aug2007$Two
Aug2007BallThree=Aug2007$Three
par(mfrow=c(1,3))
hist(Aug2007BallOne, breaks=seq(-1,9,by=1), main="August 2007, Ball One", xlab="Number Drawn")
hist(Aug2007BallTwo, breaks=seq(-1,9,by=1), main="August 2007, Ball Two", xlab="Number Drawn")
hist(Aug2007BallThree, breaks=seq(-1,9,by=1), main="August 2007, Ball Three", xlab="Number Drawn")

Sep2007=read.table("clipboard",header=T)
Sep2007BallOne=Sep2007$One
Sep2007BallTwo=Sep2007$Two
Sep2007BallThree=Sep2007$Three
par(mfrow=c(1,3))
hist(Sep2007BallOne, breaks=seq(-1,9,by=1), main="September 2007, Ball One", xlab="Number Drawn")
hist(Sep2007BallTwo, breaks=seq(-1,9,by=1), main="September 2007, Ball Two", xlab="Number Drawn")
hist(Sep2007BallThree, breaks=seq(-1,9,by=1), main="September 2007, Ball Three", xlab="Number Drawn")

Oct2007=read.table("clipboard",header=T)
Oct2007BallOne=Oct2007$One
Oct2007BallTwo=Oct2007$Two
Oct2007BallThree=Oct2007$Three
par(mfrow=c(1,3))
hist(Oct2007BallOne, breaks=seq(-1,9,by=1), main="October 2007, Ball One", xlab="Number Drawn")
hist(Oct2007BallTwo, breaks=seq(-1,9,by=1), main="October 2007, Ball Two", xlab="Number Drawn")
hist(Oct2007BallThree, breaks=seq(-1,9,by=1), main="October 2007, Ball Three", xlab="Number Drawn")

Nov2007=read.table("clipboard",header=T)
Nov2007BallOne=Nov2007$One
Nov2007BallTwo=Nov2007$Two
Nov2007BallThree=Nov2007$Three
par(mfrow=c(1,3))
hist(Nov2007BallOne, breaks=seq(-1,9,by=1), main="November 2007, Ball One", xlab="Number Drawn")
hist(Nov2007BallTwo, breaks=seq(-1,9,by=1), main="November 2007, Ball Two", xlab="Number Drawn")
hist(Nov2007BallThree, breaks=seq(-1,9,by=1), main="November 2007, Ball Three", xlab="Number Drawn")

Dec2007=read.table("clipboard",header=T)
Dec2007BallOne=Dec2007$One
Dec2007BallTwo=Dec2007$Two
Dec2007BallThree=Dec2007$Three
par(mfrow=c(1,3))
hist(Dec2007BallOne, breaks=seq(-1,9,by=1), main="December 2007, Ball One", xlab="Number Drawn")
hist(Dec2007BallTwo, breaks=seq(-1,9,by=1), main="December 2007, Ball Two", xlab="Number Drawn")
hist(Dec2007BallThree, breaks=seq(-1,9,by=1), main="December 2007, Ball Three", xlab="Number Drawn")
All 2008
> All2008=read.csv("/Users/allisoncamras/Desktop/ST 380 Project/All2008.csv", header=T)
> print(All2008)
> All2008One=All2008$Number.1
> All2008Two=All2008$Number.2
> All2008Three=All2008$Number.3
> par(mfrow=c(1,3))
> hist(All2008One, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="All 2008 Ball One")
> hist(All2008Two, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="All 2008 Ball Two")
> hist(All2008Three, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="All 2008 Ball Three")

January 2008 Evening
> print(Jan2008Evening)
> Jan2008EveningOne=Jan2008Evening$Number.1
> Jan2008EveningTwo=Jan2008Evening$Number.2
> Jan2008EveningThree=Jan2008Evening$Number.3
> par(mfrow=c(1,3))
> hist(Jan2008EveningOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="January 2008 Evening Ball One")
> hist(Jan2008EveningTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="January 2008 Evening Ball Two")
> hist(Jan2008EveningThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="January 2008 Evening Ball Three")

February 2008 Evening
> print(Feb2008Evening)
> Feb2008EveningOne=Feb2008Evening$Number.1
> Feb2008EveningTwo=Feb2008Evening$Number.2
> Feb2008EveningThree=Feb2008Evening$Number.3
> par(mfrow=c(2,2))
> hist(Feb2008EveningOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="February 2008 Evening Ball One")
> hist(Feb2008EveningTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="February 2008 Evening Ball Two")
> hist(Feb2008EveningThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="February 2008 Evening Ball Three")

March 2008 Evening
> print(March2008Evening)
> March2008EveningOne=March2008Evening$Number.1
> March2008EveningTwo=March2008Evening$Number.2
> March2008EveningThree=March2008Evening$Number.3
> par(mfrow=c(1,3))
> hist(March2008EveningOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="March 2008 Evening Ball One")
> hist(March2008EveningTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="March 2008 Evening Ball Two")
> hist(March2008EveningThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="March 2008 Evening Ball Three")
April 2008 Day

```r
> April2008Day=read.csv("/Users/allisoncamras/Desktop/ST 380 Project/April2008Day.csv", header=T)
> print(April2008Day)
> April2008DayOne=April2008Day$Number.1
> April2008DayTwo=April2008Day$Number.2
> April2008DayThree=April2008Day$Number.3
> par(mfrow=c(1,3))
> hist(April2008DayOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="April 2008 Day Ball One")
> hist(April2008DayTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="April 2008 Day Ball Two")
> hist(April2008DayThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="April 2008 Day Ball Three")
```

April 2008 Evening

```r
> April2008Evening=read.csv("/Users/allisoncamras/Desktop/ST 380 Project/April2008Evening.csv", header=T)
> print(April2008Evening)
> April2008EveningOne=April2008Evening$Number.1
> April2008EveningTwo=April2008Evening$Number.2
> April2008EveningThree=April2008Evening$Number.3
> par(mfrow=c(1,3))
> hist(April2008EveningOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="April 2008 Evening Ball One")
> hist(April2008EveningTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="April 2008 Evening Ball Two")
> hist(April2008EveningThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="April 2008 Evening Ball Three")
```

May 2008 Day

```r
> print(May2008Day)
> May2008DayOne=May2008Day$Number.1
> May2008DayTwo=May2008Day$Number.2
> May2008DayThree=May2008Day$Number.3
> par(mfrow=c(1,3))
> hist(May2008DayOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="May 2008 Day Ball One")
> hist(May2008DayTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="May 2008 Day Ball Two")
> hist(May2008DayThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="May 2008 Day Ball Three")
```

May 2008 Evening

```r
> print(May2008Evening)
> May2008EveningOne=May2008Evening$Number.1
> May2008EveningTwo=May2008Evening$Number.2
> May2008EveningThree=May2008Evening$Number.3
> par(mfrow=c(1,3))
> hist(May2008EveningOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="May 2008 Evening Ball One")
> hist(May2008EveningTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="May 2008 Evening Ball Two")
> hist(May2008EveningThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="May 2008 Evening Ball Three")
```

June 2008 Day

```r
> print(June2008Day)
> June2008DayOne=June2008Day$Number.1
> June2008DayTwo=June2008Day$Number.2
> June2008DayThree=June2008Day$Number.3
> par(mfrow=c(1,3))
> hist(June2008DayOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="June 2008 Day Ball One")
> hist(June2008DayTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="June 2008 Day Ball Two")
> hist(June2008DayThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="June 2008 Day Ball Three")
```

June 2008 Evening

```r
> print(June2008Evening)
> June2008EveningOne=June2008Evening$Number.1
> June2008EveningTwo=June2008Evening$Number.2
```
> June2008EveningThree=June2008Evening$Number.3
> par(mfrow=c(1,3))
> hist(June2008EveningOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="June 2008 Evening Ball One")
> hist(June2008EveningTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="June 2008 Evening Ball Two")
> hist(June2008EveningThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="June 2008 Evening Ball Three")

**July 2008 Day**

> print(July2008Day)
> July2008DayOne=July2008Day$Number.1
> July2008DayTwo=July2008Day$Number.2
> July2008DayThree=July2008Day$Number.3
> par(mfrow=c(1,3))
> hist(July2008DayOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="July 2008 Day Ball One")
> hist(July2008DayTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="July 2008 Day Ball Two")
> hist(July2008DayThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="July 2008 Day Ball Three")

