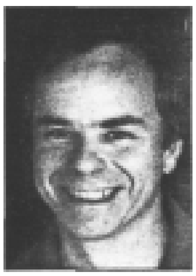
Chances Are...

What does it means to play the percentages?

by BOB JEWETT



BASEBALL FANS ARE fascinated by statistics; the neat columns of figures in the morning paper distill the essence of yesterday's games. Slumps, streaks and records are carefully noted. Why not at pool?

Ten years ago, Pat Fleming of Accu-stats published a monthly series of summaries of pro pool tournaments, including some inning-by-inning score sheets. Regrettably for those of us hooked on numbers, the pool world was not ready at that time for such detailed reports. They contained the life-blood of probability and statistics — real-life performance data.

The study of probabilities at pool is a difficult technical part of the game, but also one of the most interesting. A thorough treatment could easily fill a book; below are some highlights of how to put random events into a consistent framework.

A very common application of probability theory is to predict how often something will occur, for. example a run of ten racks of 9-ball, or a nuclear plant meltdown. Such things happen so rarely that it's hard to get a good handle on them. Usually statisticians will solve such a problem by making some assumptions, looking at history (like Fleming's data), and making a projection or extrapolation.

Let's make the following assumptions about the running a rack of 9-ball:

- Each rack is uninfluenced by previous racks, so it doesn't make any difference to the player if it's the first or tenth rack in a row.

(Reasonable for coin flips, but does it apply to people? That's why I'm calling it an 'assumption.') Further assume:

- That a player's skill doesn't vary much from day to day or table to table.

"Whoa! What about...?" you may be saying right now. I'll refine these assumptions later, but for now let's see where this leads.

Accu-stats has a specific category called "runout from the break." Top players like Earl Strickland keep their opponent seated from the break through the 9 ball about 25 percent of the time. To get the chance of two consecutive racks, just multiply that chance by itself to get one in sixteen.

$$\tfrac{1}{4} \times \tfrac{1}{4} = \tfrac{1}{16} = 6.25\%$$

For a third rack, multiply by a quarter again to get one chance in 64.

$$\tfrac{1}{4} \times \tfrac{1}{4} \times \tfrac{1}{4} = \tfrac{1}{64} = 1.563\%$$

For 10 racks, this works out to roughly one chance in a million.

$$\left(\tfrac{1}{4}\right)^{10} = \tfrac{1}{1,048,576}$$

Without more information, this is just a "best guess." For example, suppose we find out that a particular table is a little easier than average, and Earl runs out 31 percent of the time, instead of 25 percent. Maybe this is due to the 9 ball going in on the break 15 percent of the time rather than the more typical three percent. This apparently small improvement in single-rack percentages changes the 10-rack chance to roughly 1 in 100,000, a ten-fold improvement.

$$31\% = \tfrac{31}{100}$$

Running 10 racks:

$$\left(\tfrac{31}{100}\right)^{10} = \tfrac{1}{122,007}$$

In a race longer than 10 games, another factor comes in — it's possible to start the run of 10 after a miss in an early rack. It's mostly a matter of tedious bookkeeping to count up all the possible series of misses and runs that have a run of 10 in them. Using a statistical tool known as Markov's chain, it works out that in a race-to-15 match, the odds of running 10 racks improve to 1 in 22,000, and that's just for one player in one match.

In such a situation, the accuracy of the odds estimate is questionable because the conditions are so poorly known. When masses of good data are available, it's possible to estimate how close the observed averages are to the "ideal" long-term average. An everyday example of this is public opinion polls, where an estimated margin of error is often given. The more people you randomly poll, the closer the observed average will be to the national average, and the smaller your margin of error will be. A poll with a three percent-stated margin of error indicates about 1,000 people were polled.

For an example at billiards, let's look at Raymond Ceulemans' record-setting performance in the 1978 World Three-Cushion Championships.

- He scored 660 points in 393 innings.
- He made eight of 10 break shots.
- He had 145 "open innings," when he missed his first shot from his opponents' leaves.

With only 10 break shots, we might expect a normal variation from 6 of 10 to 10 of 10, for an 80 percent average. So there's a 20 percent uncertainty on this average.

$$\tfrac{6}{10} \quad \pm\tfrac{2}{10} \quad \tfrac{8}{10} \quad \pm\tfrac{2}{10} \quad \tfrac{10}{10}$$

$$60\% \quad \pm 20\% \quad 80\% \quad \pm 20\% \quad 100\%$$

Ceulemans' overall scoring accuracy was 63 percent with 660 points in 1042 shots. The margin of error is about three percent, so we can say his "innate" average* was between 60 and 66 percent. '

The runner-up in that tournament, Nobuaki Kobayashi, had a 55 percent scoring percentage, also with a three percent margin of error, so Ceulemans was clearly the best player in that tournament.

The arithmetic for calculating the "margin of error" for a percentage that's roughly 50-50 is to take one over the square root of the number of shots or people polled.

$$\frac{1}{\sqrt{1000}} = \frac{1}{31.62}$$
$$= 0.03162$$
$$= 3.162\%$$

There is about a 1 in 20 chance that the true value is more than the margin of error of the observed value. (1 won't explain all *that* arithmetic.)

In their head-to-head encounter in the final match of the tournament, Ceulemans made 72 percent of his shots, while Kobayashi slumped slightly to 52 percent. The match to 60 points was only 25 innings long, so a large fluctuation in percentages is not too surprising.

One other thing we can check from the score sheets is how well Ceulemans plays position. To do this, compare his scoring percentage from his own leaves versus the shots his opponents leave him. Having seen him play position and his opponents do their best to leave him nothing, I'd guess 75 percent and 50 percent for the two accuracy percentages. Here are the raw numbers so you can do the calculation yourself as homework: his own leave, 414 of 639; his opponent's leave, 238 of 383. Don't forget to calculate the margins of error, too.

(If you want your homework graded, send it to me c/o *Billiards Digest,* or via E-mail at *jewett@uelconi.com.)*

Where's the practical application of all this theory? That's why you've waded through all this math, isn't it? The main point you can use in your own play is that you shouldn't jump to conclusions from a small number of observations. Missing five shots in a row that are 50-50 for you does not constitute a slump. The odds of that happening are 1 in 32 — not that uncommon.

$$= \left(\tfrac{1}{2}\right)^5 = \tfrac{1}{32}$$
$$= 3.125\%$$

Miss 15 straight and you may have a problem to work on. The odds of 15 in a row solely because bad luck is 1 in 32,768.

$$= \left(\tfrac{1}{2}\right)^{15} = \tfrac{1}{32,768}$$

Check my two earlier columns on "progressive practice" for an efficient way to measure your pocketing accuracy.

Similarly, in an evenly matched race-to-five match, one of you will be on the short end of a 5-0 score about six percent of the time purely by chance. Do the math yourself and you'll see.

Relax, don't let the luck — good or bad — get to you, and look at each new game as a fresh beginning.

* "Innate" average should not be confused with what most three-cushion players call their "average." That average is determined by dividing points by innings. For example, Ceulemans' average at this event was 1.679.

$$= 660 \text{ points} \div 393 \text{ innings}$$
$$= 1.679 \ \text{Points}/\text{Inning}$$

Bob Jewett is a former ACU-I billiard champions and ***currently*** *trains BCA instructors in San Francisco.*