

**Jennifer Mateja
Andrea Scisinger
Lindsay Lacher**

Stats 2002: Probabilities for Wins and Losses of Online Gambling

Abstract:

The objective of this experiment is to determine whether online gambling is a profitable hobby. In order to achieve this goal, data was collected from nine individuals who participate in this past time. The data was then analyzed by using SPSS to create various tables and graphs that allowed us to interpret the data. We were able to conclude from the results that online gambling is not a profitable hobby, and that there was no difference in the net gains/losses between soccer, hockey, and baseball.

Background Information:

Gambling can either be an individual's rise to riches or fall to rags. With technology exponentially increasing, a new form of gambling that exists over the Internet has been developed. Individuals now have the luxury to win or lose money with the click of a mouse. In order to become involved in this new craze, individuals must register for a user name and a password at one of a few selected websites such as <www.sportsbookbetting.com>.

Following the registration process, the individual may immediately begin placing bets on his or her choice of sport. The bets range anywhere from \$2 to \$50 per game. However, the question now becomes, do the odds of winning exceed those of losing for each individual sport? In order to answer this question, we decided to perform an experiment examining the probabilities of wins and losses for the three selected sports. Inferences were then made about the data in order to determine if online gambling is a profitable hobby.

Methods:

Nine individuals living on campus at YSU who are all involved with online gambling volunteered to participate in this experiment. Over a one-week period, each individual made approximately ten to fifteen online bets for various sports and activities. Three specific sports were chosen for this particular test: soccer, hockey, and baseball. A list was compiled for a one-week period containing the amount of money gambled for each bet made and also the amount of money won or lost per bet.

The data for each sport was then inputted into SPSS in order to make several comparisons between the probabilities of wins and losses. In order to make such comparisons, several charts and graphs were formatted containing the various probabilities for each sport. Comparisons were then made between the three sports. Following this, the total wins and losses were merged and new probabilities were calculated. These new probabilities allowed us to examine online gambling as a whole rather than examining the individual sports alone.

H_0 = There is no significant difference between the net gain/loss of the three sports.

H_a = There is significant difference between the net gain/loss of each of the three sports.

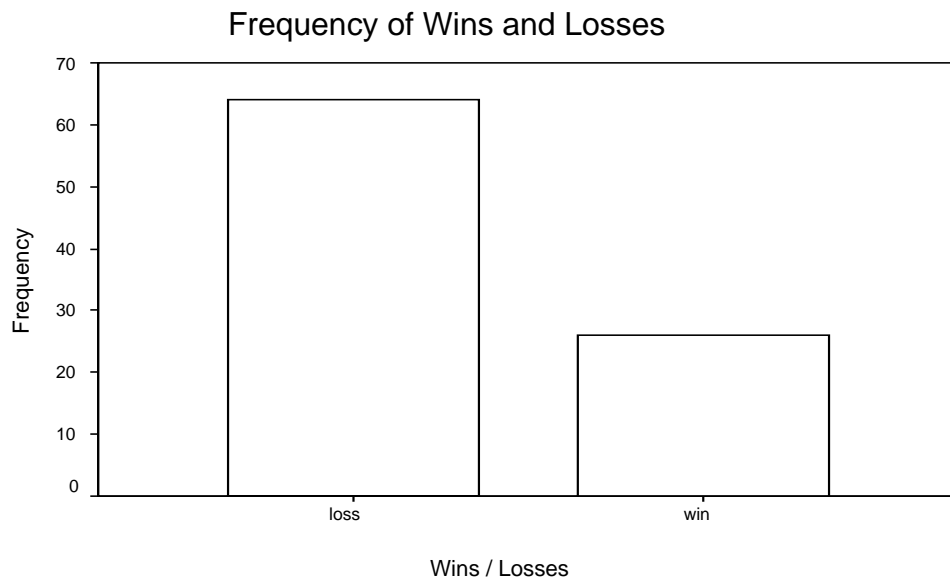
Results:

In order to analyze the data, variables were assigned to each sport when inputting the information into SPSS. Soccer was marked as '1'; hockey was marked as '2'; baseball was marked as '3.'

