Lab 4—Comparing Roulette Bets

a.)

Possible values of X:	x_1=-1	x_2 =1
Probability P(X = x)	20/38 = 0.5263	18/38 = 0.4737

b.) E(X) =

$$\sum_{i}^{2} x_i \times \mathrm{P}(X = x_i) = (x_1)\mathrm{P}(X = x_1) + (x_2)\mathrm{P}(X = x_2) = -1(.52632) + 1(.47368) = -.05263$$

This means that in the long run, choosing colors at random will yield a net gain of -$0.05263 per play, so 100 plays would have a net cost of $5.26 and yield no profit on average.

c.)

Possible values of Y:	y_1=-1	y_2 =35
Probability P(X = y)	37/38 = 0.97368	18/38 = 0.02632

d.) E(Y)=

$$\sum_{i}^{2} y_i \times \mathrm{P}(X = y_i) = (y_1)\mathrm{P}(X = y_1) + (y_2)\mathrm{P}(X = y_2) = -1(.97368) + 1(.02632) = -.05263$$

The value is the same as randomly choosing colors results.

e.) The net winnings are equal so the bets are technically equal in terms of net "gain," but the thrill of play or opinion of the player may affect how one might choose to play. Playing numbers results in more significant gains at less frequent intervals, while playing colors results in minor gains and losses throughout yielding the same results in the long run as numbers.

f.) I won 2 out of 5 times with a total gain = -1, and an average net gain of (-1-1-1+1+1)/5 = -.2. Others around me also had the same result.

g.) **Tally for Discrete Variables: C3**

```
C3  Count
-1    516
 1    484
N=   1000
```

Descriptive Statistics: C3

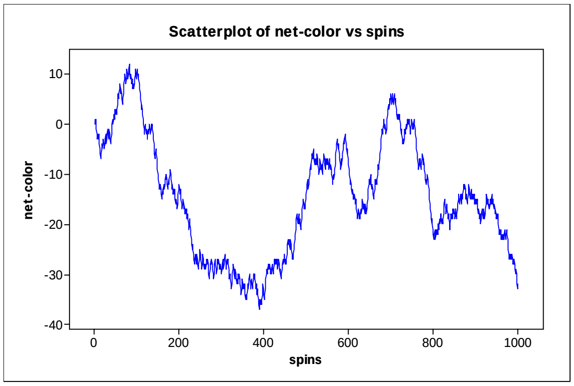
```
Variable     N  N*     Mean  SE Mean   StDev  Minimum       Q1   Median      Q3
C3        1000   0  -0.0320   0.0316  1.0000  -1.0000  -1.0000  -1.0000  1.0000

Variable  Maximum
C3         1.0000
```

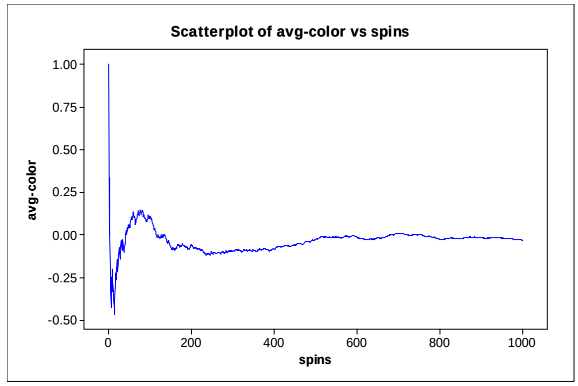
h.) P(X = x_1) = .526 ≈ Relative Frequency = 516/1000 = .516

P(X = x_2) = .474 ≈ Relative Frequency = 484/1000 = .484

The relative frequency and the probability are very similar, varying only by .01 each. The mean is -0.032 according to Minitab, while the expected value is -.05263. This may be a result of not being run enough trials for the random results to tend towards a stable mean value.

i.) `parsum` appears to add the previous total the next entry, showing a cumulative "net gain" after each event—"total gain so far."





The total net winnings varies dramatically over time and does not appear to converge to any constant number in this interval, although it should theoretically approach -.05263(1000)= -$52.63; the graph shows that this value may be reached soon, although more trials may be necessary to attain more accurate results.

