13. WINNING AND LOSING STREAKS

It was not too hard to calculate the probability of Susan Lucci's losing streak, since it started with her first Emmy nomination. It was just q^r, where $r = 28$ = number of losses in a row, and q = prob{Lose one award}.

But what is the probability that at least one losing streak of length r will occur **somewhere** among n independent rounds of a game? Assume that each round is either won or lost.

If r is less than n, this probability will be **larger** than q^r, since there are many opportunities to start a losing streak.

A good approximation for this probability, if q^r is small, is Prob{≥ 1 losing streak of length r in n trials} ≈ $[1+(n-r)p]q^r$.

This is indeed greater than q^r, as expected.

Eg: In a roulette doubling system, we bet 1 chip on Black, and each time we lose we double our bet. As soon as we finally win, we have a profit of one chip. Many people find this system appealing (I've tried it myself), although we should already be suspicious of it because of the expected value calculation we did in a previous handout.

The table maximum (which was designed with us in mind!) allows a maximum bet of 1000 chips. This allows us to play our "system" until we get 10 losses in a row. (Furthermore, the 10 losses in a row will leave us fairly "broke", since we will be down by 1023 chips from when the losing streak started).

Suppose you can play 500 rounds of this game in an evening. The probability of losing any given round is $q = 20/38 = 0.526$. Most people would calculate the probability of a losing streak of 10 to be $q^r = (0.526)^{10} = 0.00162$, or about 1.6 out of 1000.

Thus the system seems virtually unbeatable.

What they don't realize is that the probability of getting at least one losing streak of length 10 **at some point during the evening** is much higher. Higher by how much? By the first term in the formula above, $[1+(500-10)(0.474)] = 233.26$. So the actual probability of at least one streak in the evening is $(233.26)(0.00162) = 0.378$. Now we can see how dangerous the system is, since we have a 38% chance of going "broke".

(That's what happened to me when I tried the system!)

Here's a similar formula for winning streaks, assuming p^r is small:

Prob{≥ 1 winning streak of length r in n trials} ≈ $[1+(n-r)q]p^r$.

Eg: Joe DiMaggio had the longest baseball hitting streak in Major League history, successfully getting a hit in 56 games in a row. (The closest anyone else has come up until now is Pete Rose, with 44).

DiMaggio played in 1,736 games, and had a lifetime batting average of 0.325. Let's try to calculate the probability of such a streak, for such a player.

In the article called "The Great Streak" in Chance Magazine, Giles Warrack estimates the probability p of at least one hit in a given game for a player with a batting average of 0.325 to be 0.777.

(Note that p is larger than 0.325, since a batter typically has more than one "at bat", or chance to get a hit, per game).

Now, plugging into the formula with $p = 0.777$, $n = 1{,}736$ and $r = 56$, we get

Prob {a streak Like DiMaggio's} =

$$= [1+(1736-56)(0.223)](0.777^{56}) = 0.000274,$$

or about 1 in 3,700. So even for a player with DiMaggio's ability and number of at-bats, such a streak is a fairly rare event!

Table 1—Batting Averages (A), and the Probability of at Least One Hit in a Game (p)

A	p
.225	.627
.250	.671
.275	.710
.300	.745
.325	.777
.350	.806
.400	.855
.406	.860

Table 3—Batting Average (A), Games Played (n), and Number of Players Needed for an Even Chance of a 50-Game Streak Being Obtained (k)

A	n	k
.350	1000	181
.350	2000	88
.350	3000	59
.400	1000	12
.400	2000	6
.400	3000	4

Chance Magazine, vol. 8, No. 3, 1995, pp. 42-43

Table 2—Players With a 30-Game Streak

Name	Streak	Year	Batting average	Games	AB/game
DiMaggio, J.	56	1941	.325	1,736	3.93
Rose	44	1978	.303	3,652	3.94
Sisler	41	1922	.340	2,055	4.02
Cobb	40	1911	.366	3,034	3.76
Molitor	39	1978	.300	1,437	4.06
Holmes	37	1945	.302	1,320	3.78
Cobb	35	1917		See above	
DiMaggio, D.	34	1949	.298	1,399	4.03
Santiago	34	1987	.264	431	3.62
McQuinn	34	1938	.276	1,550	3.71
Sisler	34	1925		See above	
Manusch	33	1933	.330	2,009	3.81
Hornsby	33	1922	.358	2,259	3.62
Landreux	31	1980	.268	1,264	3.24
Carty	31	1970	.299	1,651	3.40
Davis	31	1969	.294	1,999	3.61
Rice	31	1924	.322	2,404	3.86
Brett	30	1970	.310	2,137	3.81
LeFlore	30	1976	.288	1,099	4.05
Musial	30	1950	.331	3,026	3.63
Goslin	30	1934	.316	2,287	3.78
Speaker	30	1912	.344	2,789	3.66