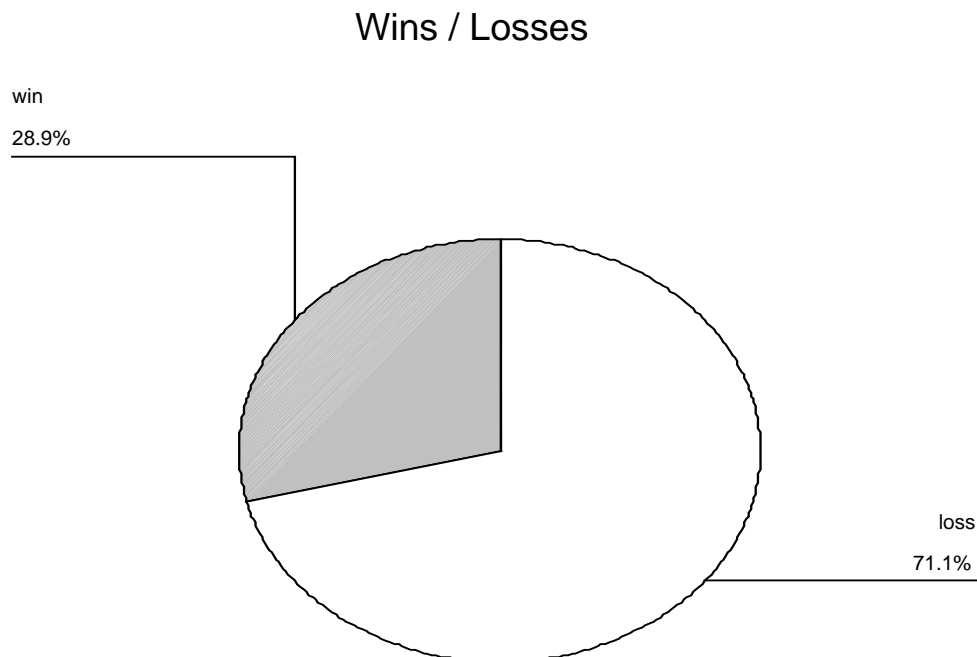
Table 1.

				Descriptives		
GAME				Statistic	Std. Error	
TRUERATE	1.00	Mean		7.667E-02	.3465	
		95% Confidence Interval for Mean	Lower Bound	-.6320		
			Upper Bound	.7853		
			5% Trimmed Mean	-.1062		
			Median	-1.0000		
			Variance	3.602		
			Std. Deviation	1.8978		
			Minimum	-1.00		
			Maximum	5.00		
			Range	6.00		
			Interquartile Range	2.3250		
			Skewness	1.390	.427	
			Kurtosis	.344	.833	
		2.00	Mean		.6944	.4526
			95% Confidence Interval for Mean	Lower Bound	-.2313	
			Upper Bound	1.6201		
			5% Trimmed Mean	.5494		
			Median	-1.0000		
			Variance	6.146		
			Std. Deviation	2.4791		
			Minimum	-1.00		
			Maximum	5.00		
			Range	6.00		
			Interquartile Range	4.3333		
			Skewness	.858	.427	
			Kurtosis	-1.187	.833	
		3.00	Mean		.1333	.3569
			95% Confidence Interval for Mean	Lower Bound	-.5965	
			Upper Bound	.8632		
			5% Trimmed Mean	-4.32E-02		
			Median	-1.0000		
			Variance	3.821		
			Std. Deviation	1.9547		
			Minimum	-1.00		
			Maximum	5.00		
			Range	6.00		
			Interquartile Range	3.2500		
			Skewness	1.280	.427	
			Kurtosis	-.064	.833	

The above chart displays the mean rate of net gain or loss for each sport per bet made. Also, the upper and lower bounds illustrate the 95% confidence interval for the gains and losses.

