Sept. '80 73

The Mathematics of Gambling

by Edward O. Thorp

In the May column, I discussed Stephen Kimmel's system for roulette(*Gambling Times*, December 1979). Subsequently, gambling expert Allan Wilson, (*The Casino Gambler's Guide)*, wrote to differ with my use of the statistical test; Chi-Square.

Discussing Wilson's letter will help to review and to clarify my objections to the Kimmel system. Kimmel's idea is: Long-time dealers tend to spin the ball and the rotor in a habitual, regular way.

Kimmel says the rotor velocity, the ball time and the distance of travel tend to be similar from spin to spin. Therefore, he expects the ball to travel the same number of revolutions relative to the rotor and to fall about the same number of pockets beyond the last number which came up from spin to spin.

The flaw is "beyond the last number which came up." It should be "beyond the number nearest the ball when it was launched." In the May column, I noted this devastating fallacy in Kimmel's argument.

To clarify the fallacy, suppose there is a perfect dealer who always spins the ball exactly the same number of revolutions and the same velocity each time. Suppose the ball is not randomly deflected by vanes on the stator.

For discussion purposes, assume the ball always travels exactly eight revolutions on the stator and the rotor always makes exactly four revolutions. The ball has gone twelve revolutions relative to the rotor when it crosses onto the rotor. Therefore, the number nearest the ball when it was launched is also nearest the ball when it crosses onto the rotor.

When the ball enters the rotor, suppose it always goes the same additional distance, twelve pockets. We are assuming it's not deflected by the dividers between the pockets and is not disturbed by other randomizing factors. If you watch roulette, you'll see that the ball is scattered considerably once it enters the rotor. The distance it travels on the rotor varies considerably from spin to spin.

Based on these "perfect" conditions, the ball comes to rest exactly twelve pockets beyond the number it was nearest when launched. If we recorded the number that was nearest the ball when it was launched, we could tell where the ball would come out by counting twelve pockets forward. Unlike Kimmel, we would have to place our bets *after the ball was spun*, not before.

On the other hand, Kimmel claims that the ball will fall out a certain number of pockets beyond the last number that came up. For this to work, the dealer would have to launch the ball just as the last number is going by. Can you imagine the dealer keeping track of the previous winning number and then waiting to launch the ball just as it goes by? This would be necessary for Kimmel's system to work.

In a casino, the ideal case we've assumed is far more perfect and favorable than Kimmel's systems. Even with this ideal situation, Kimmel's system gives no edge at all. It is absolutely worthless. To know where the ball will stop, we need to know the number that was nearest the ball at launch time. Kimmel's system does not use that information. It simply records what number won on the previous spin.

In our observation, the dealer pays no attention to what number is passing as he launches the ball. He certainly doesn't wait for a particular number to go by in order to launch the ball. Yet for Kimmel's system to work, he would have to key the launch of his ball to the previous winning number.

The number which is passing as the dealer launches the ball appears to be random. Even in our ideal case, the position that the ball stops at is twelve pockets past a random starting position which makes it random position.

In a realistic case, where the dealer is less than perfect, the distance and velocity of the ball's travel varies, as does the rotor. The final position where the ball comes to rest is even more random. Thus, there is absolutely no validity to Kimmel's system.

As I pointed out, after the ball was launched, if we were to observe the number nearest to it, and record where it came to rest, then the system would be plausible. We might hope that a graph like Kimmel's would indicate some advantage to the player. Whether such a modified system would work could be settled by observation.

Based on observation and measurements, my experience is that the level of regularity required by the dealer is far beyond human capability.

As I pointed out in the October 1979 column, the standard error around the estimated final position of the ball must be less than seventeen pockets for the player to overcome the house advantage. There are many contributing factors beyond variation in human performance such as deflection by vanes and scattering by pockets on the rotor. The error is nearly this large before we add in the human factor.

If there is no validity to Kimmel's system, what are we to make of Kimmel's data and his reported success at the tables? One interesting thing we can do is apply statistical tests to Kimmel's data to see how remarkable the result is.

Using various statistical tests, we can ask this question: "If the distribution of numbers was truly random in a sample of 199 spins, (which is what Kimmel reported), then how often are the results skewed and unusual looking as those Kimmel reported?" I used the Chi-Square test in the May coulmn. Relative to that test, I found that the chance was one

continued on page 89

Mathematics of Gambling

continued from page 70

in 80 of a result as unusual as Kimmel's.

Here's Wilson's condensed letter, stating his objection to the use of the Chi-Square test:

"In his column, 'The Mathematics of Gambling,' (May 1980 issue), my distinguished friend, Ed Thorp, expressed skepticism on the roulette prediction data given by Stephen Kimmel, (December 1979 issue). Dr. Thorp's objections stem from personal experimentation, analysis and an application of the classical Chi-Square test. Also, he suggests that Kimmel may have selected his best of five tests for illustration.

