WINNING PROBABILITIES OF LOTTO IN THE U.S.A.

Shayle R. Searle

Biometrics Unit, Cornell University, Ithaca, N. Y.

BU-1375-M October 1996

Abstract

Almost all Lotto games in the U.S.A. consist of choosing either 6 or 5 numbers from n numbers. The value of n greatly affects the probability of winning (or sharing) the jackpot; additionally, the number of "hits" for which prizes are given affects the probability of winning *some* prize (however small, e.g., a free ticket in an ensuing game). These two probabilities, and the features which determine them, are discussed and summarized for Lotto games available in 37 states (including Puerto Rico) of the U.S.A.

INTRODUCTION

Recent years have seen an increasing spread of various versions of Lotto throughout the individual states of the U.S.A.; and in overseas countries, too. Almost all these versions are adaptations of the Genoese lottery discussed by Bellhouse (1991). Earlier origins, suggests Morton (1990), go back some two thousand years, and can be traced to the Han Dynasty in China. Whatever the origins, the current forms have a player buy a ticket containing r (usually 5 or 6) numbers from the first n (54 in New York) integers. Prizes are determined by the r numbers selected at random, without replacement, from the numbers 1 through n. We call those r numbers the draw, and refer to such a lottery as r/n Lotto.

A player's ticket participates in the jackpot if its r numbers are the same as the draw. A ticket having $w(\leq r)$ of the r numbers that are in the draw is described as having w "hits". For $w = r$ the prize is the jackpot or if k such tickets have been sold, each is usually awarded one k'th part of the jackpot. This is called the parimutuel effect. For some limited values of w (less than r) prizes are also awarded. For these tickets the parimutuel effect is very dilutory of prizes because, for example, when

$r = 6$ there are six different combinations of five hits; and hence $6(n-6)$ possible different tickets each having five hits. And there are 15 different combinations of four hits and $7.5(n-6)(n-7)$ possible different tickets each having four hits. The pool of prize money for five hits is much less than the jackpot, and although the pool for four hits is usually more than that for five (because there are more possible tickets), the parimutuel effect often reduces the prize per winning ticket to a relatively small amount (e.g., $1,500 for five hits and $25 for four hits).

The parimutuel effect affects just the size of the prize that a winning ticket gets, it does not affect the probability of a ticket winning a prize, i.e., of having w hits. But, of course, the parimutuel effect cannot be quantified until after the draw, when the actual number of winning tickets is known. Furthermore, the number of tickets sold is affected by the publicized size of the jackpot, the rate of ticket sales increasing rapidly when the jackpot is known to be getting large, say $15,000,000 or greater. And of course, the total number of tickets sold is not restricted by any pre-set upper limit; it is limited solely by the time available for selling (four days in New York State for the Wednesday drawing and three for the Saturday draw). Under these circumstances one cannot calculate the expected prize value for a ticket.

Of the many forms of lottery available in most states, that which generates the largest prizes is r/n Lotto. Others include picking an integer between 000 and 999, often called "Pick 3" or, similarly, "Pick 4". And there are the scratch-off games, where one scratches a ticket to remove the covering from a set of symbols, certain combinations of which yield cash prizes. This paper is confined to r/n Lotto with attention being directed to two probabilities: that of winning (i.e., winning or sharing) the jackpot; and that of winning some prize however small it may be (e.g., as little as a free ticket in a subsequent game). These are, after all, the probabilities that appeal to the ticket buyer. For example, to an occasional buyer of Lotto tickets, New York's 1987 changing from 6/48 Lotto to 6/54 Lotto might have been though inconsequential. It is not. It reduced the probability of a ticket winning the jackpot from .000,000,0815 to .000,000,0387 or, equivalently from one chance in (approximately) 12.3 million to one in 25.8 million, a reduction of 52.5%.

To get some feel for how small these probabilities are, consider a football field: it is 160 feet wide and 360 feet long; and a dime is 11/16 inches in diameter. For New York's 6/54 Lotto, the number of possible different tickets is 25,827,165. Placing that many dimes in rows and columns touching each other would cover 1.47 football fields. If you owned just one of those dimes, the probability of having a jackpot-winning ticket in a New York Lotto game is the same as the probability that a randomly chosen dime from the 1.47 football fields is your dime. Small chance! (If instead of dimes one used dollar bills, which measure $2\frac{1}{2}$ inches $\times$ 6 inches, some 46.7 football fields would be needed -- almost 62 acres!)