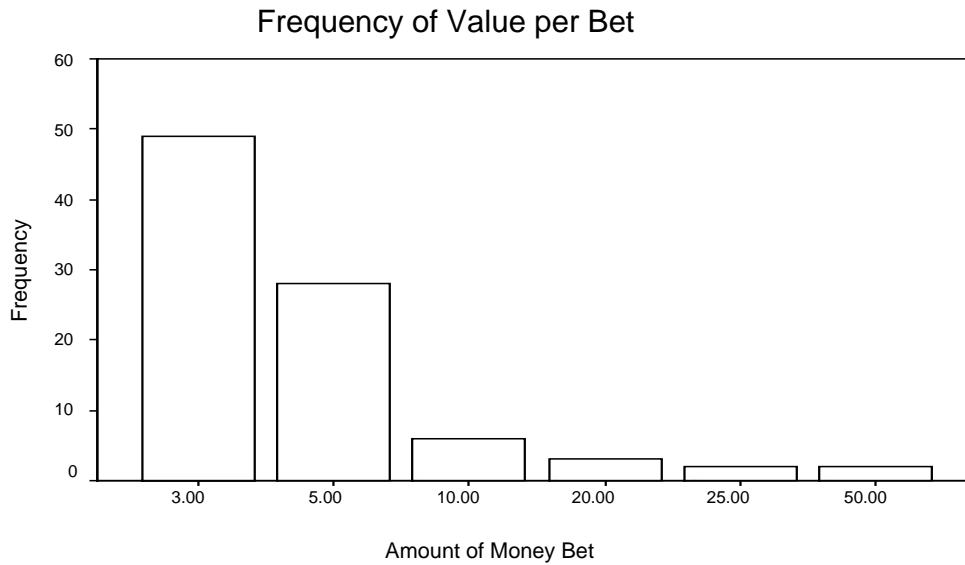
Chart 1.

The above chart displays the overall frequency distribution of wins and losses, excluding the monetary values of each bet.

Chart 2.

The above pie chart is another way to visually depict the ratio of wins to losses. It also provides percentages.

Chart 3.



The above chart displays the frequencies for the amount of money gambled per bet. The data distribution is skewed to the right meaning that most people bet smaller amounts of money.

Table 2.

TRUERATE Stem-and-Leaf Plot for
GAME= 1.00

Frequency	Stem &	Leaf
22.00	-1 .	00000000000000000000
.00	-0 .	
.00	0 .	
1.00	1 .	0
1.00	2 .	3
5.00	3 .	00333
1.00	Extremes	(>=5.0)

Stem width: 1.00
Each leaf: 1 case(s)

Table 3.

TRUERATE Stem-and-Leaf Plot for
GAME= 2.00

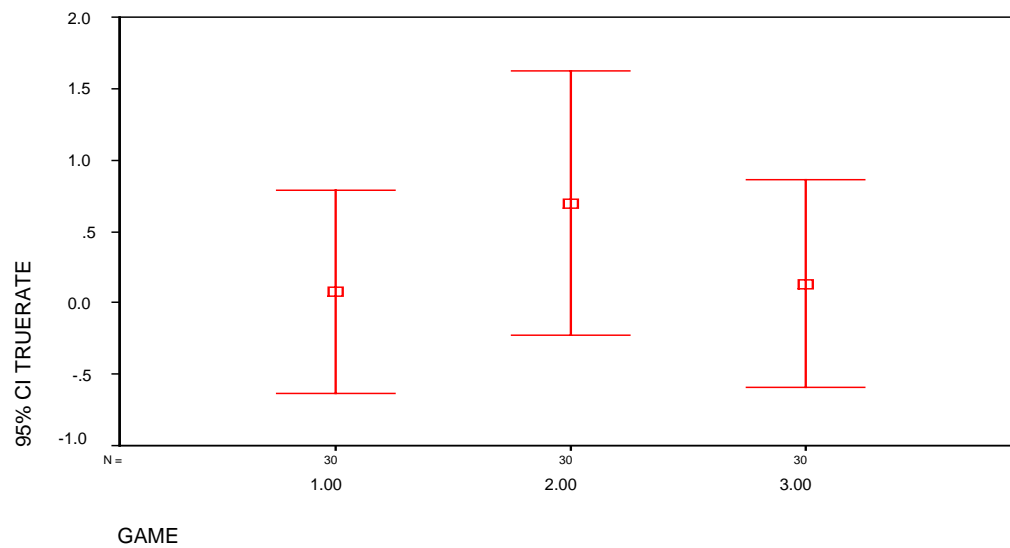
Frequency	Stem &	Leaf
20.00	-1 .	00000000000000000000
.00	-0 .	
.00	0 .	
.00	1 .	
.00	2 .	
5.00	3 .	33335
1.00	4 .	0
4.00	5 .	0000
Stem width: 1.00		
Each leaf: 1 case(s)		

Table 4.

TRUERATE Stem-and-Leaf Plot for
GAME= 3.00

Frequency	Stem &	Leaf
22.00	-1 .	00000000000000000000
.00	-0 .	
.00	0 .	
.00	1 .	
1.00	2 .	0
6.00	3 .	000333
.00	4 .	
1.00	5 .	0
Stem width: 1.00		
Each leaf: 1 case(s)		

Tables 2-4 display the frequency of the rate of gains and losses for every dollar gambled per bet. The data distribution is skewed to the right.