**July 2008 Evening**

> print(July2008Evening)
> July2008EveningOne=July2008Evening$Number.1
> July2008EveningTwo=July2008Evening$Number.2
> July2008EveningThree=July2008Evening$Number.3
> par(mfrow=c(1,3))
> hist(July2008EveningOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="July 2008 Evening Ball One")
> hist(July2008EveningTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="July 2008 Evening Ball Two")
> hist(July2008EveningThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="July 2008 Evening Ball Three")

**August 2008 Day**

> print(Aug2008Day)
> Aug2008DayOne=Aug2008Day$Number.1
> Aug2008DayTwo=Aug2008Day$Number.2
> Aug2008DayThree=Aug2008Day$Number.3
> par(mfrow=c(1,3))
> hist(Aug2008DayOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="August 2008 Day Ball One")
> hist(Aug2008DayTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="August 2008 Day Ball Two")
> hist(Aug2008DayThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="August 2008 Day Ball Three")

**August 2008 Evening**

> print(Aug2008Evening)
> Aug2008EveningOne=Aug2008Evening$Number.1
> Aug2008EveningTwo=Aug2008Evening$Number.2
> Aug2008EveningThree=Aug2008Evening$Number.3
> par(mfrow=c(1,3))
> hist(Aug2008EveningOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="August 2008 Evening Ball One")
> hist(Aug2008EveningTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="August 2008 Evening Ball Two")
> hist(Aug2008EveningThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="August 2008 Evening Ball Three")

**September 2008 Day**

> print(Sept2008Day)
> Sept2008DayOne=Sept2008Day$Number.1
> Sept2008DayTwo=Sept2008Day$Number.2
> Sept2008DayThree=Sept2008Day$Number.3
> par(mfrow=c(1,3))
> hist(Sept2008DayOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="September 2008 Day Ball One")
> hist(Sept2008DayTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="September 2008 Day Ball Two")
> hist(Sept2008DayThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="September 2008 Day Ball Three")
September 2008 Evening
> print(Sept2008Evening)
> Sept2008EveningOne=Sept2008Evening$Number.1
> Sept2008EveningTwo=Sept2008Evening$Number.2
> Sept2008EveningThree=Sept2008Evening$Number.3
> par(mfrow=c(1,3))
> hist(Sept2008EveningOne, breaks=seq(-1.9,by=1), xlab="Number Drawn", main="September 2008 Evening Ball One")
> hist(Sept2008EveningTwo, breaks=seq(-1.9,by=1), xlab="Number Drawn", main="September 2008 Evening Ball Two")
> hist(Sept2008EveningThree, breaks=seq(-1.9,by=1), xlab="Number Drawn", main="September 2008 Evening Ball Three")

October 2008 Day
> print(Oct2008Day)
> Oct2008DayOne=Oct2008Day$Number.1
> Oct2008DayTwo=Oct2008Day$Number.2
> Oct2008DayThree=Oct2008Day$Number.3
> par(mfrow=c(1,3))
> hist(Oct2008DayOne, breaks=seq(-1.9,by=1), xlab="Number Drawn", main="October 2008 Day Ball One")
> hist(Oct2008DayTwo, breaks=seq(-1.9,by=1), xlab="Number Drawn", main="October 2008 Day Ball Two")
> hist(Oct2008DayThree, breaks=seq(-1.9,by=1), xlab="Number Drawn", main="October 2008 Day Ball Three")

October 2008 Evening
> print(Oct2008Evening)
> Oct2008EveningOne=Oct2008Evening$Number.1
> Oct2008EveningTwo=Oct2008Evening$Number.2
> Oct2008EveningThree=Oct2008Evening$Number.3
> par(mfrow=c(1,3))
> hist(Oct2008EveningOne, breaks=seq(-1.9,by=1), xlab="Number Drawn", main="October 2008 Evening Ball One")
> hist(Oct2008EveningTwo, breaks=seq(-1.9,by=1), xlab="Number Drawn", main="October 2008 Evening Ball Two")
> hist(Oct2008EveningThree, breaks=seq(-1.9,by=1), xlab="Number Drawn", main="October 2008 Evening Ball Three")

November 2008 Day
> print(Nov2008Day)
> Nov2008DayOne=Nov2008Day$Number.1
> Nov2008DayTwo=Nov2008Day$Number.2
> Nov2008DayThree=Nov2008Day$Number.3
> par(mfrow=c(1,3))
> hist(Nov2008DayOne, breaks=seq(-1.9,by=1), xlab="Number Drawn", main="November 2008 Day Ball One")
> hist(Nov2008DayTwo, breaks=seq(-1.9,by=1), xlab="Number Drawn", main="November 2008 Day Ball Two")
> hist(Nov2008DayThree, breaks=seq(-1.9,by=1), xlab="Number Drawn", main="November 2008 Day Ball Three")

November 2008 Evening
> print(Nov2008Evening)
> Nov2008EveningOne=Nov2008Evening$Number.1
> Nov2008EveningTwo=Nov2008Evening$Number.2
> Nov2008EveningThree=Nov2008Evening$Number.3
> par(mfrow=c(1,3))
> hist(Nov2008EveningOne, breaks=seq(-1.9,by=1), xlab="Number Drawn", main="November 2008 Evening Ball One")
> hist(Nov2008EveningTwo, breaks=seq(-1.9,by=1), xlab="Number Drawn", main="November 2008 Evening Ball Two")
> hist(Nov2008EveningThree, breaks=seq(-1.9,by=1), xlab="Number Drawn", main="November 2008 Evening Ball Three")

December 2008 Day
> print(Dec2008Day)
> Dec2008DayOne=Dec2008Day$Number.1
> Dec2008DayTwo=Dec2008Day$Number.2
> Dec2008DayThree=Dec2008Day$Number.3
> par(mfrow=c(1,3))
December 2008 Evening
> print(Dec2008Evening)
> Dec2008EveningOne=Dec2008Evening$Number.1
> Dec2008EveningTwo=Dec2008Evening$Number.2
> Dec2008EveningThree=Dec2008Evening$Number.3
> par(mfrow=c(1,3))
> hist(Dec2008EveningOne, breaks=seq(-1.9,by=1), xlab="Number Drawn", main="December 2008 Evening Ball One")
> hist(Dec2008EveningTwo, breaks=seq(-1.9,by=1), xlab="Number Drawn", main="December 2008 Evening Ball Two")
> hist(Dec2008EveningThree, breaks=seq(-1.9,by=1), xlab="Number Drawn", main="December 2008 Evening Ball Three")

Appendix A-4
All Evening Data

```r
> AllEvening=read.csv("/Users/allisoncamras/Desktop/ST 380 Project/AllEvening.csv", header=T)
> print(AllEvening)
> AllEveningOne=AllEvening$Number.1
> AllEveningTwo=AllEvening$Number.2
> AllEveningThree=AllEvening$Number.3
> par(mfrow=c(1,3))
> hist(AllEveningOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="All Evening Ball One")
> hist(AllEveningTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="All Evening Ball Two")
> hist(AllEveningThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="All Evening Ball Three")
```
April, Ball One

April, Ball Two

April, Ball Three

May, Ball One

May, Ball Two

May, Ball Three

June, Ball One

June, Ball Two

June, Ball Three
Jan=read.table("clipboard",header=T)
JanBallOne=Jan$One
JanBallTwo=Jan$Two
JanBallThree=Jan$Three
par(mfrow=c(1,3))
hist(JanBallOne, breaks=seq(-1,9,by=1), main="January, Ball One", xlab="Number Drawn")
hist(JanBallTwo, breaks=seq(-1,9,by=1), main="January, Ball Two", xlab="Number Drawn")
hist(JanBallThree, breaks=seq(-1,9,by=1), main="January, Ball Three", xlab="Number Drawn")

Feb=read.table("clipboard",header=T)
FebBallOne=Feb$One
FebBallTwo=Feb$Two
FebBallThree=Feb$Three
par(mfrow=c(1,3))
hist(FebBallOne, breaks=seq(-1,9,by=1), main="February, Ball One", xlab="Number Drawn")
hist(FebBallTwo, breaks=seq(-1,9,by=1), main="February, Ball Two", xlab="Number Drawn")
hist(FebBallThree, breaks=seq(-1,9,by=1), main="February, Ball Three", xlab="Number Drawn")

