Riley Payung

STAT 250-003

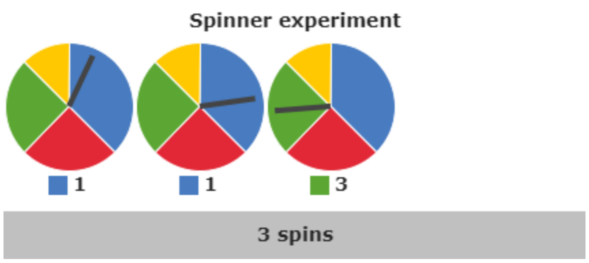
Data Analysis Assignment #2

Problem #1

1.A:

Spaces	1	2	3	4
Weight	0.375	0.25	0.125	0.125

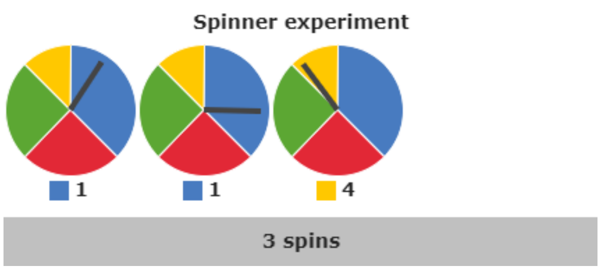
1.B:



1: 3 spaces, 2: 3 spaces, 3: 2 spaces.

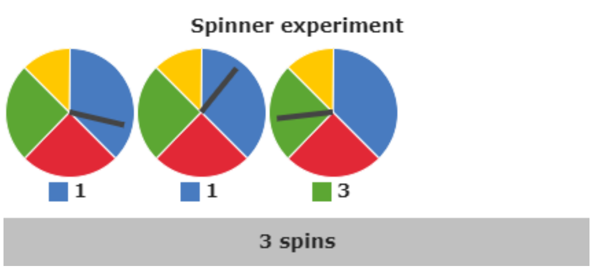
Sum: 5

1.C:



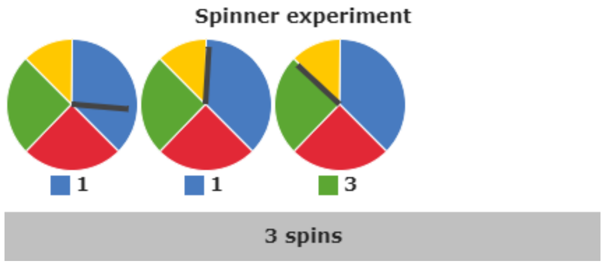
1: 3 spaces, 2: 3 spaces, 3: 1 space.

Sum: 6



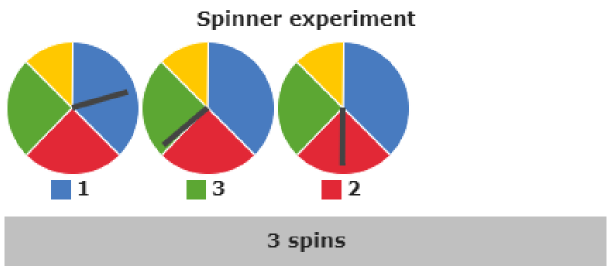
1: 3 spaces, 2: 3 spaces, 3: 2 spaces.

Sum: 5



1: 3 spaces, 2: 3 spaces, 3: 2 spaces.