The average appears to approach -.03, nearly the theoretical -0.05263. In this series of events, my results were "above average" since did not lose as much as I theoretically should have.

j.) **Tally for Discrete Variables: C13**

```
C13  Count
 -1    983
 35     17
 N=   1000
```

Descriptive Statistics: C13

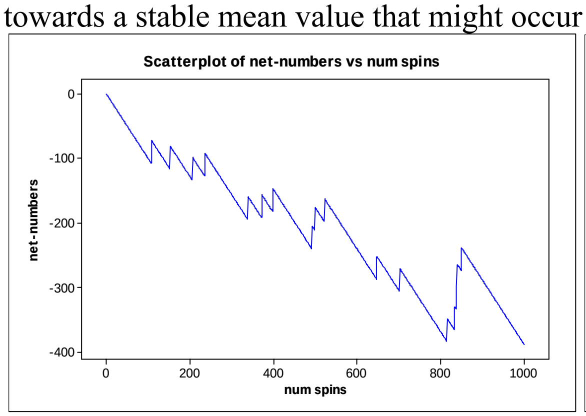
```
Variable     N  N*    Mean  SE Mean  StDev  Minimum      Q1  Median      Q3
C13       1000   0  -0.388    0.147  4.656   -1.000  -1.000  -1.000  -1.000

Variable  Maximum
C13        35.000
```

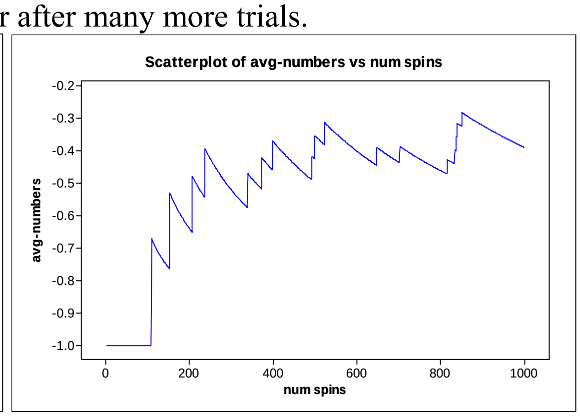
$P(X = y_1) = .97368 \approx$ Relative Frequency = 983/1000 = .983

$P(X = y_2) = .02632 \approx$ Relative Frequency = 17/1000 = .017

The relative frequencies are each within about .009 from the theoretical probabilities. While they seem fairly close, when many trials are performed, this variation can have dramatic effects on the cumulative results.

The mean is -.0388 according to Minitab, while the expected value is -.05263. This difference may be a result of not being run enough trials for the random results to tend towards a stable mean value that might occur after many more trials.



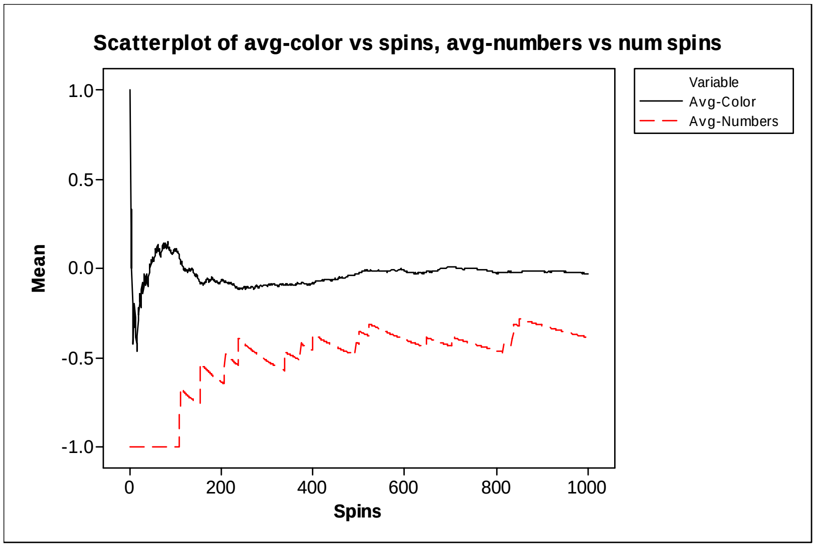


The total net winnings decreases at a constant rate until a win boosts the winnings by $35 at random intervals with a theoretical probability of .02632 of occurring. Unlike the colors net winnings, this case never broke even and only lost money, significantly more than the colors case—a difference of $356:

| (net-colors)-(net-numbers)|=|(-32)-(-388)|=$356

The average does not appear to approach a limit because the big wins create erratic behavior in the graph. The general trend appears to be upward, perhaps converging to the theoretical -.05263. Despite that possibility, it is strange that the resulting mean is -.388; a whole order of magnitude off! I would attribute this result to the fact that the relative frequency of wins is lower than theoretically predicted and the effect of such a large difference between losses and wins that can significantly alter the net gain.

k.)