Mar=read.table("clipboard",header=T)
MarBallOne=Mar$One
MarBallTwo=Mar$Two
MarBallThree=Mar$Three
par(mfrow=c(1,3))
hist(MarBallOne, breaks=seq(-1,9,by=1), main="March, Ball One", xlab="Number Drawn")
hist(MarBallTwo, breaks=seq(-1,9,by=1), main="March, Ball Two", xlab="Number Drawn")
hist(MarBallThree, breaks=seq(-1,9,by=1), main="March, Ball Three", xlab="Number Drawn")

Apr=read.table("clipboard",header=T)
AprBallOne=Apr$One
AprBallTwo=Apr$Two
AprBallThree=Apr$Three
par(mfrow=c(1,3))
hist(AprBallOne, breaks=seq(-1,9,by=1), main="April, Ball One", xlab="Number Drawn")
hist(AprBallTwo, breaks=seq(-1,9,by=1), main="April, Ball Two", xlab="Number Drawn")
hist(AprBallThree, breaks=seq(-1,9,by=1), main="April, Ball Three", xlab="Number Drawn")

May=read.table("clipboard",header=T)
MayBallOne=May$One
MayBallTwo=May$Two
MayBallThree=May$Three
par(mfrow=c(1,3))
hist(MayBallOne, breaks=seq(-1,9,by=1), main="May, Ball One", xlab="Number Drawn")
hist(MayBallTwo, breaks=seq(-1,9,by=1), main="May, Ball Two", xlab="Number Drawn")
hist(MayBallThree, breaks=seq(-1,9,by=1), main="May, Ball Three", xlab="Number Drawn")

Jun=read.table("clipboard",header=T)
JunBallOne=Jun$One
JunBallTwo=Jun$Two
JunBallThree=Jun$Three
par(mfrow=c(1,3))
hist(JunBallOne, breaks=seq(-1,9,by=1), main="June, Ball One", xlab="Number Drawn")
hist(JunBallTwo, breaks=seq(-1,9,by=1), main="June, Ball Two", xlab="Number Drawn")
hist(JunBallThree, breaks=seq(-1,9,by=1), main="June, Ball Three", xlab="Number Drawn")

Jul=read.table("clipboard",header=T)
JulBallOne=Jul$One
JulBallTwo=Jul$Two
JulBallThree=Jul$Three
par(mfrow=c(1,3))
hist(JulBallOne, breaks=seq(-1,9,by=1), main="Jul, Ball One", xlab="Number Drawn")
hist(JulBallTwo, breaks=seq(-1,9,by=1), main="Jul, Ball Two", xlab="Number Drawn")
hist(JulBallThree, breaks=seq(-1,9,by=1), main="Jul, Ball Three", xlab="Number Drawn")
Aug=read.table("clipboard",header=T)
AugBallOne=Aug$One
AugBallTwo=Aug$Two
AugBallThree=Aug$Three
par(mfrow=c(1,3))
hist(AugBallOne, breaks=seq(-1,9,by=1), main="August, Ball One", xlab="Number Drawn")
hist(AugBallTwo, breaks=seq(-1,9,by=1), main="August, Ball Two", xlab="Number Drawn")
hist(AugBallThree, breaks=seq(-1,9,by=1), main="August, Ball Three", xlab="Number Drawn")

Sep=read.table("clipboard",header=T)
SepBallOne=Sep$One
SepBallTwo=Sep$Two
SepBallThree=Sep$Three
par(mfrow=c(1,3))
hist(SepBallOne, breaks=seq(-1,9,by=1), main="September, Ball One", xlab="Number Drawn")
hist(SepBallTwo, breaks=seq(-1,9,by=1), main="September, Ball Two", xlab="Number Drawn")
hist(SepBallThree, breaks=seq(-1,9,by=1), main="September, Ball Three", xlab="Number Drawn")

Oct=read.table("clipboard",header=T)
OctBallOne=Oct$One
OctBallTwo=Oct$Two
OctBallThree=Oct$Three
par(mfrow=c(1,3))
hist(OctBallOne, breaks=seq(-1,9,by=1), main="October, Ball One", xlab="Number Drawn")
hist(OctBallTwo, breaks=seq(-1,9,by=1), main="October, Ball Two", xlab="Number Drawn")
hist(OctBallThree, breaks=seq(-1,9,by=1), main="October, Ball Three", xlab="Number Drawn")

Nov=read.table("clipboard",header=T)
NovBallOne=Nov$One
NovBallTwo=Nov$Two
NovBallThree=Nov$Three
par(mfrow=c(1,3))
hist(NovBallOne, breaks=seq(-1,9,by=1), main="November, Ball One", xlab="Number Drawn")
hist(NovBallTwo, breaks=seq(-1,9,by=1), main="November, Ball Two", xlab="Number Drawn")
hist(NovBallThree, breaks=seq(-1,9,by=1), main="November, Ball Three", xlab="Number Drawn")

Dec=read.table("clipboard",header=T)
DecBallOne=Dec$One
DecBallTwo=Dec$Two
DecBallThree=Dec$Three
par(mfrow=c(1,3))
hist(DecBallOne, breaks=seq(-1,9,by=1), main="December, Ball One", xlab="Number Drawn")
hist(DecBallTwo, breaks=seq(-1,9,by=1), main="December, Ball Two", xlab="Number Drawn")
hist(DecBallThree, breaks=seq(-1,9,by=1), main="December, Ball Three", xlab="Number Drawn")
Appendix A-7

Wekeudns, Ball One  Weekends, Ball Two  Weekends, Ball Three

Weekdays, Ball One  Weekdays, Ball Two  Weekdays, Ball Three

wekeudns=read.table("clipboard",header=T)
> one=weekends$One
> two=weekends$Two
> three=weekends$Three
> par(mfrow=c(1,3))
> hist(one,breaks=seq(-1,9,by=1),main="Weekends, Ball One",xlab="Number Drawn")
> hist(two,breaks=seq(-1,9,by=1),main="Weekends, Ball Two",xlab="Number Drawn")
> hist(three,breaks=seq(-1,9,by=1),main="Weekends, Ball Three",xlab="Number Drawn")

weekdays=read.table("clipboard",header=T)
> one=weekdays$One
> two=weekdays$Two
> three=weekdays$Three
> hist(one,breaks=seq(-1,9,by=1),main="Weekdays, Ball One",xlab="Number Drawn")
> hist(two,breaks=seq(-1,9,by=1),main="Weekdays, Ball Two",xlab="Number Drawn")
> hist(three,breaks=seq(-1,9,by=1),main="Weekdays, Ball Three",xlab="Number Drawn")
mondaysOne=mondays$One
> mondaysTwo=mondays$Two
> mondaysThree=mondays$Three
> par(mfrow=c(1,3))
> hist(mondaysOne,breaks=seq(-1,9,by=1),main="Mondays, Ball One",xlab="Number Drawn")
> hist(mondaysTwo,breaks=seq(-1,9,by=1),main="Mondays, Ball Two",xlab="Number Drawn")
> hist(mondaysThree,breaks=seq(-1,9,by=1),main="Mondays, Ball Three",xlab="Number Drawn")
> tuesdays=read.table("clipboard",header=T)
> tuesdaysOne=tuesdays$One
> tuesdaysTwo=tuesdays$Two
> tuesdaysThree=tuesdays$Three
> par(mfrow=c(1,3))
> hist(tuesdaysOne,breaks=seq(-1,9,by=1),main="Tuesdays, Ball One",xlab="Number Drawn")
> hist(tuesdaysTwo,breaks=seq(-1,9,by=1),main="Tuesdays, Ball Two",xlab="Number Drawn")
> hist(tuesdaysThree,breaks=seq(-1,9,by=1),main="Tuesdays, Ball Three",xlab="Number Drawn")
> wednesdays=read.table("clipboard",header=T)
> wednesdaysOne=wednesdays$One
> wednesdaysTwo=wednesdays$Two
> wednesdaysThree=wednesdays$Three
> par(mfrow=c(1,3))
> hist(wednesdaysOne,breaks=seq(-1,9,by=1),main="Wednesdays, Ball One",xlab="Number Drawn")
> hist(wednesdaysTwo,breaks=seq(-1,9,by=1),main="Wednesdays, Ball Two",xlab="Number Drawn")
> hist(wednesdaysThree,breaks=seq(-1,9,by=1),main="Wednesdays, Ball Three",xlab="Number Drawn")
> thursdays=read.table("clipboard",header=T)
> thursdaysOne=thursdays$One
> thursdaysTwo=thursdays$Two
> thursdaysThree=thursdays$Three
> par(mfrow=c(1,3))
> hist(thursdaysOne,breaks=seq(-1,9,by=1),main="Thursdays, Ball One",xlab="Number Drawn")
> hist(thursdaysTwo,breaks=seq(-1,9,by=1),main="Thursdays, Ball Two",xlab="Number Drawn")
> hist(thursdaysThree,breaks=seq(-1,9,by=1),main="Thursdays, Ball Three",xlab="Number Drawn")
> fridays=read.table("clipboard",header=T)
> fridaysOne=fridays$One
> fridaysTwo=fridays$Two
> fridaysThree=fridays$Three
> par(mfrow=c(1,3))
> hist(fridaysOne,breaks=seq(-1,9,by=1),main="Fridays, Ball One",xlab="Number Drawn")
> hist(fridaysTwo,breaks=seq(-1,9,by=1),main="Fridays, Ball Two",xlab="Number Drawn")
> hist(fridaysThree,breaks=seq(-1,9,by=1),main="Fridays, Ball Three",xlab="Number Drawn")
> saturdays=read.table("clipboard",header=T)
> saturdaysOne=saturdays$One
> saturdaysTwo=saturdays$Two
> saturdaysThree=saturdays$Three