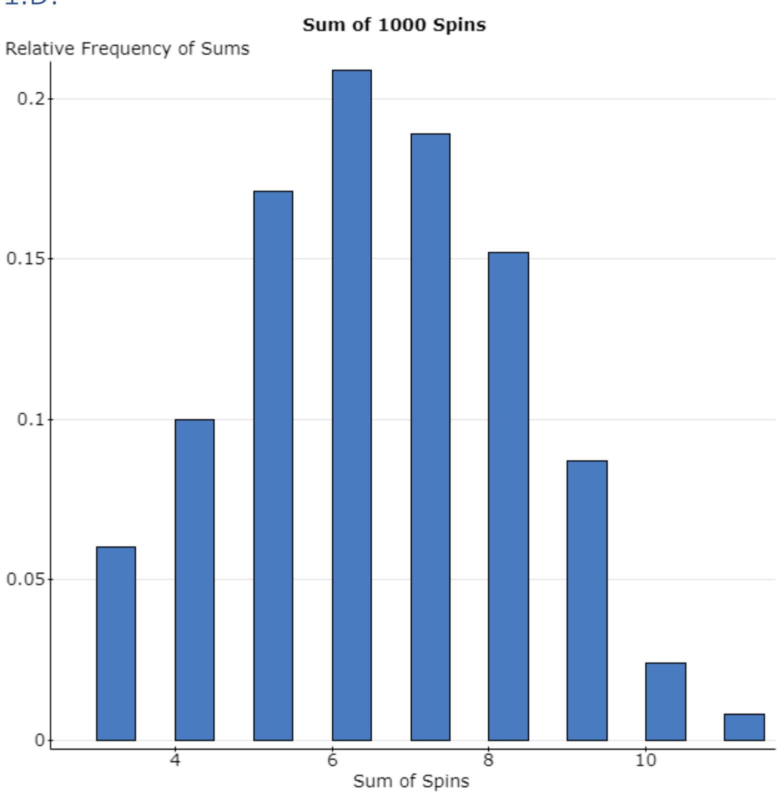
Sum: 5



1: 3 spaces, 2: 2 spaces, 3: 2 spaces.

Sum: 6

1.D:



1.E:

60 Rows of 0.4 or less / 1000 simulations = 0.06.

1.F:

Probability of rolling 1,1,1: 0.053

Probability of rolling 1,1,2: 0.035

Theoretical probability of rolling the sum of 4 or less: 0.053 + 0.035 = 0.088

Meaning that 88 out of 1000 sums of spins should result in the sum of 4 or less.

1.G:

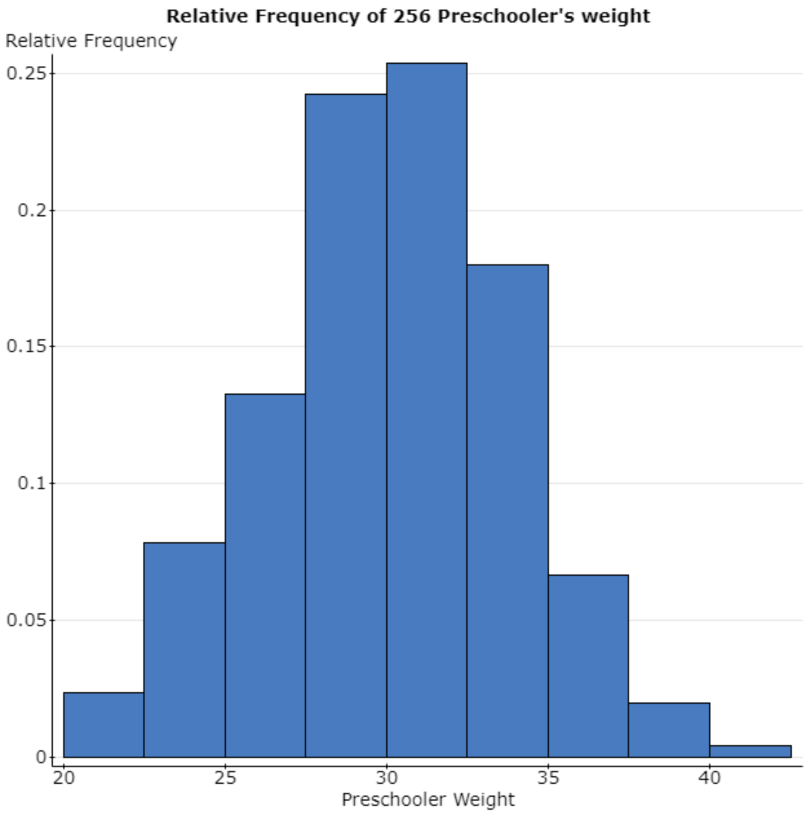
The empirical probability of getting a 4 or less is 0.06 whereas the theoretical probability is 0.088, meaning that the empirical probability is less than the theoretical, but would be closer to it given more samples.

1.H:

The empirical probability should end up being closer to the theoretical probability.