DEFINITION: A TICKET

In r/n Lotto we define each combination of r numbers out of the possible n numbers as a "ticket". The possible number of such combinations will be denoted by N. (The formula for N is given in the Appendix.) Making this definition of a ticket is necessary because almost everywhere \$1 buys a slip of paper which, in some states, shows one ticket (as just defined), whereas in other states it show two tickets. In almost all states where winning odds are published, they are the odds based on \$1, so that in the two-tickets-for-a-dollar states the published odds appear more favorable than in the one-ticket states. For purposes of comparison, we therefore consider all probabilities on a per-ticket basis, thus ignoring the cost of a ticket; we also ignore the parimutuel effect because, as already explained, it cannot be quantified before the draw.

A ticket with five hits also has five different sets of four hits. But it wins only one prize, that for five hits. And this is true in general: a winning ticket wins only one prize, that for the largest number of hits that it contains. Thus in considering prize-winning possibilities we need consider only the probability of a ticket's winning a single prize, that for its total number of hits.

THE SOURCE OF DATA

In October 1995 addresses were obtained from the Internet of the lottery offices of those states (including the District of Columbia and Puerto Rico) having lotteries. Fourteen states at that time had no lottery. They are listed in Table 1. Those with lotteries replied mostly with brochures describing their state's lottery games (including Lotto), with ancillary information about probabilities, jackpot sizes, allocation of sales receipts to the different prize-winning categories, and so on. Not all states provided information on each of these (and other) characteristics of their Lotto games. But simply by knowing the values of r and n for r/n Lotto (e.g., 6/54 in New York and 5/38 in New Jersey), together with the minimum number, m say, of hits that will win a prize, one can easily calculate the two probabilities of interest: that of winning (or sharing in) the jackpot, and that of winning a prize. These probabilities are shown in Tables 2 and 3, calculated from the formulae presented in the Appendix.

SUPPLEMENTAL HITS

The Lotto games in most states (some of which do have more than one game, e.g. Pennsylvania has 6/48, 6/33 and 5/39) are straight Lotto. But a few states have variations thereon in the form of awarding prizes not only for fewer than r hits but also for fewer than r hits in combination with also hitting some other random drawing. For example, in New York's 6/54 and Pennsylvania's 6/48 Lottos, tickets with three hits win no prize but they do if they also hit a seventh number drawn from the numbers not in the jackpot draw.

Variations of this are Pennsylvania's 6/33 wherein $r = 6$ and $n = 33$ and the 16-state "Powerball" of 5/45 with $r = 5$ and $n = 45$. Prizes in these are awarded to tickets with $r - 2$ or more hits, and to any ticket hitting the $(r + 1)$'th number drawn from 1 through n, no matter what number of hits (including none) there are with the regular draw of r numbers. The total number of tickets in these cases is denoted by $N^{+} = nN$ in Tables 2 and 3.

A different twist is the Tri-State 5/31 of Maine, New Hampshire and Vermont. It gives prizes to tickets having 5, 4 or 3 hits with or without hitting the "Wild Card" and for 2 hits with hitting the "Wild Card", the latter being a random selection from Ace, King, Queen, Jack or Ten. This leads to $N^{+} = 5N$, as seen in Table 3.

DISCUSSION

The 6/n Games

Table 2 summarizes the 6/n games reported, with n ranging from 54 to 33. For each n, the value of N is shown to the nearest 1000; and the probability of having a jackpot-winning ticket, namely $P = 1/N$, is shown correct to 9 decimal places. The smallest of these P-values is $1/N^+ = 1/30{,}550{,}000 = .000{,}000{,}027$ for Pennsylvania's 6/33 with a supplemental number. The next smallest is $1/25{,}827{,}000 = .000{,}000{,}039$ for New York's 6/54. The largest is $P = 1/1{,}948{,}000 = .000{,}000{,}513$ for the 6/36 games in Delaware, Wisconsin and New Hampshire.