mondaysOne=mondays$One
> mondaysTwo=mondays$Two
> mondaysThree=mondays$Three
> par(mfrow=c(1,3))
> hist(mondaysOne,breaks=seq(-1,9,by=1),main="Mondays, Ball One",xlab="Number Drawn")
> hist(mondaysTwo,breaks=seq(-1,9,by=1),main="Mondays, Ball Two",xlab="Number Drawn")
> hist(mondaysThree,breaks=seq(-1,9,by=1),main="Mondays, Ball Three",xlab="Number Drawn")
> tuesdays=read.table("clipboard",header=T)
> tuesdaysOne=tuesdays$One
> tuesdaysTwo=tuesdays$Two
> tuesdaysThree=tuesdays$Three
> par(mfrow=c(1,3))
> hist(tuesdaysOne,breaks=seq(-1,9,by=1),main="Tuesdays, Ball One",xlab="Number Drawn")
> hist(tuesdaysTwo,breaks=seq(-1,9,by=1),main="Tuesdays, Ball Two",xlab="Number Drawn")
> hist(tuesdaysThree,breaks=seq(-1,9,by=1),main="Tuesdays, Ball Three",xlab="Number Drawn")
> wednesdays=read.table("clipboard",header=T)
> wednesdaysOne=wednesdays$One
> wednesdaysTwo=wednesdays$Two
> wednesdaysThree=wednesdays$Three
> par(mfrow=c(1,3))
> hist(wednesdaysOne,breaks=seq(-1,9,by=1),main="Wednesdays, Ball One",xlab="Number Drawn")
> hist(wednesdaysTwo,breaks=seq(-1,9,by=1),main="Wednesdays, Ball Two",xlab="Number Drawn")
> hist(wednesdaysThree,breaks=seq(-1,9,by=1),main="Wednesdays, Ball Three",xlab="Number Drawn")
> thursdays=read.table("clipboard",header=T)
> thursdaysOne=thursdays$One
> thursdaysTwo=thursdays$Two
> thursdaysThree=thursdays$Three
> par(mfrow=c(1,3))
> hist(thursdaysOne,breaks=seq(-1,9,by=1),main="Thursdays, Ball One",xlab="Number Drawn")
> hist(thursdaysTwo,breaks=seq(-1,9,by=1),main="Thursdays, Ball Two",xlab="Number Drawn")
> hist(thursdaysThree,breaks=seq(-1,9,by=1),main="Thursdays, Ball Three",xlab="Number Drawn")
> fridays=read.table("clipboard",header=T)
> fridaysOne=fridays$One
> fridaysTwo=fridays$Two
> fridaysThree=fridays$Three
> par(mfrow=c(1,3))
> hist(fridaysOne,breaks=seq(-1,9,by=1),main="Fridays, Ball One",xlab="Number Drawn")
> hist(fridaysTwo,breaks=seq(-1,9,by=1),main="Fridays, Ball Two",xlab="Number Drawn")
> hist(fridaysThree,breaks=seq(-1,9,by=1),main="Fridays, Ball Three",xlab="Number Drawn")
> saturdays=read.table("clipboard",header=T)
> saturdaysOne=saturdays$One
> saturdaysTwo=saturdays$Two
> saturdaysThree=saturdays$Three

```r
> par(mfrow=c(1,3))
> hist(saturdaysOne,breaks=seq(-1,9,by=1),main="Saturdays, Ball One",xlab="Number Drawn")
> hist(saturdaysTwo,breaks=seq(-1,9,by=1),main="Saturdays, Ball Two",xlab="Number Drawn")
> hist(saturdaysThree,breaks=seq(-1,9,by=1),main="Saturdays, Ball Three",xlab="Number Drawn")
> sundays=read.table("clipboard",header=T)
> sundaysOne=sundays$One
> sundaysTwo=sundays$Two
> sundaysThree=sundays$Three
> par(mfrow=c(1,3))
> hist(sundaysOne,breaks=seq(-1,9,by=1),main="Sundays, Ball One",xlab="Number Drawn")
> hist(sundaysTwo,breaks=seq(-1,9,by=1),main="Sundays, Ball Two",xlab="Number Drawn")
> hist(sundaysThree,breaks=seq(-1,9,by=1),main="Sundays, Ball Three",xlab="Number Drawn")
```
Appendix A-9

```r
all=read.table("clipboard",header=T)
allOne=all$One
allTwo=all$Two
allThree=all$Three
par(mfrow=c(1,3))
hist(allOne, breaks=seq(-1,9,by=1), main="Total Lottery History, Ball One", xlab="Number Drawn")
hist(allTwo, breaks=seq(-1,9,by=1), main="Total Lottery History, Ball Two", xlab="Number Drawn")
hist(allThree, breaks=seq(-1,9,by=1), main="Total Lottery History, Ball Three", xlab="Number Drawn")
```
Appendix B-1
Lottery Play Analysis

Friday November 20th (win: 092)
Seven of the nine numbers played had at least one correct digit. Two numbers contained two correct digits. None of the numbers contained all three correct digits; no money was won. The randomly chosen number had no correct digits.

Saturday, November 21st (win: 022)
Four of eight numbers had at least one correct digit, and two of those numbers contained two correct digits. One number consisted of all three correct numbers, but they were not in the correct order. A 3-way 50/50 win was established with a pay out of $80. The randomly chosen number had no correct digits.

Sunday, November 22nd (win: 949)
Six of nine numbers had zero corresponding digits on the winning number. Two analyzed numbers had one digit that matched one of the winning digits. The random pick had zero of the three digits.

Monday, November 23rd (win: 090)
Six of ten numbers had at least one digit and one of those numbers had two digits that matched the winning number. The randomly chosen number had one digit matching the winning number. Three numbers had zero digits matching the winning number.

Tuesday, November 24th (win: 452)
Four of nine numbers had at least one digit and one of those had two digits matching the winning number. Four numbers chosen did not match any of the winning digits as did the randomly chosen ticket.

Wednesday, November 25th (win: 273)
Wednesday was the lucky number day. All nine chosen numbers had one digit that matched the winning number. Two of these nine numbers had two digits that matched. The randomly chosen number had one digit that matched the winning number.

Thursday, November 26th (win: 276)
Eight of nine numbers had one matching digit to the winning number. One other number chosen had two matching digits and the randomly chosen number also had two digits matching the winning number.

Friday, November 27th (win: 332)
Seven of ten numbers had one digit matching the winning number and two of these numbers had two matching digits. The randomly chosen number had no matching digits to the winning number.