Problem #2

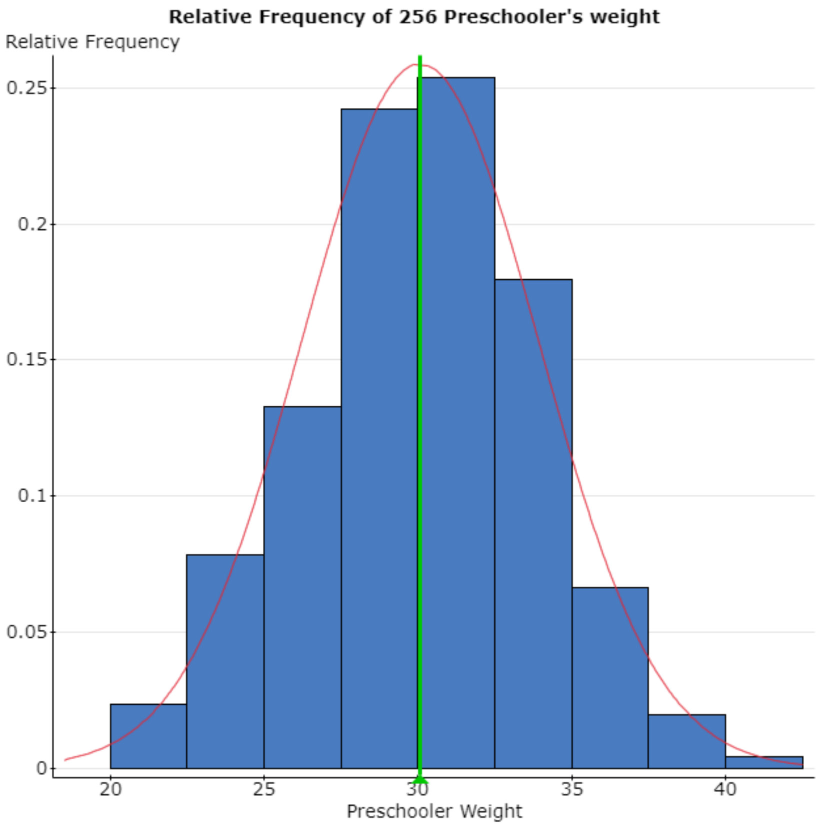
2.A:



2.B:

This is a symmetric, unimodal, bell-shaped histogram.

2.C:



2.D:

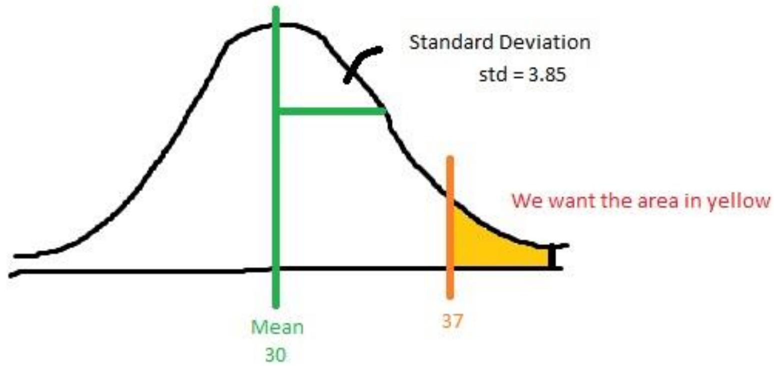
Yes, because the shape of the graph tends to be bell-shaped and the mean is in the direct center of the histogram.

2.E:

Summary statistics:

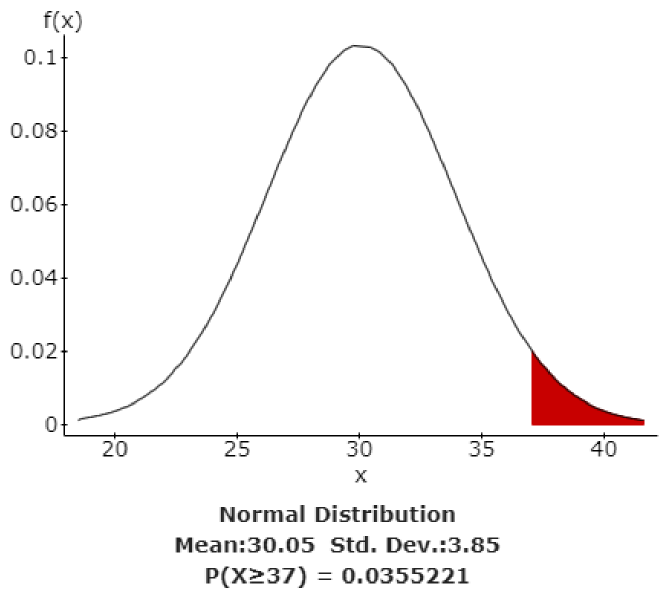
Column	n	Mean	Std. dev.
Weight	256	30.05	3.85