For various values of n, there are 10 states that give prizes for four or more hits, 27 for three or more, and two states for two or more hits. States are therefore listed in Table 2 under the headings of $m = 4$, 3, or 2, being the minimum number of hits for which a prize is given. And under the state names is given the corresponding probability (to two significant digits) as a fraction with denominator N (or N^+), the numerator being the possible number of winning tickets using the formulae in Table 4, which is in the Appendix. For example, the numerator for Illinois is $1 + 1.5(n - 6)(5n - 31)$, which with $n = 54$ is $1 + 1.5(48)239 = 17209$. The numerators for 6/54 in New York and 6/48 and 6/33 in Pennsylvania do not come directly from the formulae of Table 4 because of the supplementary hits described in the preceding section.

The largest probability of winning *some* prize is 0.16 in the 6/48 game in Indiana. The second largest is 0.058 in the 6/33 game played in Idaho, Kansas, Montana and South Dakota. The smallest probability is 0.00067 in the 6/54 game in Illinois and the second smallest is 0.0097 in the 6/49 game in Maryland and Michigan.

The 5/n Games

Table 3 summarizes the 5/n games with the same layout as Table 2 for the 6/n games. The smallest probability of winning a jackpot is $P = 1/N^+ = 1/54{,}979{,}000 = .000000018$ for the 16-state "Powerball" 5/45 game. The next smallest is $1/N^+ = 1{,}850{,}000 = .0000012$ for the Tri-State (Maine, New Hampshire and Vermont) 5/31 game with the supplementary hit of one of five playing cards. For straight 5/n Lotto the largest P is $P = 1/66{,}000 = .0000152$ for the 5/26 game offered in Florida and Louisiana.

The largest probability of winning *some* prize is 0.138 in the 5/35 games played in South Dakota, Iowa and New Hampshire, where prizes are given for two or more hits. The smallest probability is 0.010 for the 5/39 game in California, Maryland, Michigan and Minnesota, where prizes are given for 3, 4 or 5 hits.

Comparing 6/n and 5/n games

Comparisons of Tables 2 and 3 are difficult because only for n of 39 and 38 are there Lottos reported for both 6/n and 5/n. As is to be expected, the probability of a jackpot ticket is greater in 5/n than 6/n, some 5.5 times as great for n in the 39 and 38 range, and more than 8 times as great for n = 54; and some 4 times as great for n = 30. The exact ratio, N_5/N_6 is obtainable from the Appendix as $(n-5)/6$.

In contrast, the probability of winning *some* prize in 6/n is usually larger than in 5/n. This is primarily because there are more ways of winning. For example, the ratio of the probability of winning *some* prize with 5 or more hits in 6/n to that in 5/n is (from Table 5 in the Appendix) 35.4 for n = 54 and 34.8 for n = 30; for *some* prize with 4 or more hits these ratios are 8.6 and 8.2 for n = 54 and 30, respectively. This means that in 6/n one has a much more favorable probability of winning some prize than in 5/n. This is not at all unexpected because for winning some prize with 4 or more hits in 6/n means one can win with 6, 5 or 4 hits, whereas in 5/n one wins with just 5 or 4 hits. For each of the latter there are appreciably fewer ways of winning than in 6/n. Of course the situation is reversed when comparing winning with 6, 5 or 4 hits in 6/n to winning with 5, 4 or 3 hits in 5/n. The ratio then is .18 for n = 54 and .33 for n = 30. It is 1.00 for the impractical values of 6, 7 and 11, and less than 1.00 otherwise.

EXPECTED PRIZES

As has been explained in the Introduction, prizes in many of the Lotto games cannot be determined until after the draw, when the parimutuel effect can be taken into account. In these situations the expected value of a prize in such a Lotto game cannot be calculated. However, with some states information is available as to how many cents from each $1 of sales is retained for operating the lottery and how many cents are allocated to each prize pool. Then, for example, if 11c is allocated to the 4-hit pool, as in New York, the prize for each *sold* ticket that has 4 hits will be

$$\$(.11)(\text{total number of all tickets sold}) \big/ (\text{total number of 4-hit tickets sold}) .$$

This would be the 4-hit prize that would have to be multiplied by the probability of a ticket having 4 hits as part of calculating the expected prize from a ticket. And clearly, this 4-hit prize cannot itself be derived until after sales have ceased and the draw has been made.

REFERENCES

Bellhouse, D. R. (1991). The Genoese lottery. *Statistical Science* **6**, 141-148.

Morton, R. H. (1990). Lotto, then and now. *The New Zealand Statistician* **25**, 15-21.

Table 1

STATES NOT HAVING LOTTO

as of October 1995

Alabama	North Carolina
Alaska	North Dakota
Arkansas	Oklahoma
Hawaii	South Carolina
Mississippi	Tennessee
Nevada	Utah
New Mexico	Wyoming

Table 2

States having 6/n Lotto, showing probabilities of a jackpot-winning ticket, and of winning *some* prize.

n = Number of numbers in the lottery	P = Probability of having a jackpot-winning ticket (6 hits)		States and Probability (2 significant digits) of winning some sort of prize with m or more hits.		
	N(nearest 1000) for P = 1/N	P (9 dec.pts.)	m = 4	m = 3	m = 2
54	25,827,000	.000000039	IL 17209/N = .00067	NY 38829/N = .0015	
51	18,009,000	.000000056		CA 298,921/N = .017	
50	15,891,000	.000000063		TX 279,335/N = .017	
49	13,934,000	.000000072	MD, MI 13804/N = .00097	FL, MA, WA, WI 260,624/N = .019	
48	12,271,000	.000000082		PA 29568/N = .0024	IN 1921718/N = .16
47	10,738,000	.000000093		OH 225,747/N = .021	
46	9,367,000	.000000107		GA, NJ 209,541/N = .022	
44	7,059,000	.000000142	LA, MO 10774/N = .0015	CN, OR, VA 179,494/N = .025	
42	5,246,000	.000000191	KY 9667/N = .0018	AR, CO, MA, IA 152,467/N = .029	
41	4,496,000	.000000222		ID 140,036/N = .031	
40	3,838,000	.000000261		MN, NH, VT 128,300/N = .033	
39	3,263,000	.000000307	DC 8119/N = .0024		
38	2,761,000	.000000362	PR 7633/N = .0027		
36	1,948,000	.000000513	DE, WI 6706/N = .0034	NH 87906/N = .045	
33	N = 1,108,000 N^+ = 36,550,000	.000000027		ID, KA, SD, MT 63928/N=.058	PA 1,281,264/N^+=.035

Table 3

States having 5/n Lotto, showing probabilities of a jackpot-winning ticket, and of winning *some* prize.

n = Number of numbers in the lottery	P = Probability of having a jackpot-winning ticket (5 hits): N(nearest 1000) for P = 1/N	P (7 dec.pts.)	States and Probability (to 3 dec. pts.) of winning a prize m or more hits: m = 3	m = 2	m = 0
45	N = 1,222,000 N^+ = 54,979,000	.000000018			"Powerball"* 1574000/N^+ = .029
39	576,000	.0000017	CO, MD, MN, MI 5781/N = .010	PA, NY 65621/N = .114	
38	502,000	.0000020	NJ 5446/N = .011		
37	436,000	.0000022	MO 5121/N = .012		
35	325,000	.0000031	AZ, GA, MA 4501/N = .014	SD, IA, NH 45101/N = .138	
31	N = 170,000 N^+ = 850,000	.0000012		ME, NH, VT 42805/N^+ = .050	
30	143,000	.0000070	IL, MO, NE, RI 3126/N = .022		
26	66,000	.0000152	FL, LA 2206/N = .033		

* Arizona, District of Columbia, Delaware, Georgia, Idaho, Iowa, Kansas, Louisiana, Minnesota, Missouri, Montana, Nebraska, Oregon, S. Dakota, W. Virginia, Wisconsin.