Saturday, November 28th (win: 923)
Eight of nine numbers had one matching digit to the winning number. Two of those numbers had two matching digits. The randomly chosen number had two digits that matched the winning number.
### Appendix C-1

**Probability Tables**

#### Total Lottery History: 1620 draws

<table>
<thead>
<tr>
<th>number</th>
<th>Ball One</th>
<th>Ball Two</th>
<th>Ball Three</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.10556</td>
<td>0.10123</td>
<td>0.09877</td>
</tr>
<tr>
<td>1</td>
<td><strong>0.11914</strong></td>
<td>0.10309</td>
<td>0.10123</td>
</tr>
<tr>
<td>2</td>
<td>0.10494</td>
<td>0.10062</td>
<td>0.09753</td>
</tr>
<tr>
<td>3</td>
<td>0.10247</td>
<td>0.09691</td>
<td><strong>0.10741</strong></td>
</tr>
<tr>
<td>4</td>
<td>0.09012</td>
<td>0.10556</td>
<td>0.10000</td>
</tr>
<tr>
<td>5</td>
<td>0.09198</td>
<td>0.09136</td>
<td>0.10185</td>
</tr>
<tr>
<td>6</td>
<td>0.10247</td>
<td>0.10617</td>
<td>0.10123</td>
</tr>
<tr>
<td>7</td>
<td>0.10062</td>
<td><strong>0.11667</strong></td>
<td>0.09506</td>
</tr>
<tr>
<td>8</td>
<td>0.09198</td>
<td>0.09630</td>
<td>0.09506</td>
</tr>
<tr>
<td>9</td>
<td>0.09074</td>
<td>0.08210</td>
<td>0.10185</td>
</tr>
</tbody>
</table>

#### 2006: 87 draws

<table>
<thead>
<tr>
<th>number</th>
<th>Ball One</th>
<th>Ball Two</th>
<th>Ball Three</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td><strong>0.14943</strong></td>
<td>0.10345</td>
<td>0.09195</td>
</tr>
<tr>
<td>1</td>
<td>0.04598</td>
<td>0.09195</td>
<td>0.12644</td>
</tr>
<tr>
<td>2</td>
<td>0.09195</td>
<td>0.10345</td>
<td>0.06897</td>
</tr>
<tr>
<td>3</td>
<td>0.05747</td>
<td>0.09195</td>
<td>0.10345</td>
</tr>
<tr>
<td>4</td>
<td>0.11494</td>
<td>0.08046</td>
<td>0.08046</td>
</tr>
<tr>
<td>5</td>
<td>0.12644</td>
<td>0.10345</td>
<td>0.11494</td>
</tr>
<tr>
<td>6</td>
<td>0.10345</td>
<td>0.05747</td>
<td>0.11494</td>
</tr>
<tr>
<td>7</td>
<td>0.06897</td>
<td><strong>0.14943</strong></td>
<td>0.06897</td>
</tr>
<tr>
<td>8</td>
<td><strong>0.14943</strong></td>
<td>0.10345</td>
<td>0.09195</td>
</tr>
<tr>
<td>9</td>
<td>0.09195</td>
<td>0.11494</td>
<td><strong>0.13793</strong></td>
</tr>
</tbody>
</table>

#### 2007: 365 draws

<table>
<thead>
<tr>
<th>number</th>
<th>Ball One</th>
<th>Ball Two</th>
<th>Ball Three</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.11233</td>
<td>0.10411</td>
<td>0.10411</td>
</tr>
<tr>
<td>1</td>
<td><strong>0.13151</strong></td>
<td>0.09315</td>
<td>0.08219</td>
</tr>
<tr>
<td>2</td>
<td>0.10411</td>
<td>0.08767</td>
<td>0.09863</td>
</tr>
<tr>
<td>3</td>
<td>0.09041</td>
<td>0.09863</td>
<td>0.08767</td>
</tr>
<tr>
<td>4</td>
<td>0.10411</td>
<td>0.08219</td>
<td>0.09041</td>
</tr>
<tr>
<td>5</td>
<td>0.06301</td>
<td>0.12055</td>
<td>0.09315</td>
</tr>
<tr>
<td>6</td>
<td>0.09315</td>
<td><strong>0.13425</strong></td>
<td>0.09041</td>
</tr>
<tr>
<td>7</td>
<td>0.09863</td>
<td>0.11781</td>
<td>0.12603</td>
</tr>
<tr>
<td>8</td>
<td>0.09589</td>
<td>0.07945</td>
<td>0.09863</td>
</tr>
<tr>
<td>9</td>
<td>0.10685</td>
<td>0.08219</td>
<td><strong>0.12877</strong></td>
</tr>
</tbody>
</table>

#### 2008: 603 draws

<table>
<thead>
<tr>
<th>number</th>
<th>Ball One</th>
<th>Ball Two</th>
<th>Ball Three</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.11111</td>
<td>0.10448</td>
<td>0.10116</td>
</tr>
<tr>
<td>1</td>
<td><strong>0.11111</strong></td>
<td>0.10779</td>
<td>0.10779</td>
</tr>
<tr>
<td>2</td>
<td>0.10116</td>
<td>0.08955</td>
<td>0.09121</td>
</tr>
<tr>
<td>3</td>
<td><strong>0.11609</strong></td>
<td>0.09453</td>
<td>0.10448</td>
</tr>
<tr>
<td>4</td>
<td>0.07463</td>
<td><strong>0.12272</strong></td>
<td>0.09784</td>
</tr>
<tr>
<td>5</td>
<td>0.09453</td>
<td>0.08789</td>
<td><strong>0.10945</strong></td>
</tr>
<tr>
<td>6</td>
<td><strong>0.11609</strong></td>
<td>0.10116</td>
<td>0.10448</td>
</tr>
<tr>
<td>7</td>
<td>0.10282</td>
<td>0.11940</td>
<td>0.09121</td>
</tr>
<tr>
<td>8</td>
<td>0.07463</td>
<td>0.09453</td>
<td>0.09950</td>
</tr>
<tr>
<td>9</td>
<td>0.09784</td>
<td>0.07794</td>
<td>0.09287</td>
</tr>
</tbody>
</table>
### 2009: 565 draws

<table>
<thead>
<tr>
<th>number</th>
<th>Ball One</th>
<th>Ball Two</th>
<th>Ball Three</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.08850</td>
<td>0.09558</td>
<td>0.09381</td>
</tr>
<tr>
<td>1</td>
<td>0.13097</td>
<td>0.10619</td>
<td>0.10265</td>
</tr>
<tr>
<td>2</td>
<td>0.11150</td>
<td>0.12035</td>
<td>0.10796</td>
</tr>
<tr>
<td>3</td>
<td>0.10265</td>
<td>0.09912</td>
<td>0.12389</td>
</tr>
<tr>
<td>4</td>
<td>0.09381</td>
<td>0.10619</td>
<td>0.11150</td>
</tr>
<tr>
<td>5</td>
<td>0.10265</td>
<td>0.07434</td>
<td>0.09735</td>
</tr>
<tr>
<td>6</td>
<td>0.09381</td>
<td>0.10088</td>
<td>0.10265</td>
</tr>
<tr>
<td>7</td>
<td>0.10442</td>
<td>0.10796</td>
<td>0.08319</td>
</tr>
<tr>
<td>8</td>
<td>0.09912</td>
<td>0.10088</td>
<td>0.08850</td>
</tr>
<tr>
<td>9</td>
<td>0.07257</td>
<td>0.08850</td>
<td>0.08850</td>
</tr>
</tbody>
</table>

### Daytime: 498 draws

<table>
<thead>
<tr>
<th>number</th>
<th>Ball One</th>
<th>Ball Two</th>
<th>Ball Three</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.10040</td>
<td>0.11245</td>
<td>0.09639</td>
</tr>
<tr>
<td>1</td>
<td>0.11044</td>
<td>0.10843</td>
<td>0.11647</td>
</tr>
<tr>
<td>2</td>
<td>0.10442</td>
<td>0.09839</td>
<td>0.09639</td>
</tr>
<tr>
<td>3</td>
<td>0.09438</td>
<td>0.09438</td>
<td>0.10241</td>
</tr>
<tr>
<td>4</td>
<td>0.08635</td>
<td>0.10643</td>
<td>0.09639</td>
</tr>
<tr>
<td>5</td>
<td>0.09639</td>
<td>0.08032</td>
<td>0.12450</td>
</tr>
<tr>
<td>6</td>
<td>0.11245</td>
<td>0.10040</td>
<td>0.10040</td>
</tr>
<tr>
<td>7</td>
<td><strong>0.11647</strong></td>
<td><strong>0.12851</strong></td>
<td>0.08434</td>
</tr>
<tr>
<td>8</td>
<td>0.08032</td>
<td>0.10040</td>
<td>0.10241</td>
</tr>
<tr>
<td>9</td>
<td>0.09839</td>
<td>0.07028</td>
<td>0.08032</td>
</tr>
</tbody>
</table>