2.F:



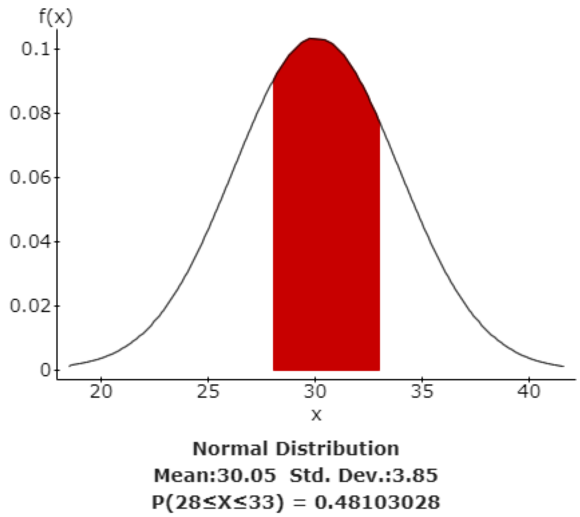
To get this probability, we would use 1 – P (x <= 37) = 1 - 0.9644779 = 0.0355221 (rounded: 0.036).

2.G:



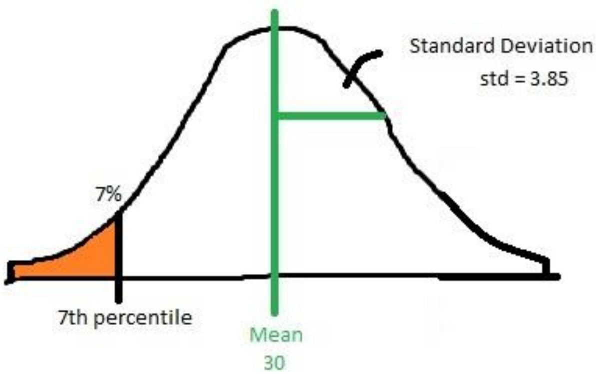
The probability means that the chance of preschoolers who weigh 37 or more pounds is 3.6%.

2.H:



The probability means that the chance of one preschooler weighing between 28 and 33 pounds is 48%.

2.I:



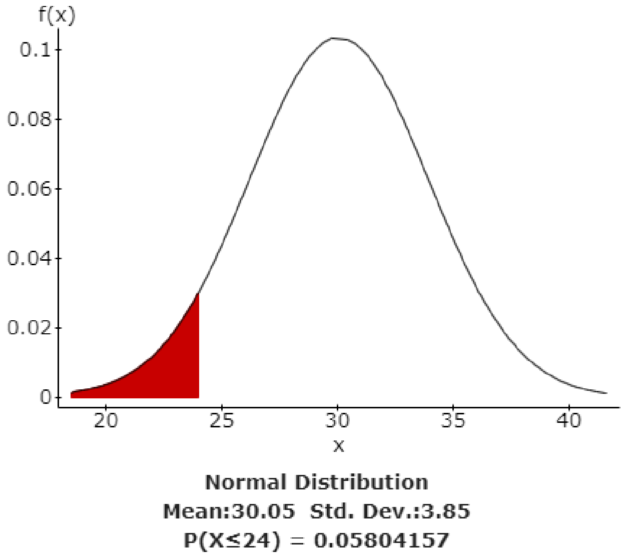
7 percent of the data lies underneath the 7th percentile.

7% of 256 is approximately 18 children. (17.92 actual)

The maximum weight is 24 pounds for someone within the 7th percentile. I found this out by sorting the data in statcrunch and going down to the 18th child. The 17th child also has the same weight of 24 pounds.

The probability that a child lands in the 7th percentile is 0.058.

2.J:



The probability that is here says that there is a 5.8% chance that a preschooler will be underweight coming in at a maximum of 24 pounds.

Problem #3

3.A:

It follows that there are a fixed number of trials, being that there are 19 houses in the sample.

Each observation can be either that it's a brown house, or its not a brown house.

