# Running a lottery, for beginners

## by John Haigh

BOOKMARK
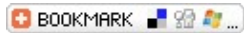
*To appreciate why Lottery formats are what they are, consider how you would design an ongoing Lottery from scratch. Your aim is to maximise the total profit, within the legal constraints that all tickets are treated equally.*

## Lottery Formats

It is reasonable to assume that the more attractive your game, the higher the gross income. Experience indicates that the single feature that attracts players is the possibility of winning an enormous sum of money. The actual chance of winning seems to be relatively unimportant. If you sell enough tickets, even though each one has a remote chance of winning a life−changing sum, it is likely that *someone* will win. And that is enough to establish the possibility as real. Another observation is that many players like to win *something* from time to time (even just a free entry to the next lottery), to keep their interest up.

If $N$ tickets are sold, each costing one unit of currency (a pound or a dollar, say), and a fraction $f$ of this sum is returned as prize money, then the profits will be $(1–f)N$ (less costs and any taxes imposed). The choice of $f$ must be *large* enough to generate the potential of a very large prize, but also *small* enough to leave a worthwhile profit. As well as fixing $f$, how the prize money should be divided among larger and smaller prizes must be decided. The value of $N$ will depend on the size of the target population, and the attractiveness of the game (in which the value of $f$ will play a role).

Modern lotteries that have proved successful include three main formats: the Genoese type (with variations); Keno games; and Numbers games. Within any of these formats, there is a another choice: shall all winners at

a given level get a fixed sum, or shall they have equal shares in a total amount allocated to that level? (In horse–race betting, both alternatives, bookmakers offering fixed odds, and the Tote using a pari mutuel system, operate in parallel.) The very highest prizes arise in lotteries that offer pari mutuel payouts.



# The Genoese format

In the principal game offered in the UK, players select six numbers from a list of 49 numbers; then six (winning) numbers are selected at random from this list, and players win prizes according to how many of the winning numbers they have chosen (see The UK National Lottery – a guide for beginners in issue 29 of *Plus*). This is the Genoese format, named in honour of lotteries used in Genoa in the Middle Ages. A game in this family will be written *m/M*; the UK game is 6/49.

The choice of *m*=6 and *M*=49 is also found in Canada, Germany, South Africa and a number of states in the USA. The *m/M* format offers great flexibility. The chance that a player chooses all the winning numbers, usually also the chance of winning a share of the maximum prize – a jackpot – is $p=1/{}^{M}C_m$. Fine–tuning of the choices of *M* and *m* enables *p* to be close to any desired quantity. For example, ${}^{32}C_6=906,192$, ${}^{33}C_6=1,107,568$, suggest lotteries whose winning chances are either side of one in a million. Taking *m*=6 and *M*=49 gives *p* close to one in 14 million, and the Michigan Rolldown is 5/33, with *p* just over one in 250,000.

# Choosing your format

If you expect to sell only about half a million tickets at £1 each, the jackpot prize cannot be enormous; and a 6/49 format here would mean that the top prize would be won rarely – once every 28 games. With such a remote prospect, interest might well evaporate quickly. The games offered reflect the sizes of the target populations.

But suppose the overall chance that nobody wins the jackpot is round about one in four, and the jackpot amount is shared equally among all who qualify – the pari–mutuel version. Then the trick of carrying unwon prize money forward to the next Lottery (a *rollover*) can boost the jackpot size, thereby enhancing the attractiveness of the next draw. Greater sales there increase the chance that this larger prize is actually won, reinforcing future sales. The UK Lottery has another method of tempting punters – the *Superdraw*, where extra money is added to the jackpot prize. (Although $f$=50% of ticket sales will eventually be returned as prize money, the amount allocated to immediate prizes is just 45% of that draw's sales; the other 5% is accumulated for these special occasions.)

Variations on this Genoese lottery give convenient ways of tweaking the chances of winning. For example, in addition to selecting $m$ numbers from $M$, players may be asked also to select one or more numbers from a separate list of $K$ numbers. Jackpot winners are those who match all winning numbers in both categories. Powerball, played across 24 states in the USA, currently asks players to select five numbers from 53, and one other (the *Powerball*) from a separate list of 42, giving $^{53}C_5$=120,526,770 different choices available.

This means the winning chance is so small that, despite the large number of players, rollovers are frequent: the potential jackpot can then exceed two hundred million dollars. The UK game, Thunderball, is a humbler affair, with fixed prizes: players choose five numbers from 34, and one from 14, so the chance of winning the top prize of $£250,000$ is $^{34}C_5 \times 14$, just better than one in four million. A Kansas game asks players to choose two numbers from 26 White numbers, and also two from 26 Red numbers, which leads to one chance in 105,625 of winning the top prize in this state with a small population. Noting that $^{14}C_4 = 1001$, a game in which players chose four White numbers from 14, and also four Red numbers from 14 would give a winning chance trivially different from one in a million.

In all these formats, lesser prizes can be won by matching just some of the winning numbers. Statistically, the hypergeometric distribution is used to compute the chances of winning at any level. If there are $m$ winning numbers, $M - m$ losing numbers, and players select $N$ of these $M$ numbers, the chance they have exactly $r$ winning numbers is

$$\frac{^mC_r \times {}^{M-m}C_{N-r}}{^MC_N},$$

provided only that the *winning numbers* are chosen at random.

## Interestingly non−random
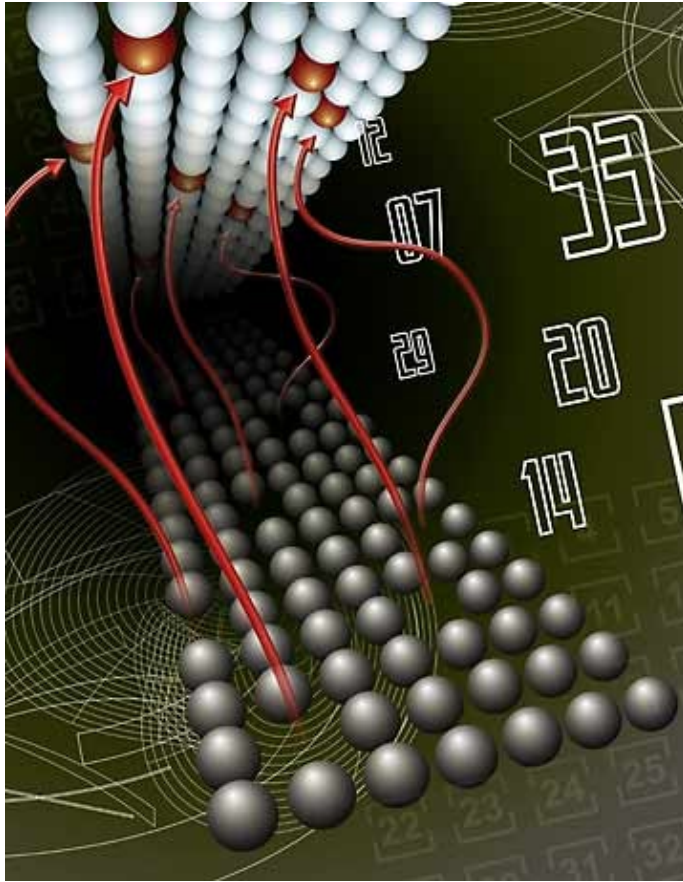


There is one other important factor to consider in designing a game: left to their own devices, players do not select all possible combinations with equal probabilities. Some types of choice – see The UK National Lottery – a guide for beginners in issue 29 of *Plus* – are far more popular than average, others less so. Because the draw itself is random, this skewness in player choice leads to MORE rollovers than a genuine random choice by players would give. As rollovers tend to increase sales, hence profits, *it is in the Lottery designers' interest* that players tend to select some combinations far more often than others.

This point is illustrated by the experience of the New Zealand Lottery, where a large proportion of players use an official random choice facility. This tends to even out the choice frequencies, leading to very few rollovers, and a relatively predictable (and disappointing) jackpot.

# Keno

Keno is often played for 16+ hours per day, with draws every few minutes. In its most common format, $m = 20$ winning numbers will be selected at random from $M = 80$ numbers, while players try to match some fixed number $R$, say $1 \leq R \leq 12$, of them. A prize might require at least $r$ of the player's numbers to be among the winning numbers; the hypergeometric formula above will give the various winning chances.

Because of Keno's frequency of play, the fixed prize format is the norm. Prizes associated with very low winning chances can be set at an eye−catching level, but with one proviso: if players are permitted to select

their own numbers, their non–random choices could mean an occasional disaster for the Lottery, when too many of them all qualify for a large sum. As a form of risk control, it is prudent to specify a maximum amount that can be paid out on any one game.

The integrity of any lottery is vital. Genoese lotteries tend to use a physical device, such as numbered balls swirling round a transparent plastic tub; if this can be shown live on TV, so much the better. But Keno, like rapid–play internet gambling games, usually invokes the pseudo–random number generator of a computer. This is inherently dangerous: not only have some such generators been found to have severe flaws, but the possibility of corruption is plain.

To illustrate Keno, the Table shows the winning chances (rounded), and prize levels for unit stake, when players select 10 numbers in the Massachusetts (M, in red) and Washington (W, in blue) versions. Note the nod towards players who consider themselves unlucky –– a prize is given for matching *no* numbers.

| *Number matched* | 0 | 5 | 6 | 7 | 8 | 9 | 10 |
|---:|---|---|---|---|---|---|---|
| *Chances: 1 in* | 22 | 19.5 | 87 | 620 | 7,400 | 160,000 | 9,000,000 |
| *Payout(M)* | 2 | 2 | 20 | 80 | 500 | 10,000 | 100,000 |
| *Payout(W)* | 3 | 2 | 5 | 50 | 500 | 5,000 | 100,000 |

It may not be immediately clear which represents better value to the player: but on average, over 69% of stake money is returned in Massachusetts, under 49% in Washington. In a different gamble, cautious players may place bets on just one number, giving themselves one chance in four of winning: even here, Washington pays out only $2, while Massachusetts pays $2.50 to a winning one dollar stake. Some Lottery designers are far more generous than others!

In a standard Numbers Game, four (or sometimes three) digits are drawn independently, and in a definite order, from the list {0,1,...,9}. There are thus 10,000 (or 1000) different possibilities. Players attempt to match all, or some, of the digits. This game is usually played with fixed payouts, with a mean return of about 50%, but the Massachusetts version is pari mutuel, and returns 60%. Working out the winning odds here is plainly straightforward. Players may "box" their bets, i.e. select the digits, without specifying their order. Thus boxing four different digits divides the stake among 24 outcomes, while boxing (say) 5,5,5,6 splits the stake among four outcomes.

# Getting the numbers wrong

Lottery designers are generally careful in their approach, but some blunders have been made, even in modern times. In the Scramble Game in Canada in 1981, players would select six digits, (as in a Numbers game) and win a fixed sum if their selection matched the winning six–digit selection *in any order*. Thus, 123456 had 720 winning chances, whereas 222222 had just one chance! In another Canadian game in 1978–9, in which the digits 0 to 9 should have had equal chances of selection, an oversight meant that for each eight times a digit from 6 to 9 appeared, a digit from 3 to 5 would arise ten times, while digits 0 to 2 would arise eleven times. And the early rules of the Massachusetts Numbers game meant that players who restricted their bets to just three of the four numbers would obtain a higher mean return than players who bet on all four.

## About the author



John Haigh is Reader in Mathematics and Statistics at the University of Sussex. His book, *Taking Chances*, (now translated into Spanish and Chinese, and in a second edition) aimed at helping the layperson understand ideas of probability, is an all–time *Plus* favourite. You can read our review in Issue 13.