Both plots are similar in that they are nearly always negative values. Otherwise they are fairly different: color follows smooth forms while numbers is rough and erratic. Also, numbers is always more negative than colors in this case. Colors tends toward a stable value on the long run, however, numbers looks to level off slightly, but is still changing rapidly even after 1000 spins—a result of the large affects of the $35 win and its small probability.

l.) $\sigma^2 = V(X) = (1-(-.05263))^2(.4737) + (-1-(-.05263))^2(.5263) = .9972\text{ dollars}^2$

$\text{SD}(X) = \sqrt{V(X)} = \sqrt{.9972} = .9986\text{ dollars}$

Units are in dollars^2 and dollars since these are the variance and standard deviation of the net gain, which is measured in dollars.

m.) $E(Y^2) = 35^2(.02632) + (-1)^2(.97368) = 33.2157$

$V(Y) = (33.2157) - (-.05263)^2 = 33.2129\text{ dollars}^2$

$\text{SD}(Y) = \sqrt{33.2129} = 5.7631\text{ dollars}$

n.) The standard deviation is larger for playing numbers. This makes sense because the difference between 1 and -1 (colors) is much smaller than 35 and -1 (numbers), therefore the difference between the expected value and the possible results will be much larger in the numbers games since the spread between the possible values is much larger—resulting in a larger standard deviation.

o.)

Standard Deviation of C3 (colors)

```
Standard deviation of C3 = 0.999988
```

Standard Deviation of C13 (numbers)

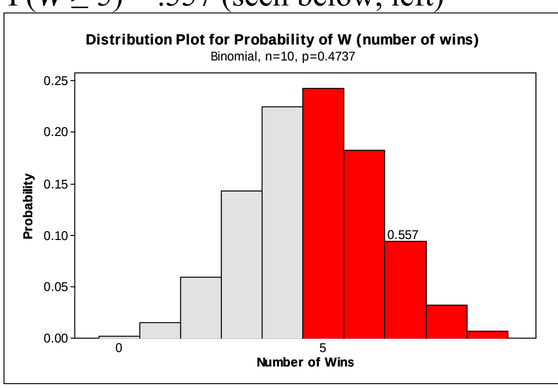
```
Standard deviation of C13 = 4.65609
```

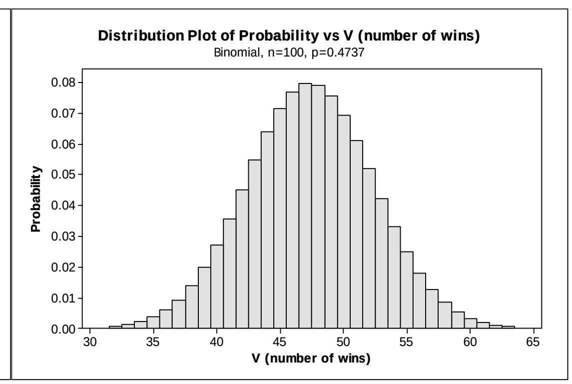
The ‘colors’ theoretical and empirical standard deviation are very similar, which makes sense because the experimental mean was similar to the theoretical one as well. ‘Numbers’ did not return a very close empirical and theoretical standard deviation, varying by more than 1. I’ll attribute this to the fact that I played an especially unlucky series of spins, creating more losses than would be expected to occur in the long run. Such a result placed more “weight” on the -1 result, pulling the mean closer to this value and creating a smaller SD.

p.) The graph in (k) reveals that the SD for ‘colors’ can be expected to be nearly 1 because it is 1 unit away from each possible result. If the scale was changed on the graph to show 35 as well it would be easy to see that the plot of ‘numbers’ remains close especially close to -1 and far from 35 indicating a larger SD. Once probabilities are considered along with the actual value of the result, the SD should lie closer to the difference between the mean and -1 than the mean and 35.

q.) *W* (number out of 10 spins that win) suits the binomial distribution because there is a discrete amount of trials (10), there are only two results (success=win or failure=lose), the trials are completely independent, and the probabilities remains constant from trial to trial.