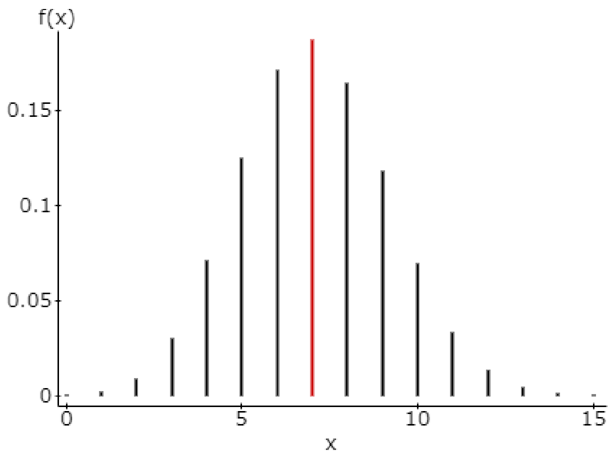
The chance of a house being brown is 37%, p = 0.37.

Each house is independent because one house being brown does not affect another house being brown (unless you have a stupid HOA).

3.B:

P(X)	X
0.00015398	0
0.00171825	1
0.00908215	2
0.0302258	3
0.07100663	4
0.12510692	5
0.17144282	6
0.18699319	7
0.16473209	8
0.11824685	9
0.06944656	10
0.03337043	11
0.01306567	12
0.00413188	13
0.00104	14
0.0002036	15
0.00002989	16
0.0000031	17
2.00E-07	18
6.25E-09	19

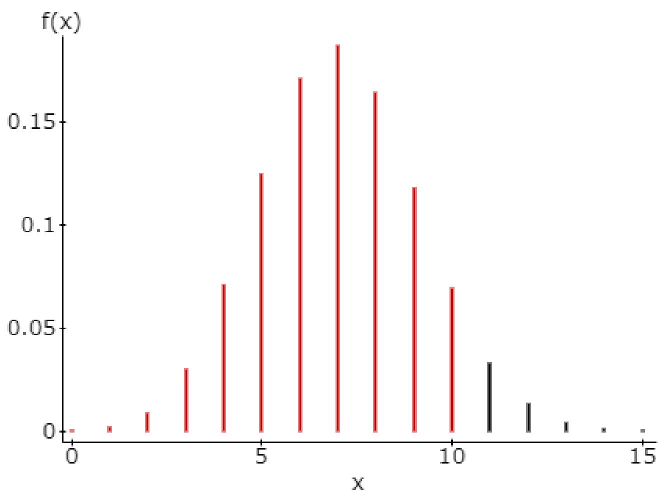
3.C:



Binomial Distribution
n:19 p:0.37
P(X=7) = 0.18699319

Out of the 19 houses, the probability that exactly 7 houses are painted brown is 0.1869.

3.D:

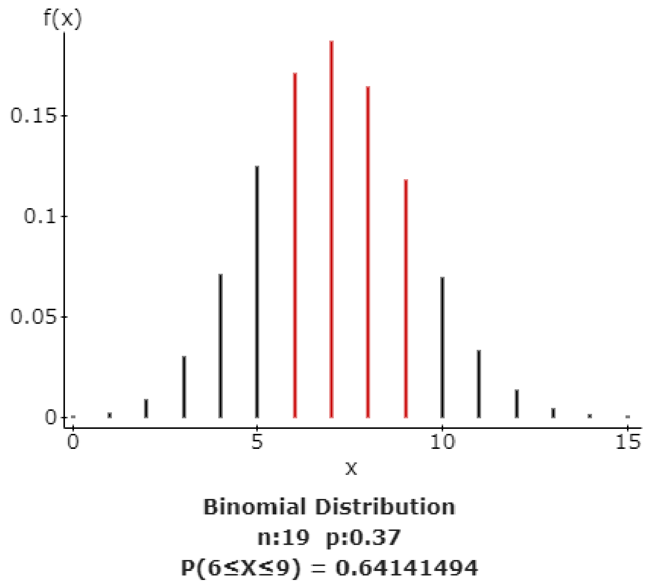


Binomial Distribution
n:19 p:0.37
P(X≤10) = 0.94815523

Out of the 19 houses, the probability that no more than 10 houses are painted brown is 0.9481.

3.E:

The probability that between 6 and 9 houses, inclusive are brown is
P(x=6)+P(x=7)+P(x=8)+P(x=9) = 0.17144282 + 0.18699319 + 0.16473209 + 0.11824685 = 0.64141494



The probability of 0.64141494 means that out of 19 houses, the probability that between 6 and 9 houses are painted brown is 0.64141494.