### Evenings: 1122 draws

<table>
<thead>
<tr>
<th>number</th>
<th>Ball One</th>
<th>Ball Two</th>
<th>Ball Three</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.10784</td>
<td>0.09626</td>
<td>0.09982</td>
</tr>
<tr>
<td>1</td>
<td><strong>0.12299</strong></td>
<td>0.10071</td>
<td>0.09447</td>
</tr>
<tr>
<td>2</td>
<td>0.10517</td>
<td>0.10160</td>
<td>0.09804</td>
</tr>
<tr>
<td>3</td>
<td>0.10606</td>
<td>0.09804</td>
<td>0.10963</td>
</tr>
<tr>
<td>4</td>
<td>0.09180</td>
<td>0.10517</td>
<td>0.10160</td>
</tr>
<tr>
<td>5</td>
<td>0.09002</td>
<td>0.09626</td>
<td>0.09180</td>
</tr>
<tr>
<td>6</td>
<td>0.09804</td>
<td>0.10873</td>
<td>0.10160</td>
</tr>
<tr>
<td>7</td>
<td>0.09358</td>
<td><strong>0.11141</strong></td>
<td>0.09982</td>
</tr>
<tr>
<td>8</td>
<td>0.09715</td>
<td>0.09447</td>
<td>0.09180</td>
</tr>
<tr>
<td>9</td>
<td>0.08734</td>
<td>0.08734</td>
<td><strong>0.11141</strong></td>
</tr>
</tbody>
</table>
Ball One Density Plots and Test

Chi-squared test for given probabilities

data: freq.one
X-squared = 12.4697, df = 9, p-value = 0.1881
Ball Two Density Plot and Test

All Lottery Data Ball Two

Chi-squared test for given probabilities

data:  freq.one
X-squared = 12.1483, df = 9, p-value = 0.2051
Ball Three Density Plot and Test

Ball Three

All Lottery Data Ball Three

> AllDataThree=AllData$Number.3
> freq.one=hist(AllDataThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="All Lottery Data Ball Three Frequencies", freq=F)$counts
> lines(density(AllDataThree))
> chisq.test(freq.one)

Chi-squared test for given probabilities

data: freq.one
X-squared = 2.0989, df = 9, p-value = 0.9898
Ball One vs. Ball Two
> chisq.test(AllDataOne,AllDataTwo)

Pearson's Chi-squared test

data:  AllDataOne and AllDataTwo
X-squared = 71.9571, df = 81, p-value = 0.7536

Ball One vs. Ball Three
> chisq.test(AllDataOne,AllDataThree)

Pearson's Chi-squared test

data:  AllDataOne and AllDataThree
X-squared = 102.6571, df = 81, p-value = 0.05249

Ball Two vs. Ball Three
> chisq.test(AllDataTwo,AllDataThree)

Pearson's Chi-squared test

data:  AllDataTwo and AllDataThree
X-squared = 81.5018, df = 81, p-value = 0.4635

Conclusion

For Ball One in the total data, there is some slight variation in the densities of the digits, but a p-value of ~0.18 is greater than 0.05, showing that there is no statistical difference between them. Balls Two and Three in the total data had high p-values as well, showing that there is not statistical difference between the digits either. The p-values increased from Ball One to Ball Three, starting with ~0.18 for Ball One, ~0.21 for Ball Two, and ~0.98 for Ball Three. This shows that Ball Three had the most uniform distribution (less variation of densities) out of the three balls, and Ball One had the least uniform distribution (most variation of densities). This difference is kind of interesting, as one would expect to find the p-values for all three balls to be very similar.

For the comparisons between the balls, the p-values for Ball One vs. Ball Two and Ball Two vs. Ball Three were both greater than 0.05. This shows that the digits picked between balls one and two are statistically independent of one another. Meaning, the digit picked for Ball One doesn’t affect the pick for Ball Two, etc. The p-value for Ball One vs. Ball Three was only slightly greater than 0.05 (it was ~0.053). Since the p-value was below 0.05, it’s still very close to showing that the digits picked for balls one and three are statistically dependent of one another, meaning that the digit picked for Ball One does affect the digit picked for Ball Three. This dependence would be strange considering that the digits for each ball are picked from separate containers. So, it could be a coincidence that the p-value is so close to 0.05, and since the p-value is still greater than 0.05 (even if by just a little), we can conclude that the digits chosen from Ball One and Ball Three are statistically independent.
Ball Frequencies for Each Year (2006-2009)

2006 Ball One Density Plot and Test

Ball One

Chi-squared test for given probabilities

data:  freq.one
X-squared = 10.0465, df = 9, p-value = 0.3467
2007 Ball One Density Plot and Test

![Density Plot](image)

Chi-squared test for given probabilities

data: freq.one
X-squared = 9.8462, df = 9, p-value = 0.3631
> All2008One=All2008$Number.1
> freq.one.2008=hist(All2008One, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="All 2008 Data Ball One", freq=F)$counts
> lines(density(All2008One))
> chisq.test(freq.one.2008)

Chi-squared test for given probabilities

data: freq.one
X-squared = 12.4942, df = 9, p-value = 0.1869
2009 Ball One Density Plot and Test

All 2009 Ball One Frequencies

> All2009=read.csv("/Volumes/COLLEGE/ST 380 Project/All2009.csv", header=T)
> All2009One=All2009$Number.1
> freq.one.2009=hist(All2009One,breaks=seq(-1,9,by=1),xlab="Number Drawn",main="All 2009 Ball One Frequencies",freq=F)$counts
> lines(density(All2009One))
> chisq.test(freq.one.2009)

Chi-squared test for given probabilities

data:  freq.one
X-squared = 11.8319, df = 9, p-value = 0.2230

---

>
Ball One Comparisons by Year

2006 vs. 2007
> chisq.test(freq.one.2006, freq.one.2007)

Pearson's Chi-squared test

data: freq.one.2006 and freq.one.2007
X-squared = 55, df = 48, p-value = 0.2267

2006 vs. 2008
> chisq.test(freq.one.2006, freq.one.2008)

Pearson's Chi-squared test

data: freq.one.2006 and freq.one.2008
X-squared = 47.5, df = 42, p-value = 0.2586

2006 vs. 2009
> chisq.test(freq.one.2006, freq.one.2009)

Pearson's Chi-squared test

data: freq.one.2006 and freq.one.2009
X-squared = 50, df = 48, p-value = 0.3939

2007 vs. 2008
> chisq.test(freq.one.2007, freq.one.2008)

Pearson's Chi-squared test

data: freq.one.2007 and freq.one.2008
X-squared = 62.5, df = 56, p-value = 0.2564

2007 vs. 2009
> chisq.test(freq.one.2007, freq.one.2009)

Pearson's Chi-squared test

data: freq.one.2007 and freq.one.2009
X-squared = 70, df = 64, p-value = 0.2833

2008 vs. 2009
> chisq.test(freq.one.2008, freq.one.2009)

Pearson's Chi-squared test

data: freq.one.2008 and freq.one.2009
X-squared = 60, df = 56, p-value = 0.3329
Ball Two

2006 Ball Two Density Plot and Test

All 2006 Data Ball Two

Chi-squared test for given probabilities

data: freq.two.2006
X-squared = 4.4651, df = 9, p-value = 0.8782
> All2007Two = All2007$Number.2
> freq.two.2007 = hist(All2007Two, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="All 2007 Data Ball Two", freq=F)$counts
> lines(density(All2007Two))
> chisq.test(freq.two.2007)