Graph 1.

The above graph displays the mean rate of gains or losses. It also shows error bars for each of the individual sports.

Table 5.**ANOVA**

NETGAIN					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	249.489	2	124.744	.348	.707
Within Groups	31170.167	87	358.278		
Total	31419.656	89			

The above table displays an analysis performed on SPSS that reveals if the net gain/losses between the three sports are different from one another.

Table 6.

Tests of Normality							
GAME	Kolmogorov-Smirnov ^a			Shapiro-Wilk			
	Statistic	df	Sig.	Statistic	df	Sig.	
TRUERATE	1.00	.448	30	.000	.605	30	.010**
	2.00	.420	30	.000	.646	30	.010**
	3.00	.452	30	.000	.602	30	.010**

** . This is an upper bound of the true significance.

a. Lilliefors Significance Correction

Because the stem and leaf plots were drastically skewed, we decided to perform a normality to test. The above table displays the whether or not the true rate for each sport follows a normal distribution.

Table 7.

Kruskal-Wallis Test

Test Statistics^{a,b}

	TRUERATE
Chi-Square	1.440
df	2
Asymp. Sig.	.487

a. Kruskal Wallis Test

b. Grouping Variable: GAME

The above table illustrates the results of a Kruskal-Wallis test, which is used to show comparisons for data that is not normally distributed.

Analyses:

Figure 1 shows the mean rate value for each of the three sports. For sport one, soccer, there is an average gain of 7.667 cents per bet. For sport two, baseball, an average of 69.4 cents is gained per bet. Lastly, for sport three, hockey, there is a net gain of about 13.3 cents per bet. Based on this data alone, none of the three sports analyzed shows a large margin of profit.

Chart 2 and 3 depict the frequency of the total number of wins and losses. By examining this data, we noticed that the difference in percentages between the two was quite large. Overall, 71.1% of the total bets made were wins, and 28.9% of the bets were losses, leading us to conclude that online gambling is unprofitable.

Chart 4 displays the frequency of the average amount of money gambled per bet. The data is skewed to the right meaning that the majority of the bets were for a smaller amount. As the value of the bets increased, the number of bets made at that value decreased.

Figures 2-4 are stem and leaf plots that display the frequency of the gains and losses made per dollar bet. The distribution pattern for each of the three plots is skewed to the right. This means that there is a much higher frequency of losses than gains. These results also drew us to the conclusion that online gambling is not a profitable hobby.

Graph 1 displays the mean rate of gain or loss per bet for each sport. Because the three means are very close to one another and they are all close to zero, we can conclude that none of the three sports analyzed produce a large net profit per bet.

Table 5 illustrates a comparison of the relationship between the three sports. Because the p-value was .707 which is greater than .05, we fail to reject the null hypothesis because there is not significant evidence supporting the alternative hypothesis.

Because the stem and leaf plots were drastically skewed, we decided to perform a normality test. Table 6 displays the results of this test. Because the p-values were all very small, the true rate values for the three sports are not normally distributed. Therefore, a nonparametric test was used as a second method to test the null hypothesis.

Table 7 illustrates the results to a Kruskal-Wallis test, which is a non-parametric test. This test is used to test the relationship between data that is not normally distributed. Because the p-value was .487, which is greater than .05, we failed to reject the null hypothesis again because there was not significant evidence to support the alternative. Therefore, there is no significant difference between the true rates for each of the three sports.

Discussion:

By performing this experiment, we were able to conclude that online gambling is an unprofitable hobby. By examining the various charts and graphs produced on SPSS, an obvious trend was revealed illustrating that online gambling is not profitable. By looking at the frequency distribution of wins and losses, the amount of losses exceeded that of wins by nearly 1/3. The Kruskal-Wallis test, which is used to test data that is not normally distributed, provides very important results for this experiment. Because the null hypothesis was .487, we failed to reject the null hypothesis because there was not significant evidence to support the alternative hypothesis. Therefore, there is no significant difference between the average net gain of three sports.

In our experiment, possible biases may have affected the outcome. Since the data was received from college students only, the results may not be a true representation of online gambling. Also the sample size was fairly small because we only collected data from nine

individuals. However, there was still a large quantity of values to work with. If more individuals had been used, the results may have been more accurate. In such an experiment as this, one way to avoid bias is by using a large sample size. Overall, the results indicate that an individual's safest bet is to not bet at all.