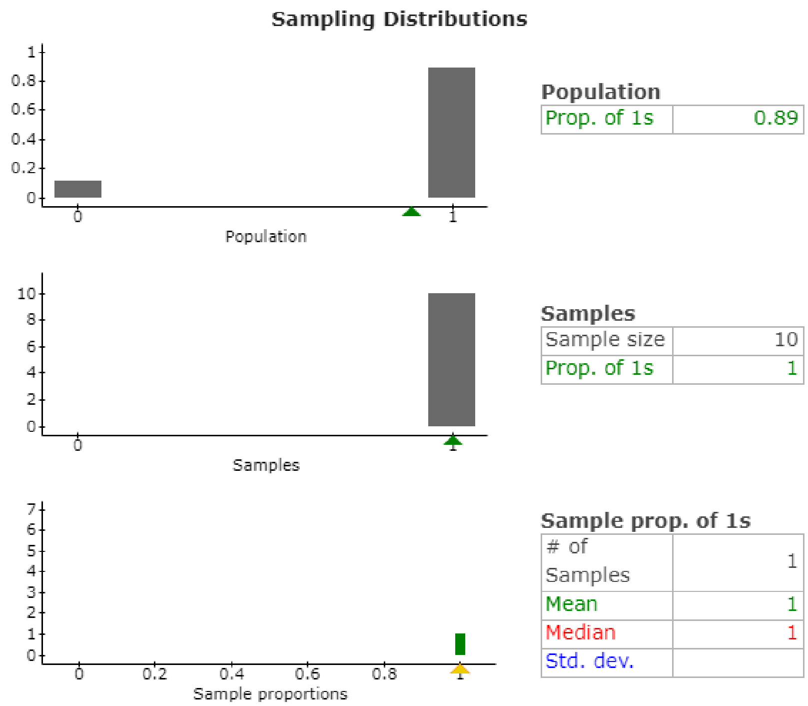
3.F:

Mean = np(1-p) = 19(0.37)(0.63) = 4.4289

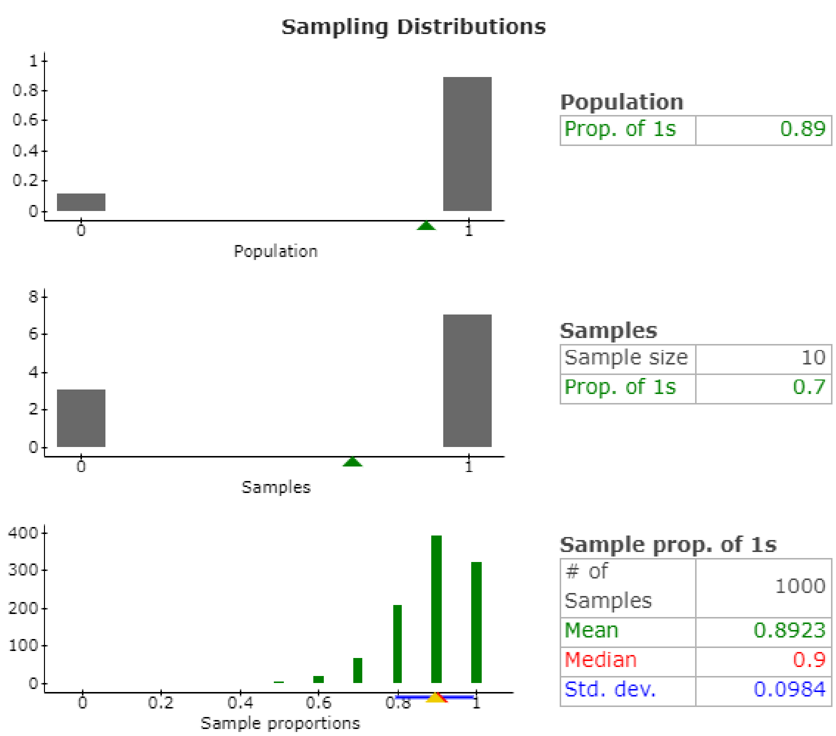
Std = sqrt(np(1-p)) = sqrt(19(0.37)(0.63)) = 2.1045

Problem #4

4.A:



4.B:



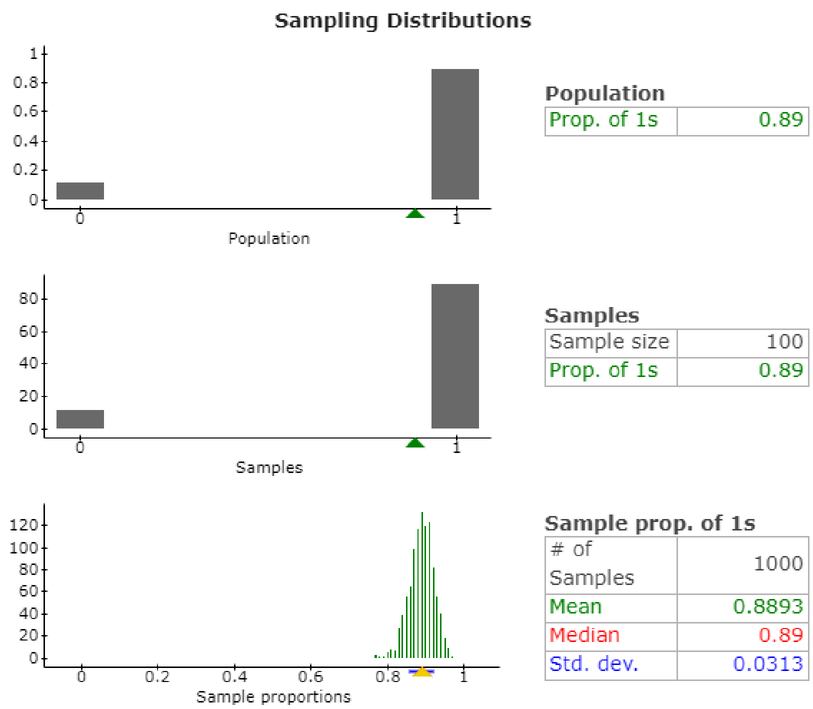
4.C:

The shape of the sampling distribution in 4.B is a left-skewed distribution that is unimodal.

4.D:

The reason that this graph is not normally shaped is because the probability is 0.89, and therefore will be closer to that number, and since the proportion cannot be greater than 1, we won't see a normal shape; If we were able to see past 1, we would be able to likely see a normally shaped graph. We don't have enough data to show that the graph would be normally distributed.

4.E:



4.F:

The shape of the graph in 4.E is a normally shaped graph that is unimodal.

4.G:

The graph is normally shaped because we have a much larger sample size of 100 than in 4.B with only 10 per sample.

4.H:

Mean = 0.8893

Std. Dev. = 0.0313

4.I:

The mean of the samples of size 100 in 4.E is just slightly less than the population proportion 0.89, at 0.8893.

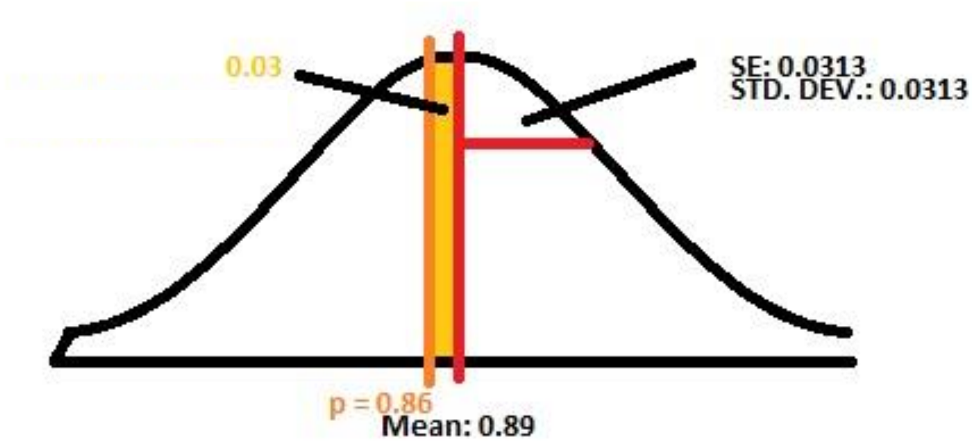
4.J:

Standard Error = sqrt((0.89(0.11))/100) = 0.03129

4.K:

The standard deviation in 4.H is 0.0313, where the calculated standard error is almost the same at 0.03129, and statcrunch probably just rounded it up. They are the same.

4.L:



Z = 0.86 – 0.89 / 0.0313 = 0.9584 -> 0.96

Z = 0.96. P (Z = 0.96) = 0.3315.

4.M: