Project 1: Intro to Descriptive Statistics

Erik Quiroga

05/20/2017

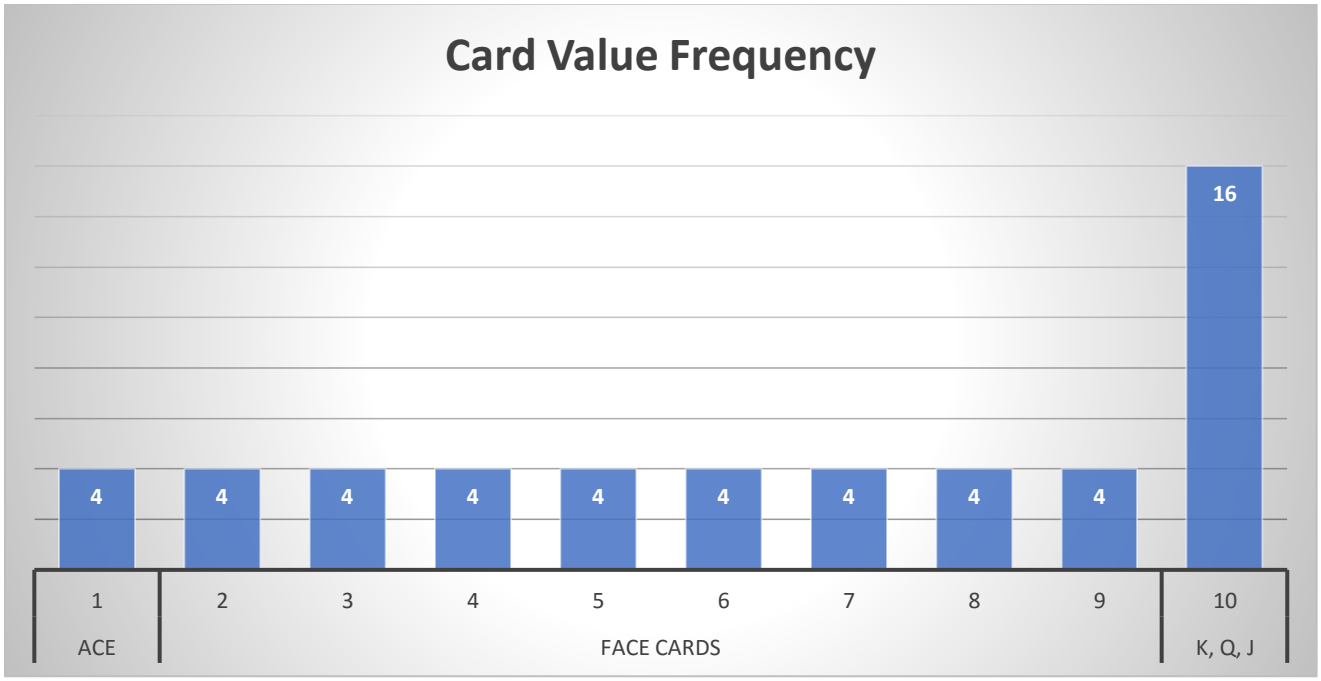
This experiment will require the use of a standard deck of playing cards. This is a deck of fifty-two cards divided into four suits (spades (♠), hearts (♥), diamonds (♦), and clubs (♣)), each suit containing thirteen cards (Ace, numbers 2-10, and face cards Jack, Queen, and King). You can use either a physical deck of cards for this experiment or you may use our data generator in the *Generate Data* section.
For the purposes of this task, assign each card a value: The Ace takes a value of 1, numbered cards take the value printed on the card, and the Jack, Queen, and King each take a value of 10.

Questions for the investigation:

1. First, create a histogram depicting the relative frequencies of the card values for a single draw. Report the mean, median, and standard deviation of the value distribution. (You should have performed this step in the Distribution of Card Values section.)
2. Take a look at the distribution of the three-card sums from the samples that you obtained, either from Generate Data, or from your own collection. Report descriptive statistics for the samples you have drawn. Include at least two measures of central tendency and two measures of variability.
3. Create a histogram of the sampled three-card sums. Compare its shape to that of the original distribution. How are they different, and can you explain why this is the case?
4. Make some estimates about values you would get on future draws. Within what range will you expect approximately 90% of your draw values to fall? What is the approximate probability that you will get a draw value of at least 20? Make sure you justify how you obtained your values.