He concludes from his Chi-Square calculation that there is a one in 80 probability that Kimmel's distribution was due to chance alone. This is 'unusual, but it is certainly not remarkable.'

I contend that the Chi-Square test is virtually inapplicable to this case. The probability of a distribution like Kimmel's occurring, due to chance alone, is far less than one in 80. The flaw is that Chi-Square would generally pertain to an ensemble of deviations, (positive and negative), scattered anywhere along the abscissa of 38 pockets.

In no way would it focus on the unlikely occurrence of long strings of deviations of the same polarity. Kimmel shows 9 side-by-side "pockets" of positive deviation, followed immediately by 13 negative (or 0) deviation on one side and 7 of negative deviation (or 0). That is remarkable, if it's true.

A proper theoretical approach for inference from this sample still eludes me. It is tempting to apply some simplistic notions from the theory of runs. I would willingly bet Thorp my last buck that in ten's of thousands of samples of 200 roulette plays on a computer, he would rarely (or maybe never) see a deviation pattern comparable to the long strings of deviations of like polarity in the Kimmel sample!" —Allan N. Wilson

I agree with Wilson's objection to my use of the Chi-Square test. The Chi-Square test measures the overall extent of the deviations from the expected frequency of occurrence in each of the individual pockets. If individual pockets or numbers were randomly biased, Chi-Square test would be appropriate in picking that up.

This is somewhat different than the effect that we're trying to find in Kimmel's data. There, we would expect strings (or runs) of pockets that occur more often than average and less often than average. We want to see if the graph which shows this is really unusual.

Wilson suggested that I use a runs test on Kimmel's data to see how remarkable it looked from that viewpoint. I found that the unusualness of the data varied anywhere from one chance in several hundred to one chance in a few thousand depending on how the runs test was applied.

Here is the problem in using the runs test: We started by choosing some level like 5 occurrences of a number. We decided that above that level, a number would be called a "head" if it showed up. Below that number of occurrences, a number would be called a "tail."

In that case, whenever a number in Kimmel's sample occurred four times or less we wrote down "T" for tail. Whenever it occurred six times or more, we wrote down "H" for head. If it occurred exactly five times, we omitted it. Then, we looked at the series of T's and H's, and used standard runs tests from statistics, to determine how likely the strings of T's and H's would occur by chance.

This illustrates the first problem in applying the runs test: Why pick five as our point of separation between success and failure? Why not pick 4½ or 5½ or 4 or 6? The division point is arbitrary and the significance of the data, according to the runs test, depends on that arbitrary choice.

Another problem with the runs test is it does not measure the degree of deviation from average. If all numbers on one-half of the wheel occurred six times and on the other half of the wheel occurred four times, than that would be just as remarkable according to the runs test.

It would be just as remarkable if numbers on one-half of the wheel occurred 10 times for each number, and occured 0 times on the other half of the wheel. Nonetheless, the runs test does show us that Kimmel's data is more remarkable than 1 chance in 80.

Subsequently, I have devised the best possible test which I think captures the effect we are after. Imagine a picture of the roulette rotor drawn as a circle with a one unit radius or one inch. The 38 pockets are evenly spaced around the circle.

From the center of the circle, the angle from one pocket to the next increases in steps of 360/38 degrees. If single 0 is at 0 degree position, by measuring angles counter-clockwise, the next number will be at 360/38 degrees and so on.

For the first number in Kimmel's sample, start at the center of the circle. Draw a little arrow out to that number. For the second number, draw a similar arrow and add it to the first arrow by fastening its tail to the head of the first arrow. Be sure the second arrow points in the same direction it was first drawn.

For each new number, continue to do this with little arrows, (or vectors), one inch in length. Point them at various angles, adding up, until there is a path of segments one inch long bending and twisting on a flat sheet of paper. If there is a pile up of outcomes on one side of the wheel, there will be a preponderance of arrows pointing in that direction. The grand sum of all the arrows, represented by the distance from the starting point to the end of the path, will point in the direction where the numbers are piling up.

My statistical test considers a result remarkable when the distance from the starting point to the end of the path is unusually large in respect to the number of observations that have been made. This is a problem discussed in William Feller's classic book of probability theory.

There seems to be no exact formula for determining how rare the Kimmel data is viewed from this standpoint. There is an approximate formula which says it would happen by chance about 1 time in 6,000.

I cross checked this result with computer simulation. I found that the Kimmel data seems more rare than this. At this point, I would guess that the likelihood of it occurring by chance would be less than 1 in 10,000.

I believe this solves the problem of testing the statistical significance of Kimmel's data. It is quite significant. According to my reasoning, the system is nonsense. I believe that the data is either a rare, lucky event that Kimmel chanced upon or. . . 9