Chi-squared test for given probabilities

data: freq.two.2007
X-squared = 12.0989, df = 9, p-value = 0.2078
2008 Ball Two Density Plot and Test

All 2008 Data Ball Two

Chi-squared test for given probabilities

data: freq.two.2008
X-squared = 10.4975, df = 9, p-value = 0.3117
2009 Ball Two Density Plot and Test

All 2009 Ball Two Frequencies

> All2009Two=All2009$Number.2
> freq.two.2009=hist(All2009Two,breaks=seq(-1,9,by=1),xlab="Number Drawn",main="All 2009 Ball Two Frequencies",freq=F)$counts
> lines(density(All2009Two))
> chisq.test(freq.one)

Chi-squared test for given probabilities

data:  freq.one
X-squared = 9.2832, df = 9, p-value = 0.4116
Ball Two Comparisons by Year

2006 vs. 2007
> chisq.test(freq.two.2006, freq.two.2007)

Pearson's Chi-squared test

data:  freq.two.2006 and freq.two.2007
X-squared = 40, df = 40, p-value = 0.4703

2006 vs. 2008
> chisq.test(freq.two.2006, freq.two.2008)

Pearson's Chi-squared test

data:  freq.two.2006 and freq.two.2008
X-squared = 50, df = 45, p-value = 0.2815

2006 vs. 2009
> chisq.test(freq.two.2006, freq.two.2009)

Pearson's Chi-squared test

data:  freq.two.2006 and freq.two.2009
X-squared = 36.6667, df = 35, p-value = 0.3914

2007 vs. 2008
> chisq.test(freq.two.2007, freq.two.2008)

Pearson's Chi-squared test

data:  freq.two.2007 and freq.two.2008
X-squared = 80, df = 72, p-value = 0.2424

2007 vs. 2009
> chisq.test(freq.two.2007, freq.two.2009)

Pearson's Chi-squared test

data:  freq.two.2007 and freq.two.2009
X-squared = 62.5, df = 56, p-value = 0.2564

2008 vs. 2009
> chisq.test(freq.two.2008, freq.two.2009)

Pearson's Chi-squared test

data:  freq.two.2008 and freq.two.2009
X-squared = 70, df = 63, p-value = 0.2544
Ball Three

2006 Ball Three Density Plot and Test

All 2006 Data Ball Three

> All2006Three=All2006$Number.3
> freq.three.2006=hist(All2006Three, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="All 2006 Data Ball Three", freq=F)$counts
> lines(density(All2006Three))
> chisq.test(freq.three.2006)

Chi-squared test for given probabilities

data:  freq.three.2006
X-squared = 4.6977, df = 9, p-value = 0.8598
2007 Ball Three Density Plot and Test

All 2007 Data Ball Three

Chi-squared test for given probabilities

data:  freq.three.2007
X-squared = 8.4176, df = 9, p-value = 0.4927
> All2008Three=All2008$Number.3
> freq.three.2008=hist(All2008Three, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="All 2008 Data Ball Three", freq=F)$counts
> lines(density(All2008Three))
> chisq.test(freq.three.2008)

Chi-squared test for given probabilities

data:  freq.three.2008
X-squared = 2.6439, df = 9, p-value = 0.9768
2009 Ball Three Density Plot and Test

> All2009Three=All2009$Number.3
> freq.three.2009=hist(All2009Three,breaks=seq(-1,9,by=1),xlab="Number Drawn",main="All 2009 Ball Three Frequencies",freq=F)$counts
> lines(density(All2009Three))
> chisq.test(freq.three.2009)

Chi-squared test for given probabilities

data:  freq.three.2009
X-squared = 8.0796, df = 9, p-value = 0.5261
Ball Three Comparisons by Year

2006 vs. 2007
> chisq.test(freq.three.2006, freq.three.2007)

Pearson's Chi-squared test

data:  freq.three.2006 and freq.three.2007
X-squared = 47.5, df = 42, p-value = 0.2586

2006 vs. 2008
> chisq.test(freq.three.2006, freq.three.2008)

Pearson's Chi-squared test

data:  freq.three.2006 and freq.three.2008
X-squared = 45, df = 42, p-value = 0.3474

2006 vs. 2009
> chisq.test(freq.three.2006, freq.three.2009)

Pearson's Chi-squared test

data:  freq.three.2006 and freq.three.2009
X-squared = 42.5, df = 42, p-value = 0.4494

2007 vs. 2008
> chisq.test(freq.three.2007, freq.three.2008)

Pearson's Chi-squared test

data:  freq.three.2007 and freq.three.2008
X-squared = 57.5, df = 49, p-value = 0.1894

2007 vs. 2009
> chisq.test(freq.three.2007, freq.three.2009)

Pearson's Chi-squared test

data:  freq.three.2007 and freq.three.2009
X-squared = 55, df = 49, p-value = 0.2579

2008 vs. 2009
> chisq.test(freq.three.2008, freq.three.2009)

Pearson's Chi-squared test

data:  freq.three.2008 and freq.three.2009
X-squared = 55, df = 49, p-value = 0.2579

Conclusion

For the comparisons of Ball One between the years, all of the p-values came out to be greater than 0.05. This means that the digits drawn for Ball One for each year is not statistically dependent on the digits drawn for Ball One for the other years, i.e. digit for Ball One in 2006 does not depend on the digit for Ball One in 2007. Also, all of the p-values for each year comparison were around this same value (0.23-0.39). This similarity and independence is definitely to be expected, otherwise the lottery would be biased, and that wouldn’t be fair.

For the Ball Two year comparisons, the p-values were also greater than 0.05, showing that the digits drawn for Ball Two are statistically independent between the years, just like Ball One. There was slightly more variation
between the p-values for Ball Two than there was for Ball One. Ball Two had a range of 0.22 while Ball One had a range of 0.16. This is most likely due to the difference of digits drawn between the balls, and probably has no statistical significance.

The same situation occurred with the Ball Three year comparisons—all of the p-values were greater than 0.05, meaning that the digits drawn for Ball Three between the years are statistically independent of one another. And, again, there was more variation between the Ball Three p-values than there was in Ball One and Ball Two. Ball Three had a range of 0.27 (compared to 0.22 for Ball Two and 0.16 for Ball One).

**SIDE NOTE:** Ball Three in most of the analyses seems to have the highest p-value, most uniform distribution, and the least variation between digit densities. This find has been fairly consistent. The p-values have seemed to increase from Ball One to Ball Three.
Ball One Daytime Density Plot and Test

```r
> AllDay=read.csv("/Volumes/COLLEGE/ST 380 Project/AllDay.csv", header=T)
> AllDayOne=AllDay$Number.1
> freq.one.day=hist(AllDayOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="All Daytime Data Ball One Frequencies", freq=F)$counts
> lines(density(AllDayOne))
> chisq.test(freq.one.day)
```

Chi-squared test for given probabilities
data:  freq.one.day
X-squared = 5.8554, df = 9, p-value = 0.7543
AllEvening= read.csv("/Volumes/COLLEGE/ST 380 Project/AllEvening.csv", header=T)
> AllEveningOne=AllEvening$Number.1
> freq.one.evening=hist(AllEveningOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="All Evening Data Ball One Frequencies", freq=F)$counts
> lines(density(AllEveningOne))
> chisq.test(freq.one.evening)
  Chi-squared test for given probabilities
data:  freq.one.evening
X-squared = 11.6007, df = 9, p-value = 0.2368

Ball One Comparison
> chisq.test(freq.one.day, freq.one.evening)
Pearson's Chi-squared test
data:  freq.one.day and freq.one.evening
X-squared = 90, df = 81, p-value = 0.2313
Ball Two Daytime Density Plot and Test

All Daytime Data Ball Two Frequencies

> AllDayTwo=AllDay$Number.2
> freq.two.day=hist(AllDayTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="All Daytime Data Ball Two Frequencies", freq=F)$counts
> lines(density(AllDayTwo))
> chisq.test(freq.two.day)

Chi-squared test for given probabilities

data:  freq.two.day
X-squared = 11.8795, df = 9, p-value = 0.2202
AllEveningTwo = AllEvening$Number.2
> freq.two.evening = hist(AllEveningTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="All Evening Data Ball Two Frequencies", freq=F)$counts
> lines(density(AllEveningTwo))
> chisq.test(freq.two.evening)

Chi-squared test for given probabilities
data:  freq.two.evening
X-squared = 5.148, df = 9, p-value = 0.821

Ball Two Comparison

> chisq.test(freq.two.day, freq.two.evening)

Pearson's Chi-squared test
data:  freq.two.day and freq.two.evening
X-squared = 70, df = 64, p-value = 0.2833
Ball Three Daytime Density Plot and Test

All Daytime Data Ball Three Frequencies

```
AllDayThree = AllDay$Number.3
> freq.three.day = hist(AllDayThree, breaks = seq(-1,9,by=1), xlab = "Number Drawn", main = "All Daytime Data Ball Three Frequencies", freq=F)$counts
> lines(density(AllDayThree))
> chisq.test(freq.three.day)

Chi-squared test for given probabilities

data: freq.three.day
X-squared = 7.743, df = 9, p-value = 0.5602
```
Ball Three Evening Density Plot and Test

AllEvening$Number.3
> freq.three.evening=hist(AllEveningThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="All Evening Data Ball Three Frequencies", freq=F)$counts
> lines(density(AllEveningThree))
> chisq.test(freq.three.evening)

Chi-squared test for given probabilities
data: freq.one
X-squared = 4.4528, df = 9, p-value = 0.8792
Ball Three Comparison

Frequencies”, freq=F$counts
> chisq.test(freq.three.day, freq.three.evening)
   Pearson's Chi-squared test
data:  freq.three.day and freq.three.evening
X-squared = 39.1667, df = 36, p-value = 0.3296

Conclusion

None of our p-values for these tests were less than 0.05. This means that within daytime and evening separately, there is no statistical difference in the probabilities between each digit (0-9). So while yes, there is some variation in the densities of the digits, there’s no statistical evidence suggesting that the lottery drawings are biased. Also, the digits drawn from daytime and evening are independent of each other. For instance, drawing a “1” for ball one in the daytime does not affect the digit drawn in the evening. One would expect the p-values for the evening draws to be higher since we have more data and therefore, a more uniform distribution.
Appendix E-1

November Data and Analysis

Density Graphs

All November Evening Ball One

All November Evening Ball Two

All November Evening Ball Three

November Evening-Mondays Ball One

November Evening-Mondays Ball Two

November Evening-Mondays Ball Three

November Evening-Tuesdays Ball One

November Evening-Tuesdays Ball Two

November Evening-Tuesdays Ball Three
R Code

> AllNovEvening=read.csv("/Volumes/COLLEGE/ST 380 Project/AllNovemberEvening.csv", header=T)
> AllNovEveningOne=AllNovEvening$Number.1
> AllNovEveningTwo=AllNovEvening$Number.2
> AllNovEveningThree=AllNovEvening$Number.3
> par(mfrow=c(1,3))
> hist(AllNovEveningOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="All November Evening Ball One", freq=F)
> lines(density(AllNovEveningOne))
> hist(AllNovEveningTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="All November Evening Ball Two", freq=F)
> lines(density(AllNovEveningTwo))
> hist(AllNovEveningThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="All November Evening Ball Three", freq=F)
> lines(density(AllNovEveningThree))

> NovEveningMonday=read.csv("/Volumes/COLLEGE/ST 380 Project/NovemberEveningMonday.csv", header=T)
> NovEveningMondayOne=NovEveningMonday$Number.1
> NovEveningMondayTwo=NovEveningMonday$Number.2
> NovEveningMondayThree=NovEveningMonday$Number.3
> par(mfrow=c(1,3))
> hist(NovEveningMondayOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Mondays Ball One", freq=F)
> lines(density(NovEveningMondayOne))
> hist(NovEveningMondayTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Mondays Ball Two", freq=F)
> lines(density(NovEveningMondayTwo))
> hist(NovEveningMondayThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Mondays Ball Three", freq=F)
> lines(density(NovEveningMondayThree))

> NovEveningTues=read.csv("/Volumes/COLLEGE/ST 380 Project/NovemberEveningTuesday.csv", header=T)
> NovEveningTuesOne=NovEveningTues$Number.1
> NovEveningTuesTwo=NovEveningTues$Number.2
> NovEveningTuesThree=NovEveningTues$Number.3
> par(mfrow=c(1,3))
> hist(NovEveningTuesOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Tuesdays Ball One", freq=F)
> lines(density(NovEveningTuesOne))
> hist(NovEveningTuesTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Tuesdays Ball Two", freq=F)
> lines(density(NovEveningTuesTwo))
> hist(NovEveningTuesThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Tuesdays Ball Three", freq=F)
> lines(density(NovEveningTuesThree))

> NovEveningWed=read.csv("/Volumes/COLLEGE/ST 380 Project/NovemberEveningWed.csv", header=T)
> NovEveningWedOne=NovEveningWed$Number.1
> NovEveningWedTwo=NovEveningWed$Number.2
> NovEveningWedThree=NovEveningWed$Number.3
> par(mfrow=c(1,3))
> hist(NovEveningWedOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Wednesdays Ball One", freq=F)
> lines(density(NovEveningWedOne))
> hist(NovEveningWedTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Wednesdays Ball Two", freq=F)
> lines(density(NovEveningWedTwo))
> hist(NovEveningWedThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Wednesdays Ball Three", freq=F)
> lines(density(NovEveningWedThree))

> NovEveningThursday=read.csv("/Volumes/COLLEGE/ST 380 Project/NovemberEveningThursday.csv", header=T)
> NovEveningThursdayOne=NovEveningThursday$Number.1
> NovEveningThursdayTwo=NovEveningThursday$Number.2
> NovEveningThursdayThree=NovEveningThursday$Number.3
> par(mfrow=c(1,3))
> hist(NovEveningThursdayOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Thursdays Ball One", freq=F)
> lines(density(NovEveningThursdayOne))
> hist(NovEveningThursdayTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Thursdays Ball Two", freq=F)
> lines(density(NovEveningThursdayTwo))
> hist(NovEveningThursdayThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Thursdays Ball Three", freq=F)
> lines(density(NovEveningThursdayThree))

> NovEveningFriday=read.csv("/Volumes/COLLEGE/ST 380 Project/NovemberEveningFriday.csv", header=T)
> NovEveningFridayOne=NovEveningFriday$Number.1
> NovEveningFridayTwo=NovEveningFriday$Number.2
> NovEveningFridayThree=NovEveningFriday$Number.3
> par(mfrow=c(1,3))
> hist(NovEveningFridayOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Fridays Ball One", freq=F)
> lines(density(NovEveningFridayOne))
> hist(NovEveningFridayTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Fridays Ball Two", freq=F)
> lines(density(NovEveningFridayTwo))
> hist(NovEveningFridayThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Fridays Ball Three", freq=F)
> lines(density(NovEveningFridayThree))

> NovEveningSat=read.csv("/Volumes/COLLEGE/ST 380 Project/NovemberEveningSat.csv", header=T)
> NovEveningSatOne=NovEveningSat$Number.1
> NovEveningSatTwo=NovEveningSat$Number.2
> NovEveningSatThree=NovEveningSat$Number.3
> par(mfrow=c(1,3))
> hist(NovEveningSatOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Saturdays Ball One", freq=F)
> lines(density(NovEveningSatOne))
> hist(NovEveningSatTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Saturdays Ball Two", freq=F)
> lines(density(NovEveningSatTwo))
> hist(NovEveningSatThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Saturdays Ball Three", freq=F)
> lines(density(NovEveningSatThree))

> NovEveningSun=read.csv("/Volumes/COLLEGE/ST 380 Project/NovemberEveningSun.csv", header=T)
> NovEveningSunOne=NovEveningSun$Number.1
> NovEveningSunTwo=NovEveningSun$Number.2
> NovEveningSunThree=NovEveningSun$Number.3
> par(mfrow=c(1,3))
> hist(NovEveningSunOne, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Sundays Ball One", freq=F)
> lines(density(NovEveningSunOne))
> hist(NovEveningSunTwo, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Sundays Ball Two", freq=F)
> lines(density(NovEveningSunTwo))
> hist(NovEveningSunThree, breaks=seq(-1,9,by=1), xlab="Number Drawn", main="November Evening-Sundays Ball Three", freq=F)
> lines(density(NovEveningSunThree))