Summary of Findings:

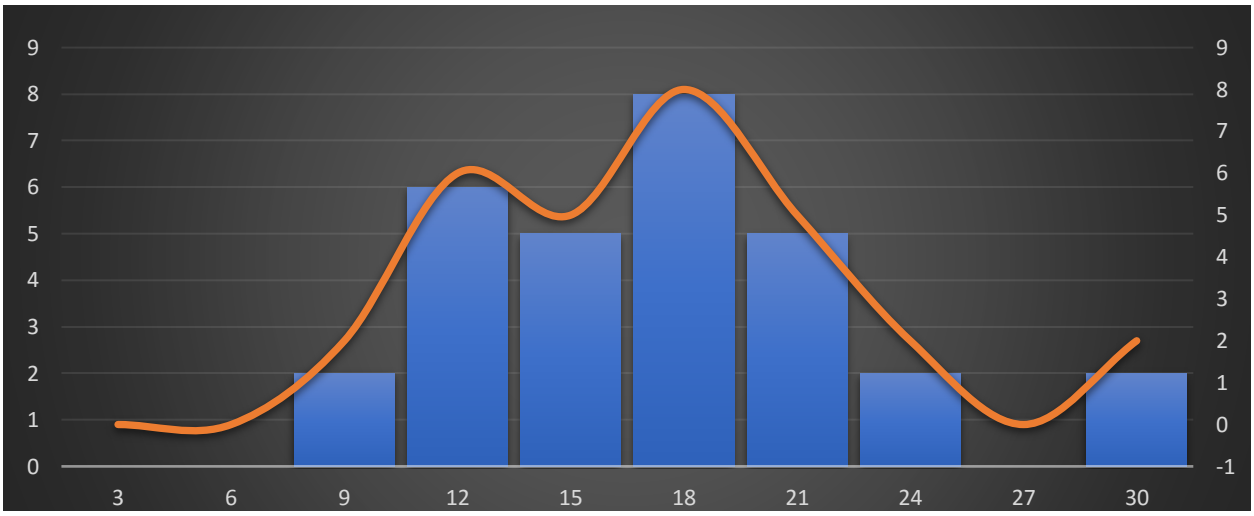
1. Histogram Card Values



Single Draw Results:

μ:	6.54
Median:	7.00
σ:	3.15

2. Sample Distribution



3-Card Draw Results:

Measures of Central Tendency:	
μ:	16.47
Median:	17.00
Mode:	18
Measures of Variability:	
Variance:	29.36
σ:	5.42

3. Analysis of Histograms:
   Compare its shape to that of the original distribution. How are they different, and can you explain why this is the case?

   Both graphs are substantially different due to the number of observations recorded, which is an important reason to generate and record a large enough sample to generate accurate results. The first graph is highly negatively skewed due to small sample sized observations (n=52) composed of just one draw of the deck. The 3-card result was composed of n=90 observations and because of the Central Limit Theorem, which states given a large sample size and a finite level of variance, the samples will follow a normal distribution as show on the 3-card histogram. Yet, the graph is slightly negatively skewed due to the how the measures of Central Tendency are arranged: Mean> Median> Mode and the number of observations. Without question, if we continue to sample more draws from the card deck the shape of the distribution will be a standard distribution curve.

4. Estimation of Future Draws:
   a) Within what range will you expect approximately 90% of your draw values to fall?
   b) What is the approximate probability that you will get a draw value of at least 20? Make sure you justify how you obtained your values?

   a) The Empirical Rule for a normal distribution states: approximately 95% of all values are within 2 standard deviations from the mean. So the range for 90% of the values will fall between 5 and 95 percentile, respectively. (1-.95= .05). The z-score will be -1.645, and 1.645, respectively. The 3-card sample mean is 16.47 with a standard deviation of 5.42. Calculating for the z-scores yields a range of 7.58 and 25.37. We conclude that 90% of draw values will fall within a range of 7.58 and 25.37.

b) The approximate probability draw with a value of at least 20 will result in a z-score of 0.65 (20-16.47/5.42). The probability of obtaining a value of at least 20 results in .47 or 47% (1- .5239).