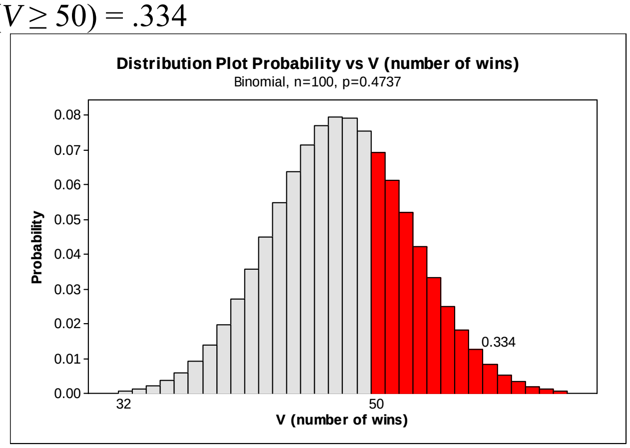
r.) P($W \geq 5$) = .557 (seen below, left)



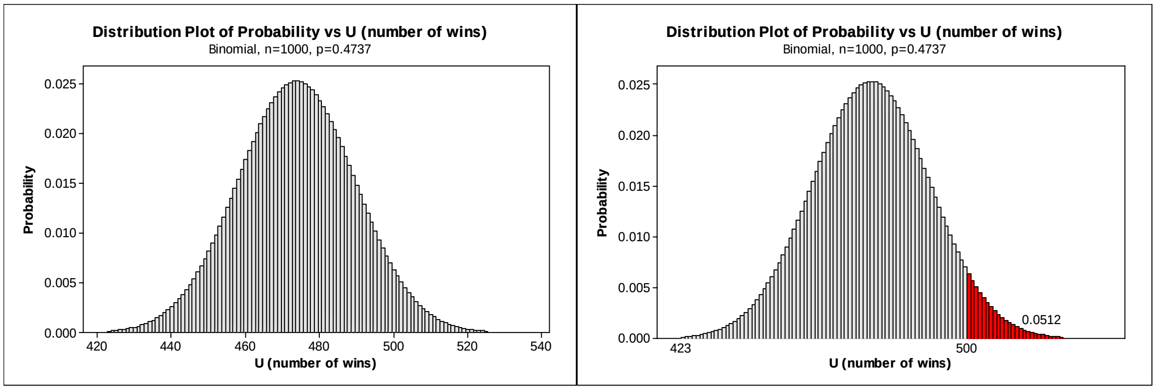


s.) The graph above (right) shows the binomial probability distribution of *V*, with *p*=.4737 and *n*=100. This is shows the probability of winning for each possible number of wins.

t.) As seen below, P($V \geq 50$) = .334



u.) The binomial probability distribution (n=1000, p=.4737) peaks at the 473^{rd} and 474^{th} plays according to Ministat, following a symmetrical shape centered around said peak with a range of about 105 plays. Winning 500 or more of the 1000 plays yields a probability of .0512.



v.)

Spins	10	100	1000
Probability of winning half	.557	.334	.0512

With increasing spins, the probability of winning more than half of the spins decreases considerably. So 10 spins has the highest probability at .557 and 1000 has the lowest probability at .0512. This makes intuitive sense because the more bins there are on the graph the more accurate the results will be (like using large rectangles vs. integrating to find the area under a graph)—likewise, the more spins have been played the closer the true odds will align with the theoretical probability thereby lowering the probability of winning half the spins.

w.) The lessons in this lab prove that high play levels bring empirical probabilities very close to theoretical probabilities, among other results. With these probabilities, it is extremely unlikely for the gambler to come away with any big gains. Although some individuals may get lucky and be among the small probability that wins, the net gain of all gamblers combined always dips into the negative as a result of the cumulative effect of many plays over time. Casinos ensure profits in many other ways as highlighted on Las Vegas travel shows, such as pricy accommodations, food, and services in addition to convincing gamblers to spend more time in the casino through psychological methods. Increased exposure to the games continues the likelihood that the casino will keep more of their guests' money.