APPENDIX: PROBABILITY FORMULAE

The Lotto games summarized in Tables 2 and 3 are either 6/n or 5/n games. For 6/n, the number of combinations of six numbers that can be chosen from n numbers, denoted by $\binom{n}{6}$ in the mathematical literature, is what we will call N:

$$N = N_6 = \binom{n}{6} = n(n-1)(n-2)(n-3)(n-4)(n-5)/720 . \tag{1}$$

One of these N combinations gets chosen at random to be designated the jackpot-winning number. Hence

$$P = \Pr\left\{\text{a purchased ticket wins or shares the jackpot}\right\} = 1/N .$$

For n in the range 30-55, the value of N is large (e.g., for n = 54, in New York, N is more than twenty-five million), and so the probability 1/N is very small: e.g., 0.000000039.

The same is true for 5/n Lotto, where

$$N = N_5 = \binom{n}{5} = n(n-1)(n-2)(n-3)(n-4)/120 . \tag{2}$$

Values of N and P = 1/N are shown in Tables 2 and 3 for values of n used throughout the U.S.A.

In r/n Lotto, denote the number of possible tickets having w hits by $N_w(r/n)$. Then it is a straightforward combinatoric argument to show that

$$N_w(r/n) = \binom{r}{w}\binom{n-r}{r-w} = \frac{r(r-1)\ldots(r-w+1)}{1(2)\ldots(w)} \frac{(n-r)(n-r-1)\ldots(n-2r+w+1)}{1(2)\ldots(r-w)}$$

and with $N_r(r/n) = 1$ (the jackpot-winning ticket), the relationship between $N_w(r/n)$ and $N_{w-1}(r/n)$ is, for $w = r, r-1, \cdots, 0$,

$$N_{w-1}(r/n) = \frac{w(n-2r+w)}{(r-w+1)^2} N_w(r/n) .$$

This in turn leads to the formulae of Table 4. Dividing those expressions by N leads, in most cases, to the probabilities of winning shown in Tables 2 and 3.

Table 4

Number of ways of winning in r/n Lotto.

No. of hits w	r = 6: Possible number of winning tickets: With w hits	r = 6: With w or more hits	r = 5: Possible number of winning tickets: With w hits	r = 5: With w or more hits
2	$5(n-6)(n-7)(n-8)(n-9)/8$	$1+(n-6)(15n^3-280n^2+1845n-4196)/24$	$5(n-5)(n-6)(n-7)/3$	$1+5(n-5)(n^2-10n+27)/3$
3	$10(n-6)(n-7)(n-8)/3$	$1+(n-6)(20n^2-255n-841)/6$	$5(n-5)(n-6)$	$1+5(n-5)^2$
4	$7.5(n-6)(n-7)$	$1+1.5(n-6)(5n-31)$	$5(n-5)$	$1+5(n-5)$
5	$6(n-6)$	$1+6(n-6)$	1	1
6	1	1	Not applicable	

Table 5

Ratios of probabilities in 6/n with those in 5/n.

	Values for n = 54	Values for n = 30
Probability of a jackpot-winning ticket		
$(1/N_6) \Big/ (1/N_5) = N_5/N_6 = \dfrac{n(n-1)(n-2)(n-3)(n-4)/120}{n(n-1)(n-2)(n-3)(n-4)(n-5)/720} = \dfrac{6}{n-5}$	.122	.240
Hence this probability for 6/n is approximately 10%-20% of that for 5/n.		
Probability of winning some prize with 5 or more hits		
$\dfrac{6n-35}{N_6} \Big/ \dfrac{1}{N_5} = \dfrac{6(6n-35)}{n-5}$	35.4	34.8
Thus this probability for 6/n is approximately 35 times that for 5/n.		
Probability of winning some prize with 4 or more hits		
$\dfrac{1+1.5(n-6)(5n-31)}{N_6} \Big/ \dfrac{5n-24}{N_5} = \dfrac{6+9(n-6)(5n-31)}{(n-5)(5n-24)}$	8.6	8.2
This probability for 6/n is approximately 8 times that of 5/n.		
Probability of winning some prize with 6, 5 or 4 hits in 6/n compared to 5, 4 or 3 hits in 5/n		
$\dfrac{1+1.5(n-6)(5n-31)}{N_6} \Big/ \dfrac{1+5(n-5)^2}{N_5} = \dfrac{6+9(n-6)(5n-31)}{(n-5)[1+5(n-5)^2]}$	.18	.33
This probability is 1.0 for n = 6.7 and 11 (of no practical use); otherwise it is less than 1.0.		